# Iterative Symbol-by-Symbol Decoding of LDPC Codes and Constellation Mapping for Multilevel Modulation

Hongtao Jiang*, Zhongfeng Wang*, and Huaping Liu†
*Broadcom Corp., Irvine, CA 92617, USA, email: {hongtaoj, zfwang}@broadcom.com
†School of EECS, Oregon State University, Corvallis, OR 97331 USA. E-mail: hliu@eecs.oregonstate.edu

***Abstract*—Iterative probabilistic decoding of binary low-density parity-check (LDPC) codes have been studied extensively. Non-binary LDPC codes have recently attracted an increasing attention. Most of the existing non-binary codes are built over GF($2^q$), and decoding methods developed for binary LDPC codes cannot be used directly with multilevel modulations. In this paper, we first extend the binary parity-check codes to the case with multiple symbols over modular arithmetic. Then, we develop a sum-product algorithm to decode this new type of codes at the symbol level. Finally, we propose an effective constellation mapping method for multilevel modulations. Error performances of this type of codes with 4-PAM, 4-PSK, and 16-PSK modulations over AWGN channels are provided. Compared with uncoded systems, the coding gain of a medium-size regular LDPC code of rate 8/9 with 4-PAM and 4-PSK modulations is about 5 dB. With 16-PSK and an appropriate constellation mapping at a code rate 3/4, the proposed code's performance is comparable to that of trellis codes.**

***Index Terms*—Parity-check code, multilevel modulation, iterative probabilistic decoding, and constellation mapping.**

## I. Introduction

Low-density parity-check (LDPC) codes were first introduced by Gallager in 1961 [1]. They were rediscovered in the late 1990's [2], [3]. It has been shown that the performance of irregular LDPC codes could be very close to the Shannon limit [4], [5]. Most of existing studies are restricted to LDPC codes over GF($2$), which cannot be used directly with multilevel modulations. Non-binary LDPC codes [6]–[8] have recently attracted a growing research interest. Most of these codes are defined over GF($q$). In [7], binary LDPC codes with multilevel-symbol mapping are discussed. In this scheme, the *a posteriori probability* (APP) at the bit level is computed first and then the sum-product algorithm [1], [2] is used for soft decision iterative decoding. Mapping mechanisms based on Gray coding are proposed in [8].

In this paper, we first show that binary LDPC codes can be extended to the case with multiple symbols and the sum-product algorithm could be modified for decoding at the symbol level over modular arithmetic. Then we propose a new constellation mapping scheme for multilevel modulations. We develop this type of codes and the decoding algorithm, and compare their performances when 4-PSK, 4-PAM, and 16-PSK modulations are employed.

CHINACOM 2010, August 25-27, Beijing, China

DOI 10.4108/chinacom.2010.7

## II. Parity-Check Codes For Multilevel Modulation

The conventional low-density parity-check codes are built over GF($2$), that is, they are over the set $\{0, 1\}$ with addition defined by

$$a + b = c \text{ mod } 2. \quad (1)$$

This can be easily extended to the modulo-$N$ case, where the addition is defined over the set $\{0, 1, \cdots N - 1\}$ as

$$a + b = c \text{ mod } N. \quad (2)$$

A parity-check code is defined through a $k \times n$ parity-check matrix $\boldsymbol{H}$. An $n$-tuple $\boldsymbol{c}$ is a code word if and only if it is orthogonal to every row vector of $\boldsymbol{H}$, that is,

$$\boldsymbol{c}\boldsymbol{H}^T = 0 \text{ mod } N \quad (3)$$

where $N$ is the number of multilevel signals in our case, and each element of $\boldsymbol{c}$ is in $\{0, 1, \cdots N - 1\}$.

For a low-density parity-check code, $\boldsymbol{H}$ is a sparse matrix. In this paper we only consider regular LDPC codes, for which $\boldsymbol{H}$ has $d_c$ non-zero elements in each column and $d_r$ non-zero elements in each row. Theoretically, a non-zero element of $\boldsymbol{H}$ does not have to be one[1], but for simplicity we restrict the non-zero element of $\boldsymbol{H}$ to be 1. For simplicity and clarity of description, we will call this type LDPC codes $N$-LDPC codes; consequently the conventional binary LDPC code would be a special case, 2-LDPC codes.

For a multilevel modulation with $N$ different levels, every signal level $A_i$ can be mapped to a single digit $i$, $i \in \{0, 1, 2, \cdots N - 1\}$. Therefore, a code word $\boldsymbol{c}$ defined by the parity-check matrix $\boldsymbol{H}$ can be mapped to the corresponding sequence of multilevel signals. Theoretically, the best signal constellation mapping should maximize the Euclidean distance among code words.

