entropy

MDPI

Article

Error Exponents of LDPC Codes under Low-Complexity Decoding

Pavel Rybin [1,*], Kirill Andreev [1,2] and Victor Zyablov [3]

[1] Center for Computational and Data-Intensive Science and Engineering, Skolkovo Institute of Science and Technology, 121205 Moscow, Russia; k.andreev@skoltech.ru
[2] Sirius University of Science and Technology, 1 Olympic Ave, 354340 Sochi, Russia
[3] Laboratory №3—Transmission, Protection and Analysis of Information, Institute for Information Transmission Problems, Russian Academy of Sciences, 119991 Moscow, Russia; zyablov@iitp.ru
* Correspondence: p.rybin@skoltech.ru

Abstract: This paper deals with the specific construction of binary low-density parity-check (LDPC) codes. We derive lower bounds on the error exponents for these codes transmitted over the memoryless binary symmetric channel (BSC) for both the well-known maximum-likelihood (ML) and proposed low-complexity decoding algorithms. We prove the existence of such LDPC codes that the probability of erroneous decoding decreases exponentially with the growth of the code length while keeping coding rates below the corresponding channel capacity. We also show that an obtained error exponent lower bound under ML decoding almost coincide with the error exponents of good linear codes.

Keywords: low-density parity check (LDPC) codes; Gallager's LDPC codes; binary LDPC codes; decoding algorithm; low-complexity; error exponent; capacity

Citation: Rybin, P.; Andreev, K.; Zyablov, V. Error Exponents of LDPC Codes Under Low-Complexity Decoding. *Entropy* **2021**, *23*, 253. https://doi.org/10.3390/e23020253

Academic Editor: Balazs Matuz, Alexey Frolov and Aaron Gulliver

Received: 17 December 2020
Accepted: 19 February 2021
Published: 22 February 2021

1. Introduction

Low-density parity-check (LDPC) codes [1] are known for their very efficient low-complexity decoding algorithms. This paper's central question is: Are there LDPC codes that asymptotically achieve the capacity of binary-symmetric channel (BSC) under a low-complexity decoding algorithm? The following results help us construct LDPC code with specific construction and develop a decoding algorithm to answer yes to this question. So, Zyablov and Pinsker showed in [2] that the ensemble of LDPC codes, proposed by Gallager (G-LDPC codes), includes codes that can correct a number of the errors that grow linearly with the code length n while the decoding complexity remains $\mathcal{O}(n \log n)$. Later the lower bound on this fraction of errors was improved in [3–5]. Thus, the main idea of LDPC code construction and decoding algorithm, considered in this paper, is as follows. We need to introduce to the construction of G-LDPC code some "good" codes that reduce the number of errors from the channel in such a way that the low-complexity majority decoding can correct the rest errors. As "good" codes, we select the codes with the error exponent of good codes under ML decoding [6]. To introduce these "good" codes to the construction, we compose the parity-check matrix layer with the parity-check matrices of "good" codes. To meet the requirements on low-complexity ($\mathcal{O}(n \log n)$) of decoding algorithm we impose restrictions on the length of "good" codes (it must be small ($\log \log(n)$) compared to the length of whole construction). To show that the proposed construction asymptotically achieves the capacity of BSC we consider the estimation on error-exponent under the proposed low-complexity decoding algorithm.

It worth mention that papers [7,8] introduce expander codes achieving the BSC capacity under an iterative decoding algorithm with low complexity. But in this paper, we are interested in LDPC-code construction and corresponding decoding algorithm.

To show that the proposed construction of LDPC code is good, we also estimate the error exponent under ML decoding and compare it with the error exponent under the proposed low-complexity decoding algorithm. Previously the authors of [9,10] have derived the upper and lower bounds on the G-LDPC codes error exponent under the ML decoding assumption. Moreover, one can conclude from [10] that the lower bound on the error exponent under ML decoding of G-LDPC codes almost coincides with the lower bound obtained for good linear codes (from [6]) under ML.

Some parts of this paper were previously presented (with almost all of the proofs omitted) in the conference paper [11]. The low-complexity decoding algorithm that we use for our analysis was proposed in [12,13]. Unlike in previous papers, Corollary 1 is significantly enhanced and proved in detail in this paper. Moreover, the results for the error-exponent bound under ML decoding and corresponding proofs are added. We compare the obtained lower bounds on the error exponents under the low-complexity decoding and the ML decoding. We evaluate the error exponents numerically for different code parameters.

2. LDPC Code Construction

Let us briefly consider the LDPC code construction from [11,12]. First, let us consider the G-LDPC code parity-check matrix $\mathbf{H}_2$ of size $\ell \times b_0 n_0$ from [1]:

$$\mathbf{H}_2 = \begin{pmatrix} \pi_1(\mathbf{H}_{b_0}) \\ \pi_2(\mathbf{H}_{b_0}) \\ \vdots \\ \pi_\ell(\mathbf{H}_{b_0}) \end{pmatrix}.$$

Here we denote $\pi_l(\mathbf{H}_{b_0})$, $l = 1, \ldots, \ell$, as a random column permutation of $\mathbf{H}_{b_0}$, which is given by

$$\mathbf{H}_{b_0} = \begin{pmatrix} \underbrace{\begin{matrix} \mathbf{H}_0 & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_0 & \ldots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \ldots & \mathbf{H}_0 \end{matrix}}_{b_0} \end{pmatrix},$$

where $\mathbf{H}_0$ is the parity-check matrix of the constituent single parity check (SPC) code of length n_0.

The elements of the Gallager's LDPC codes ensemble $\mathscr{E}_G(\ell, n_0, b_0)$ are obtained by independently selecting the equiprobable permutations π_l, $l = 1, 2, ..., \ell$.

One can write a lower bound on the G-LDPC code rate $\mathscr{E}_G(\ell, n_0, b_0)$ as

$$R_2 \geq 1 - \ell(1 - R_0), \tag{1}$$

where $R_0 = \frac{n_0 - 1}{n_0}$ is a SPC code rate.

The equality achieved if and only if the matrix $\mathbf{H}_2$ has full rank.

