Project	**IEEE 802.16 Broadband Wireless Access Working Group <http://ieee802.org/16>**
Title	**Multi-Rate LDPC code for OFDMA PHY**
Date Submitted	**2004-08-17**
Source(s)	Dr. Eli Shasha shasha@runcom.co.il Runcom Technologies Ltd. Prof. Simon Litsyn, litsyn@eng.tau.ac.il Eran Sharon, eranb@eng.tau.ac.il Tel-Aviv University
Re:	This is a response to a call for comments on IEEE P802.16e-D4
Abstract	Proposal of multi-rate LDPC coding scheme for OFDMA PHY mode.
Purpose	Review by the 802.16e working group
Notice	This document has been prepared to assist IEEE 802.16. It is offered as a basis for discussion and is not binding on the contributing individual(s) or organization(s). The material in this document is subject to change in form and content after further study. The contributor(s) reserve(s) the right to add, amend or withdraw material contained herein.
Release	The contributor grants a free, irrevocable license to the IEEE to incorporate material contained in this contribution, and any modifications thereof, in the creation of an IEEE Standards publication; to copyright in the IEEE's name any IEEE Standards publication even though it may include portions of this contribution; and at the IEEE's sole discretion to permit others to reproduce in whole or in part the resulting IEEE Standards publication. The contributor also acknowledges and accepts that this contribution may be made public by IEEE 802.16.
Patent Policy and Procedures	The contributor is familiar with the IEEE 802.16 Patent Policy and Procedures <http://ieee802.org/16/ipr/patents/policy.html>, including the statement "IEEE standards may include the known use of patent(s), including patent applications, provided the IEEE receives assurance from the patent holder or applicant with respect to patents essential for compliance with both mandatory and optional portions of the standard." Early disclosure to the Working Group of patent information that might be relevant to the standard is essential to reduce the possibility for delays in the development process and increase the likelihood that the draft publication will be approved for publication. Please notify the Chair <mailto:chair@wirelessman.org> as early as possible, in written or electronic form, if patented technology (or technology under patent application) might be incorporated into a draft standard being developed within the IEEE 802.16 Working Group. The Chair will disclose this notification via the IEEE 802.16 web site <http://ieee802.org/16/ipr/patents/notices>.

Multi-Rate LDPC code for OFDMA PHY

Dr. Eli Shasha, Runcom Technology

Prof. Simon Litsyn, Tel-Aviv University

Eran Sharon, Tel-Aviv University

1 Introduction

This contribution describes a construction of a structured Multi-Rate Low-Density Parity-Check (MR-LDPC) code for OFDMA PHY. A basic mother code is used for deriving various code rates by merging parity checks of the mother code. The mother code is constructed using a block parity-check matrix with cyclic permutation blocks, similar to the codes suggested by Samsung. All codes derived from the mother code can be implemented on the same hardware of the mother code without additional cost. Since the basic code is structured, the implementation complexity is low. The codes' parity check matrices have a lower triangular parity-bits section, enabling efficient linear encoding. Our flexible matrix's support rates, from ½ to ¾, for both large and small block size. This design of the matrixes improves the performance of both small and long blocks relative to the rigid matrixes proposed by Motorola and Intel, while using less hardware resources. We achieve decoding time linear with block length for better utilization of hardware resources. The matrixes are designed to converge in a small number of iteration to reduce latency and hardware complexity. These properties are important for minimization of power consumption. Matrixes structure is designed to accommodate the 802.16 OFDMA sub-channel structures of 48 carriers and multiples.

2 Construction of the mother code

The mother code is a length N_m rate ½ systematic structured LDPC code. The parity-check matrix of the mother code H_m is constructed using a $\frac{N_m}{2Z} \times \frac{N_m}{Z}$ block matrix H_b consisting of small $Z \times Z$ blocks, which are either the zero matrixes or a cyclic permutation matrix. For brevity of notation we denote the dimensions of the block matrix by $M_b = \frac{N_m}{2Z}$ and $N_b = \frac{N_m}{Z}$.

