

Article

On Decoder Ties for the Binary Symmetric Channel with Arbitrarily Distributed Input

Ling-Hua Chang ^{1,*}, Po-Ning Chen ²  and Fady Alajaji ³ 

¹ Department of Electrical Engineering, Yuan Ze University, Taoyuan 32003, Taiwan

² Institute of Communications Engineering, National Yang-Ming Chiao-Tung University, Taipei 112304, Taiwan; poningchen@nycu.edu.tw

³ Department of Mathematics and Statistics, Queen's University, Kingston, ON K7L 3N6, Canada; fa@queensu.ca

* Correspondence: iamjaung@saturn.yzu.edu.tw

Abstract: The error probability of block codes sent under a non-uniform input distribution over the memoryless binary symmetric channel (BSC) and decoded via the maximum a posteriori (MAP) decoding rule is investigated. It is proved that the ratio of the probability of MAP decoder ties to the probability of error grows most linearly in blocklength when no MAP decoding ties occur, thus showing that decoder ties do not affect the code's error exponent. This result generalizes a similar recent result shown for the case of block codes transmitted over the BSC under a uniform input distribution.

Keywords: binary symmetric channel; block codes; non-uniformly distributed channel inputs; joint source-channel coding; maximum a posteriori (MAP) decoding; decoder ties; error probability; error exponent



Citation: Chang, L.-H.; Chen, P.-N.; Alajaji, F. On Decoder Ties for the Binary Symmetric Channel with Arbitrarily Distributed Input. *Entropy* **2023**, *25*, 668. <https://doi.org/10.3390/e25040668>

Academic Editors: Jun Chen and Sadaf Salehkalaibar

Received: 20 March 2023

Revised: 13 April 2023

Accepted: 15 April 2023

Published: 16 April 2023



Copyright: © 2023 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 (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Consider the classical channel coding context, where we send a block code through the memoryless binary symmetric channel (BSC) with crossover probability $0 < p < 1/2$. Given a sequence of binary codes $\{\mathcal{C}_n\}_{n \geq 1}$ with n being the blocklength, we denote the sequence of corresponding minimal probabilities of decoding error under *maximum a posteriori (MAP) decoding* by $\{a_n\}_{n \geq 1}$. The following result was recently shown in [1] when the channel input selects codewords from \mathcal{C}_n according to a uniform distribution.

Theorem 1 ([1]). For any sequence of codes $\{\mathcal{C}_n\}_{n \geq 1}$ of blocklength n and size $|\mathcal{C}_n| = M$ with $\mathcal{C}_n \subseteq \{0, 1\}^n$, sent over the BSC with crossover probability $0 < p < 1/2$ under a uniform channel input distribution over \mathcal{C}_n , its minimum probability of decoding error a_n satisfies

$$b_n \leq a_n \leq \left(1 + \frac{(1-p)}{p}n\right)b_n, \quad (1)$$

where

$$b_n = P_{X^n, Y^n} \left\{ (x^n, y^n) \in \mathcal{X}^n \times \mathcal{Y}^n : P_{X^n|Y^n}(x^n|y^n) < \max_{u^n \in \mathcal{C}_n \setminus \{x^n\}} P_{X^n|Y^n}(u^n|y^n) \right\}, \quad (2)$$

where P_{X^n, Y^n} is the joint input–output distribution that $X^n = (X_1, X_2, \dots, X_n) \in \mathcal{X}^n \triangleq \{0, 1\}^n$ is sent over the BSC (via n uses) and $Y^n = (Y_1, Y_2, \dots, Y_n) \in \mathcal{Y}^n \triangleq \{0, 1\}^n$ is received.

Noting that b_n in (2) is the probability that a decoding error occurs without inducing decoder ties (which occur when two or more codewords in \mathcal{C}_n are identified by the decoder as the estimated transmitted codeword; i.e., when more than one codeword in \mathcal{C}_n maximize

$P_{X^n|Y^n}(\cdot|y^n)$ for a given received word y^n , the above result in (1) directly implies that decoder ties do not affect the error exponent of a_n . The error exponent or reliability function of a block coding communication system represents the largest rate of exponential decay of the system's probability of decoding error as the coding blocklength grows to infinity (e.g., see [2–14]).

It is known that uniformly distributed data achieves the largest entropy rate and leaves no room for data compression. Thus, ideally compressed data should exhibit uniform distribution for all blocklengths n . However, this setting is often impractical due to the sub-optimality of the implemented data compression schemes. Instead, we generally have non-uniformly distributed data after compression in the form of residual redundancy such as in speech or image coding (e.g., [15,16]). Furthermore, one may have a compressed source that can be divided into several groups, within each of which the symbols are equally probable. Decoder ties can thus occur with respect to two (or more) codewords corresponding to symbols within the same group.

In this paper, we consider a non-uniform prior distribution over C_n and prove that decoder ties, under optimal MAP decoding, still have linear and hence sub-exponential impact on the error probability a_n , thus extending Theorem 1 established for the case of a uniform prior distribution over C_n . Since our problem falls within the general framework of joint source-channel coding for point-to-point communication systems, we refer the reader to [14–21] (Section 4.6) and the references therein for theoretical studies on this subject as well as practical designs that outperform separate source and channel coding under complexity or delay constraints.

The proof technique used in [1] to show (1) above is based on the observation that there are two types of decoding errors. One is that the received tuple at the channel output induces no decoder ties but the corresponding decoder decision is wrong. The other is that the received tuple at the channel output causes a decoder tie, but the decoder picks the wrong codeword. As a result, the MAP error probability a_n can be upper bounded by the sum of two terms, b_n and δ_n , where b_n is the probability of the first type of decoding errors as given in (2), and δ_n is the probability of decoder ties regardless of whether the tie breaker misses the correct codeword or not. Under the assumption that the channel input is uniformly distributed over block code C_n for each blocklength n and an arbitrary sequence of codes $\{C_n\}_{n \geq 1}$, it was shown in [1] that flipping a properly selected bit component of the channel output that causes a decoder tie can produce a unique channel output that leads to the first type of decoding errors. An analysis of this bit-flipping manipulation shows that the ratio δ_n/b_n grows at most linearly in n and hence yields the upper bound in (1). However, this flipping technique no longer works when non-uniform channel inputs are considered. To tackle this problem, we judiciously separate the channel output tuples that induce decoder ties into two groups, one group consisting of output tuples that do not fulfill the above flipping manipulation property and the other group composed of the remaining output tuples (i.e., the complement group). We then show that the probability of the former group is upper bounded by that of the latter group, and therefore δ_n/b_n remains growing at most linearly in blocklength n under arbitrary channel input statistics. Note that the group that fails the flipping property is an empty set when channel input is uniformly distributed over C_n , thereby making the result of Theorem 1 a special case of the extended result in this paper. The rest of the paper is organized as follows. Section 2 presents the main result and highlights the key steps of the proof to facilitate its understanding. The proof is then provided in full detail, along with illustrative examples, in Section 3 and Appendices A and B. Finally, conclusions and future directions are given in Section 4.

Throughout the paper, we denote $[M] \triangleq \{1, 2, \dots, M\}$ for positive integer M and set $d(x^n, y^n | \mathcal{S})$ to be the Hamming distance between n -tuples $x^n = (x_1, x_2, \dots, x_n)$ and $y^n = (y_1, y_2, \dots, y_n)$ with the indices of the tuples restricted to $\mathcal{S} \subseteq [n]$. By convention, we set $d(x^n, y^n | \mathcal{S}) = 0$ when $\mathcal{S} = \emptyset$ and use $d(x^n, y^n)$ to represent $d(x^n, y^n | [n])$.

2. Main Result

Consider a binary code $\mathcal{C}_n \subseteq \{0, 1\}^n$ with fixed blocklength n and size M to be used over the memoryless BSC with crossover probability $0 < p < \frac{1}{2}$. Denote the prior probability on \mathcal{C}_n by P_{X^n} and hence $P_{X^n}(\mathcal{C}_n) = 1$. Without loss of generality, we assume that all codewords in \mathcal{C}_n occur with positive probability, i.e., $P_{X^n}(x^n) > 0$ for all $x^n \in \mathcal{C}_n$; hence, \mathcal{C}_n is the support of P_{X^n} .

It is known the minimal probability of decoding error is achieved by the MAP decoder, which upon the reception of the channel output $y^n \in \{0, 1\}^n$ estimates the codeword $x^n \in \mathcal{C}_n$ according to

$$e(y^n) = \arg \max_{u^n \in \mathcal{C}_n} P_{X^n|Y^n}(u^n|y^n), \tag{3}$$

where $P_{X^n|Y^n}$ is the posterior conditional distribution of X^n given Y^n . We can see from (3) that if more than one $u^n \in \mathcal{C}_n$ achieves the maximum value of $P_{X^n|Y^n}(u^n|y^n)$ for a given y^n , a decoder tie occurs, in which case the set of these u^n , denoted conveniently as $\{e(y^n)\}$, contains more than one element. As a result, an erroneous MAP decision is made if one of the two situations occurs: *i*) the transmitted codeword does not belong to $\{e(y^n)\}$; *ii*) the transmitted codeword belong to $\{e(y^n)\}$ and $|\{e(y^n)\}| > 1$, but the tie breaker picks the wrong one from $\{e(y^n)\}$. By conveniently denoting

$$\mathcal{C}_n = \{c_1, c_2, \dots, c_M\}, \tag{4}$$

the probability of the first situation acts as a lower bound b_n for a_n (i.e., $b_n \leq a_n$), where b_n is given in (2) and can be written as

$$b_n = \sum_{i=1}^M P_{X^n}(c_i) P_{Y^n|X^n} \left(\left\{ y^n \in \{0, 1\}^n : P_{X^n|Y^n}(c_i|y^n) < \max_{r \in [M] \setminus \{i\}} P_{X^n|Y^n}(c_r|y^n) \right\} \right). \tag{5}$$

It is shown in [22] that b_n exactly equals the generalized Poor–Verdú (lower) bound [23,24] as its tilting parameter approaches infinity. The probability of the second situation is bounded above by the probability that the transmitted codeword belongs to $\{e(y^n)\}$ and $|\{e(y^n)\}| > 1$, disregarding whether the tie breaker picks the wrong codeword or not, and this upper bound can be expressed as

$$\delta_n \triangleq \sum_{i=1}^M P_{X^n}(c_i) P_{Y^n|X^n} \left(\left\{ y^n \in \{0, 1\}^n : P_{X^n|Y^n}(c_i|y^n) = \max_{r \in [M] \setminus \{i\}} P_{X^n|Y^n}(c_r|y^n) \right\} \right). \tag{6}$$

We thus have

$$b_n \leq a_n \leq b_n + \delta_n. \tag{7}$$

By proving the inequality

$$\delta_n \leq 2qn b_n, \tag{8}$$

where

$$q \triangleq \frac{1-p}{p} > 1, \tag{9}$$

we have our main result as follows.

Theorem 2. For any sequence of binary codes $\{\mathcal{C}_n\}_{n \geq 1}$ and prior probabilities $\{P_{X^n}\}_{n \geq 1}$ used over the BSC, we have

$$b_n \leq a_n \leq (1 + 2qn)b_n. \tag{10}$$

Remark 1. Theorem 2 implies that the relative deviation of a_n from b_n is at most linear in the blocklength n and the impact of decoder ties in (6) to a_n is only sub-exponential. Consequently, a_n and b_n must have the same error exponent. Note also that the upper bound in (10) differs from the result in Theorem 1 by an additional multiplicative factor of 2 in the qn term. As explained in the

introduction section, this is a consequence of the fact that the probability of the group of channel output tuples that cause decoder ties but fail the flipping manipulation property is upper bounded by that of the remaining tie-inducing channel outputs. The full technical details are provided in Section 3.2. Finally, we emphasize that Theorem 2 holds for arbitrary binary codes, including “bad” codes for which high probability codewords have small Hamming distance between them. Hence, tightening the upper bound in (10) by restricting the analysis for “sufficiently good” codes, in the sense that their most likely codewords sit “sufficiently” far apart in $\{0, 1\}^n$, is an interesting future direction.

List of Main Symbols: Before providing an overview of the main steps of the proof of Theorem 2 (which is presented in full detail in the next section), we describe in Table 1 the main symbols used in the paper and indicate the equation where they are first introduced. We emphasize that sets $\mathcal{T}_{j|i}$, $\mathcal{N}_{j|i}$ and $\mathcal{S}_{1,j}^{(m)}$ are defined differently from their counterparts in [1] that use the same notation.

We also visually illustrate in Figure 1 some of the main sets defined in Table 1 under the setting of Example 1, which is presented in Section 3 below for a non-uniformly distributed binary code with $M = 4$ codewords and blocklength $n = 4$ given by $\mathcal{C}_4 = \{c_1, c_2, c_3, c_4\} = \{0000, 0101, 0110, 0111\}$. More specifically, we only show the non-empty component subsets in $\mathcal{Y}^n = \{0, 1\}^4$ corresponding to codewords c_1 and c_2 ; refer to Table A2 in Appendix A for a detailed listing of all component subsets in $\{0, 1\}^4$ (including empty ones).

Table 1. Summary of the main symbols used in this paper.