Consider a G-LDPC parity check matrix with an additional layer consisting of linear codes (LG-LDPC code). Let us denote this matrix as $\mathbf{H}$:

$$\mathbf{H} = \begin{pmatrix} \pi_1(\mathbf{H}_{b_0}) \\ \pi_2(\mathbf{H}_{b_0}) \\ \vdots \\ \pi_\ell(\mathbf{H}_{b_0}) \\ \pi_{\ell+1}(\mathbf{H}_{b_1}) \end{pmatrix},$$

where $\mathbf{H}_{b_1}$ is given by

$$\mathbf{H}_{b_1} = \begin{pmatrix} \mathbf{H}_1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_1 & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \underbrace{\mathbf{0} \quad \mathbf{0} \quad \dots \quad \mathbf{H}_1}_{b_1} \end{pmatrix},$$

where b_1 is such that $b_1 n_1 = b_0 n_0$. As soon as the first ℓ layers of matrix $\mathbf{H}$ is the G-LDPC parity-check matrix, we can write $\mathbf{H}$ as

$$\mathbf{H} = \begin{pmatrix} \mathbf{H}_2 \\ \pi_{\ell+1}(\mathbf{H}_{b_1}) \end{pmatrix}.$$

For a given SPC code with the code length n_0 and the parity-check matrix $\mathbf{H}_0$ and for a given linear code with the code length n_1 and the parity-check matrix $\mathbf{H}_1$, the elements of the LG-LDPC codes ensemble $\mathscr{E}_{LG}(\ell, n_0, b_0, R_1, n_1, b_1)$ are obtained by independently selecting the equiprobable permutations π_l, $l = 1, 2, ..., \ell + 1$.

The length of the constructed LG-LDPC code is $n = b_0 n_0 = b_1 n_1$, and the code rate R is lower bounded by

$$R \geq R_1 - \ell(1 - R_0),$$

According to (1),

$$R \geq R_1 + R_2 - 1.$$

3. Decoding Algorithms

In this paper, we consider two decoding algorithms of the proposed construction. The first one is the well-known maximum likelihood decoding algorithm $\mathscr{A}_{ML}$. Under the second decoding algorithm $\mathscr{A}_C$, the LG-LDPC code is decoded as a concatenated code. In other words, in the first step, we decode the received sequence using linear codes with the parity-check matrix $\mathbf{H}_1$ from the $\ell + 1$ layer of $\mathbf{H}$. Then, in the second step, we decode the sequence obtained in the previous step using the G-LDPC code with the parity-check matrix $\mathbf{H}_2$. Thus, this algorithm $\mathscr{A}_C$ consists of the following two steps:

1. The received sequence is separately decoded with the well-known maximum likelihood algorithm by b_1 linear codes with the parity-check matrix $\mathbf{H}_1$ from the $\ell + 1$ layer of $\mathbf{H}$.
2. The tentative sequence is decoded with the well-known bit-flipping (majority) decoding algorithm $\mathscr{A}_M$ by the G-LDPC code with the parity-check matrix $\mathbf{H}_2$.

An important note here is that $\mathscr{A}_C$ is a two-step decoding algorithm, and each step is performed only once. It first decodes the received sequence once by the ML decoding algorithm using linear codes $\mathbf{H}_1$. Then it applies the iterative bit-flipping (majority) algorithm $\mathscr{A}_M$ using the G-LDPC code to the tentative sequence.

It is also worth noting that the complexity of the proposed decoding algorithm $\mathscr{A}_C$ is $\mathcal{O}(n \log n)$ with some restrictions on the length of the linear codes with the parity-check matrix $\mathbf{H}_1$ (see Lemma 3). At the same time the complexity of ML decoding is exponential.

4. Main Results

Consider a BSC with a bit error probability p. Let a decoding error probability P be the probability of the union of decoding denial and the erroneous decoding events. In this paper, we consider the decoding error probability P in the following form

$$P \leq \exp\{-nE(\cdot)\},$$

with the $E(\cdot)$ being the required error exponent.

Let us define two error exponents: $E_C(\cdot)$ and $E_{ML}(\cdot)$ corresponding to the $\mathscr{A}_C$ decoding algorithm (having $\mathcal{O}(n \log n)$ complexity) and the $\mathscr{A}_{ML}$ decoding algorithm (having an exponential complexity) respectively. Let us consider first the error exponent $E_C(\cdot)$.

Theorem 1. *Let there exist in the ensemble $\mathscr{E}_G(\ell, n_0, b_0)$ of the G-LDPC codes a code with the code rate R_2 that can correct any error pattern of weight up to $\lfloor \omega_t n \rfloor$ while decoding with the bit-flipping (majority) algorithm $\mathscr{A}_M$.*

Let there exist a linear code with code length n_1, code rate R_1 and an error exponent under maximum likelihood decoding lower bounded by $E_0(R_1, p)$.

Then, in the ensemble $\mathscr{E}_{LG}(\ell, n_0, b_0, R_1, n_1, b_1)$ of the LG-LDPC codes, there exists a code with code length n,

$$n = n_0 b_0 = n_1 b_1,$$

code rate R,

$$R \geq R_1 + R_2 - 1,$$

and an error exponent over the memoryless BSC with BER p under the decoding algorithm $\mathscr{A}_C$ with complexity $\mathcal{O}(n \log n)$ lower bounded by $E_C(\cdot)$:

$$E_C(R_1, n_1, \omega_t, p) = \min_{\omega_t \leq \beta \leq \beta_0} \left\{ \beta E_0(R_1, p) + E_2(\beta, \omega_t, p) - \frac{1}{n_1} H(\beta) \right\}, \tag{2}$$

where $\beta_0 = \min\left(\frac{\omega_t}{2p}, 1\right)$, $H(\beta) = -\beta \ln \beta - (1-\beta)\ln(1-\beta)$ – an entropy function – and $E_2(\beta, \omega_t, p)$ is given by

$$E_2(\beta, \omega_t, p) = \frac{1}{2}\left(\omega_t \ln \frac{\omega_t}{p} + (2\beta - \omega_t) \ln \frac{2\beta - \omega_t}{1-p} \right) - \beta \ln(2\beta),$$

where n_1 satisfies the following condition:

$$\frac{-\ln \beta_0}{E_0(R_1, p)} \leq n_1 \leq \frac{1}{R_1} \log_2 \log_2(n). \tag{3}$$

Corollary 1. *$E_C(\cdot) > 0$, if $R \to C$, where C is the capacity of a memoryless BSC with error probability p, such that $R_1 \to C$ and $R_2 < 1$.*

Thus, according to Corollary 1, we can state that there exists an LG-LDPC code such that the error probability of the low-complexity decoding algorithm $\mathscr{A}_C$ exponentially decreases with the code length for all code rates below the channel capacity C.

Remark 1. *We have obtained the lower bound on $E_C(R_1, n_1, \omega_t, p)$ assuming $n \to \infty$, where $n_0 = const$, $n_1 = const$, $b_0 \to \infty$, and $b_1 \to \infty$. As a result, the complexity of $\mathscr{A}_C$ algorithm equals to $\mathcal{O}(n \log n)$, and we have the right inequality of condition (3) for n_1.*

Theorem 1 was obtained in [12]. The main idea of the proof is based on the following results. Our previous results [3,4] show that in the ensemble $\mathscr{E}_G(\ell, n_0, b_0)$ of G-LDPC codes, there exists a code that can correct any error pattern of weight up to $\lfloor \omega_t n \rfloor$ under the algorithm $\mathscr{A}_M$ with complexity $\mathcal{O}(n \log n)$. In [6], it was shown that there exists a linear code for which the error exponent under ML decoding is lower bounded by $E_0(R, p)$, where $E_0(R, p) > 0$ for $\forall R < C$.