Let P^i for $i = 0,\ldots,Z-1$ denote the Z cyclic $Z \times Z$ permutation matrices, such that P^0 is the identity matrix.

Let 0 denote the zero $Z \times Z$ matrix. Let $P^{st} = \begin{bmatrix} 1 & & & & \\ 1 & 1 & & \mathbf{0} & \\ & 1 & \ddots & & \\ & & \ddots & 1 & \\ \mathbf{0} & & & 1 & 1 \end{bmatrix}$ denote a $Z \times Z$ stairs matrix.

The block parity-check matrix is constructed based on the following rules:

1) The $M_b \times M_b$ right had side of the matrix is lower triangular

2) The lower rightmost block of the matrix is P^{st}

3) Each row in the block matrix contains 6 or 7 non-zero block .

4) H_b is an irregular matrix, such that the block columns weights spectrum is optimized for providing good code performance in a small number of decoding iterations. (a block column's weight is the number of non-zero block entries in a column).

5) Rows j and $j + \dfrac{M_b}{2}$ for $j = 1, \ldots, \dfrac{M_b}{2}$ have no overlapping non-zero block entries.

6) Let $H_{m/2}$ denote the resulting parity-check matrix that is obtained by summing up the lower and upper $\dfrac{M_b}{2} \times N_b$ parts the block matrix H_b. Then, the resulting parity check matrices H_m and $H_{m/2}$ do not contain short cycles.

An example of such a block matrix for $N_m = 384, Z = 24$ is:

$$
H_b = \begin{bmatrix}
0 & 0 & P^{20} & P^{20} & P^2 & P^{22} & P^{23} & P^8 & P^{20} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & P^{21} & P^7 & P^{22} & P^{20} & P^{13} & P^{13} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & P^{21} & P^{10} & P^{21} & P^{11} & P^{18} & P^9 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & P^{14} & P^9 & P^{10} & P^{20} & P^7 & P^4 & 0 & 0 & 0 & 0 \\
P^0 & P^{21} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & P^{12} & P^9 & P^{10} & P^{12} & 0 & 0 & 0 \\
P^5 & P^{17} & P^9 & P^6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & P^{13} & P^4 & P^3 & 0 & 0 \\
0 & P^1 & P^1 & P^{12} & P^{10} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & P^{22} & P^3 & P^{13} & 0 \\
0 & 0 & 0 & P^{15} & P^{23} & P^{18} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & P^{22} & P^7 & P^{st}
\end{bmatrix}
$$

3 Deriving codes of various rates from the mother code

Codes of various rates between ½ and ¾ can be derived from the mother code by summing up pairs of rows of the block matrix H_b. A code of rate $\dfrac{1}{2} + \dfrac{\alpha}{N_b}$ for any $\alpha \in \{0, \ldots, \dfrac{N_b}{4}\}$ can be obtained by summing up the rows j and $j + \dfrac{M_b}{2}$ for $j = 1, \ldots, \alpha$. The resulting block matrix after summation is a $(M_b - \alpha) \times N_b$, hence the code rate is $1 - \dfrac{M_b - \alpha}{N_b} = \dfrac{1}{2} + \dfrac{\alpha}{N_b}$. Note that due to property (5) of the constructed mother code, the resulting matrix after summation remains a block matrix with permutation blocks as its entries. All derived codes are of length N_m. All derived codes have the same number of 1's in their parity-check matrix (again, due to property (5)).

3 Encoding

All the constructed codes have a lower triangular block parity-check matrix; hence they have a very simple linear encoding based on Gaussian elimination.

Actually, encoding can be performed using the decoder by initializing the information bits part of the decoder input messages with the information bits and the parity bits part of the decoder input messages with erasures. Then the decoder can (always) recover the erased bits due to the lower triangular structure of the parity check matrix. If the parity checks are processed one after another by the decoder, then in each check only a single bit is unknown and the decoder sets this bit to be the XOR of the 6 (or 5) known bits in the check.

