

## On the Asymptotic Weight and Stopping Set Distribution of Regular LDPC Ensembles

Vishwambhar Rathi, *Student Member, IEEE*

**Abstract**— In this correspondence, we estimate the variance of weight and stopping set distribution of regular low-density parity-check (LDPC) ensembles. Using this estimate and the second moment method we obtain bounds on the probability that a randomly chosen code from regular LDPC ensemble has its weight distribution and stopping set distribution close to respective ensemble averages. We are able to show that a large fraction of total number of codes have their weight and stopping set distribution close to the average.

**Index Terms**—Low-density parity-check (LDPC) codes, second moment method, stopping set distribution, weight distribution.

### I. INTRODUCTION

The weight distribution is an important characterization of a code. For a code  $G$  of block length  $n$ , we define  $N(G, n\omega)$  as the weight distribution function, denoting the number of codewords with normalized weight  $\omega$ .<sup>1</sup> In general,  $N(G, n\omega)$  is hard to compute for a specific code. In fact, even the determination of the minimum distance is NP-complete [1]. On the other hand, for some ensembles of codes it is easy to compute the *expected* weight distribution function, i.e.,  $\mathbb{E}[N(G, n\omega)]$ . This is true for e.g., Shannon's random ensemble but also for suitably defined low-density parity-check (LDPC) ensembles. A possible approach to study the weight distribution of *individual* codes is to first compute the ensemble average and then to show that most codes have a weight distribution close to this average. For LDPC codes, it has been conjectured that for regular ensembles most codes have a weight distribution close to the ensemble average [2], [3].

In 1989, Sourlas showed that there is a strong connection between error-correcting codes and disordered spin models [4], [5]. To this end, let us consider the exponent  $\frac{1}{n} \ln N(G, n\omega)$  and define

$$W_{\text{sp}}(\omega) := \lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{E}[\ln N(G, n\omega)]$$

$$W_{\text{com}}(\omega) := \lim_{n \rightarrow \infty} \frac{1}{n} \ln \mathbb{E}[N(G, n\omega)]$$

where sp stands for “statistical physics,” since  $W_{\text{sp}}(\omega)$  can be computed by statistical physics methods and com stands for “combinatorics,” as  $W_{\text{com}}(\omega)$  can easily be computed by combinatorial methods. From Jensen's inequality we know that  $W_{\text{sp}}(\omega) \leq W_{\text{com}}(\omega)$ . It has been shown in [2], [3] that for regular LDPC ensembles  $W_{\text{sp}}(\omega) = W_{\text{com}}(\omega)$ . However for irregular LDPC ensembles this is not the case [6]. The equality between  $W_{\text{com}}(\omega)$  and  $W_{\text{sp}}(\omega)$  for regular ensembles suggests that a randomly chosen code should have  $N(G, n\omega)$  “close” to  $\mathbb{E}[N(G, n\omega)]$  with high probability. In this paper we obtain an asymptotic lower bound on this probability using the second moment method by estimating the variance of  $N(G, n\omega)$ . However, to estimate the variance we need to verify that the solution set of a certain

system of polynomial equations satisfies some properties (see Lemma 3.4 for details). Assuming that these properties are satisfied, we show that for a regular LDPC ensemble with left degree 1 and right degree  $r$ , any  $\epsilon > 0$  and for all  $\omega$  such that  $W_{\text{com}}(\omega)$  is positive

$$\lim_{n \rightarrow \infty} \mathbb{P} \left( 1 - \epsilon \leq \frac{N(G, n\omega)}{\mathbb{E}[N(G, n\omega)]} \leq 1 + \epsilon \right) \geq 1 - \frac{\delta(\omega, 1, r)}{\epsilon^2} \quad (1)$$

where  $\delta(\omega, 1, r)$  is a function of  $\omega$  and can be evaluated by solving a polynomial equation. If  $W_{\text{com}}(\omega) \leq 0$ , then  $\mathbb{E}[N(G, n\omega)] = o(1)$  (Lemma 3.2) and by Markov's inequality we have

$$\lim_{n \rightarrow \infty} \mathbb{P}(N(G, n\omega) = 0) = 1.$$

Clearly for  $\omega$  such that  $W_{\text{com}}(\omega) \leq 0$ , the convergence of  $N(G, n\omega)$  to  $\mathbb{E}[N(G, n\omega)]$  follows trivially. Hence, in the rest of the paper, our focus will be on the weight  $\omega$  such that  $W_{\text{com}}(\omega) > 0$ .

Note that the convergence implied by the bound in (1) is pointwise. For a fixed  $\omega$ , (1) implies that asymptotically at least a fraction  $1 - \frac{\delta(\omega, 1, r)}{\epsilon^2}$  of codes in the ensemble have their weight distribution function in a window of width  $\epsilon$  around the ensemble average. In Fig. 1, we plot the bound in (1) for regular codes with  $\frac{1}{r} = 0.75$  and  $0.5$ . We observe that if we fix the ratio  $\frac{1}{r}$  and let  $1, r$  increase then the bound converges to 1. This implies that for large left and right degrees, almost all the codes in the ensemble have their weight distribution very close to the ensemble average. Note that in this case it is well known that the weight distribution converges to the weight distribution of Shannon's random ensemble [7].

Another important property of LDPC codes is the stopping set distribution. Stopping sets determine the performance of LDPC codes under iterative decoding over erasure channel. The bound obtained in (1) can be easily extended to stopping set distribution. This is because of the fact that the method of determining the moments in both the cases is same. Hence we focus on the weight distribution and in the end we briefly describe the computation for stopping set distribution. In Fig. 2, we plot the bound in (1) for stopping set distribution. Again we observe that if we fix the ratio  $\frac{1}{r}$  and let  $1, r$  increase then the bound converges to 1.

The paper is organized in the following way. A brief introduction to LDPC codes and second moment method is given in Section II. In Section III, we use the second moment method to prove the bound in (1) for weight distribution. We apply the second moment method to stopping set distribution in Section IV. A discussion in Section V concludes the paper.

### II. PRELIMINARIES

#### A. LDPC Ensembles

LDPC codes, originally invented by Gallager [8], are usually defined in terms of ensembles of *bipartite graphs*. A graph consists of a set of *variable* nodes and a set of *check* nodes, together with edges connecting both sets giving rise to a code of block length  $n$  in the following way: a vector  $(x_1, \dots, x_n) \in \text{GF}(2)^n$  is a codeword if and only if for each check node the sum (modulo 2) of the values of its adjacent variable nodes is zero. The coordinates of a codeword are indexed by the variable nodes  $1, \dots, n$ . A stopping set is a subset of the set of variable nodes such that its neighboring check nodes are connected to it at least twice.

An ensemble of bipartite graphs can be defined in terms of a pair of *degree distributions*. A degree distribution is a real valued polynomial with nonnegative coefficients and it evaluates to unity at unity. Associated with the ensemble is a degree distribution pair  $(\lambda(x) = \sum_i \lambda_i x^{i-1}, \rho(x) = \sum_j \rho_j x^{j-1})$ , shorthand  $(\lambda, \rho)$ , where

Manuscript received July 17, 2004; revised July 10, 2005. The material in this correspondence was presented in part at International Symposium on Information Theory, Adelaide, Australia, 2005.

The author is with Communication Theory Laboratory (LTHC), EPFL-Swiss Federal Institute of Technology, Lausanne, CH-1015, Switzerland (e-mail: vishwambhar.rathi@epfl.ch).

Communicated by A. Ashikhmin, Associate Editor for Coding Theory.  
Digital Object Identifier 10.1109/TIT.2006.880065

<sup>1</sup>Here and in what follows it is understood that  $\omega$  is such that  $n\omega$  is an integer.