Let us now consider the lower bound on the error exponent $E_{ML}(\cdot)$.

Theorem 2. *In the ensemble $\mathscr{E}_{LG}(\ell, n_0, b_0, R_1, n_1, b_1)$, there exists an LG-LDPC code such that the error exponent of this code over the memoryless BSC with BER p under the decoding algorithm $\mathscr{A}_{ML}$ is lower bounded by*

$$E_{ML}(p) = \max_{\omega_0 \le \omega_c \le 1} \{\min(E_\delta(\omega_c, p), E_{\omega_c}(\omega_c, p))\},$$

where $\omega_0 = \max(\delta, p)$, δ is the code distance of this code, and $E_\delta(\omega_c, p)$ is given by

$$E_\delta(\omega_c, p) = \max_{\delta \le \omega \le \omega_c} \left\{ \nu(\omega) + \omega \left(\ln 2 + \ln \sqrt{p(1-p)} \right) \right\},$$

where $\nu(\omega)$ is an asymptotic spectrum of the LG-LDPC code:

$$\nu(\omega) = \lim_{n \to \infty} \frac{\ln \overline{N}(\omega n)}{n},$$

where $\overline{N}(\omega n)$ is the average number of codewords and $E_{\omega_c}(\omega_c, p)$ is given by

$$E_{\omega_c}(\omega_c, p) = (1 - \omega_c) \ln \frac{1 - \omega_c}{1 - p} + \omega_c \ln \frac{\omega_c}{p}.$$

We have obtained Theorem 2 using the methods developed in [14] to estimate the error exponent under the ML decoding of codes with the given spectrum. We have taken ideas of [1] for G-LDPC codes to construct the upper bound on the code spectrum and the lower bound on the code distance of the proposed LDPC construction (see Lemmas 1 and 2).

Lemma 1. *The value of $\nu(\omega)$ for the codes from the ensemble $\mathscr{E}_{LG}(\ell, n_0, b_0, R_1, n_1, b_1)$ of LG-LDPC codes is upper bounded by*

$$\nu(\omega) \le \nu_0(\omega) = -(\ell - 1)H(\omega) + \min_{s>0} \left\{ \frac{\ell - 1}{n_0} \ln g_0(s, n_0) + \frac{1}{n_1} \ln g_1(s, R_1, n_1) - \omega \ell \ln s \right\},$$

where $g_0(s, n_0)$ is a spectrum function of the constituent SPC code with length n_0,

$$g_0(s, n_0) = \sum_{i=0}^{n_0} \frac{(1+s)^{n_0} + (1-s)^{n_0}}{2},$$

and $g_1(s, R_1, n_1)$ is a spectrum function of the constituent linear code with a good spectrum, code rate R_1 and length n_1 obtained in [14]:

$$g_1(s, R_1, n_1) \le 1 + n_1 2^{-(1-R_1)n_1} \sum_{i=\lceil \delta_{VG} n_1 \rceil}^{n_1} \binom{n_1}{i} s^i,$$

where δ_{VG} is given by the Varshamov-Gilbert bound.

Lemma 2. *Let the positive root δ_0 of the following equation exist:*

$$\nu_0(\delta_0) = 0.$$

Then, a code with minimum code distance $\delta \ge \delta_0$ exists in ensemble $\mathscr{E}_{LG}(\ell, n_0, b_0, R_1, n_1, b_1)$.

5. Numerical Results

One can see from Theorems 1 and 2 that the obtained lower-bounds $E_C(\cdot)$ and $E_{ML}(\cdot)$ depend on the set of parameters: error probability p of BSC, code rate R_1 and length n_1 of linear code from added layer, code rate R_2 and constituent code length n_0 of G-LDPC code (the value of ω_t, used in $E_C(\cdot)$ bound, depends on these parameters), wherein the code rate R of whole construction depends on R_1 and R_2.

Thus, to simplify the analysis let us first fix the parameters $R_1 = 0.85$, $n_1 = 2000$, $R = 0.5$ and $p = 10^{-3}$ and find how $E_C(\cdot)$ and $E_{ML}(\cdot)$ depend on the SPC code length n_0 (see Figure 1). Then, let us consider the maximized $E_{ML}(\cdot)$ and $E_C(\cdot)$ over the values of n_0 (see Figure 2).

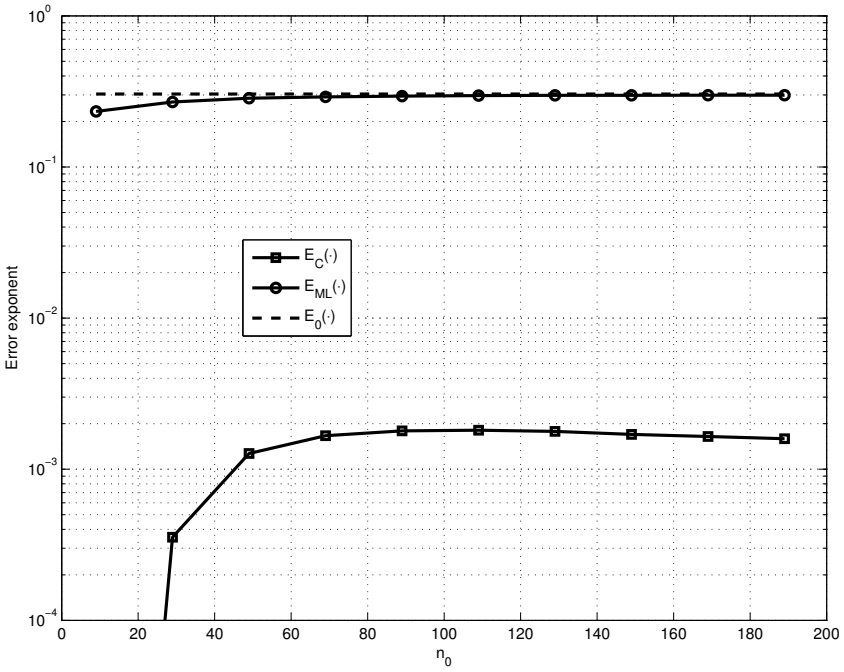

Figure 1. Comparison of the dependence on n_0 of $E_C(\cdot)$, $E_{ML}(\cdot)$ and $E_0(\cdot)$ for fixed $R_1 \approx 0.85$, $n_1 = 2000$, $R = 0.5$ and $p = 10^{-3}$.

