On Undetected Error Probability of Binary Matrix Ensembles

Tadashi Wadayama†

Abstract—In this paper, an analysis of the undetected error probability of ensembles of \( m \times n \) binary matrices is presented. The ensemble called the Bernoulli ensemble whose members are considered as matrices generated from i.i.d. Bernoulli source is mainly considered here. 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. Furthermore, as a byproduct of the proof of the variance formula, simple covariance formula of the weight distribution is 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 an ensemble of sparse matrices. 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 the average weight distributions of LDPC codes, such as those described by Litsyn and Shevelev [4][5], Burshtein and Miller [6], Richardson and Urbanke [9], 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 [7][8] 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], [3] 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. 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.

II. AVERAGE UNDETECTED ERROR PROBABILITY

A. Notation

For a given \( m \times n \) binary parity check matrix \( H \), let \( C(H) \) be the binary linear code of length \( n \) defined by \( H \), namely, \( C(H) \triangleq \{ x \in F_2^n : Hx^t = 0^m \} \) where \( F_2 \) is the Galois field with two elements \( \{0, 1\} \) (the addition over \( F_2 \) is denoted by \( \oplus \)). The notation \( 0^m \) denotes the zero vector of length \( m \). In this paper, a boldface letter, such as \( x \) for example, denotes a binary row vector.

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

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

\[
P_U(H) = \sum_{e \in C(H), e \neq 0^m} \epsilon^{w(e)} (1 - \epsilon)^{n - w(e)}
\]

where \( w(x) \) denotes the Hamming weight of vector \( x \). The

†Nagoya Institute of Technology, email:wadayama@nitech.ac.jp. A part of this work was presented at ITA workshop in UCSD, Feb. 2007.
above equation can be rewritten as
\[ P_U(H) = \sum_{w=1}^{n} A_w(H) e^w (1 - \epsilon)^{n-w}, \] (2)
where \( A_w(H) \) is defined by
\[ A_w(H) \triangleq \sum_{x \in Z^{(n,w)}} I[Hx^t = 0^m]. \] (3)
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 \( I[\text{condition}] \) is the indicator function such that \( I[\text{condition}] = 1 \) if \( \text{condition} \) is true; otherwise, it evaluates to 0.

Suppose that \( G \) is a set of binary \( m \times n \) matrices \( (m, n \geq 1) \). Note that \( G \) 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 \( H \in G \). Thus, \( G \) can be considered as an ensemble of binary matrices. Let \( f(H) \) be a real-valued function which depends on \( H \in G \). The expectation of \( f(H) \) with respect to the ensemble \( G \) is defined by
\[ E_G[f(H)] \triangleq \sum_{H \in G} P(H) f(H). \] (4)
The average weight distribution of a given ensemble \( G \) is given by \( E_G[A_w(H)] \). This quantity is very useful for analyzing the performance of binary linear codes, including analysis of the undetected error probability.

B. Bernoulli ensemble

In this paper, we will focus on a parameterized ensemble \( B_{m,n,k} \) which is called the Bernoulli ensemble because the Bernoulli ensemble is amenable to ensemble analysis. The Bernoulli ensemble \( B_{m,n,k} \) 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 \( p \triangleq k/n \). 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 \( H \in B_{m,n,k} \) 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 \( H \in B_{m,n,k} \) is associated with the probability
\[ P(H) = p^{\bar{w}(H)} (1 - p)^{mn - \bar{w}(H)}, \] (5)
where \( \bar{w}(H) \) is the number of ones in \( H \) (i.e., Hamming weight of \( H \)). The average weight distribution of the Bernoulli ensemble is given by
\[ E_{B_{m,n,k}}[A_w(H)] = \left( \frac{1 + z^w}{2} \right)^m \binom{n}{w} \] (6)
for \( w \in [0, n] \) where \( z \triangleq 1 - 2p \). The notation \([a, b]\) denotes the set of consecutive integers from \( a \) to \( b \). The average weight distribution of this ensemble was first discussed by Litsyn and Shevelev [4].

If \( k \) is a constant (i.e., not a function of \( n \)), this ensemble can be considered as an ensemble of sparse matrices. In the special case where \( k = n/2 \), equal probability \( 1/2^{mn} \) is assigned to every matrix in the Bernoulli ensemble. As a simplified notation, we will denote \( R_{m,n} \triangleq B_{m,n,n/2} \), where \( R_{m,n} \) is called the random ensemble. Since a typical instance of \( R_{m,n} \) contains \( \Theta(mn) \) ones, the ensemble can be regarded as an ensemble of dense 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 \( G \), we have
\[ E_G[P_U(H)] = E_G \left[ \sum_{w=1}^{n} A_w(H) e^w (1 - \epsilon)^{n-w} \right] \]
\[ = \sum_{w=1}^{n} E_G[A_w(H)] e^w (1 - \epsilon)^{n-w}. \] (7)
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 \( R_{m,n} \), the average undetected error probability has a simple closed form.