## III. Decoding Algorithm

Now let $y$ be the received signal in an additive white Gaussian noise (AWGN) channel expressed as

$$y = A + n \quad (4)$$

[1]It can be shown that as long as the element is relatively prime to $N$, with minimum modifications, the code design will still be valid. This will be discussed in our future research.

where $n$ is Gaussian noise with variance $\sigma_n^2$. The *posteriori* probability of code symbol $x$ being equal to $i$ in $\{0, 1, \cdots N-1\}$ is

$$\text{pr}(x = i|y) = \frac{e^{-(y-A_i)^2/(2\sigma_n^2)}}{\sum_{k=0}^{N-1} e^{-(y-A_k)^2}/(2\sigma_n^2)}. \quad (5)$$

Let $\text{Pr}(k|y)$ be the discrete Fourier transform (DFT) of $\text{pr}(x = i|y)$ expressed as

$$\text{Pr}(k|y) = \sum_{k=0}^{N-1} \text{pr}(x = i|y) e^{-j2\pi ik/N}. \quad (6)$$

We have the following theorem:

*Theorem 1:* Let $x_m$ $(m = 0, 1, \cdots M)$ be a single parity-check code

$$\sum_{k=0}^{N-1} x_m = 0 \text{ mod } N. \quad (7)$$

Let the probability of $x_m = i$ $(i = 0, 1, \cdots, N-1)$ be $\text{pr}(i, m)$ and its DFT be denoted as $\text{Pr}(k, m)$. The extrinsic probability $\text{pr}_{\text{ext}}(i_0, 0)$ with known probability $\text{pr}(i, m)$, $(m = 1, 2, \cdots, M)$ can be computed as

$$\text{pr}_{\text{ext}}(i_0, 0) = \frac{1}{N}\text{DFT}\left(\prod_{m=1}^{M} \text{Pr}(k, m)\right) \text{ at } i_0. \quad (8)$$

*Proof:* The extrinsic probability of code symbol in position $0, x_0 = i_0$ in $\{0, 1, \cdots, N-1\}$ is

$$\text{pr}_{\text{ext}}(i_0, 0) = \sum_{(i_1+\cdots,i_M)=-i_0 \text{ mod } N} \left(\prod_{m=1}^{M} \text{pr}(i_m, m)\right). \quad (9)$$

The product of $\text{Pr}(k, m)$ is

$$\begin{aligned} \prod_{m=1}^{M} \text{Pr}(k, m) &= \prod_{m=1}^{M} \sum_{i_m=0}^{N-1} \text{pr}(i_m, m) e^{-j2\pi k i_m/N} \\ &= \sum_{i_1,\cdots,i_M} \prod_{m=1}^{M} \text{pr}(i_m, m) e^{-j2\pi \sum \frac{i_m k}{N}}. \quad (10) \end{aligned}$$

The DFT of the extrinsic probabilities $\text{pr}_{\text{ext}}(i_0, 0), i_0 \in \{0, 1, \cdots, N-1\}$, at value $k$, is computed as

$$\begin{aligned} &\sum_{i_0=0}^{N-1} \text{pr}_{\text{ext}}(i_0, 0) e^{-i2\pi k i_0/N} \\ &= \sum_{i_0}^{N-1} \sum_{i_1+\cdots+i_M=-i_0 \text{ mod } N} \prod_{m=1}^{M} \text{pr}(i_m, m) e^{-j2\pi i_0 k/N} \\ &= \sum_{i_1,\cdots,i_M} \prod_{m=1}^{M} \text{pr}(i_m, m) e^{-j2\pi i_m k/N} \\ &= \prod_{m=1}^{M} \text{Pr}(k, m). \quad (11) \end{aligned}$$

Hence,

$$\begin{aligned} \text{pr}_{\text{ext}}(i_0, 0) &= \frac{1}{N} \sum_{k=0}^{N-1} \prod_{m=1}^{M} \text{Pr}(k, m) e^{-j2\pi k i_0/N} \\ &= \frac{1}{N}\text{DFT}\left(\prod_{m=1}^{M} \text{Pr}(k, m)\right) \text{ at } i_0. \quad (12) \end{aligned}$$

This concludes the proof.

It is easy to see that when $N = 2$, the result is the same as Gallager's result [1], which is the core of the sum-product algorithm.

