On Concentration of Measures for LDPC Code Ensembles

Igal Sason Ronen Eshel
sason@ee.technion.ac.il reshel@gmail.com
Department of Electrical Engineering
Technion – Israel Institute of Technology
Haifa 32000, Israel

***Abstract*—This work considers the concentration of measures for low-density parity-check (LDPC) code ensembles. The two results derived in this paper follow from Azuma's inequality for Doob martingales with bounded differences. The first result is a tightened concentration inequality for the conditional entropy (originally derived by Méasson et al.), and the second result is a concentration inequality for the cardinality of the fundamental systems of cycles of a bipartite graph from the ensemble.**

***Index Terms*—Azuma's inequality, concentration of measures, low-density parity-check (LDPC) codes, martingales.**

I. Introduction

The concentration of measure phenomenon is a principle that is applied in measure theory, probability and combinatorics, and it has consequences in some other fields such as functional analysis (see, e.g., [3], [5] and [15]).

The basic concentration theorem of iterative message-passing decoding (see [10, pp. 487–490]) asserts that all except an exponentially (in the block length) small fraction of codes perform within an arbitrary small δ from the ensemble average (where δ is a positive number that can be chosen arbitrarily small). Therefore, assuming a sufficiently large block length, the ensemble average forms a good indicator for the performance of individual codes. In general, all the concentration inequalities which have been proved in the setting of iterative message-passing decoding so far are rather loose, and much stronger concentration phenomena can be observed in practice for moderate to large block lengths. Therefore, to date, these concentration inequalities serve mostly to justify theoretically the ensemble approach, but they are not tight for finite block lengths.

In the following, we present briefly relevant mathematical background that is essential to this work.

A. Doob's Martingales

This sub-section provides a short background on martingales to set definitions and notation. For a more thorough study of martingales, the reader it referred to, e.g., [11, Chapter 14].

Definition 1: **[Doob's Martingale]** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A Doob's martingale sequence is a sequence $X_0, X_1, \ldots$ of random variables (RVs) and corresponding sub σ-algebras $\mathcal{F}_0, \mathcal{F}_1, \ldots$ that satisfy the following conditions:

1) $X_i \in \mathbb{L}^1(\Omega, \mathcal{F}_i, \mathbb{P})$ for every i, i.e., each X_i is defined on the same sample space Ω, it is measurable with respect to the corresponding σ-algebra $\mathcal{F}_i$ (i.e., X_i is $\mathcal{F}_i$-measurable) and $\mathbb{E}[|X_i|] = \int_\Omega |X_i(\omega)| d\mathbb{P}(\omega) < \infty$.
2) $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \ldots$ (where this sequence of σ-algebras is called a filtration).
3) The equality $X_i = \mathbb{E}[X_{i+1}|\mathcal{F}_i]$ holds almost surely (a.s.) for every i.

Remark 1: For every i

$$\mathbb{E}[X_{i+1}] = \mathbb{E}\big[\mathbb{E}[X_{i+1}|\mathcal{F}_i]\big] = \mathbb{E}[X_i]$$

so the expectation of a martingale stays constant.

Remark 2: One can generate martingale sequences by the following procedure: Given a RV $X \in \mathbb{L}^1(\Omega, \mathcal{F}, \mathbb{P})$ and an arbitrary filtration of sub σ-algebras $\{\mathcal{F}_i\}$, let

$$X_i = \mathbb{E}[X|\mathcal{F}_i] \quad i = 0, 1, \ldots.$$

Then, the sequence $X_0, X_1, \ldots$ forms a martingale since

1) The RV $X_i = \mathbb{E}[X|\mathcal{F}_i]$ is $\mathcal{F}_i$-measurable, and also $\mathbb{E}[|X_i|] \leq \mathbb{E}[|X|] < \infty$ (since conditioning reduces the expectation of the absolute value).
2) By construction $\{\mathcal{F}_i\}$ is a filtration.
3) For every i

$$\begin{aligned}
&\mathbb{E}[X_{i+1}|\mathcal{F}_i] \\
&= \mathbb{E}\big[\mathbb{E}[X|\mathcal{F}_{i+1}]|\mathcal{F}_i\big] \\
&= \mathbb{E}[X|\mathcal{F}_i] \quad (\text{since } \mathcal{F}_i \subseteq \mathcal{F}_{i+1}) \\
&= X_i \quad \text{a.s.}
\end{aligned}$$