Symbol	Description	Defined in
$[M]$	A shorthand for $\{1, 2, \dots, M\}$	
\mathcal{C}_n	The code $\{c_1, c_2, \dots, c_n\}$ with c_1 being the all-zero codeword	
$d(u^n, v^n \mathcal{S})$	The Hamming distance between the portions of u^n and v^n with indices in \mathcal{S}	
<i>All terms below are functions of \mathcal{C}_n (this dependence is not explicitly shown to simplify notation)</i>		
\mathcal{T}_i	The set of channel outputs y^n inducing a decoder tie when c_i is sent	(12)
\mathcal{N}_i	The set of channel outputs y^n leading to a tie-free decoder decision error when c_i is sent	(15)
$\mathcal{I}_i(y^n)$	The set $\{m \in [M] \setminus \{i\} : y^n \in \mathcal{T}_m\}$ for $y^n \in \mathcal{T}_i$	(21)
$\mathcal{S}_{i,j}$	The set of indices for which the components of c_i and c_j differ	
$\ell_{i,j}$	The size of $\mathcal{S}_{i,j}$, i.e., $ \mathcal{S}_{i,j} $	
$\mathcal{T}_{j i}$	The subset of \mathcal{T}_i consisting of channel outputs y^n such that j is the minimal number r in $\mathcal{I}_i(y^n)$ satisfying $d(c_i, y^n \mathcal{S}_{i,r}) < \ell_{i,r}$	(22a)
$\mathcal{N}_{j i}$	The subset of \mathcal{N}_i consisting of channel outputs y^n that satisfy $P_{X^n, Y^n}(c_i, y^n) \cdot q = P_{X^n, Y^n}(c_j, y^n) \cdot \frac{1}{q}$ and that are not included in $\mathcal{N}_{r i}$ for $r \in [j-1] \subset \{i\}$	(22b)
$\mathcal{T}_{j i}^*$	The subset of $\mathcal{T}_i \setminus \left(\bigcup_{h \in [M] \setminus \{i\}} \mathcal{T}_{h i}\right)$ consisting of channel outputs y^n such that j is the minimal number in $\mathcal{I}_i(y^n)$	(23)
$\mathcal{S}_{1,j}^{(m)}$	The subset of $\mathcal{S}_{1,j}$ defined according to whether each index in $\mathcal{S}_{1,j}$ is in each of $\mathcal{S}_{1,2}, \dots, \mathcal{S}_{1,j-1}, \mathcal{S}_{1,j+1}, \dots, \mathcal{S}_{1,M}$	(43)
$\mathcal{S}_{1,j}^{(m)}$	The union of $\mathcal{S}_{1,j}^{(1)}, \mathcal{S}_{1,j}^{(2)}, \dots, \mathcal{S}_{1,j}^{(m)}$	(48)
$\ell_{1,j}^{(m)}$	The size of $\mathcal{S}_{1,j}^{(m)}$, i.e., $ \mathcal{S}_{1,j}^{(m)} $	
η_k	The mapping from $k \in \{0, 1, \dots, \ell_j - 1\}$ to $[2^{M-2}]$ used for partitioning $\mathcal{T}_{j 1}$ into $\ell_{1,j}$ subsets $\{\mathcal{T}_{j 1}(k)\}_{0 \leq k < \ell_{1,j}}$	(49)
$\mathcal{T}_{j 1}(k)$	The k th partition of $\mathcal{T}_{j 1}$ for $k = 0, 1, \dots, \ell_{1,j} - 1$	(52a)
$\mathcal{N}_{j 1}(k)$	The k th subset of $\mathcal{N}_{j 1}$ for $k = 0, 1, \dots, \ell_{1,j} - 1$	(52b)
$\mathcal{U}_{j 1}(k)$	The set of representative elements in $\mathcal{T}_{j 1}(k)$ for partitioning $\mathcal{T}_{j 1}(k)$	
$\mathcal{T}_{j 1}(u^n; k)$	The subset of $\mathcal{T}_{j 1}(k)$ associated with $u^n \in \mathcal{U}_{j 1}(k)$	(55a)
$\mathcal{N}_{j 1}(u^n; k)$	The subset of $\mathcal{N}_{j 1}(k)$ associated with $u^n \in \mathcal{U}_{j 1}(k)$	(55b)

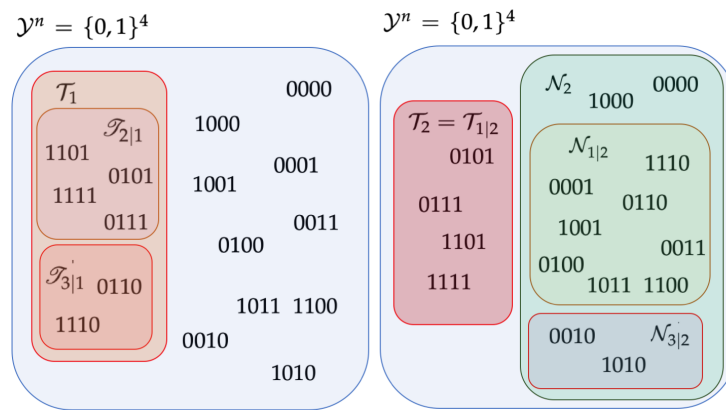


Figure 1. An illustration, based on the setting in Example 1 for a non-uniformly distributed binary code (with $M = n = 4$) given by $\mathcal{C}_4 = \{c_1, c_2, c_3, c_4\} = \{0000, 0101, 0110, 0111\}$ of the non-empty component subsets of \mathcal{Y}^n defined in Table 1 and corresponding to codewords $c_1 = 0000$ (left figure) and $c_2 = 0101$ (right figure).

Overview of the Proof: Given that codeword c_i is sent over the channel, $i \in [M]$, let \mathcal{T}_i denote the set of output tuples y^n that result in MAP decoding ties:

$$\mathcal{T}_i \triangleq \left\{ y^n \in \{0, 1\}^n : P_{X^n|Y^n}(c_i|y^n) = \max_{r \in [M] \setminus \{i\}} P_{X^n|Y^n}(c_r|y^n) \right\} \quad (11)$$

$$= \left\{ y^n \in \{0, 1\}^n : P_{X^n, Y^n}(c_i, y^n) = \max_{r \in [M] \setminus \{i\}} P_{X^n, Y^n}(c_r, y^n) \right\}, \quad (12)$$

where (12) holds because $P_{X^n|Y^n}(x^n|y^n) = \frac{P_{X^n, Y^n}(x^n, y^n)}{P_{Y^n}(y^n)}$. Then, δ_n in (6) can be rewritten as

$$\delta_n = \sum_{i \in [M]} P_{X^n}(c_i) P_{Y^n|X^n}(\mathcal{T}_i|c_i) = \sum_{i \in [M]} P_{X^n, Y^n}(c_i, \mathcal{T}_i). \quad (13)$$

Similarly, let \mathcal{N}_i denote the set of output tuples y^n which guarantee a tie-free MAP decoding error when c_i is transmitted over the channel:

$$\mathcal{N}_i \triangleq \left\{ y^n \in \{0, 1\}^n : P_{X^n|Y^n}(c_i|y^n) < \max_{r \in [M] \setminus \{i\}} P_{X^n|Y^n}(c_r|y^n) \right\} \quad (14)$$

$$= \left\{ y^n \in \{0, 1\}^n : P_{X^n, Y^n}(c_i, y^n) < \max_{r \in [M] \setminus \{i\}} P_{X^n, Y^n}(c_r, y^n) \right\}. \quad (15)$$

Hence, b_n in (5) can be rewritten as:

$$b_n = \sum_{i \in [M]} P_{X^n}(c_i) P_{Y^n|X^n}(\mathcal{N}_i|c_i) = \sum_{i \in [M]} P_{X^n, Y^n}(c_i, \mathcal{N}_i). \quad (16)$$

Note if $\delta_n = 0$, then (7) is tight and (10) holds trivially; so, without loss of generality, we assume in the proof that $\delta_n > 0$, which implies that there exists at least one non-empty \mathcal{T}_i for $i \in [M]$. Then, according to (13) and (16), we have that

$$\frac{\delta_n}{b_n} = \frac{\sum_{i \in [M]} P_{X^n, Y^n}(c_i, \mathcal{T}_i)}{\sum_{i \in [M]} P_{X^n, Y^n}(c_i, \mathcal{N}_i)} \leq \frac{\sum_{i \in [M]: \mathcal{T}_i \neq \emptyset} P_{X^n, Y^n}(c_i, \mathcal{T}_i)}{\sum_{i \in [M]: \mathcal{T}_i \neq \emptyset} P_{X^n, Y^n}(c_i, \mathcal{N}_i)}. \quad (17)$$

We can upper-bound (17) by

$$\frac{\sum_{i \in [M]: \mathcal{T}_i \neq \emptyset} P_{X^n, Y^n}(c_i, \mathcal{T}_i)}{\sum_{i \in [M]: \mathcal{T}_i \neq \emptyset} P_{X^n, Y^n}(c_i, \mathcal{N}_i)} \leq \max_{1 \in [M]: \mathcal{T}_i \neq \emptyset} \frac{P_{X^n, Y^n}(c_i, \mathcal{T}_i)}{P_{X^n, Y^n}(c_i, \mathcal{N}_i)}, \quad (18)$$

where for convenience we will refer to an inequality of the form given in (18) as the *ratio-sum* inequality. As a result, Theorem 2 holds if we can substantiate that $2qn$ is an upper bound for (18). To this end, we will find a *proper partition* of \mathcal{T}_i and an *equal number* of disjoint subsets of \mathcal{N}_i , of which the individual probabilities can be evaluated. For ease of notation, we denote the individual probabilities corresponding to the K -partition of \mathcal{T}_i and K disjoint subsets of \mathcal{N}_i by $\{\alpha_k\}_{k=1}^K$ and $\{\beta_k\}_{k=1}^K$, respectively. Then, we obtain that

$$\frac{P_{X^n, Y^n}(\mathbf{c}_i, \mathcal{T}_i)}{P_{X^n, Y^n}(\mathbf{c}_i, \mathcal{N}_i)} \leq \frac{\sum_{k=1}^K \alpha_k}{\sum_{k=1}^K \beta_k}. \tag{19}$$

By showing that each individual ratio $\alpha_k/\beta_k, k \in [K]$ is bounded above by $2qn$, the ratio-sum inequality can again be applied to complete the proof.

3. Proof of Theorem 2

In [1], where a uniformly distributed prior probability P_{X^n} over \mathcal{C}_n is assumed, one can flip a properly selected bit in the output $y^n \in \mathcal{T}_i$ to convert it to a corresponding element in \mathcal{N}_i . In light of this connection, one can evaluate the ratio $\frac{P_{Y^n|X^n}(\mathcal{T}_i|\mathbf{c}_i)}{P_{Y^n|X^n}(\mathcal{N}_i|\mathbf{c}_i)}$. This approach, however, no longer works when a non-uniformly distributed prior probability is considered. Therefore, we have to devise a more judicious approach to extend the result in [1] for a general prior probability.

3.1. A Partition of Non-empty \mathcal{T}_i and Corresponding Disjoint Subsets of \mathcal{N}_i

In this section, instead of finding a disjoint covering of the set of decoder ties \mathcal{T}_i as in [1], we establish a proper partition of \mathcal{T}_i from Definitions 1 and 2. This is one of the key differences from the techniques used in [1]. Example 1 is given after Proposition 1 to illustrate Definitions 1 and 2.

Given $y^n \in \mathcal{T}_i$ defined in (12), there exists at least one $m \in [M] \setminus \{i\}$ such that

$$P_{X^n, Y^n}(\mathbf{c}_i, y^n) = P_{X^n, Y^n}(\mathbf{c}_m, y^n) = \max_{r \in [M] \setminus \{i\}} P_{X^n, Y^n}(\mathbf{c}_r, y^n). \tag{20}$$

We collect the indices m that satisfy (20) in $\mathcal{I}_i(y^n)$ as follows:

$$\mathcal{I}_i(y^n) \triangleq \left\{ h \in [M] \setminus \{i\} : P_{X^n, Y^n}(\mathbf{c}_i, y^n) = P_{X^n, Y^n}(\mathbf{c}_h, y^n) = \max_{r \in [M] \setminus \{i\}} P_{X^n, Y^n}(\mathbf{c}_r, y^n) \right\}. \tag{21}$$

Remark 2. First, we note that $\mathcal{I}_i(y^n)$ is not empty as long as $y^n \in \mathcal{T}_i$. Also, for any $y^n \in \mathcal{T}_i$, we can infer from (21) that $h \in \mathcal{I}_i(y^n)$ if and only if $y^n \in \mathcal{T}_h$.

In Definitions 1 and 2 that follow, we will assign each $y^n \in \mathcal{T}_i$ to a subset indexed by $j \in \mathcal{I}_i(y^n)$. These subsets will form a partition of \mathcal{T}_i as stated in Proposition 1.

Definition 1. For $j \in [M] \setminus \{i\}$, denoting by $\mathcal{S}_{i,j}$ the set of indices where the bit components of \mathbf{c}_i and \mathbf{c}_j differ, we define

$$\left\{ \mathcal{T}_{j|i} \triangleq \left\{ y^n \in \mathcal{T}_i : j = \min_{r \in \mathcal{I}_i(y^n) : d(\mathbf{c}_i, y^n | \mathcal{S}_{i,r}) < |\mathcal{S}_{i,r}|} r \right\}; \tag{22a}$$

$$\left\{ \mathcal{N}_{j|i} \triangleq \left\{ y^n \in \mathcal{N}_i : P_{X^n, Y^n}(\mathbf{c}_i, y^n) \cdot q = P_{X^n, Y^n}(\mathbf{c}_j, y^n) \cdot \frac{1}{q} \right. \right. \\ \left. \left. \text{and } P_{X^n, Y^n}(\mathbf{c}_j, y^n) \neq P_{X^n, Y^n}(\mathbf{c}_r, y^n) \text{ for } r \in [j-1] \setminus \{i\} \right\} \right\}. \tag{22b}$$

Since there may exist $y^n \in \mathcal{T}_i$ satisfying $d(c_i, y^n | \mathcal{S}_{i,r}) = |\mathcal{S}_{i,r}|$ for all $r \in \mathcal{I}_i(y^n)$, the collection of all elements in $\bigcup_{j \in [M] \setminus \{i\}} \mathcal{T}_{j|i}$ may not exhaust the elements in \mathcal{T}_i (see Example 1). We thus go on to collect the remaining elements in $\mathcal{T}_i \setminus \bigcup_{j \in [M] \setminus \{i\}} \mathcal{T}_{j|i}$ as follows.

Definition 2. Define for $j \in [M] \setminus \{i\}$,

$$\mathcal{T}_{j|i} \triangleq \left\{ y^n \in \mathcal{T}_i \setminus \left(\bigcup_{h \in [M] \setminus \{i\}} \mathcal{T}_{h|i} \right) : j = \min_{r \in \mathcal{I}_i(y^n)} r \right\}. \tag{23}$$

With the sets defined in Definitions 1 and 2, a partition of \mathcal{T}_i and disjoint subsets of \mathcal{N}_i are constructed as proven in the following proposition.

Proposition 1. For non-empty \mathcal{T}_i , the following two properties hold.

- (i) The collection $\{\mathcal{T}_{j|i} \cup \mathcal{T}_{j|i}\}_{j \in [M] \setminus \{i\}}$ forms a (disjoint) partition of \mathcal{T}_i .
- (ii) $\{\mathcal{N}_{j|i}\}_{j \in [M] \setminus \{i\}}$ is a collection of disjoint subsets of \mathcal{N}_i .

Before proving Proposition 1, we provide the following example to illustrate the above sets.