**Lemma 1:** The average undetected error probability of the random ensemble \( R_{m,n} \) is given by
\[ E_{R_{m,n}}[P_U(H)] = 2^{-m} (1 - (1 - \epsilon)^n). \] (8)
(Proof) By using (7), we have
\[ E_{R_{m,n}}[P_U(H)] = \sum_{w=1}^{n} E_{R_{m,n}}[A_w(H)] e^w (1 - \epsilon)^{n-w} \]
\[ = \sum_{w=1}^{n} 2^{-m} \binom{n}{w} e^w (1 - \epsilon)^{n-w} \]
\[ = 2^{-m} (1 - (1 - \epsilon)^n). \] (9)
The second equality is based on the well known result [1]:
\[ E_{R_{m,n}}[A_w(H)] = 2^{-m} \binom{n}{w}. \] (10)
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 \( \{G_n\}_{n>0} \) be a series of ensembles such that \( G_n \) 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 \( \{G_n\}_{n>0} \) is defined as follows.

\[
T_{G_n} = \lim_{n \to \infty} \frac{1}{n} \log_2 E_{G_n}[P_U]
\]

where \( f(\ell) \) is the asymptotic growth rate of \( \{G_n\}_{n>0} \).

(Proof) Based on the definition of asymptotic growth rate, we can rewrite \( T_{G_n} \) in the form

\[
T_{G_n} = \lim_{n \to \infty} \frac{1}{n} \log_2 E_{G_n}[P_U]
= \lim_{n \to \infty} \frac{1}{n} \log_2 \sum_{w=1}^{n} E_{G_n}[A_w] e^n(1-\epsilon)^{n-w}
= \lim_{n \to \infty} \frac{1}{n} \log_2 \sum_{w=1}^{n} 2^n f(\frac{w}{n}) + \kappa(n, w) + o(1),
\]

where \( \kappa(\epsilon, n, w) \) is defined by

\[
K(\epsilon, n, w) \triangleq \frac{w}{n} \log_2 \epsilon + \left(1 - \frac{w}{n}\right) \log_2 (1-\epsilon).
\]

Using a conventional technique for bounding summation, we have the following upper bound on \( T_{G_n} \):

\[
T_{G_n} \leq \lim_{n \to \infty} \frac{1}{n} \log_2 \sum_{w=1}^{n} 2^n f(\frac{w}{n}) + \kappa(n, w) + o(1)
= \lim_{n \to \infty} \frac{1}{n} \log_2 \max_{n=1}^{\infty} [f(\frac{w}{n}) + \kappa(\epsilon, n, w) + o(1)]
= \sup_{0<\ell \leq 1} [f(\ell) + \ell \log_2 \epsilon + (1-\ell) \log_2 (1-\epsilon)].
\]

We can also show that \( T_{G_n} \) is greater than or equal to the right-hand side of the above inequality \((\ref{eq:asymptotic_growth})\) in a similar manner. This means that the right-hand side of the inequality is asymptotically tight.

The next example discusses the case of the random ensemble.

**Example 2**: Let us again consider the series of the random ensembles given by \( \{R_n\}_{n \geq 0} \). 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

\[
h(x) \triangleq -x \log_2 x - (1-x) \log_2 (1-x).
\]

In this case, by using Theorem 1 we have

\[
T_{R_{1-R}} = \sup_{0<\ell \leq 1} [h(\ell) - (1-R) + \ell \log_2 \epsilon + (1-\ell) \log_2 (1-\epsilon)].
\]

Let

\[
D_{\ell, \epsilon} \triangleq \ell \log_2 \left( \frac{\ell}{\epsilon} \right) + (1-\ell) \log_2 \left( \frac{1-\ell}{1-\epsilon} \right).
\]

By using \( D_{\ell, \epsilon} \), we can rewrite \((\ref{eq:asymptotic_growth})\) as

\[
T_{R_{1-R}} = \sup_{0<\ell \leq 1} [- (1-R) - D_{\ell, \epsilon}].
\]

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

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

where \( f(\ell) \) is the asymptotic growth rate of \( \{G_n\}_{n>0} \).
which is identical to the exponent obtained in expression (12).
Let \( g_{(rnd)}(\ell) = h(\ell) - (1 - R) + \ell \log_2 \epsilon + (1 - \ell) \log_2 (1 - \epsilon) \). Figure 1 displays the behavior of \( g_{(rnd)}(\ell) \) when \( R = 0.5 \). This figure confirms the result that the maximum (sup\(0<\ell\leq1\)) \( g_{(rnd)}(\ell) = -0.5 \) is attained at \( \ell = \epsilon \).

\[ f(\ell) = h(\ell) + (1 - R) \log_2 \left( \frac{1 + e^{-2k\ell}}{2} \right). \] (24)

This formula is presented in [4]. The error exponent of this ensemble shows a different behavior from that for random ensembles.