Remark 3: In continuation to Remark 2, one can choose

$$\mathcal{F}_0 = \{\Omega, \emptyset\}, \quad \mathcal{F}_n = \mathcal{F}$$

so that $X_0, X_1, \ldots, X_n$ is a martingale sequence where

$$\begin{aligned}
X_0 &= \mathbb{E}[X|\mathcal{F}_0] = \mathbb{E}[X] \quad (\text{since } X \text{ is independent of } \mathcal{F}_0) \\
X_n &= \mathbb{E}[X|\mathcal{F}_n] = X \ \text{ a.s.} \quad (\text{since } X \text{ is } \mathcal{F}\text{-measurable}).
\end{aligned}$$

In this case, one gets a martingale sequence where the first element is the expected value of X, and the last element of the sequence is X itself (a.s.). This has the following interpretation: At the beginning, we don't know anything about X, so it is initially estimated by its expectation. We then reveal at each step more and more information about X until we can specify it exactly (a.s.).

B. Azuma's Inequality

Definition 2: **[d-Lipschitz martingales]** Let $X_0, \ldots, X_n$ be a martingale, and $d = (d_1, \ldots, d_n)$ be a vector with positive entries. The sequence $X_0, \ldots, X_n$ is said to be a d-Lipschitz martingale if

$$|X_i - X_{i-1}| \leq d_i \quad \forall\, i = 1, 2, \ldots, n \quad \text{a.s.}$$

Azuma's inequality [1] forms a useful concentration inequality for d-Lipschitz martingales. In the following, this inequality is introduced. The interested reader is referred to [2] for a survey on concentration inequalities for (super/ sub) martingales.

Theorem 1: **[Azuma's inequality]** Let $X_0, \ldots, X_n$ be a d-Lipschitz martingale, then

$$\mathbb{P}(|X_n - X_0| \geq r) \leq 2 \exp\left(-\frac{r^2}{2\sum_{i=1}^n d_i^2}\right), \quad \forall\, r > 0.$$

This theorem is proved, e.g., in [2].

C. Concentration of the Conditional Entropy

In the following, we consider ensembles of binary low-density parity-check (LDPC) codes. Following standard notation, let λ_i and ρ_i denote the fraction of edges attached, respectively, to variable and parity-check nodes of degree i. The LDPC code ensemble that is denoted by LDPC(n, λ, ρ) is characterized by the block length n of the codes, and the pair $\lambda(x) \triangleq \sum_i \lambda_i x^{i-1}$ and $\rho(x) \triangleq \sum_i \rho_i x^{i-1}$ which represent, respectively, the left and right degree distributions from the edge perspective.

The concentration of the conditional entropy for random ensembles of LDPC codes was stated in [7, Theorem 4] and [8, Theorem 1]. The following theorem was proved in [10, Appendix C] based on Azuma's inequality:

Theorem 2: **[Concentration of conditional entropy]** Let $\mathcal{C}$ be chosen uniformly at random from the ensemble LDPC(n, λ, ρ). Assume that the transmission of the code $\mathcal{C}$ takes place over an MBIOS channel. Let $H(\mathbf{X}|\mathbf{Y})$ designate the conditional entropy of the transmitted codeword $\mathbf{X}$ given the received sequence $\mathbf{Y}$ from the channel. Then for any $\xi > 0$,

$$\mathbb{P}(|H(\mathbf{X}|\mathbf{Y}) - \mathbb{E}_{\text{LDPC}(n,\lambda,\rho)}[H(\mathbf{X}|\mathbf{Y})]| \geq \sqrt{n}\,\xi) \leq 2\exp(-B\xi^2)$$

where $B \triangleq \frac{1}{2(d_{\text{c}}^{\max}+1)^2(1-R_d)}$, $d_{\text{c}}^{\max}$ is the maximal check-node degree, and R_{d} is the design rate of the ensemble.
The conditional entropy scales linearly with n, and this concentration inequality considers deviations at the order of the square root of n.

The proof of Theorem 2 is revisited in Section II-A for the derivation of a tightened concentration inequality (see Theorem 3).

II. Derivation of Two Concentration Results for LDPC Code Ensembles

A. A Tightened Concentration Inequality for the Conditional Entropy

In the following, we revisit the proof of Theorem 2 in [7, Appendix I] in order to derive a tightened version of this concentration result.

