ML and Near-ML Decoding Performance of LDPC Codes over BEC: Bounds and Decoding Algorithms

Irina E. Bocharova, *Senior Member, IEEE,* Boris D. Kudryashov, *Senior Member, IEEE,*
Vitaly Skachek, *Member, IEEE,* Eirik Rosnes, *Senior Member, IEEE,*
and Øyvind Ytrehus, *Senior Member, IEEE*

Abstract

Performance of maximum-likelihood (ML) decoding on the binary erasure channel for finite length LDPC codes from two random ensembles is studied. The theoretical average spectrum of the Gallager ensemble is computed by using a recurrent procedure and compared to the empirically found average spectrum for the same ensemble as well as to the empirical average spectrum of the Richardson-Urbanke ensemble and spectra of selected codes from both ensembles. Distance properties of the random codes from the Gallager ensemble are discussed. A tightened union-type upper bound on the ML decoding error probability based on the precise coefficients of the average spectrum is presented. A new upper bound on the ML decoding performance of the LDPC codes from the Gallager ensemble based on computing rank of submatrices of the code parity-check matrix is derived. A new low-complexity near-ML decoding algorithm for quasi-cyclic LDPC codes is proposed and simulated. Its performance is compared to the upper bounds on the ML decoding performance.

I. Introduction

A binary erasure channel (BEC) is one of the simplest for analysis and consequently well-studied models of communication channels. In spite of its simplicity, during the last decades the BEC started playing more important role thanks to the emergence of new applications. For example, in communication networks virtually all errors occurring at physical level can be detected using a rather small redundancy. Data packets with detected errors can be viewed as symbol erasures.

As it was already mentioned, an important reason for the BEC popularity is that analysis of decoding performance over the BEC is simpler than for other channel models. On the other hand, it is expected that ideas and findings for the BEC might be useful for constructing codes and developing decoding algorithms for other important communication channels, such as, for example, binary symmetric channel (BSC) or additive white Gaussian noise (AWGN) channels.

A remarkable property of the BEC is that maximum-likelihood (ML)-decoding of any linear code over this channel is reduced to solving a system of linear equations. This means that ML decoding of an $[n, k]$ LDPC code (where n is the code length and k is the number of information symbols) with ν erasures can be performed by Gaussian elimination with time complexity at most $O(\nu^3)$. Exploiting the sparsity of the parity-check matrix of the LDPC codes can lower the complexity to approximately $O(\nu^2)$ (see overview and analysis in [2] and references therein). Practically feasible algorithms with a thorough complexity analysis can be found in [3]–[5]. This makes ML-decoded LDPC codes strong competitors for scenarios with strong restrictions on the decoding delay. It is worth noting that ML decoding allows for achieving low error probability at rates strictly above so-called belief propagation (BP) decoding thresholds [6].

However, ML decoding of long LDPC codes of lengths of a few thousand bits over the BEC is still considered impractical.

Low-complexity suboptimal decoding techniques for LDPC codes over a BEC are based on the following two approaches. The first approach consists of adding redundant rows to the original code parity-check matrix (see, for example, [**?**], [7], [8]). The second approach implies a post-processing in case of BP decoding failure [9]–[11].

In [9], [12] a post-processing step is based on the concept of guessing the bits. Applying bit guessing based algorithms to decoding of LDPC codes improves the performance of BP decoding, but does not provide near-ML performance. A new low-complexity near-ML decoding algorithm is presented in this paper.

A commonly used approach to the analysis of decoding techniques is based on studying a behavior of random linear codes over a given channel model. As for the ensembles of LDPC codes, two most intensively studied code ensembles are the classical Gallager ensemble [13] and more general ensemble presented in [6]. The Gallager ensemble is historically the first thoroughly studied ensemble of regular LDPC codes. The ensemble in [6] can be described by random Tanner graphs with given parity-check and symbol node degree distributions. We call this ensemble and LDPC codes from this ensemble *Richardson-Urbanke*

I. Bocharova and B. Kudryashov are with Dept. of Information Systems, St.-Petersburg University of Information Technologies, Mechanics and Optics, St.-Petersburg, 197101, Russia, (e-mail: {iebocharova, bdkudryashov}@corp.ifmo.ru).

I. Bocharova and V. Skachek are with Institute of Computer Science University of Tartu (e-mail: {irinaboc, vitaly}@ut.ee).

E. Rosnes is with Simula@UiB, and Ø. Ytrehus is with Simula@UiB and with University of Bergen (e-mail: {eirik, oyvind}@ii.uib.no).

This paper was presented in part at the 9th International Symposium on Turbo Codes and Iterative Information Processing, Brest, France, September, 2016 [1] and at the IEEE International Symposium on Information Theory, Aachen, Germany, June 2017 [**?**].

This work is supported in part by the Norwegian-Estonian Research Cooperation Programme under the grant EMP133, and by the Estonian Research Council under the grant PUT405.

(RU) ensemble and RU LDPC codes, respectively. Several ensembles of regular LDPC codes are described and analyzed in [14].

It is shown in [13, Appendix B] that asymptotic weight enumerators of the random (J, K)-regular LDPC codes approach asymptotic weight enumerators of random linear codes if J and K grow. Similar behavior for other ensembles of regular LDPC codes is confirmed in [14]. Thus, regular ensembles are good enough to achieve near optimal performance.

On the other hand, it is well known that both asymptotically [6] and in a finite length regime irregular LDPC codes outperform their regular counterparts and more often are recommended for real-life applications [15], [16]. Finite-length analysis of the RU LDPC codes under BP and (to a lesser degree) under ML decoding was performed in [17]. In this paper we further develop ML decoding case for regular codes. New error probability bounds for regular codes are presented.

For general linear codes the detailed overviews of lower and upper bounds for the AWGN channel and BSC are presented by Sason and Shamai in their tutorial [18] and for the AWGN channel, BSC, and BEC in Polyanskiy et al. [19] and by Di et al. in [17].

For computing upper bounds for LDPC codes there exist two approaches. One approach is based on the union-type bound. It requires knowledge of the code weight enumerators or their estimates. The second approach used in [17] implies estimating rank of submatrices of an LDPC code parity-check matrix.

Notice, that for infinitely long codes, the analysis of the BER performance of BP decoding based on the density evolution (see e.g. [6], [20]) can be applied. However, the density evolution technique is not suitable for analysis of finite length codes since dependencies caused by cycles in the Tanner graph associated with the code are not taken into account.

In this paper, first we consider a tightened union-type bound based on precise average spectra of random finite length LDPC codes. The difference between our approach and other techniques is the way of calculating the bound. Instead of manipulating with hardly computable coefficients of series expansions of generating functions we compute the spectra by using efficient recurrent procedures. This allows for obtaining precise average weight enumerators with complexity growing linearly with the code length. New bounds, based on computing rank, are derived for the RU and the Gallager ensembles of regular LDPC codes.

As it is mentioned above, in this paper, we propose a decoding algorithm which provides near-ML decoding of long quasi-cyclic (QC) LDPC block codes. The decoding complexity is polynomial in the window length, but it is linear in the code length. It is well known that a QC LDPC block code can be represented as a tail-biting (TB) parent convolutional LDPC code. Thereby, decoding techniques developed for TB codes are applicable to QC LDPC codes. The proposed algorithm resembles a wrap-around suboptimal decoding of TB convolutional codes [21], [22]. Decoding of a TB code requires identification of the correct starting state, and, thus, in the ML decoding the Viterbi algorithm has to be applied to each possible starting state. In contrast, wrap-around decoding applies the Viterbi algorithm once to the *wrapped-around* trellis diagram with all starting state metrics initialized to zero. This decoding approach yields near-ML performance at a typical complexity of a few times the complexity of the Viterbi algorithm.

The new algorithm is based on a combination of BP decoding of the QC LDPC code followed by so-called "quasi-cyclic sliding-window" ML (SWML) decoding. The latter technique is applied "quasi-cyclically" to a relatively short sliding window, where the decoder performs ML decoding of a zero-tail terminated (ZT) LDPC convolutional code. Notice that unlike sliding-window near-ML decoding of convolutional codes considered in [23], the suggested algorithm working on the parent LDPC convolutional code has significantly lower computational complexity due to the sparsity of the code parity-check matrix [24]. On the other hand, it preserves almost all advantages of the convolutional structure in the sense of erasure correcting capability.

The rest of the paper is organized as follows. Preliminaries are given in Section II. A recurrent algorithm for computing average spectrum for the Gallager ensemble of binary LDPC codes is presented in Section III. Empirical average spectra for the Gallager and the Richardson-Urbanke ensembles are computed and compared to the theoretical average spectra as well as to the spectra of selected codes in the same section. Distance properties of the Gallager ensemble are discussed in Appendix A. In Section IV, two types of upper bounds on the error probability of ML decoding over the BEC are considered. The corresponding proofs are given in Appendices B and C for the RU and the Gallager ensembles, respectively. A new algorithm for near ML decoding for long QC LDPC codes based on interpretation of these codes as TB convolutional codes and using wrap-around sliding-window decoding is proposed in Section V. Simulation results in Section VI confirm the efficiency of the algorithm. Asymptotic thresholds which stem from the derived rank bounds are presented in Section VII.

Conclusions are drawn.

Remark: The following text will be removed from the final version.

Previously unpublished contributions appearing in this paper are the following:

- Numerical comparison of the precise average spectra of short codes from the RU and the Gallager ensembles with Gallager's bound and random coding bound for linear codes.
- Lower bound on error probability on the BEC for codes with known estimate on the minimum distance.
- New rank-computing based upper bound on the error probability on the BEC for the Gallager ensemble
- Comparison of the newly derived bounds with estimated average error probability over the sets of randomly generated codes from the ensembles.

- More thorough simulation results of the new decoding algorithm and performance comparison with newly derived error probability bounds.

II. Preliminaries

In this paper, we study both average performance of ensembles of random LDPC codes determined by randomly constructed matrices H and QC LDPC codes widely used in practical schemes.

For a binary linear $[n, k]$ code $\mathcal{C}$ of rate $R = k/n$ let $r = n - k$ be the code redundancy. Denote by $\{A_{n,w}\}$, $w = 0, 1, ..., n$, the code weight enumerators, where $A_{n,w}$ is the number of codewords of weight w. By an abuse of notation, we use $A_{n,w}$ for both weight enumerators of specific code and for random weight enumerators in code ensembles.

We denote by H an $r \times n$ parity-check matrix which determines $\mathcal{C}$. Random matrices H of size $r \times n$ do not necessarily have full rank $\rho = r$ which means that "true" dimension of the code is equal to $k' = n - \rho \geq k$. Following commonly accepted assumption we ignore the difference between k' and k when deriving bounds on the error probability, but we take it into account when considering code examples.

Two ensembles of random regular LDPC codes are studied below. First one is the Gallager ensemble [13] of (J, K)-regular LDPC codes. Codes of this ensemble are determined by random parity-check matrices H which consist of the strips H_i of width $M = r/J$ rows each, $i = 1, 2, \ldots, J$. All strips are random column permutations of the strip where the jth row contains K ones in positions $(j-1)K + 1, (j-1)K + 2, \ldots, jK$, for $j = 1, 2, \ldots, n/K$.

The second ensemble of the RU (J, K)-LDPC codes is a special case of the ensemble described in [6, Definition 3.15]. Notice that the Gallager ensemble and the RU ensemble are called in [14] the ensemble $\mathcal{B}$ and $\mathcal{H}$, respectively.