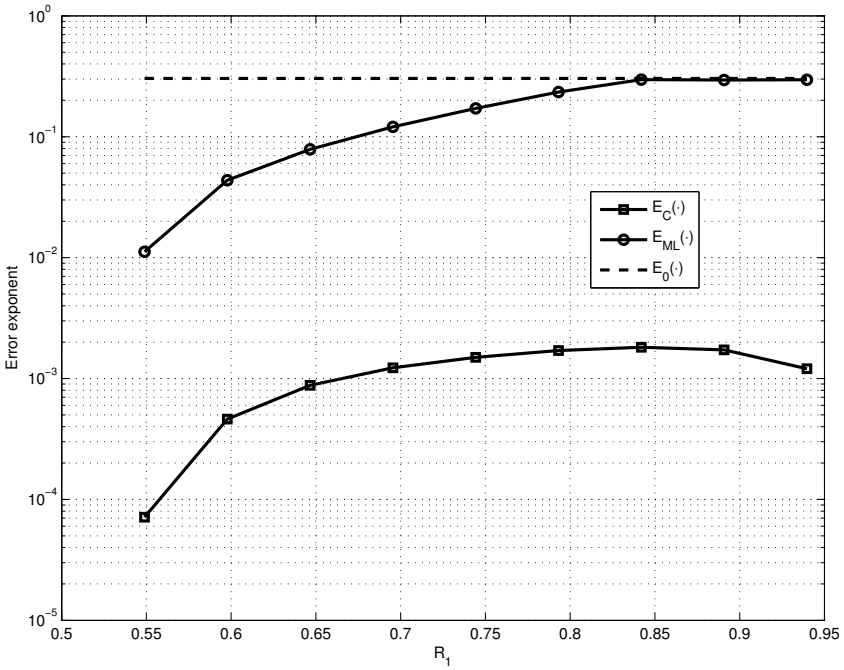

Figure 2. Comparison of the dependences on R_1 of $E_C(\cdot)$, $E_{ML}(\cdot)$ and $E_0(\cdot)$ for fixed $n_1 = 2000$, $R = 0.5$ and $p = 10^{-3}$.

We can explain the different behaviors of $E_{ML}(\cdot)$ and $E_C(\cdot)$ shown in Figures 1 and 2 by the following: the value of $E_{ML}(\cdot)$ significantly depends on the value of the code distance δ of the LG-LDPC code and the value of $E_C(\cdot)$ depends on the value of the error

fraction ω_t, which is guaranteed to be corrected by the G-LDPC code. And it is known that for the fixed code rate R code distance of LDPC code increases with the growth of constituent code length n_0 and guaranteed corrected error fraction ω_t has the maximum for the certain parameters n_0 and ℓ.

In Figure 3, we compare the dependencies on R for fixed $p = 10^{-3}$ of the obtained lower bound $E_{ML}(\cdot)$, maximized over the values of n_0 and R_1, and of the lower bound $E_0(\cdot)$.

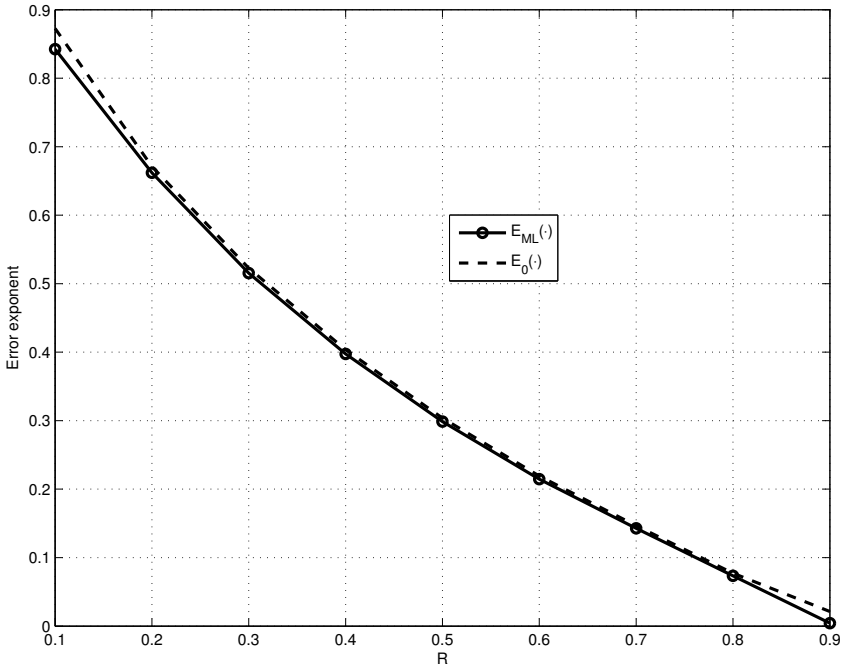

Figure 3. Comparison of the dependencies on R for fixed $p = 10^{-3}$ of $E_{ML}(\cdot)$, maximized over the values of n_0 and R_1 for fixed $n_1 = 2000$, and of $E_0(\cdot)$.

Figure 4 shows the dependencies on R of the maximum values of $E_{ML}(\cdot)$ and $E_C(\cdot)$ for fixed $p = 10^{-3}$ (the maximization was performed over the values of n_0 and R_1).

As observed from Figure 4, $E_C(\cdot)$ is approximately two orders of magnitude smaller than $E_{ML}(\cdot)$, which almost reaches the lower bound on the error exponent $E_0(\cdot)$ of the good linear code (see Figure 3). However, it is important to note that $E_{ML}(\cdot)$ encounters only exponential decoding complexity and $E_C(\cdot)$ encounters the decoding complexity of $\mathcal{O}(n \log n)$.

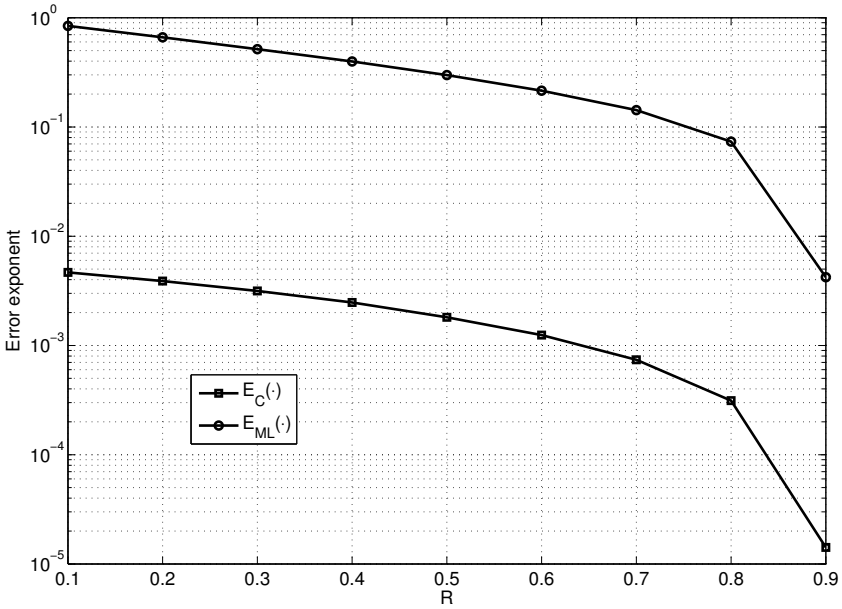

Figure 4. Comparison of the dependencies on R of $E_C(\cdot)$ and $E_{ML}(\cdot)$, maximized over the values of n_0 and R_1 for fixed $n_1 = 2000$ and $p = 10^{-3}$.