Following the proof in [7, Appendix I], let $\mathcal{G}$ be a bipartite graph which represents a code chosen uniformly at random from the ensemble LDPC(n, λ, ρ). Define the RV

$$Z = H_{\mathcal{G}}(\mathbf{X}|\mathbf{Y})$$

which forms the conditional entropy when the transmission takes place over an MBIOS channel whose transition probability is given by $P_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^n p_{Y|X}(y_i|x_i)$ where $p_{Y|X}(y|1) = p_{Y|X}(-y|0)$. Fix an arbitrary order for the $m = n(1-R_{\text{d}})$ parity-check nodes where R_{d} forms the design rate of the LDPC code ensemble. Let $\{\mathcal{F}_t\}_{t\in\{0,1,\ldots,m\}}$ form a filtration of the σ-algebras that are generated by the first t parity-check equations. Accordingly, based on Remark 2, let us define the following martingale sequence

$$Z_t = \mathbb{E}[Z|\mathcal{F}_t] \quad t \in \{0, 1, \ldots, m\}.$$

By construction, $Z_0 = \mathbb{E}[H_{\mathcal{G}}(\mathbf{X}|\mathbf{Y})]$ is the expected value of the conditional entropy for the LDPC code ensemble, and Z_m is the RV that is equal (a.s.) to the conditional entropy of the particular code from the ensemble (see Remark 3). Similarly to [7, Appendix I], we obtain upper bounds on the differences $|Z_{t+1} - Z_t|$ and then rely on Azuma's inequality in Theorem 1.

Without loss of generality, the parity-checks are ordered in [7, Appendix I] by increasing degree. Let $\mathbf{r} = (r_1, r_2, \ldots)$ be the set of parity-check degrees in ascending order, and Γ_i be the fraction of parity-check nodes of degree i. Hence, the first $m_1 = n(1-R_{\text{d}})\Gamma_{r_1}$ parity-check nodes are of degree r_1, the successive $m_2 = n(1-R_{\text{d}})\Gamma_{r_2}$ parity-check nodes are of degree r_2, and so on. The $(t+1)$th parity-check will therefore have a well defined degree, to be denoted by r. In order to avoid any further overlap with the proof in [7, Appendix I], we note that according to this proof

$$|Z_{t+1} - Z_t| \leq (r+1)\, H_{\mathcal{G}}(\tilde{X}|\mathbf{Y}) \tag{1}$$

where $H_{\mathcal{G}}(\tilde{X}|\mathbf{Y})$ is a RV which designates the conditional entropy of a parity-bit $\tilde{X} = X_{i_1} \oplus \ldots \oplus X_{i_r}$ (i.e., $\tilde{X}$ is equal to the modulo-2 sum of some r bits in the codeword $\mathbf{X}$) given the received sequence $\mathbf{Y}$ at the channel output. The proof in [7, Appendix I] was then completed by upper bounding the parity-check degree r by the maximal parity-check degree $d_{\text{c}}^{\max}$, and also by upper bounding the conditional entropy of the parity-bit $\tilde{X}$ by 1. This gives

$$|Z_{t+1} - Z_t| \leq d_{\text{c}}^{\max} + 1 \quad t = 0, 1, \ldots, m-1. \tag{2}$$

which then proves Theorem 2 from Azuma's inequality. Note that the d_i's in Theorem 1 are equal to $d_{\text{c}}^{\max} + 1$, and n in Theorem 1 is replaced with the length $m = n(1 - R_{\text{d}})$ of the martingale sequence $\{Z_t\}$ (that is equal to the number of the parity-check nodes in the graph).

In the continuation, we deviate from the proof in [7, Appendix I] in two respects:

- The first issue is related to the upper bound on the conditional entropy $H_{\mathcal{G}}(\tilde{X}|\mathbf{Y})$ in (1) where $\tilde{X}$ is the modulo-2 sum of some r bits of the transmitted codeword $\mathbf{X}$ given the channel output $\mathbf{Y}$. Instead of taking the most trivial upper bound that is equal to 1, as was done in [7, Appendix I], a simple upper bound on the conditional

entropy is derived; this bound depends on the parity-check degree r and the capacity C (see Lemma 1).

- The second difference is minor, but it proves to be helpful for tightening the large-deviation inequality for LDPC code ensembles that are not right-regular (i.e., the case where the degrees of the parity-check nodes are not fixed to a certain value). Instead of upper bounding the term $r+1$ on the right-hand side of (1) with $d_{\mathrm{c}}^{\max}+1$, it is suggested to leave it as is since Azuma's inequality applies in general to the case where the differences of the martingale sequence in Theorem 1 are not fixed, and since the number of parity-check nodes of degree r is equal to $n(1-R_{\mathrm{d}})\Gamma_r$. The effect of this simple modification is shown in Example 2.