**Example 3:** Consider the Bernoulli ensemble with parameters \( R = 0.5 \) and \( k = 20 \). Let

\[ g_{(spm)}(\ell) = H(\ell) + (1 - R) \log_2 \left( \frac{1 + e^{-2k\ell}}{2} \right) + \ell \log_2 \epsilon + (1 - \ell) \log_2 (1 - \epsilon). \] (25)

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

Figure 3 presents the error exponent of Bernoulli 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 \( \epsilon \) is smaller than (around) 0.3, the error exponent is a monotonically decreasing function of \( \epsilon \).

The examples suggest that a sparse ensemble has less powerful error detection performance than that of a dense ensemble (such as the 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 Bernoulli ensemble in Fig. 3 is almost equal to that of the random ensemble when \( \epsilon \) is larger than (around) 0.3.

The above properties of the error exponents of the Bernoulli 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 Bernoulli ensemble.

The weight of typical error vectors is very close to \( cn \) when \( n \) is sufficiently large. For a large value of \( \epsilon \), such as \( \epsilon = 0.4 \), the average weight distribution around \( w = 0.4n \), namely \( E_2[A_{4n}] \), dominates the undetected error probability. In such a range, the difference in the average weight distributions corresponding to the random and the Bernoulli 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 Bernoulli ensemble with constant $k$. A sparse matrix offers almost same error detection performance of a dense matrix with linear time complexity if $\epsilon$ is sufficiently large.

III. VARIANCE OF UNDETECTED ERROR PROBABILITY

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 analyze the variance of undetected error probability.

A. Covariance formula

The covariance between two real-valued functions $f(\cdot), g(\cdot)$ defined on an ensemble $\mathcal{G}$ is given by

$$
\text{Cov}_\mathcal{G}[f, g] \triangleq E_\mathcal{G}[fg] - E_\mathcal{G}[f]E_\mathcal{G}[g].
$$

(26)

The next theorem forms the basis of the derivation of the variance of the undetected error probability for the Bernoulli ensemble. The covariance of the weight distribution for the Bernoulli ensemble is given in the following theorem.

Theorem 2: The covariance of the weight distribution for the Bernoulli ensemble $B_{m,n,k}$ is given by

$$
\text{Cov}_{B_{m,n,k}}(A_{w_1}, A_{w_2}) \triangleq \left( \frac{1 + z^{w_1}}{2} \right)^m \left( \frac{1 + z^{w_2}}{2} \right)^m 
\times \sum_{v=\max(0,w_1+w_2-n)} \binom{n}{w_1} \binom{n-w_1}{v} \binom{n-w_1-w_2}{w_2-v}
\times \left( \left( 1 + \frac{z^{w_1+w_2-2v} - z^{w_1+w_2}}{1 + z^{w_1}} \right) - 1 \right)
$$

for $1 \leq w_1 \leq w_2 \leq n$ and

$$
\text{Cov}_{B_{m,n,k}}(A_{w_1}, A_{w_2}) = \text{Cov}_{B_{m,n,k}}(A_{w_2}, A_{w_1})
$$

for $1 \leq w_2 < w_1 \leq n$ where $z = 1 - 2p$ and $p = k/n$.

(Proof) See Appendix.

Remark 1: When $k = n/2$, $B_{m,n,k}$ becomes the random ensemble $R_{m,n}$. We discuss this case here.

We first assume that $1 \leq w_1 \leq w_2 \leq n$. Let $p = 1/2$ (i.e., $k = n/2$). In such a case, we have $z = 1 - 2p = 0$. Define $L$ by

$$
L \triangleq \left( 1 + \frac{z^{w_1+w_2}}{1 + z^{w_1}} \right).
$$

(29)

The variable $L$ takes the following values:

$$
L = \begin{cases} 1, & w_1 < w_2 \\ 1, & w_1 = w_2, \ v < w_1 \\ 2, & w_1 = w_2, \ v = w_1. 
\end{cases}
$$

(30)

Substituting $z = 0$ into equation (27) and using the identity (28), we get

$$
\text{Cov}_{B_{m,n,k}}(A_{w_1}, A_{w_2}) = \left\{ \begin{array}{ll} 0, & 1 \leq w_1 \neq w_2 \leq n \\ 2^{-2m(n^2)}(2m-1), & 1 \leq w_1 = w_2 \leq n. \end{array} \right.
$$

(31)

Another proof of this formula is presented in [10].

