

# Bounds on the Threshold of Linear Programming Decoding

Pascal O. Vontobel

Dept. of EECS

Massachusetts Institute of Technology

Cambridge, MA 02139, USA

pascal.vontobel@ieee.org

Ralf Koetter

CSL and Dept. of ECE

University of Illinois at Urbana-Champaign

Urbana, IL 61801, USA

koetter@uiuc.edu

**Abstract**—Whereas many results are known about thresholds for ensembles of low-density parity-check codes under message-passing iterative decoding, this is not the case for linear programming decoding. Towards closing this knowledge gap, this paper presents some bounds on the thresholds of low-density parity-check code ensembles under linear programming decoding.

## I. INTRODUCTION

Message-passing iterative (MPI) decoding and linear programming (LP) decoding are very efficient methods to achieve excellent decoding performance of low-density parity-check (LDPC) codes on a variety of channels. While an enormous amount of work has been devoted to the understanding of LDPC codes under MPI decoding (see e.g. [1], [2], [3], [4], [5], [6] for results on thresholds), comparably few results on the performance of LDPC codes using the more recent LP decoding are known. In this paper we provide analytical bounds on thresholds of LP decoding; these bounds establish necessary conditions for the existence of LP decoding thresholds.

We note that the existence of such thresholds is not clear *a priori*. In contrast to [7], where we discuss cases where we can guarantee a threshold under LP decoding, here we will show cases where an LP decoding threshold does not exist. E.g., the sequence of random codes where each entry in a parity-check matrix is drawn independently from a Bernoulli distribution with arbitrary nonzero parameter  $\theta$  does not possess an SNR-threshold (in the AWGNC case) or an  $\varepsilon$ -threshold (in the BSC case). Also, there are ensembles of codes that do not show arbitrary low probability of decoding error for any SNR (AWGNC) or any positive  $\varepsilon$  (BSC) even if the code rate is allowed to approach zero and the variable degree is allowed to grow exponentially fast in the block length.

Because LP decoding and MPI decoding are essentially equivalent in the case of the binary erasure channel (BEC), well-known results about MPI decoding for the BEC (see e.g. [2], [3]) can be used for making statements about LP decoding thresholds and therefore we will not deal any further with the BEC case here.

The paper is structured as follows. After having concluded this introduction with some remarks about our notation, we will review ML and LP decoding in Sec. II. The main part

of the paper will be Sec. III where we present so-called 0-neighborhood-based bounds on the LP decoding threshold. Finally, in Sec. IV we will present so-called 2-neighborhood-based bounds.

Let us fix some notation. We let  $\mathbb{R}$ ,  $\mathbb{R}_+$ , and  $\mathbb{R}_{++}$  be the set of real numbers, the set of non-negative real numbers, and the set of positive real numbers, respectively. Moreover, we will use the canonical embedding of the set  $\mathbb{F}_2 = \{0, 1\}$  into  $\mathbb{R}$ . The convex hull (see e.g. [8]) of a set  $\mathcal{A} \subseteq \mathbb{R}^n$  is denoted by  $\text{conv}(\mathcal{A})$ . If  $\mathcal{A}$  is a subset of  $\mathbb{F}_2^n$  then  $\text{conv}(\mathcal{A})$  denotes the convex hull of the set  $\mathcal{A}$  after  $\mathcal{A}$  has been canonically embedded in  $\mathbb{R}^n$ . Similarly, the conic hull (see e.g. [8]) of a set  $\mathcal{A} \subseteq \mathbb{R}^n$  will be denoted by  $\text{conic}(\mathcal{A})$  and if  $\mathcal{A}$  is a subset of  $\mathbb{F}_2^n$  then  $\text{conic}(\mathcal{A})$  denotes the conic hull of the set  $\mathcal{A}$  after  $\mathcal{A}$  has been canonically embedded in  $\mathbb{R}^n$ . The  $i$ -th component of a vector  $\mathbf{x}$  will be called  $[\mathbf{x}]_i$  or  $x_i$  and the element in the  $j$ -th row and  $i$ -th column of a matrix  $\mathbf{A}$  will be called  $[\mathbf{A}]_{j,i}$ .