6. Conclusions

The main result of this paper is that we prove (see Corollary 1) the existence of such LDPC code with specific construction that the probability of erroneous decoding with low-complexity algorithm ($\mathcal{O}(n \log n)$) decreases exponentially with the growth of the code length for all code rates below BSC capacity. We also obtain the lower bound on error exponent under ML decoding for proposed construction (see Theorem 2) and show with numeric results that obtained lower-bound almost coincide with the error exponents of good linear codes for the certain parameters.

As a future work to improve the lower-bound for the low-complexity decoder, we plan to consider error-reducing codes instead of good linear codes and generalize our results for channels with reliability (e.g., channels with additive white Gaussian noise (AWGN) and "soft" reception).

7. Proofs of the Main Results

7.1. Error Exponent for Decoding Algorithm $\mathscr{A}_C$

Theorem 1 was proved in [12]. Here, we provide the proof for convenience of the reader in more detail, especially for the essential Corollary 1.

Let us first consider the complexity of the decoding algorithm $\mathscr{A}_C$ of an LG-LDPC code.

Lemma 3. *The complexity of the decoding algorithm $\mathscr{A}_C$ of an LG-LDPC code with length n is of order $O(n \log n)$ if the length of the linear code satisfies the inequality $n_1 \leq \frac{1}{R_1} \log_2 \log_2(n)$.*

Proof. Since the length of the linear code is equal to n_1 and the code rate is R_1, the complexity of the maximum likelihood decoding algorithm for the single code is of order $O(2^{R_1 n_1})$. The total number of codes is b_1, which is proportional to n, and then, the complexity of decoding all of the codes is of order $O(n 2^{R_1 n_1})$.

In [4], it was shown that the complexity of the bit-flipping decoding algorithm of LDPC codes is $O(n \log n)$.

Therefore, the complexity of decoding algorithm $\mathscr{A}_C$ is of order $O(n \log_2 n)$ if the following condition is satisfied:

$$n 2^{R_1 n_1} \leq n \log_2(n).$$

Here we find the condition on n_1:

$$n_1 \leq \frac{1}{R_1} \log_2 \log_2(n). \tag{4}$$

□

Let us now consider the proof of Theorem 1.

Proof. Assume that in the first step of the decoding algorithm $\mathscr{A}_C$ of LG-LDPC code, the decoding error occurred exactly in i linear codes. Since each code contains no more than n_1 errors, the total number of errors W after the first step of decoding is no greater than in_1. Let $i = \beta b_1$, where β is the fraction of linear codes in which the decoding failure occurred; then,

$$W \leq \beta b_1 n_1 = \beta n.$$

According to [4], LDPC code is capable of correcting any error pattern with weight less than W, that is,

$$W < W_0 = \lfloor \omega_t n \rfloor,$$

where ω_t is the fraction of errors guaranteed corrected by the G-LDPC code [4] (Theorem 1). Consequently, for the case where $\beta < \omega_t$, the decoding error probability P for LG-LDPC code under decoding algorithm $\mathscr{A}_C$ is equal to 0:

$$P = 0, \quad \beta < \omega_t.$$

At $\beta > \omega_t$, the error decoding probability is defined as

$$P = \sum_{i=\lfloor \omega_t b_1 \rfloor}^{b_1} \binom{b_1}{i} P_2(W \geq W_0 | i) P_1^i (1 - P_1)^{b_1 - i}, \tag{5}$$

where P_1 is the error decoding probability of linear code,

$$P_1 \leq \exp\{-n_1 E_0(R_1, p)\},$$

and $P_2(W \geq W_0 | i)$ is the probability that the number of errors after the first step of the decoding algorithm $\mathscr{A}_C$ is not less than W_0 under the condition that the decoding error occurred exactly in i linear codes.

Since the number of errors no more than doubles in a block in the case of error decoding with the maximum likelihood decoding algorithm, it must be greater than $\frac{W_0}{2}$ errors before the first step in i erroneous blocks to have more than W_0 errors after the first step of decoding algorithm $\mathscr{A}_C$. Then, we can write $P_2(W \geq W_0 | i)$ as

$$P_2(W \geq W_0 | i) = \sum_{j=\lfloor \frac{\omega_t n}{2} \rfloor}^{in_1} \binom{in_1}{j} p^j (1 - p)^{in_1 - j}.$$

Using the Chernoff bound, we obtain

$$P_2(W \geq W_0 | i) \leq \exp\{-n E_2(\beta, \omega_t, p)\},$$

where

$$E_2(\beta, \omega_t, p) = \begin{cases} \frac{1}{2}\left(\omega_t \ln \frac{\omega_t}{p} + (2\beta - \omega_t) \ln \frac{2\beta - \omega_t}{1 - p}\right) - \beta \ln 2\beta, \ \beta < \beta_0, \\ 0, \ \beta \geq \beta_0. \end{cases} \tag{6}$$

Here $\beta = \frac{i}{b_1} > \omega_t$, and

$$\beta_0 = \min\left(\frac{\omega_t}{2p}, 1\right)$$

because $\beta > 1$ has no sense.

In accordance with (6), the probability $P_2(W \geq W_0|i)$ can be replaced with the trivial estimation $P_2(W \geq W_0|i) \leq 1$ for $i \geq \lceil \beta_0 b_1 \rceil$, and then, sum (5) is upper bounded as follows:

$$P \leqslant \sum_{i=\lfloor \omega_t b_1 \rfloor}^{\lfloor \beta_0 b_1 \rfloor} \left(\begin{array}{c} b_1 \\ i \end{array} \right) P_2(W \geq W_0|i) P_1^i (1-P_1)^{b_1 - i} + \sum_{i=\lceil \beta_0 b_1 \rceil}^{b_1} \left(\begin{array}{c} b_1 \\ i \end{array} \right) P_1^i (1-P_1)^{b_1 - i}.$$

Let P_{II} denote the first sum in the right part of this inequality and P_{I} denote the second sum. Let us consider each sum separately.

