Complexity of the McEliece Cryptosystem based on GDBF Decoder for QC-LDPC Codes

Omran Al Rasheed[1], Dejan Drajić[2], Predrag Ivaniš[3] and Goran Đorđević[4]

Abstract – **In this paper, we propose the application of Gradient Descent Bit Flipping (GDBF) algorithm in decryption stage of McEliece cryptosystem based on Quasi-Cyclic Low-Density Parity Check (QC-LDPC) codes. Comparing of complexity between GDBF and Gallager B decoders is given by binary operations for each decrypted bit.**

Keywords – **Low Density Parity Check, Quasi-Cyclic, McEliece cryptosystem.**

I. Introduction

The main purpose of the cryptosystem is to guarantee confidentiality to the message. McEliece cryptosystem is important public key code, derived by Robert J. McEliece in 1978 [1]. It is based on binary Goppa codes [2], it is still unbroken, and worth mentioning that no polynomial-time algorithm is discovered to recover the secret message in the McEliece cryptosystem even using quantum computers. This property is unique from other worldwide implemented solutions for public key cryptosystem and digital signatures as RSA [3], DSA, ECDSA [4].

The McEliece cryptosystem demands the sender to encode the information into a codeword through the receiver's public key, then to introduce a certain number of intentional errors in the encoded word. This allows the receiver that is able to correct those errors through his secret private key, to recover the secret information. The McEliece shows some drawbacks, among of them, the large of the key that prevented its widespread adoption and low transmission rate. There are many ways to overcome these drawbacks of the McEliece cryptosystem, and the most solutions are based on replacing of Goppa codes with other families of codes. However, it is difficult to replace Goppa codes with other codes in the absence incurring into serious security flaws. The families of code must achieve to some conditions to ensure the security, must to be large enough to keep away from the enumeration, a generator matrix of a permutation equivalent code must be obscure that no permit to appear any construction for the secret code and this code has efficient algorithm to correct the codeword form the errors, this means that receiver can able to read the transmitted ciphertext over the unsecure channel.

On the other hand, it is well known that Low Density Parity Check (LDPC) codes are powerful error correction codes that achieve performance near the Shannon limit [5]. The effective iterative decoding algorithms for LDPC codes make these codes have high attraction comparing with other class of codes. The first supposition to use LDPC codes in the public-key cryptosystem was in [6]. In that paper, it has been clarified that employment of LDPC codes instead of Goppa codes does not allow decreasing the key length. However, it have been presented that QC-LDPC code based cryptosystem is immune to any known attack [7]. The new features are appeared with QC-LDPC code, less key size and higher code rates and keep or respect the original version.

In this paper, we consider the required computational effort for gradient descent bit flipping (GDBF) decoder [8] in BSC with McEliece cryptosystem. Complexity of the decoder is measured by number of needed binary operations for each iteration of GDBF algorithm. We compare the cost of GDBF decoder with Gallager B decoder [5]. In Section II McEliece cryptosystem based on QC-LDPC is described. The GDBF decoder over BSC is described in Section III. In Section IV a calculation cost of GDBF decoder for each iteration. Finally, some concluding remarks and future research directions are given in Section V.

II. Description McEliece Cryptosystem based on QC-LDPC Codes

In the McEliece cryptosystem, the receiver makes the private key that is created by the sparse parity-check matrix **H**, that is selected randomly and has the form

$$\mathbf{H} = [\mathbf{H}_0 \,|\, \mathbf{H}_1 \,|\, ... \,|\, \mathbf{H}_{n_0-1}], \tag{1}$$

where $\mathbf{H}_i$ is a circulant block, and each row (column) has weight d_v. It is important to avoid short cycles (4-length) in matrix **H**. Selecting randomly n_0 vectors h_i (disjoint set of different modulo p), give us huge families of codes with codes identical parameters [11] as a condition to replace Goppa. Systematic generator matrix for the code is **G**=[**I** |**P**], where **I** is a kxk identity matrix and **P** is given in form

$$\mathbf{P} = \begin{bmatrix} (\mathbf{H}_{n_0-1}^{-1}\mathbf{H}_0)^{\mathrm{T}} \\ (\mathbf{H}_{n_0-1}^{-1}\mathbf{H}_1)^{\mathrm{T}} \\ \vdots \\ (\mathbf{H}_{n_0-1}^{-1}\mathbf{H}_{n_0-2})^{\mathrm{T}} \end{bmatrix}, \tag{2}$$

[1] Omran Al Rasheed is with the School of Electrical Engineering, University of Belgrade, Bul. kralja Aleksandra 73, 11120 Belgrade, Serbia (e-mail: omrano84@hotmail.com).