The following upper bound is related to the first item above:

Lemma 1: Let $\mathcal{G}$ be a bipartite graph which corresponds to a binary linear block code whose transmission takes place over an MBIOS channel. Let $\mathbf{X}$ and $\mathbf{Y}$ designate the transmitted codeword and received sequence at the channel output. Let $\tilde{X} = X_{i_1} \oplus \ldots \oplus X_{i_r}$ be a parity-bit of some r code bits of $\mathbf{X}$. Then, the conditional entropy of $\tilde{X}$ given $\mathbf{Y}$ satisfies

$$H_{\mathcal{G}}(\tilde{X}|\mathbf{Y}) \leq h_2\left(\frac{1-C^{\frac{r}{2}}}{2}\right). \tag{3}$$

Further, for a binary symmetric channel (BSC) or a binary erasure channel (BEC), this bound can be improved to

$$h_2\left(\frac{1-\left[1-2h_2^{-1}(1-C)\right]^r}{2}\right) \tag{4}$$

and

$$1-C^r \tag{5}$$

respectively, where h_2^{-1} in (4) designates the inverse of the binary entropy function on base 2.

Note that if the MBIOS channel is perfect (i.e., its channel capacity is $C=1$ bit per channel use) then (3) holds with equality (both sides of (3) are equal to zero), whereas the trivial upper bound is 1.

Proof: Let us upper bound the conditional entropy $H(\tilde{X}\big|\mathbf{Y})$ with $H(\tilde{X}\big|Y_{i_1},\ldots,Y_{i_r})$, where the latter conditioning refers to the intrinsic information for the bits $X_{i_1},\ldots X_{i_r}$ which are used to calculate the parity-bit $\tilde{X}$. Then, from [12, Eq. (17) and Appendix I], the conditional entropy of the bit $\tilde{X}$ given the n-length received sequence $\mathbf{Y}$ satisfies the inequality

$$H(\tilde{X}\big|\mathbf{Y}) \leq 1 - \frac{1}{2\ln 2}\sum_{p=1}^{\infty}\frac{(g_p)^r}{p(2p-1)} \tag{6}$$

where (see [12, Eq. (19)])

$$g_p \triangleq \int_0^{\infty} a(l)(1+e^{-l})\tanh^{2p}\left(\frac{l}{2}\right)dl, \quad \forall\, p\in\mathbb{N} \tag{7}$$

and $a(\cdot)$ denotes the symmetric *pdf* of the log-likelihood ratio at the output of the MBIOS channel, given that the channel input is equal to zero. From [12, Lemmas 4 and 5], it follows that $g_p \geq C^p$ for every $p\in\mathbb{N}$. Substituting this inequality in (6) gives that

$$\begin{aligned} H(\tilde{X}\big|\mathbf{Y}) &\leq 1-\frac{1}{2\ln 2}\sum_{p=1}^{\infty}\frac{C^{pr}}{p(2p-1)} \\ &= h_2\left(\frac{1-C^{\frac{r}{2}}}{2}\right) \end{aligned} \tag{8}$$

where the last transition follows from the power series expansion of the binary entropy function around one-half (see [12, Eq. (16)]):

$$h_2(x) = 1-\frac{1}{2\ln 2}\sum_{p=1}^{\infty}\frac{(1-2x)^{2p}}{p(2p-1)}, \quad 0\leq x\leq 1. \tag{9}$$

The tightened bound on the conditional entropy for the BSC is obtained from (6) and the equality

$$g_p = \left(1-2h_2^{-1}(1-C)\right)^{2p}, \quad \forall\, p\in\mathbb{N}$$

which holds for the BSC (see [12, Eq. (97)]). This replaces C on the right-hand side of (8) with $\left(1-2h_2^{-1}(1-C)\right)^2$, thus leading to the tightened bound in (4).

