ISSN 1064-2269, Journal of Communications Technology and Electronics, 2018, Vol. 63, No. 12, pp. 1518–1523. © Pleiades Publishing, Inc., 2018.
Original Russian Text © P.S. Rybin, F.I. Ivanov, 2017, published in Informatsionnye Protsessy, 2017, Vol. 17, No. 4, pp. 284–293.

On Estimation of the Error Exponent for Finite Length Regular Graph-Based LDPC Codes[1]

P. S. Rybin[a, b, *] and F. I. Ivanov[a, b, **]

[a]*Kharkevich Institute for Information Transmission Problems, Russian Academy of Sciences, Moscow, 127051 Russia*
[b]*National Research University Higher School of Economics, Moscow, 101000 Russia*
**e-mail: prybin@iitp.ru*
***e-mail: fii@iitp.ru*

Received December 15, 2017

Abstract—The error exponent of the regular graph-based binary low-density parity-check (LDPC) codes under the maximum likelihood (ML) decoding algorithm in the binary symmetric channel (BSC) is analyzed. Unlike most other papers where error exponents are considered for the case when the length of LDPC codes tends to infinity (asymptotic analysis), the finite length case (finite length analysis) is considered. In this paper, a method of deriving the lower bound on the error exponent for a regular graph-based LDPC code with finite length under ML decoding is described. Also we analyze Dependences of the error exponent on various LDPC code parameters are also analyzed. The numerical results obtained for the considered lower bound are represented and analyzed at the end of the paper.

Keywords: LDPC code, error exponent, finite length

DOI: 10.1134/S1064226918120197

1. INTRODUCTION

The low-density parity-check (LDPC) code was proposed by Gallager in [1], where the lower bound on the code distance was also obtained. In [2] and [3], the asymptotic upper and lower bounds on the error exponent under the maximum likelihood (ML) decoding of LDPC codes were investigated. It is also known from [2] that the lower bound on the error exponent under the ML decoding of LDPC codes asymptotically tends to the lower bound on the error exponent under the ML decoding of good linear codes obtained in [4].

It should be noted that all works mentioned above, where error exponents were considered, were obtained for the asymptotic case: it was assumed that the code length tends to infinity. This assumption allows one to use some well-known methods of asymptotic behavior estimation.

There is a limited number of papers where nonasymptotic behavior of LDPC codes is considered. In [5], the performance of finite length LDPC codes in the waterfall onset region (such values of the signal-to-noise ratio for which the output probability of error per codeword after decoding begins to decrease) was considered. The authors proposed an algorithm (based on the density evolution technique) that predicts the error performance of finite-length LDPC codes transmitted through various binary symmetric memoryless channels (BSMCs). Using a combinatorial characterization of decoding failures, expressions for the average bit and block erasure probabilities were derived in [6], [7], and [8] for various LDPC ensembles and iterative message-passing decoding algorithms. Some modifications of these methods was presented in [9]. Most of these works deal with only a binary erasure channel and low-complexity decoding (e.g., bit-flipping or belief propagation). Only a few papers are devoted to channels with errors.

This work was inspired partly by studies [10] and [11], where methods for the analysis of graph-based LDPC codes were developed. It should also be noted that the average weight distribution of finite-length LDPC codes was estimated in [7] but, in this paper, an error exponent was not considered. In this paper, we analyze the decoding performance of regular LDPC codes with finite length under the conditions of the ML decoding algorithm. We describe a method of deriving the lower bound on the error exponent for the LDPC code with this decoding algorithm using finite length analysis based on generating functions and some other combinatorial methods. The error exponents are computed numerically and analyzed for various code parameters.