4 Accommodating various code lengths

Various code lengths can be accommodated in 3 ways:

1) Shortening the mother code and its derivatives: shorter codes can be obtained by shortening each of the length N_m codes of rates ½ to ¾ that are obtained from the mother code. Any pair of code length N_c and code rate R_c maintaining $\frac{N_m}{4} \le N_c(1-R_c) \le \frac{N_m}{2}$ can be obtained by shortening one of the length N_m codes. Hence, larger flexibility in code length and rate is obtained compared to using a single rate mother code as suggested by Intel and Motorola.

2) Changing the size Z of the permutation blocks, which inflates/deflates the parity-check matrix, as suggested by Samsung.

Both schemes suggested above suffer from the same disadvantage: the decoding time remains the same regardless of the code length. When shortening is used the decoder still works on the complete code even though the actual code length might be very small and the decoding time of short and long blocks is the same. When the permutation block size Z is used for changing the code length, the effect is the same. The block matrix structure allows to process Z rows or columns of the parity check matrix simultaneously using Z processors. When we "deflate" the permutation block size by factor α we obtain a shorter code of length αN_m, but it also means that only αZ processors are active. Hence, the hardware is not used efficiently, and the decoding time of the length αN_m code remains the same as the decoding time of a length N_m code.

For this reason we believe that the following third method is preferred:

3) Using several mother codes with different lengths: We suggest using 4 mother codes of lengths $N_m = 2304, 1728, 1152, 576$ with $Z = 12$. The code lengths are chosen as multiples of 48 bits to accommodate an integer number of OFDMA sub-channels. The permutation block size is chosen to be $Z = 12$ for providing the possibility for sufficient parallelism to allow enough decoding iterations at high throughput and sever latency conditions as described in the next section. Note that all the mother codes and the codes derived from them can be implemented on the same hardware. Adding each additional mother code requires only a ROM for maintaining its block parity check matrix H_b. The ROM size required for maintaining the matrix H_b is $\frac{7N_m}{2Z} \times (\lceil \log_2 \frac{N_m}{Z} \rceil + \lceil \log_2 Z \rceil)$ bits. For the largest length 2304 code this amounts to 4032bit ROM. For maintaining all 4 mother codes a 9744bit ROM is needed. Thus, at a low cost of an increased ROM size we obtain full rate and code length flexibility. The same encoder/decoder hardware supports lengths 2304, 1728, 1152 and 576 codes of any rate between ½ and ¾. Obtaining other code length between the 4 mother codes' lengths is done through the shortening mechanism.

The advantage of this method is that the decoding time is linear with the code length since all mother codes are constructed using 12×12 permutation blocks. This provides a good solution for both streaming applications and immediate ACK applications. In streaming applications, codes of different length perform the same number of iterations for a given throughput (while in the first two methods the performance of shorter codes would be severely impaired because they will be able to perform less decoding iterations). In immediate ACK applications shorter codes can be used when lower latency is needed.

5 Structured Vs. Random code construction
Assume the following code parameters and system parameters:
N – code length
R – code rate
$M = N(1\text{-}R)$ – number of rows in the parity-check matrix
d_c – average parity-check matrix row weight
I – desired maximal number of iterations (assuming no statistical multiplexing is used)
f – system clock
Z – decoder parallelism (number of processors)
L – decoding latency