For $a \in \{1, 2, ...\}$ denote by a^m the sequence $(a, a, ..., a)$ of m identical symbols a. In order to construct an $r \times n$ parity-check matrix H of an LDPC code from the RU ensemble one does the following

- construct the sequence $\boldsymbol{a} = (1^J, 2^J, ..., n^J)$;
- apply a random permutation $\boldsymbol{b} = \pi(\boldsymbol{a})$ to obtain a sequence $\boldsymbol{b} = (b_1, ..., b_N)$, where $N = Kr = Jn$;
- elements $b_1, ..., b_K$ show locations of nonzero elements of the first row of H, elements $b_{K+1}, ..., b_{2K}$ show locations of nonzero elements of the second row of H, etc.

A code from the RU ensemble is (J, K)-regular if for a given permutation π all elements of subsequences $(b_{iK-K+1}, ..., b_{iK})$ are different for all $i = 1, ..., r$, otherwise it is irregular. The regular RU codes belong to the ensemble $\mathcal{A}$ in [14] which is defined by equiprobable parity-check matrices with row weight K and column weight J. It is shown in [14] that the three ensembles $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{H}$, have the same asymptotic average weight enumerators.

It is known (see [14, Theorem 3]) that for large n the total number of (J, K)-regular $[n, n-r]$-codes (ensemble $\mathcal{A}$ in [14]) is equal to

$$\frac{(Jn)!}{(K!)^r (J!)^n} \exp\left\{-\frac{(K-1)(J-1)}{2}\right\} \left(1 + o(n^{-1+\delta})\right) \ ,$$

where $\delta > 0$ and $o(x) \to 0$ if $x \to 0$. At the same time the number of different codes from the RU ensemble constructed as described above is

$$\frac{(Jn)!}{(K!)^r (J!)^n} \ .$$

Thus, a fraction of (J, K)-regular LDPC codes in the RU ensemble is

$$\exp\left\{-\frac{(K-1)(J-1)}{2}\right\} \left(1 + o(n^{-1+\delta})\right) \ ,$$

that is, most of the "(J, K)-regular" RU codes are indeed irregular.

As a performance measure we use word (block, frame) error rate (FER) P_e, which for the BEC is defined as the probability that the decoder cannot recover the information of a received word uniquely.

Consider ML decoding over the BEC. Let $\varepsilon > 0$ denote the channel symbol erasure probability. Assume that a codeword $\boldsymbol{x} = (x_1, ..., x_n)$ is transmitted and that ν erasures occurred. Then we denote by I a set of indices of the erased positions, that is, $I = (i_1, ..., i_\nu) \subseteq \{1, 2, ..., n\}$, $|I| = \nu$ and by $x_I = (x_{i_1}, ..., x_{i_\nu})$ a set of unknowns corresponding to the erased positions. Let $I^{\mathrm{c}} = \{1, 2, ..., n\} \setminus I$ and $x_{I^{\mathrm{c}}}$ be a set of indices of unerased positions and a set of values in the unerased positions of $\boldsymbol{x}$, respectively.

We denote by H_I a submatrix of H consisting of columns indexed by I. From $\boldsymbol{x} H^{\mathrm{T}} = \boldsymbol{0}$ follows

$$\boldsymbol{x}_I H_I^{\mathrm{T}} = \boldsymbol{x}_{I^{\mathrm{c}}} H_{I^{\mathrm{c}}}^{\mathrm{T}} \triangleq \boldsymbol{s} \ , \tag{1}$$

where $\boldsymbol{s}$ is the syndrome vector or, equivalently,

$$\boldsymbol{y}_I H_I^{\mathrm{T}} = \boldsymbol{0} \ , \tag{2}$$

where $\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{x}'$ is a codeword.

If a solution of (2) is not unique, that is,

$$\text{rank}\,(H_I) < |I|\,, \tag{3}$$

then the corresponding set of erasures cannot be (uniquely) corrected. Otherwise, the set of erasures I is correctable. Thus, the ML decoding error probability (for the BEC) is the probability such that a set of erasures I satisfies

$$P_e = \Pr\left(\text{rank}\,(H_I) < |I|\right)\,. \tag{4}$$

III. Average spectra for ensembles of regular LDPC codes

A. Weight enumerator generating functions

In this section, we study average weight enumerators for different ensembles of LDPC codes. The weight distribution of any linear code can be represented via its weight generating function

$$G_n(s) = \sum_{w=0}^{n} A_{n,w} s^w\,,$$

where $A_{n,w}$ is the random variable representing the number of binary codewords of weight w and length n, and s is a formal variable. Our goal is to find $\mathrm{E}\{A_{n,w}\}$, where $\mathrm{E}\{\cdot\}$ is the expected value over the code ensemble. In general, computing coefficients $A_{n,w}$ is a rather difficult task. If a generating function can be represented as a degree of another generating function (or expressed via a degree of such function) then for numerical computations we can use the following simple recursion.

Lemma 1: Let $f(s) = \sum_{l\geq 0} f_l s^l$ be a generating function. Then the coefficients in series expansion of the generating function $F_L(s) = [f(s)]^L = \sum_{l\geq 0} F_{l,L} s^l$ satisfy the following recurrent equation

$$F_{l,L} = \begin{cases} f_l, & L = 1 \\ \sum_{i=0}^{l} f_i F_{l-i,L-1}, & L > 1 \end{cases}\,. \tag{5}$$

B. General linear codes

For completeness, we present the average spectrum for the ensemble of random linear codes determined by equiprobable $r \times n$ parity-check matrices, where $r = n - k$, k and n are the code dimension and length, respectively. The weight generating function of all binary sequences of length n is $G_n(s) = (1+s)^n$. Then the average spectrum coefficients are

$$\mathrm{E}\{A_{n,w}\} = \binom{n}{w} 2^{-r}\,,\ w > 0, \tag{6}$$

where 2^{-r} is the probability that a binary sequence $\boldsymbol{x}$ of length n and weight $w > 0$ satisfies $\boldsymbol{x}H^{\mathrm{T}} = \mathbf{0}$.

If a random linear code contains only codewords of even weight then its generating function has the form

$$G_n(s) = \sum_{w\ \text{even}} \binom{n}{w} s^w = \frac{(1+s)^n + (1-s)^n}{2}$$

and the average spectrum coefficients are

$$\mathrm{E}\{A_{n,w}\} = \begin{cases} 2^{-r+1}\binom{n}{w}, & w \text{ is even }, w > 0 \\ 0, & w \text{ is odd.} \end{cases} \tag{7}$$

C. The Gallager binary (J, K)-regular random LDPC codes

Generating functions of the number of sequences satisfying the nonzero part of one parity check are equal

$$g(s) \quad = \quad \sum_{i\ \text{even}} \binom{K}{i} s^i = \frac{1}{2}\left[(1+s)^K + (1-s)^K\right]\,. \tag{8}$$

The generating function for the strip is

$$G_{J,K}(s) = g(s)^M = \sum_{w=0}^{n} N_{n,w} s^w, \tag{9}$$

where $N_{n,w}$ denotes the total number of binary sequences of weight w satisfying $\boldsymbol{x}H_1^{\mathrm{T}} = 0^M$. Due to Lemma 1 we can compute $N_{n,w}$ precisely. The probability that $\boldsymbol{x}H_1^{\mathrm{T}} = 0^M$ is valid for random binary $\boldsymbol{x}$ of weight w is equal to

$$p_1(w) = \frac{N_{n,w}}{\binom{n}{w}}\,.$$

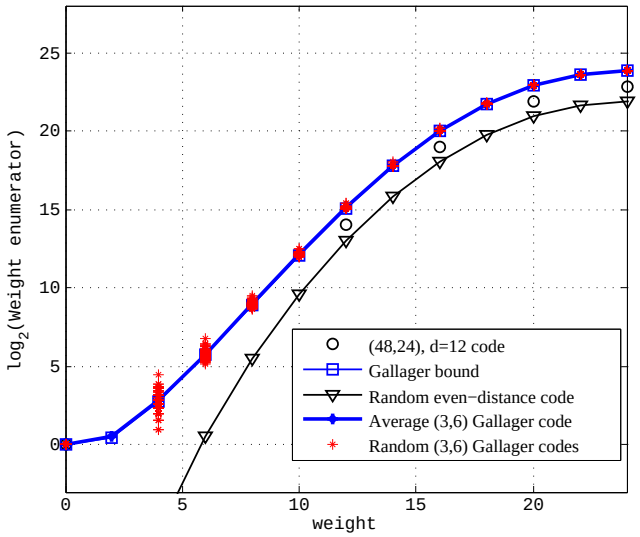

Fig. 1. The theoretical and empirical spectra of the Gallager $(3, 6)$-regular LDPC codes of length $n = 48$. The Gallager bound and the random coding bound are defined by (10) and (7), respectively.

Since submatrices H_j, $j = 1, ..., J$ are obtained by independent random column permutations of H_1, the expected number of codewords among $\binom{n}{w}$ sequences of weight w is

$$\mathrm{E}\{A_{n,w}\} = \binom{n}{w} p_1(w)^J = \binom{n}{w}^{1-J} N_{n,w}^J \ , \tag{10}$$

where $N_{n,w}$ are computed recursively using Lemma 1.

D. The empirical and theoretical average spectra for the Gallager and the Richardson-Urbanke ensembles

In this section, we compare the theoretical average spectra computed according to (7) and (10) with the empirically obtained average spectra for the Gallager and the RU ensembles. Furthermore, we compare the average spectra with the spectra of both randomly generated and constructed LDPC and linear block codes.

We remark that there is a weakness in the Gallager proof of Theorem 2.4 in [13], which is similar to that in the derivations (8) – (10) above. Formula (2.17) in [13] (and (10) in this paper) states that the average number of weight w binary sequences which satisfy parity checks of all J strips simultaneously is obtained by computing the Jth degree of $N_{n,w}/\binom{n}{w}$, that is, the probability of weight w vectors satisfying the parity checks of the first strip. This formula relies on the assumption that parity checks of strips are statistically independent. Strictly speaking, this statement is not true because they are always linearly dependent (sum of parity checks of any two strips is all-zero sequence). The detailed discussion and examples can be found in the Appendix A.

Since in our derivations of the bounds on the error probability in Section IV we rely on the same assumption it is important to compare the empirical and the theoretical average spectra. Moreover, as it is shown in Appendix A, in the Gallager ensemble there is no code whose spectrum coincides with the average spectrum. Thus, estimating deviation of the spectrum of a particular code from the average spectrum is an interesting issue.

One more question that we try to answer in this section is how close are the average spectra for the Gallager and the RU ensembles. It is known [14] that the Gallager and the RU ensembles have the same asymptotic average spectrum. However, the relations between these spectra for finite lengths are unknown.

In Figs. 1-4, the distance spectra of 100 randomly selected rate $R = \frac{1}{2}$ codes of length $n = 48$ ("Random codes" in the legends) and their empirical average spectrum ("Average" in the legends) are compared with the theoretical average spectra. We take into account that all codewords of a (J, K)-regular LDPC code from the Gallager ensemble have even weight. If K is even then the all-one codeword belongs to any code from the ensemble. It is easy to see that in this case weight enumerators $A_{n,w}$ are symmetric functions of w, i.e. $A_{n,w} = A_{n,n-w}$. By using this symmetry, we show only half of the spectra in these figures.

In Figs. 1, 2 we present the average spectra for the Gallager ensembles and the average spectrum for the ensemble of random linear codes with only even-weight codewords, computed by using formulas from subsection III-B, spectra of 100 random codes from the Gallager ensemble and the empirical average spectrum computed over 100 random codes from the same ensemble. The spectrum of quasi-perfect (48,24), $d = 12$ linear code is presented in the same figures as well. In Figs.