2. CONSTRUCTION OF LDPC CODES

Sometimes it is convenient to represent LDPC codes in terms of the Tanner graph [12]. This is a

[1] The article was translated by the authors.

bipartite graph, where $m > 0$ check nodes (parity-checks) are associated with the $n > m$ variable nodes (codeword bits), where n is the length of the corresponding LDPC code. Although only regular LDPC codes are considered in this paper, we provide a more general way for description of LDPC code ensembles. In this case, any LDPC code can be described in terms of bipartite graphs that are characterized by two probability vectors

$$\tilde{\lambda} = (\tilde{\lambda}_2, \ldots, \tilde{\lambda}_l),$$

$$\tilde{\rho} = (\tilde{\rho}_1, \ldots, \tilde{\rho}_{n_0}),$$

where $\tilde{\lambda}_k$ is the portion of variable nodes with the degree k, and $\tilde{\rho}_t$ is the fraction of check nodes with degree t. For convenience, we also define the polynomials

$$\tilde{\lambda}(x) = \sum_{k=2}^{l} \tilde{\lambda}_k x^{k-1},$$

$$\tilde{\rho}(x) = \sum_{t=2}^{n_0} \tilde{\rho}_t x^{t-1}.$$

Let E denotes the total number of edges, n denotes the number of variable nodes and m denotes the number of check nodes. Then $n\tilde{\lambda}_k k$ is equal to the number of edges outgoing from variable nodes with the degree k and $m\tilde{\rho}_t t$ is equal to the number of edges emanating from check nodes with the degree t. Thus,

$$n = \frac{E}{\sum_{k=2}^{l} \tilde{\lambda}_k k} = \frac{E}{1 + \tilde{\lambda}'(1)},$$

$$m = \frac{E}{\sum_{t=2}^{n_0} \tilde{\rho}_t t} = \frac{E}{1 + \tilde{\rho}'(1)},$$

where $\tilde{\lambda}'(1)$ and $\tilde{\rho}'(1)$ are derivatives of functions $\tilde{\lambda}(x)$ and $\tilde{\rho}(x)$ of variable x calculated at the point $x = 1$.

For each variable node with degree i, we assign i variable sockets. Similarly, for each check node with the degree i, we assign i check sockets. The total number of variable sockets and the total number of check sockets are both equal to the total number of edges E. The ensemble of bipartite graphs is obtained by choosing randompermutation π with the uniform probability from the space of all permutations of size E. For each $1 \le i \le E$, we connect the variable node associated with the ith variable socket to the check node associated with the π_ith check socket. Note that, in this way, multiple edges may link a pair of nodes. The mapping from the bipartite graph space into the space of the parity-check matrix $\mathbf{H}$ of the LDPC code is such that element $\mathbf{H}_{i,j}$ in the matrix corresponding to the ith check node and the jth variable node is set to 1 if there is an odd number of edges between the two nodes and to 0 otherwise.

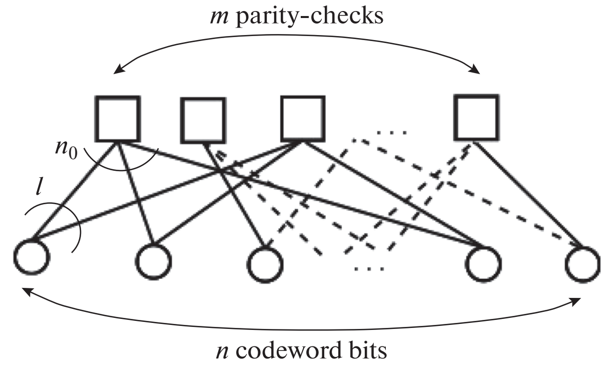

Fig. 1. Tanner graph of the regular (2,3) LDPC code.

The rate R of each code in the ensemble satisfies inequality $R \ge R'$, where

$$R' = 1 - \frac{m}{n} = 1 - \frac{\sum_{k=2}^{l} \tilde{\lambda}_k k}{\sum_{t=2}^{n_0} \tilde{\rho}_t t} = 1 - \frac{1 + \tilde{\lambda}(1)}{1 + \tilde{\rho}'(1)}$$

is the designed rate of the code (the inequality is due to a possible degeneracy in m parity-check equations of matrix $\mathbf{H}$).

A special case of the irregular code ensemble that was described above is obtained when all variable nodes have a constant degree l and all check nodes have a constant degree n_0. In this case, the ensemble is called regular, and there is constant $m \in \mathbb{N}$ such that $nl = mn_0$, $R \ge 1 - l/n_0$, and

$$\tilde{\lambda}(x) = x^{l-1},$$

$$\tilde{\rho}(x) = x^{n_0-1}.$$

In this paper, only regular LDPC codes are considered. The Tanner graph of a regular (l, n_0) LDPC code is presented in the Fig. 1.

