

A low-floor bit-mapping scheme for LDPC coded BICM for 5G and beyond systems

Mingyang ZHU¹, Ming JIANG^{1,2*}, Chunming ZHAO^{1,2},
Lijie HU³ & Xiaodong XU³

¹National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China;

²Purple Mountain Laboratories, Nanjing 211111, China;

³China Mobile Research Institute, Beijing 100032, China

Received 17 October 2022/Revised 9 December 2022/Accepted 26 December 2022/Published online 19 September 2023

Abstract This paper proposes a low-floor bit-mapping (LFBM) scheme for bit-interleaved coded modulation (BICM) systems to meet more stringent quality of service requirements of 5G and beyond. For high-efficiency transmissions, we consider the 5G low-density parity-check codes with high-order 2^m -quadrature amplitude modulations (QAMs). The proposed LFBM scheme overcomes plenty of trapping sets induced by the bit interleaver, which focuses on the waterfall performance too aggressively. When the high-order QAM is used, the row-column interleaver specified by the 5G standard is such a bit interleaver. The LFBM scheme only optimizes the rule of mapping an m -bit tuple output by the row-column interleaver to a modulation symbol, rather than the entire bit interleaver. Therefore, the optimized bit mapper is actually an m -bit permutation module concatenated with the original bit interleaver employed in the current 5G BICM systems. The simulation results confirm that the LFBM scheme can lower the error floor of the 5G BICM system by approximately two orders of magnitude, while with negligible performance loss in the waterfall region.

Keywords bit-interleaved coded modulation, low-density parity-check codes, error floors, bit mapping, high-order modulation

Citation Zhu M Y, Jiang M, Zhao C M, et al. A low-floor bit-mapping scheme for LDPC coded BICM for 5G and beyond systems. *Sci China Inf Sci*, 2023, 66(10): 202301, <https://doi.org/10.1007/s11432-022-3670-5>

1 Introduction

Low-density parity-check (LDPC) codes [1] have been extensively investigated due to their capacity-approaching performance and high-throughput implementation. In the past twenty years, protograph-based LDPC codes [2, 3] have attracted considerable attentions due to their structured construction. Based on a small protograph, the lifting [4] operation can give a larger Tanner graph, which entirely defines an LDPC code. Moreover, protograph-based raptor-like (PBRL) LDPC codes [4] have been adopted into the 5G new radio (NR) standard [5] as the channel coding scheme. As the advantages of PBRL codes, 5G LDPC codes have a rate-compatible structure and support many lifting factors (i.e., various code lengths). Thus, a wide range of blocklengths and coding rates can be achieved by 5G systems, which gives flexible choices for different channel conditions. Further, the 5G LDPC code is one of the most powerful error-correction codes, particularly in the waterfall region. However, when designing 5G LDPC codes or other PBRL LDPC codes, their waterfall performances are the most important evaluation criterion. Therefore, these LDPC codes may not perform very well in the error-floor region, i.e., the high signal-to-noise ratio (SNR) region. Poor error-floor performance will limit the applications of these LDPC codes in many high-reliability applications in 5G and beyond, such as industrial automation systems and vehicle-to-everything (V2X) communications [6–8]. For instance, the reliability requirement for industrial automation systems may be higher than 99.999999% [7]. It is widely recognized that the error floor of LDPC codes under message-passing decoding is mainly caused by trapping sets [9].

* Corresponding author (email: jiang_ming@seu.edu.cn)

Richardson [9] proposed the concept of the trapping set, in which a trapping set is defined as a set of incorrect variable nodes (VNs) and the associated unsatisfied check nodes (CNs). He also presented a semi-analytical methodology to estimate the error floors with the assumption of binary phase-shift keying (BPSK) transmission over the binary-input additive white Gaussian noise (BIAWGN) channel.

There is extensive literature on the error floors of LDPC codes with the assumption of the BIAWGN channel [9–15]. However, bit-interleaved coded modulation (BICM) is a more pragmatic approach to achieve high spectrum efficiency by high-order modulations [16, 17]. Our previous work corroborated that the error-floor problem is also significant for LDPC coded BICM systems [18]. In a BICM system, code bits usually need to be interleaved before modulation. A specially designed interleaver can utilize the unequal mutual information (MI) [19] of the label positions and the code structure to further improve the error-rate performance [20]. In most existing literature [20–22], the decoding threshold is of prime importance in the design of a bit interleaver. However, we observed that the high-order modulation with a threshold-optimized interleaving scheme usually raises the error floor of an LDPC code. For example, the 5G LDPC codes with the largest lifting factor 384 in general have error floors lower than the frame error rate (FER) of 10^{-8} in the BICM systems without bit interleaving (the BPSK can also be included into the BICM systems without bit interleaving), but the bit interleaving scheme specified in the 5G standard will raise the error floor to about the FER of 10^{-5} – 10^{-6} . In this paper, we are interested in these error floors induced by bit interleaving. Generally, evaluating the error floors for long 5G LDPC codes (e.g., more than 10000 bits) with a high-order modulation by the Monte Carlo simulation is difficult. This is one of the reasons why rare literature focuses on the error floors of such long LDPC codes, further, with the BICM, which is more complicated than the BIAWGN channel. Thanks to the multi-thread C++ simulation and the field programmable gate array (FPGA) assisted simulation, we can evaluate the error floor of LDPC coded BICM systems and find numerous abnormal trapping sets that never appear in the case of the BIAWGN channel. It is demonstrated that short cycles play an important role in a trapping set that is dominant in the BIAWGN channel [10, 13, 14]. Nonetheless, these abnormal trapping sets do not contain any short cycles and are inferred to be induced by the interleaving scheme. By analysing the structures of these abnormal trapping sets, we ascertain their common weakness, which is caused by the joint effect of the 5G LDPC structure and the threshold-optimized interleaving scheme.

In order to overcome the aforementioned common weakness, we propose a low-floor bit-mapping (LFBM) scheme for the 5G LDPC codes constructed from the base graph 1 (BG1). The 5G LDPC codes constructed from the base graph 2 (BG2) have many poor structures [23] so that their error floors are relatively high, particularly for long codes. In general, the long BG2 LDPC codes have error floors at the FER of 10^{-4} – 10^{-5} [24], which is unsatisfactory for ultra-reliable communications. Also, since this paper focuses on the error floors induced by interleaving schemes rather than the code itself, only the BG1 LDPC codes are considered. Moreover, the BG1 supports longer and higher-rate LDPC codes, which are more widely used in 5G communication systems (see Subsections 6.2.2 and 7.2.2 in [5]). The fundamental principles and the benefits of the proposed LFBM scheme are briefly summarized as follows.

- We find that the abnormal trapping sets induced by the interleaving are essentially caused by mapping all the core parity-check bits (corresponding to the double-diagonal part of the parity-check matrix) to less reliable signal label positions. The proposed LFBM scheme preferentially protects the core parity check bits by more reliable signal label positions and then optimizes the interleaving pattern for the remaining code bits based on the decoding threshold.
- The row-column interleaver specified by the 5G standard [5] usually raises the error floor. The proposed LFBM scheme can be attached to this row-column interleaver to lower the error floor. In other words, the LFBM scheme only changes the rule of mapping an m -bit tuple (output by the row-column interleaver) to a modulation symbol, which can be seen as a permutation for m bits. Therefore, the LFBM scheme can be easily implemented for 5G BICM systems to achieve enhanced performance.
- The proposed LFBM scheme in general lowers the error floor of a 5G LDPC coded BICM system by two orders of magnitude, and the performance loss in the waterfall region does not exceed 0.3 dB. (In most cases, the waterfall-performance loss is less than 0.1 dB.)

The rest of this paper is organized as follows. Section 2 reviews the protograph-based LDPC codes, the BICM, and the extrinsic information transfer analysis. Section 3 analyses the so-called abnormal trapping sets and proposes an LFBM scheme to overcome these trapping sets. Additionally, the section discusses the extra complexity caused by the LFBM. Section 4 exhibits some simulation results, including the decoding thresholds and the improvement of the error floor. Section 5 concludes the paper.

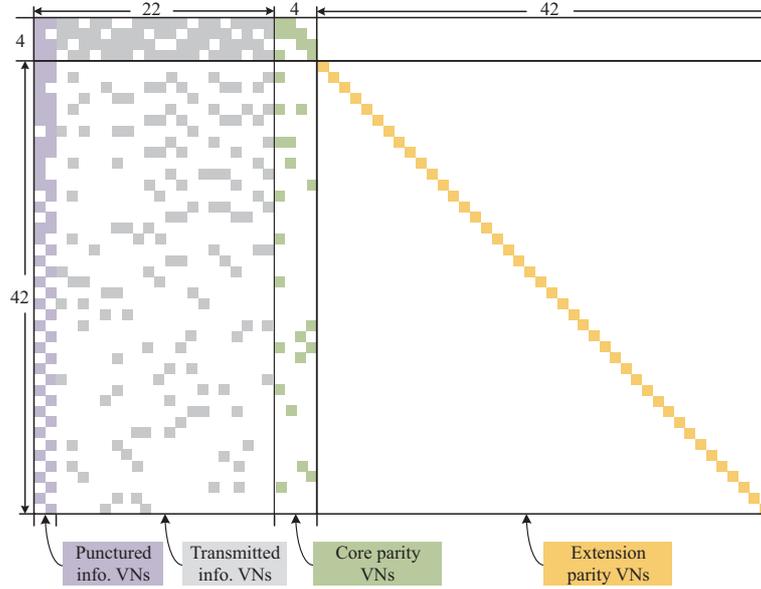


Figure 1 (Color online) The base matrix of the 5G LDPC codes constructed from the BG1.