The tightened result for the BEC holds since from (7), $g_p = C$ for all $p\in\mathbb{N}$ (see [12, Appendix II]), and substituting this equality in (6) gives (4) (from (9), $\sum_{p=1}^{\infty}\frac{1}{p(2p-1)} = 2\ln 2$). This completes the proof of Lemma 1. ∎

From Lemma 1 and (1)

$$|Z_{t+1}-Z_t| \leq (r+1)\, h_2\left(\frac{1-C^{\frac{r}{2}}}{2}\right) \tag{10}$$

with the corresponding two improvements for the BSC and BEC (where the second term on the right-hand side of (10) is replaced by (4) and (5), respectively). This improves the loosened bound $(d_{\mathrm{c}}^{\max}+1)$ in [7, Appendix I].

From (10) and Theorem 1, we obtain the following tightened version of the concentration inequality in Theorem 2.

Theorem 3: **[Tightened concentration result for the conditional entropy]** Let $\mathcal{C}$ be chosen uniformly at random from the ensemble LDPC(n,λ,ρ). Assume that the transmission of the code $\mathcal{C}$ takes place over an MBIOS channel. Let $H(\mathbf{X}|\mathbf{Y})$ designate the conditional entropy of the transmitted codeword $\mathbf{X}$ given the received sequence $\mathbf{Y}$ at the channel output. Then

$$\mathbb{P}(|H(\mathbf{X}|\mathbf{Y}) - \mathbb{E}_{\mathrm{LDPC}(n,\lambda,\rho)}[H(\mathbf{X}|\mathbf{Y})]| \geq \sqrt{n}\,\xi) \leq 2\exp(-B\xi^2)$$

for every $\xi>0$, and

$$B \triangleq \frac{1}{2(1-R_{\mathrm{d}})\sum_{i=1}^{d_{\mathrm{c}}^{\max}}(i+1)^2\Gamma_i\left[h_2\left(\frac{1-C^{\frac{i}{2}}}{2}\right)\right]^2} \tag{11}$$

where $d_{\mathrm{c}}^{\max}$ is the maximal check-node degree, Γ_i is the fraction of parity-check nodes of degree i, R_{d} is the design rate of the ensemble, and C is the channel capacity (in bits per channel use). For the BSC and BEC, the parameter B in the exponent can be respectively improved (increased) to

$$B \triangleq \frac{1}{2(1-R_{\mathrm{d}})\sum_{i=1}^{d_{\mathrm{c}}^{\max}}(i+1)^2\Gamma_i\left[h_2\left(\frac{1-\left[1-2h_2^{-1}(1-C)\right]^i}{2}\right)\right]^2}$$

and

$$B \triangleq \frac{1}{2(1-R_{\mathrm{d}})\sum_{i=1}^{d_{\mathrm{c}}^{\max}}(i+1)^2\Gamma_i\,(1-C^i)^2}. \tag{12}$$

Remark 4: From (11), Theorem 3 indeed yields a stronger concentration inequality than Theorem 2.

Remark 5: In the limit where $C \to 1$ bit per channel use, it follows from (11) that if $d_{\mathrm{c}}^{\max} < \infty$ then $B \to \infty$. This is in contrast to the value of B in Theorem 2 which does not depend on the channel capacity and is finite. Note that B should be indeed infinity for a perfect channel, and therefore Theorem 3 is tight in this case.

In case that $d_{\mathrm{c}}^{\max} = \infty$, it is easy to show that if $\rho'(1) < \infty$ then $B \to \infty$ in the limit where $C \to 1$. This is in contrast to the value of B in Theorem 2 which vanishes when $d_{\mathrm{c}}^{\max} = \infty$, and therefore Theorem 2 does not enable to assert concentration in this case (see Example 2).

Example 1: [Comparison of Theorems 2 and 3 for right-regular LDPC code ensembles] In the following, we exemplify the improvement in the tightness of Theorem 3 for right-regular LDPC code ensembles. Consider the case where the communications takes place over a binary-input additive white Gaussian noise channel (BIAWGNC) or a BEC. Let us consider the $(2, 20)$ regular LDPC code ensemble whose design rate is equal to 0.900 bits per channel use. For a BEC, the threshold of the channel bit erasure probability under belief propagation (BP) decoding is given by

$$p_{\mathrm{BP}} = \inf_{x \in (0,1]} \frac{x}{1-(1-x)^{19}} = 0.0531$$

which corresponds to a channel capacity of $C = 0.9469$ bits per channel use. For the BIAWGNC, the threshold under BP decoding is equal to $\sigma_{\mathrm{BP}} = 0.4156590$. From [10, Example 4.38] which expresses the capacity of the BIAWGNC in terms of the standard deviation σ of the Gaussian noise, the minimum capacity of a BIAWGNC over which it is possible to communicate with vanishing bit error probability under BP decoding is $C = 0.9685$ bits per channel use. Accordingly, let us assume that for reliable communications on both channels, the capacity of the BEC and BIAWGNC is set to 0.98 bits per channel use.