B. Variance of undetected error probability

The variance of the undetected error probability is a straightforward consequence of Theorem 2.

Corollary 1: The variance of the undetected error probability of the Bernoulli ensemble, $\sigma_{B_{m,n,k}}^2$ is given by

$$
\sigma_{B_{m,n,k}}^2 = \sum_{w_1=1}^{n} \sum_{w_2=1}^{n} \text{Cov}_{B_{m,n,k}}(A_{w_1}, A_{w_2}) \times \epsilon^{w_1+w_2}(1-\epsilon)^{2n-w_1-w_2}.
$$

(32)

(Proof) The variance of the undetected error probability $P_U$ is given by

$$
\sigma_{B_{m,n,k}}^2 = E_{B_{m,n,k}}[(P_U - \mu)^2] = E_{B_{m,n,k}}[P_U^2] - E_{B_{m,n,k}}[P_U]^2.
$$

(33)

We first consider the second moment of the undetected error probability:

$$
E_{B_{m,n,k}}[P_U^2] = E_{B_{m,n,k}} \left[ \left( \sum_{w=1}^{n} A_w \epsilon^w (1-\epsilon)^{n-w} \right)^2 \right]
$$

$$
= E_{B_{m,n,k}} \left\{ \sum_{w_1=1}^{n} \sum_{w_2=1}^{n} A_{w_1} A_{w_2} \epsilon^{w_1+w_2} (1-\epsilon)^{2n-w_1-w_2} \right\}
$$

$$
= \sum_{w_1=1}^{n} \sum_{w_2=1}^{n} E_{B_{m,n,k}}[A_{w_1} A_{w_2}] \epsilon^{w_1+w_2} (1-\epsilon)^{2n-w_1-w_2}.
$$

(34)

The squared average undetected error probability can be expressed as

$$
E_{B_{m,n,k}}[P_U]^2 = E_{B_{m,n,k}} \left[ \left( \sum_{w=1}^{n} A_w \epsilon^w (1-\epsilon)^{n-w} \right) \right]^2
$$

$$
= \sum_{w_1=1}^{n} \sum_{w_2=1}^{n} E_{B_{m,n,k}}[A_{w_1}] E_{B_{m,n,k}}[A_{w_2}] \epsilon^{w_1+w_2} (1-\epsilon)^{2n-w_1-w_2}.
$$

(35)
Combining these equalities and the covariance of the weight distribution, the variance of undetected error probability \( \sigma^2_{B_{m,n,k}} \) is obtained.

Remark 2: The covariance of the weight distribution for a given ensemble \( B_{m,n,k} \) is useful not only for the evaluation of the variance of \( P_U \). Let \( X \) be a random variable represented by

\[
X = \sum_{w=0}^{n} \alpha(w)A_w,
\]

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

\[
\sigma^2_X = \sum_{w_1=0}^{n} \sum_{w_2=0}^{n} \text{Cov}_{B_{m,n,k}}(A_{w_1}, A_{w_2}) \alpha(w_1) \alpha(w_2).
\]

A specialized version (the case where \( X = P_U \)) of this equation has been derived in the previous corollary.

Example 4: Let us consider the Bernoulli ensemble with \( m = 1, n = 2 \) and \( k = 1/2 (p = 1/4) \). Table I displays the weight distributions and undetected error probabilities for the 4 matrices in \( B_{1,2,1/2} \).

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

From the definition of a Bernoulli 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:

\[
E_{B_{1,2,1/2}}[P_U] = \frac{2}{3} \epsilon - \frac{7}{8} \epsilon^2
\]

\[
E_{B_{1,2,1/2}}[P^2_U] = \frac{21}{8} \epsilon^2 - \frac{3}{8} \epsilon^3 + \epsilon^4.
\]

From these moments, the variance can be derived:

\[
\sigma^2_{B_{1,2,1/2}} = E_{B_{1,2,1/2}}[P^2_U] - E_{B_{1,2,1/2}}[P_U]^2
\]

\[
= \frac{3}{8} \epsilon^2 - \frac{3}{8} \epsilon^3 + \frac{15}{64} \epsilon^4.
\]