2 Preliminaries

2.1 Protograph-based LDPC codes and 5G LDPC codes

A protograph-based LDPC code [2–4] is constructed from a small bipartite graph that connects a set of J CNs $\{C^0, C^1, \dots, C^{J-1}\}$ to a set of K VNs $\{V^0, V^1, \dots, V^{K-1}\}$. The protograph can be represented by its base bi-adjacency matrix \mathbf{B} . In a base matrix, each row or column corresponds to a CN or VN in the protograph, respectively. Then, a larger graph (the Tanner graph of an LDPC code) can be obtained by the lifting operation. After lifting, an LDPC code is completely defined by the Tanner graph $\mathcal{G} = (\mathcal{V}, \mathcal{C}, \mathcal{E})$, where \mathcal{V} , \mathcal{C} , and \mathcal{E} are the VN, CN, and edge sets, respectively. When the lifting factor is Z , each entry in a base matrix corresponds to a $Z \times Z$ circulant permutation matrix. Thus, a base matrix can be seen as a block matrix and then its rows and columns are also called block rows and block columns, respectively. 5G LDPC codes are constructed from two protographs (BG1 and BG2), which have rate-compatible structures.

A 5G LDPC code is also a type of PBRL LDPC code. For example, the base matrix \mathbf{B} of the BG1 is shown in Figure 1. 5G LDPC codes have two punctured VNs corresponding to two leftmost columns of the base matrix. In the protograph of 5G LDPC codes, the VNs can be divided into three types: the information VNs, the core parity VNs, and the extension parity VNs [25]. In Figure 1, the information VNs are further separated into the punctured and transmitted parts. Given the number of transmitted bits n , the number of information bits k , and the selection of the base graph B (1 or 2), an (n, k, B) 5G LDPC code can be obtained from the lowest-rate mother code by puncturing [5]. Note that puncturing the degree-1 VNs is equivalent to directly removing the related columns and rows from the parity-check matrix. To simplify the presentation, we only consider the 5G LDPC codes without shortening, which is commonly used for rate matching. And, the chosen codes only puncture the leftmost two block columns and some degree-1 block columns (i.e., n is divisible by Z). By the way, all the proposed techniques and provided results can be easily generalized to the 5G LDPC codes with arbitrary puncturing and shortening. With the above assumptions, an (n, k, B) 5G LDPC code can be alternatively specified by (R, Z, B) , where $R = \frac{k}{n}$ is the code rate and Z is the lifting factor. To clearly show the code rate, we will say an (R, Z, B) 5G LDPC code in the following.

2.2 Bit interleaved coded modulation in the 5G-NR systems

After encoding and puncturing, the transmitted code bits $\mathbf{v} = (v_0, v_1, \dots, v_{n-1})$ are sent to an bit interleaver. When the high-order quadrature amplitude modulation (QAM) is applied for the 5G-NR systems, a row-column interleaver will be attached. In the 5G standard [5], a row-write column-read interleaver with m rows is matched to the 2^m -QAM, which is denoted by the mode- m row-column

interleaver in this paper. Assume that the interleaving pattern is $\boldsymbol{\pi} = \{\pi_0, \pi_1, \dots, \pi_{n-1}\}$, thus the interleaved sequence is $\boldsymbol{v}' = \{v'_0, v'_1, \dots, v'_{n-1}\}$, where $v'_i = v_{\pi_i}$. Each m bits in one column of the interleaver will be mapped to a modulation symbol $\boldsymbol{x} = \{x^0, x^1, \dots, x^{m-1}\}$ in the constellation \mathcal{X} , where $\{x^i\}_{i=0}^{m-1}$ are m label positions or bit-levels of a modulation symbol. For a Gray-mapping QAM, we assume that

$$I_B(x^0) = I_B(x^1) \geq I_B(x^2) = I_B(x^3) \geq \dots \geq I_B(x^{m-2}) = I_B(x^{m-1}), \quad (1)$$

where $I_B(x^i) = I(x^i; Y)$ is the MI between the bit-level x^i and the underlying random variable Y of the received signal, which is also called the bit-level MI (BMI) of the bit-level x^i in this paper. Actually, the bit-levels of the QAM constellations defined in the 5G standard [26] satisfy the inequation (1).

In general, the interleaved bits are divided into m bits each and then directly mapped to a modulation symbol. For instance, one can arbitrarily choose that the most significant bit of a modulation symbol corresponds to the leftmost or rightmost position of the m bits $(b_0, b_1, \dots, b_{m-1})$ (e.g., $b_i \rightarrow x^i$ for $0 \leq i < m$), which is called natural mapping in this paper. While, we can also add a bit mapping function $\Phi(\cdot)$ after bit interleaving, which maps m bits $(b_0, b_1, \dots, b_{m-1})$ to a modulation symbol according to a specially designed rule such as $b_j \rightarrow x^i$. Note that $\Phi(\cdot)$ is equivalent to performing an extra interleaving for the m -tuples before the natural mapping. The interleaving pattern of the bit mapper is named the bit-mapping pattern. In most literature, the bit mapper is ignored because its role can be entirely replaced by an equivalent bit interleaver with the natural mapping. Assume that the bit-mapping pattern is $\boldsymbol{\beta} = \{\beta_0, \beta_1, \dots, \beta_{m-1}\}$, which indicates that $b_{\beta_i} \rightarrow x^i$ for m bits $(b_0, b_1, \dots, b_{m-1})$. The combination of the bit interleaver and the bit mapper gives an equivalent bit interleaver with the interleaving pattern $\boldsymbol{\Pi} = \{\Pi_0, \Pi_1, \dots, \Pi_{n-1}\}$, where $\Pi_i = \pi_{m\lfloor i/m \rfloor + \beta_{i \bmod m}}$. $\boldsymbol{\Pi}$ is referred to as the underlying interleaving pattern. The two equivalent transmitter models are shown in Figure 2.

In this paper, we separate the bit interleaver and the bit mapper into two independent modules. In the later section, we only optimize the bit mapper to improve error-floor performance and the row-column interleaver defined in the 5G standard is still kept, which only brings negligible extra hardware complexity. For example, the bit interleaving for a 5G BICM system employing the 256-QAM is shown in Figure 3. Two bit mapping schemes are included in this figure: the often-used natural mapping and the proposed designed mapping. Since no bit-mapping schemes are specifically defined in the 5G standard, the natural mapping will be applied in general for 5G BICM systems. Note that when $n \bmod m \neq 0$, $(m \lfloor n/m \rfloor - n)$ zero bits need to be filled. However, this zero-filling operation has no essential influence, so we can assume $n \bmod m = 0$ for simplicity.

The decoding threshold is usually the key quantity to be optimized in designing an interleaving scheme, such as the variable degree matched mapping (VDMM) [20] and its generalized versions [21, 22]. Due to the protograph structure (especially the column weight distribution) of 5G LDPC codes, the mode- m row-column interleaver defined in the 5G standard naturally becomes a VDMM-like interleaving scheme. Therefore, the mode- m row-column interleaver usually further improves the waterfall performance when high-order modulations are applied. However, these threshold-optimized interleaving schemes may incur high error floors, since the trade-off between the waterfall and error-floor performances is a well-known fact. For comparison, we also consider the transmission scheme without bit interleaving¹⁾, denoted by the mode-1 interleaving scheme (using a 1-row interleaver). Figure 4 is used as an example to show the gain of the waterfall performance and the degradation of the error-floor performance caused by the mode- m interleaving. In Figure 4, we can clearly see that for the 64- and 256-QAM, the mode- m interleaving gives remarkable gain in the waterfall region but induces much higher error floors. While, for the relatively low-order 16-QAM, both the gain of the waterfall performance and the degradation of the error-floor performance are not significant. This can be easily understand, because a higher-order modulation has more diversity (shown as (1)) for the bit-levels, which implies that the interleaving can more significantly affect the performance.

2.3 Protograph-based extrinsic information transfer analysis

The extrinsic information transfer analysis for protograph-based LDPC codes (PEXIT) was proposed in [27] to analyze the decoding threshold, and this method has been extended to design the LDPC codes for coded modulation systems [22]. In [28], a more detailed PEXIT method for the BICM system with the bit-metric decoding was presented, in which the hypothesis of surrogate bit-channels is used. The PEXIT

1) For LDPC codes, no bit interleaving can also be seen as a random interleaving scheme.

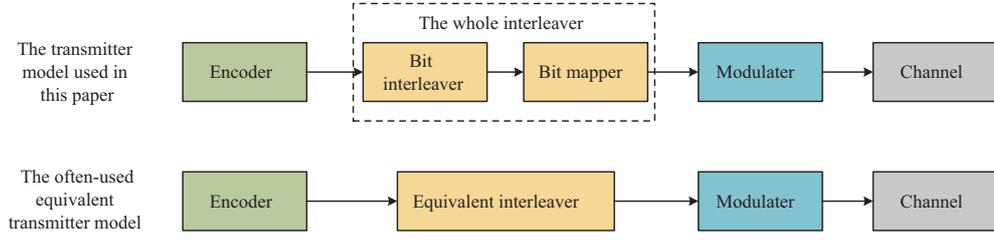


Figure 2 (Color online) Two equivalent models of the transmitter in a BICM system.