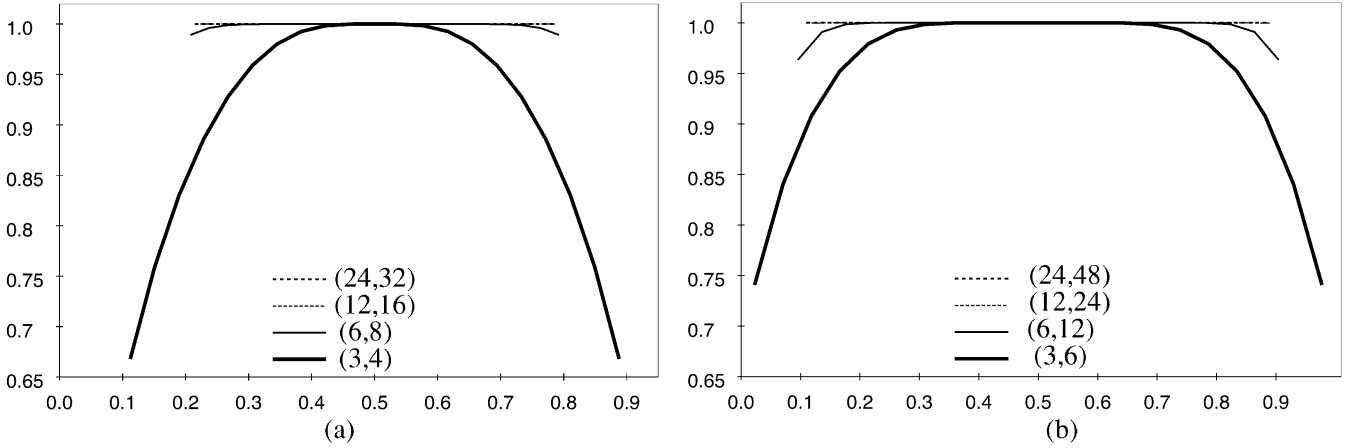


Fig. 1. The x-axis is the relative weight  $\omega$  such that  $W_{\text{com}}(\omega) > 0$  and y-axis is the bound  $1 - \frac{\delta(\omega, \frac{1}{r})}{\epsilon^2}$  with  $\epsilon = 0.95$ : (a) for ensembles with rate= 0.25 and (b) for ensembles with rate= 0.5.

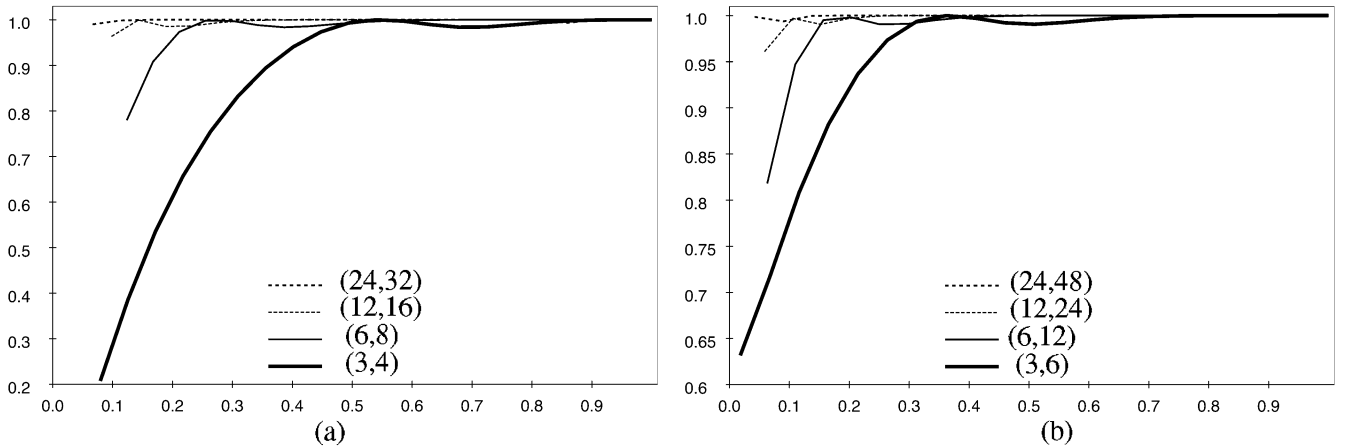


Fig. 2. The x-axis is the relative stopping set size  $s$  and y-axis is the bound  $1 - \frac{\delta(s, \frac{1}{r})}{\epsilon^2}$  with  $\epsilon = 0.95$ : (a) for ensembles with rate= 0.25 and (b) for ensembles with rate= 0.5.

$\lambda_i(\rho_j)$  denotes the fraction of the total number of edges connected to a variable (check) node of degree  $i(j)$ . Given a pair  $(\lambda, \rho)$  of degree distributions and the block length  $n$ , an *ensemble* of bipartite graphs  $\mathbb{G}(n, \lambda, \rho)$  is defined by running over all possible permutations of edges connecting variable and check nodes according to  $\lambda$  and  $\rho$ , respectively. For a  $(1, r)$ -regular code ensemble  $\mathbb{G}(n, 1, r)$  we have:  $\lambda(x) = x^{1-1}$ ,  $\rho(x) = x^{r-1}$ . Let  $G$  be a graph chosen at random from  $\mathbb{G}(n, 1, r)$ . Let  $N(G, n\omega)$  be the weight distribution function denoting the number of codewords of weight  $n\omega$  in  $G$ , where  $\omega = W/n$  is the normalized weight with  $W$  denoting the weight. Let  $\sigma^2(G, n\omega)$  denote the variance of  $N(G, n\omega)$  over the ensemble  $\mathbb{G}(n, 1, r)$ ,  $\sigma^2(G, n\omega) = \mathbb{E}[N(G, n\omega)^2] - \mathbb{E}[N(G, n\omega)]^2$ . Similarly let  $S(G, ns)$  denote the number of stopping sets of size  $ns$  and let  $\sigma^2(G, ns)$  denote the variance of  $S(G, ns)$  over the ensemble  $\mathbb{G}(n, 1, r)$ . The *support set* of a word is the set of its non zero bits. The *overlap* between two words is the intersection of their support sets. We denote a vector  $(x_1, x_2, x_3)$  by  $\underline{x}$ , the transpose of  $\underline{x}$  by  $\underline{x}^T$ , the dot product between  $\underline{x}$  and  $\underline{y}$  is denoted by  $\underline{x} \cdot \underline{y}^T$ ,  $\underline{x} \underline{y}$  denotes the component wise multiplication, i.e., the vector  $(x_1 y_1, x_2 y_2, x_3 y_3)$ . We use the notation that a vector to the power of a vector and also a scalar to the power of a vector is a vector, i.e.,  $\underline{x}^{\underline{k}} := (x_1^{k_1}, x_2^{k_2}, x_3^{k_3})$  and  $e^{\underline{x}} := (e^{x_1}, e^{x_2}, e^{x_3})$ . Finally,  $x^+ := \max(x, 0)$  and  $f'(t)$  denotes the derivative of the function  $f(x)$  evaluated at  $t$ .

### B. Second Moment Method

Let  $\{X_n\}$  be a sequence of random variables indexed by  $n$ ,  $n \in \mathbb{N}$ . Let  $\sigma_n^2 = \mathbb{E}[(X_n - \mathbb{E}[X_n])^2]$  be the variance of  $X_n$ . Then by Chebyshev's inequality we have for any  $a \geq 0$

$$\mathbb{P}(|X_n - \mathbb{E}[X_n]| \geq a) \leq \frac{\sigma_n^2}{a^2}.$$

If we choose  $a = \epsilon \mathbb{E}[X_n]$  and if  $\lim_{n \rightarrow \infty} \frac{\sigma_n^2}{\mathbb{E}[X_n]^2} = \delta$ , then we can draw the conclusion that

$$\lim_{n \rightarrow \infty} \mathbb{P}\left(1 - \epsilon \leq \frac{X_n}{\mathbb{E}[X_n]} \leq 1 + \epsilon\right) \geq 1 - \frac{\delta}{\epsilon^2}.$$

In order to apply this bound to  $N(G, n\omega)$ , we need to compute the ratio  $\lim_{n \rightarrow \infty} \frac{\sigma^2(G, n\omega)}{\mathbb{E}[N(G, n\omega)]^2} = \lim_{n \rightarrow \infty} \frac{\mathbb{E}[N^2(G, n\omega)]}{\mathbb{E}[N(G, n\omega)]^2} - 1$ .

### III. MOMENT CALCULATIONS FOR WEIGHT DISTRIBUTION

We start with the first moment. As shown in [9]–[11],

$$\mathbb{E}[N(G, n\omega)] = \frac{\binom{n}{n\omega}}{\binom{n}{n1}} \text{Coeff}\left(p(x)^{\frac{1}{r}}, x^{n1\omega}\right) \quad (2)$$

where  $\text{Coeff}(p(x)^{\frac{n-1}{2}}, x^{n-1\omega})$  denotes the coefficient of  $x^{n-1\omega}$  in the Taylor series expansion of  $p(x)^{\frac{n-1}{2}}$  and  $p(x) = ((1+x)^{\mathbf{r}} + (1-x)^{\mathbf{r}})/2$ . We note that  $p(x)$  has only even powers of  $x$ . To remove this periodicity of powers, we define the polynomial  $q(y) = p(x)$ , where  $y = x^2$ . Now,  $\text{Coeff}(p(x)^{\frac{n-1}{2}}, x^{n-1\omega}) = \text{Coeff}(q(y)^{\frac{n-1}{2}}, y^{\frac{n-1\omega}{2}})$ . In the next lemma, we recall the Hayman method to approximate  $\text{Coeff}(q(y)^{\frac{n-1}{2}}, y^{\frac{n-1\omega}{2}})$  for large values of  $n$ , a proof of which can be found in [12], [13].

