



KATEDRA MATEMATIKY A  
TEORETICKEJ INFORMATIKY

TECHNICAL UNIVERSITY OF KOŠICE  
Faculty of Electrical Engineering and  
Informatics

Department of Mathematics and Theoretical  
Informatics

RNDr. Štefan Berezňý, PhD.

Address: Nĕmcovej 32  
042 00 Košice  
Slovak Republic

Telephone number: +421/55/602 2447  
E-mail: Stefan.Berezny@tuke.sk  
Web: <http://people.tuke.sk/stefan.berezny/>

---

**Review of the thesis**  
**LDPC Codes – New Methodologies**  
**by Ing. Jan Broulím**

Dissertation presented by Ing. Jan Broulím is divided into seven chapters and six annexes. In the first chapter called “Introduction” author introduces the reader to the problem he is dealing with. In the second chapter “Error correction coding”, the author gives a historical overview of “Error correction coding” and gives the basic terms and definitions. The third chapter is dedicated to the LDPC code (Background, Encoding, Tanner Graphs and Decoding). In the fourth chapter “Construction of LDPC Codes”, the author devotes to generation of LDPC matrices, optimization of LDPC codes, and to optimization tasks applications. In the fifth chapter “Mapping LDPC decoder on parallel architectures”, he deals with the parallelization of LDPC decoders and their implementation using OpenCL and CUDA frameworks. In the sixth chapter “Improving performance of LDPC decoders”, the author focuses on the most important parts of his work. How to improve the performance of decoders and also on the results obtained from the experiments. In the last, seventh, chapter “Conclusion” he summarizes the information presented and points out other possible directions. Finally, there are six annexes marked as A, B, C, D, E, and F.

The topic of the thesis is current and follows the current effort of analyzing existing procedures, to find more effective solutions and optimal procedures. The same applies in the area of codes, coding, and decoders. Analysis and optimization of LDPC codes are among the most current topics of research.

The procedures chosen by the author are appropriate and sufficient. They are described in this thesis, as well as in the articles, which are co-authored by the author of this thesis. The measurements and comparisons have been appropriately chosen so that it is easy to see the improvements after applying the procedures modified by author.

Results presented are original and sufficient to meet the objectives of the thesis and are in the frame of the topic as well. The results were published and their list is presented at the end of this thesis.

The author has provided sufficient insight into the subject, as evidenced by the literature cited and by the literature review on pages 117–124. The author’s publications are on pages 125–127. He lists 22 publications. Articles are chronologically organized from 2012 to 2017, with the last 4 being in preparation. Most articles deal with issues that correspond to the topic of thesis. Articles [1] to [18] are publications in Proceedings of papers. By the last 4 articles [19]–[22], which are in preparation, the author does not indicate where they should be published. It is a pity that no author’s publication has been published in a domestic or foreign scientific journal dealing with the theme of LDPC (Scopus, Web of science, etc.). Notwithstanding this claim, I consider the submitted publications by Jan Broulím to be sufficient (most of them are in indexed in Scopus and Web of Science).

The formal aspect of work is very good. The work is written in a clear and comprehensible manner. The text is complemented by tables and charts that facilitate understanding of the facts. However, some formulations contain inaccuracies, typos, and sometimes even vague formulations.

I have the following comments on the text of the thesis:

- p. 9: row 1 → the same designation should be used “State-of-the-art”.
- p. 23: in Definition 2.3.3 → it is not entirely clear what is the matrix  $\mathbf{P}$ .
- p. 28: in part 3.2 → dimensions of matrices  $\mathbf{G}$  and  $\mathbf{H}$  are missing. The value range of  $k$  is not determined and from which set can be selected parameter  $k$ .
- p. 28: in part 3.3 → missing Fig. 3-2.
- p. 29: in part 3.3, Eq. (3.3) → symbol  $\triangleq$  si not listed in the List of Symbols. What are  $c_i$  and  $v_j$ ? What denotes  $\mathbf{H}_{i,j}$ ? In Eqs. (3.5) and (3.6) it is not said, what denotes variable  $x$ .
- p. 32: Algorithm 1, row 2 → the set  $\mathcal{M}$  is not defined.
- p. 32: Algorithm 1, row 3 → the set  $\mathcal{N}$  is not defined.
- p. 32: Algorithm 1, row 14 → the set  $\mathcal{M}$  is not defined.
- p. 32: Algorithm 1, row 15 → the set  $\mathcal{N}$  is not defined.
- p. 35: in Eqs. (3.30), (3.31), and (3.32) is incorrect usage of indices  $i$  and  $j$ .
- p. 36: in Eq. (3.35) is not stated using which indices is calculated min. For calculation of Eq. (3.34) is used Eq. (3.36) and vice versa.