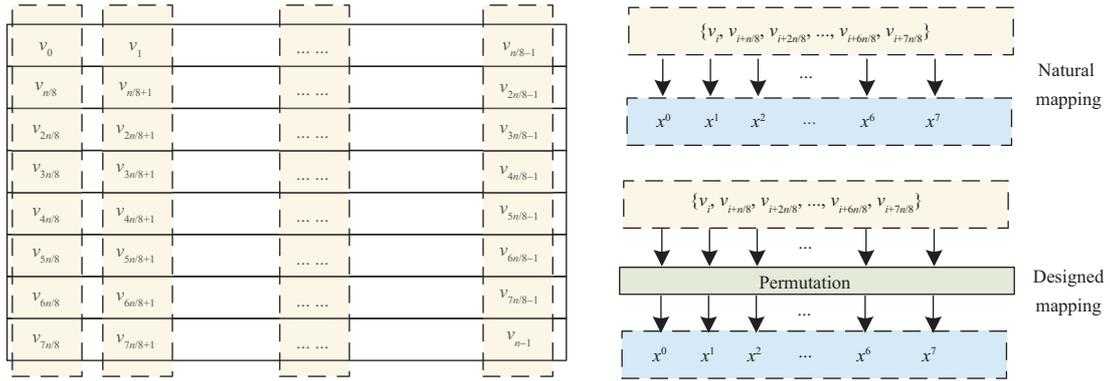


Figure 3 (Color online) The mode-8 row-column interleaver and the bit mappers.

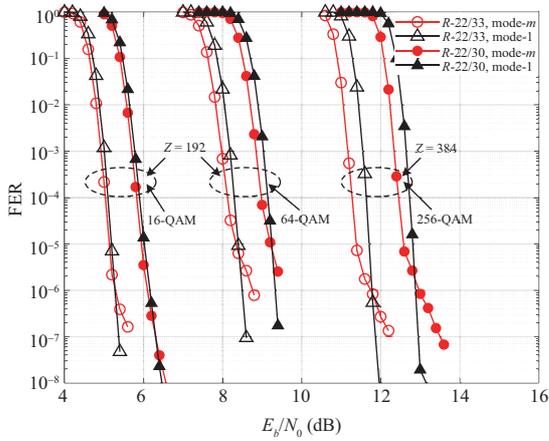


Figure 4 (Color online) FER comparison for the mode-1 and mode- m interleaving schemes in the 16-, 64-, and 256-QAM systems, where the rate-22/30 and 22/33 5G LDPC codes are used.

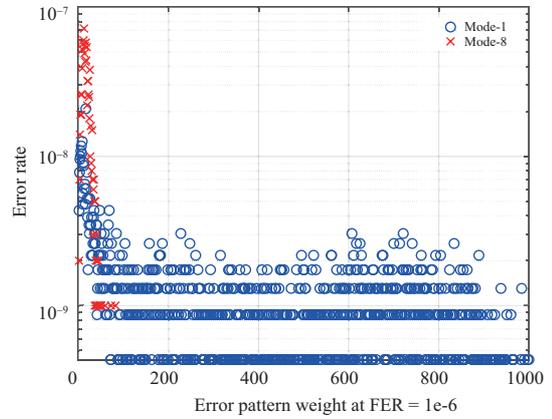


Figure 5 (Color online) Comparison of the weights of the error patterns for the mode-1 and the mode-8 interleaving schemes. The error patterns are collected at FER = 10^{-6} and the (22/33, 384, 1) 5G LDPC coded 256-QAM is used.

method used here is similar to that used in [22]. First, let $J(\sigma)$ denote the MI between a binary random variable X , with $\Pr\{X = \frac{\sigma^2}{2}\} = \Pr\{X = -\frac{\sigma^2}{2}\} = \frac{1}{2}$, and a Gaussian random variable $Y \sim \mathcal{N}(X, \sigma^2)$. For a BIAWGN channel, the log-likelihood ratios (LLRs) can be modeled as independent Gaussian random variables $L \sim \mathcal{N}(\pm \frac{\sigma^2}{2}, \sigma^2)$. Therefore, $J(\sigma)$ represents the capacity of a BIAWGN channel with the parameter σ , and it is given by [29]

$$J(\sigma) = 1 - \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-\sigma^2/2)^2}{2\sigma^2}} \cdot \log_2(1 + e^{-y}) dy. \quad (2)$$

To construct m surrogate bit-channels, we should guarantee that the capacity of each bit-channel equals the BMI of the corresponding bit-level. Therefore, the surrogate BIAWGN bit-channels can be constructed with

$$J(\sigma_L^i) = I_B(x^i), \quad 0 \leq i < m, \quad (3)$$

where σ_L^i is the equivalent variance of the LLRs, $L \sim \mathcal{N}(\pm \frac{\sigma_L^i}{2}, (\sigma_L^i)^2)$, at the bit-level i .

Specifically, a VN V^j of a protograph may be mapped to different bit-levels, so we calculate the ratios λ_{ij} for each V^j , where λ_{ij} denotes the ratio of the bits belonging to V^j mapped to bit-level i . We initialize the MI I_{ch}^j between the channel observation (or LLR) and the code bit for each VN V^j as

$$I_{\text{ch}}^j = \sum_{i=0}^{m-1} \lambda_{ij} J(\sigma_L^i), \quad 0 \leq j < K. \quad (4)$$

The remaining steps are the same as those in [22]. Then, we simply use the function $\Psi(\mathbf{\Pi})$ to denote the PEXIT process for calculating the decoding threshold \mathcal{T} , where $\mathbf{\Pi}$ is the underlying interleaving pattern (i.e., combination of the interleaving patterns of the bit interleaver and mapper). The decoding threshold of an LDPC coded BICM system with a specific interleaving pattern $\Psi(\mathbf{\Pi})$ is computed as $\mathcal{T} = \Psi(\mathbf{\Pi})$.

2.4 Simulation settings

Since the non-converged iterative decoding will produce decoding failures that are not trapping sets (they can be corrected if we continue the iterative decoding), we should guarantee that the maximum number of iterations is enough. Thus, all the simulation results about error floors are done by a 6-bit quantized layered min-sum algorithm (MSA) [30] decoder with a maximum number of iterations 30. The scaling factor is set to 0.75 and the saturation ratio of the quantized values is adaptively adjusted for different codes and channels. In addition, a floating-point sum-product algorithm (SPA) decoder is used to verify the decoding thresholds of various interleaving schemes. This is because the PEXIT method is based on the SPA.

3 Lowering the error floor of 5G LDPC coded BICM systems

3.1 Error floors induced by interleaving: a class of abnormal trapping sets

Recall that an LDPC code is defined by its Tanner graph $\mathcal{G} = (\mathcal{V}, \mathcal{C}, \mathcal{E})$, where \mathcal{V} , \mathcal{C} , and \mathcal{E} are the VN, CN, and edge sets, respectively. For a subset $\mathcal{T} \subset \mathcal{V}$, let $\Gamma(\mathcal{T})$ denote the set of neighbors of \mathcal{T} in \mathcal{C} . Using \mathcal{T} and $\Gamma(\mathcal{T})$, the included subgraph of \mathcal{T} , denoted by $\mathcal{G}(\mathcal{T})$, is defined as $\mathcal{G}(\mathcal{T}) = (\mathcal{T}, \Gamma(\mathcal{T}), \mathcal{E}')$, where $\mathcal{E}' \subset \mathcal{E}$ is the subset containing edges connecting \mathcal{T} and $\Gamma(\mathcal{T})$. The subset of $\Gamma(\mathcal{T})$ with odd degree is denoted by $\Gamma_o(\mathcal{T})$. Likewise, the subset of $\Gamma(\mathcal{T})$ with even degree is denoted by $\Gamma_e(\mathcal{T})$. Then, several concepts about trapping sets is defined as follows.

Definition 1. \mathcal{T} is an (a, b) trapping set if $|\mathcal{T}| = a$ and $|\Gamma_o(\mathcal{T})| = b$. a is the size of the trapping set, and b is usually referred to as the number of unsatisfied CNs.

Definition 2. An (a, b) trapping set is called an elementary trapping set if all the CNs in $\Gamma(\mathcal{T})$ have degree of 1 or 2.

Definition 3. An (a, b) trapping set is called an absorbing set if all the VNs in \mathcal{T} are connected to more CNs in $\Gamma_e(\mathcal{T})$ than to CNs in $\Gamma_o(\mathcal{T})$.

Extensive literature shows that, for an LDPC coded BPSK or quadrature phase shift keying (QPSK) system, elementary trapping sets and absorbing sets play an important role in the error-floor performance [10–12]. In fact, in an overwhelmingly large number of cases, the dominant trapping sets satisfy both the definitions of elementary trapping and absorbing sets. Moreover, it has been demonstrated that, for the BPSK and QPSK cases, most dominant trapping sets contain at least one short cycle and these trapping sets can be found by some search algorithms starting from a short cycle [10, 13, 14]. However, a quite different phenomenon is observed for the high-order modulation system. We observed that the high-order modulation system with a threshold-optimized²⁾ interleaving scheme suffers from high error floors and most error patterns are not dominant trapping sets in the BPSK or QPSK system. These special trapping sets, which do not contain any short cycles, are referred to as abnormal trapping sets. We find that abnormal trapping sets will appear when applying the mode- m or other threshold-optimized interleaving scheme. On the contrary, when mode-1 or random interleaving is applied, the

²⁾ We find that the mode- m row-column interleaving adopted in the 5G standard is a near threshold-optimized interleaving scheme for 5G LDPC codes expanded from the BG1.