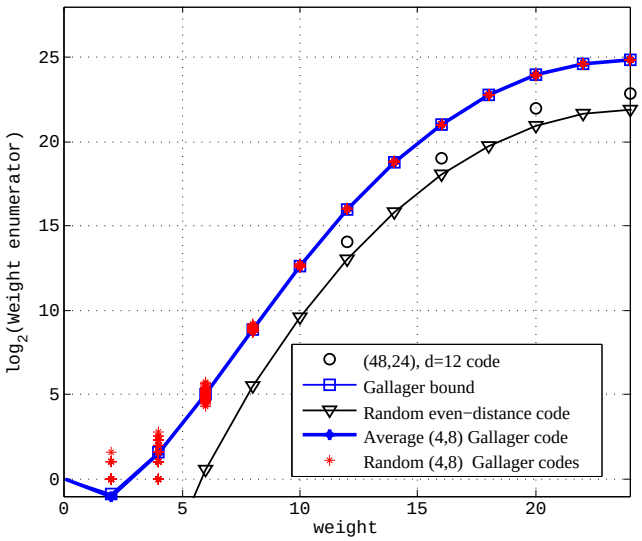

Fig. 2. The theoretical and empirical spectra of the Gallager $(4, 8)$-regular LDPC codes of length $n = 48$. The Gallager bound and the random coding bound are defined by (10) and (7), respectively.

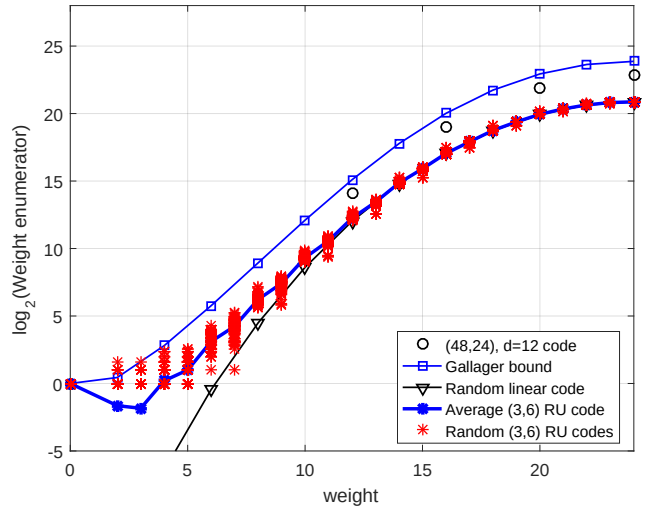

Fig. 3. The theoretical spectra of the Gallager $(3, 6)$-regular LDPC codes and the empirical spectra of the RU $(3, 6)$-regular LDPC codes of length $n = 48$. The Gallager bound and the random coding bound are defined by (10) and (7), respectively.

3, 4 the average spectrum for the corresponding Gallager ensemble and the average spectrum for the ensemble of random linear codes with only even-weight codewords are compared with spectra of 100 random codes from the RU ensemble and the empirical average spectrum computed over 100 random codes from the RU ensemble.

Observations regarding the Gallager LDPC codes:

- For the Gallager $(3, 6)$ and $(4, 8)$-regular LDPC codes their empirical average spectra perfectly match with the theoretical average spectra computed for the corresponding Gallager ensembles.
- For all codes from the Gallager ensemble the number of high-weight codewords is perfectly predicted by the theoretical average spectrum.
- The number of low-weight codewords has large variation.

Remarks about the RU codes:

- Most of the RU LDPC codes are irregular and have codewords of both even and odd weight.
- Typically, parity-check matrices of random codes from the RU ensemble have full rank and these codes have lower rate than LDPC codes from the Gallager ensemble. For this reason, the empirical average spectrum of the RU ensemble lies below the theoretical average spectrum computed for the Gallager ensemble.
- Average distance properties of the RU codes are better than those of the Gallager codes.

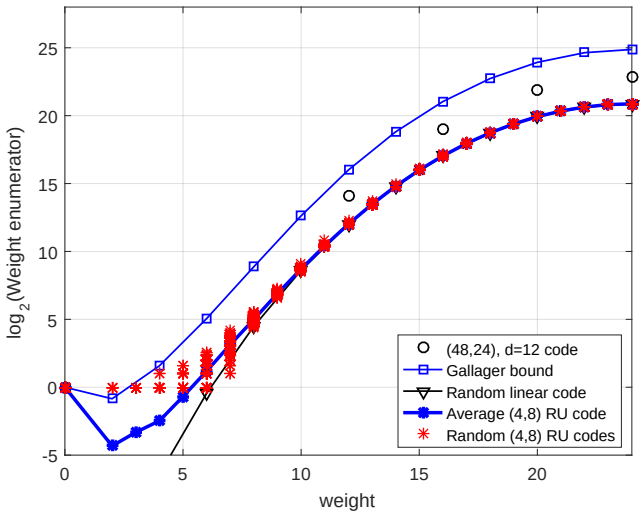

Fig. 4. The theoretical spectra of the Gallager $(4, 8)$-regular LDPC codes and the empirical spectra of the RU $(4, 8)$-regular LDPC codes of length $n = 48$. The Gallager bound and the random coding bound are defined by (10) and (7), respectively.

- Variation of the number of low-weight codewords is even larger than that for the Gallager codes.

Since for all considered ensembles the low-weight codewords have much larger probability than for the general linear codes, we expect to observe the phenomenon of the error floor.

IV. Error probability bounds on a BEC

A. Lower bounds

In this section, we consider bounds on the ML decoding error probability. We start with a simple lower bound which is true for any linear code

Theorem 1:

$$P_e \geq P_e(n, k, \varepsilon) \triangleq \sum_{i=r+1}^{n} \binom{n}{i} \varepsilon^i (1-\varepsilon)^{n-i} \,. \tag{11}$$

Remark: In [19, Theorem 38] the lower bound on the error probability of ML decoding is given. It differs from (11) by a multiplier which is close to 1. This difference appears because of different definitions of the frame error event in this paper and in [19, Theorem 38].

Proof. It readily follows from the definition of the decoding error probability and from condition in (4) that if the number of erasures $\nu > r \geq \text{rank}\,(H)$ then the decoding error probability is equal to one. □

The bound (11) ignores erasure combinations of weight less than or equal to r. Such combinations lead to an error if they cover all nonzero elements of a codeword.

Theorem 2: Let the code minimum distance be $d_{\min} \leq d_0$. Then,

$$P_e \geq P_e(n, k, \varepsilon) + \sum_{w=d_0}^{r} \binom{n-d_0}{w-d_0} \varepsilon^w (1-\varepsilon)^{n-w} \,. \tag{12}$$

Proof. There is at least one nonzero codeword $\boldsymbol{c}_0$ of weight at most d_0 in the code. Each erasure combination of weight $w \geq d_0$ which covers nonzero positions of $\boldsymbol{c}_0$ leads to additional decoder failures taken into account as the sum in the RHS of (12). □

We remark that the upper bounds on the minimum distance of linear codes with given $n \leq 256$ and k can be found in [25]. Lower bounds (11) and (12) for some code examples are plotted in Figs. 5, 6, 7 and 13 and discussed in Section VI.

B. Upper bounds for general linear codes

Next, we consider the ensemble-average ML decoding block error probability $\mathrm{E}\{P_e\}$ over the BEC with erasure probability $\varepsilon > 0$. This average decoding error probability can be interpreted as an upper bound on the achievable error probability for codes from the ensemble. In other words, there exists at least one code in the ensemble whose ML decoding error probability is upper-bounded by $\mathrm{E}\{P_e\}$. For the ease of notation, in the sequel we use P_e for the ensemble-average error probability. For the ensemble of random binary $[n, n-r]$ linear codes

$$\begin{aligned} P_e &= \sum_{\nu=r+1}^{n} \binom{n}{\nu} \varepsilon^{\nu} (1-\varepsilon)^{n-\nu} \\ &+ \sum_{\nu=1}^{r} \binom{n}{\nu} \varepsilon^{\nu} (1-\varepsilon)^{n-\nu} P_{e|\nu} \ , \end{aligned} \tag{13}$$

where $P_{e|\nu}$ denotes the conditional ensemble-average error probability given ν erasures occurred.

By using the approach based on estimating ranks of submatrices of random matrices [26] the following expression for $P_{e|\nu}$ was obtained in [17], [27], [28]

$$P_{e|\nu} = \Pr\left(\mathrm{rank}(H_I) < \nu\right) = 1 - \prod_{j=0}^{\nu-1} \left(1 - 2^{j-r}\right) \leq 2^{\nu-r} \ , \tag{14}$$

where H_I is the $r \times \nu$ submatrix of a random $r \times n$ parity-check matrix H.

The bound combining (13) and (14) is used as a benchmark to compare the FER performance of ML decoding of LDPC codes to the same performance of general linear codes in Figs. 5 – 16.

The alternative upper bound for a specific linear code with known weight enumerator has the form [27]

$$P_e \leq \sum_{i=d}^{n} \min\left\{ \binom{n}{i}, \sum_{w=d}^{i} A_{n,w} \binom{n-w}{i-w} \right\} \varepsilon^i (1-\varepsilon)^{n-i} \ . \tag{15}$$

In particular, this bound can be applied to random ensembles of codes with known average spectra (see Section III). We call this bound *S-bound*. It is presented for several ensembles in Figs. 5–16 and is discussed in Section VI.

C. Random coding upper bounds for regular (J, K)-LDPC codes

In this subsection, we derive an upper bound on the ensemble-average error probability of ML decoding for the RU and Gallager ensembles of (J, K)-regular LDPC codes. Similarly to the approach in [17], we estimate the rank of the submatrix H_I.

Theorem 3: The (J, K)-RU ensemble-average ML decoding error probability for $[n, n-r]$ codes, $n = MK$, $r = MJ$, $M \gg 1$, is bounded from above by

$$\begin{aligned} P_e &\leq \sum_{\nu=r+1}^{n} \binom{n}{\nu} \varepsilon^{\nu} (1-\varepsilon)^{n-\nu} \\ &+ \sum_{i=1}^{r} 2^{i-r} \left(1 + \frac{\binom{n-i}{K}}{\binom{n}{K}}\right)^{r} \binom{n}{i} \varepsilon^i (1-\varepsilon)^{n-i} \quad . \end{aligned} \tag{16}$$

Proof. See Appendix B. □

The same technique leads to the following bound for the Gallager ensemble of random LDPC codes:

Theorem 4: The (J, K)-Gallager ensemble-average ML decoding error probability for $[n, n-r]$ codes, $n = MK$, $r = MJ$, $M \gg 1$, is bounded from above by

$$\begin{aligned} P_e &\leq \sum_{i=1}^{r} \sum_{\mu=0}^{J(n-i)/K} \min\left\{1, 2^{\mu+i-r}\right\} \min\left\{1, \binom{\mu+J-1}{J-1} \binom{M}{\mu/J}^{J} \left(\frac{n-i}{n}\right)^{\mu K}\right\} \binom{n}{i} \varepsilon^i (1-\varepsilon)^{n-i} \\ &+ \sum_{i=r+1}^{n} \binom{n}{i} \varepsilon^i (1-\varepsilon)^{n-i} \quad . \end{aligned} \tag{17}$$

Proof. See Appendix C. □

We refer to bounds (16) and (17) as *R-bounds* since they are based on estimating ranks of submatrices of H.

Computations show that while for rates close to the capacity these bounds are rather tight, for small ε (or for rates significantly lower than the capacity) these bounds are weak. The reason for the bound untightness is related to the Gallager ensemble properties discussed in detail in Appendix A.