Let  $\mathcal{C}$  be a binary linear code defined by a parity-check matrix  $\mathbf{H}$  of size  $m$  by  $n$ . Based on  $\mathbf{H}$ , we define the sets  $\mathcal{I} \triangleq \mathcal{I}(\mathbf{H}) \triangleq \{1, \dots, n\}$ ,  $\mathcal{J} \triangleq \mathcal{J}(\mathbf{H}) \triangleq \{1, \dots, m\}$ ,  $\mathcal{I}_j \triangleq \mathcal{I}_j(\mathbf{H}) \triangleq \{i \in \mathcal{I} \mid [\mathbf{H}]_{j,i} = 1\}$  for each  $j \in \mathcal{J}$ , and  $\mathcal{J}_i \triangleq \mathcal{J}_i(\mathbf{H}) \triangleq \{j \in \mathcal{J} \mid [\mathbf{H}]_{j,i} = 1\}$  for each  $i \in \mathcal{I}$ . Moreover, for each  $j \in \mathcal{J}$  we define the codes  $\mathcal{C}_j \triangleq \mathcal{C}_j(\mathbf{H}) \triangleq \{\mathbf{x} \in \mathbb{F}_2^n \mid \mathbf{h}_j \mathbf{x}^T = 0 \pmod{2}\}$ , where  $\mathbf{h}_j$  is the  $j$ -th row of  $\mathbf{H}$ . Note that the code  $\mathcal{C}_j$  is a code of length  $n$  where all positions not in  $\mathcal{I}_j$  are unconstrained. For simplicity of notation, we will never indicate the parity-check matrix as an argument of  $\mathcal{I}$ ,  $\mathcal{J}$ , etc.; it will be clear from the context to what parity-check matrix we are referring to. Finally, by a family of codes we will mean a sequence of (deterministically or randomly generated) codes where the block length goes to infinity.

## II. ML AND LP DECODING

Let us use the above-mentioned code  $\mathcal{C}$  for data transmission over a binary-input discrete memoryless channel with input alphabet  $\mathcal{X} \triangleq \{0, 1\}$ , output alphabet  $\mathcal{Y}$ , and channel law  $P_{Y|X}(y|x)$ . Because the channel is memoryless,  $P_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x}) = \prod_{i \in \mathcal{I}} P_{Y|X}(y_i|x_i)$ , where  $\mathbf{X} \triangleq (X_1, \dots, X_n)$ , where  $\mathbf{Y} \triangleq (Y_1, \dots, Y_n)$ , where the random variable  $X_i$  denotes the channel input at time index  $i$ , and where the random variable  $Y_i$  denotes the channel output at time index  $i$ . Upon observing  $\mathbf{Y} = \mathbf{y}$ , the maximum-likelihood (ML) decoding rule decides

for

$$\hat{\mathbf{x}}(\mathbf{y}) = \arg \max_{\mathbf{x} \in \mathcal{C}} P_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x}).$$

Let the  $i$ -th log-likelihood ratio  $\Gamma_i$ ,  $i \in \mathcal{I}$ , be the random variable

$$\Gamma_i \triangleq \Gamma_i(Y_i) \triangleq \log \left( \frac{P_{Y|X}(Y_i|0)}{P_{Y|X}(Y_i|1)} \right) \in \mathbb{R} \cup \{\pm\infty\}$$

with realization  $\gamma_i \triangleq \gamma_i(y_i)$ . Then, noting that

$$\log P_{Y|X}(y_i|x_i) = -\gamma_i x_i + \log P_{Y|X}(y_i|0),$$

ML decoding can also be written as

$$\hat{\mathbf{x}}(\mathbf{y}) = \arg \min_{\mathbf{x} \in \mathcal{C}} \sum_{i \in \mathcal{I}} \gamma_i x_i.$$

Because the cost function is linear, and a linear function attains its minimum at the extremal points of a convex set, this is essentially equivalent to

$$\hat{\mathbf{x}}(\mathbf{y}) = \arg \min_{\mathbf{x} \in \text{conv}(\mathcal{C})} \sum_{i \in \mathcal{I}} \gamma_i x_i.$$

Although this is a linear program, it can usually not be solved efficiently because its description complexity is usually exponential in the block length of the code.

However, one might try to solve a relaxation of the above minimization problem. Noting that  $\text{conv}(\mathcal{C}) \subseteq \bigcap_{j \in \mathcal{J}} \text{conv}(\mathcal{C}_j)$  (which follows from the fact that  $\mathcal{C} = \bigcap_{j \in \mathcal{J}} \mathcal{C}_j$ ), Feldman, Wainwright, and Karger [9], [10] defined the linear programming decoder (LP decoder) to be given by the solution of the linear program

$$\hat{\omega}(\mathbf{y}) = \arg \min_{\omega \in \bigcap_{j \in \mathcal{J}} \text{conv}(\mathcal{C}_j)} \sum_{i \in \mathcal{I}} \gamma_i \omega_i. \quad (1)$$

Because of its importance, we will abbreviate the set  $\bigcap_{j \in \mathcal{J}} \text{conv}(\mathcal{C}_j)$  by  $\mathcal{P} \triangleq \mathcal{P}(\mathbf{H})$  and call it the fundamental polytope [11], [12]. The fundamental polytope  $\mathcal{P}$  can be expressed using inequalities as follows [9], [10], [11], [12]

$$\mathcal{P} = \left\{ \omega \in \mathbb{R}^n \left| \begin{array}{l} \forall i \in \mathcal{I} : 0 \leq \omega_i \leq 1 \text{ and} \\ \forall j \in \mathcal{J}, \forall \mathcal{I}'_j \subseteq \mathcal{I}_j, |\mathcal{I}'_j| \text{ odd} : \\ \sum_{i \in \mathcal{I}'_j} \omega_i + \sum_{i \in (\mathcal{I}_j \setminus \mathcal{I}'_j)} (1 - \omega_i) \leq |\mathcal{I}_j| - 1 \end{array} \right. \right\}.$$

