Zbornik radova 49. Konferencije za ETRAN, Budva, 5-10. juna 2005, tom II
Proc. 49th ETRAN Conference, Budva, June 5-10, 2005, Vol. II

ON THE LP DECODING REGIONS

Dejan Vukobratović, Nemanja Delić, Vojin Šenk, *Faculty of Technical Sciences, University of Novi Sad*

Abstract – *Recent introduction of Linear Programming (LP) decoding and connections established between this method and iterative Belief Propagation (BP) decoding, motivates further exploration of LP decoding. Unlike BP decoding on code graphs with cycles, where convergence behavior of decoding process is still unexplained, LP decoding offers clear (geometric) interpretation. Some aspects of this interpretation are examined, commented and verified.*

1. INTRODUCTION

Codes defined on graphs and iterative BP decoding proved to be a capacity-approaching combination. The decade of theoretical research and simulation studies brought many explanations and insights into the BP decoding process. However, convergence behavior of BP decoding on a general code graphs with cycles (where BP decoding is known to be suboptimal) is very hard to characterize. In other words, for a given code, channel and BP decoder limited by some fixed number of iterations, the mapping of the decoder's input space (i.e. channel output space) into the set of corresponding outputs characterized by their decoding regions is still not resolved using analytical tools.

The great breakthrough in this area has been accomplished by complete characterization of BP decoding behavior in the case of the binary erasure channel (BEC), by introducing the notion of *stopping sets* [1]. Based on similar ideas, in the AWGN channel case, the concept of *near-codewords* or *trapping sets* was introduced [2]. Trapping sets yielded significantly to the general understanding of BP decoding and pointed out to the structures (subgraphs) in code graph to be avoided in code construction, but the complete characterization of the decoding regions in the AWGN case is still lacking.

A related direction in understanding of the BP decoding behavior, that in a way generalizes the notions of stopping and trapping sets, is to investigate the pseudo-codewords of a particular graphs constructed appropriately from the code graph. This method was pioneered by Wiberg [3] who introduced the pseudo-codewords on computation trees, and later refined by Koetter and Vontobel [4] using graph covers. The pseudo codewords of the finite graph covers, that interfere with the "original" codewords defined by the code graph, have a simple geometric characterization given by a so called fundamental cone [4].

LP decoding [5] is a novel method for decoding codes on graphs. Formulated as a linear program, the decoding problem description is as complex as the description of the convex domain of optimization is. If we want optimal (ML decoding), the domain of optimization is the convex hull of all the codewords, which is of exponential descriptive complexity. Usual approach in efficiently solving such a problems is to introduce relaxations, i.e. to optimize over a less complex domain. In [5], appropriate relaxations are proposed for several classes of codes on graphs, including low-density parity-check (LDPC) codes that particularly attracts our attention. A new, relaxed domain of optimization is defined by a number of inequality constraints which is linear in the code length, which enables solving LP optimization even for a moderately sized code lengths, using quadratic complexity LP solving algorithms (simplex, interior-point). This complexity, however, is still not comparable with BP decoding.

The reason why LP decoder is particularly interesting is in its convex domain of optimization, referred to as the fundamental polytope. It represents a convex polytope that includes all codewords as its vertices, but introduces even more (non-codeword) vertices that can be produced at the output of LP decoder. Together with the original codewords they form a set of LP pseudocodewords. Additionally, the fundamental cone associated with the fundamental polytope obtained by usual LP relaxation method for binary linear codes [5,6] is exactly the same fundamental cone that characterizes pseudocodewords on graph covers [4]. Tight relations between suboptimal behavior of LP and BP decoders is topic that receives a lot of attention [4,7] and deserves to be further investigated.

The organization of this paper is as follows. Section 2 gives a short introduction to the LP decoding. In section 3 geometric interpretation of LP decoding regions is given. In section 4 fundamental aspects of LP pseudocode are analyzed and some illustrative examples are given. Concluding remarks are given in section 5.

2. LP DECODING [5]

Standard problem of communication over the channel with noise is considered. As a channel model we could use any memoryless binary-input channel, but we will focus mainly on standard binary-input AWGN channel.