V. Sliding window near-ML decoding for QC LDPC codes

A binary QC LDPC block code can be considered as a tail-biting (TB) parent convolutional code determined by a polynomial parity-check matrix whose entries are monomials or zeros.

A rate $R = b/c$ *parent* LDPC convolutional code can be determined by its polynomial parity-check matrix

$$H(D) = \begin{pmatrix} h_{11}(D) & h_{12}(D) & \dots & h_{1c}(D) \\ h_{21}(D) & h_{22}(D) & \dots & h_{2c}(D) \\ \vdots & \vdots & \ddots & \vdots \\ h_{(c-b)1}(D) & h_{(c-b)2}(D) & \dots & h_{(c-b)c}(D) \end{pmatrix}, \tag{18}$$

where D is a formal variable, $h_{ij}(D)$ is either zero or a monomial entry, that is, $h_{ij}(D) \in \{0, D^{w_{ij}}\}$ with w_{ij} being a nonnegative integer, and $\mu = \max_{i,j}\{w_{ij}\}$ is the syndrome memory.

The polynomial matrix (18) determines an $[M_0 c, M_0 b]$ QC LDPC block code using a set of polynomials modulo $D^{M_0} - 1$. If $M_0 \to \infty$ we obtain an LDPC convolutional code which is considered as a parent convolutional code with respect to the QC LDPC block code for any finite M_0. By tailbiting the parent convolutional code to length $M_0 > \mu$, we obtain the binary parity-check matrix

$$H_{\mathrm{TB}} = \begin{pmatrix} H_0 & H_1 & \dots & H_{\mu-1} & H_\mu & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & H_0 & H_1 & \dots & H_{\mu-1} & H_\mu & \dots & \mathbf{0} \\ \vdots & & \ddots & \vdots & \vdots & \vdots & \ddots & \\ H_\mu & \mathbf{0} & \dots & \mathbf{0} & H_0 & H_1 & \dots & H_{\mu-1} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ H_1 & \dots & H_\mu & \mathbf{0} & \dots & \mathbf{0} & \dots & H_0 \end{pmatrix}, \tag{19}$$

of an equivalent (in the sense of column permutation) TB code (all matrices H_i including H_{TB} should have a transpose operator to get the exact TB code [29]), where H_i, $i = 0, 1, \dots, \mu$, are binary $(c-b) \times c$ matrices in the series expansion

$$H(D) = H_0 + H_1 D + \cdots + H_\mu D^\mu.$$

If every column and row of $H(D)$ contains J and K nonzero entries, respectively, we call $\mathcal{C}$ a (J, K)-regular QC LDPC code and irregular otherwise.

Notice that by zero-tail termination [29] of (18) at length $W > \mu$, we can obtain a parity-check matrix of a $[Wc, (W-\mu)b]$ zero-tail terminated (ZT) QC LDPC code.

Consider a BEC with erasure probability $\varepsilon > 0$. Let H be an $(c-b)M_0 \times cM_0$ parity-check matrix of a binary $[n = M_0 c, k = M_0 b, d_{\min}]$ QC LDPC block code, where $d_{\min}$ is the minimum Hamming distance of the code. An ML decoder corrects any pattern of ν erasures if $\nu \le d_{\min} - 1$. If $d_{\min} \le \nu \le n - k$ then a unique correct decision can be obtained for some erasure patterns. The number of such correctable patterns depends on the code structure.

As explained in Section II, ML decoding over the BEC is reduced to solving (1). Its complexity for sparse parity-check matrices is of order ν^2, $\nu = |I|$, that is, still computationally intractable for LDPC codes of large lengths.

In order to reduce decoding complexity, we apply a sliding-window decoding algorithm which is modified for QC LDPC block codes. This decoder is determined by a binary parity-check matrix H_{W}

$$H_{\mathrm{W}} = \begin{pmatrix} H_0 & \dots & H_\mu & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & H_0 & \dots & H_\mu & \mathbf{0} & \dots & \mathbf{0} \\ \vdots & \ddots & \ddots & & \ddots & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{0} & H_0 & \dots & H_\mu & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & H_0 & \dots & H_\mu \end{pmatrix} \tag{20}$$

of size $(W-\mu)(c-b) \times Wc$, where $W \ge 2\mu + 1$ denotes the size of the decoding window in blocks. The matrix (20) determines a ZT LDPC parent convolutional code. We start decoding with BP decoding applied to the original QC LDPC block code of length $n = M_0 c$, and then apply ML decoding to the ZT LDPC parent convolutional code determined by the parity-check matrix (20). It implies solving a system of linear equations

$$\boldsymbol{x}_{I,\mathrm{W}} H_{I,\mathrm{W}}^{\mathrm{T}} = \boldsymbol{s}_{\mathrm{W}}\ , \tag{21}$$

where $\boldsymbol{x}_{I,\mathrm{W}} = (x_{I,\mathrm{W},i}, x_{I,\mathrm{W},i+1 \mod n}, \ldots, x_{I,\mathrm{W},i+Wc-1 \mod n})$, $i = 0, s, 2s, \ldots, \mod n$ is a subvector of $\boldsymbol{x}_I$ corresponding to the chosen window, s denotes the size of the window shift, and $\boldsymbol{s}_{\mathrm{W}}$ and $H_{I,\mathrm{W}}$ are the corresponding subvector of $\boldsymbol{s}$ and submatrix of H_I, respectively. The final decision is made after $\alpha n/s$ steps, where α denotes the number of passes of sliding-window decoding. The formal description of the decoding procedure is given below as Algorithms 1 and 2.

Notice, that the choice of s affects both the performance and the complexity. By increasing s we can speed up the decoding procedure at the cost of some performance loss. In the sequel, we use $s = c$ bits that corresponds to the lowest possible FER.

Algorithm 1 BP-BEC

while there exist parity checks with only one erased symbol **do**
Assign to the erased symbol the modulo-2 sum of all nonerased symbols participating in the same parity check.
end while

Algorithm 2 Wrap-around algorithm for near-ML decoding of QC LDPC codes over the BEC

Input: BEC output $\boldsymbol{y}$.
Perform BP decoding for $\boldsymbol{y}$.
$\mathrm{wstart} \leftarrow 0$;
$\mathrm{wend} \leftarrow W - 1$;
$\mathrm{corrected} \leftarrow 1$;
while $\mathrm{corrected} > 0$ **do**
$\mathrm{corrected} \leftarrow 0$;
Apply ML decoding to the window $(y_{\mathrm{wstart}}, \ldots, y_{\mathrm{wend}})$;
$\mathrm{wstart} \leftarrow \mathrm{wstart} + s \mod n$;
$\mathrm{wend} \leftarrow \mathrm{wend} + s \mod n$;
if $\mathrm{wstart} = 0$ **then**
corrected $\leftarrow$ number of corrected erasures after a full round:
end if
end while
return $\boldsymbol{y}$

VI. Numerical results and simulations

A. Short codes

In Figs. 5 and 6 we compare upper and lower bounds on the error probability of ML decoding for short codes on the BEC.

First, notice that the lower bounds in Theorems 1 and 2 (sphere packing and the tightened sphere packing bound) almost coincide with each other near the channel capacity. However, the new bound is significantly tighter than the known one at the low erasure probability region. For further comparisons we use the new bound whenever information on the code minimum distance is available.

Upper bounds in Theorems 3 and 4 are also presented in Figs. 5 and 6. These two bounds are indistinguishable at high symbol erasure probabilities but the difference between them is visible at the low ε region where all bounds are rather weak.

Notice, that at high ε random coding bounds for LDPC codes are close to those of random linear codes. For the (J, K)-regular LDPC codes with $J = 4$ the random bound is almost as good as the random bound for general linear codes of the same length and dimension in a wide range of ε values.

In Fig. 7 we compare upper bounds on the ML decoding FER performance for the Gallager ensemble of (J, K)-regular LDPC codes with different pairs (J, K) to the upper bounds for general linear codes of different code rates. Interestingly, the convergence rate of the bounds for LDPC codes to the bounds for linear codes depends on the code rate. For rate $R = 1/3$ even for very sparse codes with column weight $J = 3$ their FER performance under ML decoding is very close to the FER performance of general linear codes.

Next, we present simulation results for two ensembles of LDPC codes. The first ensemble is the Gallager ensemble.

In Figs. 8, 9 we present the FER performance of ML decoding simulated for the 10 randomly selected $(3, 6)$ and $(4, 8)$-regular Gallager codes of length $n = 96$. The FER performance for these codes has large variation and most of the FER performance curves are located between the rank R-bound and the spectral S-bound.

In Figs. 10, 11 we present the FER performance of ML decoding simulated for the 10 randomly selected $(3, 6)$ and $(4, 8)$ RU codes of length $n = 96$. The FER performance variation in both cases is smaller than for the Gallager codes and the $(4, 8)$-regular RU codes are concentrated around the R-bound.

There are two reasons for better behavior of the RU codes. First, the code rate for the RU codes is typically equal to $R = 1/2$ whereas for the Gallager codes the rate is $R \geq 52/96 = 0.5417$.

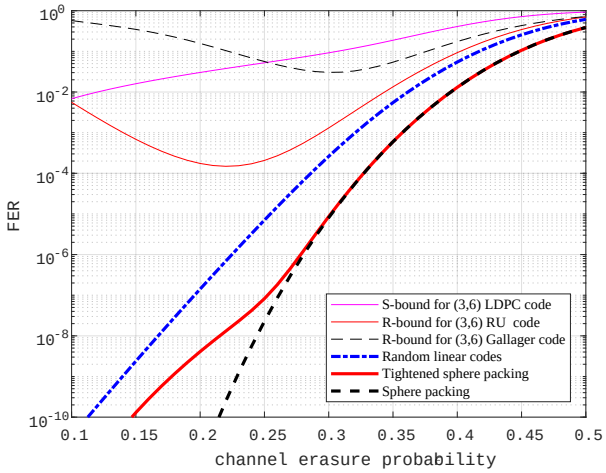

Fig. 5. Bounds for the binary $(3, 6)$-regular LDPC codes $n = 96$, $R = 1/2$. Here S-bound is defined by (15), R-bounds for RU and Gallager codes are defined by (16) and (17), respectively. Random coding bound is computed by (13)–(14). Sphere packing bound and the tightened sphere packing bound are computed according (11) and (12), respectively.

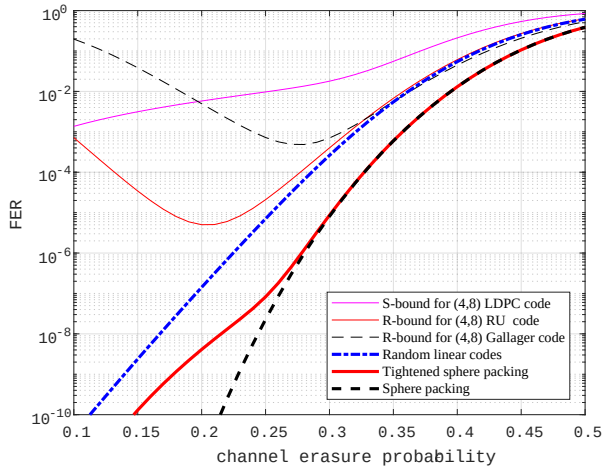

Fig. 6. Bounds for the binary $(4, 8)$-regular LDPC codes $n = 96$, $R = 1/2$ Here S-bound is defined by (15), R-bounds for RU and Gallager codes are defined by (16) and (17), respectively. Random coding bound is computed by (13)–(14). Sphere packing bound and the tightened sphere packing bound are computed according to (11) and (12), respectively.