When analyzing the decoding performance of LP decoding of a binary linear code that is used for data transmission over a binary-input output-symmetric channel, we can without loss of generality assume that the all-zeros codeword was sent. (See also [9] and [10] that discuss this so-called ‘‘ $\mathcal{C}$ -symmetry’’ property.) We observe that a necessary (but usually not sufficient) condition that one decides for the all-zeros codeword in (1) is that<sup>1</sup>

$$\sum_{i \in \mathcal{I}} \gamma_i \omega_i \geq 0 \quad \text{for all } \omega \in \mathcal{P} \setminus \{\mathbf{0}\}. \quad (2)$$

<sup>1</sup>Actually, without changing the content of the statement in (2), we can replace  $\omega \in \mathcal{P} \setminus \{\mathbf{0}\}$  by  $\omega \in \mathcal{P}$ .

It can easily be seen that this condition is equivalent to the condition that

$$\sum_{i \in \mathcal{I}} \gamma_i \omega_i \geq 0 \quad \text{for all } \omega \in \text{conic}(\mathcal{P}) \setminus \{\mathbf{0}\}. \quad (3)$$

The set  $\text{conic}(\mathcal{P})$ , which is the conic hull of the fundamental polytope, is called the fundamental cone  $\mathcal{K} \triangleq \mathcal{K}(\mathbf{H})$ . In terms of inequalities,  $\mathcal{K}$  can be written as [9], [10], [11], [12]

$$\mathcal{K} = \left\{ \omega \in \mathbb{R}^n \left| \begin{array}{l} \forall i \in \mathcal{I} : 0 \leq \omega_i \text{ and} \\ \forall j \in \mathcal{J}, \forall i' \in \mathcal{I}_j : \\ \omega_{i'} - \sum_{i \in (\mathcal{I}_j \setminus \{i'\})} \omega_i \leq 0 \end{array} \right. \right\}.$$

The condition in (3) can then be stated as

$$\sum_{i \in \mathcal{I}} \gamma_i \omega_i \geq 0 \quad \text{for all } \omega \in \mathcal{K} \setminus \{\mathbf{0}\}. \quad (4)$$

We will use the following definition for the block decoding error event under LP decoding: it is the complement of the event that the all-zeros vector is the unique solution in (1).

### III. 0-NEIGHBORHOOD-BASED BOUNDS ON THE THRESHOLD FOR REGULAR LDPC CODES

We focus our attention on  $(w_{\text{col}}, w_{\text{row}})$ -regular LDPC codes, i.e. codes defined by parity-check matrices that have uniform column weight  $w_{\text{col}}$  and uniform row weight  $w_{\text{row}}$ . For these type of codes, we present a technique to obtain bounds on the threshold under LP decoding which we will call 0-neighborhood-based bounds; the choice for this name will become clear later on.

*Assumption 1:* In the following, we will always assume that the all-zeros codeword was sent, i.e. we will not explicitly write the conditioning on  $\mathbf{X} = \mathbf{0}$  when making statements involving probabilities.  $\square$

Under this assumption, the log-likelihood ratios  $\Gamma_1, \dots, \Gamma_n$  are i.i.d. random variables and so there is a random variable  $\Gamma$  such that  $\Gamma_i \sim \Gamma$  for all  $i \in \mathcal{I}$ . Moreover, let  $\mathcal{G} \subseteq (\mathbb{R} \cup \{\pm\infty\})$  be the support of the pdf of  $\Gamma$ .

*Example 2:* Let us discuss the random variable  $\Gamma$  for three channels: the binary-input additive white Gaussian channel (AWGNC), the binary symmetric channel (BSC), and the binary erasure channel (BEC).

- AWGNC (with modulation map  $0 \mapsto +\sqrt{E_c}$ ,  $1 \mapsto -\sqrt{E_c}$  and where the added noise has variance  $\sigma^2$ ):  $\mathcal{G} = \mathbb{R}$  and  $\Gamma$  is a continuous random variable that is normally distributed with mean  $2E_c/\sigma^2$  and variance  $4E_c/\sigma^2$ .
- BSC (with cross-over probability  $\varepsilon$ ):  $\mathcal{G} = \{\pm G\}$  and  $\Gamma$  is a discrete random variable that takes on the value  $G$  with probability  $1 - \varepsilon$  and the value  $-G$  with probability  $\varepsilon$ , where  $G \triangleq \log\left(\frac{1-\varepsilon}{\varepsilon}\right)$ .
- BEC (with erasure probability  $\varepsilon$ ):  $\mathcal{G} = \{0, +\infty\}$  and  $\Gamma$  is a discrete random variable that takes on the value  $+\infty$  with probability  $1 - \varepsilon$  and the value  $0$  with probability  $\varepsilon$ .  $\square$