We can also consider another route to derive the variance by using Corollary I. The covariances of \( B_{1,2,1/2} \) are given by

\[
\text{Cov}_{B_{1,2,1/2}}(1,1) = 3/8
\]

\[
\text{Cov}_{B_{1,2,1/2}}(1,2) = \text{Cov}_{B_{1,2,1/2}}(2,1) = 3/16
\]

\[
\text{Cov}_{B_{1,2,1/2}}(2,2) = 15/64.
\]

From Corollary I we obtain the variance

\[
\sigma^2_{B_{1,2,1/2}} = \sum_{w_1=1}^{n} \sum_{w_2=1}^{n} \text{Cov}_{B_{m,n,k}}(A_{w_1}, A_{w_2}) \times \epsilon^{w_1+w_2}(1-\epsilon)^{4-w_1-w_2}
\]

\[
= (3/8) \epsilon^2(1-\epsilon)^2 + (3/16) \epsilon^3(1-\epsilon)
\]

\[
+ (3/16) \epsilon^3(1-\epsilon) + (15/64) \epsilon^4
\]

\[
= \frac{3}{8} \epsilon^2 - \frac{3}{8} \epsilon^3 + \frac{15}{64} \epsilon^4,
\]

that is identical to expression [40].

In the case of \( k = n/2 \) (i.e. the case of a random ensemble), we can derive a closed form expression for the variance.

Corollary 2: For the random ensemble \( R_{m,n} \), the variance of the undetected error probability \( P_U \) is given by

\[
\sigma^2_{R_{m,n}} = (1 - 2^{-m})2^{-m} \left( (\epsilon^2 + (1-\epsilon)^2)^n - (1-\epsilon)^{2n} \right).
\]

(Proof) The variance of undetected error probability \( \sigma^2_{R_{m,n}} \) can be obtained in the following way:

\[
\sigma^2_{R_{m,n}} = E_{R_{m,n}}[P^2_U] - E_{R_{m,n}}[P_U]^2
\]

\[
= \sum_{w_1=1}^{n} \sum_{w_2=1}^{n} \text{Cov}_{R_{m,n}}(A_{w_1}, A_{w_2}) \epsilon^{w_1+w_2}(1-\epsilon)^{2n-w_1-w_2}
\]

\[
= (1 - 2^{-m})2^{-m} \sum_{w=1}^{n} (\frac{n}{w}) \epsilon^2(1-\epsilon)^{2n-2w}.
\]

The second equality is due to Corollary I. The last equality are due to Eq. [51]. We can further simplify the expression using the binomial theorem:

\[
\sigma^2_{R_{m,n}} = (1 - 2^{-m})2^{-m} \sum_{w=0}^{n} \left( \frac{n}{w} \right) (\epsilon^2)^w ((1-\epsilon)^2)^{n-w}
\]

\[
- (1 - 2^{-m})2^{-m} (1-\epsilon)^{2n}
\]

\[
= (1 - 2^{-m})2^{-m} \times ((\epsilon^2 + (1-\epsilon)^2)^n - (1-\epsilon)^{2n}).
\]

The last equality is the claim of the theorem.

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

Example 5: We consider the random ensemble with \( m = 20, n = 40 \), and the Bernoulli ensemble with \( m = 20, n = 40, k = 5 \) (labeled “Sparse” in Fig. 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 \( \epsilon \) decreases. In contrast, the curve for the Bernoulli ensemble has a peak around \( \epsilon \approx 0.025 \). 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.

C. Asymptotic behavior

We here discuss the asymptotic behavior of the covariance of the weight distribution and the variance of \( P_U \) for the Bernoulli ensemble. The following corollary explains the
The function $\alpha(\mu, \nu)$ is defined by
\[
\alpha(\mu, \nu) \triangleq h\left(\frac{\mu}{1 - R}\right) + \mu \log_2 \left(\frac{e^{-2k(\ell_1 + \ell_2 - 2\nu)} - e^{-2k(\ell_1 + \ell_2)}}{1 - R - \mu}\right) \\
+ \left(1 - R - \mu\right) \log_2 \left(\frac{1 + e^{-2k(\ell_1)}(1 + e^{-2k\ell_2})}{1 + e^{-2k(\ell_1 + \ell_2)}}\right).
\] (50)

(Proof) We here rewrite the covariance formula (27) into asymptotic form. By using the Binomial theorem, we have
\[
\left(1 + \frac{z^{w_1+w_2-2\nu} - z^{w_1+w_2}}{1 + z^{w_1}}\right)^m - 1 \\
= \sum_{i=1}^{m} \binom{m}{i} \left(\frac{z^{w_1+w_2-2\nu} - z^{w_1+w_2}}{1 + z^{w_1}}\right)^i.
\] (51)