Let the set of bits for check node $m$ be $C_s(m) = \{i : H_{mi} = 1\}$. Let the set of checks for bit $i$ be $B_s(i) = \{m : H_{mi} = 1\}$. We denote a set $B_s(i)$ excluding check $m$ as $B_s(i)\backslash m$. Now the sum-product algorithm for the $N$-LDPC codes can be stated as following:

**Initialization:** let the log-likelihood of the input probability be

$$ll_{\text{input},i}(k) = \log\left(\text{pr}(k, i)\right), \; k = 0, 1, \cdots, N-1 \quad (13)$$

where $\text{pr}(k, i)$ can be calculated by using Eq. (5) for the AWGN channel, and the extrinsic log-likelihood of the probability is computed as

$$ll_{\text{ext},mi}(k) = 0, \; k = 0, 1, \cdots, N-1. \quad (14)$$

Notice that the target is the log-likelihood ratio. Therefore, we can add or subtract a common factor for log-likelihood. For example, for the AWGN case of Eq. (5), we can choose

$$ll_{\text{input},i}(k) = (2yA_k - A_k * A_k)/(2\sigma^2). \quad (15)$$

**Vertical Step:** for the check $m$, for each bit $i \in C_s(m)$, calculate the following:

$$ll_{mi}(k) = ll_{\text{input},i}(k) + \sum_{m' \in B_s \backslash m} ll_{\text{ext},m'i(k)}. \quad (16)$$

**Horizontal Step:** from the $ll_{mi}(k)$ for the check $m$, calculate the probability $\text{pr}(k, i)$ for each $k$. From the main theorem, calculate the extrinsic probability $\text{pr}_{\text{ext},m}(k, i)$ for the parity check $m$. Then the new $ll_{\text{ext},mn}(k)$ is calculated through the following equation:

$$ll_{\text{ext},mn}(k) = \log\left(\text{pr}_{\text{ext},m}(k, n)\right). \quad (17)$$

**Output:** At the end of iteration, the output log-likelihood can be calculated as

$$ll_{\text{output},i}(k) = ll_{\text{input},i}(k) + \sum_{m \in B_s(i)} ll_{\text{ext},mi(k)}. \quad (18)$$

From the output log-likelihood, we choose the $k$ which has the biggest $ll_{\text{output},i}(k)$ as the decoder output.

### IV. Constellation Mapping

For multilevel modulations, there are various constellation mapping schemes between symbols and the constellation. For each pair of code words, the Euclidean distance between them can be calculated. It is essential to keep this distance as large as possible. In reality, the number of code words could be huge and it is not easy to determine the constellation mapping by maximizing the Euclidean distance between each pair of code words.

Instead of considering all code words of the original LDPC codes, we propose to consider the single parity check code with two symbols:

$$c_0 + c_1 = 0 \text{ mod } 16. \tag{19}$$

We then try to find the constellation mapping to maximize the Euclidean distance for the above code words. Because of the very small set of code words for the above single parity check code, constellation mapping can be found through an exhaustive search or via a Monte-Carlo method.

### V. Results

In the simulation, after each vertical step and horizontal step, the output is calculated. If the output satisfies the parity check, the simulation stops. A failure is declared if no valid code word is found after 20 iterations.

Fig. 1 compares the performance of 4-LDPC codes with 4-PAM modulation and that of the uncoded 4-PAM system. The code rate is 8/9, and $d_r = 27, d_c = 3$ are chosen. The signal constellation mapping is $\{0 \to -3, 1 \to -1; 2 \to -1; 3 \to 3\}$. For the $(1134\ 126)$ 4-LDPC code, the gain is about 5 dB at a symbol error rate of $10^{-5}$, and the $(2268\ 252)$ 4-LDPC code performs about 0.5 dB better than the $(1134\ 126)$ 4-LDPC code at the same symbol error rate.

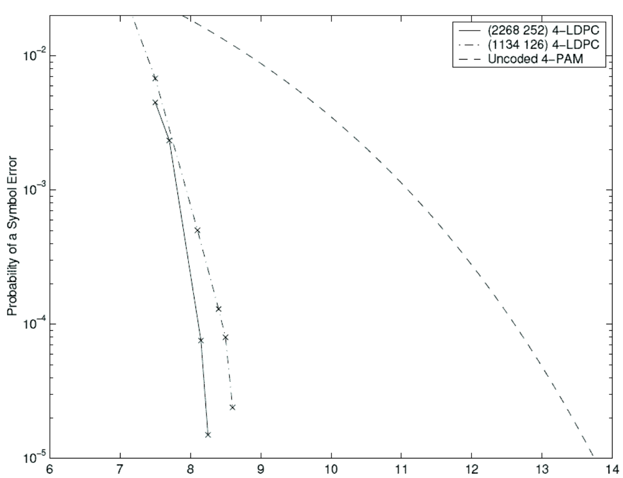


Fig. 1. Comparison of symbol error rates for $(2268\ 252)$ 4-LDPC, $(1136\ 126)$ 4-LDPC with 4-PAM modulation (code rate 8/9), and uncoded 4-PAM.