*Definition 3:* Let

$$\Gamma_{\text{pos}} \triangleq + \sum_{\substack{i \in \mathcal{I} \\ \Gamma_i \geq 0}} \Gamma_i \quad \text{and} \quad \Gamma_{\text{neg}} \triangleq - \sum_{\substack{i \in \mathcal{I} \\ \Gamma_i < 0}} \Gamma_i$$

be random variables with realizations  $\gamma_{\text{pos}}$  and  $\gamma_{\text{neg}}$ , respectively. Note that  $\Gamma_{\text{pos}} \geq 0$  and  $\Gamma_{\text{neg}} \geq 0$  w.p. 1.  $\square$

*Lemma 4:* With probability one we have

$$\lim_{n \rightarrow \infty} \frac{\Gamma_{\text{pos}}}{n} = +E[\Gamma | \Gamma \geq 0] \cdot \Pr(\Gamma \geq 0) = + \int_{0^-}^{+\infty} \gamma p_{\Gamma}(\gamma) d\gamma,$$

$$\lim_{n \rightarrow \infty} \frac{\Gamma_{\text{neg}}}{n} = -E[\Gamma | \Gamma < 0] \cdot \Pr(\Gamma < 0) = - \int_{-\infty}^{0^-} \gamma p_{\Gamma}(\gamma) d\gamma.$$

*Proof:* Follows easily from the weak law of large numbers.  $\square$

*Lemma 5:* Consider a code with a  $(w_{\text{col}}, w_{\text{row}})$ -regular parity-check matrix. Let  $\gamma \in \mathcal{G}^n$ . A necessary condition that the LP decoder decides in favor of the all-zeros codeword is

$$\frac{\gamma_{\text{pos}}}{\gamma_{\text{neg}}} \geq w_{\text{row}} - 1.$$

*Proof:* We saw in (4) that a necessary condition for LP decoding to decide in favor of the all-zeros codeword is that  $\sum_{i \in \mathcal{I}} \gamma_i \omega_i \geq 0$  for all  $\omega \in \mathcal{K} \setminus \{\mathbf{0}\}$ . Let us construct a vector  $\omega \in \mathbb{R}^n$  as follows:

$$\omega_i \triangleq \begin{cases} \frac{1}{w_{\text{row}} - 1} & \text{if } \gamma_i \geq 0 \\ 1 & \text{if } \gamma_i < 0 \end{cases}. \quad (5)$$

It can easily be seen that  $\omega \in \mathcal{K}(\mathbf{H})$ .<sup>2</sup> We obtain the necessary condition

$$0 \leq \sum_{i \in \mathcal{I}} \gamma_i \omega_i = -1 \cdot \gamma_{\text{neg}} + \frac{1}{w_{\text{row}} - 1} \cdot \gamma_{\text{pos}},$$

which is equivalent to the necessary condition in the lemma statement.  $\square$

*Theorem 6:* Consider a family of codes that have  $(w_{\text{col}}, w_{\text{row}})$ -regular parity-check matrices. In the limit  $n \rightarrow \infty$ , a necessary condition such that the LP decoder decides in favor of the all-zeros codeword with probability one is

$$-\frac{E[\Gamma | \Gamma \geq 0]}{E[\Gamma | \Gamma < 0]} \cdot \frac{\Pr(\Gamma \geq 0)}{\Pr(\Gamma < 0)} \geq w_{\text{row}} - 1,$$

or, equivalently,

$$-\frac{\int_{0^-}^{+\infty} \gamma p_{\Gamma}(\gamma) d\gamma}{\int_{-\infty}^{0^-} \gamma p_{\Gamma}(\gamma) d\gamma} \geq w_{\text{row}} - 1.$$

*Proof:* This follows upon combining Lemmas 4 and 5.  $\square$

*Corollary 7:* Consider the setup of Th. 6. If the memoryless channel is a BSC with cross-over probability  $\varepsilon$  then the necessary condition in Th. 6 reads

$$\varepsilon \leq \frac{1}{w_{\text{row}}}.$$

*Proof:* For a BSC with cross-over probability  $\varepsilon$  we obtain

$$E[\Gamma | \Gamma \geq 0] \cdot \Pr(\Gamma \geq 0) = +G \cdot (1 - \varepsilon),$$

$$-E[\Gamma | \Gamma < 0] \cdot \Pr(\Gamma < 0) = -(-G) \cdot \varepsilon = +G \cdot \varepsilon$$

where  $G$  is defined as in Ex. 2. Therefore, the condition in Th. 6 is that  $\frac{G(1-\varepsilon)}{G\varepsilon} \geq w_{\text{row}} - 1$ , i.e.  $\varepsilon \leq \frac{1}{w_{\text{row}}}$ .  $\square$

<sup>2</sup>The vector  $\omega$  can be seen as a generalization of the so-called canonical completion [11], [12], however instead of assigning values according to the graph distance with respect to a single node, we assign values according to the graph distance with respect to the set of nodes where  $\gamma_i$  is negative. Moreover, we assign only the values 1 and  $1/(w_{\text{row}} - 1)$ .