Since the considered code ensembles is right-regular (i.e., the parity-check degree is fixed to $d_{\mathrm{c}} = 20$), then B in Theorem 3 is improved by a factor of $\left[h_2\left(\frac{1-C^{\frac{d_{\mathrm{c}}}{2}}}{2}\right)\right]^{-2} = 5.134$. This implies that the concentration inequality in Theorem 3 is satisfied with a block length that is 5.134 times shorter than the block length which corresponds to Theorem 2. For the BEC, the result is improved by a factor of $\frac{1}{\left(1-C^{d_{\mathrm{c}}}\right)^2} = 9.051$ due to the tightened value of B in (12) as compared to Theorem 2.

Example 2: [Comparison of Theorems 2 and 3 for a heavy-tail Poisson distribution (Tornado codes)] In the following, we compare Theorems 2 and 3 for Tornado LDPC code ensembles. This capacity-achieving sequence for the binary erasure channel (BEC) refers to the heavy-tail Poisson distribution, and it was introduced in [6, Section IV], [14] (see also [10, Problem 3.20]). We rely in the following on the analysis in [12, Appendix VI].

Suppose that we wish to design Tornado code ensembles that achieve a fraction $1-\varepsilon$ of the capacity of a BEC under iterative message-passing decoding (where ε can be set arbitrarily small). Let p designate the bit erasure probability of the channel. The parity-check degree is Poisson distributed, and therefore the maximal degree of the parity-check nodes is infinity. Hence, $B = 0$ according to Theorem 2, and this theorem therefore is useless for showing a concentration phenomenon for the considered code ensemble. On the other hand, from Theorem 3

$$\begin{aligned}
&\sum_i (i+1)^2 \Gamma_i \left[h_2\left(\frac{1-C^{\frac{i}{2}}}{2}\right)\right]^2 \\
&\overset{\text{(a)}}{\leq} \sum_i (i+1)^2 \Gamma_i \\
&\overset{\text{(b)}}{=} \frac{\sum_i \rho_i (i+2)}{\int_0^1 \rho(x)\, \mathrm{dx}} + 1 \\
&\overset{\text{(c)}}{=} (\rho'(1)+3) d_{\mathrm{c}}^{\mathrm{avg}} + 1 \\
&\overset{\text{(d)}}{=} \left(\frac{\lambda'(0)\rho'(1)}{\lambda_2} + 3\right) d_{\mathrm{c}}^{\mathrm{avg}} + 1 \\
&\overset{\text{(e)}}{\leq} \left(\frac{1}{p\lambda_2} + 3\right) d_{\mathrm{c}}^{\mathrm{avg}} + 1 \\
&\overset{\text{(f)}}{=} O\left(\log^2\left(\frac{1}{\varepsilon}\right)\right)
\end{aligned}$$

where inequality (a) holds since the binary entropy function on base 2 is bounded between zero and one, equality (b) holds since

$$\Gamma_i = \frac{\frac{\rho_i}{i}}{\int_0^1 \rho(x)\, \mathrm{d}x}$$

where Γ_i and ρ_i denote the fraction of parity-check nodes and the fraction of edges that are connected to parity-check nodes of degree i respectively (and also since $\sum_i \Gamma_i = 1$), equality (c) holds since $d_{\mathrm{c}}^{\mathrm{avg}} = \left(\int_0^1 \rho(x)\, \mathrm{d}x\right)^{-1}$ where $d_{\mathrm{c}}^{\mathrm{avg}}$ denotes the average parity-check node degree, equality (d) holds since $\lambda'(0) = \lambda_2$, inequality (e) is due to the stability condition for the BEC (where $p\lambda'(0)\rho'(1) < 1$ is a necessary condition for reliable communication on the BEC under BP decoding), and finally equality (f) follows from the analysis in [12, Appendix VI] (an upper bound on λ_2 is derived in [12, Eq. (120)], and the average parity-check node degree scales like $\log \frac{1}{\varepsilon}$). Hence, from the above chain of inequalities and (11), it follows that for a small gap to capacity, the parameter B in Theorem 3 scales (at least) like $O\left(\frac{1}{\log^2\left(\frac{1}{\varepsilon}\right)}\right)$. Theorem 3 therefore asserts that a concentration phenomenon for this LDPC code ensemble exists. As shown above, the parameter B in (11) tends to zero rather slowly as we let the fractional gap ε tend to zero (which therefore demonstrates a rather fast concentration in Theorem 3).