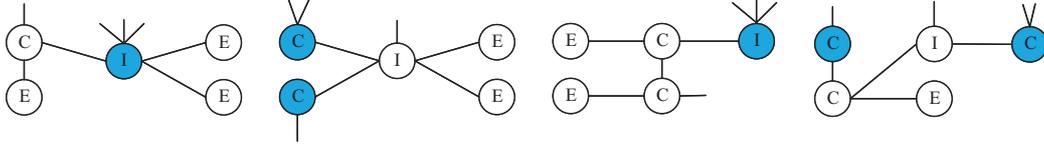


Figure 6 (Color online) Quasi-normal hypergraphs of $(5, 4)$ abnormal trapping sets collected for the $(22/33, 384, 1)$ 5G LDPC code with 256-QAM and mode-8 interleaving. (I, C, and E in the circle represent the information, core parity, and extension parity VNs, respectively. The VNs that do not satisfy the definition of absorbing sets are marked in blue.)

dominant trapping sets become similar to those that appear in the BPSK system, which include several short cycles.

We first give a high-level comparison, for mode-1 and mode- m , in terms of the weight of error patterns. In Figure 4, we can see that the $(22/33, 384, 1)$ 5G LDPC code has an error floor below the FER of 10^{-5} when 256-QAM and mode-8 interleaving are applied, while mode-1 does not have an obvious error floor until the FER of 10^{-8} . Thus, we collected 1000 error patterns for each interleaving mode at the FER of 10^{-6} to compare the weight of the error patterns occurred in the two interleaving modes. In Figure 5, all error patterns are classified according to their weights and the error rate of each weight is calculated by $\frac{W_i}{1000} \cdot 10^{-6}$, where W_i is the number of the occurrences of the weight- i error patterns within 1000 collected error patterns. We can see that there is a significant difference in the error patterns between the two types of interleaving. The mode-8 interleaving has a huge number of low-weight errors, which are in general harmful to the error-floor performance [31]. In contrast, the weight distribution of the error patterns in mode-1 is more random. Therefore, we can state that the high error floor of mode-8 is induced³⁾ by the weakness of the interleaving.

Further, we proceed to explore the characteristics of the trapping sets induced by the threshold-optimized interleaving. Among the collected 1000 error patterns for mode-8, more than 800 non-isomorphic trapping sets are found. Most of these trapping sets are abnormal trapping sets, which do not contain any cycles and are not absorbing sets. The quasi-normal hypergraphs [14] of four $(5, 4)$ abnormal trapping sets collected in the simulation are shown in Figure 6, where each circle represents a VN and each degree-one or two CN is replaced by a single edge. It is easy to verify every abnormal trapping set in the figure is not an absorbing set since at least one VN is connected to equal or more CNs in $\Gamma_o(\mathcal{T})$ than to CNs in $\Gamma_e(\mathcal{T})$.

Since more than 800 non-isomorphic trapping sets are found in a total of 1000 decoding failures (which implies that the number of dominant non-isomorphism trapping sets is prohibitively large), conquering a few specific trapping sets can hardly improve the error-floor performance. Thus, we need to find out the weakness that is owned in common by these trapping sets. Since a trapping set of a quasi-cyclic LDPC (QC-LDPC) code has at least $Z - 1$ isomorphisms, using the block columns and rows can simplify the representation of such a trapping set. To do so, we count the occurrences of each block column within the collected 1000 error patterns, as shown in Figure 7. One can see that the 23–26, 28, 31, 34-th block columns have a significantly higher probability to be in error, where the 23–26-th block columns are the core parity VNs and the other three block columns are the extension parity VNs. Further, one can find that the three extension parity VNs are connected to the four core parity VNs in the Tanner graph. The Tanner graph involving these VNs is shown in Figure 8. It implies that the errors located at the extension parity VNs are probably propagated from the core parity VNs. On the other hand, the core parity VNs have larger degrees than the extension ones, thus correcting the core parity VNs preferentially is more profitable than correcting the extension parity VNs (which can propagate more correct messages and restrain more error propagations). So far, we have found that the common weakness of all trapping sets induced by the threshold-optimized interleaving is the unreliable core parity VNs. Therefore, we propose a LFBM scheme in the next subsection, which protects the core parity VNs first and then optimizes the decoding threshold.

3.2 An LFBM scheme

We propose an LFBM scheme for 5G LDPC coded BICM with 2^m -QAM to improve the error-floor performance while keeping good waterfall performance. In the proposed LFBM scheme, the mode- m row-column interleaver is still applied, which means that we only optimize β and π is fixed. We do not jointly

³⁾ Here we only say “induced” rather than “caused”, since the causes of the error floor are more complicated, which are related to the code structure, the channel, and the interleaving scheme.

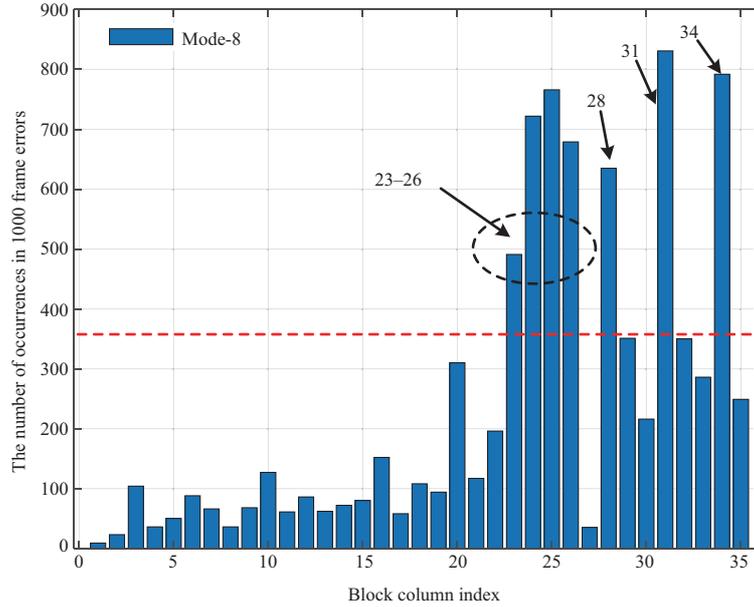


Figure 7 (Color online) The number of the occurrences of every block column within the 1000 error patterns.

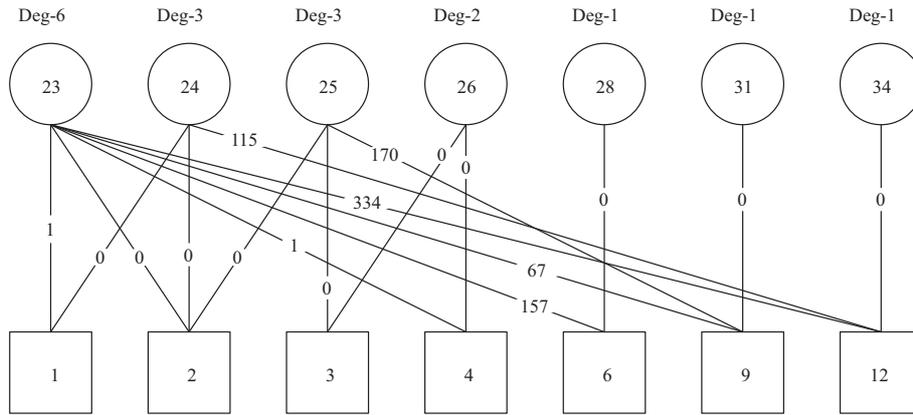


Figure 8 The Tanner graph involving the 23–26, 28, 31, 34-th block columns. The number at the edge is the cyclic shift value.

optimize the bit interleaver and the bit mapper due to the following two reasons. First, we observed that the mode- m row-column interleaver has already given a good waterfall performance. Second, considering the ease of implementation, adding an optimized mapper only brings negligible additional hardware complexity. Therefore, the proposed LFBM scheme can be easily and flexibly implemented to the existing 5G BICM systems to achieve enhanced performance.

Consider the 2^m -QAM with the mode- m row-column interleaver. Then, all the transmitted VNs in the protograph⁴⁾ (excluding the punctured and shortened VNs) are divided into m VN groups. Let \mathcal{S}_v and \mathbb{V} denote the transmitted-VN set and the VN group, respectively. Therefore, we have $\mathcal{S}_v = \{\mathbb{V}_0, \mathbb{V}_1, \dots, \mathbb{V}_{m-1}\}$ for the mode- m row-column interleaver. For a practical implementation, the code bits corresponding to \mathbb{V}_i , $0 \leq i < m$, are filled into the i -th row (also labeled from 0 to $m - 1$) of the interleaver. Generally, \mathbb{V}_0 is on the information side and \mathbb{V}_{m-1} is on the extension parity side.

First, we represent the operation of mapping a VN group to a bit-level as an event, denoted by $\mathbb{V}_j \rightarrow x^i$, $0 \leq i, j < m$. After the mode- m row-column interleaving, all VNs have been divided into m VN groups.

4) In the following, we describe the proposed LFBM scheme from the perspective of a protograph. In other words, puncturing and shortening operations should be taken in the protograph. However, in some special scenarios, a 5G LDPC code may be punctured and shortened only with a part of bits corresponding to a VN in the protograph. For these scenarios, we can change the concept of “VNs in the protograph” to “VNs in the expanded Tanner graph” (i.e., code bits). Then, the proposed LFBM scheme can also be applied.

With an abuse of notation, we alternatively represent the interleaving pattern $\mathbf{\Pi}$ as

$$\mathbf{\Pi} = \bigcup_{0 \leq i < m} \mathbb{V}_{\beta_i} \rightarrow x^i, \tag{5}$$

where m VN groups \mathbb{V} are determined by the row-column interleaver and $\mathbb{V}_{\beta_i} \rightarrow x^i$ is the operation of the bit mapper with the bit-mapping pattern $\beta = \{\beta_0, \dots, \beta_{m-1}\}$. Although the LFBM will change the underlying interleaving pattern $\mathbf{\Pi}$, the design method presented in this section only affects β .