Example 1. This example illustrates the necessity of introducing $\mathcal{T}_{j|i}$ as a companion to $\mathcal{T}_{j|i}$. Suppose $\mathcal{C}_4 = \{c_1, c_2, c_3, c_4\} = \{0000, 0101, 0110, 0111\}$. Let $P_{X^4}(c_1) = \frac{q^2}{2+q^2+q^{-2}}$, $P_{X^4}(c_2) = P_{X^4}(c_3) = \frac{1}{2+q^2+q^{-2}}$ and $P_{X^4}(c_4) = \frac{q^{-2}}{2+q^2+q^{-2}}$. Then, $y^4 = 0111$ satisfies

$$\underbrace{P_{X^4, Y^4}(c_1, y^4)}_{\frac{q^2}{(2+q^2+q^{-2})} p^4 q^1} = \max_{r \in [4] \setminus \{1\}} P_{X^4, Y^4}(c_r, y^4) = \underbrace{P_{X^4, Y^4}(c_2, y^4)}_{\frac{1}{(2+q^2+q^{-2})} p^4 q^3} = \underbrace{P_{X^4, Y^4}(c_3, y^4)}_{\frac{1}{(2+q^2+q^{-2})} p^4 q^3} > \underbrace{P_{X^4, Y^4}(c_4, y^4)}_{\frac{q^{-2}}{(2+q^2+q^{-2})} p^4 q^4}, \tag{24}$$

where the probabilities $P_{X^n, Y^n}(x^n, y^n)$ are written in the form

$$P_{X^n, Y^n}(x^n, y^n) = P_{X^n}(x^n) P_{Y^n | X^n}(y^n | x^n) = P_{X^n}(x^n) p^n q^{n-d(x^n, y^n)}. \tag{25}$$

Note that the first equality in (24) indicates $0111 \in \mathcal{T}_1$ and the last two equalities and the right-most inequality jointly imply $\mathcal{I}_1(0111) = \{2, 3\}$. In light of Proposition 1, this 0111 must lie in one and only one of $\{\mathcal{T}_{j|1} \cup \mathcal{T}_{j|1}\}_{j \in [4] \setminus \{1\}}$ as shown in Tables A1 and A2 of Appendix A. Since there exist no integers h in $\mathcal{I}_1(0111)$ fulfilling $d(c_1, 0111 | \mathcal{S}_{1,h}) < |\mathcal{S}_{1,h}|$, this 0111 belongs to $\mathcal{T}_{j|1}$ with $j = \min_{r \in \mathcal{I}_1(0111)} r = 2$. Recall that in [1], an element in $\mathcal{N}_{j|1}$ can be obtained if flipping a zero of $y^n \in \mathcal{T}_{j|1}$ can make it further away from c_1 but closer to c_j . However, for $y^4 = 0111$ in this example if we flip the only zero to one, it gets further away from both c_1 and c_h for any $h = 2, 3, 4$. Therefore, the bit-flipping manipulation fails.

With $y^4 = 0111$, we also have

$$\underbrace{P_{X^4, Y^4}(c_2, y^4)}_{\frac{1}{(2+q^2+q^{-2})} p^4 q^3} = \max_{r \in [4] \setminus \{2\}} P_{X^4, Y^4}(c_r, y^4) = \underbrace{P_{X^4, Y^4}(c_1, y^4)}_{\frac{q^2}{(2+q^2+q^{-2})} p^4 q} = \underbrace{P_{X^4, Y^4}(c_3, y^4)}_{\frac{1}{(2+q^2+q^{-2})} p^4 q^3} > \underbrace{P_{X^4, Y^4}(c_4, y^4)}_{\frac{1}{(2+q^2+q^{-2})} p^4 q^4}, \tag{26}$$

where the first equality indicates $0111 \in \mathcal{T}_2$ and the remaining parts in (26) jointly imply that $\mathcal{I}_2(0111) = \{1, 3\}$. Proposition 1 then states that this 0111 lies in one and only one of $\{\mathcal{T}_{j|2} \cup \mathcal{T}_{j|2}\}_{j \in [4] \setminus \{2\}}$. Since $1 \in \mathcal{I}_2(0111) = \{1, 3\}$ and $d(c_2, 0111 | \mathcal{S}_{2,1}) = 0 < |\mathcal{S}_{2,1}| = 2$, we have $0111 \in \mathcal{T}_{1|2}$ according to (22a). Thus, we can flip a bit in 0111 to get further away from c_2 and closer to c_j simultaneously. More specifically, the bit-flipping manipulation produces either 0110 or 0011, which lies in $\mathcal{N}_{1|2}$ as $y^4 = 0111$ is in $\mathcal{T}_{1|2}$. Therefore, we can associate the element in $\mathcal{T}_{1|2}$ with an element in $\mathcal{N}_{1|2}$ via a single flipping operation. For completeness, a full list of the sets \mathcal{T}_i , \mathcal{N}_i , $\mathcal{T}_{j|i}$, $\mathcal{T}_{j|i}$ and $\mathcal{N}_{j|i}$ for $i \in [4]$ and $j \in [4] \setminus \{i\}$, is given in Appendix A.

Proof of Proposition 1. First, we note that by the definitions in (22a) and (23), $\{\mathcal{T}_{j|i}\}_{j \in [M] \setminus \{i\}}$ are disjoint and so is $\{\mathcal{S}_{j|i}\}_{j \in [M] \setminus \{i\}}$. Additionally, (23) implies $\mathcal{T}_{j|i} \cap \mathcal{S}_{h|i} = \emptyset$ for arbitrary $j, h \in [M] \setminus \{i\}$. Furthermore, according to Definitions 1 and 2, for any $y^n \in \mathcal{T}_i$, we have either $y^n \in \mathcal{T}_{h|i}$ or $y^n \in \mathcal{S}_{h|i}$ for some $h \in [M] \setminus \{i\}$. Consequently, $\{\mathcal{T}_{j|i} \cup \mathcal{S}_{j|i}\}_{j \in [M] \setminus \{i\}}$ forms a partition of \mathcal{T}_i .

On the other hand, the inequality in (22b) prevents multiple inclusions of an element from the previous collections. Therefore, $\{\mathcal{N}_{j|i}\}_{j \in [M] \setminus \{i\}}$ are a collection of disjoint subsets of \mathcal{N}_i . \square

Remark 3. When channel inputs are uniformly distributed as considered in [1], it follows that

$$\mathcal{I}_i(y^n) = \left\{ h \in [M] \setminus \{i\} : d(c_i, y^n) = d(c_h, y^n) = \max_{r \in [M] \setminus \{i\}} d(c_r, y^n) \right\}, \tag{27}$$

and $d(c_i, y^n | \mathcal{S}_{i,j}) = \frac{|\mathcal{S}_{i,j}|}{2} < |\mathcal{S}_{i,j}|$ for every $j \in \mathcal{I}_i(y^n)$. Therefore, (22a) is reduced to

$$\mathcal{T}_{j|i} = \left\{ y^n \in \mathcal{T}_i : j = \min_{r \in \mathcal{I}_i(y^n)} r \right\}, \tag{28}$$

and

$$\mathcal{S}_{j|i} = \emptyset. \tag{29}$$

We then have the following two remarks. First, we note that the $\mathcal{T}_{j|i}$ newly defined via (22a) and reduced to (28) in the regime considered in [1] is more restrictive than the $\mathcal{T}_{j|i}$ introduced in [1] (Equation (16a)). As a consequence, $\{\mathcal{T}_{j|i}\}_{j \in [M] \setminus \{i\}}$ forms a partition of \mathcal{T}_i in this paper while those introduced in [1] (Equation (16a)) are a disjoint covering of \mathcal{T}_i under uniform channel inputs. Second, (29) shows that [1] does not need to consider a companion $\mathcal{S}_{j|i}$ to $\mathcal{T}_{j|i}$, but this paper does.

Based on Proposition 1, we continue the derivation from (17) and obtain:

$$\frac{\delta_n}{b_n} \leq \frac{\sum_{i \in [M]} P_{X^n, Y^n} \left(c_i, \bigcup_{j \in [M] \setminus \{i\}} (\mathcal{T}_{j|i} \cup \mathcal{S}_{j|i}) \right)}{\sum_{i \in [M]} P_{X^n, Y^n} \left(c_i, \bigcup_{j \in [M] \setminus \{i\}} \mathcal{N}_{j|i} \right)} \tag{30}$$

$$= \frac{\sum_{i \in [M]} \sum_{j \in [M] \setminus \{i\}} P_{X^n, Y^n} (c_i, \mathcal{T}_{j|i}) + \sum_{i \in [M]} \sum_{j \in [M] \setminus \{i\}} P_{X^n, Y^n} (c_i, \mathcal{S}_{j|i})}{\sum_{i \in [M]} \sum_{j \in [M] \setminus \{i\}} P_{X^n, Y^n} (c_i, \mathcal{N}_{j|i})}, \tag{31}$$

where (31) holds because $\{\mathcal{T}_{j|i}\}_{j \in [M] \setminus \{i\}}$ and $\{\mathcal{S}_{j|i}\}_{j \in [M] \setminus \{i\}}$ are disjoint, and the same applies to $\{\mathcal{N}_{j|i}\}_{j \in [M] \setminus \{i\}}$. An additional upper bound for (31) requires the verification of the inequality:

$$\sum_{i \in [M]} \sum_{j \in [M] \setminus \{i\}} P_{X^n, Y^n} (c_i, \mathcal{S}_{j|i}) \leq \sum_{j \in [M]} \sum_{i \in [M] \setminus \{j\}} P_{X^n, Y^n} (c_j, \mathcal{T}_{i|j}), \tag{32}$$

which is an immediate consequence of the proposition to be proven in the next section (Proposition 2), stating that

$$y^n \in \mathcal{S}_{j|i} \text{ and } h \in \mathcal{I}_i(y^n) \Rightarrow y^n \in \mathcal{T}_{\ell|h} \text{ for some } \ell \in \mathcal{I}_h(y^n) \text{ and } P_{X^n, Y^n} (c_i, y^n) = P_{X^n, Y^n} (c_h, y^n). \tag{33}$$

3.2. Verification of (32)

Recall that the main technique used in [1] is to associate every element in \mathcal{T}_i with a corresponding element in \mathcal{N}_i via the bit-flipping manipulation. By this bit-flipping association, the probability ratio of the elements and corresponding elements respectively in \mathcal{T}_i and \mathcal{N}_i can be evaluated. However, as Example 1 indicates, for an element in $\mathcal{S}_{j|i}$,

the bit-flipping association no longer works. This reveals the challenge of generalizing the results in [1] from uniform channel inputs to arbitrarily distributed channel inputs. A solution is to subdivide the elements in \mathcal{T}_i into two groups $\{\mathcal{T}_{ji}\}_{j \in [M] \setminus \{i\}}$ and $\{\mathcal{S}_{ji}\}_{j \in [M] \setminus \{i\}}$, where the bit-flipping association to $\{\mathcal{N}_{ji}\}_{j \in [M] \setminus \{i\}}$ works for the former group but not for the latter. The inequality in (32) can then be used to exclude the latter group with an upper bound:

$$\frac{\delta_n}{b_n} \leq \frac{\sum_{i \in [M]} \sum_{j \in [M] \setminus \{i\}} P_{X^n, Y^n}(\mathbf{c}_i, \mathcal{T}_{ji}) + \sum_{i \in [M]} \sum_{j \in [M] \setminus \{i\}} P_{X^n, Y^n}(\mathbf{c}_i, \mathcal{S}_{ji})}{\sum_{i \in [M]} \sum_{j \in [M] \setminus \{i\}} P_{X^n, Y^n}(\mathbf{c}_i, \mathcal{N}_{ji})} \tag{34}$$

$$\leq 2 \frac{\sum_{i \in [M]} \sum_{j \in [M] \setminus \{i\}} P_{X^n, Y^n}(\mathbf{c}_i, \mathcal{T}_{ji})}{\sum_{i \in [M]} \sum_{j \in [M] \setminus \{i\}} P_{X^n, Y^n}(\mathbf{c}_i, \mathcal{N}_{ji})}. \tag{35}$$

Since uniform channel inputs as considered in [1] guarantee (29), it can be seen from (35) that the multiplicative factor of 2 can be reduced to 1 as observed in Remark 1. For general arbitrary channel inputs, we have the factor of 2 since the set \mathcal{S}_{ji} may not be empty. The validity of (32) can be confirmed by the next proposition.

Proposition 2. *Suppose $y^n \in \mathcal{S}_{ji}$. Then, for every $h \in \mathcal{I}_i(y^n)$, we have*

$$y^n \in \mathcal{T}_{\ell|h} \text{ for some } \ell \in \mathcal{I}_h(y^n) \text{ and } P_{X^n, Y^n}(\mathbf{c}_i, y^n) = P_{X^n, Y^n}(\mathbf{c}_h, y^n). \tag{36}$$

Proof. Suppose $y^n \in \mathcal{S}_{ji}$. Then, $d(\mathbf{c}_i, y^n | \mathcal{S}_{i,h}) = |\mathcal{S}_{i,h}|$ for every $h \in \mathcal{I}_i(y^n)$. We therefore have:

$$P_{X^n, Y^n}(\mathbf{c}_i, y^n) = P_{X^n, Y^n}(\mathbf{c}_h, y^n) = \max_{r \in [M] \setminus \{i\}} P_{X^n, Y^n}(\mathbf{c}_r, y^n). \tag{37}$$

We can rewrite (37) as

$$P_{X^n, Y^n}(\mathbf{c}_h, y^n) = P_{X^n, Y^n}(\mathbf{c}_i, y^n) = \max_{r \in [M] \setminus \{h\}} P_{X^n, Y^n}(\mathbf{c}_r, y^n), \tag{38}$$

implying $y^n \in \mathcal{T}_h$ and $i \in \mathcal{I}_h(y^n)$. Noting that $d(\mathbf{c}_h, y^n | \mathcal{S}_{h,i}) = 0 < |\mathcal{S}_{h,i}|$ because $d(\mathbf{c}_i, y^n | \mathcal{S}_{i,h}) = |\mathcal{S}_{i,h}|$ and $S_{h,i} = S_{i,h}$, we conclude that the smallest integer $\ell \in \mathcal{I}_h(y^n)$ satisfying $d(\mathbf{c}_h, y^n | \mathcal{S}_{h,\ell}) < |\mathcal{S}_{h,\ell}|$ exists, and therefore $y^n \in \mathcal{T}_{\ell|h}$. \square

Remark 4. *Two observations can be made based on Proposition 2. First, Proposition 2 indicates that every $y^n \in \mathcal{S}_{ji}$ must appear at least once in the sum $\sum_{h \in [M]} \sum_{\ell \in [M] \setminus \{h\}} P_{X^n, Y^n}(\mathbf{c}_h, \mathcal{T}_{\ell|h})$, contributing the same probability mass $P_{X^n, Y^n}(\mathbf{c}_h, y^n)$ as $P_{X^n, Y^n}(\mathbf{c}_i, y^n)$. Second, Proposition 2 also implies that every $y^n \in \mathcal{S}_{ji}$ cannot be contained in $(\bigcup_{h \in [M]} \bigcup_{r \in [M] \setminus \{h\}} \mathcal{T}_{r|h}) \setminus \mathcal{S}_{ji}$. This observation can be substantiated as follows. For every $h \in \mathcal{I}_i(y^n)$, Proposition 2 implies $y^n \in \mathcal{T}_{\ell|h}$ for some $\ell \in \mathcal{I}_h(y^n)$ and hence Definition 2 immediately gives $y^n \notin \mathcal{T}_{r|h}$ for all $r \in [M] \setminus \{h\}$. For $h \notin \mathcal{I}_i(y^n)$, we have $y^n \notin \mathcal{T}_h$ and therefore $y^n \notin \mathcal{T}_{r|h} \subseteq \mathcal{T}_h$ for all $r \in [M] \setminus \{h\}$ as pointed out in Remark 2. As a result, every $y^n \in \mathcal{S}_{ji}$ appears exactly once in the sum $\sum_{h \in [M]} \sum_{r \in [M] \setminus \{h\}} P_{X^n, Y^n}(\mathbf{c}_h, \mathcal{T}_{r|h})$. Combining the two observations leads to:*

$$\sum_{i \in [M]} \sum_{j \in [M] \setminus \{i\}} \sum_{y^n \in \mathcal{S}_{ji}} P_{X^n, Y^n}(\mathbf{c}_i, y^n) \leq \sum_{j \in [M]} \sum_{\ell \in [M] \setminus \{j\}} \sum_{y^n \in \mathcal{T}_{\ell|j}} P_{X^n, Y^n}(\mathbf{c}_j, y^n). \tag{39}$$

To flesh out the above inequality, we give the next example.

