Newly developed quasi-cyclic low density parity check codes for hard disks record data

ABID YAHYA*, FARID GHANI, R. BADLISHAH AHMED
School of Computer and Communication Engineering, University Malaysia Perlis (UniMAP), 02000 Kuala Perlis, Perlis, Malaysia

This paper presents a new technique for constructing the Quasi-Cyclic low density parity check (QC-LDPC) codes based on row division method. The new codes offer more flexibility in term of girth, code rates and codeword length. In this method of code construction, the rows are used to form as the distance graph. Then they are transformed to a parity-check matrix in order to acquire the desired girth. Simulation results show that the proposed QC-LDPC codes achieve a 0.1dB coding gain over randomly constructed codes and perform 1.3 dB from the Shannon-limit at a BER of 10^{-6} with a code rate of 0.89 for block length of 1332.

(Received July 5, 2011; accepted July 25, 2011)

Keywords: Bit error rate, Communication channels, Encoding, QC-LDPC

1. Introduction

Hard Disks (HDDs) record data by magnetizing ferromagnetic material in the way of magnetization represent patterns of binary data bits. The recorded data are read from the disk by detecting the transitions in magnetization and decoding the original data. Different encoding systems, such as Modified Frequency Modulation, group code recording, run-length limited encoding, and others are employed.

The opponent of hard drives is thermostat. The devices store data in bits, which are microscopic spots on a hard drive platter. The bits themselves are made up of about 50 to 100 cobalt-platinum grains. When the grains get magnetized in a particular direction, the bit represents either a "1" or "0". To increase the areal density, which is the amount of data a single platter inside a hard drive can hold, technologists have squinched the size of bits and grains over the years. This has assisted computer manufacturers to enhance the competence of hard drives from a few megabytes to more than 100 gigabytes.

Sequent years of shrinkage, nevertheless, have led to magnetic grains that measure about 8 nanometers long. Reducing the grains more in size could cause them to flip at room temperature and so damage the data--an aspect of the "super paramagnetic effect," first recognized in the mid-1990s by Stan Charap of Carnegie Mellon University and bringing down on the number of grains inside each bit, absent additional changes, would increase noise and decrease reliability.

Barium titanate, BaTiO3 is the first ferroelectric ceramics, which is a good candidate for a variety of applications, such as piezoelectric actuators, multiplayer ceramic capacitors and positive temperature coefficient resistors, due to its excellent dielectric, ferroelectric and piezoelectric properties [1].

Drive manufacturers have bought time with perpendicular drives, which stack the bits in vertical position. But the "no more shrinkage" problem has yet solved by that solution. The heat-assisted site requires to modify the grains. Unlike cobalt-platinum grains, iron-platinum grains will not flip at room temperature. To record or erase data, a laser integrated into the drive would heat a specific bit. The data would get recorded or erased, and the bit would quickly cool.

Recent drives also construct wide use of Error Correcting Codes (ECCs), especially Reed–Solomon error correction codes. These methods store extra bits for each block of data that are found out by mathematical formulas. The additional bits permit many errors to be determined. Though these extra bits adopt space on the hard drive, they permit higher recording densities to be employed, follow-on in much larger storage capacity for user data. In this proposal the newest drives, low-density parity-check codes (LDPC) are supplanting Reed-Solomon. Since LDPC codes facilitate performance close to the Shannon Limit and therefore permit for the highest storage density obtainable.

Channel coding plays key role in providing a reliable communication method that can overcome signal degradation in practical channels. The breakthrough of convolutional codes [2] led off a new field of study into non-algebraic codes based on linear transformations using generator and parity-check matrices. Convolutional codes are encoded using a finite-state process, which generates them a linear order encoding scheme. Subsequently, convolutional codes led to the discovery of a class of codes called Turbo codes [3], which are the class of concatenated convolutional codes and randomize the order of some of the bits by using an interleaver.