In the proposed LFBM scheme, \mathcal{S}_v is divided into three subsets \mathcal{S}_c , \mathcal{S}_e , and \mathcal{S}_r as follows. \mathcal{S}_c denotes the set of VN groups, which mainly includes core parity VNs (may also include some information VNs and extension parity VNs). \mathcal{S}_e denotes the set of VN groups, which entirely consists of extension parity VNs. \mathcal{S}_r denotes the remaining subset as $\mathcal{S}_r = \mathcal{S}_v \setminus (\mathcal{S}_c \cup \mathcal{S}_e)$. We use $|\mathcal{S}|$ to represent the size of a set, i.e., the number of entries in the set. Recall that the m bit-levels $\{x^0, x^1, \dots, x^{m-1}\}$ are rearranged in decreasing order in terms of the corresponding BMI. Thus, x^0 is the most protected (i.e., reliable) position and x^{m-1} is the least protected one.

The LFBM scheme can be described as a three-stage bit-mapping process. In the first stage, we map the VN groups $\mathbb{V}_i \in \mathcal{S}_e$ to the bit-levels with the lowest BMIs, denoted by

$$\Phi_1(\mathcal{S}_e) = \bigcup_{\mathbb{V}_i \in \mathcal{S}_e} \mathbb{V}_i \rightarrow x^i. \tag{6}$$

The second stage is to search for the combinations (i.e., possible selections) of the VN groups in the \mathcal{S}_c and map them to some more protected bit-levels. The total number of the combinations is $N_{\text{com}} = \sum_{i=1}^{|\mathcal{S}_c|} \binom{|\mathcal{S}_c|}{i}$. Then, $\{\mathcal{S}_t\}_{t=1}^{N_{\text{com}}}$ is used to denote the set of all possible combinations of VN groups in the \mathcal{S}_c . Note that for 5G LDPC codes, N_{com} is quite small since $|\mathcal{S}_c|$ is usually small. After selecting a combination \mathcal{S}_t , we will map these VN groups in \mathcal{S}_t to the bit-levels $\{x^0, x^2, \dots, x^{2(|\mathcal{S}_t|-1)}\}$, which are chosen from the most reliable position. It is important that we select the bit-levels at intervals⁵⁾ in this stage, since the performance will degrade rapidly at the waterfall region if too many strongly protected bit-levels are selected. If there are more than one VN groups in \mathcal{S}_t (usually $|\mathcal{S}_t| = 1, 2$), we use a simple mapping rule like the VDMM [20], which in general improves the waterfall performance for protograph LDPC codes. In this way, we map the VN group with greater average degree to the bit-levels with higher BMI. Therefore, we can reorder the VN groups in \mathcal{S}_t and give a reordered set \mathcal{S}'_t as follows:

$$\mathcal{S}'_t = \{\mathbb{V}_{\kappa_0}, \mathbb{V}_{\kappa_1}, \dots, \mathbb{V}_{\kappa_{|\mathcal{S}'_t|-1}}\}, \tag{7}$$

where

$$\text{Deg}(\mathbb{V}_{\kappa_0}) \geq \text{Deg}(\mathbb{V}_{\kappa_1}) \geq \dots \geq \text{Deg}(\mathbb{V}_{\kappa_{|\mathcal{S}'_t|-1}}), \tag{8}$$

and $\text{Deg}(\mathbb{V})$ is the average degree (i.e., column weight) of the VN group \mathbb{V} .

Then, we can represent the mapping operation of the second stage as a function $\Phi_2(\cdot)$:

$$\Phi_2(\mathcal{S}'_t) = \bigcup_{0 \leq j < |\mathcal{S}'_t|} \mathbb{V}_{\kappa_j} \rightarrow x^{2j}. \tag{9}$$

There may exist a conflict that the bit-level selected in the second stage has already been chosen in the first stage. If the conflict appears, we can randomly choose another achievable bit-level instead. Fortunately, we numerically verify that this conflict will never occur if the puncturing satisfying the assumption given in Subsection 2.1 is performed. The explanation for this is given in Appendix A.

Since \mathcal{S}_t is a subset of \mathcal{S}_c , the VN groups $\mathbb{V}_i \in \mathcal{S}_c \setminus \mathcal{S}_t$ have not been mapped to some specific bit-levels. Thus, these VN groups should be added to the remaining set \mathcal{S}_r , so we update \mathcal{S}_r as

$$\mathcal{S}_r \leftarrow \mathcal{S}_r \cup (\mathcal{S}_c \setminus \mathcal{S}_t). \tag{10}$$

The final stage is to map the remaining VN groups $\mathbb{V}_i \in \mathcal{S}_r$ to the remaining bit-levels \mathcal{B}_r , where

$$\mathcal{B}_r = \{x^i\}_{i=0}^{m-1} \setminus \left(\{x^{2i}\}_{i=0}^{|\mathcal{S}'_t|-1} \cup \{x^i\}_{i=m-|\mathcal{S}_e|}^{m-1} \right). \tag{11}$$

5) For 16-QAM, since there are only two pairs of equivalent bit-levels $\{x^0, x^1\}$ and $\{x^2, x^3\}$, it is allowed to map the two VN groups (if have two) in \mathcal{S}_t to the two consecutive bit-levels x^0 and x^1 if the error floor is not satisfactory when using the interval mapping rule. The only exception breaking the interval mapping rule in our examples is the (22/44, 192, 1) 5G LDPC code with 16-QAM.

In this stage, the PEXIT method is used to search for the best decoding threshold \mathcal{T} . If the combination \mathcal{S}_t is chosen in the second stage, a total of N_t mapping patterns will be searched in the third stage, where

$$N_t = \frac{(m - |\mathcal{S}_t| - |\mathcal{S}_e|)!}{2^{(m-2|\mathcal{S}_t|-2\lceil|\mathcal{S}_e|/2\rceil)/2}}. \quad (12)$$

It is obvious that the search range is relatively small since $m/2$ real component bit-levels and $m/2$ imaginary component bit-levels in the 2^m -QAM with the Gray labeling are equivalent. Combining the mapping functions Φ_1 and Φ_2 , and a possible mapping $\mathcal{S}_r \rightarrow \mathcal{B}_r$ in this stage, an underlying interleaving pattern $\mathbf{\Pi}'$ is given as

$$\mathbf{\Pi}' = \Phi_1(\mathcal{S}_e) \cup \Phi_2(\mathcal{S}'_t) \cup (\mathcal{S}_r \rightarrow \mathcal{B}_r). \quad (13)$$

By searching over all possible $\mathbf{\Pi}'$, the third-stage bit-mapping function $\Phi_3(\mathcal{S}_r)$ outputs the $\mathbf{\Pi}'$ with the best decoding threshold, thus can be written as

$$\Phi_3(\mathcal{S}_r) = \arg \min_{\mathcal{S}_r \rightarrow \mathcal{B}_r} \Psi(\mathbf{\Pi}'). \quad (14)$$

So far, we know that once a \mathcal{S}_t is chosen from \mathcal{S}_c , the β as well as $\mathbf{\Pi}$ can be completely determined by $\Phi_1(\cdot)$, $\Phi_2(\cdot)$, and $\Phi_3(\cdot)$. Then, the Monte Carlo simulation with the specific β is required to evaluate the performance. If the waterfall or error-floor performance is not satisfactory, we will choose another \mathcal{S}_t and repeat the above process. In the worst case, we need to exhaustively try all possible \mathcal{S}_t and the maximum number of simulations is $\sum_{i=1}^{|\mathcal{S}_c|} \binom{|\mathcal{S}_c|}{i} = 2^{|\mathcal{S}_c|} - 1$. In the next subsection, it is demonstrated that $2^{|\mathcal{S}_c|} - 1$ is in general less than or equal to three. If we directly use the Monte Carlo simulation to optimize the bit-mapping pattern, we require $\frac{m!}{2^{m/2}}$ (the denominator is due to the equivalence of bit-levels) simulations for verifying the error floor, which is infeasible for $m \geq 6$. The LFBM scheme is described in Algorithm 1. In addition, an example of how to optimize the LFBM scheme is shown in Example 1.

Algorithm 1 LFBM scheme

```

1: Initialize the VN-group sets  $\mathcal{S}_e$ ,  $\mathcal{S}_c$ , and  $\mathcal{S}_r$ ;
2: Map the VN groups  $\mathbb{V}_i \in \mathcal{S}_e$  using  $\Phi_1(\mathcal{S}_e)$ ;
3: Set  $t = 1$ ;
4: while  $t \leq N_{\text{com}}$  do
5:   Choose a subset  $\mathcal{S}_t \subseteq \mathcal{S}_c$  and its reordered set  $\mathcal{S}'_t$ ;
6:   Map the VN groups  $\mathbb{V}_i \in \mathcal{S}'_t$  using  $\Phi_2(\mathcal{S}'_t)$ ;
7:   Update the remaining VN-group set  $\mathcal{S}_r \leftarrow \mathcal{S}_r \cup (\mathcal{S}_c \setminus \mathcal{S}_t)$ ;
8:   Map the VN groups  $\mathbb{V}_i \in \mathcal{S}_r$  using  $\Phi_3(\mathcal{S}_r)$ ;
9:   Perform the Monte Carlo simulation with  $\mathbf{\Pi} = \Phi_1(\mathcal{S}_e) \cup \Phi_2(\mathcal{S}'_t) \cup \Phi_3(\mathcal{S}_r)$  to evaluate the performance;
10:  if both the error-floor and waterfall performances are satisfactory then
11:    Break;
12:  else
13:     $t \leftarrow t + 1$ ;
14:  end if
15: end while
    
```