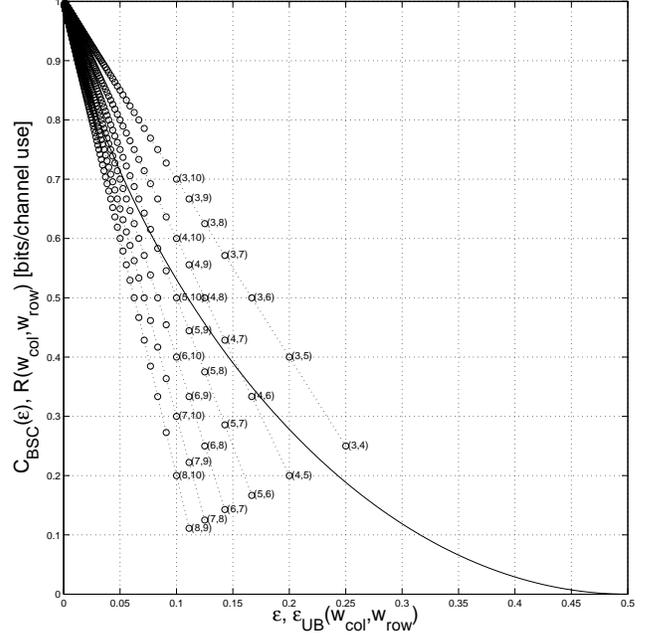


Fig. 1. The solid line shows the capacity  $C_{\text{BSC}}(\varepsilon)$  of a BSC as a function of the cross-over probability  $\varepsilon$ . The circles have the following meaning: the circle with label  $(w_{\text{col}}, w_{\text{row}})$  shows the point  $(R(w_{\text{col}}, w_{\text{row}}), \varepsilon_{\text{UB}}(w_{\text{col}}, w_{\text{row}}))$ , where  $R(w_{\text{col}}, w_{\text{row}})$  and  $\varepsilon_{\text{UB}}(w_{\text{col}}, w_{\text{row}})$  are the designed rate and the threshold upper bound from Cor. 7, respectively, for  $(w_{\text{col}}, w_{\text{row}})$ -regular LDPC codes.

*Example 8:* Fig. 1 tries to capture some of the implications of Th. 6 / Cor. 7. The circles in the plot that are to the left of the capacity curve yield non-trivial upper bounds on the error correction capability of LP decoding. (Note that the designed rate of a  $(w_{\text{col}}, w_{\text{row}})$ -regular LDPC code is  $1 - \frac{w_{\text{col}}}{w_{\text{row}}}$  and that the actual rate is lower bounded by this quantity.)  $\square$

*Example 9:* It is a surprising fact that the bounds in Th. 6 and Cor. 7 do not depend on the variable degree  $w_{\text{col}}$  at all. In particular, this implies that decoding would fail even if we choose extremely large variable degrees. For example we might consider a sequence of codes defined by parity-check matrices that contain *all* rows of a given weight  $w_{\text{row}}$ . Clearly, a specific code  $\mathcal{C}_n$  of this sequence is a  $\left(\binom{n-1}{w_{\text{row}}-1}, w_{\text{row}}\right)$ -regular code which contains either one ( $\mathcal{C}_n = \{\mathbf{0}\}$ ) or two codewords ( $\mathcal{C}_n = \{\mathbf{0}, \mathbf{1}\}$ ) depending on  $k$  being odd or even. Thus, while the rate of this code sequence approaches zero, LP decoding will not succeed for  $\varepsilon > \frac{1}{w_{\text{row}}}$ .  $\square$

*Example 10:* Th. 6 and Cor. 7 can easily be extended to families of codes where the row weight grows as a function of  $n$ ; let us call these extensions Th. 6' and Cor. 7'. It is clear from Cor. 7' that there cannot be an LP decoding threshold for the BSC if the row weight grows unboundedly. Moreover, coming back to the code family in Ex. 9, if  $w_{\text{row}}$  is allowed to grow with  $n$ , LP decoding will fail as  $n \rightarrow \infty$  despite the variable degree being exponentially larger than the check degree.

*Example 11:* Similarly, the family of random codes where entries in a parity-check matrix are drawn independently from

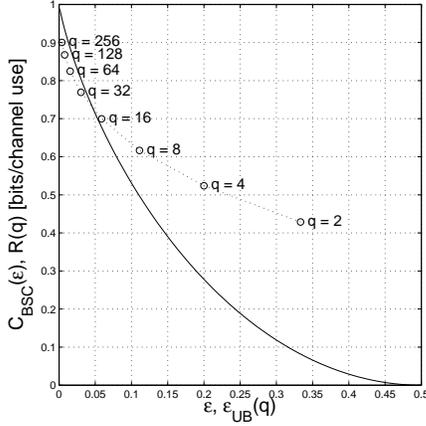


Fig. 2. The solid line shows the capacity  $C_{\text{BSC}}(\varepsilon)$  of a BSC as a function of the cross-over probability  $\varepsilon$ . The circles have the following meaning: the circle with label  $q$  shows the point  $(R(q), \varepsilon_{\text{UB}}(q))$ , where  $R(q)$  is the rate of the PG(2,  $q$ )-based code and where  $\varepsilon_{\text{UB}}(q)$  is the threshold upper bound from Cor. 7 for a family of  $(w_{\text{col}}(q), w_{\text{row}}(q))$ -regular LDPC codes. (See also the main text for further explanations.)

