Design of a Labeling Scheme for 32-QAM Delayed Bit-Interleaved Coded Modulation

Zhe Zhang 1, Liang Zhou 1,*, Junyi Du 2 and Yue Zhao 1

1 National Key Lab on Communication, University of Electronic Science and Technology of China, Chengdu 611731, China; 17781480151@163.com (Z.Z.); zhaoyue@std.uestc.edu.cn (Y.Z.)
2 Southwest China Institute of Electronic Technology, Chengdu 610000, China; jydu1989@163.com
* Correspondence:lzhou@uestc.edu.cn

Received: 13 May 2020; Accepted: 19 June 2020; Published: 22 June 2020

Abstract: It is very challenging to design the capacity-approaching labeling schemes for large constellations, such as 32-QAM, in delayed bit-interleaved coded modulation (DBICM). In this paper, we investigate the labeling design for 32-QAM DBICM with various numbers of bits delayed by one time slot. In particular, we aim to obtain the labeling schemes with a high DBICM channel capacity by searching the possible labeling schemes. To reduce the search space of the candidate labeling schemes, we propose the criteria that are necessary for good labeling. Based on the proposed criteria, a three-step search algorithm is proposed to obtain the candidate labeling efficiently. Numerical results demonstrate that the DBICM with our proposed labeling scheme can approach the capacity of 32-QAM within 0.015 dB at an information rate greater than 2.5 bits/symbol.

Keywords: bit-interleaved coded modulation; delayed bit-interleaved coded modulation; gray labeling; depth-first search algorithm

1. Introduction

Bit-interleaved coded modulation (BICM) [1] has been widely used in practical systems for its simplicity. It is a bandwidth efficient coded modulation scheme and has been well investigated in point-to-point communications [1–7]. Compared to trellis-code modulation [8], BICM introduces a random interleaver between the encoder and modulator, which enables researchers to design the encoder and the modulator separately [7]. Therefore, BICM reduces the implementation complexity and attracts enormous attention.

In BICM, the dependence between transmitted bits is ignored by introducing the ideal bit-interleaver. This character is convenient for code design. However, due to the loss of dependence between transmitted bits, the channel capacity of BICM is less than that of coded modulation (CM), which is the constellation capacity. For squared constellations that have Gray labeling schemes, the gap between BICM capacity and constellation capacity is negligible [1]. For non-squared constellations, e.g., eight-quadrature amplitude modulation (QAM) and 32-QAM, Gray labeling is impossible, and the capacity gap is significant. The work [1] proposed the concept of quasi-Gray labeling, and this labeling means to be the best choice for cross constellations. The works [9,10] proposed a design method for quasi-Gray labeling based on the edge-profile. The other newer related works about the labeling focus on the BICM-iterative decoding (ID) [11–14] or the other constellation [15–19].

To reduce the gap, an improved BICM scheme, namely delayed BICM (DBICM), was proposed in [20]. DBICM maps bits from multiple consecutive codewords to a transmit symbol, which introduces dependence among the codewords. In decoding the consecutive codewords, the decoded result of a codeword is fed back to the demapper as a priori information for the next codeword. Therefore, the DBICM scheme can achieve a considerable gain over BICM. In [20], DBICM was verified to have a
higher channel capacity and achieved about 0.7 dB and 0.5 dB gain over the corresponding BICM in eight-amplitude-shift keying (ASK) and half 16-QAM, respectively. It was pointed out in [20] that the labeling scheme is essential to DBICM design, the same as BICM [11,12,15–18]. However, different from BICM, delay pattern can also affect the performance of DBICM. The DBICM capacity with various delay pattern was investigated in [21]. In [22], the authors proposed the criteria for designing the bit labeling scheme for DBICM with iterative decoding. One of their criteria was based on the harmonic mean of the minimum squared Euclidean distance with a priori information. The other criterion was based on the bit-wise mutual information metrics. Numerical results showed that a 0.5 dB gain was achieved by the optimal labeling scheme designed with their criteria compared to BICM-ID in 16-QAM. A low density parity check (LDPC) code design for DBICM based on the paragraph extrinsic information transfer (EXIT) was proposed in [23]. They achieved a signal-to-noise ratio gain of 0.5 dB to 0.1 dB over the BICM counterparts at a code rate ranging from 0.25 to 0.5. In [24], the authors proposed a pipeline decoding structure of LDPC-DBICM in a two-way relay channel. In the 8-PSK constellation, DBICM with specific delay patterns had about 0.8 dB and 0.35 dB over the BICM scheme, respectively.

Though the DBICM scheme can achieve considerable gain over BICM, it is a great challenge to find the optimal labeling scheme for large constellations, such as 32-QAM. In this work, we attempt to tackle this problem. We consider the cross 32-QAM constellation for its wide application in practical systems and its high energy efficiency [25–28]. To design the labeling scheme for 32-QAM DBICM in order to maximize the DBICM capacity, a straightforward way is to enumerate all the 32-QAM labeling schemes for each number of delayed bits. However, this is computationally infeasible for the huge number of labeling schemes. Therefore, the search space should be reduced to make the search feasible. Our work focuses on how to reduce the search space and find the good labeling schemes for the 32-QAM constellation. To the best of our knowledge, this is the first work focused on designing labeling for such a large constellation. We summarize the main contributions of this work as follows:

1. We propose three criteria for good labeling schemes that have a high DBICM capacity for 32-QAM based on the Gray labeling rule. In particular, Criterion 1 requires that the partial labels of delayed bits satisfy the Gray labeling rule in the 32-QAM constellation and the partial labels of undelayed bits satisfy the Gray labeling rule in the sub-constellations determined by the values of delayed bits. Criterion 2 requires that both the partial labels of delayed bits and undelayed bits satisfy the Gray labeling rule in the 32-QAM constellation. Criterion 3 requires that 4 bit partial labeling satisfy the Gray labeling in 32-QAM. These proposed criteria significantly reduce the search space for good labeling schemes.

2. We propose an efficient three-step search algorithm to find all candidate labeling schemes based on the proposed criteria. In the first step, we propose a depth-first search algorithm to find the partial labeling based on Criterion 3. Then, we construct the full labeling from that partial labeling. In the end, by eliminating the labeling not satisfying Criterion 1, 2, we obtain the candidate labeling. For 32-QAM, the proposed algorithm can find the candidate labeling schemes in a short period of time.

3. We propose to use the sum Hamming distance between all constellation pairs with the minimum Euclidean distance to evaluate the bit error rate (BER) performance of two capacity equivalent labeling schemes. The labeling with a smaller sum Hamming distance is expected to have a better BER performance.

4. Numerical results show that our designed labeling scheme can approach the capacity of 32-QAM within 0.015 dB at an information rate larger than 2.5 bits/symbol. In addition, for the labeling scheme with the same DBICM capacity, the labeling scheme with a smaller sum Hamming distance has a better BER performance than that with a larger Hamming distance.

The remaining parts of this paper are as follows. Section 2 introduces the concept of DBICM. Section 3 analyzes the channel capacity of DBICM and explains why we focus on the delay patterns with various bits delayed by one time slot. Section 4 introduces the proposed criteria for good labeling schemes and the proposed labeling search algorithm. Section 5 demonstrates the designed labeling
schemes, the DBICM channel capacity, and the BER results of LDPC coded DBICM schemes. Section 6 draws the conclusion.

2. Delayed Bit-Interleaved Coded Modulation

Figure 1 depicts the system model of a DBICM system. At the transmitter side, an information sequence \( \mathbf{b} \), of length \( k \) is encoded and interleaved in time slot \( t \). The resultant codeword \( \mathbf{c}_t \) is of length \( n \). Then, for a constellation of size \( 2^m \), \( \mathbf{c}_t \) is passed through a one-to-\( m \) serial-to-parallel converter, leading to \( m \) length \( n/m \) sub-sequences \( \mathbf{c}'_i \), \( i = 0, \ldots, m-1 \). Here, we assume that \( n \) is a multiple of \( m \) for simplicity. All sub-sequences \( \mathbf{c}'_i \) are sent to a bit delay module, in which \( \mathbf{c}'_i \) is delayed by \( T_i \) symbol periods. Finally, every \( m \) bit, in which the \( i^{th} \) bit is drawn from \( c'_{i-T_i} \), is mapped to a symbol \( x \) and transmitted through an additive white Gaussian noise (AWGN) channel. In DBICM, the vector \( D = \{ T_0, T_1, \ldots, T_{m-1} \} \) of length \( m \) is defined as the delay pattern. It determines the dependence between multiple consecutive codewords.