3. MAIN RESULT

Investigating the decoding error probability P, we will consider a memoryless binary-symmetric channel (BSC) with crossover probability p. The decoding error probability P is estimated in the following way:

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

where $E(\cdot)$ is the required error exponent.

In this paper, we consider the error exponent $E(\cdot)$ of (l, n_0) regular LDPC code of length n under the ML decoding algorithm, which has exponential complexity.

According to [10], we can write the following:

Theorem 1. *The average number $\bar{N}(W)$ of codewords of weight W of the irregular LDPC codes determined by polynomials $(\tilde{\lambda}(x), \tilde{\rho}(x))$ satisfies the relationship:*

$$\bar{N}(W) = \sum_{j=0}^{E} \frac{t(W,j)q(j)}{\binom{E}{j}}, \qquad (1)$$

where t(W, j) is defined in the following way:

$$t(W,j) = \left(\prod_k (1+xy^k)^{\tilde{\lambda}_k n}\right)\left[x^W \;\; y^j\right],$$

where notation $f_i = f(x)[x^i]$ means that $f(x) = \sum_i f_i x^i$, and q(j) is given by

$$q(j) = \left(\prod_k g_0^{\tilde{\rho}_k(1-R)n}(x,k)\right)\left[x^j\right],$$

where $g_0 = \frac{1}{2}\left((1+x)^k + (1-x)^k\right)$.

In the case of regular (l, n_0) LDPC code, we can easily calculate the exact value of $t(W,j)$. Indeed, since $\tilde{\lambda}_k = 1$ when $k = l$ and $\tilde{\lambda}_k = 0$ otherwise, then

$$t(W,j) = (1+xy^l)^n\left[x^W \;\; y^j\right],$$

$$(1+xy^l)^n\left[x^W \;\; y^j\right] = \sum_{k=0}^{n}\binom{n}{k}x^k y^{lk}\left[x^W \;\; y^j\right],$$

$$t(W,j) = \begin{cases}\binom{n}{W}, & j = lW\\ 0, & j \neq lW.\end{cases}$$

We can also simplify the expression for $q(j)$:

$$q(j) = (q_0^m(x,n_0))\left[x^j\right],$$

$$q(j) = \left(\sum_{k=0}^{\left\lfloor\frac{n_0}{2}\right\rfloor}\binom{n_0}{2k}x^{2k}\right)^m\left[x^{lW}\right].$$

Thus,

$$\bar{N}(W) = \frac{\binom{n}{W}q(lW)}{\binom{E}{lW}}. \qquad (2)$$

Due to the fact that it is hard to perform exact calculation of $q(lW)$, we will use the following theorem [13] to estimate it:

Theorem 2 (Hayman, 1956). *Let $f(x) = \sum_k a_k x^k$ and r_k is a positive root of equation $a(r_k) = k$ for all k = 1, 2, ..., where $a(x) = x\frac{f'(x)}{f(x)}$.*

Then

$$a_k \approx \frac{f(r_k)}{r_k^k\sqrt{2\pi b(r_k)}}, \qquad (3)$$

where $b(x) = xa'(x)$.

In our case, $f(x) = (g_0(x, n_0))^m$ and $k = lW$. Therefore,

$$a(x) = x\frac{m(g_0(x,n_0))^{m-1}g_0'(x,n_0)}{(g_0(x,n_0))^m}.$$

By definition $f(x) \neq 0$ for $x \geq 0$, then

$$a(x) = mx\frac{g_0'(x,n_0)}{g_0(x,n_0)}.$$

Thus,

$$mr_{lW}\frac{g_0'(r_{lW},n_0)}{g_0(r_{lW},n_0)} = lW.$$

Finally, we obtain

$$q(lW) \leq a_{lW},$$

where a_{lW} is obtained from (3). Thereby, we can estimate the average number $\bar{N}(W)$ of codewords of weight W of a regular (l, n_0) LDPC code as

$$\bar{N}(W) \leq \frac{a_{lW}\binom{n}{W}}{\binom{E}{lW}}. \qquad (4)$$

Inequality (4) allows us to estimate not only the spectrum of the LDPC code but also its relative minimum distance $\delta = \frac{d}{n}$. We will present corresponding results in the next section.