**Lemma 3.1:** [Hayman Method] Let  $q(y) = \sum_i q_i y^i$  be a polynomial with non negative coefficients such that  $q_0 \neq 0$  and  $q_1 \neq 0$ . Define  $a_q(y) := y^{\frac{dq(y)}{dy} \frac{1}{q(y)}}$  and  $b_q(y) := y^{\frac{d^2 a_q(y)}{dy^2}}$ . Then for  $n$  tending to infinity so that  $\frac{n-1\omega}{2} \in \mathbb{N}$

$$\text{Coeff}\left(q(y)^{\frac{n-1}{2}}, y^{\frac{n-1\omega}{2}}\right) = \frac{q(y_\omega)^{\frac{n-1}{2}}}{(y_\omega)^{\frac{n-1\omega}{2}} \sqrt{2\pi \frac{n-1}{2} b_q(y_\omega)}} (1 + o(1)) \quad (3)$$

where the term  $o(1)$  converges to zero and  $y_\omega$  is the unique positive solution of  $a_q(y) = \frac{\mathbf{r}\omega}{2}$ .

Since  $q(y) = p(x)$  and  $y = x^2$ , we have  $a_q(y) = a_p(x)/2$ , where  $a_p(x) = x \frac{dp(x)}{dx} \frac{1}{p(x)}$ . Similarly,  $b_q(y) = b_p(x)/4$ , where  $b_p(x) = x \frac{d^2 a_p(x)}{dx^2}$ . Also  $y_\omega = x_\omega^2$ , where  $x_\omega$  is the unique positive solution of  $a_p(x) = \mathbf{r}\omega$  which simplifies to

$$x \frac{(1+x)^{\mathbf{r}-1} - (1-x)^{\mathbf{r}-1}}{(1+x)^{\mathbf{r}} + (1-x)^{\mathbf{r}}} = \omega. \quad (4)$$

Thus by substituting these relationships in Lemma 3.1, we get

$$\text{Coeff}\left(p(x)^{\frac{n-1}{2}}, x^{n-1\omega}\right) = \frac{2p(x_\omega)^{\frac{n-1}{2}}}{(x_\omega)^{n-1\omega} \sqrt{2\pi \frac{n-1}{2} b_p(x_\omega)}} (1 + o(1)). \quad (5)$$

We summarize our results thus far.

**Lemma 3.2:** [Ensemble Average of Weight Distribution] Consider the regular LDPC ensemble  $\mathbb{G}(n, 1, \mathbf{r})$ . Then for  $\omega \in (0, 1)$  such that  $1n\omega \in 2\mathbb{N}$ ,

$$\begin{aligned} \mathbb{E}[N(G, n\omega)] &= \frac{2\sqrt{\mathbf{r}} e^{n\left(\frac{1}{2}\ln(p(x_\omega)) - (1-\omega)h(\omega) - \ln(x_\omega)\right)}}{\sqrt{2\pi n b_p(x_\omega)}} (1 + o(1)) \quad (6) \end{aligned}$$

where  $h(\omega) = -(\omega \ln \omega + (1-\omega) \ln(1-\omega))$ ,  $\ln \omega$  is the natural logarithm of  $\omega$  and  $x_\omega$  is the unique positive solution of equation (4). If  $n-1\omega$  is odd, then  $\mathbb{E}[N(G, n\omega)] = 0$ .

*Proof:* We note that  $n-1\omega$  must be even, otherwise  $\mathbb{E}[N(G, n\omega)] = 0$  as  $\text{Coeff}(p(x)^{\frac{n-1}{2}}, x^{n-1\omega}) = 0$  in (2). When  $n-1\omega$  is even, using Stirling's approximation we get:

$$\binom{n}{n-1\omega} = \frac{e^{nh(\omega)}}{\sqrt{2\pi n \omega (1-\omega)}} (1 + o(1)).$$

By substituting this and (5) in (2), we get the desired result.  $\square$

To compute the second moment, we note that

$$\mathbb{E}[N^2(G, n\omega)] = \mathbb{E}\left[\sum_{w, w'} I_{w, w'}(G, n\omega)\right]$$

where  $w, w'$  are both words of length  $n$  and weight  $n\omega$  and

$$I_{w, w'}(G, n\omega) = \begin{cases} 1, & \text{if } w, w' \text{ are codewords of } G \\ 0, & \text{otherwise.} \end{cases}$$

By definition of the ensemble, the expectation  $\mathbb{E}[I_{w, w'}(G, n\omega)]$  does not depend on the specific choice of the pair  $w, w'$  but only on the cardinality of the overlap between the support sets of  $w$  and  $w'$ . In particular we can fix  $w$  to be a codeword of weight  $n\omega$  with support set  $\mathcal{W} = \{1, 2, \dots, n\omega\}$ , so that

$$\mathbb{E}[N^2(G, n\omega)] = \binom{n}{n\omega} \sum_{w'} \mathbb{E}[I_{w, w'}(G, n\omega)]$$

where we have dropped the subscript  $w$  as  $w$  is fixed. We can also fix  $w'$  to  $w'(i)$  for a given cardinality of overlap  $i$  with  $w$ .  $w'(i)$  has support set  $\mathcal{W}' = \{1, 2, \dots, i, n\omega + 1, \dots, 2n\omega - i\}$ . Then

$$\mathbb{E}[N^2(G, n\omega)] = \binom{n}{n\omega} \sum_{i=0}^{n\omega} \binom{n\omega}{i} \binom{n-n\omega}{n\omega-i} \mathbb{E}[I_{w'(i)}(G, n\omega)].$$

The binomials inside the summation correspond to the number of words having cardinality of overlap with  $w$  equals to  $i$ . To calculate  $\mathbb{E}[I_{w'(i)}(G, n\omega)]$ , we note that there are 3 different types of edges taking value 1. These types are: edges connected to  $\mathcal{W} \cap \mathcal{W}'$ , edges connected to  $\mathcal{W} \setminus (\mathcal{W} \cap \mathcal{W}')$  and finally, edges connected to  $\mathcal{W}' \setminus (\mathcal{W} \cap \mathcal{W}')$ . A placement of edges is *valid* if each check node is connected to an even number of edges from  $\mathcal{W}$  as well as from  $\mathcal{W}'$ , i.e., if the number of edges from each of the 3 different classes are *all even* or *all odd*. A moment's thought shows that the generating function for the number of valid placement is given by  $f(x_1, x_2, x_3)^{\frac{n-1}{2}} = f(\underline{x})^{\frac{n-1}{2}}$ , where  $x_1$  corresponds to the number of edges connected to  $\mathcal{W} \setminus (\mathcal{W} \cap \mathcal{W}')$ ,  $x_2$  corresponds to the number of edges connected to  $\mathcal{W} \cap \mathcal{W}'$  and  $x_3$  corresponds to the number of edges connected to  $\mathcal{W}' \setminus (\mathcal{W} \cap \mathcal{W}')$ , and where  $f(\underline{x})$  is the summation of the terms in the expansion of  $(1+x_1+x_2+x_3)^{\mathbf{r}}$  which have powers of  $x_1, x_2$  and  $x_3$  either all even or all odd. Explicitly

$$\begin{aligned} f(\underline{x}) &= \frac{1}{4} ((1+x_1+x_2+x_3)^{\mathbf{r}} \\ &\quad + (1+x_1-x_2-x_3)^{\mathbf{r}} \\ &\quad + (1-x_1+x_2-x_3)^{\mathbf{r}} \\ &\quad + (1-x_1-x_2+x_3)^{\mathbf{r}}). \quad (7) \end{aligned}$$