By using this identity, the covariance in (27) can be rewritten in the following form:
\[
\text{Cov}_{m,n,k}(A_{w_1}, A_{w_2}) \\
= 2^{-2m} \sum_{v=1}^{m} \left(\frac{n}{w_1}\left(\frac{n - w_1}{n - w_2 - v}\right)^\nu \Theta, \right.
\]
where $\Theta$ is defined by
\[
\Theta \triangleq \sum_{i=1}^{m} \binom{m}{i} \left(\frac{z^{w_1+w_2-2\nu} - z^{w_1+w_2}}{1 + z^{w_1}}\right)^i \\
\times \left((1 + z^{w_1})(1 + z^{w_2})\right)^{m-i}.
\] (52)

Letting $w_1 = \ell_1 n, w_2 = \ell_2 n, \nu = \nu m, m = (1 - R)n$, we have
\[
\lim_{n \to \infty} \frac{1}{n} \log_2 2^{-2m} = -2(1 - R) \quad (53)
\]
and
\[
\lim_{n \to \infty} \frac{1}{n} \log_2 \left(\frac{n}{w_1}\right)^\nu \left(\frac{n - w_1}{w_2 - v}\right)^\gamma \\
= h(\ell_1) + h\left(\frac{\nu}{\ell_1}\right) + h\left(\frac{\ell_2 - \nu}{1 - \ell_1}\right).
\] (54)

If $k$ is a constant and $0 \leq \ell \leq 1$, then, making use of the identity [4]
\[
\lim_{n \to \infty} \left(1 - 2\left(\frac{k}{n}\right)^{\ell n} = \lim_{n \to \infty} z^{\ell n} = e^{-2k\ell}
\] (55)
we get
\[
\lim_{n \to \infty} \frac{1}{n} \log_2 \Theta = \sup_{0 < \mu \leq 1 - R} \alpha(\mu).
\] (56)

Combining these asymptotic expressions, the claim of the corollary is derived.

The following corollary gives the asymptotic growth rate of the variance of the undetected error probability.

**Corollary 4:** The asymptotic growth rate of the variance of the undetected error is given by
\[
\lim_{n \to \infty} \frac{1}{n} \log_2 \sigma^2_{S_{m,n,(1-R)n,k}} = \sup_{0 < \ell_1, 0 < \ell_2 \leq 1} S(\ell_1, \ell_2),
\] (57)
where $S(\ell_1, \ell_2)$ is given by

$$S(\ell_1, \ell_2) \triangleq (\ell_1 + \ell_2) \log_2 \epsilon + (2 - \ell_1 - \ell_2) \log_2(1 - \epsilon) + T(\ell_1, \ell_2).$$

(Proof) It is evident that

$$\lim_{n \to \infty} \frac{1}{n} \log_2 \left( e^{\ell_1 + \ell_2 - \epsilon} (1 - \epsilon)^{2n - \ell_1 - \ell_2 - n} \right) = (\ell_1 + \ell_2) \log_2 \epsilon + (2 - \ell_1 - \ell_2) \log_2(1 - \epsilon).$$

holds. Combining this identity and Corollaries 1 and 3, we immediately have the claim of the corollary.

IV. APPENDIX

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

$$E_G[A_{w_1}, A_{w_2}] = E_G \left[ \sum_{x \in Z^{(n, w_1)}} \sum_{y \in Z^{(n, w_2)}} I[Hx^t = 0^m]I[Hy^t = 0^m] \right].$$

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

$$I[Hx^t = 0^m]I[Hy^t = 0^m] = I[Hx^t = 0^m, Hy^t = 0^m],$$

we have

$$E_G[A_{w_1}, A_{w_2}] = E_G \left[ \sum_{x \in Z^{(n, w_1)}} \sum_{y \in Z^{(n, w_2)}} I[Hx^t = 0^m, Hy^t = 0^m] \right].$$

We here encounter a problem of evaluating probability of occurrence of both $Hx^t = 0^m$ and $Hy^t = 0^m$. In preparation to solve this problem, we will introduce some notation:

**Definition 3**: For a given pair $(x, y) \in Z^{(n, w_1)} \times Z^{(n, w_2)}$, the index sets $I_1, I_2, I_3, I_4$ are defined as follows:

$$I_1 \triangleq \{ k \in [1, n] : x_k = 1, y_k = 0 \}$$

$$I_2 \triangleq \{ k \in [1, n] : x_k = 1, y_k = 1 \}$$

$$I_3 \triangleq \{ k \in [1, n] : x_k = 0, y_k = 1 \}$$

$$I_4 \triangleq \{ k \in [1, n] : x_k = 0, y_k = 0 \},$$

where $x = (x_1, x_2, \ldots, x_n)$ and $y = (y_1, y_2, \ldots, y_n)$. 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 $h = (h_1, h_2, \ldots, h_n)$ be a binary $n$-tuple. The partial weight of $h$ corresponding to an index set $I_k(k = 1, 2, 3, 4)$ is denoted by $w_k(h)$, namely