Example 1. In this example, we show the detailed process of the LFBM scheme for the (22/33, 384, 1) 5G LDPC code with 256-QAM, where the mode-8 row-column interleaver is applied. For this code, we have $\mathcal{S}_c = \{\mathbb{V}_4, \mathbb{V}_5\}$, $\mathcal{S}_e = \{\mathbb{V}_6, \mathbb{V}_7\}$, and $|\mathcal{S}_c| = |\mathcal{S}_e| = 2$. In the first bit-mapping stage, $\Phi_1(\mathcal{S}_e) = \{\mathbb{V}_6 \rightarrow x^6\} \cup \{\mathbb{V}_7 \rightarrow x^7\}$. In the second bit-mapping stage, $\{\mathbb{V}_4\}$, $\{\mathbb{V}_5\}$ and $\{\mathbb{V}_4, \mathbb{V}_5\}$ are the only $N_{\text{com}} = 3$ combinations (actually, N_{com} is quite small for 5G LDPC codes). Here we assume that $\mathcal{S}_t = \{\mathbb{V}_4, \mathbb{V}_5\}$ is chosen, then, the reordered set $\mathcal{S}'_t = \mathcal{S}_t$ due to $\text{Deg}(\mathbb{V}_4) = 5.3$ and $\text{Deg}(\mathbb{V}_5) = 2.7$. Then, $\Phi_2(\mathcal{S}'_t) = \{\mathbb{V}_4 \rightarrow x^0\} \cup \{\mathbb{V}_5 \rightarrow x^2\}$. In the third bit-mapping stage, the remaining VN-group and bit-level sets are $\mathcal{S}_r = \{\mathbb{V}_0, \mathbb{V}_1, \mathbb{V}_2, \mathbb{V}_3\}$ and $\mathcal{B}_r = \{x^1, x^3, x^4, x^5\}$, respectively. Then, we use $\Phi_3(\mathcal{S}_r)$ to map these remaining VN groups \mathcal{S}_r to the remaining bit-levels \mathcal{B}_r , leading to the best decoding threshold.

3.3 Complexity discussion

The complexity of the proposed scheme can be divided into two aspects: the complexities of the hardware implementation and bit-mapping optimization process. Note that the bit-mapping optimization presented in the last subsection is done offline and the optimized bit-mapping patterns will be stored in both the transmitter and the receiver. Given a bit-mapping pattern, its implementation is an m -bit

permutation circuit, which takes negligible hardware resources. Thus, the major complexity for the hardware implementation is the storage for all bit-mapping patterns. It is clear that each bit-mapping pattern needs $\lceil \log_2 m \rceil m$ bits to be stored in look up tables. In the 5G standard, a total of 51 lifting factors are defined. Each lifting factor specifies the lowest-rate 5G LDPC code and the higher-rate codes can be realized by puncturing. Assuming that we adopt the puncturing method introduced in Subsection 2.1, then there are 43 different code rates for the BG1. Thus, $51 \times 43 = 2193$ LDPC codes are given and the storage requirement for the bit-mapping patterns is $2193 \lceil \log_2 m \rceil m$ bits for each modulation scheme, which is quite small for today's FPGAs or application specific integrated circuits (ASICs). Moreover, we can use the same bit-mapping pattern for a code rate with any lifting factor to reduce the storage requirement, since the LDPC's performance is closely related to the protograph. This implies that the storage requirement can be reduced by a factor of 51 (the number of Z). In this way, the storage requirement is only $43 \lceil \log_2 m \rceil m$ bits for each modulation scheme.

Then, let us discuss the complexity of the optimization process specified by Algorithm 1. Recall that this complexity will not reflect into the implementation. In Algorithm 1, the complexity is mainly determined by two parts: the PEXIT search in the third stage (line 8) and the Monte Carlo simulation to verify the performance (line 9).

The complexity of the PEXIT search is determined by the number of possible bit-mapping patterns searched by (13), since we need to exhaustively compute the decoding thresholds of these bit-mapping patterns. Since there are equivalent bit-levels in QAMs, the number N_β of distinct bit-mapping patterns (which can lead to different decoding thresholds) can be further reduced. (e.g., $\mathbb{V} \rightarrow x^0$ and $\mathbb{V} \rightarrow x^1$ are equivalent) By summing over all possible \mathcal{S}_t , we have

$$N_\beta = \sum_{t=1}^{N_{\text{com}}} \frac{(m - |\mathcal{S}_t| - |\mathcal{S}_e|)!}{2^{(m-2|\mathcal{S}_t|-2\lceil |\mathcal{S}_e|/2 \rceil)/2}} = \sum_{i=1}^{|\mathcal{S}_c|} \binom{|\mathcal{S}_c|}{i} \cdot \frac{(m - i - |\mathcal{S}_e|)!}{2^{(m-2i-2\lceil |\mathcal{S}_e|/2 \rceil)/2}}. \quad (15)$$

Note that $|\mathcal{S}_c|$ and $|\mathcal{S}_e|$ are given, respectively, by (A1) and (A2), and n_i , n_c , n_e and v are defined in Appendix A. Since $n_i - v \lfloor n_i/v \rfloor < v$, we can write that

$$|\mathcal{S}_c| \leq \lceil (v + n_c)/v \rceil = 1 + \lceil n_c/v \rceil, \quad (16)$$

where \leq is due to the `ceil` function. It is easy to find $v \geq (n_i + n_c)/m$, hence, $|\mathcal{S}_c|$ is upper bounded by

$$|\mathcal{S}_c| \leq 1 + \lceil n_c m / (n_i + n_c) \rceil. \quad (17)$$

In (17), n_i and n_c are two constants which are determined by the base graph, and m is influenced by the modulation scheme. For the BG1, we have $n_i = 20$ and $n_c = 4$. Using (17), one can verify that $|\mathcal{S}_c| \leq 1 + \lceil \frac{m}{6} \rceil$. In practice, we only need to consider the cases where m is an even number. Thus, we know that $|\mathcal{S}_c| \in \{1, 2\}$ for $m \leq 6$ and $|\mathcal{S}_c| \in \{1, 2, 3\}$ for $8 \leq m \leq 12$. Consequently, the value of N_β in (15) will be relatively small since it is the sum over all possible combinations in \mathcal{S}_c . For example, when we search for the best decoding threshold for the (22/33, 384, 1) code with 256-QAM, we have $|\mathcal{S}_c| = 2$, so at most $N_\beta = 72$ bit-mapping patterns need to be considered in the PEXIT stage in Algorithm 1.

Moreover, the complexity of the Monte Carlo simulation needs to be evaluated. Although multi-thread programming or FPGA-assisted simulations can speed up the Monte Carlo simulation, we still hope that the times to perform the simulation is very limited. We know that the worst case requires $\sum_{i=1}^{|\mathcal{S}_c|} \binom{|\mathcal{S}_c|}{i} = 2^{|\mathcal{S}_c|} - 1$ simulations. Fortunately, since $|\mathcal{S}_c|$ is very small, $2^{|\mathcal{S}_c|} - 1$ will also be a small value. In an overwhelmingly large number of cases, we have $|\mathcal{S}_c| \leq 2$, thus the maximum number of Monte Carlo simulations required to evaluate the error floor is only three.

4 Numerical results

In this section, we give some simulation results, including decoding thresholds and FER performance. Recall that the LFBM mapper is always concatenated with a mode- m row-column interleaver. In the following, we will use "mode- m & LFBM" to represent the proposed interleaving scheme, or only say "LFBM scheme" for simplicity.

Table 1 Thresholds calculated by the PEXIT method

Scheme	β (22/33, 384, 1)	\mathcal{T} (22/33, 384, 1) (dB)	β (22/30, 384, 1)	\mathcal{T} (22/30, 384, 1) (dB)
Mode-1	Random	10.685	Random	11.782
Mode-8(5G)	{0,1,2,3,4,5,6,7}	10.512	{0,1,2,3,4,5,6,7}	11.635
E-PEXIT	{0,1,4,3,2,5,6,7}	10.471	{0,1,4,3,2,5,6,7}	11.633
LFBM	{5,2,0,3,4,1,6,7}	10.545	{5,2,6,1,4,3,7,0}	11.733

Table 2 Bit-mapping patterns for the codes used in the simulations

Code rate	β 16-QAM	β 64-QAM	β 256-QAM
22/30	{2,0,1,3}	{4,1,0,3,2,5}	{5,2,6,1,4,3,7,0}
22/33	{2,0,1,3}	{3,1,4,0,2,5}	{5,2,0,3,4,1,6,7}
22/38	{2,0,1,3}	{3,0,1,2,4,5}	{4,3,2,5,0,1,6,7}
22/44	{1,2,0,3}	{3,0,1,2,4,5}	{2,4,0,3,1,5,6,7}

4.1 Comparison of the decoding thresholds

We first present some numerical results of the decoding thresholds of several interleaving schemes, which are computed by the PEXIT method. The (22/33, 384, 1) and (22/30, 384, 1) 5G LDPC codes are used in this subsection, and the 256-QAM with Gray labeling is applied. The bit interleaving schemes used for comparison are listed as follows:

- The mode-1 interleaver, i.e., no bit interleaving, which may have the best error-floor performance while with poor waterfall performance.
- The mode-8 interleaver with a natural mapper, which is used in 5G BICM systems.
- The mode-8 interleaver with the proposed LFBM mapper, which gives a good trade-off between the waterfall and error-floor performances.
- The mode-8 interleaver with a threshold-optimized mapper, which is carried out by the exhaustive search of the bit-mapping patterns to find the threshold-optimized pattern. We use the shorthand notation “E-PEXIT” to denote this bit interleaving scheme.