One more reason for the difference in the FER performance is the Gallager approximation discussed in the Appendix A. In our derivations we used the Gallager approximation for estimating the number of non-full-rank submatrices of size $r \times \nu$, where ν denotes the number of erasures. For decreasing values of ν, the number of independent parity checks in "independent" strips decreases, and, therefore, the bound gets loose. Since the parity-check matrix for the RU codes is not split into strips, the inter-row dependencies for this submatrix are weaker than for the Gallager codes.

In Fig. 12 we compare the best among 10 randomly selected $(3, 6)$-regular LDPC codes and the best among 10 randomly selected $(4, 8)$-regular LDPC codes of the two ensembles. As predicted by bounds, the ML decoding performance of the $(4, 8)$ codes is much better than that for the $(3, 6)$ codes in both ensembles. Due to the rate bias and imperfectness of the Gallager approximation, codes from the Gallager ensemble are weaker than the RU codes with the same parameters. The performance of the RU codes perfectly matches the R-upper bound. Moreover, the best $(4, 8)$ RU codes show even better FER performance than the average FER performance over general linear codes at the high erasure probability region.

B. Codes of moderate length

The FER performance for relatively long codes of length $n = 1008$ are shown in Fig. 13. Notice that the difference between the lower and upper bounds for general linear codes which was noticeable for short codes became very small for $n = 1008$. Since the lower bound (11) is simply the probability of more than r erasures, the fact that the upper and the lower bounds almost coincide leads us to the conclusion that even for rather low channel erasure probabilities ε achieving ML decoding performance requires correcting almost all combinations of erasures of weight close to the code redundancy r.

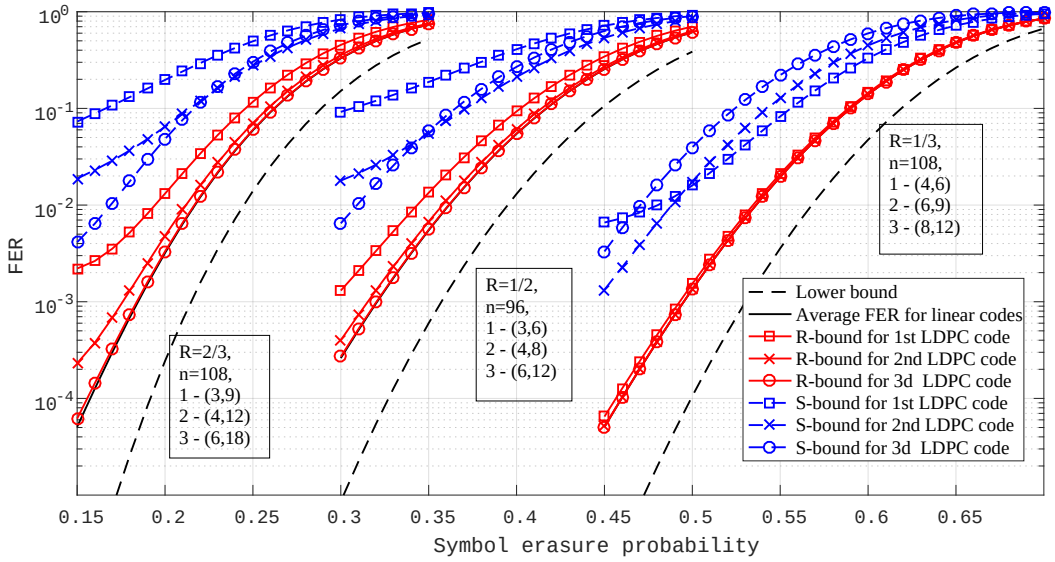

Fig. 7. Error probability bounds for the binary (J, K)-regular LDPC codes of length $n \approx 100$. Here S- and R-bounds are defined by (15) and (16), respectively. The average ML decoding FER performance for linear codes is computed by (13) – (14). The lower bound is (12).

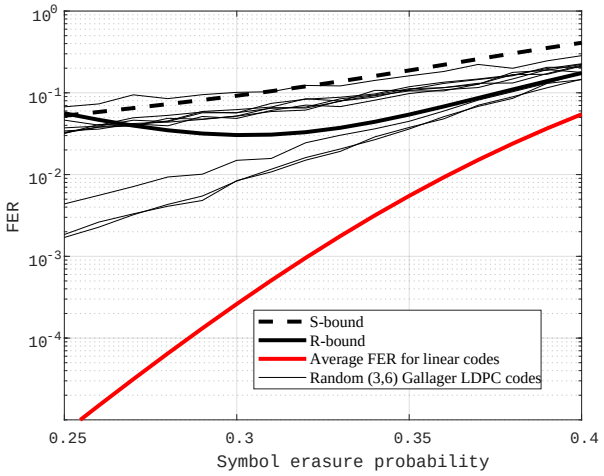

Fig. 8. Error probability bounds and simulation results for the Gallager $(3, 6)$-regular codes of length $n = 96$. S- and R- bounds are defined by (15) and (17), respectively. The average ML decoding FER performance for linear codes is computed by (13)–(14).

Notice that according to the S-bounds in Fig. 13 error floors are expected at the low erasure probability region. The error floor level strongly depends on the J and K and rapidly decreases with increasing J.

In Figs. 14 and 15 we show simulation results for 5 randomly selected $(3, 6)$ and $(4, 8)$ codes of length $n = 1008$ from the Gallager and the RU ensembles, respectively. At the same plots the rank and spectral bounds for the corresponding ensembles are shown.

We observe that for rates close to the capacity, the rank and spectral bounds demonstrate approximately the same behavior as the average simulated FER. For the low channel erasure probabilities the spectral bound predicts error floors. As expected, the spectral bound is weak for rates far below the channel capacity.

In the same plots, we show the average BP decoding performance of all 5 codes. Since all 10 codes (5 for the Gallager and 5 for the RU ensembles) show the identical FER performance, we present only one of the FER BP curves in each plot.

Finally, we demonstrate the FER performance of the ML decoding for 2 non-random codes. First one is the irregular QC LDPC code with base matrix of size 12×24 optimized for the BEC channel using approach in [30]. The second code is the $(4, 8)$-regular so-called double-Hamming code from [31]. Simulations show that the irregular code has better ML decoding FER performance than any randomly generated codes and almost everywhere outperforms the double-Hamming code. The double-Hamming code is better than the randomly generated RU and Galllager codes but it mimics their error-floor behavior.

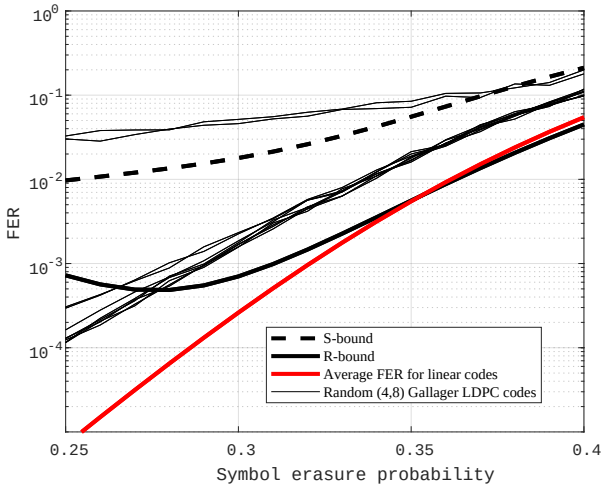

Fig. 9. Error probability bounds and simulation results for the Gallager $(4, 8)$-regular codes of length $n = 96$. S- and R- bounds are defined by (15) and (17), respectively. The average ML decoding FER performance for linear codes is computed by (13)–(14).

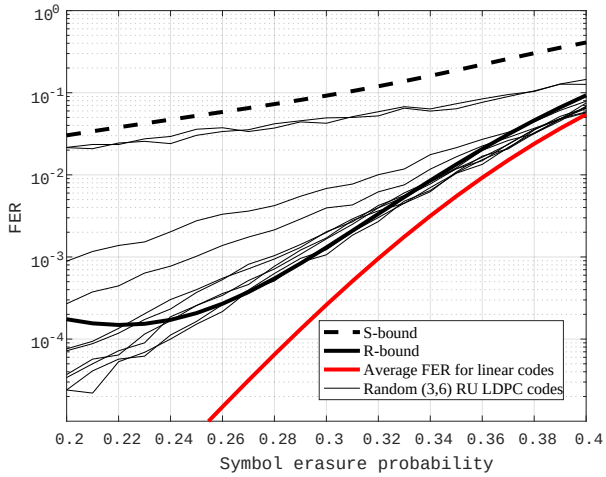

Fig. 10. Error probability bounds and simulation results for the $(3, 6)$-RU codes of length $n = 96$. S- and R- bounds are defined by (15) and (16), respectively. The average ML decoding FER performance for linear codes is computed by (13)–(14)

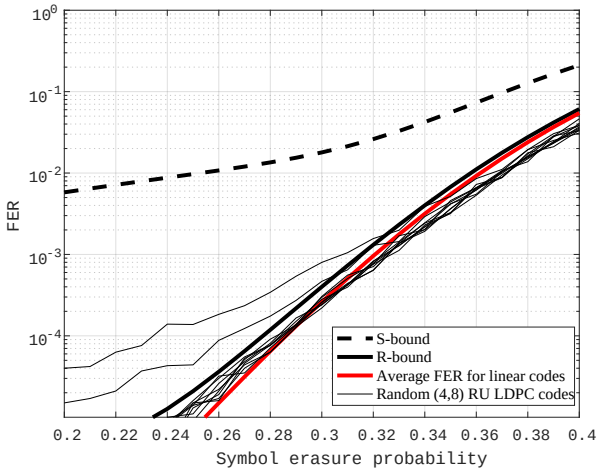

Fig. 11. Error probability bounds and simulation results for the $(4, 8)$-RU codes of length $n = 96$. S- and R- bounds are defined by (15) and (16), respectively. The average ML decoding FER performance for linear codes is computed by (13)–(14).

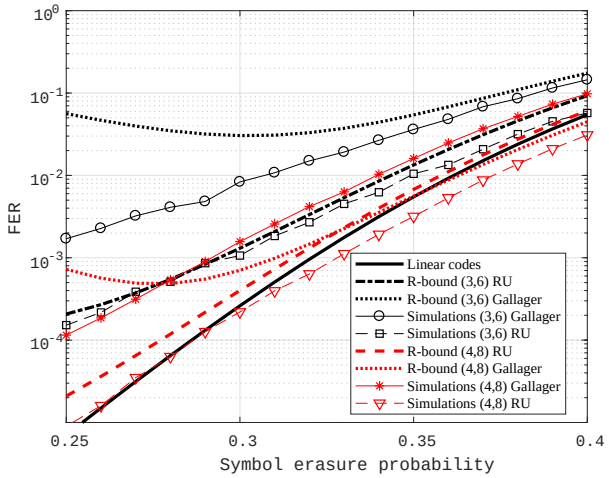

Fig. 12. Simulation results for the $(3,6)$-regular and $(4,8)$-regular codes of length $n = 96$ from the Gallager and RU ensembles. S-bound is defined by (15), R- bounds for RU and Gallager codes are are defined by (16) and (17), respectively. The average ML decoding FER performance for linear codes is computed by (13)–(14).