[2] Dejan D. Drajić is with IRITEL A.D., Batajnicki drum 23, Belgrade, Serbia. Also, he is with the School of Electrical Engineering, University of Belgrade, Bul. kralja Aleksandra 73, 11120 Belgrade, Serbia (e-mail: dejan.d.drajic@gmail.com).

[3] Predrag N. Ivaniš is with the School of Electrical Engineering, University of Belgrade, Bul. kralja Aleksandra 73, 11120 Belgrade, Serbia (e-mail: predrag.ivanis@etf.rs).

[4] Goran T. Đorđević is with the Faculty of Electronic Engineering at University of Niš, Aleksandra Medvedeva 14, Niš 18000, Serbia, E-mail: goran.djordjevic@elfak.ni.ac.rs.

where $\mathbf{H}_{n_0-1}$ is non-singular matrix and operator ^{T} denotes transposition.

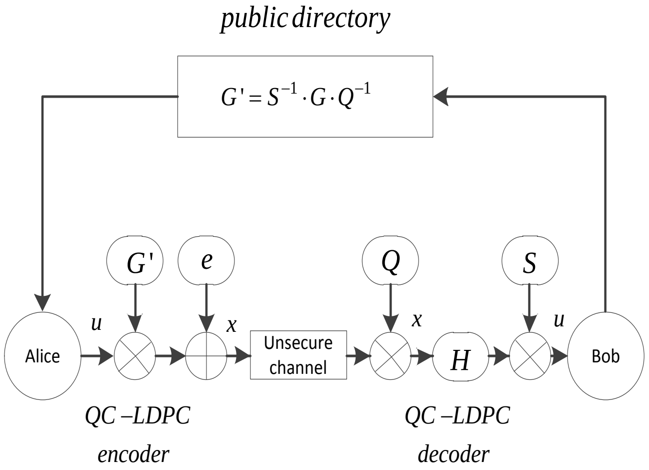

Fig. 1. McEliece cryptosystem with QC-LDPC codes

Procedure of McEliece cryptosystem with QC-LDPC codes is shown in Fig. 1. The receiver (Bob) chooses other two matrices such that increase the code secrecy. The first matrix is *k*x*k* scrambling matrix **S** and the second is a *n*x*n* non-singular matrix **Q**. Finally the public key is given by

$$\mathbf{G}' = \mathbf{S}^{-1} \times \mathbf{G} \times \mathbf{Q}^{-1}. \qquad (3)$$

It should be noticed that low density of matrix $\mathbf{G}'$ helps to avoid attacks to the dual code. Matrix **H** can be recognized by one row of each circulant block, so **H** is mapped into a new matrix such $\mathbf{H}' = \mathbf{H}\mathbf{Q}^T$. The sender obtains the public key from the public directory, and anyone can get this key. Alice encrypts the cleartext ***u*** to obtain a ciphertext as $\boldsymbol{x} = \boldsymbol{u}\mathbf{G}' + \boldsymbol{e}$. and **e** is a random vector of t' intentional errors. The first requirement that Bob does it when the ciphertext is received, is reverse transformation as:

$$\boldsymbol{x}' = \boldsymbol{x}\mathbf{Q} = \boldsymbol{u}\mathbf{S}^{-1}\mathbf{G} + \boldsymbol{e}\mathbf{Q}, \qquad (4)$$

where $\boldsymbol{x}'$ is a codeword and is affected by the vector error $\boldsymbol{e}\mathbf{Q}$, with weight less or equal to $t = t'm$, and *m* is column and row of matrix **Q**. Finally, Bob can obtain the cleartext after decoding process using any LDPC decoder and then multiplication by matrix **S**.

The capability of the receiver to correct all errors depends on the type of the LDPC decoder and in the other hand on the structure of the private key. The receiver has a lot of chances to select the better structure for private key to avoid the short cycles in the Tanner graph and to avoid any attack. LDPC decoder has an important role to eliminate the intentional error vector. For the decoding algorithm, there are many algorithms based on soft decision and message passing between variable and check nodes, such as Sum Product Algorithm (SPA) [12], SPA algorithm is very complexity and has the best performance. The running time required by each algorithm depends highly on its computational complexity and the processing platform used.