Since there are  $1(n\omega - i)$  edges connected to  $\mathcal{W} \setminus (\mathcal{W} \cap \mathcal{W}')$ ,  $1i$  edges connected to  $\mathcal{W} \cap \mathcal{W}'$  and  $1(n\omega - i)$  edges connected to  $\mathcal{W}' \setminus (\mathcal{W} \cap \mathcal{W}')$ , we have

$$\begin{aligned} \mathbb{E}[I_{w'(i)}(G, n\omega)] &= \text{Coeff}\left(f(\underline{x})^{\frac{n-1}{2}}, x_1^{1(n\omega-i)} x_2^{1i} x_3^{1(n\omega-i)}\right) \\ &\quad \cdot \frac{1}{(n-1)!} ((1(n\omega - i))!)^2 (1i)! (n-1-2n\omega + 1i)!. \end{aligned}$$

As all the edges are labeled, the factor  $(n-1)!$  corresponds to the total number of graphs in the ensemble  $\mathbb{G}(n, 1, \mathbf{r})$ . The term  $(1(n\omega - i))!$  corresponds to interchanging the positions of edges connected to  $\mathcal{W} \setminus (\mathcal{W} \cap \mathcal{W}')$ , as well as to  $\mathcal{W}' \setminus (\mathcal{W} \cap \mathcal{W}')$ ,  $(1i)!$  corresponds to interchanging the positions of edges connected to  $\mathcal{W} \cap \mathcal{W}'$ , and  $(1(n - 2n\omega + i))!$  corresponds to interchanging of the positions of edges taking value 0. Hence, see the first equation at the bottom of the following page. Let  $S_i$  be the  $i^{\text{th}}$  summation term in (8), so  $S_i = F_i C_i$ . Note that  $S_i = 0$  for  $i < (2n\omega - n)^+$  as there can not exist two words of length  $n$  and weight  $n\omega$  such that the cardinality

of their overlap is less than  $(2n\omega - n)^+$ . A property of the term  $S_{n\omega}$  that we will need later is

$$S_{n\omega} = \mathbb{E}[N(G, n\omega)]. \quad (8)$$

This simply follows from the fact that for  $i = n\omega$ , the words  $w$  and  $w'(i)$  are identical. Now to get a closed-form expression for  $\mathbb{E}[N^2(G, n\omega)]$ , we use Stirling's formula to approximate the factorial terms and to approximate the Coeff function we use the following multidimensional extension of Lemma 3.1 as given in [14, Th. 2].

**Lemma 3.3:** [Multidimensional Saddle Point Method] Let  $\underline{i} := (1(n\omega - i), 1i, 1(n\omega - i))$ ,  $\underline{j} := (1(n\omega - j), 1j, 1(n\omega - j))$  and  $0 < \lim_{n \rightarrow \infty} \frac{i}{n} < \omega$ ,  $f(\underline{x})$  be as defined in (7) and  $\underline{t} = (t_1, t_2, t_3)$  be a positive solution of  $a_f(\underline{x}) = \frac{\underline{r}\underline{i}}{n\underline{1}}$ , where  $a_f(\underline{x}) = (\frac{x_i \partial f}{\partial x_i})_{i=1}^3$ .

Then  $\text{Coeff}(f(\underline{x})^{\frac{n\underline{1}}{\underline{r}}}, \underline{x}^{\underline{i}})$  can be approximated using the saddle point method for multivariate polynomials

$$\text{Coeff}\left(f(\underline{x})^{\frac{n\underline{1}}{\underline{r}}}, \underline{x}^{\underline{i}}\right) = \frac{4f(\underline{t})^{\frac{n\underline{1}}{\underline{r}}}}{(\underline{t})^{\underline{i}} \sqrt{\left(2\pi \frac{n\underline{1}}{\underline{r}}\right)^3 |B_f(\underline{t})|}} (1 + o(1))$$

where  $B_f(\underline{x})$  is a  $3 \times 3$  matrix whose elements are given by  $B_{f(i,j)} = x_j \frac{\partial a_i}{\partial x_j} = B_{f(j,i)}$ . Also,  $\text{Coeff}(f(\underline{x})^{\frac{n\underline{1}}{\underline{r}}}, \underline{x}^{\underline{j}})$  can be approximated in terms of  $\text{Coeff}(f(\underline{x})^{\frac{n\underline{1}}{\underline{r}}}, \underline{x}^{\underline{i}})$ . This approximation is called the *local limit theorem* of  $\underline{j}$  around  $\underline{i}$ . Explicitly, if  $\underline{u} := \sqrt{\frac{\underline{r}}{n\underline{1}}}(\underline{j} - \underline{i})$  and  $\|\underline{u}\| = O((\ln n)^{\frac{1}{3}})$ , then

$$\text{Coeff}\left(f(\underline{x})^{\frac{n\underline{1}}{\underline{r}}}, \underline{x}^{\underline{j}}\right) = \underline{t}^{\underline{i}-\underline{j}} \text{Coeff}\left(f(\underline{x})^{\frac{n\underline{1}}{\underline{r}}}, \underline{x}^{\underline{i}}\right) \cdot \exp\left(-\frac{1}{2}\underline{u} \cdot B_f(\underline{t})^{-1} \cdot \underline{u}^T\right) (1 + o(1)).$$

*Proof:* The proof of the lemma follows from a modification of the proof of [14, Th. 2]. This is rather tedious and is therefore relegated to the appendix  $\square$ .

The system of equations corresponding to  $a(\underline{x}) = \frac{\underline{r}\underline{i}}{n\underline{1}}$  is symmetric in  $x_1$  and  $x_3$ . Hence a positive solution  $\underline{x}$  of this system of equations satisfies  $x_1 = x_3$  and the system reduces to the following:

$$x_1 \frac{(1 + 2x_1 + x_2)^{\underline{r}-1} - (1 - 2x_1 + x_2)^{\underline{r}-1}}{(1 + 2x_1 + x_2)^{\underline{r}} + 2(1 - x_2)^{\underline{r}} + (1 - 2x_1 + x_2)^{\underline{r}}} = \omega - \alpha \quad (9)$$

$$x_2 \frac{(1 + 2x_1 + x_2)^{\underline{r}-1} - 2(1 - x_2)^{\underline{r}-1} + (1 - 2x_1 + x_2)^{\underline{r}-1}}{(1 + 2x_1 + x_2)^{\underline{r}} + 2(1 - x_2)^{\underline{r}} + (1 - 2x_1 + x_2)^{\underline{r}}} = \alpha \quad (10)$$

where  $\alpha = \frac{i}{n}$ .

In order to evaluate the second moment, we need to find the dominant terms of the summation in (8). To find all the dominant terms, let the term corresponding to  $i = i_m$ , i.e.,  $S_{i_m} = F_{i_m} C_{i_m}$  be a local maximum of  $\{S_i\}_{i=0}^{n\omega}$ . We first check if the end terms  $S_{(2n\omega - n)^+}$  and  $S_{n\omega}$  can be dominant. The Assumption 2 of the Lemma 3.4 eliminates the possibility that  $S_{(2n\omega - n)^+}$  is a dominant term. In the proof of Lemma 3.4 we will see that  $\ln(S_{n\omega^2})/n = 2W_{\text{com}}(\omega)$ . This with (8) implies that  $S_{n\omega}$  is not a dominant term. So we consider  $i_m$  such that  $0 < \lim_{n \rightarrow \infty} \frac{i_m}{n} < \omega$ . Let  $\Delta = i - i_m$  and  $\alpha_m = \frac{i_m}{n}$ . We expand  $F_i$  and  $C_i$  for  $\Delta \in (-\sqrt{n}(\ln n)^{\frac{1}{3}}, \sqrt{n}(\ln n)^{\frac{1}{3}})$  in terms of  $F_{i_m}$  and  $C_{i_m}$  using Stirling's approximation and the local limit theorem of Lemma 3.3, respectively. Then

$$\begin{aligned} F_i &= F_{i_m} \exp\left(\Delta(1-1)\ln\left(\frac{i_m(n-2n\omega+i_m)}{(n\omega-i_m)^2}\right)\right) \\ &\quad \cdot \exp\left(\Delta^2\left(\frac{1-1}{n\omega-i_m} + \frac{1-1}{2i_m} + \frac{1-1}{2(n-2n\omega+i_m)}\right)\right) \\ &\quad \cdot \left(1 + O\left(\frac{\Delta^3}{n^2}\right)\right) \\ C_i &= C_{i_m} \exp\left(\Delta \ln\left(\frac{t_1^2}{t_2}\right) - \frac{\Delta^2}{2n\sigma_c^2(\alpha_m)}\right) (1 + o(1)) \end{aligned}$$

where (11) is shown at the bottom of the page. Hence

$$\begin{aligned} \frac{S_i}{S_{i_m}} &= \exp\left(\Delta\left((1-1)\ln\left(\frac{i_m(n-2n\omega+i_m)}{(n\omega-i_m)^2}\right)\right.\right. \\ &\quad \left.\left.+ \ln\left(\frac{t_1^2}{t_2}\right)\right)\right) \\ &\quad \cdot \exp\left(\Delta^2\left(\frac{1-1}{n\omega-i_m} + \frac{1-1}{2i_m} + \frac{1-1}{2(n-2n\omega+i_m)} - \frac{1}{2n\sigma_c^2(\alpha_m)}\right)\right) \\ &\quad \times (1 + o(1)) \end{aligned} \quad (12)$$