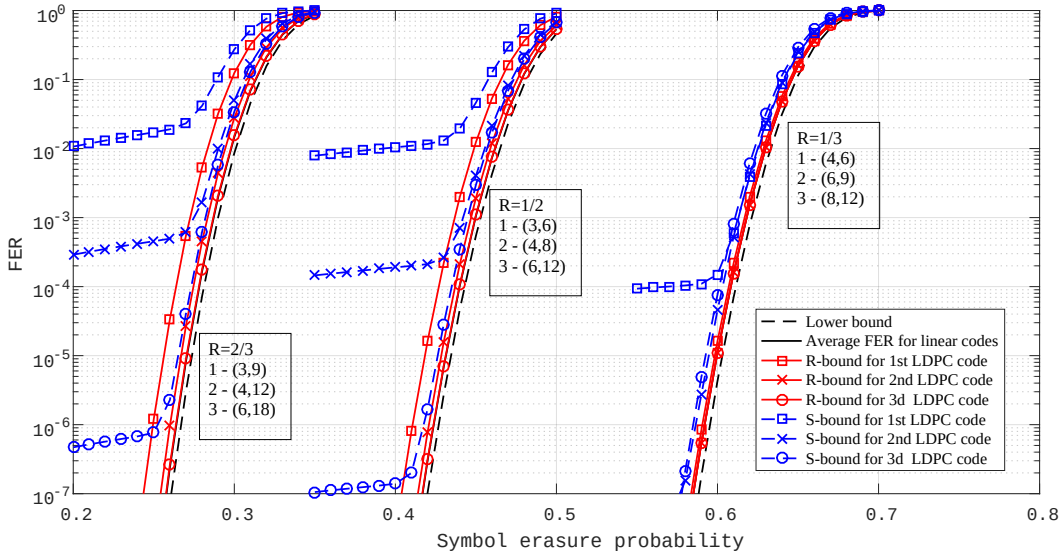

Fig. 13. Error probability bounds for the binary (J,K)-regular LDPC codes of length $n = 1008$. S- and R- bounds are defined by (15) and Theorem 3, respectively. The average ML decoding FER performance for linear codes is computed by (13)–(14). The lower bound is (12).

C. Sliding window near-ML decoding for QC LDPC codes

Simulation results for the irregular rate $12/24$ LDPC code and for the double-Hamming regular LDPC code are presented in Fig. 16. Parameters of the codes and the sliding window near-ML (SWML) decoder are summarized in Table I.

The FER performance of SWML decoding for the double-Hamming code is very close to the theoretical bounds despite very low decoding complexity. In contrast, the FER performance of BP decoding for this code is extremely inefficient.

For the irregular code, BP decoding is much more efficient than for the double-Hamming code, but its SWML decoding FER performance is worse than that for the double-Hamming code.

In general, both codes show very good error-correcting performance.

VII. Thresholds

In order to obtain the asymptotic upper bound on the error probability for $(J,K)-$regular LDPC codes we use the following inequality

$$\frac{\binom{n-\nu}{K}}{\binom{n}{K}} \leq \left(\frac{n-\nu}{n}\right)^K . \tag{22}$$

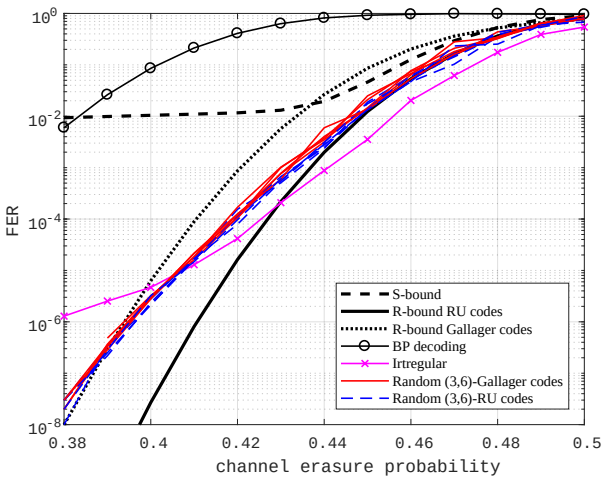

Fig. 14. Error probability bounds and simulation results for the $(3, 6)$-regular LDPC codes of length $n = 1008$. Here S-bound is defined by (15), R-bounds for the RU and Gallager codes are defined by (16) and (17), respectively.

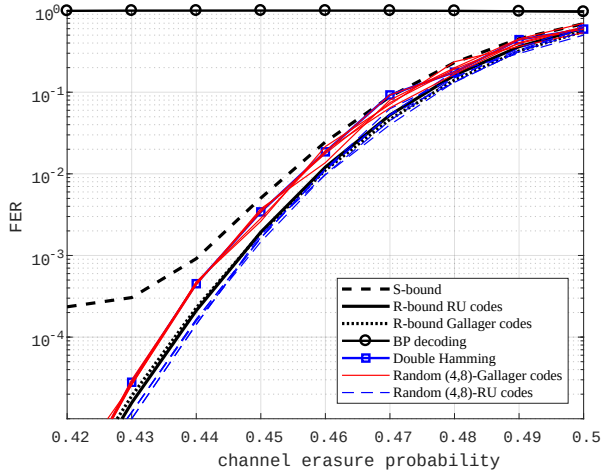

Fig. 15. Error probability bounds and simulation results for the $(4, 8)$-regular LDPC codes of $n = 1008$. Here S-bound is defined by (15), R-bounds for the RU and Gallager codes are defined by (16) and (17), respectively.

For the RU ensemble, it follows from (16) that

$$P_e \leq \sum_{\nu=1}^{n} \min\{1, B_\nu\} \binom{n}{\nu} \varepsilon^\nu (1-\varepsilon)^{n-\nu} \tag{23}$$

$$\leq n \cdot \exp\left\{-\min_{\nu}\{\max\{T_1, T_1 - \log B_\nu\}\}\right\}, \tag{24}$$

where

$$B_\nu = 2^{\nu - r}\left(1 + \frac{\binom{n-\nu}{K}}{\binom{n}{K}}\right)^r,$$

$$T_1 = -\log\binom{n}{\nu} - \nu \log\varepsilon - (n-\nu)\log(1-\varepsilon)\,.$$

Denote by $\alpha = \nu/n$ the normalized number of erasures. The asymptotic error probability exponent can be written as

$$E(\varepsilon) = \lim_{n\to\infty}\left\{-\frac{\log P_e}{n}\right\}$$

$$= \min_{\alpha\in[0,1]}\{\max\{F_1(\alpha,\varepsilon)\,, F_2(\alpha,\varepsilon)\}\}\,, \tag{25}$$

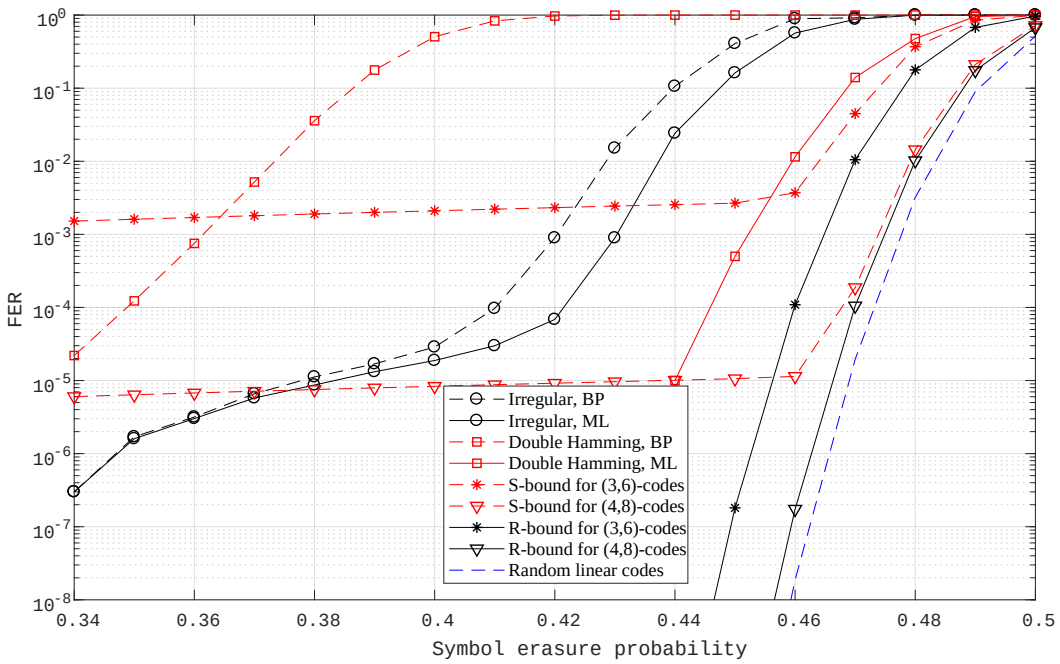

Fig. 16. Error probability bounds and the simulated FER performance of SWML decoding for QC LDPC codes of length $n = 4800$. S- and R- bounds are defined by (15) and (16), respectively. The average FER performance of ML decoding for linear codes is computed by (13)–(14).

TABLE I
EXAMPLE PARAMETERS OF SWML DECODERS FOR CODES OF LENGTH $n = 4800$

Code parameters	Codes	
	Irregular	Double-Hamming
Base matrix size	12×24	8×16
Lifting degree	24	32
Decoding window	51	68
Window Shift	24	16
Overall TB length	200	300
Maximum number of passes	15	15

where

$$F_1(\alpha,\varepsilon) = -h(\alpha) - \alpha\log\varepsilon - (1-\alpha)\log(1-\varepsilon)\ , \tag{26}$$

$$F_2(\alpha,\varepsilon) = F_1(\alpha,\varepsilon) - F_3(\alpha)\ , \tag{27}$$

$$F_3(\alpha) = \left(\alpha - \frac{J}{K}\right)\log 2 + \frac{J}{K}\left[\log\left(1+(1-\alpha)^K\right)\right]\ , \tag{28}$$

$$h(\alpha) = -\alpha\log\alpha - (1-\alpha)\log(1-\alpha)\ . \tag{29}$$

In (25) – (29) all logarithms are to the base of e. The asymptotic decoding threshold is defined as the maximum ε providing $E(\varepsilon) > 0$, or as the minimum ε providing $E(\varepsilon) = 0$. It is easy to see that $F_1(\alpha,\varepsilon)$ is always positive except at the point $\alpha = \varepsilon$ where $F_1(\alpha,\varepsilon) = 0$ and $F_2(\alpha,\varepsilon) > 0$ for $\alpha < \varepsilon$ and $F_2(\alpha,\varepsilon) = 0$ at $\alpha = \varepsilon$ if $F_3(\alpha) = F_3(\varepsilon) = 0$.

In other words, the ML decoding threshold can be found as an unique solution of the equation

$$\varepsilon = \frac{J}{K}\left[1 - \frac{\log\left(1+(1-\varepsilon)^K\right)}{\log 2}\right]\ . \tag{30}$$

Notice that increasing K leads to the simple expression for the threshold

$$\varepsilon \xrightarrow[K\to\infty]{} \frac{J}{K} = 1 - R$$

which corresponds to the capacity of the BEC channel.

Numerical values of the ML decoding threshold for different code rates and different column weights are shown in Table II.

TABLE II
ASYMPTOTIC ML DECODING THRESHOLDS FOR THE BINARY (J, K)-REGULAR LDPC CODES ON THE BEC

R	J					
	3	4	5	6	8	9
1/4	0.746930*	—	—	0.749989	—	0.750000
1/3	—	0.665737*	—	0.666633	0.666665	—
1/2	0.491422*	0.497987	0.499507	0.499878	0.499992	0.499998
2/3	0.323598	0.330648	0.332560	0.333106	0.333314	0.333327
3/4	0.241029	0.247364	0.249191	0.249747	0.249975	0.249992

Surprisingly, the new bounds on the thresholds marked by asterisk in Table II essentially identical to the upper bounds on the ML decoding thresholds in [32, Eq. (37), Table 1], although analytical expressions for bounds are different. This confirms the tightness of the bound in Theorem 3 and bounds in [32] for rates near the channel capacity.