Example 2. *Proceeding from Example 1, we observe from Tables A1 and A2 in Appendix A that 0111 is contained in $\mathcal{T}_{2|1}$, $\mathcal{T}_{1|2}$ and $\mathcal{T}_{1|3}$. Hence, it appears once in the sum $\sum_{j \in [4]} \sum_{i \in [4] \setminus \{j\}} P_{X^n, Y^n}(\mathbf{c}_j, \mathcal{S}_{ij})$ while it contributes twice in the sum $\sum_{j \in [4]} \sum_{i \in [4] \setminus \{j\}} P_{X^n, Y^n}(\mathbf{c}_j, \mathcal{T}_{ij})$. We then confirm from (A35) that:*

$$\sum_{i \in [4]} \sum_{j \in [4] \setminus \{i\}} P_{X^4, Y^4}(c_i, \mathcal{T}_{j|i}) \geq \sum_{i \in [4]} \sum_{j \in [4] \setminus \{i\}} P_{X^4, Y^4}(c_i, \mathcal{T}_{j|i}). \tag{40}$$

We continue the derivation from (35) and obtain

$$\frac{\delta_n}{b_n} \leq 2 \frac{\sum_{i \in [M]} \sum_{j \in [M] \setminus \{i\}: \mathcal{T}_{j|i} \neq \emptyset} P_{X^n, Y^n}(c_i, \mathcal{T}_{j|i})}{\sum_{i \in [M]} \sum_{j \in [M] \setminus \{i\}: \mathcal{T}_{j|i} \neq \emptyset} P_{X^n, Y^n}(c_i, \mathcal{N}_{j|i})} \tag{41}$$

$$\leq 2 \max_{i \in [M] \text{ and } j \in [M] \setminus \{i\}: \mathcal{T}_{j|i} \neq \emptyset} \frac{P_{X^n, Y^n}(c_i, \mathcal{T}_{j|i})}{P_{X^n, Y^n}(c_i, \mathcal{N}_{j|i})}, \tag{42}$$

where we add the restriction $\mathcal{T}_{j|i} \neq \emptyset$ in (41) to exclude the cases of zero dividing by zero in (42), and (42) follows the ratio-sum inequality in (18).

In the next section, we introduce a number of delicate decompositions of non-empty $\mathcal{T}_{j|i}$ and an equal number of disjoint subsets of $\mathcal{N}_{j|i}$ to facilitate the bit-flipping association of the pairs.

3.3. Atomic Decomposition of Non-Empty $\mathcal{T}_{j|i}$ and the Corresponding Disjoint Subsets of $\mathcal{N}_{j|i}$

To simplify the exposition, we assume without loss of generality that c_1 is the all-zero codeword (It is known that we can simultaneously flip the same position of all codewords to yield a new code of equal performance over the BSC. Thus, via a number of flipping manipulations, we can transform any code to a code of equal performance with the first codeword being all-zero.) Below we present the proof for $i = 1$. The proof for general $i > 1$ follows analogously.

Since c_1 is the all-zero codeword, $\mathcal{S}_{1,j}$ is the set containing the indices of the non-zero components of c_j . To facilitate the investigation of the structure of c_j relative to the remaining codewords $\{c_r\}_{r \in [M] \setminus \{1,j\}}$, we first partition $\mathcal{S}_{1,j}$ into 2^{M-2} subsets according to whether each index in $\mathcal{S}_{1,j}$ is in $\mathcal{S}_{1,2}, \dots, \mathcal{S}_{1,j-1}, \mathcal{S}_{1,j+1}, \dots, \mathcal{S}_{1,M}$ or not, as follows:

$$\mathcal{S}_{1,j}^{(m)} \triangleq \left(\bigcap_{r=2}^{j-1} \mathcal{S}_{r,\lambda_r} \right) \cap \left(\bigcap_{r=j+1}^M \mathcal{S}_{r,\lambda_r} \right) \cap \mathcal{S}_{1,j} \quad \text{for } m \triangleq 1 + \sum_{r=2}^{j-1} \lambda_r \cdot 2^{r-2} + \sum_{r=j+1}^M \lambda_r \cdot 2^{r-3}, \tag{43}$$

where $\mathcal{S}_{r,1} \triangleq \mathcal{S}_{1,r}$ and $\mathcal{S}_{r,0} \triangleq [n] \setminus \mathcal{S}_{1,r} = \mathcal{S}_{1,r}^c$, and each $\lambda_r \in \{0, 1\}$. An example of the partition is given below.

Example 3. Suppose $\mathcal{C}_4 = \{00000, 11001, 01111, 01101\}$. For $j = 3$ and $\mathcal{S}_{1,j} = \{2, 3, 4, 5\}$, we obtain $2^{4-2} = 4$ subsets as

$$\mathcal{S}_{1,3}^{(m)} = \begin{cases} \mathcal{S}_{1,3}^{(1)} = \mathcal{S}_{1,2}^c \cap \mathcal{S}_{1,4}^c \cap \mathcal{S}_{1,3} = \{4\}, & \text{if } (\lambda_4, \lambda_2) = (0, 0); \\ \mathcal{S}_{1,3}^{(2)} = \mathcal{S}_{1,2} \cap \mathcal{S}_{1,4}^c \cap \mathcal{S}_{1,3} = \emptyset, & \text{if } (\lambda_4, \lambda_2) = (0, 1); \\ \mathcal{S}_{1,3}^{(3)} = \mathcal{S}_{1,2}^c \cap \mathcal{S}_{1,4} \cap \mathcal{S}_{1,3} = \{3\}, & \text{if } (\lambda_4, \lambda_2) = (1, 0); \\ \mathcal{S}_{1,3}^{(4)} = \mathcal{S}_{1,2} \cap \mathcal{S}_{1,4} \cap \mathcal{S}_{1,3} = \{2, 5\}, & \text{if } (\lambda_4, \lambda_2) = (1, 1). \end{cases} \tag{44}$$

As c_1 is the all-zero codeword, the components of c_r with indices in $\mathcal{S}_{1,j}^{(m)}$ can now be unambiguously identified and must all equal λ_r . As a result,

$$d(c_1, c_r | \mathcal{S}_{1,j}^{(m)}) = \begin{cases} |\mathcal{S}_{1,j}^{(m)}|, & \lambda_r = 1; \\ 0, & \lambda_r = 0. \end{cases} \tag{45}$$

Example 4. Proceeding from Example 3, we have

$$\begin{cases} d(c_1, c_2 | \mathcal{S}_{1,3}^{(1)}) = 0 & \text{because } \lambda_2 = 0; \\ d(c_1, c_2 | \mathcal{S}_{1,3}^{(2)}) = |\mathcal{S}_{1,3}^{(2)}| = 0 & \text{because } \lambda_2 = 1; \\ d(c_1, c_2 | \mathcal{S}_{1,3}^{(3)}) = 0 & \text{because } \lambda_2 = 0; \\ d(c_1, c_2 | \mathcal{S}_{1,3}^{(4)}) = |\mathcal{S}_{1,3}^{(4)}| = 2 & \text{because } \lambda_2 = 1, \end{cases} \tag{46}$$

and

$$\begin{cases} d(c_1, c_4 | \mathcal{S}_{1,3}^{(1)}) = 0 & \text{because } \lambda_4 = 0; \\ d(c_1, c_4 | \mathcal{S}_{1,3}^{(2)}) = 0 & \text{because } \lambda_4 = 0; \\ d(c_1, c_4 | \mathcal{S}_{1,3}^{(3)}) = |\mathcal{S}_{1,3}^{(3)}| = 1 & \text{because } \lambda_4 = 1; \\ d(c_1, c_4 | \mathcal{S}_{1,3}^{(4)}) = |\mathcal{S}_{1,3}^{(4)}| = 2 & \text{because } \lambda_4 = 1. \end{cases} \tag{47}$$

It should be emphasized that $\mathcal{S}_{1,j}^{(m)}$ in this paper is defined differently from that in [1]. While the one defined in [1] partitions $\mathcal{S}_{1,j}$ only according to codewords with indices less than j , the one defined in this paper considers all other $M - 2$ codewords in the partition manipulation, and hence the order of codewords becomes irrelevant.

Next, to decompose $\mathcal{T}_{j|1}$, we further define a sequence of incremental sets:

$$\mathcal{S}_{1,j}^{(m)} \triangleq \bigcup_{h=1}^m \mathcal{S}_{1,j}^{(h)}, \quad m \in [2^{M-2}], \tag{48}$$

and set $\mathcal{S}_{1,j}^{(0)} \triangleq \emptyset$. Let $\ell_{1,j} \triangleq |\mathcal{S}_{1,j}|$ and $\ell_{1,j}^{(m)} \triangleq |\mathcal{S}_{1,j}^{(m)}|$ respectively denote the sizes of $\mathcal{S}_{1,j}$ and $\mathcal{S}_{1,j}^{(m)}$ and note that $0 = \ell_{1,j}^{(0)} \leq \ell_{1,j}^{(1)} \leq \ell_{1,j}^{(2)} \leq \dots \leq \ell_{1,j}^{(2^{M-2})} = \ell_{1,j}$.

The idea behind the partition of $\mathcal{T}_{j|1}$ into $\ell_{1,j}$ subsets, indexed by $k \in [\ell_{1,j} - 1] \cup \{0\}$, is as follows. Pick one $y^n \in \mathcal{T}_{j|1}$. We start by examining whether $d(c_1, y^n | \mathcal{S}_{1,j}^{(1)})$ is strictly less than $\ell_{1,j}^{(1)} - 1$. If the answer is negative, we continue examining whether $d(c_1, y^n | \mathcal{S}_{1,j}^{(2)})$ is strictly less than $\ell_{1,j}^{(2)} - 1$. Proceed until we reach the smallest m such that $d(c_1, y^n | \mathcal{S}_{1,j}^{(m)}) < \ell_{1,j}^{(m)} - 1$ holds. Setting k to be equal to $k = d(c_1, y^n | \mathcal{S}_{1,j}^{(m)})$, we assign this y^n to the subset $\mathcal{T}_{j|1}(k)$. Notably, there exists no such number $m \in [2^{M-2}]$ that satisfies $d(c_1, y^n | \mathcal{S}_{1,j}^{(m)}) < \ell_{1,j}^{(m)} - 1$ if and only if $d(c_1, y^n | \mathcal{S}_{1,j}) = \ell_{1,j} - 1$; in this case, we find the smallest m satisfying $\mathcal{S}_{1,j}^{(m)} = \mathcal{S}_{1,j}$ and assign this element to $\mathcal{T}_{j|1}(\ell_{1,j} - 1)$ as $d(c_1, y^n | \mathcal{S}_{1,j}^{(m)}) = \ell_{1,j} - 1$. For ease of describing the above algorithmic partition process, we introduce a mapping from $k \in [\ell_{1,j} - 1] \cup \{0\}$ to $m \in [2^{M-2}]$ as follows:

$$\eta_k \triangleq \begin{cases} \min \{ m \in [2^{M-2}] : k < \ell_{1,j}^{(m)} - 1 \}, & 0 \leq k < \ell_{1,j} - 1; \\ \min \{ m \in [2^{M-2}] : k = \ell_{1,j}^{(m)} - 1 \}, & k = \ell_{1,j} - 1. \end{cases} \tag{49}$$

We can see that for $0 \leq k < \ell_{1,j} - 1$, we have $\ell_{1,j}^{(\eta_k - 1)} - 1 \leq k < \ell_{1,j}^{(\eta_k)} - 1$. Therefore, if $y^n \in \mathcal{T}_{j|1}$ is assigned to $\mathcal{T}_{j|1}(k)$ for some $k < \ell_{1,j} - 1$, we must have

$$\ell_{1,j}^{(\eta_k - 1)} - 1 \leq d(c_1, y^n | \mathcal{S}_{1,j}^{(\eta_k - 1)}) \leq d(c_1, y^n | \mathcal{S}_{1,j}^{(\eta_k)}) = k < \ell_{1,j}^{(\eta_k)} - 1. \tag{50}$$

On the other hand, if $y^n \in \mathcal{T}_{j|1}$ is collected in $\mathcal{T}_{j|1}(\ell_{1,j} - 1)$, then $\mathcal{S}_{1,j}^{(\eta_k)} = \mathcal{S}_{1,j}$ and

$$\ell_{1,j}^{(\eta_{k-1})} - 1 \leq d(\mathbf{c}_1, y^n | \mathcal{S}_{1,j}^{(\eta_{k-1})}) \leq d(\mathbf{c}_1, y^n | \mathcal{S}_{1,j}^{(\eta_k)}) = \ell_{1,j} - 1. \tag{51}$$

A formal definition of $\mathcal{T}_{j|1}(k)$ is given next, where the corresponding subsets $\mathcal{N}_{j|1}(k)$ of $\mathcal{N}_{j|1}$ are also introduced.

