Non-uniform Quantization of Soft Information for 5G LDPC Codes

Xiaohua Chang1* and Yu Zhang2
1,2 Beijing National Research Center for Information Science and Technology, Department of Electronic Engineering, Tsinghua University, Beijing 100084, China
*cxh17@mails.tsinghua.edu.cn

Abstract. This paper proposes a non-uniform quantization algorithm of soft information based on density evolution for the decoding of low-density parity check (LDPC) codes, which applies a non-linear transform function to soft information, and provides good block error rate performance. The transform function is obtained by minimizing weighted mean square quantization error for every combination of modulation type and LDPC codes and has an analytical expression of distribution of soft information. It can be realized as look-up table via quantization of transform function and has low complexity of implementation. Simulation results show that the error performance of proposed algorithm is close to no quantization condition and has a gain of around 0.3 dB compared to uniform quantization algorithm for 5G's rate 2/3 LDPC codes with APSK modulation over additive white Gaussian noise (AWGN) channel.

1. Introduction
LDPC codes [1] were introduced by Gallager in the 1960s, and have attracted lots of interest for the promising prospect in forward error correction with excellent performance and simplicity of iterative decoding. LDPC codes have been adopted as several IEEE standards such as IEEE 802.16e, IEEE 802.11n. The latest release of cellular standard in 3GPP [2] has adopted LDPC codes as data channel codes. Compared with LTE turbo codes and other commercialized LDPC codes, 5G New Radio LDPC codes have several advantages. They have better area throughout efficiency, reduced decoding complexity, better decoding latency, and improved block error rate performance.

Soft decoding of LDPC codes has the ability to achieve good performance with the iterative decoding algorithm referred to as sum-product algorithm [3]. It is based on belief propagation thus has great computation loads not suitable for hardware implementation. [4] proposed a complexity-reduced iterative decoding algorithm referred to as min-sum algorithm. It does not need signal-to-noise ratio (SNR) and complex arithmetic. With minor correction, by introducing a normalization factor $\alpha$, the performance improves a lot. Normalized min-sum (NMS) algorithm is applied in practical hardware implementation.

Quantization of soft information has great impact on BLER performance in hardware implementation. In previous works, uniform quantization with truncation is adopted [5]. This strategy works well in BPSK modulation, in which case the probability distribution of soft information is mixed Gaussian distribution. In the design of LDPC codes [6], threshold is achieved by density evolution assuming the BPSK modulation and AWGN channel. However, in high modulation mode such as Amplitude Phase Shift Keying (APSK), the probability distribution has a long tail. Generally,
by setting a maximum value $V$, soft information value higher than $V$ is truncated to $V$. This method brings performance loss, which varies with threshold value $V$.

In this paper, we proposed a non-uniform quantization method which minimizes weighted mean square quantization error. The weight is gained by density evolution which is large in high probability interval of soft information and is small in low probability interval. Generally, uniform quantization is suitable for BPSK and QPSK modulation type as the soft information is mixed Gaussian distribution which is conserved in density evolution. Non-uniform quantization is suitable for high skewness distribution which needs extra non-linear transform. The transform function is saved as look-up table and needs minor hardware resources. Simulation results show that this method has little performance loss compared with no quantization method and huge performance gap over uniform quantization method especially in high modulation type.

2. Preliminary

2.1. Soft information

In the design of communication system, decoding with soft decision has great performance gains. Suppose the constellation set is $c$, which corresponds to $m$ bits $b = \{b_0, b_1, \cdots, b_m\}$, the soft information referred to as log-likelihood ratio (LLR) is $C = \{C_0, C_1, \cdots, C_m\}$ which is defined as $C_j = \log \frac{P(c_j = 0)}{P(c_j = 1)}$. From paper [7], the soft information has the expression

$$C_j = \beta \left[ \min_{a_{a(j)}} |r - a|^2 - \min_{a_{a(j)}} |r - a|^2 \right]$$

In which $\beta$ is a constant associated with SNR, $r$ is the received signal, $c_0(j)$ is the set of constellation corresponding to $b_j = 0$, $c_1(j)$ is the set corresponding to $b_j = 1$.