VIII. CONCLUSION

Both finite-length and asymptotic analysis of ML decoding performance for LDPC codes on the BEC channel is presented. The obtained bounds are very useful since unlike other channel models, for the BEC ML decoding can be implemented for rather long codes. Moreover a sliding window decoding algorithm which is efficient for near-ML decoding of very long codes is developed.

Comparison of the presented bounds with empirical estimates of the average error probability over sets of randomly constructed codes has shown that the new bounds are rather tight at rates close to the channel capacity even for short codes. For code length $n > 1000$, the bounds are rather tight in a wide range of parameters.

The new bounds lead to a simple analytical expression for the ML decoding thresholds on the BEC for regular LDPC codes.

APPENDIX A

There is a weakness in the Gallager's proof of Theorem 2.4 in [13], analogous to that in derivations (8)-(10) above. Formulas (2.17) in [13] and (10) in this paper state that the average number of weight w binary sequences which simultaneously satisfy parity checks in J strips is

$$\mathrm{E}(A_{n,w}) = \binom{n}{w} \left[\frac{N_{n,w}}{\binom{n}{w}} \right]^J , \tag{31}$$

where $N_{n,w}$ is the number of weight w sequences satisfying the parity checks of the first strip H_1. This formula relies on the assumption that parity checks of strips are independent. It is known that this assumption is not true because the strips of the parity-check matrix are always linearly dependent (sum of parity checks of any two strips is all-zero sequence) and, as a consequence, the actual rate of the regular (J, K)-LDPC codes is higher than $1 - J/K$.

The fact that we intensively use strip-independence hypothesis in our derivations motivated us to study deeper the influence of the strip independence assumption both on the conclusions in [13] and on the derivations done in this paper.

In order to verify how this assumption influences the precision of estimates let us consider the following simple example.

Example 1: Consider (3,3)-regular code with $M = 2$. The first strip is

$$\begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix} .$$

The other two strips are obtained by random permutations of the columns of this strip. In total there exist $(6!)^2$ LDPC codes but most of the codes are equivalent. By taking into account that the first row in each strip determines the second row we obtain that the choice of each code is determined by the choice of the third and fifth rows of the parity-check matrix. Thus, there are at most $\binom{6}{3}^2 = 400$ equiprobable classes of equivalent codes. We compute the average spectra over codes with a certain code dimension and the average spectrum over all codes. The obtained results are presented in Table III.

TABLE III
SPECTRA OF (3,3)-REGULAR LDPC CODES OF LENGTH 6

Dimension	Number of codes	Average Spectrum
2	288	$\left(1 \quad 0 \quad \frac{1}{2} \quad 0 \quad \frac{5}{2} \quad 0 \quad 0\right)$
3	108	$\left(1 \quad 0 \quad 2 \quad 0 \quad 5 \quad 0 \quad 0\right)$
4	4	$\left(1 \quad 0 \quad 6 \quad 0 \quad 9 \quad 0 \quad 0\right)$
Average parameters over the ensemble		
2.29	—	$\left(1 \quad 0 \quad \frac{24}{25} \quad 0 \quad \frac{81}{25} \quad 0 \quad 0\right)$

We note that the lower bound on the code rate $R \geq 1 - J/K = 0$, but due to existance of at least two rows linearly dependent on other rows, the tightened lower bound on the code rate is $R \geq 1 - 4/6 \geq 1/3$.

Let us compare these empirical estimates with the Gallager bound. The generating function for the first strip is

$$g(s) = 1 + N_{6,2}s^2 + N_{6,4}s^4 = 1 + 6s^2 + 9s^4.$$

According to (31)

$$\mathrm{E}(A_{6,2}) = \binom{6}{2}\left(\frac{6}{\binom{6}{2}}\right)^3 = \frac{24}{25} \tag{32}$$

$$\mathrm{E}(A_{6,4}) = \binom{6}{4}\left(\frac{9}{\binom{6}{4}}\right)^3 = \frac{81}{25} \tag{33}$$

$$\tag{34}$$

which matches with the empirical average over all codes presented in Table III.

These computations lead us to the following conclusions:

- In the ensemble of the (J, K)-regular LDPC codes there are codes of different dimensions. The average spectra depend on the dimension of the code and differ from the average spectra over all codes of the ensemble.
- Average over all codes may coincide with the Gallager estimate but does not correspond to any particular linear code. Moreover, the estimated number of codewords (sum of all spectrum components) is not necessarily equal to a power of 2.

Notice that if M is large enough then influence of strip dependence on the precision of the obtained spectrum estimate is negligible. However, if $\nu \ll M$, that is for low ε region, the assumption of strip independence should be used with caution.

APPENDIX B

Proof of Theorem 3

Assume that the number of erasures is $\nu > 0$. The error probability of ML decoding over the BEC is estimated as the probability that ν columns of the random parity-check matrix H from the RU ensemble corresponding to the erased positions are linearly dependent, $\nu \leq r$.

Let H_I be a submatrix consisting of columns determined by the set I of the erased positions, $|I| = \nu$. We can write

$$\Pr\left(\mathrm{rank}(H_I) < \nu \middle| \, \nu\right) \leq \sum_{\boldsymbol{x}_I \neq \mathbf{0}} \Pr\left(\boldsymbol{x}_I H_I^{\mathrm{T}} = \mathbf{0} \middle| \, \nu\right). \tag{35}$$

Consider a random vector $\boldsymbol{s} = \boldsymbol{x}_I H_I^{\mathrm{T}}$. Denote by $\boldsymbol{s}_i^j$ a subvector $(s_i, ..., s_j)$ of the vector $\boldsymbol{s}$. The probability of the all-zero vector $\boldsymbol{s}$ is

$$p(\boldsymbol{s} = \mathbf{0}|\nu) = p(s_1 = 0|\nu) \prod_{i=2}^{r} p(s_i = 0|\boldsymbol{s}_1^{i-1} = \mathbf{0}, \nu). \tag{36}$$

Next, we prove that $p(s_1 = 0|\nu) \geq p(s_i = 0|\boldsymbol{s}_1^{i-1} = \mathbf{0}, \nu)$, $i = 2, 3..., r$. We denote by ν_i the number of erasures in nonzero positions of the ith parity check. For the choice of a random vector $\boldsymbol{x}$ and a random parity-check matrix from the RU ensemble the probability of a zero syndrom component s_i is

$$p(s_i = 0|\nu_i) = \begin{cases} 1, & \nu_i = 0 \\ \frac{1}{2}, & \nu_i > 0. \end{cases} \tag{37}$$

First, we observe that for all i

$$\begin{aligned} p(s_i = 0|\nu) &= p(s_i = 0|\nu_i = 0, \nu)p(\nu_i = 0|\nu) + p(s_i = 0|\nu_i > 0, \nu)p(\nu_i > 0|\nu) \\ &= 1 \cdot p(\nu_i = 0|\nu) + \frac{1}{2} \cdot (1 - p(\nu_i = 0|\nu)) \\ &= \frac{1 + p(\nu_i = 0|\nu)}{2}. \end{aligned} \tag{38}$$

For all $i \neq j$ let K' denote the number of row positions in which the corresponding elements either in row j or in row i of H are nonzero. Since $K' \geq K$ the following inequality holds

$$\begin{aligned} p(\nu_j = 0|\nu_i = 0, \nu) &= \frac{\binom{n-K'}{\nu}}{\binom{n}{\nu}} \\ &\leq \frac{\binom{n-K}{\nu}}{\binom{n}{\nu}} \\ &= p(\nu_j = 0|\nu). \end{aligned} \tag{39}$$

For the second parity check by using arguments similar to those in (38) we obtain

$$p(s_2=0|s_1=0,\nu) \;=\; \frac{1}{2}\left(1+p(\nu_2=0|s_1=0,\nu)\right)\ . \tag{40}$$

The conditional probability in the RHS can be estimated as follows

$$\begin{aligned} p(\nu_2=0|s_1=0,\nu) &= \sum_{\nu_1=0}^{\min\{K,\nu\}} p(\nu_2=0|s_1=0,\nu_1,\nu)p(\nu_1|s_1=0,\nu) \\ &= \sum_{\nu_1=0}^{K} p(\nu_2=0|\nu_1,\nu)\frac{p(s_1=0|\nu_1,\nu)p(\nu_1|\nu)}{p(s_1=0|\nu)}\ . \end{aligned} \tag{41}$$

Here we take into account that $p(\nu_2=0|s_1=0,\nu_1,\nu)=p(\nu_2=0|\nu_1,\nu)$.

By substituting (37) into (41) we have

$$\begin{aligned} & p(\nu_2=0|s_1=0,\nu) \\ = \;& \frac{p(\nu_2=0|\nu_1=0,\nu)p(\nu_1=0,\nu)}{p(s_1=0|\nu)} + \frac{\sum_{\nu_1=1}^{\min\{K,\nu\}} p(\nu_2=0|\nu_1,\nu)p(\nu_1|\nu)}{2p(s_1=0|\nu)} \\ = \;& \frac{p(\nu_2=0|\nu_1=0,\nu)p(\nu_1=0|\nu) + \sum_{\nu_1=0}^{\min\{K,\nu\}} p(\nu_2=0|\nu_1,\nu)p(\nu_1|\nu)}{2p(s_1=0|\nu)}\ . \end{aligned} \tag{42}$$

The second term in the nominator is equal to $p(\nu_2=0|\nu)$ and $p(\nu_i=0|\nu)$ does not depend on i. Thus, we obtain

$$\begin{aligned} p(\nu_2=0|s_1=0,\nu) &= p(\nu_2=0|\nu)\frac{p(\nu_2=0|\nu_1=0,\nu)+1}{2p(s_1=0|\nu)} \\ &\overset{(a)}{\leq} p(\nu_2=0|\nu)\frac{p(\nu_2=0|\nu)+1}{2p(s_1=0|\nu)} \\ &\overset{(b)}{=} p(\nu_2=0|\nu)\ , \end{aligned} \tag{43}$$

where the inequality (a) follows from (39) and the equality (b) follows from (38). From (38), (40) and (43) we conclude that

$$p(s_2=0|s_1=0,\nu) \leq p(s_2=0|\nu)\ .$$

Consecutively applying these derivations for $i=3,4,...,r$ we can prove that

$$p(s_i=0|\boldsymbol{s}_1^{i-1}=\mathbf{0},\nu) \leq p(s_i=0|\nu)$$

and then from (36) it follows that

$$p(\boldsymbol{s}=\mathbf{0}|\nu) \leq p(s_1=0|\nu)^r\ .$$

The probability that the row i in H_I has only zeros can be bounded from above by

$$p(\nu_i=0|\nu) = \frac{\binom{n-\nu}{K}}{\binom{n}{K}}\ . \tag{44}$$

The probability that the entire sequence of length ν is a codeword (all r components of the syndrome are equal to zero) is

$$p(\boldsymbol{s}=\mathbf{0}|\nu) \leq 2^{-r}\left(1+\frac{\binom{n-\nu}{K}}{\binom{n}{K}}\right)^r\ . \tag{45}$$

By substituting (45) into (35), we obtain

$$P_{e|\nu} = \Pr\left(\text{rank}(H_I)<\nu|\,\nu\right) \leq 2^{\nu-r}\left(1+\frac{\binom{n-\nu}{K}}{\binom{n}{K}}\right)^r\ .$$

We conclude that the statement of Theorem 3 follows from (13).