- p. 37: For calculation of Eq. (3.39) is used Eq. (3.40) and vice versa. In Eq. (3.42) is not clear, what are  $k$ ,  $t_i$ ,  $t_i^\alpha$ , and  $\sigma$ .
- p. 38: Eq. (3.44) → why matrix  $\mathbf{A}_0$  contains exactly four ones in a given row?
- pp. 42 – 45: Figs 4-1 – 4-4 → I'm missing more detailed description of figures. Which software was used for processing, visualization, and obtaining of the data?
- p. 49: Algorithm 6 → what does it mean  $u \in G$  or  $v_{\text{start}} \in G$ ?
- p. 50: Fig. 4-8 → missing more detailed description.
- p. 51: Algorithm 8 → are sets  $VN$  and  $CN$  provided as input? Algorithm doesn't produce output set  $U_{TG}$ .
- p. 52: row 2 → should read  $E_B/N_0$ .
- p. 55: Table 4.1 → should read  $E_B/N_0$ .
- p. 56: Fig. 4-12 (b) and (d) → should read  $E_B/N_0$ .
- p. 59: Fig. 4-15 (b) → should read  $E_B/N_0$ .
- p. 64: in the first point of bulleted list is wrong usage of indices  $m$  and  $n$ .
- p. 67: Algorithm 10 row 14 → where from comes value  $q_{\text{index}} = \text{value}$ , if  $\text{value}$  is uninitialized?
- p. 68: Algorithm 11 row 6 → syntactic error.
- pp. 73 – 74: Algorithm 12 and Algorithm 13 → denotations  $r_{i \rightarrow j}$  and  $q_{j \rightarrow i}$  are undefined.
- p. 81: Eq. (6.7) → symbol  $\oplus$  is not defined.
- p. 85: Eq. (6.20) → what symbolizes  $\Delta S$ . Only  $\Delta S(t)$  is defined.
- p. 87: Algorithm 17 row 2 → where  $\mathbf{p}$  and  $\mathbf{q}$  come from?
- p. 88: Algorithm 19 → in row 1 should be stated that  $m \in \{1, \dots, M\}$ . In the fifth row it is not clear what is  $l(m)$ .
- p. 90: row 2 → should be written “linearity”.
- p. 97: Fig. A-2 → missing.
- p. 100: row 2 → missing terminating parenthesis and the end of row 4 should read “ $v_n$ ”.

These comments have no significant impact on the contents of the presented work, and therefore I recommend the dissertation work of Jan Broulím for the defense.

A handwritten signature in blue ink, appearing to read 'Štefan BEREŽNÝ', is written above a horizontal line. To the right of the signature, there is a large, stylized blue checkmark or '1' symbol.

RNDr. Štefan BEREŽNÝ, PhD.

Košice, 17. 02. 2019





Report on the doctoral thesis of  
Jan Broulín  
*LDPC Codes – New Methodologies*

The work is divided into seven parts (including Introduction and Conclusions) and six appendices. Parts two and three are a kind of mild introduction into the theoretical context of LDPC codes and error correction in general. Each of the remaining three main parts describes one of author's contributions. Namely: construction of LDPC codes with suitable properties using genetic optimization; decoding of LDPC codes using parallelization; and several strategies to improve the decoding performance.

My background in mathematics and theoretical computer science would make me interested mainly in theoretical aspects of the candidate's work. However, the text does not provide much in this direction. While I am aware that the work resides mainly in the design, implementation and *empirical testing* of algorithms, I nevertheless start this report by listing several shortcomings that one could expect to be avoided even in this kind of work.