Definition 3. Define for $k = 0, 1, \dots, \ell_{1,j} - 1$,

$$\left\{ \mathcal{T}_{j|1}(k) \triangleq \left\{ y^n \in \mathcal{T}_{j|1} : \ell_{1,j}^{(\eta_{k-1})} - 1 \leq d(\mathbf{c}_1, y^n | \mathcal{S}_{1,j}^{(\eta_{k-1})}) \text{ and } d(\mathbf{c}_1, y^n | \mathcal{S}_{1,j}^{(\eta_k)}) = k \right\}; \tag{52a}$$

$$\left\{ \mathcal{N}_{j|1}(k) \triangleq \left\{ y^n \in \mathcal{N}_{j|1} : \ell_{1,j}^{(\eta_{k-1})} = d(\mathbf{c}_1, y^n | \mathcal{S}_{1,j}^{(\eta_{k-1})}) \text{ and } d(\mathbf{c}_1, y^n | \mathcal{S}_{1,j}^{(\eta_k)}) = k + 1 \right\}, \tag{52b}$$

where η_k is defined in (49).

With Definition 3, we have the following proposition.

Proposition 3. For non-empty $\mathcal{T}_{j|1}$, the following two properties hold.

- (i) $\{\mathcal{T}_{j|1}(k)\}_{k \in [\ell_{1,j}-1] \cup \{0\}}$ forms a partition of $\mathcal{T}_{j|1}$;
- (ii) $\{\mathcal{N}_{j|1}(k)\}_{k \in [\ell_{1,j}-1] \cup \{0\}}$ is a collection of disjoint subsets of $\mathcal{N}_{j|1}$.

Proof. It can be seen from the definitions of $\{\mathcal{T}_{j|1}(k)\}_{k \in [\ell_{1,j}-1] \cup \{0\}}$ and $\{\mathcal{N}_{j|1}(k)\}_{k \in [\ell_{1,j}-1] \cup \{0\}}$ that they are collections of mutually disjoint subsets of $\mathcal{T}_{j|1}$ and $\mathcal{N}_{j|1}$, respectively. It remains to argue that every element in $\mathcal{T}_{j|1}$ belongs to $\mathcal{T}_{j|1}(k)$ for some $k \in [\ell_{1,j} - 1] \cup \{0\}$. Noting that the element y^n in $\mathcal{T}_{j|1}$ satisfies $d(\mathbf{c}_1, y^n | \mathcal{S}_{1,j}) \leq \ell_{1,j} - 1$, we differentiate two cases: $d(\mathbf{c}_1, y^n | \mathcal{S}_{1,j}) \leq \ell_{1,j} - 2$ and $d(\mathbf{c}_1, y^n | \mathcal{S}_{1,j}) = \ell_{1,j} - 1$. For the former case, $d(\mathbf{c}_1, y^n | \mathcal{S}_{1,j}^{(m)}) < \ell_{1,j}^{(m)} - 1$ must hold for $m = \eta_k$; hence, this y^n will be contained in $\mathcal{T}_{j|1}(k)$. For the latter case, y^n will be included in $\mathcal{T}_{j|1}(\ell_{1,j} - 1)$. The proof is thus completed. \square

In light of Proposition 3, we can apply the ratio-sum inequality to obtain

$$\frac{P_{X^n, Y^n}(\mathbf{c}_1, \mathcal{T}_{j|1})}{P_{X^n, Y^n}(\mathbf{c}_1, \mathcal{N}_{j|1})} \leq \frac{\sum_{k=0: \mathcal{T}_{j|1}(k) \neq \emptyset}^{\ell_{1,j}-1} P_{X^n, Y^n}(\mathbf{c}_1, \mathcal{T}_{j|1}(k))}{\sum_{k=0: \mathcal{T}_{j|1}(k) \neq \emptyset}^{\ell_{1,j}-1} P_{X^n, Y^n}(\mathbf{c}_1, \mathcal{N}_{j|1}(k))} \tag{53}$$

$$\leq \max_{k \in [\ell_{1,j}-1] \cup \{0\}: \mathcal{T}_{j|1}(k) \neq \emptyset} \frac{P_{X^n, Y^n}(\mathbf{c}_1, \mathcal{T}_{j|1}(k))}{P_{X^n, Y^n}(\mathbf{c}_1, \mathcal{N}_{j|1}(k))}. \tag{54}$$

We continue to construct a fine partition of $\mathcal{T}_{j|1}(k)$ and the corresponding disjoint subsets of $\mathcal{N}_{j|1}(k)$ in Proposition 4 after giving the next definition.

Definition 4. Define for $u^n \in \mathcal{T}_{j|1}(k)$,

$$\left\{ \mathcal{T}_{j|1}(u^n; k) \triangleq \left\{ y^n \in \mathcal{T}_{j|1}(k) : d(u^n, y^n | (\mathcal{S}_{1,j}^{(\eta_k)})^c) = 0 \right\}; \tag{55a}$$

$$\left\{ \mathcal{N}_{j|1}(u^n; k) \triangleq \left\{ y^n \in \mathcal{N}_{j|1}(k) : d(u^n, y^n | (\mathcal{S}_{1,j}^{(\eta_k)})^c) = 0 \right\}, \tag{55b}$$

where η_k is given in (49).

Note from Definition 4 that for one element u^n in non-empty $\mathcal{T}_{j|1}(k)$, we can find a group of elements that have identical bit components to u^n with indices in $(\mathcal{S}_{1,j}^{(\eta_k)})^c$. We

denote this group as $\mathcal{T}_{j|1}(u^n; k)$. We continue this grouping manipulation until all elements in $\mathcal{T}_{j|1}(k)$ are exhausted as summarized below.

Proposition 4. For non-empty $\mathcal{T}_{j|1}(k)$, there exists a representative subset $\mathcal{U}_{j|1}(k) \subseteq \mathcal{T}_{j|1}(k)$ such that the following two properties hold.

- (i) $\{\mathcal{T}_{j|1}(u^n; k)\}_{u^n \in \mathcal{U}_{j|1}(k)}$ forms a (non-empty) partition of $\mathcal{T}_{j|1}(k)$;
- (ii) $\{\mathcal{N}_{j|1}(u^n; k)\}_{u^n \in \mathcal{U}_{j|1}(k)}$ is a collection of (non-empty) disjoint subsets of $\mathcal{N}_{j|1}(k)$.

Since the above proposition can be self-validated via its sequential selection manipulation of each u^n from $\mathcal{T}_{j|1}(k)$, we omit the proof. Interested readers can find the details in [1] (Section III-C).

From Proposition 4, using again the ratio-sum inequality, we obtain that for non-empty $\mathcal{T}_{j|1}(k)$,

$$\frac{P_{X^n, Y^n}(\mathbf{c}_1, \mathcal{T}_{j|1}(k))}{P_{X^n, Y^n}(\mathbf{c}_1, \mathcal{N}_{j|1}(k))} \leq \frac{\sum_{u^n \in \mathcal{U}_{j|1}(k)} P_{X^n, Y^n}(\mathbf{c}_1, \mathcal{T}_{j|1}(u^n; k))}{\sum_{u^n \in \mathcal{U}_{j|1}(k)} P_{X^n, Y^n}(\mathbf{c}_1, \mathcal{N}_{j|1}(u^n; k))} \tag{56}$$

$$\leq \max_{u^n \in \mathcal{U}_{j|1}(k)} \frac{P_{X^n, Y^n}(\mathbf{c}_1, \mathcal{T}_{j|1}(u^n; k))}{P_{X^n, Y^n}(\mathbf{c}_1, \mathcal{N}_{j|1}(u^n; k))}. \tag{57}$$

Noting that the above result can be similarly conducted for general $i > 1$, we combine (42), (54) and (57) to conclude that

$$\frac{\delta_n}{b_n} \leq 2 \max_{i \in [M] \text{ and } j \in [M] \setminus \{i\}: \mathcal{T}_{j|i} \neq \emptyset} \max_{k \in [\ell_{i,j-1}] \cup \{0\}: \mathcal{T}_{j|i}(k) \neq \emptyset} \max_{u^n \in \mathcal{U}_{j|i}(k)} \frac{P_{X^n, Y^n}(\mathbf{c}_i, \mathcal{T}_{j|i}(u^n; k))}{P_{X^n, Y^n}(\mathbf{c}_i, \mathcal{N}_{j|i}(u^n; k))}. \tag{58}$$

The final task is to evaluate $P_{X^n, Y^n}(\mathbf{c}_i, \mathcal{T}_{j|i}(u^n; k)) / P_{X^n, Y^n}(\mathbf{c}_i, \mathcal{N}_{j|i}(u^n; k))$ in order to characterize a linear upper bound for δ_n / b_n .

3.4. Characterization of a Linear Upper Bound for δ_n / b_n

We again focus on $i = 1$ with \mathbf{c}_1 being the all-zero codeword for simplicity. The definitions of $\mathcal{T}_{j|1}(u^n; k)$ in (55a) and $\mathcal{N}_{j|1}(u^n; k)$ in (55b) indicate that when dealing with the ratio $P_{X^n, Y^n}(\mathbf{c}_1, \mathcal{T}_{j|1}(u^n; k)) / P_{X^n, Y^n}(\mathbf{c}_1, \mathcal{N}_{j|1}(u^n; k))$, we only need to consider those bits with indices in $\mathcal{S}_{1,j}^{(\eta_k)}$ because the remaining bits of all tuples in $\mathcal{T}_{j|1}(u^n; k)$ and $\mathcal{N}_{j|1}(u^n; k)$ have identical values as u^n . Note that all $|\mathcal{T}_{j|1}(u^n; k)|$ elements in $\mathcal{T}_{j|1}(u^n; k)$ have exactly k ones with indices in $\mathcal{S}_{1,j}^{(\eta_k)}$, and all $|\mathcal{N}_{j|1}(u^n; k)|$ elements in $\mathcal{N}_{j|1}(u^n; k)$ have exactly $k + 1$ ones with indices in $\mathcal{S}_{1,j}^{(\eta_k)}$, we can immediately infer that

$$\frac{P_{X^n, Y^n}(\mathbf{c}_1, \mathcal{T}_{j|1}(u^n; k))}{P_{X^n, Y^n}(\mathbf{c}_1, \mathcal{N}_{j|1}(u^n; k))} = \frac{P_{X^n}(\mathbf{c}_1) \cdot P_{Y^n|\mathbf{c}_1}(\mathcal{T}_{j|1}(u^n; k)|\mathbf{c}_1)}{P_{X^n}(\mathbf{c}_1) \cdot P_{Y^n|\mathbf{c}_1}(\mathcal{N}_{j|1}(u^n; k)|\mathbf{c}_1)} = \frac{(1-p)}{p} \cdot \frac{|\mathcal{T}_{j|1}(u^n; k)|}{|\mathcal{N}_{j|1}(u^n; k)|}. \tag{59}$$

The cardinalities of $\mathcal{T}_{j|1}(u^n; k)$ and $\mathcal{N}_{j|1}(u^n; k)$ then decide the ratio in (59) as verified in the next proposition, based on which the proof of Theorem 2 can be completed from (58).

Proposition 5. For $u^n \in \mathcal{T}_{j|1}(k)$, we have

$$\frac{P_{X^n, Y^n}(\mathbf{c}_1, \mathcal{T}_{j|1}(u^n; k))}{P_{X^n, Y^n}(\mathbf{c}_1, \mathcal{N}_{j|1}(u^n; k))} \leq \frac{(1-p)}{p} n. \tag{60}$$

Proof. Recall from (22a), (52a) and (55a) that $y^n \in \mathcal{T}_{j|1}(u^n; k) \subseteq \mathcal{T}_{j|1}(k) \subseteq \mathcal{T}_{j|1}$ if and only if

$$\begin{cases} P_{X^n, Y^n}(c_1, y^n) = P_{X^n, Y^n}(c_j, y^n) = \max_{h \in [M] \setminus \{1\}} P_{X^n, Y^n}(x_{(h)}^n, y^n) \text{ and } d(c_1, y^n | \mathcal{S}_{1,j}) < |\mathcal{S}_{1,j}|; & (61a) \\ \ell_{1,j}^{(\eta_k-1)} - 1 \leq d(c_1, y^n | \mathcal{S}_{1,j}^{(\eta_k-1)}) \text{ and } d(c_1, y^n | \mathcal{S}_{1,j}^{(\eta_k)}) = k; & (61b) \\ d(u^n, y^n | (\mathcal{S}_{1,j}^{(\eta_k)})^c) = 0. & (61c) \end{cases}$$

Thus, the number of elements in $\mathcal{T}_{j|1}(u^n; k)$ is exactly the number of channel outputs y^n fulfilling the above three conditions. We then examine the number of y^n satisfying (61b) and (61c). Noting that these y^n have either $\ell_{1,j}^{(\eta_k-1)} - 1$ ones or $\ell_{1,j}^{(\eta_k-1)}$ ones with indices in $\mathcal{S}_{1,j}^{(\eta_k-1)}$, we know that there are at most

$$\binom{\ell_{1,j}^{(\eta_k-1)}}{\ell_{1,j}^{(\eta_k-1)} - 1} \binom{\ell_j^{(\eta_k)} - \ell_{1,j}^{(\eta_k-1)}}{k - (\ell_{1,j}^{(\eta_k-1)} - 1)} + \binom{\ell_{1,j}^{(\eta_k-1)}}{\ell_{1,j}^{(\eta_k-1)}} \binom{\ell_j^{(\eta_k)} - \ell_{1,j}^{(\eta_k-1)}}{k - \ell_{1,j}^{(\eta_k-1)}} \tag{62}$$

of y^n tuples satisfying (61b) and (61c). Disregarding (61a), we get that the number of elements in $\mathcal{T}_{j|1}(u^n; k)$ is upper-bounded by (62).