We know that there is a local maximum at  $\Delta = 0$ , hence the coefficient of  $\Delta$  in (12) will vanish. This gives an additional equation governing  $\alpha_m$

$$\left(\frac{\alpha_m(1-2\omega+\alpha_m)}{(\omega-\alpha_m)^2}\right)^{1-1} = \left(\frac{t_2}{t_1^2}\right)^1. \quad (13)$$

We solve (9), (10), and (13) and find all the solutions such that  $0 < \alpha_m < \omega$ ,  $t_1 > 0$ ,  $t_2 > 0$  and the coefficient of  $\Delta^2$  in (12) is negative (this ensures that  $S_{i_m}$  is a local maximum). One of the possible solution to this system of polynomial equations is  $\alpha_m = \omega^2$ . This is because  $\{C_i\}_{i=0}^{n\omega}$  and  $\{F_i\}_{i=0}^{n\omega}$  are concave and convex sequences respectively, both achieving their extreme values at  $i = n\omega^2$ . Hence  $\{S_i\}_{i=0}^{n\omega}$  also achieves an extreme value at  $i = n\omega^2$ . If  $\alpha_m = \omega^2$  is a unique global maximum in the solution set of (9), (10), and (13), then we can get a

$$\mathbb{E}[N^2(G, n\omega)] = \sum_{i=0}^{n\omega} \underbrace{\text{Coeff}\left(f(\underline{x})^{\frac{n\underline{1}}{\underline{r}}}, x_1^{1(n\omega-i)} x_2^1 x_3^{1(n\omega-i)}\right)}_{C_i} \cdot \underbrace{\frac{\binom{n}{n\omega}}{(n\underline{1})!} \binom{n\omega}{i} \binom{n-n\omega}{n\omega-i} ((1(n\omega-i))! (1i)! (1(n-2n\omega+i))!)}_{F_i}.$$

$$\begin{aligned} F_{i_m} &= \left(\frac{(n\omega-i_m)^{2(n\omega-i_m)} i_m^{i_m} (n-2n\omega+i_m)^{n-2n\omega+i_m}}{n^n}\right)^{1-1} \frac{1}{\sqrt{1}} (1 + o(1)) \\ \sigma_c^2(\alpha_m) &= \frac{1}{\mathbf{1r}((-1, 1, -1) \cdot B_f(\underline{t})^{-1} \cdot (-1, 1, -1)^T)}. \end{aligned} \quad (11)$$

closed form expression for second moment. We summarize this in the following lemma.

**Lemma 3.4:** [Second Moment Method] Consider the regular LDPC ensemble  $\mathbb{G}(n, 1, \mathbf{r})$ . Then for  $\omega \in (0, 1)$ , if  $W_{\text{com}}(\omega) > 0$  and if the following conditions are satisfied:

- 1)  $\alpha_m = \omega^2$  is the only solution of (9), (10), and (13) for which coefficient of  $\Delta^2$  in (12) is negative;
- 2)  $\lim_{n \rightarrow \infty} \frac{\ln(S_{n\omega^2})}{n} > \frac{\ln(S_{(2n\omega-n)})}{n}$ , then by the second moment method we have

$$\lim_{n \rightarrow \infty} \mathbb{P} \left( 1 - \epsilon \leq \frac{N(G, n\omega)}{\mathbb{E}[N(G, n\omega)]} \leq 1 + \epsilon \right) \geq 1 - \frac{\delta(\omega, \mathbf{1}, \mathbf{r})}{\epsilon^2}$$

where

$$\begin{aligned} \delta(\omega, \mathbf{1}, \mathbf{r}) &= \frac{b_p(x_\omega) \sqrt{\mathbf{r}} \omega (1 - \omega) \sigma_c(\omega^2)}{\sqrt{|B_f(x_\omega, x_\omega^2, x_\omega)| (\omega^2 (1 - \omega)^2 - (1 - 1) \sigma_c^2(\omega^2))}} \\ &\quad - 1 \\ &\quad \sigma_c^2(\omega^2) \\ &= \frac{1}{\mathbf{1r}((-1, 1, -1) \cdot B_f(x_\omega, x_\omega^2, x_\omega)^{-1} \cdot (-1, 1, -1)^T)} \end{aligned}$$

and  $x_\omega$  is the only positive solution of (4).

*Remark:* Note that the conditions of Lemma 3.4 are hard to verify in general but they are typically easy to verify for any given regular LDPC ensemble.

*Proof:* We observe that the solution  $t$  of (9) and (10) for  $\alpha = \omega^2$  satisfies  $t_2 = t_1^2$  and this system of equations reduces to a single equation which is identical to (4), the equation we need to solve to find  $\mathbb{E}[N(G, n\omega)]$ . Thus  $t_1 = x_\omega$ . By (12) and noting that the terms  $S_{n\omega^2 + \Delta}$  for  $\Delta \notin (-\sqrt{n}(\ln n)^{\frac{1}{3}}, \sqrt{n}(\ln n)^{\frac{1}{3}})$  are much smaller than  $S_{n\omega^2}$ , we get

$$\begin{aligned} \mathbb{E}[N^2(G, n\omega)] &= S_{n\omega^2} \sum_{\Delta = -\sqrt{n}(\ln n)^{\frac{1}{3}}}^{\sqrt{n}(\ln n)^{\frac{1}{3}}} \exp\left(\frac{-\Delta^2}{2\sigma_s^2}\right) (1 + o(1)) \\ &= S_{n\omega^2} \int_{-\infty}^{\infty} \exp\left(\frac{-x^2}{2\sigma_s^2}\right) dx (1 + o(1)) \\ &= S_{n\omega^2} \sqrt{2\pi\sigma_s^2} (1 + o(1)) \end{aligned}$$

where

$$\frac{1}{\sigma_s^2} = \frac{1}{n\sigma_c^2(\omega^2)} - \frac{1 - 1}{n\omega^2(1 - \omega)^2}.$$

Also  $f(x_\omega, x_\omega^2, x_\omega) = p(x_\omega)^2$ . To evaluate  $S_{n\omega^2}$ , we use Lemma 3.3 and Stirling's approximation for factorial terms. This gives

$$\begin{aligned} \mathbb{E}[N^2(G, n\omega)] &= \frac{4\sigma_c(\omega^2) \mathbf{r} \sqrt{\mathbf{r}} \omega (1 - \omega)}{2\pi n \sqrt{(\omega^2 (1 - \omega)^2 - (1 - 1) \sigma_c^2(\omega^2))}} \\ &\quad \cdot \frac{e^{2n \left( \frac{1}{\mathbf{r}} \ln(p(x_\omega)) - (1 - 1)h(\omega) - 1\omega \ln(x_\omega) \right)}}{\sqrt{|B_f(x_\omega, x_\omega^2, x_\omega)|}} (1 + o(1)). \end{aligned}$$

We need the condition  $W_{\text{com}}(\omega) > 0$ , as  $\lim_{n \rightarrow \infty} \frac{\ln(S_{n\omega^2})}{n} = 2W_{\text{com}}(\omega)$  and  $\lim_{n \rightarrow \infty} \frac{\ln(S_{n\omega})}{n} = W_{\text{com}}(\omega)$ . Clearly, when  $W_{\text{com}}(\omega)$  is negative,  $S_{n\omega^2}$  cannot be a global maximum. Now using Lemma 3.2, the second moment method gives us

$$\lim_{n \rightarrow \infty} \mathbb{P} \left( 1 - \epsilon \leq \frac{N(G, n\omega)}{\mathbb{E}[N(G, n\omega)]} \leq 1 + \epsilon \right) \geq 1 - \frac{\delta(\omega, \mathbf{1}, \mathbf{r})}{\epsilon^2}.$$

This proves the lemma.  $\square$

The bound obtained in Lemma 3.4 can in general only be evaluated numerically except for the cases when (4) can be solved analytically. For example, for the (3, 4)-regular code we get

$$\begin{aligned} \delta(\omega, 3, 4) &= \frac{8\omega(1 - \omega)(3 - \Omega)}{\sqrt{(81 - 27\Omega + 16\omega(1 - \omega)(8\omega - 8\omega^2 - 18 + 3\Omega))}} \\ &\quad \times \frac{1}{\sqrt{-21 + 80\omega(1 - \omega) + 9\Omega}} \end{aligned}$$

where  $\Omega = \sqrt{9 - 32\omega + 32\omega^2}$ .