A codeword $\mathbf{x}$ from a binary linear (n,k) code C is selected for transmission. The code is given by its $m\times n$ parity-check matrix $\mathbf{H}$. In a special but significant case, when $\mathbf{H}$ is sparse, code C is called an LDPC code. At the channel output a vector $\mathbf{y}$ of elements of channel output alphabet is observed in accordance with the channel transition probabilities $p(y|x)$. The decoder's task is to make a guess $\hat{\mathbf{x}}$ of the codeword $\mathbf{x}$ selected at the input of the channel, given the output $\mathbf{y}$.

Optimal decoder in the case of equiprobable selection of codewords at the input of the channel is the maximum-likelihood (ML) decoder that selects:

$$\hat{\mathbf{x}} = arg \max_{\mathbf{x}\in \mathrm{C}} p(\mathbf{y}\,|\,\mathbf{x})\,. \tag{1}$$

In a several steps [5] it is easy to transform (1) into a linear objective function, giving to the ML decoding an LP form:

$$\hat{\mathbf{x}} = arg \min_{\mathbf{x}\in \mathrm{CH(C)}} \sum_{i=1}^{n} \lambda_i \cdot x_i \tag{2}$$

where the convex domain of optimization is the convex hull of all codeword vectors of code C in R^n, CH(C), and λ_i is a log-likelihood ratio (LLR) value associated with the channel output y_i:

$$\lambda_i = log \frac{p(y_i \mid x_i = 0)}{p(y_i \mid x_i = 1)}. \quad (3)$$

Using (2), LP decoding fails to give a practical ML decoder due to an exponential number of constrains needed to describe CH(C). To achieve a practical decoder, the domain of optimization with linear number of constraints is needed. Obviously, this domain should contain all codewords from C as its vertices (and therefore, since convex, to contain the whole CH(C)).

A standard way [6] to replace (relax) CH(C) of binary linear code C with appropriate polytope is as follows. First, a local code C_i is defined as a single-parity check (SPC) code defined by $\mathbf{H}_i = [\mathbf{h}_i]$, where $\mathbf{h}_i$ is i-th row of parity-check matrix $\mathbf{H}$. CH(C_i) is convex hull of all codewords of C_i. Note that CH(C_i) is much larger object than CH(C) since its vertices are all vertices of unit n-hypercube except those that do not satisfy "parity-constraint" on a (usually) small subset of coordinates indicated by position of ones in $\mathbf{h}_i$. Then, a relaxed polytope P is defined as the intersection of convex hulls CH(C_i) of all local codes C_i, $i = 1, \ldots, m$, i.e.

$$P = \bigcap_{i=1}^{m} \mathrm{CH}(C_i), \qquad i = 1, \ldots, m \quad (4)$$

Having defined a new domain for optimization, the fundamental polytope P [6,4], an LP decoder produces $\hat{\mathbf{x}}$ given $\boldsymbol{\lambda} = \boldsymbol{\lambda}(\mathbf{y})$ as:

$$\hat{\mathbf{x}} = arg \min_{\mathbf{x} \in P} \sum_{i=1}^{n} \lambda_i \cdot x_i \quad (5)$$

Number of constraints needed to describe CH(C_i) grows exponentially with number of ones in $\mathbf{h}_i$ (i.e. with Hamming weight $w_H(\mathbf{h}_i)$). Therefore, number of ones per row of $\mathbf{H}$ of binary linear code C to be used with LP decoder should grow slower than linearly with n, as in the case of LDPC codes.

Fundamental polytope P contains CH(C) and therefore all the codewords of C as its vertices. Unfortunately, the set of all vertices of P is usually larger than C. Therefore, the LP decoder defined as (5) is suboptimal and the set of its possible outputs is the set of vertices $V(P)$ of fundamental polytope P. We will refer to this set as LP pseudocode; the set of all LP pseudocodewords. Fortunately, the set of integral vertices (vertices with all coordinates taking values from set {0,1}) of LP pseudocode is exactly C, which gives to LP decoder a *ML certificate property* [5]: if LP decoder produces integral $\hat{\mathbf{x}}$ at its output, than $\hat{\mathbf{x}}$ is ML codeword. The set $F(P) = V(P) \backslash C$ is a set of non-integral or fractional LP pseudocodewords.

3. LP DECODING REGIONS

The LP decoder is considered as a block in communication system. We assume that the input to this block is a vector $\boldsymbol{\lambda}$, so we are considering LLR input space. For an AWGN channel with parameter σ, LLR space is linearly related with the channel output space ($\boldsymbol{\lambda} = 2\mathbf{y}/\sigma^2$) so we could equivalently take an AWGN channel output space (signal-space) as the input space. However, using LLR space as input space is more general (channel independent).