a Bernoulli( $\theta$ ) distribution will not only have poor threshold performance under LP decoding but will fail with high probability as the code length approaches infinity for *any* symmetric channel for which the expression

$$-\frac{\mathbb{E}[\Gamma | \Gamma \geq 0]}{\mathbb{E}[\Gamma | \Gamma < 0]} \cdot \frac{\Pr(\Gamma \geq 0)}{\Pr(\Gamma < 0)}$$

is (upper) bounded. The result follows from the observation that the weight of the rows in  $\mathbf{H}$  is exponentially concentrated around  $\theta n$ . Indeed, given a vector of log-likelihood ratios, the vector with components  $\frac{1}{n\theta - \delta - 1}$  in positions where  $\gamma_i$  is non-negative and 1 in the remaining positions is inside  $\mathcal{K}(\mathbf{H})$  with high probability for  $\delta > 0$  and  $n \rightarrow \infty$ .  $\square$

While the above considerations give some insight in the asymptotic behavior of decoding error for LP decoding, the characterization and spirit of Th. 6 is essentially combinatorial.

*Example 12:* We saw in Ex. 10 that for any family of codes where the row weight grows as a function of  $n$ , Cor. 7' implies that there cannot be an LP decoding threshold for the BSC. A special case, though, arises when the rate of the code family under consideration goes to 1 when  $n \rightarrow \infty$  because then also the best code family under the best possible decoding algorithm can only correct a vanishing fraction of bit flips as  $n \rightarrow \infty$ . A family where the rate goes to 1 as  $n \rightarrow \infty$  is the family of type-I PG(2,  $q$ )-based codes, cf. [13]; in the context of LP decoding, these codes were analyzed in [14], [15]. A code from this family is indexed by  $q$  (where  $q \triangleq 2^s$  for some positive integer  $s$ ), has length  $n(q) \triangleq q^2 + q + 1$ , rate  $R(q) \triangleq 1 - (3^s + 1)/(q^2 + q + 1)$ , and  $w_{\text{col}}(q) = w_{\text{row}}(q) = q + 1$ .

Fig. 2 shows the following: for each  $q$  we plot the point  $(R(q), \varepsilon_{\text{UB}}(q))$ , where  $R(q)$  is the rate of the PG(2,  $q$ )-based code and where  $\varepsilon_{\text{UB}}(q)$  is the LP decoding threshold upper bound from Cor. 7 for a  $(w_{\text{col}}(q), w_{\text{row}}(q))$ -regular family of codes. Care must be taken when giving an interpretation to this figure since the PG(2,  $q$ )-based codes are finite-length codes

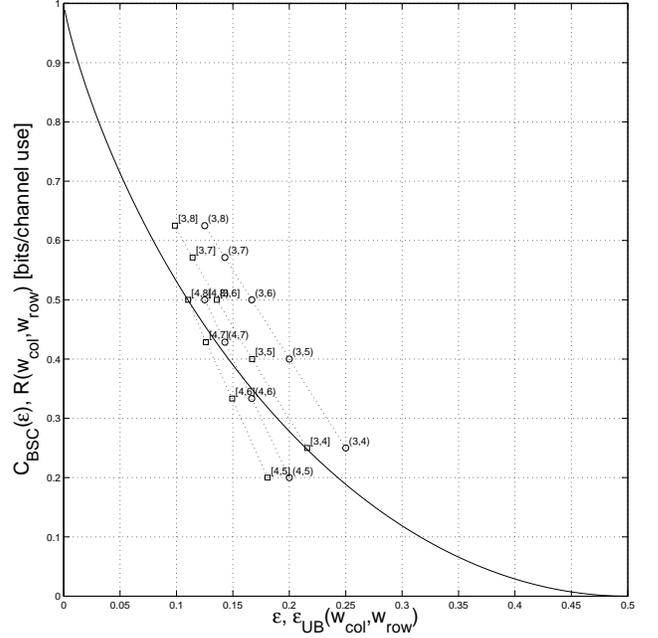


Fig. 3. The solid line shows the capacity  $C_{\text{BSC}}(\varepsilon)$  of a BSC as a function of the cross-over probability  $\varepsilon$ . The circles have the same meaning as in Fig. 1. The squares have the following meaning: the square with label  $q$  shows the point  $(R(q), \varepsilon_{\text{UB}}(q))$ , where  $R(q)$  is the rate of the PG(2,  $q$ )-based code and where  $\varepsilon_{\text{UB}}(q)$  is the 2-neighborhood-based threshold upper bound from Sec. IV for a family of  $(w_{\text{col}}(q), w_{\text{row}}(q))$ -regular LDPC codes. (See also the main text for further explanations.)

for finite  $q$ .  $\square$

We leave it as an exercise for the reader to generalize the results in this section to irregular LDPC codes.