Turbo codes are the first known capacity approaching error correction codes, which provide a powerful error correction capability when decoded by an iterative decoding algorithm [3]. The rediscovery of low density parity check (LDPC) code, which was originally proposed by Gallager [4] and was later generalized as MacKay-Neal code [5] put back Turbo coding as the forward error correction (FEC) technique. LDPC codes were neglected for a long time as their computational complexity for the hardware technology was very high. LDPC codes have acquired considerable attention due to its near-capacity error execution and powerful channel coding technique with an adequately long codeword length [6]. LDPC codes have several advantages over Turbo codes. In the decoding part, Turbo code faces difficulty to apply parallelism due to the sequential nature of the decoding algorithm. In case of LDPC, decoding can be accomplished with a high degree of parallelism to attain a very high decoding throughput. LDPC codes do not need a long interleaver, which usually causes a large delay in Turbo codes. LDPC codes can be constructed directly for a desired code rate while Turbo codes, which are based on convolutional codes, require other methods such as puncturing to acquire the desired rate.

The codes are classified into two major categories, explicitly, block codes and convolutional codes. Hamming codes, Bose- Chaudhuri-Hocquenghem (BCH) codes; Reed-Solomon codes (RS) [7-8] and newly rediscovered LDPC codes are the example of block codes. Block codes like Hamming, BCH and RS codes have structures but with limited code length. A bounded-distance decoding algorithm is usually employed in decoding block codes except LDPC codes. In general, it is hard to use soft decision decoding for block codes.

Advances in error correcting codes have revealed that, using the message passing decoding algorithm, irregular LDPC codes can accomplish consistent communication at SNR very close to the Shannon limit as Compared to Turbo codes [9]. The numerical analysis method for calculating the threshold of the LDPC codes is examined by Hou *et al,* [10] over AWGN channel with uncorrelated flat Rayleigh fading channel. Additionally, using the nonlinear optimization technique of differential evolution, the degree distribution pairs are optimized for the uncorrelated Rayleigh fading channel and it has been observed that their threshold values are very close to the capacity of this channel for moderate block size with excellent performance. The two adaptive coded modulation schemes employing LDPC codes for Rayleigh fading channels are proposed by Zhang *et al.,* [11]. It is shown from their work that the proposed schemes have made good use of the time-varying nature of Rayleigh fading channel. It is also observed that the proposed schemes perform better by employing LDPC with large code length.

The performance of irregular LDPC codes is investigated in [12] with three BP based decoding algorithms, specifically the Uniformly Most Powerful (UMP) BP-based algorithm, the Normalized BP-based algorithm, and the Offset BP-based algorithm on a fast Rayleigh fading channel by employing density evolution (DE). It is observed from the study that the performance and decoding complexity of irregular LDPC codes with the offset BP-based algorithm can be very close to that with the BP algorithm on the fast Rayleigh fading channel. After successful evolution of irregular LDPC codes, Ohhashi and Ohtsuki [12] then analyze the performance of regular LDPC codes with the normalized BP-based algorithms on the fast Rayleigh fading channel. Formulas for short and long regular LDPC codes are derived based on the probability density function (PDF) of the initial likelihood information and DE for the normalized BP-based algorithm on the fast Rayleigh fading channel. Performance of the long regular LDPC codes with the Normalized BP-based algorithm in the proposed method outperforms the BP and the UMP BP-based algorithms on fast Rayleigh fading channel. In this paper new QC-LDPC codes have been developed and then implement the newly designed codes on FPGA platform. Simulation results demonstrate that the proposed QC-LDPC codes achieve a 0.1dB coding gain over randomly constructed codes and perform 1.3 dB from the Shannon-limit at a BER of 10^{-6} with a code rate of 0.89 for block length of 1332.

2. Quasi-cyclic LDPC codes

Quasi-Cyclic LDPC codes form a large class of codes with nice encoding and decoding, which have been comprehensively premeditated in the sense that their hardware is both cheap and easy to implement. Since both encoding and decoding require less memory, which have many gains for hardware and software implementations. This memory advantage is catered by being able to illustrate the matrices employing a series of short polynomials.