III. GDBF Decoding Algorithm for BSC

A. LDPC Codes

LDPC codes are linear block codes that are designed by appropriate construction of the corresponding parity check matrix **H**, which is characterized by being sparse. LDPC code is denoted as ***C***(*n*,*k*) and matrix **H** with dimension *m*x*n*. Every code vector ***c*** satisfies the condition **H*c***=0, where operations are performed over GF(2). Irregular LDPC code is one with a sparse check matrix **H** that has a variable number of ‘1’s per row or per column. The information can be represented by bipartite or Tanner graph. The bipartite graph describes the relationship between two types of nodes, the symbol nodes v_j with column weight d_v and the parity check nodes c_i with row weight d_c. QC-LDPC codes [9] have some advantages compared with other constructions that are constructed by circulant permutation matrices such easier to implement with good performances [10]. Every row of the parity check matrix **H** corresponding to a parity check equation, and thus to a parity check node. Each bit of the code vector corresponding to a symbol node. Let v_j a variable node in the Tanner graph and *M*(*j*) denotes as a set of parity check nodes connected with v_j, and *N*(*i*) is a set of variable nodes connected with parity check node c_i. Let ***r*** is a received codeword and ***e*** is an entry error sequence determined by the probability of the BSC.

B. GDBF Decoder over BSC

GDBF is a class of BF algorithms based on the gradient descent algorithm. This algorithm is designed for transmission is done over additive white Gaussian noise (AWGN) channel and for each iteration an inversion function is calculated as

$$\Delta_j^{(t)}(x,y) = x_j^{(t)} y_j + \sum_{i \in M(j)} \prod_{l \in N(i)} x_l^{(t)}, \quad j = 1,2,....,n, \quad t \geq 0 \quad (5)$$

for AWGN x_j can be considered by polar representation (-1,+1) of the *j*th bit of the estimated codeword in the *t*th iteration, and y_j denotes the corresponding symbol of the received signal. For all variable nodes, inversion function is calculated for every particular iteration and only symbols have the minimum value of inversion function are flipped to obtain the new value for the *j*th bit for the next iteration $x_j^{(t+1)}$. Here we implement GDBF for BSC, where x_j takes binary value (0,1) and inversion function also takes integer value. We can consider an inversion function as a special case for regular LDPC code as:

$$\Delta_{M,j}^{(t)}(v,r) = v_j^{(t)} \oplus r_j + \sum_{i \in M(j)} \overset{\oplus}{\sum_{l \in N(i)}} v_l^{(t)}, \qquad (6)$$

where summation over modulo 2 is denoted by using $\oplus$ in combination with summation operator. In the above expression, the value of the inversion function is confined to the set of integer values $\Delta_{M,j}^{(t)}(v,r) \in \{0,1,...,d_v+1\}$, where large value of (6) indicates that the corresponding bit should be probably flipped.

IV. Complexity of the GDBF Decoder over BSC

Decryption stage for the McEliece cryptosystem includes LDPC decoder to correct the codeword from the errors. We suppose that receiver performs GDBF decoder. We can apprise the number of binary operation over the Tanner graph for one iteration. We consider some parameters for construction QC-LDPC codes, $n=n_0p$ length of these codes, dimension $k=k_0p$ and redundancy $r=p$, where n_0 is a small integer and $k_0=n_0$-1 with large value for p. In principle, every check node receives d_c binary values from the register which contains all estimated bits for each iteration. Check node XORs all the messages and sends the result to all neighbour variable nodes, so at the check node (d_c-1) binary sums and the total number of operations at the check nodes is $r(d_c$-1). This process is illustrated in Fig. 2. for parity check matrix

$$\mathbf{H} = \begin{bmatrix} 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \end{bmatrix}. \qquad (7)$$

Variable node receives messages from XORs, for every variable bit - one XOR between the estimated bit in tth iteration and original received bit, and d_v parity checks. The additional operations are needed for integer values. We can use full adder and for n variable nodes we have about nd_v full adders.

Majority logic (ML) gates is designed to calculate inverse function for every bit, determine the maximum value of Delta, all bits with maximum value and flipping them, so 3n binary operations are needed for these tasks. In conclusion, the cost of one iteration of GDBF can be estimated as