At the receiver side, noisy signals \( \mathbf{y} \) are received and fed to the demapper. The demapper calculates the log-likelihood ratios (LLRs) for each bit in \( \mathbf{c}'_{i-T_i} \) based on \( \mathbf{y} \) and the extrinsic information about the delayed bits of the previous codewords, denoted by \( L_i[\mathbf{c}_t] \), from the decoder. The resultant LLRs are denoted by \( L_i[\mathbf{c}'_{i}] \). The LLRs are fed to an inverse bit delay module to reconstruct the LLRs for the codeword \( \mathbf{c}_t \), which are denoted by \( L_i[\mathbf{c}_t] \). After an \( m \)-to-one parallel-to-serial converter and a deinterleaver, the LLRs of the codeword \( \mathbf{c}_t \) are fed to the decoder to obtain the estimated information sequence \( \hat{\mathbf{b}}_t \). The extrinsic information associated with the delayed bits, i.e., \( c'_{i-T_i}, T_i \neq 0 \), is fed back to the demapper to facilitate the demapping of the undelayed bits, i.e., \( c'_{i-T_i}, T_i = 0 \), of the next codeword.

It can be seen from Figure 1 that two extra modules, a bit delay module and an inverse delay module, are introduced in a DBICM system compared to a BICM system. The bit delay module introduces a coupling effect between consecutive codewords. The receiver exploits this coupling effect by demapping the undelayed bits of a codeword with the a priori information associated with the delayed bits from the previous coupled codewords. This leads to a potential SNR gain of DBICM over BICM.

3. Channel Capacity Analysis and Delay Pattern Selection

In this section, we analyze the channel capacity of BICM and DBICM. Then, the delay patterns \( D \) considered in this paper for DBICM will be specified.

The channel capacity for BICM \( C_{BICM} \) is a function of the constellation \( \mathcal{X} \), the labeling scheme \( \mathcal{L} : \{0,1\}^m \rightarrow \mathcal{X} \), and the SNR. In this work, we only consider the cross 32-QAM constellation for a given SNR. Therefore, \( \mathcal{L} \) determines the BICM channel capacity.
Denote $l(x) : \{0,1\}^m$ the label of a constellation point $x \in \mathcal{X}$. Let $l_i(x), i \in \{0,1,\ldots,m-1\}$, be the $i^{\text{th}}$ bit of $l(x)$, which is transmitted on the $i^{\text{th}}$ bit channel. Let $I(l_i;y)$ be the mutual information (MI) between the $i^{\text{th}}$ label bit and the received signal $y$. The BICM capacity for 32-QAM is written as [1]:

$$C_{\text{BICM}} = \sum_{i=0}^{4} I(l_i;y).$$  \hspace{1cm} (1)

In BICM, all label bits are considered to be independent. The reason is the introducing of the ideal bit-interleaver between the encoder and the mapper. As a result, the BICM channel capacity is less than the constellation capacity $C$ in AWGN channel, which is written as:

$$C = I(l_0;y) + I(l_1;y|l_0) + \cdots + I(l_4;y|l_0,l_1,l_2,l_3).$$  \hspace{1cm} (2)

In DBICM, the channel capacity is not only a function of the constellation $\mathcal{X}$, the labeling scheme $\mathcal{L}$, and the SNR. It is also a function of the delay pattern $\mathcal{D}$. Denote the indexes of the delayed bits, i.e., $l_i$ with $T_i \neq 0$, by $\mathcal{A}$ and the indexes of the undelayed bits, i.e., $l_i$ with $T_i = 0$, by $\mathcal{A}^c = \{0,\ldots,m-1\} \setminus \mathcal{A}$. As explained in Section 2, the demapping of delayed bits $l_\mathcal{A}$ only depends on the received signals, and therefore, the bit channel capacities for these bits are the same as those in BICM. However, the demapper utilizes the a priori information associated with the delayed bits in previous codewords to calculate the LLRs of the undelayed bits $l_{\mathcal{A}^c}$. Therefore, the bit channel capacities of the undelayed bits are different from those in BICM. In the following, we show two example delay patterns in Figure 2. By assuming that the delayed bits from previous codewords are perfectly known by the demapper, the DBICM capacities of these delay patterns are calculated.

![Figure 2](example-image.png)

**Figure 2.** Two delay patterns for 32-QAM. (a) The delay pattern $\mathcal{D} = \{T_0 = 1, T_1 = T_2 = T_3 = T_4 = 0\}$. (b) The delay pattern $\mathcal{D} = \{T_0 = 4, T_1 = 3, T_2 = 2, T_3 = 1, T_4 = 0\}$.

In Figure 2a, $l_0$ is delayed by one time slot. For this delay pattern, the DBICM channel capacity $C_{\text{DBICM}}$ is written as:

$$C_{\text{DBICM}} = I(l_0;y) + \sum_{i=1}^{4} I(l_i;y|l_0).$$  \hspace{1cm} (3)
In Figure 2b, $l_i$ is delayed by $m - i - 1$ time slots. In this case, the DBICM channel capacity $C_{DBICM}$ is written as:

$$C_{DBICM} = I(l_0;y) + I(l_1;y|l_0) + \cdots + I(l_4;y|l_0,l_1,l_2,l_3). \quad (4)$$

It can be seen from Equation (4) that the DBICM capacity for the delay pattern in Figure 2b is equal to the 32-QAM constellation capacity $C$ (as shown in Equation (2)) regardless of the labeling scheme. Therefore, it is optimal in the sense that it maximizes the DBICM capacity. However, this delay pattern has a large decoding latency since the codeword transmitted in time slot $t$ cannot be decoded until time slot $t + 4$. In this paper, we consider the delay patterns in which the delayed bits $I_A$ are delayed by only one time slot, i.e., $T_i \leq 1$ for $i \in \{0, \ldots, m - 1\}$. This minimizes the decoding latency introduced by DBICM. More importantly, we will show in Section 5 that by carefully designing the labeling scheme for this type of delay patterns, the DBICM capacity can almost achieve the constellation capacity at a moderate to high information rate. For this type of delay patterns, the DBICM capacity is written as:

$$C_{DBICM} = \sum_{i \in A} I(l_i;y) + \sum_{j \in A^c} I(l_{j};y|l_A). \quad (5)$$

4. Design of the Labeling Scheme

In this section, we consider the design of labeling schemes for the delay patterns of interest in this work. In particular, we aim to maximize the DBICM capacity by designing labeling schemes and their associated delay patterns. To achieve this, a straightforward way is to enumerate all possible combinations of the considered delay patterns and the labeling schemes and calculate the capacity for each combination. The one that maximizes the DBICM capacity is selected as the optimized labeling scheme and the associated delay pattern. However, this is computationally infeasible since there are $P_{32}^m$ labeling schemes for 32-QAM. Here, $P_{n}^m = n!/(n-m)!$ represents the permutation of $m$ elements from $n$ elements. By considering the delay patterns, the number of combinations is even larger.

In the following, we propose the criteria that should be satisfied by a good labeling scheme for a given delay pattern. Based on the proposed criteria, the size of the candidate labeling set will be small enough. Then, we propose an efficient three-step search algorithm, which utilizes the proposed criteria to obtain the candidate labeling. After we eliminate some equivalent labeling from the candidate set, we enumerate all the candidate labeling and choose the one with the highest DBICM capacity and the best BER performance as our designed labeling.

4.1. Labeling Search Criteria

To present the search criteria, let us first introduce the concepts of the partial label (PL), partial labeling scheme (PLS), and partial Gray labeling scheme (PGLS).