#### IV. 2-NEIGHBORHOOD-BASED BOUNDS ON THE THRESHOLD FOR REGULAR LDPC CODES

Because the assignment of a value to  $\omega_i$  in (5) was only based on the value of  $\gamma_i$ , we call the resulting bound in Cor. 7 a 0-neighborhood-based bound. (Of course, the way we assigned a value to every  $\omega_i$  in (5) can also be seen as a very simplistic, and usually sub-optimal way, of solving the linear program in (1).) It is natural to try to formulate more sophisticated assignments of a value to  $\omega_i$ . The next simplest approach is to formulate a rule that does not depend on  $\gamma_i$  only, but also on  $\gamma_{i'}$  where  $i'$  ranges over all variable nodes at Tanner graph distance 2 from variable node  $i$ . The resulting bounds on the threshold will therefore be called 2-neighborhood-based bounds.

For  $i \in \mathcal{I}$  let  $\mathcal{N}_i^{(2)}$  be the subset of  $\mathcal{I}$  that includes all variable nodes  $i'$  with Tanner graph distance at most 2 from  $i$ . In the following, we assume that the Tanner graph under consideration has girth at least six. In the case of a  $(w_{\text{col}}, w_{\text{row}})$ -regular LDPC codes, this implies that the set  $\mathcal{N}_i^{(2)}$  has size  $|\mathcal{N}_i^{(2)}| = 1 + w_{\text{col}}(w_{\text{row}} - 1)$ . (Fig. 4 (left) shows part of a (3, 4)-regular code, i.e. node  $i$ , all check nodes at Tanner graph distance 1 from  $i$ , and all variable nodes at Tanner graph distance 2 from  $i$ .)

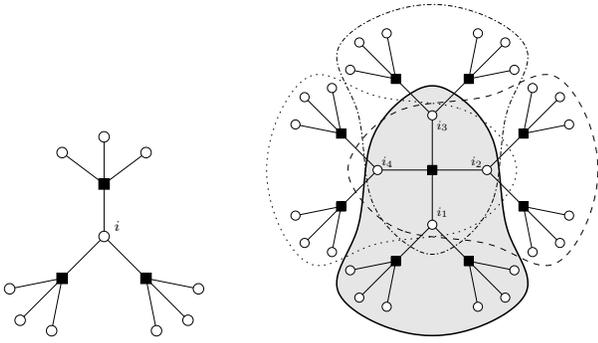


Fig. 4. Left: 2-neighborhood. Right: overlapping 2-neighborhoods. (See Sec. IV for more explanations.)

*Definition 13:* Let  $\Gamma'_i$  be the vector that contains all the random variables  $\{\Gamma_{i'}\}_{i' \in \mathcal{N}_i^{(2)}}$  and let  $\gamma'_i$  be its realization. Our new rule (that replaces (5)) for defining a vector  $\omega \in \mathbb{R}^n$  is now  $\omega_i \triangleq \alpha_{\gamma'_i}$ , where  $\alpha_{\gamma'_i} \in \mathbb{R}_+$  for all  $\gamma'_i \in \mathcal{G}^{|\mathcal{N}_i^{(2)}|}$  is chosen such that for all possible  $\gamma \in \mathcal{G}^n$  we obtain a vector  $\omega$  that lies in the fundamental cone.

*Lemma 14:* Assume that we have such values  $\{\alpha_{\gamma'_i}\}_{\gamma'_i}$  as defined in Def. 13. With probability one we have

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i \in \mathcal{I}} \gamma_i \omega_i = \sum_{\gamma'_i} p_{\gamma'_i} \alpha_{\gamma'_i},$$

where

$$p_{\gamma'_i} \triangleq \Pr(\{\Gamma_{i'}\}_{i' \in \mathcal{N}_i^{(2)}} = \{\gamma_{i'}\}_{i' \in \mathcal{N}_i^{(2)}}) = \prod_{i' \in \mathcal{N}_i^{(2)}} \Pr(\Gamma_{i'} = \gamma_{i'}).$$

*Proof:* Follows from the fact that  $\gamma_i \omega_i$  depends only on finitely many  $\gamma_{i'}$  from  $\{\gamma_{i'}\}_{i' \in \mathcal{I}}$  and from the use of the weak law of large numbers.  $\square$

Consider a BSC with cross-over probability  $\varepsilon$ . An upper bound on the LP decoding threshold for the BSC is then given by the infimum of all  $\varepsilon$  where we are able to find an assignment in Def. 13 such that  $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i \in \mathcal{I}} \gamma_i \omega_i$  is negative with probability one. Finding such assignments can e.g. be done by solving a linear program that roughly looks as follows

$$\min. \quad \sum_{\gamma'_i} p_{\gamma'_i} \alpha_{\gamma'_i}$$

$$\text{subj. to } \alpha_{\gamma'_i} = 1,$$

and for each  $\gamma$  the assignment always results in a non-zero vector that lies in the fundamental cone,

where  $\tilde{\gamma}'_i$  is an arbitrary assignment of values to  $\gamma'_i$ , e.g.  $\gamma_{i'} = -G$  for all  $i' \in \mathcal{N}_i^{(2)}$ , where  $G$  is defined as in Ex. 2.