The standard decoding algorithms for LDPC codes are based on iterative message-passing algorithms. These algorithms rely on a graph-based representation of codes, where the decoding can be understood as message passing between nodes in a bipartite graph. In each iteration, messages are passed on all edges of the graph in both directions. The number of edges in the graph is equal to the number of 1's in the code's parity-check matrix, which is $M \times d_c$. Hence, a decoding iteration involves reading $M \times d_c$ messages from a memory, processing the messages and writing the updated messages back to memory. Assuming required decoding latency L, a system clock frequency f and a desired maximal number of iterations I, a decoding iteration should be performed in $\dfrac{fL}{I}$ clocks. This means that in each clock $\dfrac{Md_c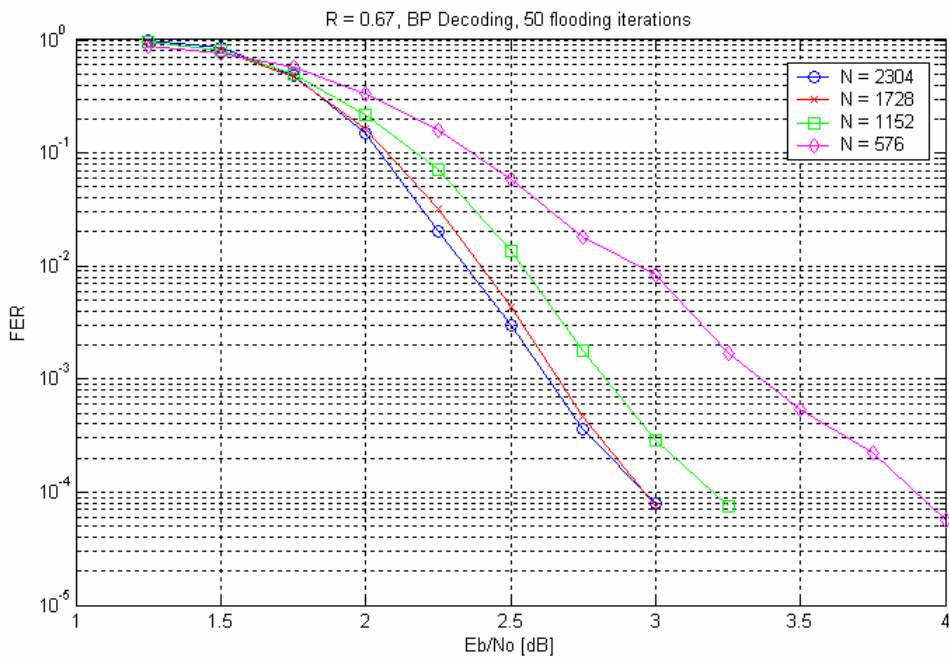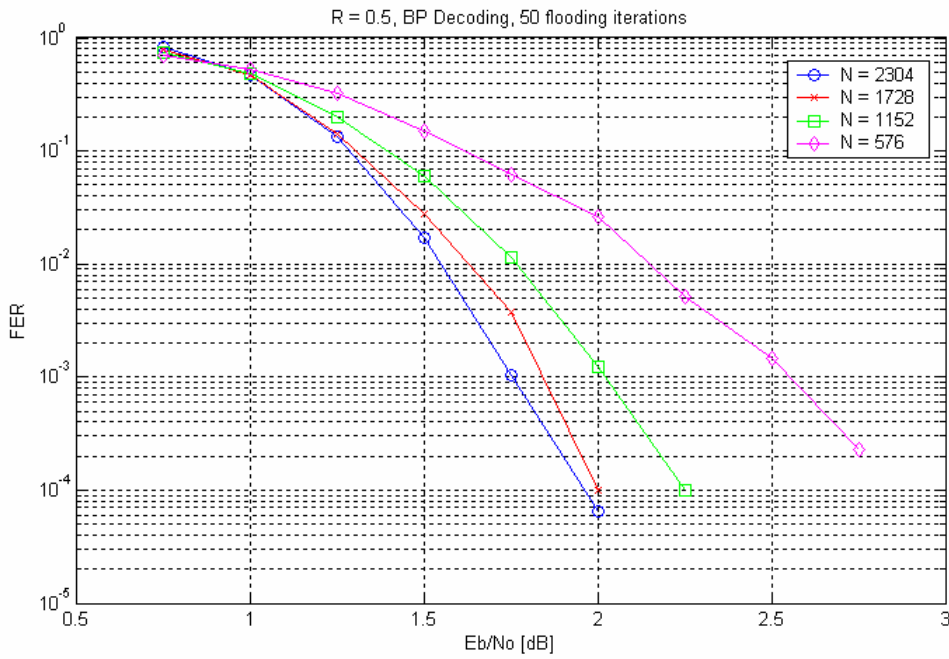I}{fL}$ messages should be read (and written) to the memory. One can then immediately see that many messages should be read each clock from the memory. This implies that using a randomly constructed LDPC code would require saving the messages in registers (since $\dfrac{Md_cI}{fL}$-port RAM do not exist) leading to increased chi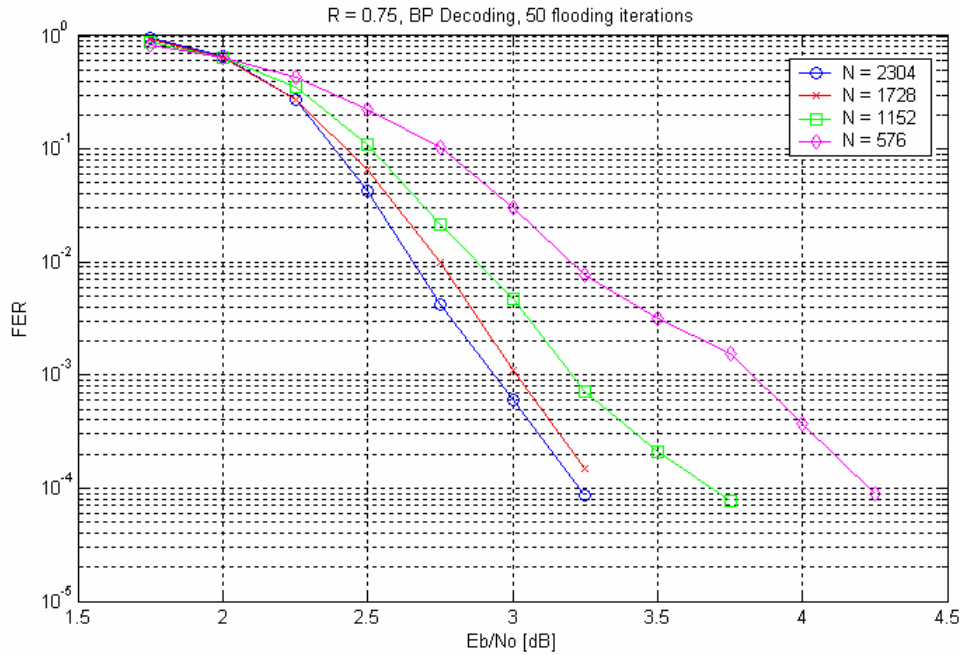p area. Furthermore, it would lead to severe routing problems of messages into processors, rendering the design impractical.
Using a structured block LDPC code can solve these problems, enabling simple simultaneous processing of sets of Z messages, by facilitating both their memory management and their routing from the memory to the processing units.

6 Simulation results

The following figures show simulation results performed over AWGN channel using standard Belief-Propagation decoder with 50 iterations. Simulations are provided for the basic code rates 1/2,2/3,3/4 (though the MR-LDPC codes support any code rate between ½ and ¾) and the four basic code lengths 2304,1728,1152,576.

R = 0.5, BP Decoding, 50 flooding iterations

R = 0.67, BP Decoding, 50 flooding iterations

7 Appendix

The block matrices H_b of MR-LDPC mother codes of length 2304,1728,1152 and 576 with $Z = 12$ permutation blocks are provided in the attached Matlab workspace file. Note that the zero matrix is described as a -1 entry in the matrix, a permutation matrix P^i is described as a i entry in the matrix and the stairs matrix P^{st} is described as a Z entry in the matrix. Codes of various rate are derived from the mother code matrix H_b by merging rows of the matrix as described in Section 3.

We also provide files containing parity-check matrices in Intel's format for codes of length 2304,1728,1152,576 and code rates 1/2,2/3,3/4.