APPENDIX C

Proof of Theorem 4

Assume that the number of erasures is $\nu > 0$. Let H_I be a submatrix consisting of columns numbered by the set I of the erased positions, $|I| = \nu$. In Section IV it is shown that the problem of estimating the FER of ML decoding can be reduced to the problem of estimating rank of submatrix H_I. Let $H_{j,I}$ denote the jth strip of H_I, $j = 1, 2, ..., J$. Denote by μ_i the number all-zero rows in $H_{i,I}$, and $\boldsymbol{\mu} = (\mu_1, ..., \mu_J)$.

Assume that the vector $\boldsymbol{x}$ is chosen uniformly at random from the set of binary vectors of length n. and let $\boldsymbol{x}_I$ be a subvector of bsx consisting of elements numbered by the set I. Then, we can write:

$$\Pr\left(\text{rank}(H_I) < \nu | \nu\right) \leq \sum_{\boldsymbol{\mu}} \Pr\left(\boldsymbol{x}_I : \boldsymbol{x}_I H_{j,I}^{\mathrm{T}} = \boldsymbol{0}, \boldsymbol{x}_I \neq 0 \text{ for all } j | \nu, \boldsymbol{\mu}\right) p(\boldsymbol{\mu}|\nu) \,. \tag{46}$$

For the Gallager ensemble, the vector $\boldsymbol{\mu}$ has the following conditional probability distribution given that the number of erasures is ν:

$$\begin{aligned} p(\boldsymbol{\mu}|\nu) &= \prod_{i=1}^{J} p(\mu_i|\nu) \\ &= \prod_{i=1}^{J} \binom{M}{\mu_i} \frac{\binom{n-\mu_i K}{\nu}}{\binom{n}{\nu}} = \prod_{i=1}^{J} \binom{M}{\mu_i} \frac{\binom{n-\nu}{\mu_i K}}{\binom{n}{\mu_i K}} \,. \end{aligned} \tag{47}$$

where we take into account that the strips are obtained by independent random permutations. Using inequality (22) we can upperbound this distribution as

$$p(\boldsymbol{\mu}|\nu) \leq \left[\prod_{i=1}^{J} \binom{M}{\mu_i}\right] \left[\prod_{i=1}^{J} \left(\frac{n-\nu}{n}\right)^{\mu_i K}\right] \tag{48}$$

$$\leq \binom{M}{\mu/J}^{J} \left(\frac{n-\nu}{n}\right)^{\sum_{i=1}^{J} \mu_i K} = \binom{M}{\mu/J}^{J} \left(\frac{n-\nu}{n}\right)^{\mu K} \,. \tag{49}$$

where $\mu = \sum_{i=1}^{J} \mu_i$ and the second inequality follows from the fact that maximum of the first product in (48) is achieved with $\mu_1 = \mu_2 = ... = \mu_J = \mu/J$.

According to (37) each of the $M - \mu_i$ nonzero rows of the i-th strip produces zero syndrome component with probability $\frac{1}{2}$. For a given set of μ_i, $i = 1, ..., J$, $\sum_{i=1}^{J} \mu_i = \mu$, $0 \leq \mu \leq r$ the union bound on the probability of the zero syndrome is

$$\begin{aligned} \Pr\left(\boldsymbol{x}_I : \boldsymbol{x}_I H_{j,I}^{\mathrm{T}} = \boldsymbol{0}, \boldsymbol{x}_I \neq 0 \text{ for all } j | \nu, \boldsymbol{\mu}\right) &\leq \min\left\{1, \sum_{\boldsymbol{x}_I \neq \boldsymbol{0}} \Pr\left(\boldsymbol{x}_I H_{j,I}^{\mathrm{T}} = \boldsymbol{0}, \boldsymbol{x}_I \neq 0 \text{ for all } j | \nu, \boldsymbol{\mu}\right)\right\} \\ &\leq \min\left\{1, (2^{\nu} - 1) \prod_{j=1}^{J} 2^{-M+\mu_j}\right\} \\ &\leq \min\left\{1, 2^{\nu - MJ + \sum_{j=1}^{J} \mu_j}\right\} = \min\left\{1, 2^{\nu - r + \mu}\right\}. \end{aligned} \tag{50}$$

From (46) it follows that

$$\Pr\left(\text{rank}(H_I) < \nu | \, \nu\right) \leq \sum_{\mu=0}^{r} \min\left\{1, 2^{\nu+\mu-r}\right\} \sum_{\boldsymbol{\mu}: \sum_{j=1}^{J} \mu_j = \mu} p(\boldsymbol{\mu}|\nu) \tag{51}$$

.

The total number of different $\boldsymbol{\mu}$ with a given sum μ is equal to $\binom{\mu+J-1}{J-1}$. From (49) we obtain

$$\sum_{\boldsymbol{\mu}: \sum_{j=1}^{J} \mu_j = \mu} p(\boldsymbol{\mu}|\nu) \leq \binom{\mu + J - 1}{J - 1} \binom{M}{\mu/J}^{J} \left(\frac{n-\nu}{n}\right)^{\mu K} \tag{52}$$

Next, we use (13) where for conditional probability $P_{e|\nu} = \Pr\left(\text{rank}(H_I) < \nu | \nu\right)$ we apply estimates (51) and (52). Thereby, we obtain (17) of Theorem 4.

REFERENCES

[1] I. E. Bocharova, B. D. Kudryashov, E. Rosnes, V. Skachek, and Ø. Ytrehus, "Wrap-around sliding-window near-ML decoding of binary LDPC codes over the BEC," in *9th International Symposium on Turbo Codes and Iterative Information Processing (ISTC), 2016*. IEEE, 2016, pp. 16–20.

[2] D. Burshtein and G. Miller, "An efficient maximum-likelihood decoding of LDPC codes over the binary erasure channel," *IEEE Trans. Inf. Theory*, vol. 50, no. 11, pp. 2837–2844, 2004.

[3] E. Paolini, G. Liva, B. Matuz, and M. Chiani, "Maximum likelihood erasure decoding of LDPC codes: Pivoting algorithms and code design," *IEEE Trans. Comm.*, vol. 60, no. 11, pp. 3209–3220, 2012.

[4] M. Cunche, V. Savin, and V. Roca, "Analysis of quasi-cyclic ldpc codes under ml decoding over the erasure channel," in *Int. Symp. on Inform. Theory and its Applications (ISITA)*, 2010, pp. 861–866.

[5] S. Kim, S. Lee, and S.-Y. Chung, "An efficient algorithm for ml decoding of raptor codes over the binary erasure channel," *IEEE Commun. Lett.*, vol. 12, no. 8, 2008.

[6] T. Richardson and R. Urbanke, *Modern coding theory*. Cambridge University Press, 2008.

[7] S. Sankaranarayanan and B. Vasic, "Iterative decoding of linear block codes: A parity-check orthogonalization approach," *IEEE Trans. on Inform. Theory*, vol. 51, no. 9, pp. 3347–3353, 2005.

[8] N. Kobayashi, T. Matsushima, and S. Hirasawa, "Transformation of a parity-check matrix for a message-passing algorithm over the bec," *IEICE Trans. on Fundamentals of electronics, communications and computer sciences*, vol. 89, no. 5, pp. 1299–1306, 2006.

[9] H. Pishro-Nik and F. Fekri, "On decoding of low-density parity-check codes over the binary erasure channel," *IEEE Trans. Inf. Theory*, vol. 50, no. 3, pp. 439–454, 2004.

[10] G. Hosoya, T. Matsushima, and S. Hirasawa, "A decoding method of low-density parity-check codes over the binary erasure channel," in *Proc. 27th Symp. on Inform. Theory and its Applications (ISITA2004)*, 2004, pp. 263–266.

[11] P. M. Olmos, J. J. Murillo-Fuentes, and F. Pérez-Cruz, "Tree-structure expectation propagation for decoding LDPC codes over binary erasure channels," in *2010 IEEE International Symposium on Information Theory Proceedings (ISIT)*, 2010, pp. 799–803.

[12] B. N. Vellambi and F. Fekri, "Results on the improved decoding algorithm for low-density parity-check codes over the binary erasure channel," *IEEE Trans. Inf. Theory*, vol. 53, no. 4, pp. 1510–1520, 2007.

[13] R. G. Gallager, *Low-density parity-check codes*. M.I.T. Press: Cambridge, MA, 1963.

[14] S. Litsyn and V. Shevelev, "On ensembles of low-density parity-check codes: asymptotic distance distributions," *IEEE Trans. Inf. Theory*, vol. 48, no. 4, pp. 887–908, 2002.

[15] *Air Interface for Fixed and Mobile Broadband Wireless Access Systems*, IEEE P802.16e/D12 Draft, Oct. 2005.

[16] *Digital Video Broadcasting (DVB)*, European Telecommunications Standards Institute ETSI EN 302 307, Rev. 1.2.1, Aug. 2009.

[17] C. Di, 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*, vol. 48, no. 6, pp. 1570–1579, 2002.

[18] I. Sason and S. Shamai, *Performance analysis of linear codes under maximum-likelihood decoding: A tutorial*. Now Publishers Inc, 2006.

[19] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2307–2359, 2010.

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

[21] B. D. Kudryashov, "Decoding of block codes obtained from convolutional codes," *Problemy Peredachi Informatsii*, vol. 26, no. 2, pp. 18–26, 1990.

[22] R. Y. Shao, S. Lin, and M. P. Fossorier, "Two decoding algorithms for tailbiting codes," *IEEE Trans. Comm.*, vol. 51, no. 10, pp. 1658–1665, 2003.

[23] V. Tomás, J. Rosenthal, and R. Smarandache, "Decoding of convolutional codes over the erasure channel," *IEEE Trans. Inf. Theory*, vol. 58, no. 1, pp. 90–108, 2012.

[24] I. E. Bocharova, B. D. Kudryashov, V. Skachek, and Y. Yakimenka, "Low complexity algorithm approaching the ML decoding of binary LDPC codes," in *IEEE International Symposium on Information Theory Proceedings (ISIT)*, 2016.

[25] M. Grassl, "Bounds on the minimum distance of linear codes and quantum codes," Online available at http://www.codetables.de, 2007, accessed on 2017-01-06.

[26] G. Landsberg, "Ueber eine anzahlbestimmung und eine damit zusammenhängende reihe." *Journal für die reine und angewandte Mathematik*, vol. 111, pp. 87–88, 1893.

[27] E. R. Berlekamp, "The technology of error-correcting codes," *Proceedings of the IEEE*, vol. 68, no. 5, pp. 564–593, 1980.

[28] S. J. MacMullan and O. M. Collins, "A comparison of known codes, random codes, and the best codes," *IEEE Trans. Inf. Theory*, vol. 44, no. 7, pp. 3009–3022, 1998.

[29] R. Johannesson and K. S. Zigangirov, *Fundamentals of convolutional coding*, 2nd ed. John Wiley & Sons, 2015.

[30] I. Bocharova, B. Kudryashov, and R. Johannesson, "Searching for binary and nonbinary block and convolutional LDPC codes," *IEEE Trans. Inf. Theory*, vol. 62, no. 1, pp. 163–183, Jan. 2016.

[31] I. E. Bocharova, F. Hug, R. Johannesson, and B. D. Kudryashov, "Double-Hamming based QC LDPC codes with large minimum distance," in *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, 2011, pp. 923–927.

[32] I. Sason and R. Urbanke, "Parity-check density versus performance of binary linear block codes over memoryless symmetric channels," *IEEE Transactions on Information Theory*, vol. 49, no. 7, pp. 1611–1635, 2003.