The Tanner graph in Fig. 4 (left) has many symmetries that can be used to simplify the above linear program. E.g. if there is a graph isomorphism that maps an assignment  $\tilde{\gamma}'_i$  to an assignment  $\overline{\gamma}'_i$ , then without loss of generality we can assume that  $\alpha_{\tilde{\gamma}'_i} = \alpha_{\overline{\gamma}'_i}$ . In this way, the dimensionality of the above linear program can be reduced significantly. Without going into the details, the requirement that the resulting assignment always results in a vector in the fundamental cone can be

simplified by introducing some auxiliary variables according to overlapping 2-neighborhoods as is sketched in Fig. 4 (right).

*Example 15:* Fig. 3 shows the improved upper bounds for  $(w_{\text{col}}, w_{\text{row}})$ -regular code families.  $\square$

Whereas the above approach results in relatively small linear programs for small  $w_{\text{col}}$  and  $w_{\text{row}}$ , similar 4-, 6-, 8-, etc., neighborhood-based approaches seem to be computationally much more demanding. We leave it as an open question to see if there are ways to handle also these cases in an efficient numerical way.

## ACKNOWLEDGMENTS

P.O.V.'s research was supported by NSF Grant CCF-0514801. R.K.'s research was supported by NSF Grant CCF-0514869.

## REFERENCES

- [1] N. Wiberg, *Codes and Decoding on General Graphs*. PhD thesis, Linköping University, Sweden, 1996.
- [2] M. G. Luby, M. Mitzenmacher, M. A. Shokrollahi, D. A. Spielman, and V. Stemann, "Practical loss-resilient codes," in *Proc. 29th Annual ACM Symp. on Theory of Computing*, pp. 150–159, 1997.
- [3] M. G. Luby, M. Mitzenmacher, and M. A. Shokrollahi, "Analysis of random processes via and-or tree evaluation," in *Proc. 9th Annual ACM-SIAM Symp. on Discrete Algorithms*, pp. 364–373, 1998.
- [4] T. J. Richardson and R. L. Urbanke, "The capacity of low-density parity-check codes under message-passing decoding," *IEEE Trans. on Inform. Theory*, vol. IT-47, no. 2, pp. 599–618, 2001.
- [5] M. Lentmaier, D. V. Truhachev, D. J. Costello, Jr., and K. Zigangirov, "On the block error probability of iteratively decoded LDPC codes," in *5th ITG Conference on Source and Channel Coding*, (Erlangen, Germany), Jan. 14–16 2004.
- [6] H. Jin and T. Richardson, "Block error iterative decoding capacity for LDPC codes," in *Proc. IEEE Intern. Symp. on Inform. Theory*, (Adelaide, Australia), pp. 52 – 56, Sep. 4–9 2005.
- [7] R. Koetter and P. O. Vontobel, "On the block error probability of LP decoding of LDPC codes," in *Proc. Inaugural Workshop of the Center for Information Theory and its Applications*, (UC San Diego, La Jolla, CA, USA), Feb. 6–10 2006.
- [8] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge University Press, 2004.
- [9] J. Feldman, *Decoding Error-Correcting Codes via Linear Programming*. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA, 2003. Available online under <http://www.columbia.edu/~jef2189/pubs.html>.
- [10] J. Feldman, M. J. Wainwright, and D. R. Karger, "Using linear programming to decode binary linear codes," *IEEE Trans. on Inform. Theory*, vol. IT-51, no. 3, pp. 954–972, 2005.
- [11] R. Koetter and P. O. Vontobel, "Graph covers and iterative decoding of finite-length codes," in *Proc. 3rd Intern. Symp. on Turbo Codes and Related Topics*, (Brest, France), pp. 75–82, Sept. 1–5 2003.
- [12] P. O. Vontobel and R. Koetter, "Graph-cover decoding and finite-length analysis of message-passing iterative decoding of LDPC codes," *submitted to IEEE Trans. Inform. Theory*, available online under <http://www.arxiv.org/abs/cs.IT/0512078>, Dec. 2005.
- [13] Y. Kou, S. Lin, and M. P. C. Fossorier, "Low-density parity-check codes based on finite geometries: a rediscovery and new results," *IEEE Trans. on Inform. Theory*, vol. IT-47, pp. 2711–2736, Nov. 2001.
- [14] P. O. Vontobel, R. Smarandache, N. Kiyavash, J. Teutsch, and D. Vukobratovic, "On the minimal pseudo-codewords of codes from finite geometries," in *Proc. IEEE Intern. Symp. on Inform. Theory*, (Adelaide, Australia), pp. 980–984, Sep. 4–9 2005. Available online under <http://www.arxiv.org/abs/cs.IT/0508019>.
- [15] P. O. Vontobel and R. Smarandache, "On minimal pseudo-codewords of Tanner graphs from projective planes," in *Proc. 43rd Allerton Conf. on Communications, Control, and Computing*, (Allerton House, Monticello, Illinois, USA), Sep. 28–30 2005. Available online under <http://www.arxiv.org/abs/cs.IT/0510043>.