4.1.1. Partial Gray Labeling Scheme

**Definition 1. (PL):** For a given non-empty index set $C = \{c_j | c_j \in \{0,1,\ldots, m-1\}\}$, the PL of a constellation point $x$ is defined as $I(x)$ excluding $\{l_j(x), j \in C\}$, which is denoted by $I^C(x)$. Here, $C$ is the complementary set of $C$ in $\{0,1,\ldots, m-1\}$, i.e., $C = \{0,1,\ldots, m-1\}\backslash C$.

**Definition 2. (PLS):** For a given non-empty index set $C = \{c_j | c_j \in \{0,1,\ldots, m-1\}\}$ and a labeling scheme $\mathcal{L}$ of $\mathcal{X}$, by replacing $I(x)$ with $I^C(x)$ for all $x \in \mathcal{X}$, the resultant labeling scheme is a PLS of $\mathcal{X}$, denoted by $\mathcal{L}^C$.

Note that in PLS $\mathcal{L}^C$, $2^{|C|}$ distinct labels are mapped to $2^m$ constellation points, where $|C|$ obtains the cardinality of $C$. Therefore, each label is mapped to $2^{m-|C|}$ points.
where we propose the criterion for optimal labeling schemes of a given delay pattern next. Note that, since each label in PLS $L^C$ is mapped to $2^n-|C|$ points, there might be pairs of constellation points $(x_i, x_j), x_i, x_j \in \mathcal{X}, i \neq j,$ which have the minimum Euclidian distance $d_{min}$ in $\mathcal{X}$, which have the same PLs. In this case, PLS $L^C$ still satisfies the Gray labeling rule as explained in Definition 3. Therefore, $L^C$ is a PGLS.

In [1], it was conjectured that Gray labeling maximizes the BICM capacity. Based on this principle, we propose the criterion for optimal labeling schemes of a given delay pattern next.

4.1.2. Criterion for Optimal Labeling Schemes

It can be seen from Equation (5) that the DBICM capacity is the sum of the bit channel capacities of delayed bits $I_{A_t}$, i.e., $\sum_{i \in A} I(l_i;y)$, and the bit channel capacities of undelayed bits $I_{A_c}$, i.e., $\sum_{j \in A^C} I(l_j;y|I_{A_c})$. To optimize the DBICM capacity, the capacities of delayed and undelayed bits should be maximized simultaneously.

For the delayed bits $I_{A_t}$, their bit channel capacities are equal to those in BICM. Based on the Gray labeling rule for BICM, the PLS for $C = A_t$, i.e., $L^{A_t}$, should be a PGLS in order to maximize $\sum_{i \in A} I(l_i;y)$.

For the undelayed bits $I_{A_c}$, if the values of the delayed bits $I_A$ are given, i.e., $I_A = \{b_0,\ldots,b_d\}$, where $d = |A|$, their bit channel capacities are equal to the BICM capacities in the sub-constellation $\mathcal{X}^{A_t}$. Here, $\mathcal{X}^{A_t} = \{x | x \in \mathcal{X}, I^A_t(x) = I_A = \{b_0,\ldots,b_d\}\}$. Therefore, in order to maximize the channel capacity of undelayed bits, i.e., $\sum_{j \in A^C} I(l_j;y|I_{A_c})$, the labeling of $I_{A_c}$ should satisfy the Gray labeling rule in the sub-constellations $\mathcal{X}^{A_t}$ for all $I_A \in \{0, 1\}^d$.

In summary, the criterion for optimal labeling schemes is as follows:

**Criterion 1:**

1. The PLS of delayed bits $I_{A_t}$, i.e., $L^{A_t}$, should be a PGLS.
2. The labeling of undelayed bits $I_{A_c}$ should satisfy the Gray labeling rule in the sub-constellations $\mathcal{X}^{A_t}$ for all $I_A \in \{0, 1\}^d$.

Let $L_{11}$ be the set of labeling schemes that satisfy the first item of Criterion 1 and $L_{12}$ be the set of labeling schemes that satisfy the second item of Criterion 1, respectively. The set of candidate optimal labeling schemes is $L_1 = L_{11} \cap L_{12}$. To obtain $L_1$, we can first find $L_{11}$ and then eliminate the labeling schemes $L$ that are not in $L_{12}$, i.e., $\{L | L \in L_{11}, L \notin L_{12}\}$.

Though Criterion 1 is optimal in the sense that the optimal labeling scheme should be in $L_1$, the search space, i.e., $|L_1|$, is still too large. To see this, let us consider the delay patterns in which one bit is delayed, i.e., $|A| = 1$. We first examine the space of $L_{11}$ and then $L_{12}$. For $|A| = 1$, for the delayed bits $I_{A_t}$, all possible PLSs $L^{A_t}$ are PGLSs since the PLs for $A$ have only one bit. Therefore, the PLs of all constellation points have at most one bit in difference, which means they satisfy the Gray labeling rule. To find a labeling scheme $L_{11}$, we first select 16 points $\{x_1, x_2, \ldots, x_{16}\} \subset \mathcal{X}$ and then the sub-constellation $\mathcal{X}_0$. There are $C_{16}^{16}$ choices for $\mathcal{X}_0$, where $C_n^m = n!/(m!(n-m)!)$ is the number of combinations of selecting $m$ elements from $n$ elements. Then, for each $x \in \mathcal{X}_0$, let PL $I^{A_t}(x) = 0$. For the undelayed bits $I_{A_c}$, $2^{16} = 16$ distinct PLs will be assigned to $\mathcal{X}_0$. There are $P_{16}^{16}$ ways to label $\mathcal{X}_0$. For the other 16 points $x \in \mathcal{X}, x \notin \mathcal{X}_0$, denote them by $\mathcal{X}_1$, and let their PLs $I^{A_t}(x) = 1$. There are also $P_{16}^{16}$ ways to label the undelayed bits of $\mathcal{X}_1$. Therefore, $|L_{11}| = C_{16}^{16} \times P_{16}^{16} \times P_{16}^{16}$. Turning to $L_{12}$, we cannot calculate the exact number of labeling schemes that satisfy Criterion 1(b). However, our simulations showed that $L_{12}$ could not be found in a reasonable time frame. Therefore, by only considering Criterion 1 for optimal labeling schemes, the size of $L_1$ is still too large.

In order to reduce the search space, we propose more criteria for good labeling schemes in the following.
4.1.3. Criteria for Good Labeling Schemes

**Criterion 2**: The PLS $L^C$ of the undelayed bits $l^C$ should be a PGLS.

Note that Criterion 2 is different from the second item of Criterion 1. In Item 2 of Criterion 1, the labeling of undelayed bits is required to satisfy the Gray labeling rule only in the sub-constellations. However, in Criterion 2, the PLS $L^C$ of the undelayed bits should satisfy the Gray labeling rule in the whole constellation so that it is a PGLS.

Denote $L_2$ as the set of labeling schemes that satisfy Criterion 2. It can be seen that $L_2 \subseteq L_{12}$ since if a labeling scheme satisfies Criterion 2, it must satisfy Item 2 of Criterion 1, but the inverse is not true. Therefore, the search space for good labeling schemes is reduced by replacing Item 2 of Criterion 1 with Criterion 2, i.e., $|L_2 \cap L_{11}| \leq |L_{12} \cap L_{11}|$. Note that by introducing Criterion 2, the optimal labeling scheme may not be found since the optimal labeling scheme is guaranteed to be within $L_{12} \cap L_{11}$, but not guaranteed to be within $L_2 \cap L_{11}$. Therefore, we say that Criterion 2 is for a good labeling scheme instead of for the optimal labeling scheme. However, the simulation results in Section 5 showed that the obtained good labeling scheme could almost approach the constellation capacity at a mid-to-high information rate regime.