- The explanation of what an *optimal* decoding would mean is lacking. That is, we are not told what is the required property of the output  $\hat{c}$  with respect to the received corrupted message. It is only mentioned in passing that the minimum distance codeword is the goal, without explaining why, and whether we would accept an output which is not a codeword but is likely to have few error bits.
- The SP algorithm is described very superficially. It is not explained what “message passing” actually means. No attempt is made to explain the main idea why such an algorithm should have a chance to converge. For example, the reader would welcome some comments on why formula (3.19) yields a probability that the bit flip will improve the syndrome. This is not just a question of the theoretical introduction, such considerations would be particularly welcome as a motivation for improvements proposed in the research part. Or are they just a result of blind guesses?
- The use of linear algebra (the indispensable core of *linear* codes) feels sometimes shaky. A symptomatic example is the chapter 3.2 where the two identity matrices in  $\mathbf{G}$  and  $\mathbf{H}$  have different ranks in general, and the formula (3.1) obviously is not an equivalence (in the sense that many different generating matrices exist). A place where this may have some practical impact is for example formula (6.7) which is not correct in general, since the resulting matrix may have smaller rank. With respect to the mutational decoding as a whole, it is easy to observe that it actually means considering relations between several syndrome bits. This again should naturally provoke some theoretical comments on how this may be useful. Instead, we are just told that “it can be shown that the information from several decoders can be combined together for achieving better performance”. Shown experimentally, or even proved?
- The fitness value (4.4) is once more not explained in terms of its usefulness, and the formula itself is difficult to understand. Is  $\left\lfloor \frac{E_B}{N_0} \right\rfloor$  the number of iterations needed? And what is then  $\text{BER} \left( \left\lfloor \frac{E_B}{N_0} \right\rfloor \right)$ ?

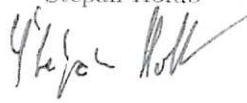
- The description of the parallel message passing algorithm in section 5.1 is completely incomprehensible due to inconsistent notation. As a little example, on the third line, p. 64,  $\mathbf{H}_{i,j}$  should be replaced with  $\mathbf{H}_{c_i,v_j}$ . One can only rely here on the example in Appendix A.
- Also the pseudocode cannot be always trusted. As far as I can see, a termination condition is missing on page 76. Symptomatic is the following statement on p. 65: “Both tables are particularly useful for understanding the principle and checking the correctness of the implementation.” Indeed, interested readers have to check the correctness themselves from examples.

Altogether, this work would not be satisfactory in a theoretical study program. However, I appreciate that it is submitted to the Department of *Applied* Electronics and Telecommunications, and that therefore its main value is in practical implementation and empirical testing. On the other hand, from this perspective, my ability to verify the presented results is severely limited, as well as my ability, assuming they are correct, to compare them with a similar work of others. The fact that results are by rule published in venues of rather local character does not help.

All this notwithstanding, I can state that the main three contributions are reasonable in terms of proposed concepts, and the obtained results, if correct and correctly compared to other similar work, seem convincing.

I therefore believe I can recommend this work to be accepted for a defense as a doctoral dissertation.

Prague, March 14, 2019

Štěpán Holub  


## Posudek doktorské práce

„LDPC codes – new methodologies“

autor: Ing. Jan Broulím

Předložená doktorská práce se věnuje obecně důležité problematice algoritmů detekčních a opravných kódů, a to konkrétně studiu a vlastnímu návrhu LDPC kódů. Výzkum byl prováděn na několika pracovištích, především na FEL ZČU ve skupině pod vedením doc. Georgieva a částečně i v SÚJV Dubna, kde byl pro potřeby dizertace využit výpočetní komplex HybriLIT. Ve spolupráci s několika partnery se v rámci skupiny řeší široká oblast úkolů, skupina má významné postavení ve vývoji řídicí elektroniky a příslušného programového vybavení (pro pokročilé detektory, aplikace do kosmu), což je spojeno často s přenosem velkých objemů dat ve složitých podmínkách (velké vzdálenosti, přítomnost vysoké radiace). Elektronika při přítomnosti různých typů radiace je náchylná k chybám (Single Event Effects). I v těchto podmínkách je potřeba zabezpečit bezchybný přenos dat, a proto jsou výsledky práce pro širokou vědeckou komunitu velmi potřebné. Potřebnost tematiky dokládá fakt, že v průběhu posledních 5 let je na WOS registrováno celkem 2 703 výsledků s danou tematikou.