The sum P_{I} can be easily estimated as a tail of the binomial distribution with probability P_1 using the Chernoff bound:

$$P_{\mathrm{I}} \leq \exp\{-nE_{\mathrm{I}}(R_1, n_1, \omega_t, p)\},$$

where

$$E_{\mathrm{I}}(R_1, n_1, p) = \beta_0 E_0(R_1, p) - \frac{1}{n_1} H(\beta_0),$$

and P_1 satisfies the condition

$$P_1(n_1, R_1, p) \leq \beta_0.$$

Thus,

$$n_1 \geq \frac{-\ln \beta_0}{E_0(R_1, p)}. \tag{7}$$

Let us now consider the sum P_{II}:

$$P_{II} \leq \lceil (\beta_0 - \omega_t) b_1 \rceil \times \max_{\omega_t \leq \beta \leq \beta_0} \left\{ \left(\begin{array}{c} b_1 \\ \beta b_1 \end{array} \right) P_2(W \geq W_0|\beta b_1) P_1^{\beta b_1} (1-P_1)^{(1-\beta)b_1} \right\}.$$

Hence, at $n \to \infty$ ($b_1 \to \infty$ and $b_0 \to \infty$), we obtain

$$E_{\mathrm{II}}(R_1, n_1, \omega_t, p) = \min_{\omega_t \leq \beta \leq \beta_0} \left\{ E_2(\beta, \omega_t, p) + \beta E_0(R_1, p) - \frac{1}{n_1} H(\beta) \right\}. \tag{8}$$

Let us note that if a minimum is achieved at β_0 in the right part of equality (8), then according to (6), we obtain $E_{\mathrm{II}} = E_{\mathrm{I}}$. Consequently, $E_{\mathrm{II}} \leq E_{\mathrm{I}}$.

It is easy to see that at $n \to \infty$, the following inequality is satisfied:

$$P \leq \exp\{-nE(R_1, n_1, \omega_t, p)\},$$

where $E(R_1, n_1, \omega_t, p) = \min\{E_{\mathrm{II}}, E_{\mathrm{I}}\} = E_{\mathrm{II}}$.

According to the proved lemma, the complexity of the decoding algorithm $\mathscr{A}_{\mathrm{C}}$ is of order $O(n \log n)$ if the condition (4) is satisfied, but for the obtained estimation, the condition (7) must also be satisfied. Thus,

$$\frac{-\ln \beta_0}{E_0(R_1, p)} \leq n_1 \leq \frac{1}{R_1} \log_2 \log_2 n.$$

This completes the proof. □

Before proving Corollary 1, we need to consider the lower-bound behavior of the error fraction ω_t guaranteed corrected by G-LDPC code. In [4], the new estimation of the error

fraction ω_t guaranteed corrected by generalized LDPC code with a given constituent code was obtained. Let us formulate this result for G-LDPC code:

Theorem 3. *Let the root ω_0 exist for the following equation:*

$$h(\omega_0) - \ell F_e(\omega_0, n_0) = 0, \tag{9}$$

where $F_e(\omega_0, n_0)$ is given by

$$F_e(\omega_0, n_0) = h(\omega_t) + \max_{s>0, 0<v<1} \left\{ \omega_0 \log_2 sv - \frac{1}{n_0} \log_2 (g_e(s, v, n_0) + g_0(s, n_0)) \right\},$$

where $g_0(s, n_0)$ and $g_e(s, v, n_0)$ have the following forms:

$$g_0(s, n_0) = \frac{(1+s)^{n_0} + (1-s)^{n_0}}{2}, g_e(s, v, n_0) = g_d\left(sv^2, n_0\right),$$

where

$$g_d(s, n_0) = (1+s)^{n_0} - g_0(s, n_0).$$

Let for the found value ω_0, the root α_0 exist for the following equation:

$$h(\omega_0) - \ell F_{\mathbf{s}}(\alpha, \omega_0, n_0, \ell) = 0, \tag{10}$$

where $F_{\mathbf{s}}(\alpha, \omega_0, n_0, \ell)$ is given by

$$F_{\mathbf{s}}(\alpha, \omega_0, n_0, \ell) = h(\omega_0) + \max_{s>0, 0<v<1} \left\{ \omega_0 \left(\log_2 s + \frac{\ell - \frac{1-\alpha}{\alpha}}{\ell} \log_2 v \right) - \frac{1}{n_0} \log_2 (g_d(s, n_0) v + g_0(s, n_0)) \right.$$

Then, there exists a code (with $p_n : \lim_{n\to\infty} p_n = 1$) in the ensemble $\mathscr{E}_G(\ell, n_0, b_0)$ of G-LDPC codes that can correct any error pattern with weight less than $\lfloor \omega_t n \rfloor$, where $\omega_t = \alpha_0 \omega_0$, with decoding complexity $\mathcal{O}(n \log n)$.

For Theorem 3, we obtain the following:

Corollary 2. *For the given code rate $R < 1$, there exists a G-LDPC code in the ensemble $\mathscr{E}_G(\ell, n_0, b_0)$ with $\ell > 2$ such that equation (9) has a positive root $\omega_t > 0$.*

The proof of Theorem 3 was given in a more generalized form in [4]. Here, we consider only the proof of Corollary 2. For this purpose, let us formulate some useful facts proved in [4].

First, let us formulate the condition of the existence of a symbol that upon inversion, reduces the number of unsatisfied checks:

Lemma 4. *At least one such symbol exists that will be inverted during one iteration of decoding algorithm $\mathscr{A}_M$ for G-LDPC code if the following condition is satisfied:*

$$E_{\Sigma}^{(W)} = 2 \sum_{j=1}^{W} e_{A_{1\to 0}}^{(i_j)} + \sum_{j=1}^{W} e_{A_{1\to 1}}^{(i_j)} > W\ell, \tag{11}$$

where W is the number of errors in the received sequence, $i_1, i_2, ..., i_W$ are indices of erroneous symbols, $e_{A_{1\to 0}}^{(i)}$ is the number of edges emanating from the ith variable-node to the set of check-nodes for which the checks become satisfied after the inversion of this symbol, and $e_{A_{1\to 1}}^{(i)}$ is the number of the edges emanating from the ith variable-node to the set of check-nodes for which the checks remain unsatisfied after the inversion of this symbol.

Now, let us consider the estimation of the probability that the above condition is not satisfied:

Lemma 5. *The probability* $P_W\left(E_\Sigma^{(W)} \leqslant W\ell\right)$ *for the fixed pattern of errors of weight* W *that condition (11) is not satisfied, e.g.,* $E_\Sigma^{(W)} \leqslant W\ell$*, is upper bounded as follows:*

$$P_W\left(E_\Sigma^{(W)} \leqslant W\ell\right) \leqslant 2^{-n\ell F_e(\omega,n_0)+o(n)},\ \omega = \frac{W}{n}.$$

Now, let us consider the proof of Corollary 2.

Proof. Let us select an arbitrary small value ε' and write the following condition:

$$\lim_{n\to\infty} \sum_{W=1}^{\lfloor \varepsilon' n \rfloor} 2^{-n\left(\ell F_e\left(\frac{W}{n},n_0\right)-h\left(\frac{W}{n}\right)\right)} < 1.$$

In the left part of the inequality is the upper bound on the probability of the code that condition (11) is not satisfied for some sequences.