Fig. 2 compares the performance of 4-LDPC codes with 4-PSK modulation and that of the uncoded 4-PAM system. The LDPC code is the same as the previous one. The signal constellation mapping is $\{i \to e^{j2\pi/4}\}$. For the $(1134\ 126)$ 4-LDPC code, the gain is about 5 dB at a symbol error rate of $10^{-5}$, and the performance of the $(2268\ 252)$ 4-LDPC code is about 0.5 dB better than that of the $(1134\ 126)$ 4-LDPC code at the same error rate.

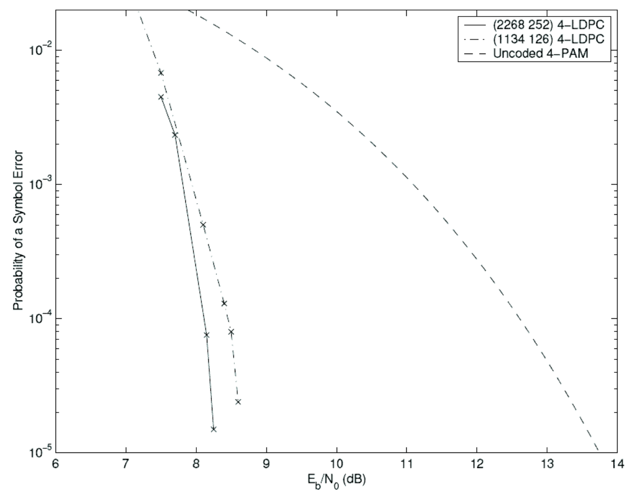


Fig. 2. Symbol error rates for $(2268\ 252)$ 4-LDPC, $(1134\ 126)$ 4-LDPC with 4-PSK modulation (code rate 8/9), and uncoded 4-PSK.

For the 16-LDPC codes, we have tried several different signal constellation mappings. We have also conducted a computer search to find the constellation mapping that has the largest minimum Euclidean distance among the code words for a single parity-check code $c_0 + c_1 = 0$ mod 16. The resulting constellation mapping is shown in Fig. 3 and the minimum distance among the code words is 1.18.

It is interesting to compare the performances of $N$-LDPC codes and trellis codes [9], since both of them can use large sets of signals for data transmission. Fig. 4 compares the performance of 16-LDPC codes with 16-PSK modulation and that of the uncoded 8-PSK system. The code rate is $3/4$, and $d_r$ and $d_c$ equal 12 and 3, respectively. The constellation mapping shown in Fig. 3 is applied for the 16-LDPC code. Compared with the uncoded system, the gain of the $(1004\ 251)$ code is about 4.6 dB at a symbol error rate of $10^{-5}$, which is comparable to the performance of the trellis code providing the same spectral efficiency. We have also tried other signal constellation mappings; the coding gains range from 1 dB to about 5 dB over the uncoded system at a symbol error rate of $10^{-5}$. From these results for this example, we observe that it is critical to optimize the signal constellation mapping to maximize the coding gain for the $N$-LDPC codes.

### VI. Conclusion

We developed a new multilevel LDPC code and proposed an iterative symbol-by-symbol decoding algorithm for this code. We have also proposed an effective constellation mapping method. Compared with the uncoded 4-PSK, the proposed