A code is pronounced to be quasi-cyclic (QC) if a cyclic shift of any codeword by p positions is still a codeword. Thus a cyclic code is a QC code with $p = 1$. The block length n of a QC code is a multiple of p, as a result

$$n = m \times p \tag{1}$$

Circulants, or cyclic matrices, are indispensable components in the generator matrix for a QC code.

A circulant matrix is defined as a square matrix, such that each row is incurred by a cyclic shift of the preceding. The parity-check matrix of a QC code is decomposed into $m \times m$ blocks of circulant matrices, with submatrix dimensions $vm \times \mu m$, where ν and $\mu$are the circulants number in row and column, respectively.

Let C be a $m \times m$ matrix. It can be said that C is circulant if its rows are incurred by successive shifts. This means that a QC code can be distinguished by a generator matrix of the form [13-16]:

$$C = [C_0, C_1, C_2, C_3, \dots, C_{m-1}] \tag{2}$$

where $C_i,\ \ i = 0,1,2, \dots, m-1,$are circulants of order $m \times m$ of the form:

$$C = \begin{bmatrix} C_0 & C_1 & C_2 & \cdots & C_{m-1} \\ C_{m-1} & C_0 & C_1 & \cdots & C_{m-2} \\ C_{m-2} & C_{m-1} & C_0 & \cdots & C_{m-3} \\ \vdots & \vdots & \vdots & & \vdots \\ C_1 & C_2 & C_3 & \cdots & C_0 \end{bmatrix} \quad (3)$$

Consider weight-1 and weight-2 circulants, in polynomial form, by referring the positions of the 1's in the first row of the circulant. Circulant with weight-2 is denoting in the polynomial form:

$$P(x) = x^{\alpha} + x^{\beta} \quad (4)$$

Great care has to be taken while dealing with the values of α and β, since efficient decoding and removal of short cycles depend on these values.

2.1 New algorithm for quasi-cyclic LDPC codes

The proposed algorithm for the construction of the new QC-LDPC (N, j, k) code is as described in the following steps:

1. The codeword can be encoded as follows;

$$mod2(\mathbf{H} \times C^{T}) = 0 \quad (5)$$

where H represents parity check matrix and *C* denotes codeword.

2. For an efficient encoding the codeword's rows are split into sub rows with respect to group size.

$$C = (b, |p_1, p_2, \ldots, p_k) \quad (6)$$

where b and p represent the information and parity check bits respectively.

3. The constraint of group's number on row weight size persists the row-column (RC) connections to generate variety of codes.

$$\Psi = \sum_{i=1}^{k} \chi \quad (7)$$

where Ψ represents group size and χ stands for number of row.

4. In each sub row the number of 1-component is selected in order to maintain the concentrated degree of distribution which results in random selection. Otherwise non-concentrated degree of distribution will appear.

$$\mathbf{H}_{QC} = \begin{bmatrix} H_{cir} & Z & H_{cir} & Z & H_{cir} & H_{cir} & Z \\ H_{cir} & H_{cir} & Z & H_{cir} & H_{cir} & Z & Z \;\vdots \\ Z & H_{cir} & Z & H_{cir} & H_{cir} & Z & H_{cir} \\ \cdots\;\cdots & & & & & & \end{bmatrix} \quad (8)$$

$$H_{cir} = \begin{bmatrix} I_1 & \cdots & I_{m^{k-1}} \\ I_n & \cdots & \vdots \\ \vdots & \ddots & \vdots \\ I_{n^{j-1}} & \cdots & I_{m^{k-1} n^{j-1}} \end{bmatrix} \quad (9)$$

A $p \times p$ $\mathbf{H}_{QC}$ array is obtained from the aforesaid equations, where $H_{cir}$is the permutation matrix with the location vector of the field elements, cyclically shifting codeword by one position. Therefore, each row of $H_{cir}$is obtained from shifting the rows of the identity matrix to the left. Hence $H_{cir}$is $p \times p$ circulant permutation matrix and $\mathbf{H}_{QC}$ is $p \times p$ array of $p \times p$ circulant permutation matrix.