With random information bits and AWGN channel, the soft information has simple distribution form for BPSK and QPSK modulation. However, in APSK modulation, the $C_j$ can vary in a large range with low probability.

2.2. Decoding

We assume that an $(N, K)$ LDPC code with information length $K$ and block length $N$. Parity check matrix is denoted as $H = [h_{nm}]$ of $M = N - K$ rows and $N$ columns. The code words satisfy $H \cdot c = 0$. The parity check matrix can be treated as bipartite graph, with every row as a check node, every column as a variable node, and $h_{mj} = 1$ corresponds to a connection of check node $m$ and variable node $c_j$. $M(j)$ is the set of check nodes connected with variable node $c_j$, and $N(m)$ is the set of variable nodes connected with check node $m$. We define the following notations in one iteration.

- $C_j$: The a-priori soft information of bit $c_j$ computed from received signal.
- $R_{mj}$: The LLRs sent from check node $m$ to variable node $c_j$.
- $L(q_{mj})$: The LLRs sent from variable node $c_j$ to check node $m$.
- $L(q_j)$: The a-posteriori LLRs of bit $c_j$ after each iteration.

The decoding process works as follows. The ram matrix $R_{mj}$ is initialized as 0. Then for each iteration, compute the LLRs of every bit.

$$L(q_j) = \sum_{m \in M(j)} R_{mj} + C_j$$

$$L(q_{mj}) = L(q_j) - R_{mj}$$
Then the ram matrix of min-sum algorithm is updated as follows.

\[
R_{mj} = \prod_{n \in N(m), n \neq j} \text{sgn}(L(q_{mj})) \cdot \min_{n \in N(m), n \neq j} |L(q_{mj})| \tag{4}
\]

After each iteration, the hard decision of information bits is computed by setting \( c_j = 0 \) if \( L(q_j) \geq 0 \) and setting \( c_j = 1 \) if \( L(q_j) < 0 \). If the condition \( \mathbf{H} \cdot \mathbf{e} = \mathbf{0} \) is satisfied or the number of iteration reach the maximum iteration, stop the decoding process.

3. Non-uniform quantization algorithm

3.1. Density evolution

The density evolution of min-sum algorithm can be found in [8], which is presented as follows. A specific \( R_{mj} \) is a function of \( d_c - 1 \) random variables of independent identically distribution denoted as \( L = \{L_0, L_1, \cdots, L_{d_c - 1}\} \) with probability distribution of \( p_L(x) \). \( d_c \) is the degree of check node \( m \).

Supposing \( p_L(x) \) is even function, we define \( \phi(x) = \int_{-\infty}^{\infty} p_L(z)dz + \int_{-\infty}^{x} p_L(z)dz \) for \( x > 0 \). The probability density function (pdf) of \( R_{mj} \) can be obtained as

\[
p_R(x) = (d_c - 1)p_L(x)\phi(|x|)^{d_c - 2} \tag{5}
\]

For an irregular LDPC code in which check nodes have different degrees, the pdf of ram matrix \( R \) is associated with code ensembles which can be expressed as [9]

\[
p_R(x) = \rho^i(\phi(x))p_L(x) \tag{6}
\]

In which \( \rho(x) = \sum_{i=2}^{d_c} \lambda_i x^{i-1} \) is check node distribution, and \( \lambda_i \) is the fraction of edges emanating from check node of degree \( i \).

As we can see, the larger values of soft information is less useful because of the operation min. The probability density function of check node LLRs is weighted by a function of \( \phi(|x|) \) which is monotone decreasing.