Since LP solution is always a vertex of domain of optimization, the set of outputs from LP decoder is the set of pseudocodewords $V(P)$. Therefore, LP decoder, as given by (5), is a function between following sets:

$$\mathrm{LP}(\boldsymbol{\lambda}) : D(\lambda)^n \to V(P), \quad (6)$$

where $D(\lambda)$ is the set of all possible values of λ (LLR domain) that is obviously channel dependent (e.g. for AWGN channel $D(\lambda) = R$, for BEC channel $D(\lambda) = \{0,\infty\}$, etc.)

For every LP decoder output $\hat{\mathbf{x}} \in V(P)$ the LP decoding region w.r.t. pseudocodeword $\hat{\mathbf{x}}$ is defined as:

$$\mathrm{D}_{\hat{\mathbf{x}}}^{\mathrm{LP}} = \{\boldsymbol{\lambda} \mid \mathrm{LP}(\boldsymbol{\lambda}) = \hat{\mathbf{x}}\} = \{\boldsymbol{\lambda} \mid \boldsymbol{\lambda} \cdot \hat{\mathbf{x}} \le \boldsymbol{\lambda} \cdot \mathbf{x}, \forall \mathbf{x} \ne \hat{\mathbf{x}}, \mathbf{x} \in V(P)\} \quad (7)$$

Of particular interest are the LP decoding regions w.r.t codewords from C. Since fundamental polytope P satisfies so called C-symmetry [5], it turns out that these are all the same, so it is possible to analyze only the one that corresponds to the *all-zero codeword* **0**:

$$\mathrm{D}_{\mathbf{0}}^{\mathrm{LP}} = \{\boldsymbol{\lambda} \mid \boldsymbol{\lambda} \cdot \mathbf{x} \ge \mathbf{0}, \forall \mathbf{x} \ne \mathbf{0}, \mathbf{x} \in V(P)\} \quad (8)$$

From (8) it is apparent that $\mathrm{D}_0^{\mathrm{LP}}$ is a convex cone with apex in origin. Most of the hyperplanes $\boldsymbol{\lambda}\mathbf{x} \ge 0$ given in (8) are actually redundant. The set of the hyperplanes that describe the faces of $\mathrm{D}_0^{\mathrm{LP}}$ are defined by a subset $M(P)$ of $V(P)$ that is called the set of minimal LP pseudocodewords [8], therefore:

$$\mathrm{D}_{\mathbf{0}}^{\mathrm{LP}} = \{\boldsymbol{\lambda} \mid \boldsymbol{\lambda} \cdot \mathbf{x} \ge \mathbf{0}, \forall \mathbf{x} \ne \mathbf{0}, \mathbf{x} \in M(P)\}. \quad (9)$$

It is interesting to relate the all-zero codeword LP decoding region $\mathrm{D}_0^{\mathrm{LP}}$ with the fundamental polytope P. In order to do this it is necessary to construct the so called conic hull of P. This object is known as the fundamental cone [4]:

$$K(P) = \{\alpha \mathbf{x} \mid \alpha \ge 0, \mathbf{x} \in P\} \quad (10)$$

Edges of the fundamental cone are defined by the set of rays starting at the origin and going through the set $M(P)$. Now, the convex cone $\mathrm{D}_0^{\mathrm{LP}}$ is exactly the dual cone of the fundamental cone $K(P)$. The relations between different convex objects are sketched on a Figure 1. Additionally, it is important to note that every set defined so far is a function of only one mathematical object, the parity check matrix **H** [8].

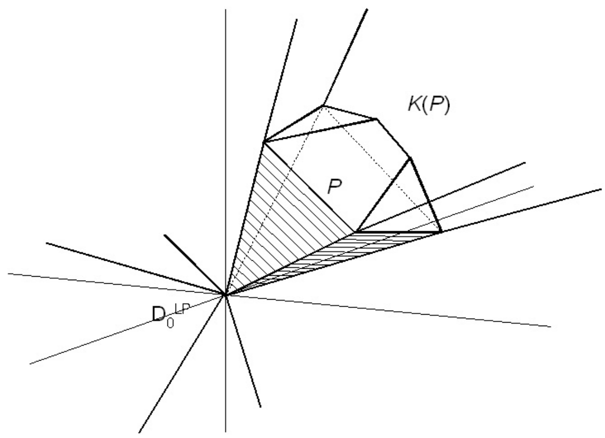


Figure 1. *LP decoding region*

4. LP PSEUDOCODE PARAMETERS

From the previous section it is evident that word error rate performance (WER) of LP decoder is fully defined by the set $M(P)$. However, for the bit error rate (BER) performance it is necessary to take into account the whole set $V(P)$.