Let $I_1, \ldots, I_{m^{k-1}}$ denote the rows of the parity check matrix **H**. First in this work consider splitting each row of **H** into the same number of sub rows. All the new sub rows have the same length as the original row. The weight (or "1s") of the original row is distributed among the sub rows. A regular row weight distribution can be done as follows. Let q be a positive integer, split each row $I_i$of **H** into q sub rows $I_{i,1}, I_{i,2}, \ldots, I_{i,q^{k-1}}$. The distribution of k "ones" of I_i into $I_{i,1}, I_{i,2}, \ldots, I_{i,q^{k-1}}$ is carried out in a rotating manner. In the first rotation, the first "1" of I_i is put in I_{i1}, the second "1" of I_i is put in $I_{i,2}$ and so on. In the second rotation, the $(q+1)^{\text{th}}$ "one" of I_i is put in $I_{i,1}$, the $(q+2)^{\text{th}}$ "one" of I_i is put in I_{i+2} and so on. This rotating distribution of the "ones" of I_i continues until all the "ones" of I_i have been distributed into the q sub rows.

The above row's splitting results in a new parity check matrix **H** withqm^{k-1}rows which has the following structural properties: (1) each row has weight k (2) every column has weight j (3) any two rows (or columns) have at most one 1-component in common. Such a sparse parity-check matrix is said to be (j, k) regular and the code generated by it is called (j, k) a regular code. The constraint on the rows and columns of **H** given by property (3) is called the row-column (RC) constraint. Therefore, the above row's splitting results in a new parity check matrix with smaller density. Moreover, when the parity check matrix is systematic, one can easily use it to extract the density of the code.

5. Select the rows (a) find the row with the least distance (b) Store shift values of the submatrices (c) UNION ('rows')

The shift values of the submatrices are stored in order to reduce the memory requirement by a factor $1/p$, when $p \times p$ circulant permutation matrices are employed. This lead us to a point that a location of 1 is fixed in the first row and determine the location of other 1 uniquely therefore the required memory for storing the parity check matrix of the QC-LDPC code can be reduced by a factor $1/p$.

3. Results and discussion

3.1 Performance of Large Girth QC-LDPC codes

Fig. 1 compares the BER performance of QC-LDPC code with girth 12 with that of the random constructed code. Both codes have a block length 2041 and have similar performance in low E_b / N_o region. In high region, the new QC-LDPC code with girth 12 performs better than the randomly constructed code approximately by 0.4 dB at a BER of 10^{-5} with 45 iterations.

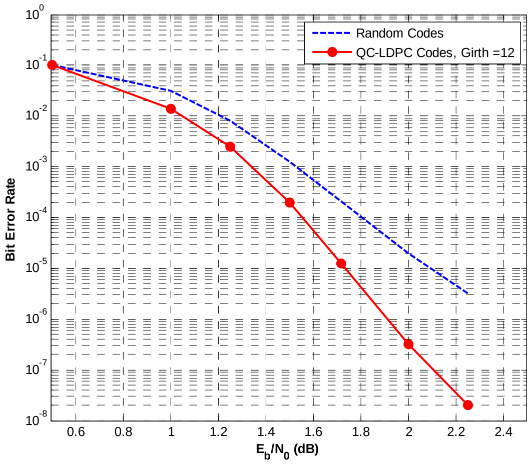

Fig. 1. BER performance of regular girth-12 QC-LDPC codes, with a block length 2041.

Fig. 2 compares the BER performance of QC-LDPC code with girth 16 with that of the randomly constructed code. Both codes have a block length 2947 and maximum number of iterations is set to 25. The performance difference of the randomly constructed and QC-LDPC code are minor in low region of E_b / N_o. However, in the high E_b / N_o region, the proposed code with girth 16 achieves 0.056 dB again at a BER of 10^{-5} over the randomly constructed code.