Now let us denote the average spectrum function $\psi(z)$ for a (l, n_0) regular LDPC code as

$$\psi(z) = \sum_{W=d}^{n}\bar{N}(W)z^W.$$

Applying (4), we can obtain the following obvious estimate for $\psi(z)$:

$$\psi(z) \leq n\sum_{W=d}^{n}\bar{N}(W)z^W \leq n\sum_{W=d}^{n}\frac{a_{lW}\binom{n}{W}}{\binom{E}{lW}}z^W.$$

This estimate of the spectrum function will be used in the following theorem obtained in [14], which helps us to estimate the error probability of a given (l, n_0) regular LDPC code under the ML decoding.

Theorem 3. *The error probability P(R, p) of a linear code with spectrum $\psi(z)$, rate R, and length n under ML* *decoding in a BSC with crossover probability p is upper-bounded in the following way:*

$$P(R,p) \le \exp\{-nE(R,p)\},$$

where $E(R,p)$ is given by:

$$E(R,p) = \max_{s\ge 0, t\le 0, r\le 0}\left\{\frac{r}{s-r}\ln g(s) - \frac{s}{s-r}\left(\ln g(r) + \frac{1}{n}\ln\psi\left(\frac{g(r,t)}{g(r)}\right)\right)\right\},$$

where $g(s)$ and $g(r,t)$ are defined as follows:

$$g(s) = (1-p)^{1-s} + p^{1-s},$$

$$g(r,t) = (1-p)^{1-r+t}p^{-t} + (1-p)^{-t}p^{1-r+t}.$$

4. NUMERICAL RESULTS

First of all let us present an estimate of the average number of codewords $\bar{N}(W)$ of weight W of regular LDPC codes with fixed parameters l and n_0 but different lengths n. For convenient representation of numerical results, we introduce a special function $\nu(\omega) = \frac{\log\bar{N}(\omega n)}{n}$, where $\omega = \frac{W}{n}$, $0 \le \omega \le 1$.

In Fig. 2, we can see the obtained spectrum estimate in the case of $(l, n_0) = (7, 14)$ regular LDPC codes with different lengths $n = 100$, $n = 300$, and $n = 1000$ and fixed rate $R = 0.5$. One can see that $\nu(\omega)$ grows with the growth of code length n.

Figure 3 illustrates dependences of $\nu(\omega)$ on l of for the fixed length $n = 300$ and rate $R = 0.5$.

One can notice that $\nu(\omega)$ grows with decreasing l (especially for small and large ω). This fact means that the relative minimum distance $\delta = \frac{d}{n}$ is a function of l, which monotonically grows with l (for fixed n and R). This fact is more pronounced in Fig. 4.

Numerical results presented in Fig. 4 also allow us to conclude that the relative distance decreases with code length n. One can also notice that LDPC codes with parameters $n \approx 100$, $R = 0.5$ and $l \in \{5, 7, 9\}$ have $\delta \approx 0.12$ while in the asymptotic case $\delta_{VG}(0.5) \approx 0.11$, where $\delta_{VG}(R)$ is the Varshamov–Gilbert bound. Thus, in the case of small lengths, these codes lie above the Varshamov–Gilbert bound.

Let us now present numerical results for error exponent $E(R,p)$, which was obtained according to Theorem 3.

Figure 5 presents the dependence of the error exponent $E(R,p)$ on the length of the LDPC code n and the channel crossover probability p for the fixed parameters $l = 7$, $n_0 = 14$, and $R = 0.5$. We can note that the error exponent grows with n and decreases with increasing p (when n is fixed).