In both cases, the influence of the particular pseudocodeword on the error performance for binary-input output-symmetric channels should be described by its "distance" from the transmitted codeword. We will consider only binary linear codes and symmetric decoders so the transmitted codeword is the all zero codeword **0**, and the word "distance" is replaced with the word "weight", i.e. distance from **0**. In the case of binary linear code C and ML decoding, Hamming weight $w_H(\mathbf{x})$, $\mathbf{x} \in C$, is the distance metric of interest. For the pseudocodewords of suboptimal decoders (e.g. BP and LP) a generalization of this metric, called pseudo-weight, had to be introduced [3]:

$$w_P^{\mathrm{AWGN}}(\mathbf{x}) = \frac{\|\mathbf{x}\|_1^2}{\|\mathbf{x}\|_2^2}, \qquad (11)$$

where $\|\mathbf{x}\|_1$ and $\|\mathbf{x}\|_2$ are L_1 and L_2 norm of **x** respectively. It is clear that for integral **x,** the pseudo-weight reduces to a Hamming weight.

For each pseudocodeword from $M(P)$ or $V(P)$ we have associated real number, the pseudo-weight. If the pseudo-codewords of $M(P)$ or $V(P)$ are classified according to this parameter, what results is the so called minimal pseudo-weight spectrum and pseudo-weight spectrum respectively, usually presented in the form of histograms [8]. These are, from the code performance viewpoint, very useful graphical representations of an LP pseudocode.

As a running example we have analyzed a particularly adequate dual Hamming (7,3,4) code defined by its parity check matrix **H** that does not have any length-4 cycles:

$$\mathbf{H} = \begin{matrix} 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 \end{matrix} \qquad (12)$$

Starting with local convex hulls $\mathrm{CH}(C_i)$ it is necessary to take each $\mathbf{h}_i$ and translate it into the set of hyperplanes defining $\mathrm{CH}(C_i)$. To do this, it is necessary to define $N(\mathbf{h}_i)$ as the set of positions of non-zero elements of $\mathbf{h}_i$, i.e. $N(\mathbf{h}_i) = \{j | h_{ij} = 1\}$ (e.g. $N(\mathbf{h}_1) = \{1,2,5\}$). For every odd-element number subset $T \subseteq N(\mathbf{h}_i)$ we have a new hyperplane [5]:

$$\mathrm{CH}(C_i) = \{\mathbf{x} \mid \sum_{j\in T} x_j + \sum_{j\in N(\mathbf{h}_i)\setminus T} x_j \leqslant |N(\mathbf{h}_i)|-1, \forall \mathbf{h}_i\} \quad (13)$$

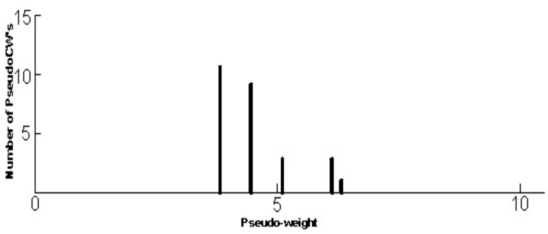
The fundamental polytope P is obtained by taking into account all hyperplanes defined by (13) for all SPC local codes C_i, as given in (4). Finding all the vertices of P produces the set of all pseudocodewords $V(P)$. This task is done for the dual Hamming (7,3,4) code and the results are listed in Table 1. In this table, the vertices of $V(P)$ are given with their coordinates rescaled, in order to have integer values. This does not affect the pseudo-weight as defined in (11). To be more precise, all points (pseudo-codewords) belonging to the same edge of the fundamental cone K are characterized with the same pseudo-weight value and are therefore considered equivalent.