Though the proposed Criterion 2 reduced the search space for good labeling schemes, it was still too large. To see this, let us consider the delay patterns in which two bits are delayed as an example. For $|A| = 2$, in order to find $L_2$, we first search for all PLSs $L^C$ that satisfy Criterion 2. Note that for each PLS $L^C$, $2^{||A||} = 8$ unique PLs are assigned to the 32 constellation points, and each of the unique PLs are assigned to $2^{||A||} = 4$ constellation points. Then, for each $L^C$, the full labeling schemes are constructed by assigning the labels of the delayed bits to the four points with the same PL. For each group of four points with the same PL, there are $P_4^4 = 24$ ways to label the delayed bits. By considering that there are eight groups of points with different PLs for the delayed bits, each $L^C$ can generate $(P_4^4)^8$ full labeling schemes. We found the number of PLSs $L^C$ that satisfy Criterion 2 by simulations. It was approximately equal to $1.43 \times 10^8$. Therefore, $|L_2| \approx 1.43 \times 10^8 \times (P_4^4)^8$ for $|A| = 2$.

To find $L_{11}$, we first found all PLSs $L^A$ that satisfied Item 1 of Criterion 1. Then, for each $L^A$, the full labeling schemes could be constructed in the same way as that for finding $L_2$. Finally, we found that $|L_{11}| \approx 3.7 \times 10^9 \times (P_8^8)^4$. Obviously, the sizes of $L_2$ and $L_{11}$ were still very large. In order to further reduce the search space, we propose the following criterion.

**Criterion 3**: For a good labeling scheme $L$, there exists an index set $A^* = \{i_1, i_2, i_3, i_4\} \subset \{0, 1, \ldots, 4\}$, and PLS $L^{A^*}$ is a PGLS.

Denote $L_3$ as the set of the labeling that satisfies the Criterion 3. Now, our target for obtaining the set of good labeling candidates is to find $L = L_{11} \cap L_2 \cap L_3$. To show that Criterion 3 must be satisfied by a good labeling, we will prove that the following proposition holds.

**Proposition 1.** For each undelayed bit $l_j, j \in A^*$, its channel capacity in DBICM is larger than or equal to its channel capacity in BICM, i.e.,

$$I(l_j; y) \leq I(l_j; y|1_A).$$  \hspace{1cm} (6)

**Proof.** The left-hand-side of Equation (6) can be written as:

$$I(l_j; y|1_A) = H(l_j|1_A) - H(l_j|y, 1_A) = H(l_j) - (H(l_j) - H(l_j|1_A)) - H(l_j|y, 1_A)$$

$$= H(l_j) - I(l_j; 1_A) - H(l_j|y, 1_A).$$  \hspace{1cm} (7)
where \( H(\cdot) \) is the commentropy function. In both DBICM and BICM, for using an ideal bit-interleaver, all bits are independent, i.e., \( I(l_i; l_j) = 0 \) for all \( i, j \in \{0, \ldots, m-1\} \) and \( i \neq j \). Therefore, \( I(l_i; l_A) = 0 \) and Equation (7) can be written as:

\[
I(l_i; y | l_A) = H(l_i) - H(l_i | y, l_A) \\
= H(l_i) - H(l_i | y) + H(l_i | y) - H(l_i | y, l_A) \\
= I(l_i; y) + I(l_i; l_A | y) \geq I(l_i; y).
\]  

(8)

Note that \( I(l_i; l_A | y) = 0 \) does not imply \( I(l_i; l_A | y) = 0 \) since \( y \) can be regarded as a function of \( l_j \) and \( l_A \). When \( y \) is known, the MI between \( l_j \) and \( l_A \) depends on the structure of \( y \).

Based on Proposition 1 and Equation (5), the following proposition holds.

**Proposition 2.** For a given constellation \( X \), labeling scheme \( \mathcal{L} \), and SNR, the DBICM capacity is larger than or equal to the BICM capacity, i.e.,

\[
C_{DBICM} = \sum_{i \in A} I(l_i; y) + \sum_{j \in A^c} I(l_j; y | l_A) \\
\geq \sum_{i \in A} I(l_i; y) + \sum_{j \in A^c} I(l_i; y) \\
= C_{BICM}
\]  

(9)

It can be seen from Proposition 2 that the BICM capacity is the lower bound of the DBICM capacity. If a labeling scheme satisfies Criterion 3, it can maximize the BICM capacities of four bits and expect to have a high BICM capacity. Therefore, Criterion 3 can optimize the lower bound of the DBICM capacity and should be satisfied by a good labeling scheme.

### 4.2. Good Labeling Search Guideline

The proposed criteria for good labeling are summarized in Table 1. To find the candidate labeling set \( L = L_{11} \cap L_2 \cap L_3 \), one can find the set with the minimum number of elements in \( L_{11}, L_2, \) and \( L_3 \) first. Then, the labeling schemes in the smallest set are checked against the other two criteria. The labeling that does not belong to \( L \) is eliminated.

| Criteria Number | Description |
|-----------------|-------------|
| 1               | The PLS of delayed bits \( l_A \), i.e., \( \mathcal{L}^A \), should be a PGLS. |
| 2               | The PLS \( \mathcal{L}^A \) of the undelayed bits \( l_A \) should be a PGLS. |
| 3               | PLS \( \mathcal{L}^{A^*} \) for which \( A^* = \{i_1, i_2, i_3, i_4\} \subset \{0, 1, \ldots, 4\} \) is a PGLS. |

In this paper, we find \( L_3 \) first. To achieve this, all the four bit PGLS \( L^{A^*} \) are enumerated and are denote by \( L^* \). Then, the full labeling schemes are constructed from \( L^* \), and \( L_3 \) is obtained. Then, the labeling that does not belong to \( L_{11}, L_2 \) is eliminated from \( L_3 \).

Note that with \( L_3 \), various delayed patterns should be considered. For each delay pattern, all labeling schemes in \( L_3 \) are checked against Item 1 of Criteria 1 and Criterion 2. Take the delay patterns with one delayed bit as an example. For each labeling \( \mathcal{L} \in L_3 \), there are \( C_1^1 \) delay patterns \( \mathcal{D}_i = \{T_i = 1, \{T_j = 0\}, j = \{0, 1, \ldots, m-1\} \ \backslash \{i\}, i \in \{0, 1, \ldots, m-1\} \) to be considered. If no delay patterns can let the labeling \( \mathcal{L} \) satisfy Item 1 of Criteria 1 and Criterion 2 simultaneously, \( \mathcal{L} \) is eliminated from \( L_3 \). With this process, all the candidate labeling schemes \( \mathcal{L} \in L \) and the associated delay patterns are obtained. Then, for each \( \mathcal{L} \in L \) and the associated delay patterns, the DBICM capacity is calculated.
The labeling \( L \) and its associated delay pattern that achieve the highest DBICM capacity for a given SNR are considered to be the designed bit labeling scheme. In the following, an efficient three-step algorithm is proposed to implement this search guideline for good labeling schemes.

4.3. Search Algorithm for Good Labeling Schemes

In this section, we propose a three-step algorithm to find the candidate labeling schemes. In the PGLS search step, we find the set of 4 bit PGLSs \( L^* \) by utilizing an efficient depth-first search algorithm. In the full labeling generation step, we utilize the PGLSs \( \mathcal{L}^{A^*} \in L^* \) to construct the full labeling. The final step is to eliminate the labeling that does not belong to \( L_{1u}, L_2 \) simultaneously. Furthermore, we eliminate some strictly equivalent labeling schemes (SELS), which are defined as below, from the candidate set.

For two DBICM labeling schemes \( \mathcal{L}_1 \) and \( \mathcal{L}_2 \) of a constellation \( X \), the sets of delayed and undelayed bits are \( \mathcal{A}_1, \mathcal{A}_2 \), respectively. The partial labels of a point \( x \) for \( \mathcal{A}_1 \) associated with \( \mathcal{L}_1 \) and \( \mathcal{L}_2 \) are denoted by \( \ell_{1}^{A_1}(x), \ell_{2}^{A_1}(x) \), respectively. The partial labels of a point \( x \) for \( \mathcal{A}_2 \) associated with \( \mathcal{L}_1 \) and \( \mathcal{L}_2 \) are denoted by \( \ell_{1}^{A_2}(x), \ell_{2}^{A_2}(x) \), respectively. Let \( |y,z|_H \) be the Hamming distance between two (partial) labels \( y \) and \( z \) of \( x \).