On the other hand, from (22b), (52b) and (55b), we obtain that $w^n \in \mathcal{N}_{j|1}(u^n; k) \subseteq \mathcal{N}_{j|1}(k) \subseteq \mathcal{N}_{j|1}$ if and only if

$$\begin{cases} P_{X^n, Y^n}(c_1, w^n) \cdot q^2 = P_{X^n, Y^n}(c_j, w^n); & (63a) \\ P_{X^n, Y^n}(c_1, w^n) \cdot q^2 \neq P_{X^n, Y^n}(c_r, w^n) \text{ for } r \in [j-1] \setminus \{1\}; & (63b) \\ \ell_{1,j}^{(\eta_k-1)} = d(c_1, w^n | \mathcal{S}_{1,j}^{(\eta_k-1)}) \text{ and } d(c_1, w^n | \mathcal{S}_{1,j}^{(\eta_k)}) = k + 1; & (63c) \\ d(u^n, w^n | (\mathcal{S}_{1,j}^{(\eta_k)})^c) = 0. & (63d) \end{cases}$$

We then claim that any w^n satisfying (63c) and (63d) directly validate (63a) and (63b). Note that the validity of the claim, which we prove in Appendix B, immediately implies that the number of elements in $\mathcal{N}_{j|1}(u^n; k)$ can be determined by (63c) and (63d), and hence

$$|\mathcal{N}_{j|1}(u^n; k)| = \binom{\ell_j^{(\eta_k)} - \ell_{1,j}^{(\eta_k-1)}}{k + 1 - \ell_{1,j}^{(\eta_k-1)}}. \tag{64}$$

Under this claim, (62) and (64) result in

$$\frac{|\mathcal{T}_{j|1}(u^n; k)|}{|\mathcal{N}_{j|1}(u^n; k)|} \leq \frac{\binom{\ell_{1,j}^{(\eta_k-1)}}{\ell_{1,j}^{(\eta_k-1)} - 1} \binom{\ell_j^{(\eta_k)} - \ell_{1,j}^{(\eta_k-1)}}{k - (\ell_{1,j}^{(\eta_k-1)} - 1)} + \binom{\ell_{1,j}^{(\eta_k-1)}}{\ell_{1,j}^{(\eta_k-1)}} \binom{\ell_j^{(\eta_k)} - \ell_{1,j}^{(\eta_k-1)}}{k - \ell_{1,j}^{(\eta_k-1)}}}{\binom{\ell_j^{(\eta_k)} - \ell_{1,j}^{(\eta_k-1)}}{k + 1 - \ell_{1,j}^{(\eta_k-1)}}} \tag{65}$$

$$= \ell_{1,j}^{(\eta_k-1)} + \frac{k + 1 - \ell_{1,j}^{(\eta_k-1)}}{\ell_j^{(\eta_k)} - k} \tag{66}$$

$$\leq \ell_{1,j}^{(\eta_k-1)} + \frac{\ell_j^{(\eta_k)} - \ell_{1,j}^{(\eta_k-1)}}{1} \tag{67}$$

$$\leq n, \tag{68}$$

where (67) holds because $k \leq \ell_j^{(\eta_k)} - 1$ by (49), and (68) follows from $\ell_{1,j}^{(\eta_k)} \leq \ell_{1,j} \leq n$. The proof of the proposition is thus completed by (59) and (68). \square

4. Conclusions

In this paper, we analyzed the error probability of block codes sent over the memoryless BSC under an arbitrary (not necessarily uniform) input distribution and used in conjunction with (optimal) MAP decoding. We showed that decoder ties do not affect the error exponent of the probability of error, thus extending a similar result recently established in [1] for uniformly distributed channel inputs. This result was obtained by proving that the relative deviation of the error probability from the probability of error grows no more than linearly in blocklength when no MAP decoding ties occur, directly implying that decoder ties have only a sub-exponential effect on the error probability as blocklength grows without bound. Future work includes further extending this result for more general channels used under arbitrary input statistics, such as non-binary symmetric channels (Note that the result of Theorem 1 can be extended for non-binary (q -ary, $q > 2$) codes sent over q -ary symmetric memoryless channels under a uniform input distribution; see [25] (Theorem 2).) and binary non-symmetric channels. Studying how to sharpen the upper bound derived in (10) for “sufficiently good” codes as highlighted in Remark 1 and for codes with small blocklengths are other worthwhile future directions.

Author Contributions: Conceptualization, L.-H.C., P.-N.C. and F.A.; Formal analysis, L.-H.C.; Writing—original draft, L.-H.C.; Writing—review & editing, P.-N.C. and F.A. All authors have read and agreed to the published version of the manuscript.

Funding: The work of Ling-Hua Chang is supported by the Ministry of Science and Technology, Taiwan under Grant MOST 109-2221-E-155-035-MY3. The work of Po-Ning Chen is supported by the Ministry of Science and Technology, Taiwan, under Grant MOST 110-2221-E-A49-024-MY3. The work of Fady Alajaji is supported by the Natural Sciences and Engineering Research Council (NSERC) of Canada.

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

Appendix A. Supplement to Example 1

Under distribution

$$P_{X^4}(c_1) = \frac{q^2}{2 + q^2 + q^{-2}} \tag{A1}$$

$$P_{X^4}(c_2) = P_{X^4}(c_3) = \frac{1}{2 + q^2 + q^{-2}} \tag{A2}$$

$$P_{X^4}(c_4) = \frac{q^{-2}}{2 + q^2 + q^{-2}} \tag{A3}$$

over the code $C_4 = \{c_1, c_2, c_3, c_4\} = \{0000, 0101, 0110, 0111\}$, we obtain:

$$\mathcal{T}_1 = \left\{ y^4 \in \{0, 1\}^4 : P_{X^4, Y^4}(c_1, y^4) = \max_{r \in [4] \setminus \{1\}} P_{X^4, Y^4}(c_r, y^4) \right\} \tag{A4}$$

$$= \left\{ y^4 \in \{0, 1\}^4 : \frac{P_{X^4}(c_1)}{q^{d(c_1, y^4)}} = \max \left(\frac{P_{X^4}(c_2)}{q^{d(c_2, y^4)}}, \frac{P_{X^4}(c_3)}{q^{d(c_3, y^4)}}, \frac{P_{X^4}(c_4)}{q^{d(c_4, y^4)}} \right) \right\} \tag{A5}$$

$$= \left\{ y^4 \in \{0, 1\}^4 : d(c_1, y^4) - 2 = \min \left(d(c_2, y^4), d(c_3, y^4), d(c_4, y^4) + 2 \right) \right\} \tag{A6}$$

$$= \{0101, 0110, 0111, 1101, 1110, 1111\}, \tag{A7}$$

where (A5) follows from (25), and

$$\mathcal{T}_{j|1} \triangleq \left\{ y^4 \in \mathcal{T}_1 : j = \min_{r \in \mathcal{I}_1(y^4) : d(c_1, y^4 | S_{1,r}) < |S_{1,r}|} r \right\} = \emptyset \text{ for } j = 2, 3, 4, \tag{A8}$$

and

$$\mathcal{T}_{j|1} \triangleq \left\{ y^4 \in \mathcal{T}_1 \setminus \left(\bigcup_{h \in [4] \setminus \{1\}} \mathcal{T}_{h|1} \right) : j = \min_{r \in \mathcal{I}_1(y^n)} r \right\} \tag{A9}$$

$$= \begin{cases} \{0101, 0111, 1101, 1111\}, & j = 2; \\ \{0110, 1110\}, & j = 3; \\ \emptyset, & j = 4. \end{cases} \tag{A10}$$

The above derivations are verified via Table A1. Continuing with the same setting, we obtain

$$\mathcal{T}_2 = \left\{ y^4 \in \{0, 1\}^4 : P_{X^4, Y^4}(c_2, y^4) = \max_{r \in [4] \setminus \{2\}} P_{X^4, Y^4}(c_r, y^4) \right\} \tag{A11}$$

$$= \left\{ y^4 \in \{0, 1\}^4 : \frac{P_{X^4}(c_2)}{q^{d(c_2, y^4)}} = \max \left(\frac{P_{X^4}(c_1)}{q^{d(c_1, y^4)}}, \frac{P_{X^4}(c_3)}{q^{d(c_3, y^4)}}, \frac{P_{X^4}(c_4)}{q^{d(c_4, y^4)}} \right) \right\} \tag{A12}$$

$$= \left\{ y^4 \in \{0, 1\}^4 : d(c_2, y^4) = \min \left(d(c_1, y^4) - 2, d(c_3, y^4), d(c_4, y^4) + 2 \right) \right\} \tag{A13}$$

$$= \{0101, 0111, 1101, 1111\}, \tag{A14}$$

$$\mathcal{T}_{j|2} \triangleq \left\{ y^4 \in \mathcal{T}_2 : j = \min_{r \in \mathcal{I}_2(y^4) : d(c_2, y^4 | S_{2,r}) < |S_{2,r}|} r \right\} = \begin{cases} \mathcal{T}_2, & j = 1; \\ \emptyset, & j = 3, 4 \end{cases} \tag{A15}$$

and

$$\mathcal{T}_{j|2} \triangleq \left\{ y^4 \in \mathcal{T}_2 \setminus \left(\bigcup_{h \in [4] \setminus \{1\}} \mathcal{T}_{h|2} \right) : j = \min_{r \in \mathcal{I}_2(y^n)} r \right\} = \emptyset \text{ for } j = 1, 3, 4, \tag{A16}$$

where the above derivations are also confirmed via Table A1. Based on Table A1, we further have

$$\mathcal{T}_3 = \left\{ y^4 \in \{0, 1\}^4 : P_{X^4, Y^4}(c_3, y^4) = \max_{r \in [4] \setminus \{3\}} P_{X^4, Y^4}(c_r, y^4) \right\} \tag{A17}$$

$$= \left\{ y^4 \in \{0, 1\}^4 : \frac{P_{X^4}(c_3)}{q^{d(c_3, y^4)}} = \max \left(\frac{P_{X^4}(c_1)}{q^{d(c_1, y^4)}}, \frac{P_{X^4}(c_2)}{q^{d(c_2, y^4)}}, \frac{P_{X^4}(c_4)}{q^{d(c_4, y^4)}} \right) \right\} \tag{A18}$$

$$= \left\{ y^4 \in \{0, 1\}^4 : d(c_3, y^4) = \min \left(d(c_1, y^4) - 2, d(c_2, y^4), d(c_4, y^4) + 2 \right) \right\} \tag{A19}$$

$$= \{0110, 0111, 1110, 1111\}, \tag{A20}$$

$$\mathcal{T}_{j|3} \triangleq \left\{ y^4 \in \mathcal{T}_3 : j = \min_{r \in \mathcal{I}_3(y^4) : d(c_3, y^4 | S_{3,r}) < |S_{3,r}|} r \right\} = \begin{cases} \mathcal{T}_3, & j = 1; \\ \emptyset, & j = 2, 4, \end{cases} \tag{A21}$$

and

$$\mathcal{T}_{j|3} \triangleq \left\{ y^4 \in \mathcal{T}_3 \setminus \left(\bigcup_{h \in [4] \setminus \{3\}} \mathcal{T}_{h|3} \right) : j = \min_{r \in \mathcal{I}_3(y^n)} r \right\} = \emptyset \text{ for } j = 1, 2, 4. \tag{A22}$$

Furthermore, we establish from Table A1 that

$$\mathcal{T}_4 = \left\{ y^4 \in \{0, 1\}^4 : P_{X^4, Y^4}(c_4, y^4) = \max_{r \in [4] \setminus \{4\}} P_{X^4, Y^4}(c_r, y^4) \right\} \tag{A23}$$

$$= \left\{ y^4 \in \{0, 1\}^4 : \frac{P_{X^4}(c_4)}{q^{d(c_4, y^4)}} = \max \left(\frac{P_{X^4}(c_1)}{q^{d(c_1, y^4)}}, \frac{P_{X^4}(c_2)}{q^{d(c_2, y^4)}}, \frac{P_{X^4}(c_3)}{q^{d(c_3, y^4)}} \right) \right\} \tag{A24}$$

$$= \left\{ y^4 \in \{0, 1\}^4 : d(c_4, y^4) + 2 = \min \left(d(c_1, y^4) - 2, d(c_2, y^4), d(c_3, y^4) \right) \right\} \tag{A25}$$

$$= \emptyset, \tag{A26}$$

$$\mathcal{T}_{j|4} \triangleq \left\{ y^4 \in \mathcal{T}_4 : j = \min_{r \in \mathcal{I}_4(y^4) : d(c_4, y^4 | S_{4,r}) < |S_{4,r}|} r \right\} = \emptyset \text{ for } j = 1, 2, 3, \tag{A27}$$

and

$$\mathcal{S}_{j|4} \triangleq \left\{ y^4 \in \mathcal{T}_4 \setminus \left(\bigcup_{h \in [4] \setminus \{4\}} \mathcal{T}_{h|4} \right) : j = \min_{r \in \mathcal{I}_4(y^4)} r \right\} = \emptyset \text{ for } j = 1, 2, 3. \tag{A28}$$

After summarizing all sets derived above in Table A2, we remark that

$$\begin{cases} \mathcal{S}_{2|1} \subseteq \mathcal{T}_{1|2} & \text{and } P_{X^4, Y^4}(c_1, y^4) = P_{X^4, Y^4}(c_2, y^4) & \text{for every } y^4 \in \mathcal{S}_{2|1}; \\ \mathcal{S}_{3|1} \subseteq \mathcal{T}_{1|3} & \text{and } P_{X^4, Y^4}(c_1, y^4) = P_{X^4, Y^4}(c_2, y^4) & \text{for every } y^4 \in \mathcal{S}_{3|1}. \end{cases} \tag{A29}$$