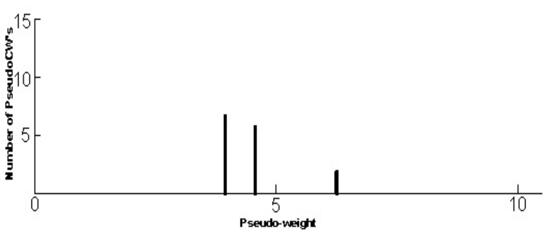
N	Pseudocodeword	Pseudo-weight	Set
1	0 0 0 0 0 0 0	-	C, $M(P)$
2	1 1 1 0 0 1 0	4	C, $M(P)$
3	0 1 1 1 0 0 1	4	C, $M(P)$
4	1 0 1 1 1 0 0	4	C, $M(P)$
5	0 1 0 1 1 1 0	4	C, $M(P)$
6	0 0 1 0 1 1 1	4	C, $M(P)$
7	1 0 0 1 0 1 1	4	C, $M(P)$
8	1 1 0 0 1 0 1	4	C, $M(P)$
9	0 1 2 1 1 1 0	4.5	$F(P)$, $M(P)$
10	0 1 1 1 2 0 1	4.5	$F(P)$, $M(P)$
11	0 1 1 1 1 0 0	4	$F(P)$, $M(P)$
12	1 0 1 1 1 0 2	4.5	$F(P)$, $M(P)$
13	0 1 0 1 1 1 2	4.5	$F(P)$, $M(P)$
14	0 1 1 1 2 0 3	4	$F(P)$, $M(P)$
15	1 0 1 1 1 2 0	4.5	$F(P)$, $M(P)$
16	0 1 1 1 0 2 1	4.5	$F(P)$, $M(P)$
17	0 1 2 1 1 3 0	4	$F(P)$, $M(P)$
18	2 1 1 1 0 0 1	4.5	$F(P)$, $M(P)$
19	2 1 0 1 1 1 0	4.5	$F(P)$, $M(P)$
20	3 2 1 1 1 0 0	4	$F(P)$, $M(P)$
21	1 2 1 1 1 0 0	4.5	$F(P)$, $M(P)$
22	0 2 1 2 1 3 3	5.14	$F(P)$
23	3 1 2 2 1 0 3	5.14	$F(P)$
24	3 1 1 2 2 3 0	5.14	$F(P)$
25	3 2 2 1 2 3 3	6.4	$F(P)$
26	2 1 1 1 2 2 1	6.25	$F(P)$
27	2 1 2 1 1 1 2	6.25	$F(P)$
28	1 2 1 1 1 2 2	6.25	$F(P)$

Table 1 *List of all pseudocodewords V(P) of dual Hamming code*

An interesting behavior of LP pseudocode noted in [5] is that adding redundant (linearly dependant with the existing set) parity-check tightens the fundamental polytope P, therefore increasing the performance of LP decoding. Since a larger number of parity-checks introduces more constraints into linear program, the cost for a better performance is increased complexity.

Similar effect exists in the case of BP decoding. Adding more parity checks increases the reliability information gathered about each codeword bit, but processing that they introduce adds to the complexity of decoding process.





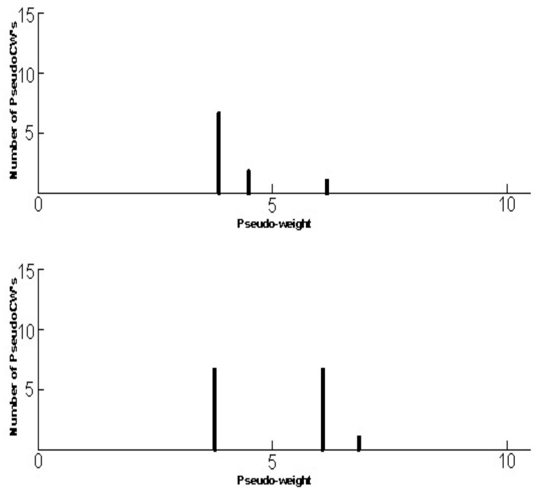


Figure 2. *LP decoding region*

To present this effect visually we have appended one-by-one the cyclic shifts of rows of **H** (12) that are redundant into a new **H** matrices, and observed the changes in the pseudo-weight spectrum of the pseudocodes. The results are given in the Figure 2, starting with the pseudo-weight spectrum of the pseudocode given in Table 1.