#### IV. MOMENT CALCULATIONS FOR STOPPING SET DISTRIBUTION

As shown in [15], the first moment of stopping set distribution is given by

$$\mathbb{E}[S(G, ns)] = \frac{\binom{n}{ns}}{\binom{n\mathbf{1}}{n\mathbf{1}s}} \text{Coeff} \left( (\beta(x))^{\frac{1}{\mathbf{r}}}, x^{n\mathbf{1}s} \right)$$

where  $\beta(x) = (1 + x)^{\mathbf{r}} - \mathbf{r}x$ . Applying the Lemma 3.1 and using Stirling's approximation, we get

$$\mathbb{E}[S(G, ns)] = \frac{\sqrt{\mathbf{r}} e^{n \left( \frac{1}{\mathbf{r}} \ln(\beta(x_s)) - (1 - 1)h(s) - 1s \ln(x_s) \right)}}{\sqrt{2\pi n b_\beta(x_s)}} (1 + o(1))$$

where  $x_s$  is the only positive solution of

$$x \frac{(1 + x)^{\mathbf{r} - 1} - 1}{(1 + x)^{\mathbf{r}} - \mathbf{r}x} = s. \quad (14)$$

The second moment is  $\mathbb{E}[S^2(G, ns)] = \mathbb{E}[\sum_{\mathbf{s}, \mathbf{s}'} I_{\mathbf{s}, \mathbf{s}'}(G, ns)]$ , where  $\mathbf{s}, \mathbf{s}'$  are both stopping sets of cardinality  $ns$ . By definition of the ensemble, the expectation  $\mathbb{E}[I_{\mathbf{s}, \mathbf{s}'}(G, ns)]$  depends only on the cardinality of the overlap between  $\mathbf{s}$  and  $\mathbf{s}'$ . Hence like in the previous section for weight distribution we can fix  $\mathbf{s}$  to be equal to  $\mathbf{s} = \{1, 2, \dots, ns\}$  and for a given cardinality of overlap  $i$  we can fix  $\mathbf{s}'$  to be equal to  $\mathbf{s}'(i) = \{1, 2, \dots, i, ns + 1, \dots, 2ns - i\}$ . As  $\mathbf{s}$  is fixed, we drop the subscript  $\mathbf{s}$  in  $I_{\mathbf{s}, \mathbf{s}'}(G, ns)$ . This gives

$$\mathbb{E}[S^2(G, ns)] = \binom{n}{ns} \sum_{i=0}^{ns} \binom{ns}{i} \binom{n - ns}{ns - i} \mathbb{E}[I_{\mathbf{s}'(i)}(G, ns)].$$

For  $I_{\mathbf{s}'(i)}(G, ns) = 1$ , we need that every check node in  $G$  is either not connected to  $\mathbf{s}$  or connected to  $\mathbf{s}$  by more than one edge. Similarly every check node is either not connected to  $\mathbf{s}'(i)$  or connected to  $\mathbf{s}'(i)$  by more than one edge. This implies

$$\begin{aligned} \mathbb{E}[I_{\mathbf{s}'(i)}(G, ns)] &= \text{Coeff} \left( g(\mathbf{x})^{\frac{1}{\mathbf{r}}}, x_1^{1(ns-i)} x_2^i x_3^{1(ns-i)} \right) \\ &\quad \cdot \frac{1}{(n\mathbf{1})!} ((1(ns - i))!)^2 (i!) (n\mathbf{1} - 2n\mathbf{1}s + i)! \end{aligned}$$

where

$$\begin{aligned} g(\mathbf{x}) &= (1 + x_1 + x_2 + x_3)^{\mathbf{r}} \\ &\quad - \mathbf{r}(1 + x_1)^{\mathbf{r} - 1} (x_2 + x_3) \\ &\quad - \mathbf{r}x_1 ((1 + x_3)^{\mathbf{r} - 1} - (\mathbf{r} - 1)x_3) \\ &\quad - \mathbf{r}x_2 ((1 + x_3)^{\mathbf{r} - 1} - 1). \end{aligned} \quad (15)$$

Thus, see (16) shown at the bottom of the following page. To evaluate Coeff in (16), we use [14, Th. 2]. Again applying the same line of ar-

TABLE I  
 $\lim_{\omega \rightarrow \omega_{\min}^+} 1 - \frac{\delta(\omega, \mathbf{1}, \mathbf{r})}{\epsilon^2}$  FOR RATE =  $\frac{1}{2}$  AND  $\epsilon = 0.95$

( $\mathbf{1}, \mathbf{r}$ )-code	$\omega_{\min}$	$1 - \frac{\delta(\omega_{\min}, \mathbf{1}, \mathbf{r})}{0.95^2}$
(3, 6)	0.0227334	0.740611
(6, 12)	0.0956337	0.963306
(12, 24)	0.109404	0.999617
(24, 48)	0.110026	$\sim 1.0$

TABLE II  
 $\lim_{\omega \rightarrow \omega_{\min}^+} 1 - \frac{\delta(\omega, \mathbf{1}, \mathbf{r})}{\epsilon^2}$  FOR RATE =  $\frac{1}{4}$  AND  $\epsilon = 0.95$

( $\mathbf{1}, \mathbf{r}$ )-code	$\omega_{\min}$	$1 - \frac{\delta(\omega_{\min}, \mathbf{1}, \mathbf{r})}{0.95^2}$
(3, 4)	0.112159	0.667889
(6, 8)	0.207437	0.989098
(12, 16)	0.214428	0.999994
(24, 32)	0.214502	$\sim 1.0$

guments as for the weight distribution and if the conditions of Lemma 3.4 are true in the setting of stopping set distribution, then we get

$$\begin{aligned} & \mathbb{E}[S^2(G, ns)] \\ &= \frac{\sigma_c(s^2) \mathbf{r} \sqrt{\mathbf{r} s (1-s)}}{2\pi n \sqrt{(s^2(1-s)^2 - (1-1)\sigma_c^2(s^2))}} \\ & \cdot \frac{e^{2n \left( \frac{1}{\mathbf{r}} \ln(\beta(x_s)) - (1-1)h(s) - 1s \ln(x_s) \right)}}{|B_g(x_s, x_s^2, x_s)|} (1 + o(1)) \end{aligned}$$

where  $B_g(\underline{x})$  is same as defined in Lemma 3.3 with respect to  $g(\underline{x})$  and  $x_s$  is the positive solution of (14). Hence by the second moment method we have

$$\lim_{n \rightarrow \infty} \mathbb{P} \left( 1 - \epsilon \leq \frac{S(G, ns)}{\mathbb{E}[N(G, ns)]} \leq 1 + \epsilon \right) \leq 1 + \epsilon \geq 1 - \frac{\delta(s, \mathbf{1}, \mathbf{r})}{\epsilon^2}$$

where

$$\begin{aligned} \delta(s, \mathbf{1}, \mathbf{r}) &= \frac{b_\beta(x_s) \sqrt{\mathbf{r} s (1-s)} \sigma_c(s^2)}{\sqrt{|B_g(x_s, x_s^2, x_s)| (s^2(1-s)^2 - (1-1)\sigma_c^2(s^2))}} - 1 \\ \sigma_c^2(s^2) &= \frac{1}{\mathbf{1r} |(-1, 1, -1) \cdot B_g(x_s, x_s^2, x_s)^{-1} \cdot (-1, 1, -1)^T|}. \end{aligned}$$

## V. DISCUSSION

Fix the relative weight  $\omega$ . If  $\epsilon \in (0, 1)$  then we conclude that asymptotically for at least a fraction  $1 - \frac{\delta(\omega, \mathbf{1}, \mathbf{r})}{\epsilon^2}$  of codes, the number of codewords  $N(G, n\omega)$  (for a fixed  $\omega$ ) is at most a constant factor away from the ensemble average.