The decoding thresholds calculated by the PEXIT method are shown in Table 1. Also included in Table 1 are the corresponding interleaving patterns. In the given interleaving patterns in Table 1, the sets \mathcal{S}_t in the LFBM schemes for (22/33, 384, 1) and (22/30, 384, 1) 5G LDPC codes are $\{\mathbb{V}_5\}$ and $\{\mathbb{V}_5, \mathbb{V}_6\}$, respectively. Clearly, the E-PEXIT scheme outperforms the other schemes in terms of the decoding threshold. We can see that the mode-8 scheme used in 5G-NR also has a good decoding threshold close to that of the E-PEXIT scheme. For the proposed LFBM scheme, its decoding threshold is slightly poorer than the mode-8 scheme, but the LFBM scheme significantly improves the error-floor performance. Moreover, the mode-1 scheme without bit interleaving has the poorest decoding threshold in this example.

In addition, we use a floating-point SPA decoder to verify the waterfall performances of the aforementioned four interleaving schemes. This is because the PEXIT method is based on the SPA decoding. The FER performances of (22/33, 384, 1) and (22/30, 384, 1) 5G LDPC codes employing these four interleaving schemes are shown in Figure 9. It can be seen that the LFBM scheme outperforms the mode-1 scheme in the waterfall region for both codes. Both the E-PEXIT and mode-8 schemes incur high error floors although their waterfall performances are better than the LFBM scheme.

4.2 Comparison of the error-floor performance

In most communication systems employing LDPC codes for forward error correction, the MSA is used to simplify the implementation of a hardware decoder. Also, the error-rate performance (especially the error floor) under the MSA decoding is slightly different from the SPA decoding. Therefore, we realized a 6-bit quantized layered MSA decoder using the multi-thread C++ programming to evaluate the FER performances for several 5G LDPC codes with different modulation and interleaving schemes. In our experiments, the (22/30, 384, 1), (22/33, 384, 1), (22/38, 384, 1), and (22/44, 384, 1) 5G LDPC codes with the largest Z are combined with the 256-QAM and the (22/30, 192, 1), (22/33, 192, 1), (22/38, 192, 1), and (22/44, 192, 1) 5G LDPC codes are employed in the 64-QAM and 16-QAM BICM systems. For each combination of a code and a modulation order, the FER performances of the mode- m (5G), the mode-1

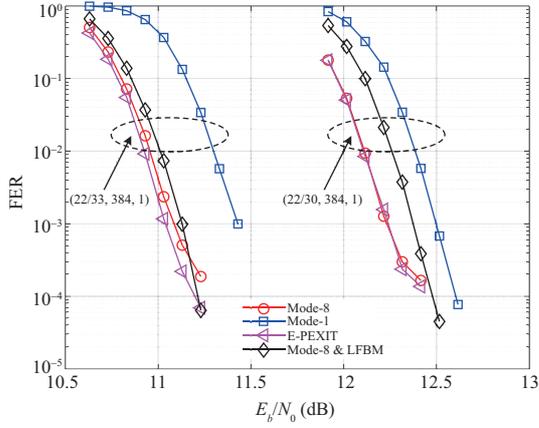


Figure 9 (Color online) FER Performance comparison of the four interleaving schemes applying to the (22/30, 384, 1) and (22/33, 384, 1) 5G LDPC codes.

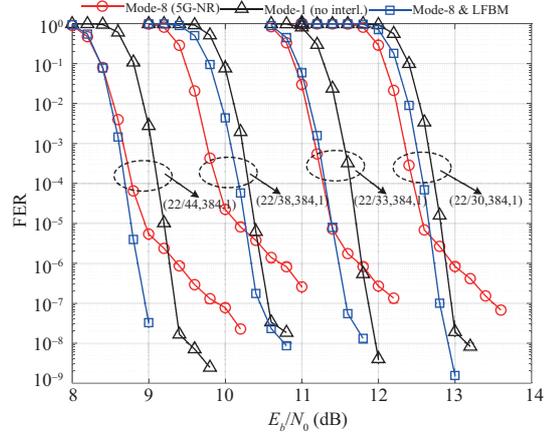


Figure 10 (Color online) FER performances of the (22/30, 384, 1), (22/33, 384, 1), (22/38, 384, 1), and (22/44, 384, 1) 5G LDPC codes with 256-QAM.

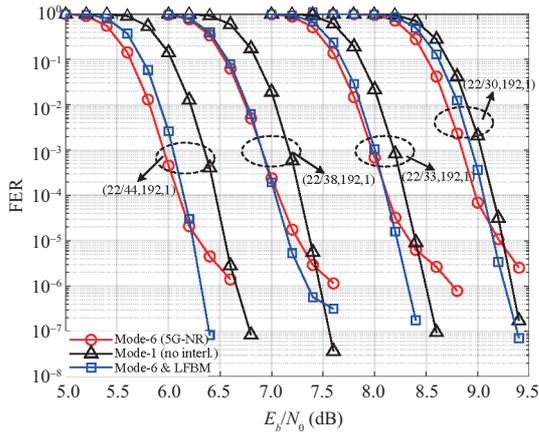


Figure 11 (Color online) FER performances of the (22/30, 192, 1), (22/33, 192, 1), (22/38, 192, 1), and (22/44, 192, 1) 5G LDPC codes with 64-QAM.

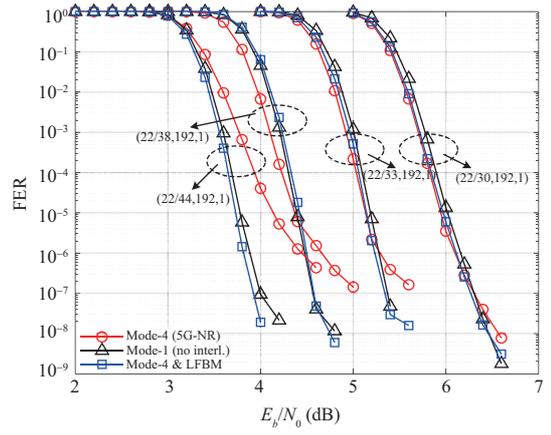


Figure 12 (Color online) FER performances of the (22/30, 192, 1), (22/33, 192, 1), (22/38, 192, 1), and (22/44, 192, 1) 5G LDPC codes with 16-QAM.

(no interleaving), and the mode- m & LFBM schemes are evaluated. All optimized bit-mapping patterns for the LFBM scheme are given in Table 2.

The FER performances for the 256-QAM, the 64-QAM, and the 16-QAM BICM systems are shown in Figures 10–12, respectively. As mentioned in Preliminaries section, for 64-QAM and 256-QAM, the mode- m interleaving significantly improves the waterfall performance over the mode-1 interleaving but raises the error floor. In Figures 10 and 11, we can see that the LFBM scheme lowers the error floor by two orders of magnitude for most cases and the degradation of the waterfall performance does not exceed 0.3 dB. Particularly, for the (22/33, 384, 1) and (22/44, 384, 1) codes used for 256-QAM, the LFBM scheme even has the same waterfall performance as the mode-8. More interestingly, one can also see the trade-off between the waterfall and the error floor in these two figures. For those LFBM schemes having better error-floor performance than the mode-1's (e.g., (22/30, 384, 1) and (22/38, 384, 1)), they cause a relatively obvious degradation of the waterfall performance. In contrast, the LFBM schemes that do not degrade the waterfall performance have, in general, a worse error-floor performance than the mode-1's. For the 16-QAM shown in Figure 12, although the gap between the mode-1 and the mode-4 becomes smaller, our proposed LFBM scheme still lowers the error floor by 1–3 orders of magnitude for all the selected codes. Moreover, the (22/44, 192, 1) has a very high error floor (its FER curve has a reduced slope at $\text{FER} \approx 10^{-1}$), thus it looks like that the mode-4 even performs worse than the mode-1 in the waterfall region. However, by connecting the proposed LFBM with the mode-4, one can see that the

mode-4 & LFBM outperforms the mode-1 and the mode-4 everywhere.

5 Conclusion

In this paper, we propose an LFBM scheme to improve the error-floor performance for high-efficiency LDPC coded BICM in 5G and beyond systems. The proposed scheme only optimizes the bit mapper while it renders the mode- m row-column interleaver unchanged, which leads to an easy implementation on hardware. From the simulation results, we can see that the LFBM scheme further improves the error-floor performance of LDPC coded BICM systems and also gives good waterfall performance. The improvement of the error-floor performance renders the 5G LDPC codes feasible for future ultra-reliable communications. Moreover, the proposed scheme can also be applied to other PBRL LDPC codes with a base matrix similar to that of 5G LDPC codes.

Acknowledgements This work was supported by National Key Research and Development Program of China (Grant No. 2018YFB1801103), Jiangsu Province Basic Research Project (Grant No. BK20192002), and Southeast University China Mobile Research Institute Joint Innovation Center.