$$C^{(1)}_{GDBF} = r(d_c - 1) + 4n + nd_v = n(4 + 2d_v) - r\,. \qquad (8)$$

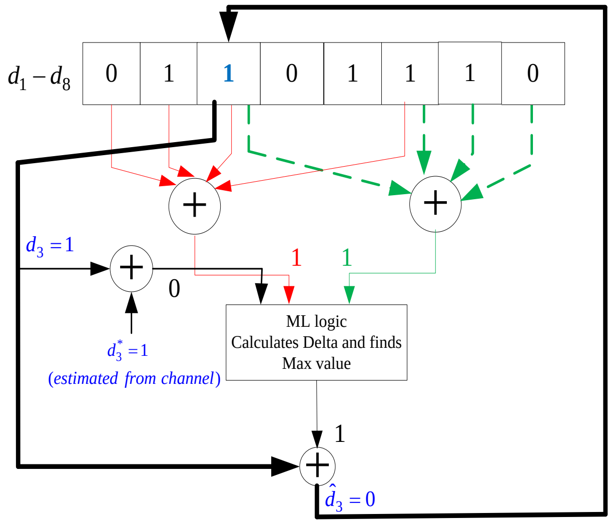

Fig. 2. Structure of GDBF decoder for BSC

To calculate the cost of the decryption process, we also estimate the cost for $x\mathbf{Q}$ and $\boldsymbol{u}'\mathbf{S}$. The traditional multiplication is used for $x\mathbf{Q}$ due to that $\mathbf{Q}$ is a sparse matrix and $n \times m$ binary operations is needed, while $\mathbf{S}$ is a density matrix. Therefore, we must use efficient algorithm as Winograd convolution to reduce the computations [13]. The choice both d_v and m adjusts the computational complexity of the decryption stage. When d_v is increased then cost of $x\mathbf{Q}$ is decreased while cost of the decoding process is increased [14]. In construct, increasing m (matrix $\boldsymbol{Q}$ must be sparse in order to allow correcting all intentional errors) will increase the cost of $x\mathbf{Q}$ while decrease the cost for decoding process. It is important to choose optimal between all defined parameters for trade-off between security and complexity.

V. Numerical Results

In this section the performance for GDBF will be compared with the typical hard-decision algorithms - bit flipping (BF) and Gallager-B (as simplest possible massage passing algorithm). Then, the complexity analysis of the both considered algorithms will be given.

Fig. 3 shows FER performance for simple BF, GDBF and Gallager B decoders for QC732, R=0.7527 with maximum number of iterations $t_{\max}$=100. It can be observed that the gap between BF and GDBF is very large, and that GDBF decoder has better performance than Gallager B decoder especially for higher values of crossover probability. This characteristic increases immunity of cryptosystem to avoid decoding attacks due to the low weight (t') of the error vector at the sender and it is often more dangerous than structural. Computing both $x\mathbf{Q}$ and $\boldsymbol{u}'\mathbf{S}$ and by using (8), we can estimate the cost (in term of the binary operation) of the decryption process for McEliece cryptosystem. If we consider that m=7, the average number of iteration is t_{av}=10 for the GBDF decoder. The values in Table I represent the binary operations needed for each decrypted bit using GDBF decoding and Gallager B decoders [15].

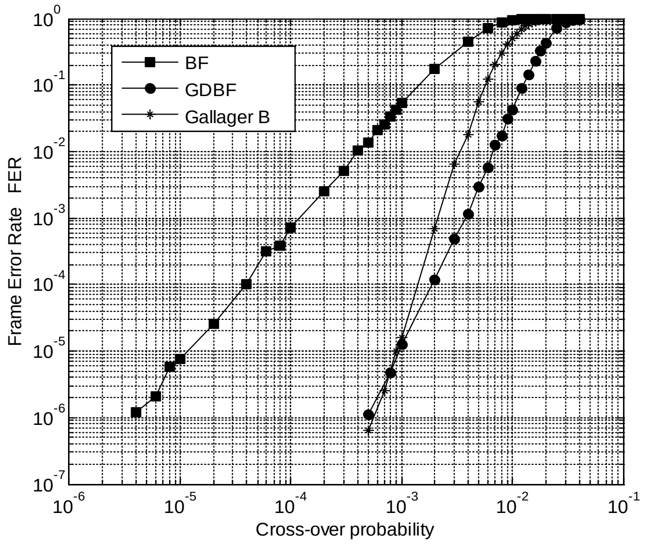

Fig. 3. FER performances for different decoders over BSC for QC732 with code rate R=0.7527

Table I
Compassion between Gallager B and GDBF Decoders used in the Decrypted Stage

p [bits]	d_v	13	15	13	15
	n_0	3	3	4	4
4096	GDBF	961	1011	1131	1185
	Gallager B	1476	1626	1598	1731
5120	GDBF	1019	1079	1227	1281
	Gallager B	1544	1694	1694	1828
6144	GDBF	1086	1146	1323	1377
	Gallager B	1611	1761	1790	1924
7168	GDBF	1143	1203	1410	1463
	Gallager B	1668	1818	1877	2010
8192	GDBF	1201	1261	1496	1550
	Gallager B	1726	1876	1963	2097
9216	GDBF	1259	1319	1583	1636
	Gallager B	1784	1934	2050	2183
14336	GDBF	1489	1549	1914	1964
	Gallager B	2014	2164	2381	2515
15360	GDBF	1605	1665	2102	2155
	Gallager B	2130	2280	2569	2702
16384	GDBF	1576	1636	2044	2098
	Gallager B	2101	2251	2511	2644

It can be noticed that GDBF decoder has better performance than Gallager B decoder. Furthermore, GDBF decoder has less computational complexity. Increasing length of the clear text and variable degree will increase the cost.