Also from Fig. 1 we see that  $1 - \frac{\delta(\omega, \mathbf{1}, \mathbf{r})}{\epsilon^2}$  is an increasing function of  $\omega$  for  $\omega \in (\omega_{\min}, 0.5)$  and is a decreasing function for  $\omega > 0.5$ . It is equal to 1 for  $\omega = 0.5$ . This implies that asymptotically in almost all the codes there are  $\mathbb{E}[N(G, \frac{n}{2})] (1 \pm \epsilon)$  codewords of weight  $\frac{n}{2}$ . For  $\omega$  close to the typical minimum distance  $\omega_{\min}$ , the bound stays nontrivial. In Tables I and II,  $\lim_{\omega \rightarrow \omega_{\min}^+} 1 - \frac{\delta(\omega_{\min}, \mathbf{1}, \mathbf{r})}{0.95^2}$  is given for regular codes of rate 0.5 and 0.25 respectively. We observe that if we fix the rate and let  $\mathbf{1}$  and  $\mathbf{r}$  increase then the bound approaches 1

for all  $\omega$  for which  $W_{\text{com}}(\omega)$  is positive. This implies that for regular ensembles with large left and right degree almost all the codes have a weight distribution which is very close to the ensemble average. We observe the same phenomenon for stopping set distribution. We see that the second moment method can capture the concentration property of the weight distribution and stopping set distribution for regular ensembles with large left and right degrees. However for the regular ensembles in general it fails to do so. Potentially by applying more sophisticated methods one could obtain better bounds, e.g., the second moment method with conditioning [16] or other methods given in [17].

## APPENDIX I

We modify the proof of [14, Th. 2] to prove Lemma 3.3. Let  $\varphi_n(\underline{z}) = f(\underline{z}) \frac{\mathbf{1}}{\mathbf{r}}, R = [-\pi, \pi]^3$  and  $I = \sqrt{-1}$ . We also expand  $\varphi_n(\underline{z})$  as  $\varphi_n(\underline{z}) = \sum_{\underline{k}} a_n(\underline{k}) \underline{z}^{\underline{k}}$ . Let  $\underline{t}$  be the positive solution of  $a_f(\underline{x}) = \frac{\mathbf{r}}{\mathbf{1}}$ . From the inverse Fourier transform, we get

$$\frac{1}{(2\pi)^3} \int_R \frac{\varphi_n(\underline{t} e^{I\underline{v}})}{\varphi_n(\underline{t})} e^{-I\underline{j} \cdot \underline{v}^T} d\underline{v} = \frac{a_n(\underline{j}) \underline{t}^{\underline{j}}}{\varphi_n(\underline{t})}. \quad (17)$$

We recall that the Fourier transform of a Gaussian is again a Gaussian

$$\begin{aligned} & \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-I\underline{u} \cdot \underline{s}^T - \frac{\underline{s} \cdot B_f(\underline{t}) \cdot \underline{s}^T}{2}} d\underline{s} \\ &= \sqrt{\frac{(2\pi)^3}{|B_f(\underline{t})|}} e^{-\frac{1}{2} \underline{u} \cdot B_f(\underline{t})^{-1} \cdot \underline{u}^T}. \quad (18) \end{aligned}$$

Also from the proof of [14, Th. 2], for any function  $K(n)$  growing with  $n$ , we have

$$\begin{aligned} & \left| \int_{[-K(n), K(n)]^3} e^{-I\underline{u} \cdot \underline{s}^T - \frac{\underline{s} \cdot B_f(\underline{t}) \cdot \underline{s}^T}{2}} d\underline{s} \right. \\ & \left. - \int_{(-\infty, \infty)^3} e^{-I\underline{u} \cdot \underline{s}^T - \frac{\underline{s} \cdot B_f(\underline{t}) \cdot \underline{s}^T}{2}} d\underline{s} \right| = O\left(\frac{1}{K(n)}\right). \quad (19) \end{aligned}$$

We would like to show that for  $n \rightarrow \infty$

$$\begin{aligned} & \left| \left( \frac{n\mathbf{1}}{\mathbf{r}} \right)^{\frac{3}{2}} \int_R \frac{\varphi_n(\underline{t} e^{I\underline{v}})}{\varphi_n(\underline{t})} e^{-I\underline{j} \cdot \underline{v}^T} d\underline{v} \right. \\ & \left. - 4 \sqrt{\frac{(2\pi)^3}{|B_f(\underline{t})|}} e^{-\frac{1}{2} \underline{u} \cdot B_f(\underline{t})^{-1} \cdot \underline{u}^T} \right| = o(1). \quad (20) \end{aligned}$$

To prove this, we write  $\varphi_n(\underline{t} e^{I\underline{v}})$  in exponential-log form and take the Taylor series expansion of the exponent around  $\underline{v} = 0$

$$\varphi_n(\underline{t} e^{I\underline{v}}) = e^{\left( \frac{n\mathbf{1}}{\mathbf{r}} (\ln(f(\underline{t})) + I a_f(\underline{t}) \cdot \underline{v}^T - \frac{\underline{v} \cdot B_f(\underline{t}) \cdot \underline{v}^T}{2} + O(\|\underline{v}\|^3)) \right)}. \quad (21)$$

Note that as  $\ln(\varphi_n(\underline{t}))$  is analytic, so all the third-order partial derivative of  $\varphi_n(\underline{t} e^{I\underline{v}})$  are bounded. Now we partition  $R$  as  $R = \cup_{i=1}^5 R_i$ , where  $R_1 = [-\delta, \delta]^3$ ,  $R_2 = [-\delta, \delta] \times [\pi - \delta, \pi + \delta]^2$ ,  $R_3 = [\pi - \delta, \pi + \delta] \times [-\delta, \delta] \times [\pi - \delta, \pi + \delta]$ ,  $R_4 = [\pi - \delta, \pi + \delta]^2 \times [-\delta, \delta]$ ,  $R_5 = R \setminus (R_1 \cup R_2 \cup R_3 \cup R_4)$ . Here  $\delta$  can be any decaying function of  $n$  which satisfies that as  $n \rightarrow \infty$  then  $n\delta^2 \rightarrow \infty$  and  $n\delta^3 \rightarrow 0$ . We choose  $\delta = n^{-\frac{2}{5}}$ . This ensures that the term  $O(\|\underline{v}\|^3)$  is negligible. By the symmetry of

$$\mathbb{E}[S^2(G, ns)] = \sum_{i=0}^{ns} \underbrace{\text{Coeff} \left( g(\underline{x})^n \frac{\mathbf{1}}{\mathbf{r}}, x_1^{1(n-s-i)} x_2^1 x_3^{1(n-s-i)} \right)}_{C_i} \cdot \underbrace{\binom{n}{ns} \binom{ns}{i} \binom{n-ns}{ns-i} ((1(n-s-i))! (1)! (1(n-2ns+i))!)}_{F_i}. \quad (16)$$

$f(\underline{x}), \varphi_n(x_1, x_2, x_3) = \varphi_n(x_1, -x_2, -x_3) = \varphi_n(-x_1, x_2, -x_3) = \varphi_n(-x_1, -x_2, x_3)$ . This implies

$$\int_{R_1} \frac{\varphi_n(\underline{t}e^{I\underline{v}})}{\varphi_n(\underline{t})} e^{-I\underline{j}\cdot\underline{v}^T} d\underline{v} = \int_{R_k} \frac{\varphi_n(\underline{t}e^{I\underline{v}})}{\varphi_n(\underline{t})} e^{-I\underline{j}\cdot\underline{v}^T} d\underline{v}$$

where  $k \in \{2, 3, 4\}$ . Now

$$\int_{R_1} \frac{\varphi_n(\underline{t}e^{I\underline{v}})}{\varphi_n(\underline{t})} e^{-I\underline{j}\cdot\underline{v}^T} d\underline{v} \stackrel{\alpha_f(\underline{x}) = \frac{\underline{r}\underline{i}}{n\underline{1}}}{(21)} \int_{R_1} e^{I(\underline{i}-\underline{j})\cdot\underline{v}^T - \frac{n\underline{1}}{2\underline{r}}\underline{v}\cdot B_f(\underline{t})\cdot\underline{v}^T + O(n\delta^3)} d\underline{v}. \quad (22)$$