Doktorská práce je přehledně rozdělena do 7 kapitol, přehledu použité literatury, povinných příloh a dalších 6 příloh. První kapitola je všeobecný úvod. Druhá kapitola pojednává o historii opravných kódů, základním definicím, vysvětluje pojmy jako paritní matice, matematické modely komunikačních kanálů apod. V třetí kapitole se autor soustřeďuje na problematiku LDPC kódů, historii, principy jejich konstrukce, kódování a dekódování. Kapitoly 4, 5 a 6 tvoří hlavní část doktorské práce. Kapitola 4 obsahuje metodologii pro návrh LDPC matic, genetický algoritmus pro vytváření LDPC matic na základě jejich grafické zobrazení (tzv. Tannerovy grafy). Použití genetických algoritmů při návrhu je největším přínosem práce. V této kapitole je také uvedena aplikace paralelních výpočtů pro urychlení optimalizace (včetně zhodnocení výsledků). Kapitola 5 popisuje paralelní přístup k vytvoření dekódovacích algoritmů LDPC kódů, implementaci na různé platformy umožňující paralelní výpočty (GPU, FPGA) a srovnání výsledků dekódovacích algoritmů pro různé parametry kódů na systémech na bázi programovacích jazyků OpenCL či CUDA. V kapitole 6 jsou shrnuty výsledky autora práce v oblasti vylepšení výkonu dekódovacích algoritmů LDPC kódů. Autor v ní používá dvě metody, Back Tracking (BT) a MLDP (mutational LDPC kódování). Byla provedena srovnání pomocí testů na standardních kódech (MSxMS-BT, SPxSP-BT, BFxBF-BT). Dále bylo zjištěno, že použití kombinace několika dekodérů poskytuje nejlepší výsledky (optimální BER-MLDP). Kapitola 7 shrnuje vykonané činnosti a dosažené výsledky (nové vytvářecí techniky pro optimalizační algoritmy, aplikace dekódování pro paralelní architekturu, dva algoritmy pro vylepšení LDPC dekodérů).

J. Broulím je uveden jako autor či spoluautor v celkem 18 publikacích, které jsou uvedeny v databázi WOS. Z těchto prací má 8 z nich přímou souvislost s tématem dizertační práce (přímo citované v přehledu literatury jsou tři z nich – publikace č. 10, 11 a 12). Kromě již opublikovaných prací je J. Broulím autorem či spoluautorem dalších 4 prací připravených k vydání.



Autor práce použil při řešení znalosti z několika oborů, matematiky, programování a výpočetní techniky. Metodika použitá při řešení je jak teoretická tak i praktická. Výsledky dizertační práce jsou zajímavé pro rozvoj celého oboru, detekci a opravu chyb při přenosu dat. Podle pořadí a počtu spoluautorů 8 původních publikací je zřejmé, že doktorand J. Broulím měl hlavní podíl na těchto výsledcích. Formální úprava dizertační práce je v pořádku, jazyková úroveň (práce je psána v AJ) je dobrá (jsou zde pouze menší překlepy v jednotlivých slovech).

Mám k doktorandovi dvě otázky: 1) Jaké jsou výhody a nevýhody CPU vs. GPU a FPGA v této oblasti? 2) Proč jste si vybral metody Bit-Flipping a Sum-Product a ne Hard-Decision nebo LogLikelihood?

Závěrem konstatuji, že odborná práce doktoranda Jana Broulíma při řešení dané problematiky je na vysoké úrovni, její zaměření je vysoce aktuální a že získal důležité prakticky využitelné výsledky. Na základě výše zmíněných faktů doporučuji, aby mu byl po obhajobě přiznán titul Ph.D.



Doc. Ing. Ivan Štekl, CSc.

**Západočeská univerzita v Plzni**

Doručeno: 13.12.2018

**ZCU 030902/2018**

listy: 3

přílohy:

druh:



zcupes1136966