VI. Conclusion

In this paper we have described McEliece cryptosystem based on QC-LPDC, with GDBF decoder. We have estimated the cost of the binary operation for each decrypted bit. Although GDBF is hard decision decoder, it has high capability to correct errors for BSC. Also, the complexity of GDBF algorithm is reduced when compared with Gallager B decoder.

Acknowledgement

This work was supported by the technology development project of the Serbian Ministry of education, science and technology development, under Grant Agreement number TR32007 - "Multiservice optical transport platform OTN10/40/100 Gbps with DWDM/ROADM and CARRIER ETHERNET".

References

[1] R. J. McEliece, “A Public-Key Cryptosystem based on Algebraic Coding Theory.” DSN Progress Report, pp.114–116, 1978.

[2] V. D. Goppa, “A New Class of Linear Error-Correcting Codes,” Probl.Peredachi Inf., vol. 6, no. 3, pp. 24–30, Sep. 1970.

[3] R. L. Rivest, A. Shamir, and L. Adleman, “A Method for Obtaining Digital Signatures and Public-Key Cryptosystems,” Commun. ACM, vol. 21, no. 2, pp. 120–126, Feb. 1978.

[4] “Digital Signature Algorithm,” 1993. [Online]. Available: http://csrc.nist.gov/publications/fips/fips186-3/fips_186-3.pdf.

[5] R. G. Gallager, *Low Density Parity Check Codes*, Cambridge, MA: M.I.T Press, 1963.

[6] C. Monico, J. Rosenthal, and A. Shokrollahi, “Using Low Density Parity Check Codes in the McEliece Cryptosystem,” Proc. IEEE ISIT 2000, pp. 215, Sorrento, Italy, Jun. 2000.

[7] M. Baldi, M. Bodrato, and F. Chiaraluce, “A New Analysis of the McEliece Cryptosystem based on QC-LDPC Codes,” Security and Cryptography for Networks, ser. Lecture Notes in Computer Science. Springer Verlag, vol. 5229, pp. 246–262, 2008.

[8] T. Wadayama, K. Nakamura, M. Yagita, Y. Funahashi, S. Usami, and I. Takumi, “Gradient Descent Bit Flipping Algorithms for Decoding LDPC Codes,” IEEE Trans. Communications, vol. 58, no. 6, pp. 1610–1614, Jun. 2010.

[9] M. Tanner, D. Sridhara, A. Sridharan, T. E. Fuja, and D. J. Costello, “LDPC Block and Convolutional Codes based on Circulant Matrices,” IEEE Transactions on Information Theory, vol. 50, no. 12, pp. 2966–2984, Dec. 2004.

[10] L. Chen, J. Xu, I. Djurdjevic, and S. Lin, “Near-Shannon-Limit Quasi-Cyclic Low-Density Parity-Check Codes, IEEE Trans. Commun., vol. 52, no. 7, pp. 1038–1042, Jul. 2004.

[11] M. Baldi and F. Chiaraluce, “Cryptanalysis of a New Instance of McEliece Cryptosystem based on QC-LDPC Codes”, *Proc.* IEEE International Symposium on Information Theory (ISIT 2007), pp. 2591–2595, Nice, France, Jun. 2007.

[12] D. J. C. MacKay and R. M. Neal, “Near Shannon Limit Performance of Low Density Parity Check Codes”, Electronics Letters, vol. 32, no. 18, pp. 1645-1646, Aug. 1996.

[13] S. Winograd, “Arithmetic Complexity of Computations”, ser. CBMS-NSF Regional Conference Series in Mathematics. SIAM, vol. 33, 1980.

[14] T. J. Richardson and R. L. Urbanke, “The Capacity of Low-Density Parity-Check Codes nder Message-Passing Decoding”, IEEE Trans. Inform. Theory, vol. 47, no. 2, pp. 599–618, Feb. 2001.

[15] Baldi, Marco, Marco Bianchi, and Franco Chiaraluce, “Security and Complexity of the McEliece Cryptosystem based on Quasi Cyclic Low-Density Parity-Check Codes”, IET Information Security 7, no.3, pp. 212-220, 2013.