B. Concentration of the Cardinality of the Fundamental System of Cycles

It is well known that linear block codes which are represented by cycle-free bipartite (Tanner) graphs have poor performance even under ML decoding [4]. The bipartite graphs of capacity-approaching LDPC codes should have cycles. In [12] and [13], we address the following question:

Question: Consider an LDPC ensemble whose transmission takes place over a memoryless binary-input output symmetric channel, and refer to the bipartite graphs which represent codes from this ensemble where every code is chosen uniformly at random from the ensemble. How does the average cardinality of the fundamental system of cycles of these bipartite graphs scale as a function of the achievable gap to capacity ?

The notion of "the cardinality of the fundamental system of cycles of bipartite graphs" was formally introduced in [12, Section II-E]. An information-theoretic lower bound on the number of fundamental cycles for bipartite graphs of LDPC code ensembles was proved in [12, Corollary 1 on p. 8] (see also [13]). This bound was expressed in terms of the achievable gap to capacity when the communication takes place over an MBIOS channel. More explicitly, it has shown that the number of fundamental cycles should grow at least like $\log \frac{1}{\varepsilon}$ where ε designates the gap in rate to capacity. Hence, this lower bound is unbounded as the gap to capacity vanishes. Consistently with the study made in [4] about cycle-free codes, the lower bound on the cardinality of the fundamental system of cycles in [12, Corollary 1] showed the necessity of cycles in bipartite graphs which represent good LDPC code ensembles.

As a continuation to this work, we derive in the following a concentration result for the cardinality of the fundamental system of cycles for LDPC code ensembles. For preliminary material that is relevant for the derivation of following concentration result, the reader is referred to [12, Section II-E].

Let (n, λ, ρ) be an LDPC code ensemble, and let $\mathcal{G}$ be a bipartite graph that represents a code from this ensemble. Then, the cardinality of the fundamental system of cycles of $\mathcal{G}$, denoted by $\beta(\mathcal{G})$, is equal to $\beta(\mathcal{G}) = |E(\mathcal{G})| - |V(\mathcal{G})| + c(\mathcal{G})$ where $E(\mathcal{G})$, $V(\mathcal{G})$ and $c(\mathcal{G})$ denote the number of edges, vertices and components of $\mathcal{G}$, respectively. Therefore, for a code from the (n, λ, ρ) LDPC code ensemble, the cardinality of the fundamental system of cycles satisfies the equality

$$\beta(\mathcal{G}) = n\big[(1 - R_{\mathrm{d}}) d_{\mathrm{c}}^{\mathrm{avg}} - (2 - R_{\mathrm{d}})\big] + c(\mathcal{G}). \tag{13}$$

Let us construct a martingale sequence $X_0, \ldots, X_E$ where X_i (for $i = 0, 1, \ldots, E$) is a RV that denotes the number of components of a bipartite graph $\mathcal{G}$, chosen uniformly at random from the ensemble, given that we already revealed its first i edges (where $i = 1, \ldots, E$ and $E \triangleq n(1 - R_{\mathrm{d}})\, d_c^{\mathrm{avg}}$ is the number of edges in the graph). This martingale sequence satisfies $|X_i - X_{i-1}| \leq 1$ for $i = 1, \ldots, E$ (since if we reveal a new edge of $\mathcal{G}$, then the number of components in the graph can change by at most 1). By Azuma's inequality and (13), we obtain that for every $\lambda > 0$

$$\mathbb{P}\left(|\beta(\mathcal{G}) - \mathbb{E}_{\mathrm{LDPC}(n,\lambda,\rho)}[\beta(\mathcal{G})]| \geq \lambda\sqrt{E}\right) \leq 2\exp\left(-\frac{\lambda^2}{2}\right)$$

Note that, from (13), the expected value $\mathbb{E}_{\mathrm{LDPC}(n,\lambda,\rho)}[\beta(\mathcal{G})]$ scales linearly with the block length n. Hence, the above inequality shows a concentration result.