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

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

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

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

A. Proof of Lemma 2 (Covariance of the Bernoulli ensemble)

Let $x \in Z^{(n, w_1)}$ and $y \in Z^{(n, w_2)}$ be binary vectors satisfying $w_1 \leq w_2$. In this proof, we first prove the following equality:

$$E_{G_{n,m,k}}[I[Hx^t = 0, Hy^t = 0]] = \left( \frac{1 + z^{w_1} + z^{w_2} + z^{w_1 + w_2 - 2n}}{4} \right)^m \quad (67)$$

where $v = \#(\text{Supp}(x) \cap \text{Supp}(y))$, $z = 1 - 2p$ and $p = k/n$. The support set $\text{Supp}(v)$ is defined by

$$\text{Supp}(v) \triangleq \{ i \in [1, n] : v_i \neq 0 \},$$

where $v = (v_1, v_2, \ldots, v_n)$.

We need to consider the following three cases: Case (i): $0 < i_2 < w_1$ (i.e., the intersection of $\text{Supp}(x)$ and $\text{Supp}(y)$ is not empty but $\text{Supp}(y)$ does not include $\text{Supp}(x)$), Case (ii): $i_2 = 0$ (i.e., the intersection of $\text{Supp}(x)$ and $\text{Supp}(y)$ is empty), Case (iii): $i_2 = w_1$ (i.e., $\text{Supp}(y)$ includes $\text{Supp}(x)$).

We first study Case (i). Suppose that a binary $n$-tuple $h$ is generated from a Bernoulli source with $Pr[h_i = 1] = p(i \in [1, n])$. Recall that $p$ is defined by $p = k/n$. In this case, $hx^t = 0$, $hy^t = 0$ holds if and only if $w_1(h)$ is even for $i = 1, 2, 3$ or $w_3(h)$ is odd for $i = 1, 2, 3$.

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

Based on the above argument, we can write the probability $Pr[hx^t = 0, hy^t = 0]$ as a function of $z$:

$$Pr[hx^t = 0, hy^t = 0] = \left( 1 + z^1 + z^{1+2} + z^{1+3} + z^2 + z^{2+1} + z^{2+3} \right) \left( 1 - z^1 \right) \left( 1 - z^2 \right) \left( 1 - z^3 \right)$$

$$= \frac{1 + z^{w_1} + z^{w_2} + z^{w_1 + w_2 - 2n}}{4}. \quad (69)$$

where $v \triangleq i_2$.

We next consider Case (ii). For this case, $v = i_2$ is assumed to be zero. In this case, $hx^t = 0$, $hy^t = 0$ holds if and only if both $w_1(h)$ and $w_3(h)$ are even. The probability that $h$
satisfies $hx^t = 0$ and $hy^t = 0$ under the condition $i_2 = 0$ is given by

$$Pr[hx^t = 0, hy^t = 0] = \left( \frac{1 + z^{i_2}}{2} \right) \left( \frac{1 + z^{i_3}}{2} \right) = \frac{1 + z^{w_1} + z^{w_2} + z^{w_1 + w_2}}{4}.$$  \hspace{1cm} (70)

Finally we consider Case (iii). Assume the case $v = i_2 = w_1, x \neq y$. In this case, $hx^t = 0, hy^t = 0$ holds if and only if both $w_2(h)$ and $w_3(h)$ are even. The probability $Pr[hx^t = 0, hy^t = 0]$ under the condition $v = w_1, x \neq y$ is thus given by

$$Pr[hx^t = 0, hy^t = 0] = \left( \frac{1 + z^{i_2}}{2} \right) \left( \frac{1 + z^{i_3}}{2} \right) = \frac{1 + z^{w_1} + z^{w_2} + z^{w_1 + w_2}}{4}.$$  \hspace{1cm} (71)

We next consider the case $x = y$. For this case, we also have

$$Pr[hx^t = 0, hy^t = 0] = \frac{1}{2}.$$  \hspace{1cm} (72)