Note that $\{\mathcal{S}_{j|i}\}_{i \in [M], j \in [M] \setminus \{i\}}$ are disjoint as confirmed in Remark 4 such that every element $y^n \in \mathcal{S}_{j|i}$ appears only once in the following summation:

$$\begin{aligned} \sum_{i \in [4]} \sum_{j \in [4] \setminus \{i\}} P_{X^4, Y^4}(c_i, \mathcal{S}_{j|i}) &= P_{X^4, Y^4}(c_1, 0101) + P_{X^4, Y^4}(c_1, 0110) \\ &\quad + P_{X^4, Y^4}(c_1, 0111) + P_{X^4, Y^4}(c_1, 1101) \\ &\quad + P_{X^4, Y^4}(c_1, 1110) + P_{X^4, Y^4}(c_1, 1111) \end{aligned} \tag{A30}$$

$$= \frac{p^4 q^2}{2 + q^2 + q^{-2}} (q^2 + q^2 + q + q + q + 1) \tag{A31}$$

$$= \frac{p^4 q^2}{2 + q^2 + q^{-2}} (2q^2 + 3q + 1). \tag{A32}$$

Additionally,

$$\begin{aligned} &\sum_{i \in [4]} \sum_{j \in [4] \setminus \{i\}} P_{X^4, Y^4}(c_i, \mathcal{T}_{j|i}) \\ &= P_{X^4, Y^4}(c_2, 0101) + P_{X^4, Y^4}(c_2, 0111) + P_{X^4, Y^4}(c_2, 1101) \\ &\quad + P_{X^4, Y^4}(c_2, 1111) + P_{X^4, Y^4}(c_3, 0110) + P_{X^4, Y^4}(c_3, 0111) \\ &\quad + P_{X^4, Y^4}(c_3, 1110) + P_{X^4, Y^4}(c_3, 1111) \end{aligned} \tag{A33}$$

$$= \frac{p^4}{2 + q^2 + q^{-2}} (q^4 + q^3 + q^3 + q^2 + q^4 + q^3 + q^3 + q^2) \tag{A34}$$

$$= \sum_{i \in [4]} \sum_{j \in [4] \setminus \{i\}} P_{X^4, Y^4}(c_i, \mathcal{S}_{j|i}) + \frac{p^4 q^2}{2 + q^2 + q^{-2}} (q + 1). \tag{A35}$$

Finally, we have

$$\mathcal{N}_1 = \left\{ y^4 \in \{0, 1\}^4 : \frac{P_{X^4}(c_1)}{q^{d(c_1, y^4)}} < \max \left(\frac{P_{X^4}(c_2)}{q^{d(c_2, y^4)}}, \frac{P_{X^4}(c_3)}{q^{d(c_3, y^4)}}, \frac{P_{X^4}(c_4)}{q^{d(c_4, y^4)}} \right) \right\} \tag{A36}$$

$$= \left\{ y^4 \in \{0, 1\}^4 : d(c_1, y^4) - 2 > \min \left(d(c_2, y^4), d(c_3, y^4), d(c_4, y^4) + 2 \right) \right\} \tag{A37}$$

$$= \emptyset, \tag{A38}$$

$$\begin{aligned} \mathcal{N}_2 &= \left\{ y^4 \in \{0,1\}^4 : d(c_2, y^4) > \min \left(d(c_1, y^4) - 2, d(c_3, y^4), d(c_4, y^4) + 2 \right) \right\} \\ &= \{0,1\}^4 \setminus \mathcal{T}_2, \end{aligned} \tag{A39}$$

$$\begin{aligned} \mathcal{N}_3 &= \left\{ y^4 \in \{0,1\}^4 : d(c_3, y^4) > \min \left(d(c_2, y^4), d(c_1, y^4) - 2, d(c_4, y^4) + 2 \right) \right\} \\ &= \{0,1\}^4 \setminus \mathcal{T}_3, \end{aligned} \tag{A40}$$

$$\begin{aligned} \mathcal{N}_4 &= \left\{ y^4 \in \{0,1\}^4 : d(c_4, y^4) + 2 > \min \left(d(c_2, y^4), d(c_3, y^4), d(c_1, y^4) - 2 \right) \right\} \\ &= \{0,1\}^4, \end{aligned} \tag{A41}$$

$$\mathcal{N}_{j|1} = \emptyset \quad \text{for } j = 2, 3, 4, \tag{A42}$$

$$\begin{aligned} \mathcal{N}_{j|2} &= \left\{ y^4 \in \mathcal{N}_2 : P_{X^4, Y^4}(c_2, y^4) \cdot q = P_{X^4, Y^4}(c_j, y^4) \cdot \frac{1}{q} \right. \\ &\quad \left. \text{and } P_{X^4, Y^4}(c_j, y^4) \neq P_{X^4, Y^4}(c_r, y^4) \text{ for } r \in [j-1] \setminus \{2\} \right\} \end{aligned} \tag{A43}$$

$$\begin{aligned} &= \left\{ y^4 \in \mathcal{N}_2 : \frac{P_{X^4}(c_2)}{q^{d(c_2, y^4)-1}} = \frac{P_{X^4}(c_j)}{q^{d(c_j, y^4)+1}} \text{ and } \frac{P_{X^4}(c_j)}{q^{d(c_j, y^4)}} \neq \frac{P_{X^4}(c_r)}{q^{d(c_r, y^4)}} \right. \\ &\quad \left. \text{for } r \in [j-1] \setminus \{2\} \right\} \end{aligned} \tag{A44}$$

$$\begin{aligned} &= \begin{cases} \left\{ y^4 \in \mathcal{N}_2 : \frac{P_{X^4}(c_2)}{q^{d(c_2, y^4)-1}} = \frac{P_{X^4}(c_1)}{q^{d(c_1, y^4)+1}} \right\}, & j = 1; \\ \left\{ y^4 \in \mathcal{N}_2 : \frac{P_{X^4}(c_2)}{q^{d(c_2, y^4)-1}} = \frac{P_{X^4}(c_3)}{q^{d(c_3, y^4)+1}} \text{ and } \frac{P_{X^4}(c_3)}{q^{d(c_3, y^4)}} \neq \frac{P_{X^4}(c_1)}{q^{d(c_1, y^4)}} \right\}, & j = 3; \\ \left\{ y^4 \in \mathcal{N}_2 : \frac{P_{X^4}(c_2)}{q^{d(c_2, y^4)-1}} = \frac{P_{X^4}(c_4)}{q^{d(c_4, y^4)+1}} \text{ and } \frac{P_{X^4}(c_4)}{q^{d(c_4, y^4)}} \neq \frac{P_{X^4}(c_r)}{q^{d(c_r, y^4)}} \right. \\ \quad \left. \text{for } r \in [3] \setminus \{2\} \right\}, & j = 4 \end{cases} \end{aligned} \tag{A45}$$

$$\begin{aligned} &= \begin{cases} \left\{ y^4 \in \mathcal{N}_2 : d(c_2, y^4) = d(c_1, y^4) \right\}, & j = 1; \\ \left\{ y^4 \in \mathcal{N}_2 : d(c_2, y^4) = d(c_3, y^4) + 2 \text{ and } d(c_2, y^4) \neq d(c_1, y^4) \right\}, & j = 3; \\ \left\{ y^4 \in \mathcal{N}_2 : d(c_2, y^4) = d(c_4, y^4) + 4, d(c_2, y^4) \neq d(c_1, y^4) \right. \\ \quad \left. \text{and } d(c_2, y^4) \neq d(c_3, y^4) + 2 \right\}, & j = 4 \end{cases} \\ &= \begin{cases} \{0001, 0100, 0011, 0110, 1001, 1100, 1011, 1110\}, & j = 1; \\ \{0010, 1010\}, & j = 3; \\ \emptyset, & j = 4, \end{cases} \end{aligned} \tag{A46}$$

$$\begin{aligned} \mathcal{N}_{j|3} &= \left\{ y^4 \in \mathcal{N}_3 : P_{X^4, Y^4}(c_3, y^4) \cdot q = P_{X^4, Y^4}(c_j, y^4) \cdot \frac{1}{q} \right. \\ &\quad \left. \text{and } P_{X^4, Y^4}(c_j, y^4) \neq P_{X^4, Y^4}(c_r, y^4) \text{ for } r \in [j-1] \setminus \{3\} \right\} \end{aligned} \tag{A47}$$

$$= \left\{ y^4 \in \mathcal{N}_3 : \frac{P_{X^4}(c_3)}{q^{d(c_3, y^4)-1}} = \frac{P_{X^4}(c_j)}{q^{d(c_j, y^4)+1}} \text{ and } \frac{P_{X^4}(c_j)}{q^{d(c_j, y^4)}} \neq \frac{P_{X^4}(c_r)}{q^{d(c_r, y^4)}} \right. \\ \left. \text{for } r \in [j-1] \setminus \{3\} \right\} \tag{A48}$$

$$= \left\{ \begin{aligned} & \left\{ y^4 \in \mathcal{N}_3 : \frac{P_{X^4}(c_3)}{q^{d(c_3, y^4)-1}} = \frac{P_{X^4}(c_1)}{q^{d(c_1, y^4)+1}} \right\}, & j = 1; \\ & \left\{ y^4 \in \mathcal{N}_3 : \frac{P_{X^4}(c_3)}{q^{d(c_3, y^4)-1}} = \frac{P_{X^4}(c_2)}{q^{d(c_2, y^4)+1}} \text{ and } \frac{P_{X^4}(c_2)}{q^{d(c_2, y^4)}} \neq \frac{P_{X^4}(c_1)}{q^{d(c_1, y^4)}} \right\}, & j = 2; \\ & \left\{ y^4 \in \mathcal{N}_3 : \frac{P_{X^4}(c_3)}{q^{d(c_3, y^4)-1}} = \frac{P_{X^4}(c_4)}{q^{d(c_4, y^4)+1}} \text{ and } \frac{P_{X^4}(c_4)}{q^{d(c_4, y^4)}} \neq \frac{P_{X^4}(c_r)}{q^{d(c_r, y^4)}} \right. \\ & \qquad \qquad \qquad \left. \text{for } r \in [3] \setminus \{3\} \right\}, & j = 4 \end{aligned} \right. \tag{A49}$$

$$= \left\{ \begin{aligned} & \left\{ y^4 \in \mathcal{N}_3 : d(c_3, y^4) = d(c_1, y^4) \right\}, & j = 1; \\ & \left\{ y^4 \in \mathcal{N}_3 : d(c_3, y^4) = d(c_2, y^4) + 2 \text{ and } d(c_3, y^4) \neq d(c_1, y^4) \right\}, & j = 3; \\ & \left\{ y^4 \in \mathcal{N}_3 : d(c_3, y^4) = d(c_4, y^4) + 4, d(c_3, y^4) \neq d(c_1, y^4) \right. \\ & \qquad \qquad \qquad \left. \text{and } d(c_3, y^4) \neq d(c_2, y^4) + 2 \right\}, & j = 4 \end{aligned} \right.$$

$$= \left\{ \begin{aligned} & \{0010, 0100, 0011, 0101, 1010, 1100, 1011, 1101\}, & j = 1; \\ & \{0001, 1001\}, & j = 3; \\ & \emptyset, & j = 4, \end{aligned} \right. \tag{A50}$$

$$\mathcal{N}_{j4} = \left\{ y^4 \in \mathcal{N}_4 : P_{X^4, Y^4}(c_4, y^4) \cdot q = P_{X^4, Y^4}(c_j, y^4) \cdot \frac{1}{q} \right. \\ \left. \text{and } P_{X^4, Y^4}(c_j, y^4) \neq P_{X^4, Y^4}(c_r, y^4) \text{ for } r \in [j-1] \setminus \{4\} \right\} \tag{A51}$$

$$= \left\{ y^4 \in \mathcal{N}_4 : \frac{P_{X^4}(c_4)}{q^{d(c_4, y^4)-1}} = \frac{P_{X^4}(c_j)}{q^{d(c_j, y^4)+1}} \text{ and } \frac{P_{X^4}(c_j)}{q^{d(c_j, y^4)}} \neq \frac{P_{X^4}(c_r)}{q^{d(c_r, y^4)}} \text{ for } r \in [j-1] \right\} \\ = \left\{ \begin{aligned} & \left\{ y^4 \in \mathcal{N}_4 : \frac{P_{X^4}(c_4)}{q^{d(c_4, y^4)-1}} = \frac{P_{X^4}(c_1)}{q^{d(c_1, y^4)+1}} \right\}, & j = 1; \\ & \left\{ y^4 \in \mathcal{N}_4 : \frac{P_{X^4}(c_4)}{q^{d(c_4, y^4)-1}} = \frac{P_{X^4}(c_2)}{q^{d(c_2, y^4)+1}} \text{ and } \frac{P_{X^4}(c_2)}{q^{d(c_2, y^4)}} \neq \frac{P_{X^4}(c_1)}{q^{d(c_1, y^4)}} \right\}, & j = 2; \\ & \left\{ y^4 \in \mathcal{N}_4 : \frac{P_{X^4}(c_4)}{q^{d(c_4, y^4)-1}} = \frac{P_{X^4}(c_3)}{q^{d(c_3, y^4)+1}} \text{ and } \frac{P_{X^4}(c_3)}{q^{d(c_3, y^4)}} \neq \frac{P_{X^4}(c_r)}{q^{d(c_r, y^4)}} \right. \\ & \qquad \qquad \qquad \left. \text{for } r \in [2] \right\}, & j = 3 \end{aligned} \right. \tag{A52}$$

$$= \left\{ \begin{aligned} & \left\{ y^4 \in \mathcal{N}_4 : d(c_4, y^4) = d(c_1, y^4) - 2 \right\}, & j = 1; \\ & \left\{ y^4 \in \mathcal{N}_4 : d(c_4, y^4) = d(c_2, y^4) \text{ and } d(c_4, y^4) \neq d(c_1, y^4) - 2 \right\}, & j = 2; \\ & \left\{ y^4 \in \mathcal{N}_4 : d(c_4, y^4) = d(c_3, y^4), d(c_4, y^4) \neq d(c_1, y^4) - 2 \right. \\ & \qquad \qquad \qquad \left. \text{and } d(c_4, y^4) \neq d(c_2, y^4) \right\}, & j = 3 \end{aligned} \right. \tag{A53}$$

$$= \emptyset \text{ for } j = 1, 2, 3. \tag{A54}$$

Table A1. Measures used in Example 1.

	$d(0000, y^4) - 2$	$d(0101, y^4)$	$d(0110, y^4)$	$d(0111, y^4) + 2$	$\mathcal{I}_1(y^4)$	$\mathcal{I}_2(y^4)$	$\mathcal{I}_3(y^4)$	$\mathcal{I}_4(y^4)$
$y^4 = 0000$	-2	2	2	5	\emptyset	\emptyset	\emptyset	\emptyset
$y^4 = 0001$	-1	1	3	4	\emptyset	\emptyset	\emptyset	\emptyset
$y^4 = 0010$	-1	3	1	4	\emptyset	\emptyset	\emptyset	\emptyset
$y^4 = 0100$	-1	1	1	4	\emptyset	\emptyset	\emptyset	\emptyset
$y^4 = 1000$	-1	3	3	6	\emptyset	\emptyset	\emptyset	\emptyset
$y^4 = 0011$	0	2	2	3	\emptyset	\emptyset	\emptyset	\emptyset
$y^4 = 0101$	0	0	2	3	{2}	{1}	\emptyset	\emptyset
$y^4 = 0110$	0	2	0	3	{3}	\emptyset	{1}	\emptyset
$y^4 = 1001$	0	2	4	5	\emptyset	\emptyset	\emptyset	\emptyset
$y^4 = 1010$	0	4	2	5	\emptyset	\emptyset	\emptyset	\emptyset
$y^4 = 1100$	0	2	2	5	\emptyset	\emptyset	\emptyset	\emptyset
$y^4 = 0111$	1	1	1	2	{2,3}	{1,3}	{1,2}	\emptyset
$y^4 = 1011$	1	3	3	4	\emptyset	\emptyset	\emptyset	\emptyset
$y^4 = 1101$	1	1	3	4	{2}	{1}	\emptyset	\emptyset
$y^4 = 1110$	1	3	1	4	{3}	\emptyset	{1}	\emptyset
$y^4 = 1111$	2	2	2	3	{2,3}	{1,3}	{1,2}	\emptyset

Table A2. List of $\mathcal{T}_i, \mathcal{N}_i, \mathcal{T}_{ji}, \mathcal{F}_{ji}$ and \mathcal{N}_{ji} for $i \in [4]$ and $j \in [4] \setminus \{i\}$ in Example 1.