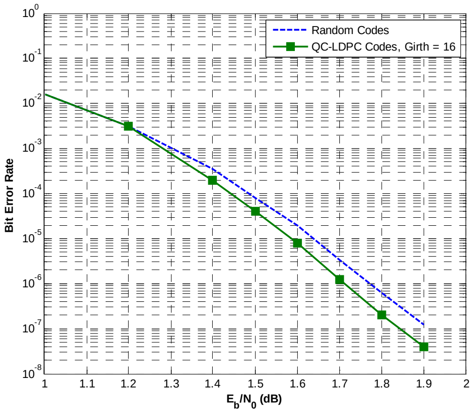

Fig. 2. BER performance of regular girth-16 QC-LDPC codes, with a block length 2947.

Fig. 3 compares the BER performance of QC-LDPC code of girth 20 with that of the randomly generated code. Both codes have a block length 3641 and maximum number of iterations is set to 30. The newly obtained QC-LDPC code of girth 20 outperforms the randomly constructed code approximately by 0.27 dB at 10^{-5} BER. Simulation results show that the new QC-LPDC codes perform better than the randomly generated codes.

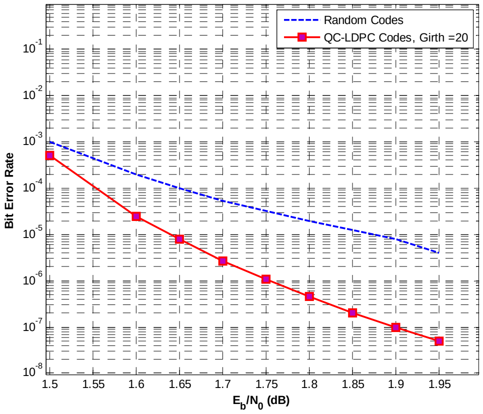

Fig. 3. BER performance of regular girth-20 QC-LDPC codes, with a block length 3641.

Figs. 1 to 3 show that the performance difference of the randomly constructed and QC-LDPC codes are minor in low region of E_b / N_o but at high region of E_b / N_o proposed QC-LDPC codes outperform randomly constructed codes. The reason is that at high SNR, random codes experience error floor due to employment of Guess and Test Algorithm (Fossorier, 2004) for high girth which consumes too much time and results in many short cycles. Table 1 shows the parameters setup and summarizes the performance of the proposed QC-LDPC codes with respect to random codes employing BPSK modulation for different girths.

Table 1. Performance comparison of the proposed QC-LDP C codes with respect to random codes employing BPSK modulation.

Girth	**Code**	**Length**	**Rate**	**Iterations**	**Coding Gain Over Random codes (BER 10^{-5}) dB**
12	(3,13)	2041	0.76	45	0.4
16	(3,7)	2947	0.57	25	0.056
20	(3,11)	3641	0.72	30	0.27

3.2 Performance enhancement of the proposed QC-LDPC codes with different code rates

Performance enhancement and comparison of the newly obtained codes is shown in Fig. 4 employing different code rates. One way to accomplish the best and high rate LDPC code is the puncturing of low rate LDPC code and also some time puncturing columns. The drawback is that it reduces the block length and also effects in the performance. The proposed QC-LDPC codes take union of the rows in order to support the multiple coding rates as mentioned in point (5) of the proposed algorithm. It provides a simple overall architecture with a smaller chip area required to support all the needed rates.

Throughout the simulation BPSK modulation scheme has been employed with 50 maximum number iterations. Simulation results in Fig. 4 portray that the proposed QC-LDPC codes achieve a 0.1dB coding gain over randomly constructed codes and perform 1.3 dB from the Shannon-limit at a BER of 10^{-6} with a code rate of 0.89 for block length of 1332. Simulation results show that the code optimization and girth conditioning not only improves the performance but also lowers the error floor for QC-LDPC codes. On the other hand random codes correspond to a smaller submatrix. The rigid weight constraints of the optimization process do not allow many variations in the random codes construction. The structure of random codes confines the construction of codes with both good cycle and girth properties and directs to error floor.