Figure 6 presents the dependence of $E(R,p)$ on n and l for fixed rate $R = 0.5$ and the channel crossover

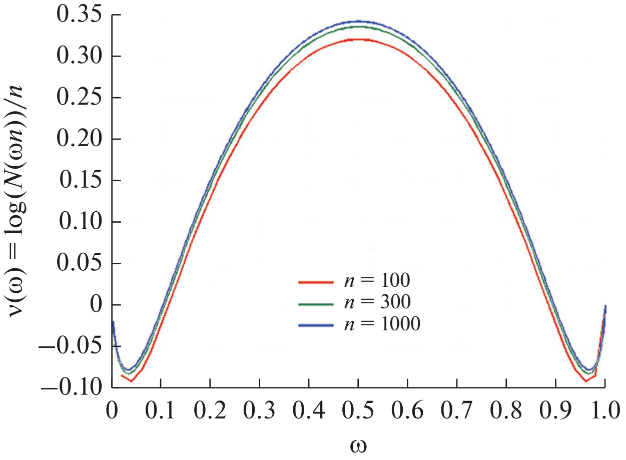

Fig. 2. Comparison of the dependences of $\nu(\omega)$ on n for fixed $l = 7$, $n_0 = 14$, and $R = 0.5$.

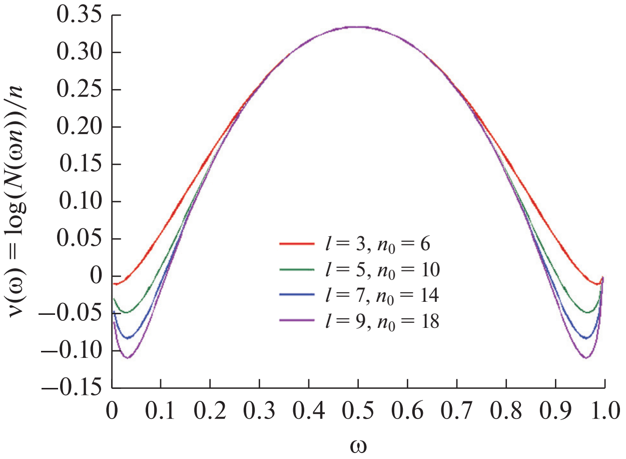

Fig. 3. Comparison of the dependences of $\nu(\omega)$ on l for fixed $n = 300$ and $R = 0.5$.

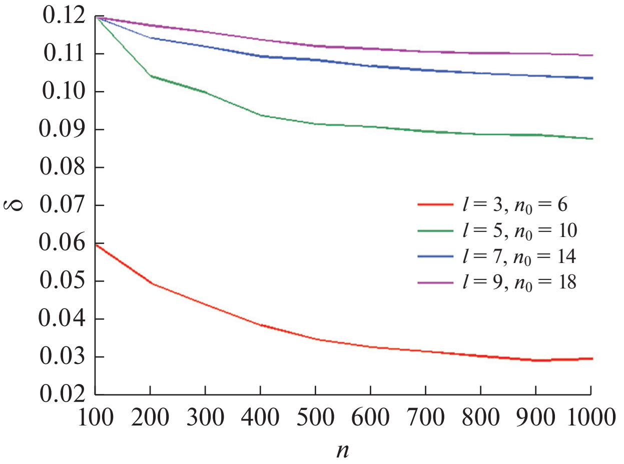

Fig. 4. Comparison of the dependences of δ on l for fixed $R = 0.5$ and different n.

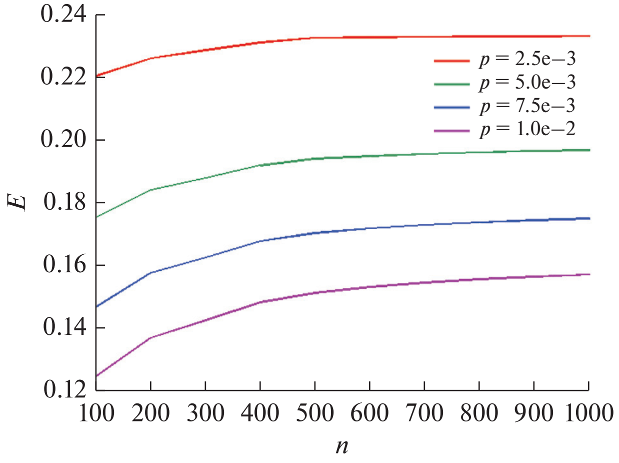

Fig. 5. Comparison of the dependences of $E(R, p)$ on n and p for fixed $R = 0.5$ and $l = 7$.