**Definition 4. (SELS):** If for any pair of constellation points \( x_1, x_2 \in X \), \( |(\ell_{1}^{A_1}(x_1), \ell_{1}^{A_1}(x_2))|_H = |(\ell_{2}^{A_2}(x_1), \ell_{2}^{A_2}(x_2))|_H \) and \( |(\ell_{1}^{A_2}(x_1), \ell_{1}^{A_2}(x_2))|_H = |(\ell_{2}^{A_2}(x_1), \ell_{2}^{A_2}(x_2))|_H \) are satisfied simultaneously, \( \mathcal{L}_1, \mathcal{L}_2 \) are called SELS (an example of the SELS is shown in Figure 3).

![Figure 3. An example of the SELS. (a) Labeling scheme \( \mathcal{L}_1 \). (b) Labeling scheme \( \mathcal{L}_2 \).](image)

Note that for a pair of SELS \( \mathcal{L}_1 \) and \( \mathcal{L}_2 \), their DBICM capacities are the same. This is because with \( |(\ell_{1}^{A_1}(x_1), \ell_{1}^{A_1}(x_2))|_H = |(\ell_{2}^{A_2}(x_1), \ell_{2}^{A_2}(x_2))|_H \), the capacities of the delayed bits for \( \mathcal{L}_1, \mathcal{L}_2 \) are the same, and with \( |(\ell_{1}^{A_2}(x_1), \ell_{1}^{A_2}(x_2))|_H = |(\ell_{2}^{A_2}(x_1), \ell_{2}^{A_2}(x_2))|_H \), the capacities of the undelayed bits for \( \mathcal{L}_1, \mathcal{L}_2 \) are the same. Based on (5), \( \mathcal{L}_1 \) and \( \mathcal{L}_2 \) have the same DBICM capacity. In addition, \( \mathcal{L}_1 \) and \( \mathcal{L}_2 \) have the same BER performance since for any pair of constellation points, the correspondence between the Hamming distance of their labels and their Euclidean distance remains the same.

The derivation of the three-step algorithm is shown in the Appendix A.

4.4. Capacity Equivalent Labeling Schemes

Although we excluded the SELS in the final step of constructing \( L \), there are still some labeling schemes that have the same DBICM capacity, but are not SELS. We call these labeling schemes capacity equivalent labeling schemes (CELS). We investigate CELS since though they are equivalent in the sense of DBICM capacity, they have different BER performance. In this work, we aim to find the labeling
scheme that not only has a high DBCIM capacity, but also a good BER performance. In the following, CELS is defined.

Definition 5. (CELS): For two DBICM labeling schemes $\mathcal{L}_1$ and $\mathcal{L}_2$ with the same PLS of delayed bits, i.e., $\mathcal{L}_1^A = \mathcal{L}_2^A$. For each sub-constellation $\mathcal{X}^A$, if for any pair of constellation points $x_1, x_2 \in \mathcal{X}^A$, $|(i^{x_1}, i^{x_2})|_H = |(i^{x_1}, i^{x_2})|_H$ holds. The two labeling $\mathcal{L}_1, \mathcal{L}_2$ are called CELS.

It can be seen from Equation (5) that the capacity of undelayed bits is the average of the BICM capacities in each sub-constellation $\mathcal{X}^A$. A pair of CELS has the same BICM capacity in each sub-constellation $\mathcal{X}^A$; therefore, they have the same capacity of undelayed bits. Their capacities of delayed bits are also the same since they have the same PLS of delayed bits. To sum up, a pair of CELS has the same DBICM capacity.

Note that a pair of SELS is a pair of CELS, but the inverse is not necessarily true. For example, two labeling schemes $\mathcal{L}_1, \mathcal{L}_2$ are shown in Figure 4a,b, respectively. The delayed bits are $A = \{l_0, l_1\}$. The undelayed bits are $A^c = \{l_2, l_3, l_4\}$, which are in the parentheses. The two labeling have the same PLS of delayed bits, i.e., $\mathcal{L}_1^A = \mathcal{L}_2^A$. Each eight points with the same color constitute a sub-constellation $\mathcal{X}^A$. It can be seen that, for any two points $x_1, x_2 \in \mathcal{X}^A$, $|(i^{x_1}, i^{x_2})|_H = |(i^{x_1}, i^{x_2})|_H$ holds. Therefore, the two labeling $\mathcal{L}_1, \mathcal{L}_2$ are CELSs. In Figure 4a, the labels of the two points in the box are $10(000), 11(100)$, and the Hamming distance of the undelayed PL is one. In Figure 4b, the labels of the same two points are $10(000), 11(000)$, and the Hamming distance of the undelayed PL is 0. Therefore, $\mathcal{L}_1, \mathcal{L}_2$ are not SELS, and they will have different BER performance.

![Figure 4. An example CELS. (a) The labeling $\mathcal{L}_1$. (b) The labeling $\mathcal{L}_2$.](image)

The calculation of DBICM capacity in Equation (5) is based on the hypothesis that all the delayed bits are perfectly known by the demapper. This holds in a high SNR region. However, in a low SNR region, the hypothesis is not valid. Therefore, if the delayed bits have errors, the demodulation structure of DBICM for the undelayed bits is destroyed, i.e., we cannot obtain the full DBICM capacities of the undelayed bits in the wrong sub-constellation. The demodulation results depend on the received value and the labeling of the whole constellation.

In this paper, we only consider the minimum Euclidean distance $d_{\text{min}}$ in the constellation as the error radius, which is held in most instances. This is because the probability of making decision errors from one point to another point decreases exponentially w.r.t the Euclidean distance between two points. In the 32QAM constellation, for a signal $s$ mapped by constellation point $x$, if the error happens in the transmission process, the received value $y$ can be demodulated to the neighbor constellation points, which have the distance $d_{\text{min}}$ away from $x$. The cost of this error is the Hamming distance of the labels between $x$ and its neighbor. Therefore, we employ the sum Hamming distance to evaluate the CELS, which is the total cost of the errors with a radius of $d_{\text{min}}$. 
The labeling schemes $\mathcal{L}_1$ and $\mathcal{L}_2$ in Figure 4 are CELS, and the sum Hamming distances for $\mathcal{L}_1$ and $\mathcal{L}_2$ are 68 and 60, respectively. Therefore, $\mathcal{L}_2$ is expected to have a better BER performance in a low SNR regime than $\mathcal{L}_1$. In a high SNR regime, $\mathcal{L}_1, \mathcal{L}_2$ will have almost the same BER. Numerical results in Figures 7 and 8 will confirm this expectation.

Note that we focus on the 32-QAM constellation in this paper for its bandwidth efficiency and wide use in practical systems [25–28], e.g., optical communication. The proposed method can also be extended to higher order modulation such as 128-QAM and 512-QAM, but the complexity becomes prohibitively high. The reason is that our search method is based on the algebraic structure of 32-QAM. For 128-QAM and 512-QAM, there is an exponential growth of the constellation points. Therefore, applying our method to these constellations will lead to the exponential growth of the complexity.

For example, the numbers of candidate labeling schemes associated with a PGLS in 128-QAM and 512-QAM are $2^{63}$ and $2^{55}$, respectively. Contrary to the number of $2^{15}$ in 32-QAM, the search spaces of 128-QAM and 512-QAM are too large.

5. Numerical Results

In this section, we first show our designed labeling schemes with the associated delay patterns. Then, the DBICM capacities of these labeling schemes are compared to the constellation capacity and that of the quasi-Gray labeling in [10]. BICM capacity with the quasi-Gray labeling is also demonstrated for comparison. Finally, we show the BER performance of LDPC coded DBICM and BICM with the proposed labeling schemes and the quasi-Gray labeling scheme.

5.1. The Proposed Labeling Schemes

We utilize the proposed labeling search algorithm in Section 4.4 to find the candidate labeling schemes that have a high DBICM capacity. In the 4 bit PGLS search step, $|L^*| = 174,792$ 4 bit PGLSs are found. By executing the full labeling generation step and eliminating the SELS, $|L_3| = 33,292,288$ candidate labeling schemes are found. For the delay patterns with 1, 2, 3, or 4 bits delayed by one time slot, the numbers of candidate labeling are $|L| = 7,891,148, 16,683,242, 16,683,242, 7,891,148$, respectively.