The change of variable  $\underline{y} := \sqrt{\frac{n\underline{1}}{\underline{r}}}\underline{v}$  in (22) gives

$$\int_{R_1} \frac{\varphi_n(\underline{t}e^{I\underline{v}})}{\varphi_n(\underline{t})} e^{-I\underline{j}\cdot\underline{v}^T} d\underline{v} = \left(\frac{\underline{r}}{n\underline{1}}\right)^{\frac{3}{2}} \times \int_{R'_1} e^{-I\underline{u}\cdot\underline{y}^T - \frac{\underline{y}\cdot B_f(\underline{t})\cdot\underline{y}^T}{2}} (1 + O(n^{-\frac{1}{5}})) d\underline{y} \quad (23)$$

where  $R'_1 = [-\sqrt{\frac{1}{\underline{r}}n\frac{1}{10}}, \sqrt{\frac{1}{\underline{r}}n\frac{1}{10}}]^3$ . Using (18), (19) in (23) gives

$$\begin{aligned} & \int_{R_1} \frac{\varphi_n(\underline{t}e^{I\underline{v}})}{\varphi_n(\underline{t})} e^{-I\underline{j}\cdot\underline{v}^T} d\underline{v} \\ &= \left(\frac{\underline{r}}{n\underline{1}}\right)^{\frac{3}{2}} O(n^{-\frac{1}{10}}) + \left(\frac{\underline{r}}{n\underline{1}}\right)^{\frac{3}{2}} \sqrt{\frac{(2\pi)^3}{|B_f(\underline{t})|}} \\ & \times e^{-\frac{\underline{u}\cdot B_f(\underline{t})\cdot\underline{u}^T}{2}} (1 + O(n^{-\frac{1}{5}})). \end{aligned} \quad (24)$$

Recall that,

$$\int_{R_5} \left| \frac{\varphi_n(\underline{t}e^{I\underline{v}})}{\varphi_n(\underline{t})} e^{-I\underline{j}\cdot\underline{v}^T} \right| d\underline{v} = \int_{R_5} \left| \frac{f(\underline{t}e^{I\underline{v}})}{f(\underline{t})} \right|^{\frac{n\underline{1}}{\underline{r}}} d\underline{v}.$$

Further,  $f(\underline{t})$  is a 3-variable polynomial of finite degree. Let  $f(\underline{t}) = \sum_k b(\underline{k})\underline{t}^{\underline{k}}$ . Then by some algebraic manipulation we get

$$\left| \frac{f(\underline{t}e^{I\underline{v}})}{f(\underline{t})} \right|^2 = 1 - \frac{\sum_{\underline{k} \neq \underline{l}} b(\underline{k})b(\underline{l})\underline{t}^{\underline{k}+\underline{l}}(1 - \cos((\underline{k} - \underline{l}) \cdot \underline{v}^T))}{f(\underline{t})^2}. \quad (25)$$

Also  $f(\underline{t})$  has  $1, t_1^2, t_2^2, t_3^2$  as its summation terms and in  $R_5$  at least one of the variable  $v_k$  satisfies  $v_k \notin [-\delta, \delta]$  where  $k \in \{1, 2, 3\}$ . This with (25) implies that for some positive constants  $c, c_1$

$$\begin{aligned} \int_{R_4} \left| \frac{f(\underline{t}e^{I\underline{v}})}{f(\underline{t})} \right|^{\frac{n\underline{1}}{\underline{r}}} d\underline{v} &\leq \pi^3 (1 - c_1 \delta^2)^{\frac{n\underline{1}}{2\underline{r}}} \\ &= \pi^3 (1 - c_1 n^{-\frac{4}{5}})^{\frac{n\underline{1}}{2\underline{r}}} \\ &= O(e^{-cn\frac{1}{5}}). \end{aligned}$$

By combining the above steps, we get

$$\begin{aligned} & \left| \left(\frac{n\underline{1}}{\underline{r}}\right)^{\frac{3}{2}} \int_{R_1} \frac{\varphi_n(\underline{t}e^{I\underline{v}})}{\varphi_n(\underline{t})} e^{-I\underline{j}\cdot\underline{v}^T} d\underline{v} \right. \\ & \left. - 4 \sqrt{\frac{(2\pi)^3}{|B_f(\underline{t})|}} e^{-\frac{1}{2}\underline{u}\cdot B_f(\underline{t})\cdot\underline{u}^T} \right| \\ &= O\left(n^{-\frac{1}{10}}\right). \end{aligned} \quad (26)$$

By using (17) in (26), we get

$$\begin{aligned} & \left| \left(\frac{n\underline{1}}{\underline{r}}\right)^{\frac{3}{2}} \frac{(2\pi)^3 a_n(\underline{j})\underline{t}^{\underline{j}}}{\varphi_n(\underline{t})} \right. \\ & \left. - 4 \sqrt{\frac{(2\pi)^3}{|B_f(\underline{t})|}} e^{-\frac{1}{2}\underline{u}\cdot B_f(\underline{t})\cdot\underline{u}^T} \right| = O(n^{-\frac{1}{10}}). \end{aligned} \quad (27)$$

The approximation of  $\text{Coeff}(f(\underline{x})^{\frac{n\underline{1}}{\underline{r}}}, \underline{x}^{\underline{j}})$  is obtained by substituting  $\underline{j} = \underline{i}$  (which implies  $\underline{u} = (0, 0, 0)$ ) in (27). Also for the local limit theorem to hold we need in (27) that  $e^{\frac{1}{2}\underline{u}\cdot B_f(\underline{t})\cdot\underline{u}^T} n^{-\frac{1}{10}} = o(1)$ . For our application choosing  $\|\underline{u}\| = O((\ln n)^{\frac{1}{3}})$  suffices.

#### ACKNOWLEDGMENT

The author wishes to thank Rüdiger Urbanke for many helpful discussions and suggestions on the preparation of the paper. He is grateful to anonymous reviewers for their comments which helped in improving the quality of the paper.

#### REFERENCES

- [1] A. Vardy, "The intractability of computing the minimum distance of code," *IEEE Trans. Inf. Theory*, vol. 43, pp. 1757–1766, 1997.
- [2] S. Condamin, Study of the Weight Enumerator Function for a Gallager Code, Project Report Cavendish Laboratory, University of Cambridge, Cambridge, U.K., Jul. 2002.
- [3] S. Van Mourik, D. Saad, and Y. Kabashima, "Critical noise levels for LDPC decoding," *Phys. Rev. E*, vol. 66, 2002.
- [4] N. Sourlas, "Spin-glass models as error-correcting codes," *Nature*, vol. 339, pp. 693–695, 1989.
- [5] N. Sourlas, "Spin glasses, error-correcting codes and finite-temperature decoding," *Europhys. Lett.*, vol. 25, pp. 159–164, 1994.
- [6] C. Di, A. Montanari, and R. Urbanke, "Weight distribution of LDPC code ensemble: Combinatorics meets statistical physics," in *Proc. Int. Symp. Inf. Theory 2004*, Chicago, IL, USA.
- [7] A. Montanari, "The Glassy Phase of Gallager Codes," *European Phys. J.*, vol. 23, 2001.
- [8] R. G. Gallager, *Low-Density Parity-Check Codes*. Cambridge, MA: M.I.T. Press, 1963.
- [9] D. Burshtein and G. Miller, "Asymptotic enumeration methods for analyzing LDPC codes," *IEEE Trans. Inf. Theory*, vol. 50, no. 6, pp. 1115–1131, June 2004.
- [10] C. Di, T. Richardson, and R. Urbanke, "Weight distribution of iterative coding systems: How deviant can you be?," in *Proc. IEEE Int. Symp. Inf. Theory*, Washington, DC, Jun. 2001, p. 50.
- [11] S. Litsyn and V. Shevelev, "On ensembles of low-density parity-check codes: Asymptotic distance distributions," *IEEE Trans. Inf. Theory*, vol. 48, no. 4, Apr. 2002.
- [12] P. Flajolet and R. Sedgewick, Analytic Combinatorics preprint.
- [13] D. Gardy, "Some results on the asymptotic behavior of coefficients of large powers of functions," *Discr. Math.*, vol. 139, pp. 189–217, 1995.
- [14] E. A. Bender and L. B. Richmond, "Central and local limit theorems applied to asymptotic enumeration II: Multivariate generating functions," *J. Combin. Theory Ser. A*, ser. A 34, pp. 255–265, 1983.
- [15] A. Orłitsky, K. Viswanathan, and J. Zhang, "Stopping set distribution of LDPC code ensembles," in *Proc. ISIT 2003*, p. 23.
- [16] S. Janson, "The second moment method, conditioning and approximation," in *Random Discrete Structures (Proceedings, Minneapolis 1993)*. New York: Springer, 1996, pp. 175–183.
- [17] C. McDiarmid, "Concentration," in *Probabilistic Methods for Algorithmic Discrete Mathematics*, M. Habib, C. McDiarmid, and J. Ramirez-Alfonsin, Eds. Berlin/Heidelberg, Germany: Springer-Verlag, 1998, pp. 195–247.