Theorem 4: **[Concentration result for the cardinality of the fundamental system of cycles]** Let $\mathcal{G}$ be a bipartite graph chosen uniformly at random from the ensemble $\mathrm{LDPC}(n, \lambda, \rho)$. Then, the cardinality of the fundamental system of cycles concentrates around its average (which scales linearly with the block length n), and it satisfies the inequality

$$\begin{aligned}&\mathbb{P}\left(|\beta(\mathcal{G}) - \mathbb{E}_{\mathrm{LDPC}(n,\lambda,\rho)}[\beta(\mathcal{G})]| \geq \lambda\sqrt{n}\right)\\ &\leq 2\exp\left(-\frac{\lambda^2}{2(1 - R_{\mathrm{d}})^2\,\left(d_c^{\mathrm{avg}}\right)^2}\right), \quad \forall\, \lambda > 0\end{aligned}$$

where R_{d} and $d_{\mathrm{c}}^{\mathrm{avg}}$ designate, respectively, the design rate and average degree of the parity-check nodes of a bipartite graph from the ensemble.

Remark 6: For various capacity-achieving sequences of LDPC code ensembles on the BEC, the average right degree scales like $\log \frac{1}{\varepsilon}$ where ε is the fractional gap to capacity [9]. This scaling is similar in essence to the lower bound on the average right degree in [12, Theorem 1]. Therefore, the exponential decay rate in the concentration inequality of Theorem 4 scales like $\left(\log \frac{1}{\varepsilon}\right)^{-2}$ and this asserts a rather strong concentration of the cardinality of the fundamental system of cycles around the average value. This concentration result complements the discussion in [12, Corollary 1] and [13] that introduced a lower bound on this average as a function of the fractional gap to capacity, and it scales like $\log \frac{1}{\varepsilon}$.

REFERENCES

[1] K. Azuma, "Weighted sums of certain dependent random variables," *Tohoku Mathematical Journal*, vol. 19, pp. 357–367, 1967.

[2] F. Chung and L. Lu, "Concentration inequalities and martingale inequalities: a survey," *Internet Mathematics*, vol. 3, no. 1, pp. 79–127, March 2006.

[3] D. P. Dubashi and A. Panconesi, *Concentration of Measure for the Analysis of Randomized Algorithms*, Cambridge University Press, 2009.

[4] T. Etzion, A. Trachtenberg and A. Vardy, "Which codes have cycle-free Tanner graphs ?," *IEEE Trans. on Information Theory*, vol. 45, no. 6, pp. 2173–2181, September 1999.

[5] M. Ledoux, *The Concentration of Measure Phenomenon*, Mathematical Surveys and Monographs (AMS), vol. 89, 2001.

[6] M. G. Luby, M. Mitzenmacher, M. A. Shokrollahi and D. A. Spielman, "Efficient erasure-correcting codes," *IEEE Trans. on Information Theory*, vol. 47, no. 2, pp. 569–584, February 2001.

[7] C. Méasson, A. Montanari and R. Urbanke, "Maxwell construction: The hidden bridge between iterative and maximum apposteriori decoding," *IEEE Trans. on Information Theory*, vol. 54, pp. 5277–5307, December 2008.

[8] A. Montanari, "Tight bounds for LDPC and LDGM codes under MAP decoding," *IEEE Trans. on Information Theory*, vol. 51, no. 9, pp. 3247–3261, September 2005.

[9] P. Oswald and A. Shokrollahi, "Capacity-achieving sequences for the erasure channel," *IEEE Trans. on Information Theory*, vol. 48, no. 12, pp. 3017–3028, December 2002.

[10] T. Richardson and R. Urbanke, *Modern Coding Theory*, Cambridge University Press, 2008.

[11] J. S. Rosenthal, *A First look at Rigorous Probability Theory*, World Scientific Publishes, second edition, 2006.

[12] I. Sason, "On universal properties of capacity-approaching LDPC code ensembles," *IEEE Trans. on Information Theory*, vol. 55, no. 7, pp. 2956–2990, July 2009.

[13] I. Sason, "On the fundamental system of cycles in the bipartite graphs of LDPC code ensembles," *Proceedings 2009 IEEE International Symposium on Information Theory (ISIT 2009)*, pp. 75–79, Seoul, Korea, June 28–July 3, 2009.

[14] A. Shokrollahi, "Capacity-achieving sequences," *IMA Volume in Mathematics and its Applications*, vol. 123, pp. 153–166, 2000.

[15] M. Talagrand, "Concentration of measure and isoperimteric inequalities in product spaces," *Publications Mathématiques de l'I.H.E.S*, vol. 81, pp. 93–205, 1995.