Then, the DBICM capacities for these labeling schemes with their delay patterns at the SNR of 7 dB and 9.2 dB were calculated. We considered this SNR since the DBICM channel capacity was around 2.5, three bits/symbol or the code rate was around 0.5, 0.6 at this point, respectively, which is practical for high order coded modulation schemes. To eliminate the CELS introduced in Section 4.4, the sum Hamming distances of these labeling schemes were calculated. The one with the minimum sum Hamming distance was selected as our designed labeling scheme.

Our designed labeling scheme for the delay patterns with 1, 2, and 3 bits delayed was the same, which is shown in Figure 4b. The designed labeling scheme for the delay patterns with 4 bits delayed is shown in Figure 5. The associated delay patterns were $D_1 = \{T_0 = T_1 = T_3 = T_4 = 0, T_2 = 1\}$, $D_2 = \{T_0 = T_1 = 1, T_2 = T_3 = T_4 = 0\}$, $D_3 = \{T_0 = T_1 = 0, T_2 = T_3 = T_4 = 1\}$, and $D_4 = \{T_1 = T_2 = T_3 = T_4 = 1, T_0 = 0\}$, respectively.

We also found the optimal delay patterns for the quasi-Gray labeling by enumerating all possible delay patterns. The associated delay patterns were $D'_1 = \{T_1 = T_2 = T_3 = T_4 = 0, T_0 = 1\}$, $D'_2 = \{T_0 = T_2 = 1, T_1 = T_3 = T_4 = 0\}$, $D'_3 = \{T_3 = T_4 = 0, T_0 = T_1 = T_2 = 1\}$, and $D'_4 = \{T_1 = T_2 = T_3 = T_4 = 1, T_0 = 0\}$, respectively.
5.2. The Channel Capacity

The DBICM channel capacities with our designed labeling schemes and the quasi-Gray labeling scheme are shown in Figure 6. The BICM channel capacity of the quasi-Gray labeling scheme is also shown in Figure 6 for comparison.

It can be seen from Figure 6 that for the same number of delayed bits, our designed labeling scheme outperformed the quasi-Gray labeling scheme in terms of DBICM capacity. It is notable that for our designed labeling scheme with three bits delayed by one time slot, the DBICM capacity was only 0.015 dB away from the constellation capacity at an information rate greater than 2.5 bits/symbol.

5.3. BER Performance

In this subsection, we evaluate the BER performance of LDPC coded DBICM and BICM with various labeling schemes. For the DBICM capacity of 2.5 bits/symbol, an LDPC code with code length $n = 10,000$ and code dimension $k = 5000$ was constructed by the progressive edge-growth (PEG) algorithm [29], where the distribution of the variable node degree is optimized by the differential evolution algorithm [30]. For the DBICM capacity of three bits/symbol, the LDPC code from DVB-S2 [31] with the code rate 3/5 and code length 64,800 was employed. The interleaver used the consecutive mapping [3] in the coded modulation. Figures 7 and 8 show the BER performance...
of the DBICM capacities of 2.5 bits/symbol and three bits/symbol, respectively. We can see that our designed labeling schemes had better BER performance compared to that of quasi-Gray labeling.

In Figures 7 and 8, we also compare the labeling schemes $L_1, L_2$ in Figure 4a,b to show the difference of the BER of CELS. We can see that in the low SNR region, $L_1$ had worse BER performance than that of $L_2$. In the high SNR regime, $L_1$ had nearly the same BER as $L_2$. This agreed with the expectation described in Section 4.4 that the one with the minimum Hamming distance would have better performance in the low SNR region.

6. Conclusions

In this paper, we considered the bit labeling design for 32-QAM DBICM. We aimed to obtain the labeling that had a high DBICM capacity and a good BER performance. To reduce the search space for the good labeling, three criteria were proposed. An efficient labeling search algorithm was proposed to obtain the labeling schemes that satisfied the proposed criteria. Then, we evaluated the candidate
labeling with DBICM capacity in a fixed SNR point and selected the best one as our designed labeling. This SNR point must correspond to the most efficient capacity of the constellation. Capacity analysis showed that the designed labeling with three bits delayed by one time slot approached the 32-QAM constellation capacity within 0.015 dB at a mid-to-high information rate regime. BER simulation results also demonstrated that the designed labeling achieved a 0.1 dB gain in terms of BER performance over the benchmark scheme.

Author Contributions: Conceptualization, Z.Z. and L.Z.; methodology, Z.Z. and J.D.; software, Z.Z. and Y.Z.; validation, Z.Z., L.Z., and J.D.; formal analysis, Z.Z. and J.D.; investigation, Z.Z.; resources, L.Z.; data curation, Y.Z.; writing, original draft preparation, Z.Z.; writing, review and editing, Z.Z., L.Z., and J.D.; visualization, Z.Z. and Y.Z.; supervision, L.Z.; project administration, L.Z.; funding acquisition, L.Z. All authors read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations
The following abbreviations are used in this manuscript:

- QAM Quadrature amplitude modulation
- DBICM Delayed bit-interleaved coded modulation
- BICM Bit-interleaved coded modulation
- CM Coded modulation
- ASK Amplitude-shift keying
- BICM-ID Bit-interleaved coded modulation iterative decoding
- BER Bit error rate
- LDPC Low density parity check
- AWGN Additive white Gaussian noise
- LLR Log-likelihood ratio
- SNR Signal-to-noise ratio
- MI Mutual information
- PL Partial label
- PLS Partial labeling scheme
- PGLS Partial Gray labeling scheme
- SELS Strictly equivalent labeling schemes
- CELS Capacity equivalent labeling schemes
- PEG Progressive edge-growth

Appendix A. The Three-Step Search Algorithm

Appendix A.1. The PGLS Search Step

In the PGLS search step, we propose a depth-first search algorithm (Algorithm 1) to obtain the set of 4 bit PGLSs \( L^* \). Let \( B = \{0000, 0000, 0001, 0001, \ldots, 1111, 1111\} \) be the set of 4 bit PLs. Denote \( I = \{i_1, i_2, \ldots, i_{32}\} \) as the PL assignment order to the 32 points in the 32-QAM constellation, i.e., \( i_1 \) is the index of a point in the 32-QAM constellation that will be assigned a 4 bit PL first, and \( i_2 \) is the index of a point in the 32-QAM constellation that will be assigned a 4 bit PL last. The algorithm finds the set of 4 bit PGLSs \( L^* \) in a few recursive steps.

In the first step, the constellation point \( x_{i_1} \) is assigned an arbitrary 4 bit PL in \( B \). Without loss of generality, let \( I^{A^*}(x_{i_1}) = (0000) \), where \( A^* \) is the indexes of the four bits that their PLS is a PGLS. Then, let \( B = B \setminus I^{A^*}(x_{i_1}) \), where \( B \setminus I^{A^*}(x_{i_1}) \) represents that the set \( B \) excludes the element \( I^{A^*}(x_{i_1}) \).

In the second step, find all legal PLs of point \( x_{i_2} \) from \( B \) and denote them by \( S^2 \). Here, legal PLs means that by assigning the PL to \( x_{i_2} \), the PLs of \( x_{i_1} \) and \( x_{i_2} \) satisfy the partial Gray labeling rule. Then, assign a PL \( I^{A^*}(x_{i_2}) \in S^2 \) to \( x_{i_2} \), and let \( B = B \setminus I^{A^*}(x_{i_2}) \).
In the $i^{th}$ step, $t = 3, \ldots, 32$, find all legal PLs of point $x_i$, and denote them by $S^t_i$. If $S^t_i$ is not empty, assign a PL $l^A(x_i) \in S^t_i$ to $x_i$, and let $B = B \setminus l^A(x_i)$. Otherwise, go back to the $t-1^{th}$ step, and assign another PL $l^A(x_{i-1}) \in S^{t-1}_i$ to $x_{i-1}$.