Let us introduce the following function $G(\omega)$:

$$G(\omega) = \ell F_e(\omega,n_0) - h(\omega) = (\ell-1)h(\omega) + \ell \max_{s>0,0<v<1}\left\{\omega\log_2 sv - \frac{1}{n_0}\log_2\left(g_e(s,v,n_0)+g_0(s,n_0)\right)\right.$$

Since the variables s and v are dummies, they can be equal to an arbitrary value if the conditions $s > 0$ and $0 < v < 1$ are satisfied. Then, let us set $s = v = \sqrt[4]{\omega}$ (this choice is justified by the fact that due to the structure of the parity-check matrix, the conditions $\ell > 2$ should be satisfied):

$$G^*(\omega) = (\ell-1)h(\omega) + \ell\left(\frac{\omega}{2}\log_2\omega - \frac{1}{n_0}\log_2\left(g_e\left(\sqrt[4]{\omega},\sqrt[4]{\omega},n_0\right)+g_0\left(\sqrt[4]{\omega},n_0\right)\right)\right).$$

Let us transform $G^*(\omega)$ as follows:

$$G^*(\omega) = -\left(\frac{\ell}{2}-1\right)\omega\log_2\omega - (\ell-1)(1-\omega)\log_2(1-\omega) - \frac{\ell}{n_0}\log_2\left(g_e\left(\sqrt[4]{\omega},\sqrt[4]{\omega},n_0\right)+g_0\left(\sqrt[4]{\omega},n_0\right)\right.$$

It is easy to show that $g_e(s,v,n_0)+g_0(s,n_0) \leqslant (1+s)^{n_0}$ for $0 < s < 1$ and $0 < v < 1$. Then, we obtain

$$G^*(\omega) = -\left(\frac{\ell}{2}-1\right)\omega\log_2\omega + \mathcal{O}(\omega).$$

It is easily noted that $G(\omega) \geqslant G^*(\omega)$ implies

$$\lim_{n\to\infty}\sum_{W=1}^{\lfloor \varepsilon' n \rfloor} 2^{-nG\left(\frac{W}{n}\right)} \leqslant \lim_{n\to\infty}\sum_{W=1}^{\lfloor \varepsilon' n \rfloor} 2^{-nG^*\left(\frac{W}{n}\right)}.$$

Since LDPC code construction requires $\ell > 2$, $\left(\frac{\ell}{2}-1\right) > 0$, and consequently,

$$G(\omega) \geqslant -c_1\omega\log_2\omega + c_2\omega + o(\omega),\ c_1 > 0.$$

$$\lim_{n\to\infty}\sum_{W=1}^{\lfloor \varepsilon' n \rfloor} 2^{-nG\left(\frac{W}{n}\right)} \leqslant \lim_{n\to\infty}\sum_{W=1}^{\lfloor \varepsilon' n \rfloor} 2^{n\cdot c_1\cdot\frac{W}{n}\cdot\log_2\frac{W}{n} - n\cdot c_2\cdot\frac{W}{n}} = \lim_{n\to\infty}\sum_{W=1}^{\lfloor \varepsilon' n \rfloor}\left(\frac{W}{n}\right)^{c_1 W} 2^{-c_2 W} \leqslant$$

$$\leqslant \lim_{n\to\infty} \sum_{W=1}^{\lfloor \varepsilon' n \rfloor} \left((\varepsilon')^{c_1} 2^{-c_2} \right)^W = \frac{(\varepsilon')^{c_1} 2^{-c_2}}{1 - (\varepsilon')^{c_1} 2^{-c_2}} = \varepsilon''.$$

It should be noted that the sign of c_2 is not important because ε'' can be made arbitrarily small by a correct choice of ε'.

Thus,

$$\lim_{n\to\infty} \sum_{W=1}^{\lfloor \varepsilon' n \rfloor} 2^{-n\left(\ell F_e\left(\frac{W}{n}, n_0\right) - h\left(\frac{W}{n}\right)\right)} \leqslant \varepsilon'' < 1.$$

Consequently, the code for which the condition (11) is satisfied for all values of $\omega_t < \varepsilon'$ exists with non-zero probability in the ensemble of G-LDPC codes. □

Finally, let us consider the proof of Corollary 1.

Proof. The correctness of the corollary is easy to see if we note that $E_0(\cdot) > 0$ for $R_1 < C$ [6] and $E_2(\cdot) \geq 0$, which follows from (6), and we can always select n_1 such that $\frac{1}{n_1} H(\beta) < \beta E_0(\cdot) + E_2(\cdot)$ because n_1 can be arbitrarily large according to condition (3). Therefore, according to Corollary 2, the construction of G-LDPC code with $\omega_t > 0$ for any code rate $R_2 < 1$ exists, helping us omit this condition in the corollary formulation (unlike the formulation of a similar corollary in [12]). □

7.2. Error Exponent for Decoding Algorithm $\mathscr{A}_{ML}$

Let us consider the proof of Theorem 2.

Proof. To simplify the proof without loss of generality, let us consider the transmission of a zero codeword over the BSC with BER p. Let the probability of the transition of the zero codeword to each of $N(w)$ codewords with weight w during decoding with algorithm $\mathscr{A}_{ML}$ be equal to $P_\delta(w)$. Moreover, let there exist a critical value w_c of the number of errors that leads to erroneous decoding with algorithm $\mathscr{A}_{ML}$. Then, we can write

$$P_{ML} = \sum_{w=d}^{w_c} N(w) P_\delta(w) + P(w \geq w_c),$$

where d is the code distance of the LG-LDPC code.

To obtain the upper bound, it is sufficient to consider the case when the zero codeword becomes the word with weight w if there are more than $w/2$ errors:

$$P_\delta(w) = \sum_{i=\frac{w}{2}}^{w} \begin{pmatrix} w \\ i \end{pmatrix} p^i (1-p)^{w-i} \leq 2^{w-1} p^{\frac{w}{2}} (1-p)^{\frac{w}{2}},$$

where $p \leq \frac{1}{2}$.

From this inequality, we easily obtain

$$E_\delta(\omega_c, p) = \max_{\delta \leq \omega \leq \omega_c} \left\{ \nu(\omega) + \omega \left(\ln 2 + \ln \sqrt{p(1-p)} \right) \right\},$$

where $\nu(\omega)$ is an asymptotic spectrum of the LG-LDPC code given by Lemma 1 and δ is the relative code distance of the LG-LDPC code given by Lemma 2.

With the help of the Chernoff bound, we obtain the exponent of the probability that more than w_c errors have occurred:

$$P(w \geq w_c) \leq \exp\{-n E_{\omega_c}(\omega_c, p)\}.$$

$$E_{\omega_c}(\omega_c, p) = (1-\omega_c) \ln \frac{1-\omega_c}{1-p} + \omega_c \ln \frac{\omega_c}{p}, \omega_c \geq p.$$

Consequently,

$$E_{ML}(p) = \max_{\omega_0 \le \omega_c \le 1} \{\min(E_\delta(\omega_c, p), E_{\omega_c}(\omega_c, p))\},$$
$$\omega_0 = \max(\delta, p).$$

□

The estimations given in Lemmas 1 and 2 were obtained by the slightly modified classical Gallager's method [1]. Thus, in this paper, we give only a sketch of the proof.