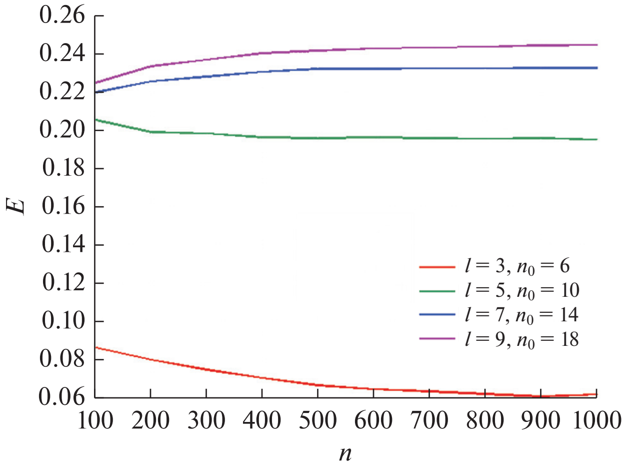

Fig. 6. Comparison of the dependences of $E(R, p)$ on n and l for fixed $R = 0.5$ and $p = 2.5 \times 10^{-3}$.

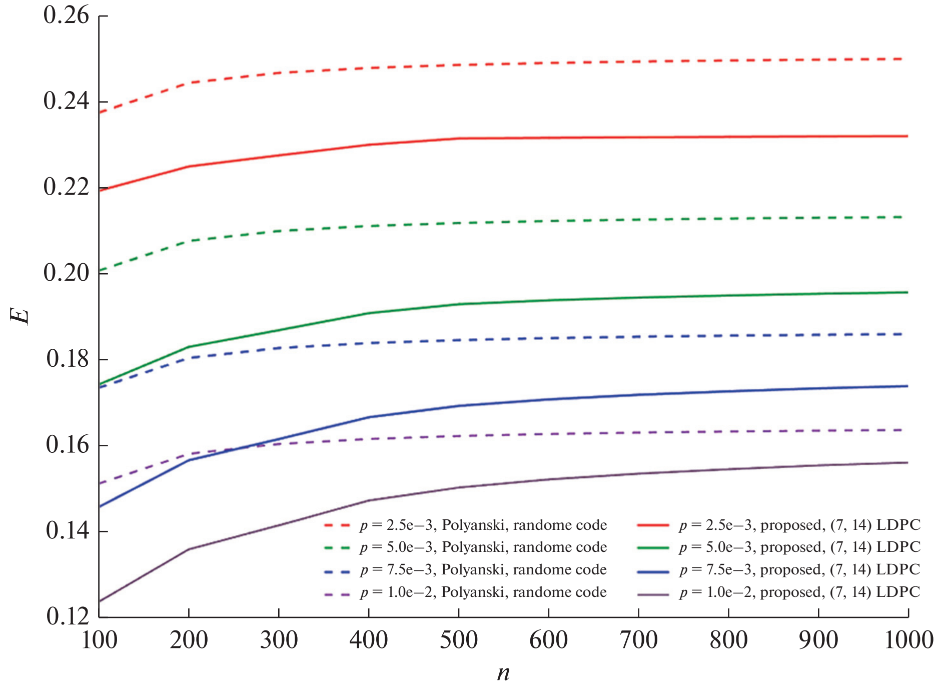

Fig. 7. Comparison of the dependences of $E(R, p)$ on n and p for fixed $R = 0.5$ and $l = 7$ for (7,14) LDPC codes and random binary codes.

probability $p = 2.5 \times 10^{-3}$. We can note that the error exponent decreases with n for $l = 3$ or $l = 5$ and increases with n for LDPC codes with $l = 7$ or $l = 9$. This fact is explained by significant reduction of the relative minimum distance for codes with $l = 3$ and $l = 5$ when the length is growing (see Fig. 4).