The above process continues. If $t = 32$ steps are executed successfully, a 4 bit PGLS is found and recorded in $L^*$. The algorithm goes back to the $31^{th}$ step and tries another PL in $S^{31}$. In the $i^{th}$ step, $S^t_i$ is empty, then go back to the $(t-1)^{th}$ step, and assign another PL in $S^{t-1}_i$ to $x_{i-1}$. When all the PLs in $S^2$ have been exhausted, the set of 4 bit PGLSs $L^*$ is obtained. The proposed depth-first search algorithm is summarized in Algorithm 1.

**Algorithm 1 Depth-first bit-labeling search.**

**Input:** The assignment order $I$;

**Output:** The 4 bit PGL schemes set $L^*$.

1. begin procedure
2. initialize: let $l^A(x_1) = (0000)$; $B = B \setminus l^A(x_1)$; set $t = 2$;
3. if $S^t_i \subset B$ then
4. go to 14;
6. end if
7. for $j = 1$; $j < |S^t_i|; j + +$ do
8. $l^A(x_i) = s_j \in S^t_i$; $B = B \setminus l^A(x_i)$, $t = t + 1$;
9. for $i = 32$ then
10. record the partial Gray labeling schemes $L^A_i : \{l^A(x_1), l^A(x_2), \ldots, l^A(x_{32})\}$ in $L^*$;
11. go to 14;
12. end if
13. return to 3;
14. $t = t - 1$, $B = B \cup \{l^A(x_i)\}$;
15. end for
16. return $L^*$.

It can be seen from Algorithm 1 that in each $i^{th}$ step, if $x_i$ is chosen so that the size of $S^t_i$ is minimized, the search complexity will be reduced significantly. This is because the PLs that do not satisfy the Gray labeling rule are excluded in an earlier stage and will not be considered in the following search process. Therefore, the label assignment order $I$ is essential for reducing the complexity of Algorithm 1.

In the following proposed Algorithm 2, we obtain the label assignment order $I$ to minimize the complexity of Algorithm 1. For Algorithm 2, the inputs are the numbered points of constellation $\mathcal{X} = \{x_0, x_1, \ldots, x_{2^{w-1}}\}$ and the set of 4 bit PL candidates $B$. The output is the assignment order $I$. The constellation point numbering rule is arbitrary. For example, in Figure A1, the points are numbered from 0 to 31 from left to right and top to bottom.

Firstly, one constellation point is selected as the first one to be labeled in an arbitrary manner. For example, in Figure A1, $x_0$ is selected. In this case, $i_1 = 5$. Let $l^A(x_{i_1}) = (0000)$. Then, eliminate $l^A(x_{i_1})$ from $B$, and eliminate $x_{i_1}$ from $\mathcal{X}$. Go to the step $t = 2$.

In the $i^{th}$ step, for each of the remaining points $x_j \in \mathcal{X}, j = 1, \ldots, |\mathcal{X}|$, find the set of legal PLs in $B$, and denote it by $S^t_{x_{i_1}}, S^t_{x_{i_2}}, \ldots, S^t_{x_{i_{|\mathcal{X}|}}}$, within all the non-empty sets $S^t_{x_j}$, choose the one with the minimum number of elements, and let the associated constellation point be the next one to be labeled. For example, if $S^t_{x_j}$ is the chosen one, let $i_t = k_j$. If there are multiple sets that have the minimum number of legal PLs, one is selected randomly. When $i_t$ is obtained, $x_{i_t}$ is assigned with a random PL $l^A(x_{i_t}) \in S^t_{x_{i_t}}$, and let $B = B \setminus l^A(x_{i_t})$, $\mathcal{X} = \mathcal{X} \setminus x_{i_t}$. When $t = 32$, the assignment order $I$ is generated.
Algorithm 2 The assignment order.

**Input:** The numbered points of constellation $\mathcal{X} = \{x_0, x_1, \ldots, x_{2^m-1}\}$; the label candidate set $B$.

**Output:** The assignment order $I = \{i_1, i_2, \ldots, i_{32}\}$.

1: begin procedure
2: initialize: let $i_1 = r, r \in \{1, 2, \ldots, 32\}, I(x_{i_1}) = (0000), B = B \setminus I(x_{i_1}), \mathcal{X} = \mathcal{X} \setminus x_{i_1}$; set $t = 2$;
3: for $j = 1, j \leq |\mathcal{X}|; j = +1$ do
4: find $S^I_{x_{i_j}}$ for $x_{i_j} \in \mathcal{X}$;
5: end for
6: find one with minimum size in $\{S^I_{x_{i_j}} | j = 1, 2, \ldots, |\mathcal{X}|\}$, and denote it as $S^I_{x_{i_1'}}$;
7: if there are multiple sets having the minimum number of legal PLs, one is selected randomly;
8: let $i_t = k';$
9: assign $x_{i_t}$ with a random label $l(x_{i_t})$ in $S^I_{x_{i_1'}}$;
10: $B = B \setminus I(x_{i_t}), \mathcal{X} = \mathcal{X} \setminus x_{i_t}$;
11: if $t = 32$ then
12: go to 16;
13: else
14: $t = t + 1$, go to 3;
15: end if
16: return $I$.

Figure A1 depicts an example of the assignment order $I$ generated by Algorithm 2. In the figure, a pair of numbers $(x_j, i_k), j = 0, 1, \ldots, 31, k = 1, 2, \ldots, 32,$ is associated with each constellation point. In particular, $x_j$ represents the numbering of a point in the 32-QAM constellation, while $i_k$ represents the numbering order of a point. For example, $(x_5, i_1)$ means that the point $x_5$ will be assigned a PL first, and $(x_0, i_2)$ means that the point $x_0$ will be assigned a PL in the second step of Algorithm 1.

![Figure A1. An example of the assignment order.](image)

In Algorithm 1, by employing this assignment order, only 1,244,160 attempts are executed for obtaining the set of 4 bit PGLSs $L^*$. This is completed in a few minutes on a personal computer.

Appendix A.1.1. The Full Labeling Generation Step

In this step, we construct the full labeling from the 4 bit PGLSs $L^* \in L^*$ by assigning the label to the bit that was not included in the PL. Without loss of generality, we assume that the 4 bit PLs are assigned to $l_0, l_1, l_2, l_3$. In this case, a value of zero or one will be assigned to $l_4$ of each constellation point to obtain the full labeling schemes.

For a 4 bit PGLS $L^*$, it maps the 4 bit PLs to 32 constellation points. Therefore, each unique PL is assigned to a pair of constellation points. For each pair of points that have the same 4 bit PL,
$l_4 = 0$ is assigned to a point, and $l_4 = 1$ is assigned to the other one. In this way, a full labeling $L$ is constructed from a PGLS $L^4$. For example, in Figure 3a, the two red points have the same 4 bit PL (0000). By assigning $l_4 = 0$ to the left one and $l_4 = 1$ to the right one, the full labels (00000) and (00001) for these two points are constructed.

For a 4 bit PGLS $L^4$, there are 16 distinct PLs. For each unique PL, $l_4$ can be assigned a value of zero or one. Therefore, there are $2^{16}$ ways to assign the values to $l_4$ of the 16 distinct PLs. Denote the set of those labeling as $L^4$. For each labeling $L_1 \in L^4$, there must be a labeling $L_2 \in L^4$ such that the values of its $l_4$ for all constellation points are the inverses of those of $L_1$.

Note that these two labeling schemes $L_1, L_2$ are SELS since the Hamming distance between any two PLs does not change by inverting the value of $l_4$. Figure 3 demonstrates an example, and the labeling schemes in Figure 3a,b are SELS. The values of $l_4$ for all points in these two labeling are inverted.

By considering the above SELS, there are only $2^{15}$ distinct labeling schemes in $L^4$. As a result, the number of full labeling schemes constructed from $L^4$ is $2^{15}|L^4|$, i.e.,

$$|L_3| = 2^{15}|L^4|.$$  \hfill (A1)

Appendix A.1.2. The Final Step

For the delay patterns with 1, 2, 3, or 4 delayed bits, each labeling in $L_3$ is checked against Item 1 of Criterion 1 and Criterion 2. If the check fails, the labeling is excluded from $L_3$. By now, the good labeling candidate set $L$ for all delayed patterns is obtained.