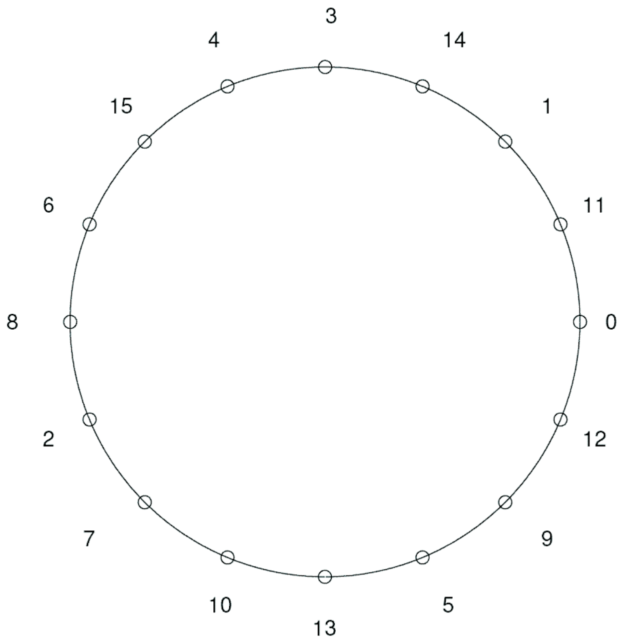


Fig. 3. The signal constellation mapping for 16-PSK obtained via computer search that has a minimum Euclidean distance of 1.18 for the single parity check code $c_0 + c_1 = 0$ mod 16.

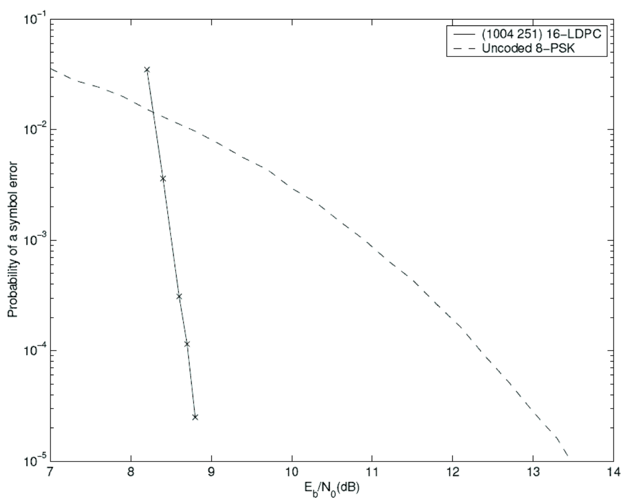


Fig. 4. Symbol error rates for (1004 251) 16-LDPC with 16-PSK modulation (code rate-3/4) and uncoded 8-PSK.

code with a coding rate of 8/9 and 4-PSK and 4-PAM modulations, which provides roughly the same spectral efficiency as the uncoded system, achieves about 5 dB coding gain. For the 16-LDPC code with 16-PSK modulation and a code rate 3/4, the error performance depends on the signal constellation mapping. With one mapping scheme obtained via computer search, the performance is comparable to that of the trellis code at the same spectral efficiency. Since this is a large family of LDPC codes, more research is needed to fully exploit the potential of these codes.

## Acknowledgment

The authors are very grateful to Dr. G. Ungerboeck for his critical comments. The authors also thank Mr. L. Longo for his support of this work.

## References

[1] R. G. Gallager, *Low Density Parity Check Codes*, Cambridge, MA, MIT Press, 1963.
[2] D. J. C. Mackay, “Good error correcting codes based on very sparse matrices,” *IEEE Trans. Inform. Theory*, vol. 45, pp. 399–431, Mar. 1999.
[3] M. Sipser and D. Spielman, “Expander codes,” *IEEE Trans. Inform. Theory*, vol. 42, pp. 1710–1722, Nov. 1996.
[4] S. Y. Chung, J. G. D. Forney, T. Richardson, and R. Urbanke, “On the design of low-density parity-check codes within 0.0045 dB of the Shannon limit,” *IEEE Commun. Lett.*, vol. 5, pp. 58–60, Feb. 2001.
[5] T. Richardson, M. A. Shokrollahi, and R. L. Urbanke, “Design of capacity-approaching irregular low-density parity-check codes,” *IEEE Trans. Inform. Theory*, vol. 47, pp. 619–637, Feb. 2001.
[6] M. C. Davey and D. J. C. MacKay, “Low density parity check codes over GF(q),” *IEEE Comm. Lett.*, vol. 2, no. 6, pp. 165–167, June 1998.
[7] E. Eleftheriou and S. Olcer, “Low-density parity-check codes for digital subscriber lines,” in *Proc. IEEE ICC’02*, vol. 3, pp. 1752–1757, May 2002.
[8] C. Luo, K. Liu, Z. Fei, and J. Kuang, “A noval mapping scheme in non-binary LDPC coded modulation system,” in *Proc. 11th IEEE Int. Conf. Communi. Tech.*, 2008.
[9] G. Ungerboeck, “Channel coding with multilevel/phase signals,” *IEEE Trans. Inform. Theory*, vol. 28, pp. 55-66, Jan. 1982.