3.2. Transform function

Soft information can be quantized with minimizing the weighted mean square quantization error criteria. Assuming the distribution function \( p_L(x) \), in which \( x \) is the LLRs, the distribution is even function and has maximum value \( V \). As stated in last section, in every iteration of decoding, small value of soft information has large weight due to min operation. Instead of minimizing quantization error of soft information, we turn to quantization error of check node LLRs. The quantization error defines as follows [9]:

\[
\sigma_q^2 = \frac{1}{12} \int_{-\infty}^{\infty} \left( \Delta(x)^2 \right) p_R(x)dx \tag{7}
\]

In which \( \Delta(x) \) is the quantization step, the constant is obtained by integration of square error.

The non-uniform quantization applies a non-linear function \( f(x) \) to the original distribution which assumes odd function and have equal maximum value \( f(V) = V \). Then a uniform quantization is applied. Assuming quantization level is \( L \), then quantization step is \( \Delta = 2V / L \), in which \( \Delta / \Delta(x) = f'(x), L \gg 1 \). Combining this equation with (6):

\[
\sigma_q^2 = \frac{\Delta^2}{6} \int_{0}^{V} \left[ f'(x) \right]^2 p_R(x)dx \tag{8}
\]

The optimization problem is stated as follows:

\[
\min_{f(x)} \int_{0}^{V} \left[ f'(x) \right] ^2 p_R(x)dx \text{s.t. } f(V) = V \tag{9}
\]
By applying Lagrangian multiplier method, we attain the optimization function.

\[ f(x) = K \int_0^1 \left[ p_x(x)^{1/3} \right] dx \]  \hspace{1cm} (10)

K is normalization factor that satisfies \( f(V) = V \). Combined with the distribution of check node LLRs, the transform function is expressed as:

\[ f(x) = K \int_0^1 \left[ \rho(\phi(x)) p_x(x)^{1/3} \right] dx \]  \hspace{1cm} (11)

Theoretically, for every combination of modulation type and code rate, if the transform function is obtained, the soft information can be transformed and quantized as new soft information \( C_j' \) and processed in decoding algorithm. However, in hardware implementation, we cannot obtain the analytical expression of transform function. Instead, it can be approximated by statistical histogram of soft information.

In the process of min-sum algorithm, the relative values rather than absolute values of LLRs have impact on performance of LDPC codes. The distribution histogram of soft information is scaled to a maximum value \( V = 1 \). When LDPC codes is chosen, the degree distribution \( \rho(x) \) can be computed from the parity check matrix of LDPC codes. From the equation above, the discrete version of transform function \( f(x) \) can be computed and stored as a look-up table. Because of the uniform quantization afterwards, only the integer part of values are stored in memory of hardware.

4. Simulation Result

Simulation is conducted in 5G LDPC code because of its good performance in all code rates and flexible coding scheme. All the performance is measured in layered structure normalized min-sum algorithm with factor \( \alpha = 0.75 \) and iteration of 20 over AWGN channel.

We choose QPSK, 8PSK, 16APSK, 32APSK, 64APSK as modulation type and (4320, 2880) rate 2/3 LDPC code respectively. As comparison, a 6-bit quantization is adopted which means \( L = 64 \). The soft information is quantized into the interval \(-32 \leq C_j \leq 31\).

The degree distribution function of the chosen LDPC code [10] is

\[ \rho(x) = \frac{5}{13} x^9 + \frac{4}{13} x^7 + \frac{3}{13} x^4 + \frac{1}{13} x^3 \]  \hspace{1cm} (12)

Figure 1 shows the error performance of the 5G LDPC code of 16APSK modulation based on 3 methods corresponding to no quantization, uniform quantization and non-uniform quantization. Figure 2 shows the SNR threshold of the (4320,2880) 5G LDPC code of QPSK, 8PSK, 16APSK, 32APSK, 64APSK with BLER = 1e-3.

![Figure 1. Error performance for iterative decoding of the (4320, 2880) 5G LDPC code of 16APSK modulation with no quantization, uniform quantization, non-uniform quantization.](image.png)
Figure 2. The SNR threshold of the (4320, 2880) 5G LDPC code of different modulation with BLER = 1e-3.