Let us first consider the proof of Lemma 1.

Proof. Let us consider the fixed word of weight W and find the probability of there being a code in the LG-LDPC code ensemble such that this word is a codeword for this code. For this purpose, let us consider the first layer of the parity-check matrix of some LG-LDPC code from the ensemble composed of the parity-check matrices of the single parity check code. We can write the probability that the considered word is a codeword for a given layer as follows:

$$P_W^{(1)} = \frac{N_1(W)}{\binom{n}{W}},$$

where $N_1(W)$ is the number of layers, and the word of weight W is a codeword.

We estimate $N_1(W)$ as

$$N_1(W) \leqslant \min_{s>0} \left\{ \frac{g_0^{b_0}(s, n_0)}{s^W} \right\},$$

where $g_0(s, n_0)$ is a spectrum function of the SPC code.

Thus,

$$P_W^{(1)} \leqslant \binom{n}{w}^{-1} \min_{s>0} \left\{ g_0^{b_0}(s, n_0) s^{-W} \right\}.$$

It is clear that the obtained estimation is the same for all $\ell - 1$ layers:

$$P_W^{(i)} \leqslant \binom{n}{w}^{-1} \min_{s>0} \left\{ g_0^{b_0}(s, n_0) s^{-W} \right\}, i = 1...\ell - 1.$$

Similarly, we can write the probability that the considered word of weight W is a codeword for the ℓth layer of the parity-check matrix composed of "optimal" linear codes:

$$P_W^{(\ell)} \leqslant \binom{n}{w}^{-1} \min_{s>0} \left\{ g_1^{b_1}(s, R_1, n_1) s^{-W} \right\},$$

where $g_1(s, R_1, n_1)$ is a spectrum of the code with a good spectrum.

Since the layer permutations are independent, we can write the probability that the given word of weight W is a codeword for the whole code construction as

$$P_W = \prod_{i=1}^{\ell} P_W^{(i)} \leqslant \binom{n}{W}^{-\ell} \min_{s>0} \left\{ g_0^{b_0(\ell-1)}(s, n_0) g_1^{b_1}(s, R_1, n_1) s^{-W\ell} \right\}.$$

Consequently, the average number of weight W codewords is given by

$$\bar{N}(W) = \binom{n}{W} P_W \leqslant \binom{n}{W}^{-(\ell-1)} \min_{s>0} \left\{ g_0^{b_0(\ell-1)}(s, n_0) g_1^{b_1}(s, R_1, n_1) s^{-W\ell} \right\}.$$

For $W = \omega n$, we obtain

$$\nu(\omega) = \lim_{n\to\infty} \frac{\ln N(\omega n)}{n} \leqslant \nu_0(\omega).$$

□

Now, let us consider the proof of Lemma 2.

Proof. If the average number of codewords $\bar{N}(W)$ in the ensemble of LG-LDPC codes satisfies the condition

$$\sum_{W=1}^{d_0} \bar{N}(W) \leqslant 1,$$

then the code with code distance $d \geqslant d_0$ exists in this ensemble.

It is easy to show that the sum of the right part of the inequality can be estimated with the last member of this sum. Therefore, using the estimation obtained in the previous lemma, we can write

$$\nu_0(\delta) \leqslant 0,$$

where $\delta = d/n$ is the relative code distance.

Thus, we can obtain the maximum value of δ_0 such that the above-considered condition is satisfied for all smaller values $\delta \leqslant \delta_0 0$ as the smallest positive root of the following equation:

$$\nu_0(\delta_0) = 0.$$

□

Author Contributions: Conceptualization, V.Z.; Investigation, P.R. and K.A.; Writing—original draft, P.R.; writing—review and editing, K.A. All authors have read and agreed to the published version of the manuscript.

Funding: The reported study was funded by RFBR, project number 19-37-51036.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Gallager, R.G. *Low-Density Parity-Check Codes*; MIT Press: Cambridge, MA, USA, 1963.
2. Zyablov, V.; Pinsker, M. Estimation of the error-correction complexity for Gallager low-density codes. *Probl. Inf. Transm.* **1975**, *11*, 18–28.
3. Rybin, P.; Zyablov, V. Asymptotic estimation of error fraction corrected by binary LDPC code. In Proceedings of the 2011 IEEE International Symposium on Information Theory Proceedings, St. Petersburg, Russia, 31 July–5 August 2011; pp. 351–355. [CrossRef]
4. Zyablov, V.; Rybin, P. Analysis of the relation between properties of LDPC codes and the Tanner graph. *Probl. Inf. Transm.* **2012**, *48*, 297–323. [CrossRef]
5. Rybin, P. On the error-correcting capabilities of low-complexity decoded irregular LDPC codes. In Proceedings of the 2014 IEEE International Symposium on Information Theory, Honolulu, HI, USA, 29 June–4 July 2014; pp. 3165–3169. [CrossRef]
6. Gallager, R.G. *Information Theory and Reliable Communication*; John Wiley & Sons, Inc.: New York, NY, USA, 1968.
7. Barg, A.; Zemor, G. Error exponents of expander codes. *IEEE Trans. Inf. Theory* **2002**, *48*, 1725–1729. [CrossRef]
8. Barg, A.; Zémor, G. Error Exponents of Expander Codes Under Linear-Complexity Decoding. *SIAM J. Discret. Math.* **2004**, *17*, 426–445. [CrossRef]
9. Burshtein, D.; Barak, O. Upper bounds on the error exponents of LDPC code ensembles. In Proceedings of the 2006 IEEE International Symposium on Information Theory, Seattle, WA, USA, 9–14 July 2006; pp. 401–405. [CrossRef]
10. Barak, O.; Burshtein, D. Lower bounds on the error rate of LDPC code ensembles. *IEEE Trans. Inf. Theory* **2007**, *53*, 4225–4236. [CrossRef]
11. Rybin, P.; Frolov, A. On the Error Exponents of Capacity Approaching Construction of LDPC code. In Proceedings of the 2018 10th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT), Moscow, Russia, 5–9 November 2018; pp. 1–5. [CrossRef]
12. Zyablov, V.; Rybin, P. Estimation of the exponent of the decoding error probability for a special generalized LDPC code. *J. Commun. Technol. Electron.* **2012**, *57*, 946–952. [CrossRef]

13. Rybin, P.S.; Zyablov, V.V. Asymptotic bounds on the decoding error probability for two ensembles of LDPC codes. *Probl. Inf. Transm.* **2015**, *51*, 205–216. [CrossRef]
14. Blokh, E.; Zyablov, V. *Linear Concatenated Codes*; Nauka: Moscow, Russia, 1982.