It is interesting to note that for dual Hamming (7,3,4) code there are no LP pseudocodewords with pseudoweight lower than 4, which is the minimum distance of this code. Additionally, by increasing the number of redundant parity-checks the effect of "pushing" the LP pseudocodewords toward larger pseudo-weights is evident. However, it is also evident that the introduced parity-checks haven't contributed to the code graph with length-4 cycles, which certainly contributed to this effect. In a full 7-by-7 parity-check matrix version, this code is a member of a wider class of codes based on a finite geometries, precisely LDPC PG(2,2) Type-I code. Good performance of this class of codes with BP algorithm is well known, but their good behavior w.r.t. minimal pseudocodeword spectrum has been recently noted in [8].

5. CONCLUSIONS

In this paper we have examined both the geometric interpretation and the error performance contribution of the LP pseudocodewords of a binary linear code *C*, defined by its parity-check matrix **H**. As in the case of ML decoding and minimal codewords, the dominant effect on error performance is due to a set of the minimal pseudocodewords that are sufficient to describe the all-zero codeword LP decoding region D_0^{LP}.

Tight relations established so far between LP decoding and BP decoding contributed to the increased interest in LP decoding. Unlike BP decoding convergence, which is still only partially understood, LP decoder maps its inputs to the set of LP pseudocodewords in a clear fashion.

Therefore, analyzing the LP pseudocode might be useful in the construction of codes to be used with BP decoding. By a careful design of the parity-check matrix **H** it should be possible to produce the fundamental polytope P with good pseudoweight spectrum, i.e. to push the pseudocodewords added by LP decoder to as larger values of pseudoweight as possible. This task might be complemented with the known efforts of optimizing LDPC code graph w.r.t. the existence of small stopping sets in order to achieve the optimal construction of LDPC codes.

ACKNOWLEDGMENT

Dejan Vukobratović would like to thank to Pascal O. Vontobel for introducing him into the interesting topic of LP decoding.

REFERENCES

[1] C. Di, D. Proietti, I.E. Telatar, T.J. Richardson, R.L. Urbanke: "Finite-length analysis of low-density parity check codes on the binary erasure channel", *IEEE Trans. On Inform. Theory*, vol. 48, no. 6, pp. 1570-1579, 2002.

[2] D.J.C. MacKay and M.S. Postol: "Weaknesses of Margulis and Ramanujan-Margulis low-density parity-check codes", preprint, available online at http://www.inference.phy.cam.ac.uk/mackay, 2003.

[3] N. Wiberg: "Codes and decoding on general graphs", Ph.D. dissertation, Linkoping University, Sweden, 1996.

[4] R. Koetter and P.O. Vontobel: "Graph covers and iterative decoding of finite-length codes", *3rd Inter. Conf. on Turbo Codes and Related Topics*, Brest, France, 2003.

[5] J. Feldman, "Decoding Error-Correcting Codes via Linear Programming", Ph.D. dissertation, M.I.T, USA, 2003.

[6] J. Feldman, D.R. Karger and M.J. Wainwright, "LP decoding", *41st Annual Allerton Conf., Monticello, Illinois*, USA, 2003.

[7] P.O. Vontobel and R. Koetter, "On the relationship between linear programming decoding and min-sum algorithm decoding", *Proc. of the ISITA Conf*, pp. 991-996, Parma, Italy, 2004.

[8] P.O. Vontobel, R. Smarandache, J. Teutsch, N. Kiyavash and D. Vukobratović, "On the Minimal Pseudo-Codewords of Codes from Finite Geometries", *submitted for publication at ISIT 2005.*

O LP REGIONIMA DEKODOVANJA

Dejan Vukobratović, Nemanja Delić, Vojin Šenk *Fakultet tehničkih nauka, Univerzitet u Novom Sadu*

Sažetak – *Nedavno objavljen metod dekodovanja linearnim programiranjem i veze utvrđene između ovog metoda i iterativnog BP dekodovanja su motivacija za proučavanje ovog metoda. Za razliku od BP dekodovanja kodova čiji grafovi sadrže petlje i čija konvergencija je još uvek nejasna, LP dekodovanje poseduje veoma jasnu (geometrijsku) interpretaciju. U ovom radu dati su komentari, potvrde i provere rezultata koji se tiču geometrijske interpretacije LP dekododovanja.*