When high order modulation type is adopted, the uniform quantization brings large performance loss of about 0.4dB, which is not acceptable in hardware implementation. Generally, the soft information is truncated to a maximum value V' such that performance of decoding can improve. However, this approach does not give a pervasive method for different LDPC codes and modulation type which means V' can only be set by hand. By bringing the non-uniform quantization, the performance loss of 6-bit quantization is approximately 0.1dB, which is moderate and can be used in practical design.

Based on the method given above, we get the transform function $f(x)$ in Figure 3. It has sparse quantization levels in large values and dense levels in small values. For high order modulation type, the curve is steeper which means small value matters more. For every combination of modulation type and code rate, we can get a transform function that do not need to adjust by hand which is useful for 5G LDPC.

Figure 3. The transform function obtained by simulation for (4320, 2880) 5G LDPC codes.

5. Conclusion
In this paper, we proposed a non-uniform quantization algorithm of soft information for LDPC codes, which has a gain of around 0.3dB performance over uniform quantization for 5G's rate 2/3 LDPC codes.
codes with APSK modulation over additive white Gaussian noise (AWGN) channel. Then the non-linear transform function is presented and can be easily implemented by storing it as a look-up table.

This algorithm is not limited to a specific LDPC code and modulation type. Moreover, it can be used for other decoding schemes with a specific modulation type and forward error correction code.

Acknowledgments
This work was supported in part by the National Natural Science Foundation of China under Grant 61790553, in part by Shenzhen Science and Technology Plan Projects under Grant JCYJ20180306170614484, and in part by the Shanghai Municipal Science and Technology Major Project under Grant 2018SHZDZX04.

References
[1] R. G. Gallager. (1963) Low-Density Parity-Check Codes. Cambridge, MA: M.I.T. Press.
[2] 3GPP. (2018). Multiplexing and channel coding. 3rd Generation Partnership Project. Sophia Antipoli, France. TS 38.212, v15.0.0, Release 15.
[3] D. J. C. MacKay. (1999) Good error-correcting codes based on very sparse matrices. In IEEE Trans. Inform. Theory, vol. 45, pp. 399–431.
[4] M. Fossorier, M. Mihaljevic, and H. Imai. (1999) Reduced complexity iterative decoding of low-density parity check codes based on belief propagation. In IEEE Trans. Commun., vol. 47, pp. 673–680.
[5] T. Richardson, A. Shokrollahi, and R. Urbanke. (2001) Design of capacity-approaching irregular low-density parity check codes. In IEEE Trans. Inform. Theory, vol. 47, pp. 619–637.
[6] T. Richardson and R. Urbanke. (2001) The capacity of low-density parity check codes under message-passing decoding. In IEEE Trans. Inform. Theory, vol. 47, pp. 599–618.
[7] J. Chen and M. P. C. Fossorier. (2002) Density evolution for two improved BP-Based decoding algorithms of LDPC codes. In IEEE Communications Letters, vol. 6, no. 5, pp. 208-210.
[8] D. Hui, S. Sandberg, Y. Blankenship, M. Andersson and L. Grosjean. (2018) Channel Coding in 5G New Radio: A Tutorial Overview and Performance Comparison with 4G LTE. In IEEE Vehicular Technology Magazine, vol. 13, no. 4, pp. 60-69.
[9] D. E. Hocevar. (2004) A reduced complexity decoder architecture via layered decoding of LDPC codes. In IEEE Workshop on Signal Processing Systems. SIPS 2004., Austin, TX, 2004, pp. 107-112.
[10] S. Jayasooriya, M. Shirvanimoghaddam, L. Ong, G. Lechner and S. J. Johnson. (2016) A New Density Evolution Approximation for LDPC and Multi-Edge Type LDPC Codes. In IEEE Transactions on Communications, vol. 64, no. 10, pp. 4044-4056.