Note that the candidate set $L$ still contains many SELS. In $L$, as the cross 32-QAM constellation is central symmetry, a labeling scheme in $L$ and the labeling schemes that are generated by rotating itself by $\frac{\pi}{4}$, $\pi$, $\frac{3\pi}{4}$ are SELS. Furthermore, a labeling scheme and its reflection are SELS. By excluding these SELS, the size of $L$ is small, which is shown in Section 5.

In the end, for each labeling in $L$, its DBICM capacity is calculated. The labeling scheme and its associated delay pattern that achieves the maximum DBICM capacity are selected as our designed labeling for 32-QAM.

References

1. Caire, G.; Taricco, G.; Biglieri, E. Bit-interleaved coded modulation. IEEE Trans. Inf. Theory 1998, 44, 927–946.
2. Du, J.; Zhou, L.; He, X. The joint demodulation and decoding for BICM-ID system with BEC approximation and optimized bit mapping. In Proceedings of the 2015 IEEE Global Communications Conference (GLOBECOM), San Diego, CA, USA, 6–10 December 2015.
3. Du, J.; Yang, L.; Yuan, J.; Zhou, L.; He, X. Bit mapping design for LDPC coded BICM schemes with multi-edge type EXIT chart. IEEE Commun. Lett. 2017, 21, 722–725.
4. Hou, J.; Siegel, P.H.; Milstein, L.B.; Pfister, H.D. Capacity-approaching bandwidth-efficient coded modulation schemes based on low-density parity-check codes. IEEE Trans. Inf. Theory 2003, 49, 2141–2155.
5. Zhao, Y.; Fang, Y.; Yang, Z. Interleaver Design for Small-Coupling-Length Spatially Coupled Protograph LDPC-Coded BICM Systems Over Wireless Fading Channels. IEEE Access 2020, 8, 33500–33510.
6. He, Y.; Jiang, M.; Ling, X.; Zhao, C. Robust BICM Design for the LDPC Coded DCO-OFDM: A Deep Learning Approach. IEEE Trans. Commun. 2020, 68, 713–727.
7. Zehavi, E. 8-PSK trellis codes for a Rayleigh channel. IEEE Trans. Commun., 1992, 40, 873–884.
8. Ungerboeck, G. Channel coding with multilevel/phase signals. IEEE Trans. Inf. Theory, 1982 28, 55–67.
9. Wesel, D.; Liu, X. Edge profile optimal constellation labeling. In Proceedings of the 2000 IEEE International Conference on Communications (ICC 2000): Global Convergence Through Communications, Conference Record, New Orleans, LA, USA, 18–22 June 2000.
10. Wesel, D.; Liu, X.; Cioffi, J.M.; Komninakis, C. Constellation labeling for linear encoders. IEEE Trans. Inf. Theory 2001, 47, 2417–2431.
11. Yang, Z.; Xie, Q.; Peng, K.; Song, J. Labeling schemes optimization for BICM-ID systems. IEEE Commun. Lett. 2010, 14, 1047–1049.
12. Muhammad, N.S.; Speidel, J. Joint optimization of signal constellation bit labeling schemes for bit-interleaved coded modulation with iterative decoding. *IEEE Commun. Lett.* **2005**, *9*, 775–777.

13. Zhang, J.; Djordjevic, I.B. Mapping design of a four-dimensional 32-ary signal constellation over an AWGN channel. *Chin. Commun.* **2014**, *11*, 775–777.

14. Navazi, H.M.; Hossain, M.J. Novel Method for Multi-Dimensional Mapping of Higher Order Modulations for BICM-ID Over Rayleigh Fading Channels. *IEEE Trans. Wirel. Commun.* **2019**, *18*, 1142–1154.

15. Yang, F.; Yan, K.; Xie, Q.; Song, J. Non-equiprobable APSK constellation labeling schemes design for BICM systems. *IEEE Comm. Lett.*, **2013**, *17*, 1276–1279.

16. Bocherer, G. Labeling Non-Square QAM Constellations for One-Dimensional Bit-Metric Decoding. *IEEE Commun. Lett.*, **2014**, *18*, 1515–1518.

17. Markiewicz, T.G. Construction and Labeling of Triangular QAM. *IEEE Commun. Lett.* **2017**, *21*, 1751–1754.

18. Colman, G.W.K.; Gohary, R.H.; El-Azizy, M.A.; Willink, T. J.; Davidson, T.N. Quasi-Gray Labelling for Grassmannian Constellations. *IEEE Trans. Wirel. Commun.* **2011**, *10*, 626–636.

19. Jia, L.; Shu, F.; Chen, M.; Zhang, W.; Li, J.; Wang, J. Joint Constellation-Labeling Optimization for VLC-CSK Systems. *IEEE Wirel. Commun. Lett.* **2019**, *8*, 1280–1284.

20. Ma, H.; Leung, W.K.; Yan, X.; Law, K.; Fosser, M. Delayed bit interleaved coded modulation. In Proceedings of the 2016 9th International Symposium on Turbo Codes and Iterative Information Processing, Brest, France, 5–9 September 2016.

21. Yan, X.; Machado, R.G.; Huang, K.; Gabry, F.; Fosser, M.; Hafermann, H.; Zhang, H.; Land, I.; Leung, W.K. Capacity Analysis of Delayed Bit Interleaved Coded Modulation. In Proceedings of the 2018 IEEE 10th International Symposium on Turbo Codes & Iterative Information Processing (ISTC), Hong Kong, China, 3–7 December 2018.

22. Wang, L.; Cai, S.; Ma, H.; Leung, W.K.; Ma, X. Bit-labeling schemes for delayed BICM with iterative decoding. In Proceedings of the 2018 IEEE International Symposium on Information Theory (ISIT), Vail, CO, USA, 17–22 June 2018.

23. Liao, Y.; Yang, L.; Yuan, J.; Huang, K.; Leung, R.; Du, J. LDPC Code Design for Delayed Bit-Interleaved Coded Modulation. In Proceedings of the 2019 IEEE Information Theory Workshop (ITW), Visby, Sweden, 25–28 August 2019.

24. Du, J.; Wang, Z.; Xiao, L.; Wang, L.; Qiao, W. LDPC coded DBICM Scheme for Two Way Relay Channel with a Pipeline Decoding Structure. In Proceedings of the 2018 IEEE International Conference on Communication Systems (ICCS), Chengdu, China, 19–21 December 2018.

25. Yang, R.; Sherratt, R.S. Enhancing MB-OFDM throughput with dual circular 32-QAM. *IEEE Trans. Consum. Electron.* **2008**, *54*, 1640–1646.

26. Wake, D.; Dupont, S.; Lethien, C.; Vilcot, J.-P.; Decoster, D. Radiofrequency transmission of 32-QAM signals over multimode fibre for distributed antenna system applications. *Electron. Lett.* **2001**, *37*, 1087–1089.

27. Zhalehpour, S.; Guo, M.; Lin, J.; Zhang, Z.; Qiao, Y.; Shi, W.; Rusch, L.A. System Optimization of an All-Silicon IQ Modulator: Achieving 100-Gbaud Dual-Polarization 32QAM. *J. Lightwave Technol.* **2020**, *38*, 256–264.

28. Hamaoka, F.; Nakamura, M.; Nagatani, M.; Kobayashi, T.; Matsushita, A.; Wakita, H.; Yamazaki, H.; Nosaka, H.; Miyamoto, H. 120-GBaud 32QAM Signal Generation using Ultra-Broadband Electrical Bandwidth Doubler. In Proceedings of the Optical Fiber Communication Conference, San Diego, CA, USA, 3–7 March 2019.

29. Hu, X.; Eleftheriou, E.; Arnold, D.M. Regular and irregular progressive edge-growth Tanner graphs. *IEEE Trans. Info. Theory* **2005**, *51*, 386–398.

30. Storn, R.; Price, K. Differential evolution simple and efficient heuristic for global optimization over continuous spaces. *J. Global Optim.* **1997**, *11*, 341–359.

31. Morello, A.; Mignone, V. DVB-S2: The Second Generation Standard for Satellite Broad-Band Services. *Proc. IEEE* **2006**, *94*, 210–227.

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).