In summary, for any cases (Cases (i), (ii), (iii)),

$$Pr[hx^t = 0, hy^t = 0] = \frac{1 + z^{w_1} + z^{w_2} + z^{w_1 + w_2} - 2v}{4}.$$  \hspace{1cm} (73)

holds. Since the rows of parity check matrices in $B_{n,m,k}$ can be independently chosen, we obtain Eq. (67) in the following way:

$$E_{B_{n,m,k}}[I[hx^t = 0, hy^t = 0]] = Pr[hx^t = 0, hy^t = 0] = \frac{1 + z^{w_1} + z^{w_2} + z^{w_1 + w_2} - 2v}{4}.$$  \hspace{1cm} (74)

Combining (60) and (67), we have

$$E_{B_{n,m,k}}[A_{w_1} A_{w_2}] = \sum_{x \in Z(n, w_1)} \sum_{y \in Z(n, w_2)} E_{B_{n,m,k}}[I[hx^t = 0^m, hy^t = 0^m]] = \sum_{x \in Z(n, w_1)} \sum_{y \in Z(n, w_2)} \left( \frac{1 + z^{w_1} + z^{w_2} + z^{w_1 + w_2 - 2v}}{4} \right)^m = \sum_{v = \max \{0, w_1 - w_2 - n\}}^{w_1} \left( \frac{n}{w_1} \right) \left( \frac{1 + z^{w_1} + z^{w_2} + z^{w_1 + w_2 - 2v}}{4} \right)^m.$$  \hspace{1cm} (75)

Since

$$E_{B_{n,m,k}}[A_{w_1} A_{w_2}] = \left( \frac{n}{w_1} \right) \left( \frac{1 + z^{w_1}}{2} \right)^m \left( \frac{1 + z^{w_2}}{2} \right)^m \left( \frac{1 + z^{w_1 + w_2}}{2} \right)^m,$$

holds [4], we thus have

$$E_{B_{n,m,k}}[A_{w_1} A_{w_2}] = \left( \frac{n}{w_1} \right) \left( \frac{1 + z^{w_1}}{2} \right)^m \left( \frac{1 + z^{w_2}}{2} \right)^m \left( \frac{1 + z^{w_1 + w_2}}{2} \right)^m.$$  \hspace{1cm} (76)

holds [4], we thus have

$$E_{B_{n,m,k}}[A_{w_1} A_{w_2}] = \left( \frac{n}{w_1} \right) \left( \frac{1 + z^{w_1}}{2} \right)^m \left( \frac{1 + z^{w_2}}{2} \right)^m \left( \frac{1 + z^{w_1 + w_2}}{2} \right)^m.$$  \hspace{1cm} (77)

The last equality is due to the following combinatorial identity:

$$\sum_{v = \max \{0, w_1 - w_2 - n\}}^{w_1} \binom{n}{w_1} \binom{w_1}{v} \binom{w_2 - v}{w_2} = \frac{n}{w_1} \binom{n}{w_2}.$$  \hspace{1cm} (78)

We are ready to derive the covariance of weight distributions for the case $w_1 \leq w_2$. Substituting (75) and (77) into

$$\text{Cov}_{B_{n,m,k}}(A_{w_1}, A_{w_2}) = E_{B_{n,m,k}}[A_{w_1} A_{w_2}] - E_{B_{n,m,k}}[A_{w_1}] E_{B_{n,m,k}}[A_{w_2}] = \text{Cov}_{B_{n,m,k}}(A_{w_1}, A_{w_1})$$

we have (77) in the claim part of the Theorem. Since the definition of covariance is commutative, $\text{Cov}_{B_{n,m,k}}(A_{w_1}, A_{w_2}) = \text{Cov}_{B_{n,m,k}}(A_{w_2}, A_{w_1})$ holds if $w_1 > w_2$. \hfill \Box

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, "Codes for Error Detection", World Scientific, 2007.
[3] T. Klove and V. Korzhik, "Error Detecting Codes: General Theory and Their Application in Feedback Communication Systems", Kluwer Academic, 1995.
[4] 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.
[5] S.Litsyn and V. Shevelev, “On ensembles of low-density parity-check codes: asymptotic distance distributions,” IEEE Trans. Inform. Theory, vol.49, pp.3140–3159, Nov. 2003.
[6] D.Burshtein and G. Miller, “Asymptotic enumeration methods for analyzing LDPC codes,” IEEE Trans. Inform. Theory, vol.50, pp.1115–1131, June 2004.
[7] 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.
[8] V. Radh, “On the asymptotic weight distribution of regular LDPC ensembles,” in Proceedings of International Symposium on Information Theory, 2005.
[9] T. Richardson, R. Urbanke, “Modern Coding Theory,” online: http://ithwww.epfl.ch/
[10] T.Wadayama, “Asymptotic concentration behaviors of linear combinations of weight distributions on random linear code ensemble,” ArXiv, arXiv:0803.1025v1 (2008).