The method to calculate the Shannon's limit is as follows (Shannon, 1949);

$$R \le B \log 2\left(1+ S/N\right) \tag{10}$$

where, R represents the code rate and replace the signal power S by RE_b and the noise power N by BN_0.

$$R/B \le \log 2\left(1+\left(RE_b / BN_0\right)\right) \tag{11}$$

R/B is called the bandwidth efficiency in units of bit/second/Hz.

substituting E_b/N_0 with $\left(E_b/N_0\right)\min$ and rearranging;

$$\left(E_b/N_0\right)\min \ge \left(2^{(R/B)}-1\right)/(R/B) \tag{12}$$

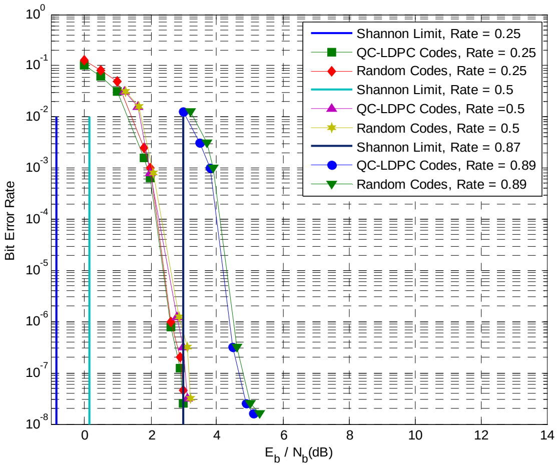

Fig. 4. Shannon's limit of proposed QC-LDPC codes and randomly constructed codes, for block length 1332.

4. Conclusion

Channel coding plays key role in providing a reliable communication method that can overcome signal degradation in practical channels. This paper investigates the potential of large girth Quasi-Cyclic low density parity check codes and compares with renowned codes. Performance evaluation of the newly obtained codes has been established in term of BER and BLER for a given value of E_b / N_o. The newly constructed QC-LDPC codes have both theoretical as well as practical implications. Simulation results demonstrate that the proposed QC-LDPC codes achieve a 0.1dB coding gain over randomly constructed codes and perform 1.3 dB from the Shannon-limit at a BER of 10^{-6} with a code rate of 0.89 for block length of 1332.

References

[1] Z. Ž. Lazarević, N. Ž. Romĉević, M. J. Romĉević, Optoelectron. Adv. Mater. – Rapid Commun., **5**(1), 30 (2011).

[2] H. L. Charles, Error-Control Convolutional Coding, 1st edition, Artech House, Inc. Norwood, MA, USA, 1997.

[3] C. Berrou, A. Glavieux, P. Thitimajshima, IEEE ICC'93, Geneva, Switzerland, 1064 (1993).

[4] R. G. Gallager, Low-Density Parity-Check Code. Cambridge, MA: MIT Press, 1963.

[5] D. Mackay, R. Neal, Electronic Letters, **32**(18), 1645 (1996).

[6] D. J. Mackay, IEEE Trans. Inform. Theory, **45**, 399 (1999).

[7] S. Lin, D. Costello, Error-Control Coding: Fundamentals and Applications, Prentice-Hall, 2nd edition, 2004.

[8] S. B. Wicker, S. Kim, Fundamentals of Codes, Graphs and Iterative Decoding. Kluver International Series in Engineering and Computer Science, 2003.

[9] T. J. Richardson, R. L. Urbanke, IEEE Transactions on Information Theory, **47**(2), 599 (2001).

[10] G. Huo, S. Alouini, IEEE Transaction on Vehicular Technology, **50**(5), 1203 (2001).

[11] H. Zhang, J. M. Moura, GLOBECOM, 4022 (2003).

[12] A. Ohhashi, T. Ohtsuki, IEEE 60th Vehicular Technology Conference, 2004. VTC2004-Fall. **4**, 2530 (2004).

[13] R. M. Tanner, IEEE Transactions on Information Theory, **IT-27**, 533 (1981).

[14] Z. Chen, IEEE Trans. Inform.Theory, **40**, 1666 (1994).

[15] M. Rossi, M. Sala, On a class of quasi-cyclic codes, Technical Report 50, University College Cork, Cork, Ireland, 2005.

[16] T. S. Rappaport, Wireless Communications – Principles and Practice. 2nd edition, by Prentice Hall, 2002.

*Corresponding author: abidusm@gmail.com