At the end of this chapter, we present the results of comparison between error exponents obtained in this paper (solid lines) and the results obtained with the use of the Polyansky bound [15] for random binary codes (dotted lines). These comparison results are shown in Fig. 7. Regular (7,14) LDPC codes with $R = 0.5$ and lengths from 100 to 1000 were chosen for this comparison. Comparison was performed for various input error probabilities of the channel.

It is easy to see that the Polyansky bound for random codes is always higher than the bound obtained in this paper. This fact suggests that regular LDPC codes (especially with short lengths) have less potential corrective performance than arbitrary binary codes.

5. CONCLUSIONS

It has been described how to estimate the spectrum function of a regular graph-based finite-length binary LDPC code. It has been shown that this estimate is useful for deriving results regarding the lower bound on the error exponent of considered codes under ML decoding in the BSC channel. Numerical results presented in this paper show that (3,6)-regular LDPC codes that are widely used in practice due to their good iterative decoding performance have the worst error exponent under the ML decoding. Moreover, this exponent decreases with increasing code length. This means that these codes should be substituted with some other codes when we deal with applications where very small error probability or very long codes are required.

ACKNOWLEDGMENTS

This study was carried out at the Kharkevich Institute for Information Transmission Problems, Russian Academy of Sciences, and was supported by the Russian Science Foundation, project no. 14-50-00150.

REFERENCES

1. R. G. Gallager, *Low-Density Parity-Check Codes* (MIT Press, Cambridge, MA, USA, 1963).
2. O. Barak and D. Burshtein, "Lower bounds on the error rate of LDPC code ensembles," IEEE Trans. Inf. Theory **53**, 4225–4236 (2007).
3. D. Burshtein and O. Barak, "Upper bounds on the error exponents of LDPC code ensembles," in *Proc. 2006 IEEE Int. Symp. on Information Theory (ISIT), Seattle, USA, 2006* (IEEE, New York, 2006), pp. 401–405.
4. R. G. Gallager, *Information Theory and Reliable Communication* (Wiley, New York, 1968).
5. M. Rahim, K. D. Nguyen, and G. Lechner, "Finite length analysis of LDPC codes," in *Proc. IEEE Wireless Comm. and Networking Conf. (WCNC), Istambul, Turkey, 2014* (IEEE, New York, 2014), pp. 206–211.
6. C. D. Proietti, I. E. Telatar, T. J. Richardson, and R. L. Urbanke, "Finite-length analysis of low-density parity-check codes on the binary erasure channel," IEEE Trans. Inf. Theory **48**, 1570–1579 (2002).
7. T. J. Richardson and R. L. Urbanke, "Finite-length density evolution and the distribution of the number of iterations on the binary erasure channel," Unpublished manuscript, available at http://lthcwww.epfl.ch/papers/RiU02.psi.
8. H. Zhang and A. Orlitsky, "Finite-length analysis of LDPC codes with large left degrees," in *Proc. Int. Symp.on Inf. Theory (ISIT), Lausanne, Switzerland, June 30–July 5, 2002* (IEEE, New York, 2002), p. 3.
9. S. J. Johnson, "A finite-length algorithm for LDPC codes without repeated edges on the binary erasure channel," IEEE Trans Inf. Theory **55**, 27–32 (2009).
10. D. Burshtein and G. Miller, "Asymptotic enumeration methods for analyzing LDPC codes," IEEE Trans. Inf. Theory **50**, 1115–1131 (2004).
11. P. S. Rybin, "On the error-correcting capabilities of low-complexity decoded irregular LDPC codes," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT), Hawaii, USA, 2014* (IEEE, New York, 2014), pp. 3165–3169.
12. R. Tanner, "A recursive approach to low complexity codes," IEEE Trans. Inf. Theory **27**, 533–547 (1981).
13. H. S. Wilf, *Generatingfunctionology* (Academic, USA, 1992).
14. E. L. Blokh and V. V. Zyablov, *Linear Concatenated Codes* (Nauka, Moscow, 1982) [in Russian].
15. Y. Polyanskiy, H. V. Poor, and S. Verdu, "Channel coding rate in the finite blocklength regime," IEEE Trans. Inf. Theory **56**, 2307–2359 (2010).