References

- Gallager R. Low-density parity-check codes. *IEEE Trans Inform Theory*, 1962, 8: 21–28
- Thorpe J. Low Density Parity Check (LDPC) Codes Constructed from Protographs. JPL INP Progress Report. 2003
- Divsalar D, Dolinar S, Jones C. Low-rate LDPC codes with simple protograph structure. In: *Proceedings of IEEE International Symposium on Information Theory*, 2005. 1622–1626
- Chen T Y, Vakili K, Divsalar D, et al. Protograph-based raptor-like LDPC codes. *IEEE Trans Commun*, 2015, 63: 1522–1532
- 3GPP. The 3rd Generation Partnership Project; Technical specification group radio access network; NR; Multiplexing and channel coding (Release 15). 3GPP TS 38.212 V15.6.0. 2019. https://www.3gpp.org/ftp/Specs/archive/38_series/38.212
- You X H, Wang C X, Huang J, et al. Towards 6G wireless communication networks: vision, enabling technologies, and new paradigm shifts. *Sci China Inf Sci*, 2021, 64: 110301
- Feng D Q, Lai L F, Luo J J, et al. Ultra-reliable and low-latency communications: applications, opportunities and challenges. *Sci China Inf Sci*, 2021, 64: 120301
- Naik G, Choudhury B, Park J M. IEEE 802.11bd & 5G NR V2X: evolution of radio access technologies for V2X communications. *IEEE Access*, 2019, 7: 70169–70184
- Richardson T. Error floors of LDPC codes. In: *Proceedings of the 41st Annual Allerton Conference*, Monticello, 2003. 1426–1435
- Karimi M, Banihashemi A H. Efficient algorithm for finding dominant trapping sets of LDPC codes. *IEEE Trans Inform Theory*, 2012, 58: 6942–6958
- Dolecek L, Lee P, Zhang Z Y, et al. Predicting error floors of structured LDPC codes: deterministic bounds and estimates. *IEEE J Sel Areas Commun*, 2009, 27: 908–917
- Farsiabi A, Banihashemi A H. Error floor analysis of LDPC row layered decoders. *IEEE Trans Inform Theory*, 2021, 67: 5804–5826
- Hashemi Y, Banihashemi A H. New characterization and efficient exhaustive search algorithm for leafless elementary trapping sets of variable-regular LDPC codes. *IEEE Trans Inform Theory*, 2016, 62: 6713–6736
- Hashemi Y, Banihashemi A H. Characterization of elementary trapping sets in irregular LDPC codes and the corresponding efficient exhaustive search algorithms. *IEEE Trans Inform Theory*, 2018, 64: 3411–3430
- Han Y, Ryan W E. Low-floor decoders for LDPC codes. *IEEE Trans Commun*, 2009, 57: 1663–1673
- Caire G, Taricco G, Biglieri E. Bit-interleaved coded modulation. *IEEE Trans Inform Theory*, 1998, 44: 927–946
- Liao Y, Qiu M, Yuan J H. Design and analysis of delayed bit-interleaved coded modulation with LDPC codes. *IEEE Trans Commun*, 2021, 69: 3556–3571
- Zhu M Y, Jiang M, Zhao C M. Error floor estimation of QC-LDPC coded modulation with importance sampling. *IEEE Commun Lett*, 2021, 25: 28–32
- Wachsmann U, Fischer R F H, Huber J B. Multilevel codes: theoretical concepts and practical design rules. *IEEE Trans Inform Theory*, 1999, 45: 1361–1391
- Divsalar D, Jones C. Protograph based low error floor LDPC coded modulation. In: *Proceedings of IEEE Military Communications Conference*, 2005. 378–385
- Tang C J, Shen H, Jiang M, et al. Optimization of generalized VDM for protograph-based LDPC coded BICM. *IEEE Commun Lett*, 2014, 18: 853–856
- Jin Y, Jiang M, Zhao C M. Optimized variable degree matched mapping for protograph LDPC coded modulation with 16QAM. In: *Proceedings of IEEE ISTC*, 2010. 161–165
- Otarinia M, Fuja T E. The 5G new radio code: elementary absorbing sets and error floor performance. In: *Proceedings of International Symposium on Information Theory and Its Applications*, Kapolei, 2020. 215–219
- He H, Jiang M, Zhu M Y, et al. Lowering the error floor of quantized NR LDPC decoders by a post-processing on trapping sets. In: *Proceedings of International Conference on Wireless Communications and Signal Processing*, Changsha, 2021. 1–5
- Richardson T, Kudekar S. Design of low-density parity check codes for 5G new radio. *IEEE Commun Mag*, 2018, 56: 28–34
- 3GPP. The 3rd Generation Partnership Project; Technical specification group radio access network; NR; Physical channels and modulation (Release 15). 3GPP TS 38.211 V15.6.0. 2019. https://www.3gpp.org/ftp/Specs/archive/38_series/38.211
- Liva G, Chiani M. Protograph LDPC codes design based on EXIT analysis. In: *Proceedings of IEEE GLOBECOM*, Washington, 2007. 3250–3254
- Steiner F, Bocherer G, Liva G. Protograph-based LDPC code design for shaped bit-metric decoding. *IEEE J Sel Areas Commun*, 2016, 34: 397–407
- Brink S T, Kramer G, Ashikhmin A. Design of low-density parity-check codes for modulation and detection. *IEEE Trans Commun*, 2004, 52: 670–678
- Chen J H, Fossorier M P C. Near optimum universal belief propagation based decoding of low-density parity check codes. *IEEE Trans Commun*, 2002, 50: 406–414
- Ranganathan S V S, Divsalar D, Wesel R D. Quasi-cyclic protograph-based raptor-like LDPC codes for short block-lengths. *IEEE Trans Inform Theory*, 2019, 65: 3758–3777

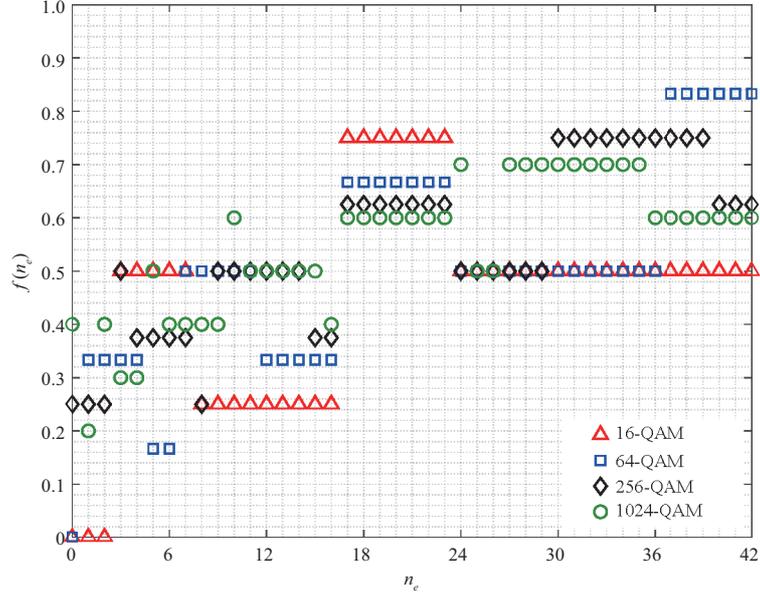


Figure A1 (Color online) All possible values of $f(n_e)$ for the LDPC codes constructed from the BG1 with the 16-, 64-, 256-, and 1024-QAM.

Appendix A

As mentioned in Subsection 3.2, some conflicted bit-levels may be chosen in the two bit-mapping steps $\Phi_1(\cdot)$ and $\Phi_2(\cdot)$. We now give an explanation to show the confliction can be completely avoided if the puncturing only performs at the block-column level. Given a Z , we can get the lowest-rate parity-check matrix of the 5G LDPC codes and then the higher-rate codes can be obtained by puncturing some degree-1 parity-check bits. Moreover, we know that puncturing degree-1 parity-check bits is equivalent to the removal of the corresponding columns and rows of the parity-check matrix. Thus, these higher-rate 5G LDPC codes can be seen as defined by a modified parity-check matrix with the removal of some columns and rows. We assume that the number of punctured degree-1 parity-check bits is a multiple of Z , which implies that the block property is still hold for the modified parity-check matrix. In this case, we can also obtain a modified protograph from the base graph by removing some CNs and VNs. Then, let n_i , n_c , and n_e denote, respectively, the number of transmitted information VNs, core parity VNs, and extension parity VNs in this modified protograph. For BG1, we have $n_i = 20$, $n_c = 4$, and $n_e \in \{0, 1, \dots, 42\}$. The number of VNs in a VN group \mathbb{V} is $v = (n_i + n_c + n_e)/m$. The numbers of core parity VN groups and extension parity VN groups can be computed, respectively, as

$$|\mathcal{S}_c| = \lceil (n_i - v \lfloor n_i/v \rfloor + n_c) / v \rceil, \quad (\text{A1})$$

$$|\mathcal{S}_e| = \lfloor n_e / v \rfloor. \quad (\text{A2})$$

Recall that x^0, x^1, \dots, x^{m-1} are arranged in descending order in terms of the reliability. According to (6) and (9), we know that the most protected bit-level chosen in $\Phi_1(\cdot)$ is $x^{m-|\mathcal{S}_e|}$ and the least protected bit-level chosen in $\Phi_2(\cdot)$ is $x^{2(|\mathcal{S}_c|-1)}$. Therefore, no conflicted bit-levels will be selected provided that $2(|\mathcal{S}_c| - 1) < m - |\mathcal{S}_e|$, which is also equivalent to the condition that $\frac{1}{m}(2(|\mathcal{S}_c| - 1) + |\mathcal{S}_e|) < 1$. Since n_i and n_c are constants for a given base graph, we can define a function with respect to n_e to describe the above condition. So let $f(n_e) = \frac{1}{m}(2(|\mathcal{S}_c| - 1) + |\mathcal{S}_e|)$. Since the values of n_e are limited in the set $\{0, 1, \dots, 42\}$, we can compute $f(n_e)$ exhaustively. By substituting n_i , n_c , m , and n_e into $f(n_e)$, we numerically verify that the inequality $f(n_e) < 1$ is hold, which indicates that the conflict will never happen under the supposed puncturing method. For 16-, 64-, 256-, and 1024-QAM, all possible values of $f(n_e)$ are given in Figure A1.