\mathcal{T}_1	{0101, 0110, 0111, 1101, 1110, 1111}			\mathcal{N}_1	\emptyset
\mathcal{T}_2	{0101, 0111, 1101, 1111}			\mathcal{N}_2	$\{0, 1\}^4 \setminus \mathcal{T}_2$
\mathcal{T}_3	{0110, 0111, 1110, 1111}			\mathcal{N}_3	$\{0, 1\}^4 \setminus \mathcal{T}_3$
\mathcal{T}_4	\emptyset			\mathcal{N}_4	$\{0, 1\}^4$
$\mathcal{T}_{2 1}$	\emptyset	$\mathcal{F}_{2 1}$	{0101, 0111, 1101, 1111}	$\mathcal{N}_{2 1}$	\emptyset
$\mathcal{T}_{3 1}$	\emptyset	$\mathcal{F}_{3 1}$	{0110, 1110}	$\mathcal{N}_{3 1}$	\emptyset
$\mathcal{T}_{4 1}$	\emptyset	$\mathcal{F}_{4 1}$	\emptyset	$\mathcal{N}_{4 1}$	\emptyset
$\mathcal{T}_{1 2}$	{0101, 0111, 1101, 1111}	$\mathcal{F}_{1 2}$	\emptyset	$\mathcal{N}_{1 2}$	$\mathcal{N}_2 \setminus \{0000, 0010, 1000, 1010\}$
$\mathcal{T}_{3 2}$	\emptyset	$\mathcal{F}_{3 2}$	\emptyset	$\mathcal{N}_{3 2}$	{0010, 1010}
$\mathcal{T}_{4 2}$	\emptyset	$\mathcal{F}_{4 2}$	\emptyset	$\mathcal{N}_{1 2}$	\emptyset
$\mathcal{T}_{1 3}$	{0110, 0111, 1110, 1111}	$\mathcal{F}_{1 3}$	\emptyset	$\mathcal{N}_{1 3}$	$\mathcal{N}_3 \setminus \{0000, 0001, 1000, 1001\}$
$\mathcal{T}_{2 3}$	\emptyset	$\mathcal{F}_{2 3}$	\emptyset	$\mathcal{N}_{2 3}$	{0001, 1001}
$\mathcal{T}_{4 3}$	\emptyset	$\mathcal{F}_{4 3}$	\emptyset	$\mathcal{N}_{4 3}$	\emptyset
$\mathcal{T}_{1 4}$	\emptyset	$\mathcal{F}_{1 4}$	\emptyset	$\mathcal{N}_{1 4}$	\emptyset
$\mathcal{T}_{2 4}$	\emptyset	$\mathcal{F}_{2 4}$	\emptyset	$\mathcal{N}_{2 4}$	\emptyset
$\mathcal{T}_{3 4}$	\emptyset	$\mathcal{F}_{3 4}$	\emptyset	$\mathcal{N}_{3 4}$	\emptyset

Appendix B. The Proof of the Claim Supporting Proposition 5

We validate the claim that (63c) and (63d) imply (63a) and (63b) via the construction of an auxiliary $v^n \in \mathcal{N}_{j|1}(u^n; k)$ from $u^n \in \mathcal{T}_{j|1}(u^n; k)$. This auxiliary v^n will be defined differently according to whether $d(c_1, u^n | \mathcal{S}_{1,j}^{(\eta_k-1)})$ equals $\ell_{1,j}^{(\eta_k-1)}$ or $\ell_{1,j}^{(\eta_k-1)} - 1$ as follows.

- (i) $d(c_1, u^n | \mathcal{S}_{1,j}^{(\eta_k-1)}) = \ell_{1,j}^{(\eta_k-1)}$: In this case, u^n has no zero components with indices in $\mathcal{S}_{1,j}^{(\eta_k-1)}$. Moreover, $d(c_1, u^n | \mathcal{S}_{1,j}^{(\eta_k)}) = k \leq \ell_j^{(\eta_k)} - 1$ indicates that

$$u^n \text{ has at least one zero component with its index in } \mathcal{S}_{1,j}^{(\eta_k)} \setminus \mathcal{S}_{1,j}^{(\eta_k-1)} = \mathcal{S}_{1,j}^{(\eta_k)}. \tag{A55}$$

Therefore, we flip arbitrarily a zero component of u^n with its index in $\mathcal{S}_{1,j}^{(\eta_k)}$ to construct a v^n such that

$$d(c_1, v^n) = d(c_1, u^n) + 1 \quad \text{and} \quad d(c_j, v^n) = d(c_j, u^n) - 1, \tag{A56}$$

which implies

$$P_{X^n, Y^n}(c_1, v^n) = P_{X^n, Y^n}(c_1, u^n) \cdot \frac{1}{q} \quad \text{and} \quad P_{X^n, Y^n}(c_j, v^n) = P_{X^n, Y^n}(c_j, u^n) \cdot q. \quad (\text{A57})$$

Then, v^n must fulfill (63a), (63c) and (63d) (with w^n replaced by v^n) as u^n satisfies (61a), (61b) and (61c). We next declare that v^n also fulfills (63b) and will prove this declaration by contradiction.

Proof of the declaration: Suppose there exists a $r \in [j - 1] \setminus \{1\}$ satisfying

$$P_{X^n, Y^n}(c_1, v^n) \cdot q^2 = P_{X^n, Y^n}(c_r, v^n). \quad (\text{A58})$$

We then recall from (45) that $d(c_1, c_r | \mathcal{S}_{1,j}^{(\eta_k)})$ is either 0 or $|\mathcal{S}_{1,j}^{(\eta_k)}|$. Thus, (A58) can be disproved by differentiating two subcases: (1) $d(c_1, c_r | \mathcal{S}_{1,j}^{(\eta_k)}) = 0$, and (2) $d(c_1, c_r | \mathcal{S}_{1,j}^{(\eta_k)}) = |\mathcal{S}_{1,j}^{(\eta_k)}|$ (Since $\ell_{1,j}^{(\eta_{k-1})} < \ell_{1,j}^{(\eta_k)}$ as can be seen from (50) and (51), we have $|\mathcal{S}_{1,j}^{(\eta_k)}| = \ell_{1,j}^{(\eta_k)} - \ell_{1,j}^{(\eta_{k-1})} > 0$, i.e., $\mathcal{S}_{1,j}^{(\eta_k)}$ non-empty).

In Subcase (1), v^n that is obtained by flipping a zero component of u^n with index in $\mathcal{S}_{1,j}^{(\eta_k)}$ must satisfy $d(c_1, v^n) = d(c_1, u^n) + 1$ and $d(c_r, v^n) = d(c_r, u^n) + 1$, which is equivalent to

$$P_{X^n, Y^n}(c_1, v^n) \cdot q = P_{X^n, Y^n}(c_1, u^n) \quad \text{and} \quad P_{X^n | Y^n}(c_r | v^n) \cdot q = P_{X^n | Y^n}(c_r | u^n). \quad (\text{A59})$$

Then, (A58) implies

$$P_{X^n, Y^n}(c_1, u^n) \cdot q^2 = P_{X^n, Y^n}(c_r, u^n). \quad (\text{A60})$$

Hence,

$$P_{X^n, Y^n}(c_1, u^n) < P_{X^n, Y^n}(c_r, u^n) \leq \max_{h \in [M] \setminus \{1\}} P_{X^n, Y^n}(x_{(h)}^n, u^n). \quad (\text{A61})$$

A contradiction to the fact that $u^n \in \mathcal{T}_{j|1}(u^n; k)$ satisfies (61a) (with y^n replaced by u^n) is obtained.

In Subcase (2), we note that $d(c_1, c_r | \mathcal{S}_{1,j}^{(\eta_k)}) = |\mathcal{S}_{1,j}^{(\eta_k)}|$ implies $\mathcal{S}_{1,j}^{(\eta_k)} \subseteq \mathcal{S}_{1,r}$. Therefore, (A55) leads to

$$d(c_1, u^n | \mathcal{S}_{1,r}) < |\mathcal{S}_{1,r}|. \quad (\text{A62})$$

The flipping manipulation on u^n results in $d(c_1, v^n) = d(c_1, u^n) + 1$ and $d(c_r, v^n) = d(c_r, u^n) - 1$, which is equivalent to

$$P_{X^n, Y^n}(c_1, v^n) \cdot q = P_{X^n, Y^n}(c_1, u^n) \quad \text{and} \quad P_{X^n, Y^n}(c_r, v^n) = P_{X^n, Y^n}(c_r, u^n) \cdot q. \quad (\text{A63})$$

Therefore, (A58) implies

$$P_{X^n, Y^n}(c_1, u^n) = P_{X^n, Y^n}(c_r, u^n), \quad (\text{A64})$$

which together with $\max_{h \in [M] \setminus \{1\}} P_{X^n, Y^n}(c_h, u^n) = P_{X^n, Y^n}(c_1, u^n)$ and (A62) result in $u^n \in \mathcal{T}_{r|1}$ because $r < j$. This contradicts $u^n \in \mathcal{T}_{j|1}$. Accordingly, v^n must also fulfill (63b); hence, $v^n \in \mathcal{N}_{j|1}(u^n; k)$. This completes the proof of the declaration.

With this auxiliary v^n , we are ready to prove that every w^n satisfying (63c) and (63d) also validates (63a) and (63b). Toward this end, we need to prove

$$P_{X^n, Y^n}(c_r, w^n) = P_{X^n, Y^n}(c_r, v^n) \quad \text{for all } r \in [M]. \quad (\text{A65})$$

Note that

$$\begin{cases} d(w^n, v^n | \mathcal{S}_{1,j}^{(\eta_k-1)}) = 0; & \text{(A66a)} \\ d(c_r, w^n | \mathcal{S}_{1,j}^{(\eta_k)}) = d(c_r, v^n | \mathcal{S}_{1,j}^{(\eta_k)}) \text{ for all } r \in [M]; & \text{(A66b)} \\ d(w^n, v^n | (\mathcal{S}_{1,j}^{(\eta_k-1)})^c) = 0, & \text{(A66c)} \end{cases}$$

where (A66a) holds because both v^n and w^n satisfy (63c), implying that all components of v^n and w^n with indices in $\mathcal{S}_{1,j}^{(\eta_k-1)}$ are equal to one; (A66b) holds because when considering only those portions with indices in (non-empty) $\mathcal{S}_{1,j}^{(\eta_k)}$, c_r gives either all ones or all zeros according to (45), and both w^n and v^n have exactly $k + 1 - \ell_{1,j}^{(\eta_k-1)}$ ones according to (63c); and (A66c) is valid since both v^n and w^n satisfy (63d). Based on (A66a)–(A66c), we remark that $d(c_r, w^n) = d(c_r, v^n)$ for all $r \in [M]$, which implies $P_{Y^n|X^n}(w^n|c_r) = P_{Y^n|X^n}(v^n|c_r)$ (equivalently, $P_{X^n,Y^n}(c_r, w^n) = P_{X^n,Y^n}(c_r, v^n)$) for all $r \in [M]$.

- (ii) $d(c_1, u^n | \mathcal{S}_{1,j}^{(\eta_k-1)}) = \ell_{1,j}^{(\eta_k-1)} - 1$: In this case, there is only one zero component of u^n with its index in $\mathcal{S}_{1,j}^{(\eta_k-1)}$. Suppose the index of such zero component lie in $\mathcal{S}_{1,j}^{(h)} \subseteq \mathcal{S}_{1,j}^{(\eta_k-1)}$, where $h \leq \eta_k - 1$. The flipping manipulation to u^n leads to v^n , which has all one components with respect to $\mathcal{S}_{1,j}^{(\eta_k-1)}$. Then, v^n must fulfill (63a), (63c), and (63d) as u^n satisfies (61a), (61b), and (61c). With the components of c_r with respect to (non-empty) $\mathcal{S}_{1,j}^{(h)}$ being either all zeros or all ones, the same contradiction argument between (A58) and (A64), with η_k replaced by h , can disprove the validity of (A58) for this v^n and for any $r \in [j - 1] \setminus \{1\}$. Therefore, v^n also fulfills (63b), implying $v^n \in \mathcal{N}_{j|1}(u^n; k)$. With this auxiliary v^n , we can again verify (A66a)–(A66c) via the same argument. The claim that w^n satisfying (63c) and (63d) validates (63a) and (63b) is thus confirmed.

References

1. Chang, L.H.; Chen, P.N.; Alajaji, F.; Han, Y.S. Decoder Ties Do Not Affect the Error Exponent of the Memoryless Binary Symmetric Channel. *IEEE Trans. Inf. Theory* **2022**, *68*, 3501–3510. [CrossRef]
2. Shannon, C.E.; Gallager, R.G.; Berlekamp, E.R. Lower bounds to error probability for coding on discrete memoryless channels—I. *Inf. Control* **1967**, *10*, 65–103. [CrossRef]
3. Shannon, C.E.; Gallager, R.G.; Berlekamp, E.R. Lower bounds to error probability for coding on discrete memoryless channels—II. *Inf. Control* **1967**, *10*, 522–552. [CrossRef]
4. McEliece, R.J.; Omura, J.K. An improved upper bound on the block coding error exponent for binary-input discrete memoryless channels. *IEEE Trans. Inf. Theory* **1977**, *23*, 611–613. [CrossRef]
5. Gallager, R.G. *Information Theory and Reliable Communication*; Wiley: New York, NY, USA, 1968.
6. Viterbi, A.J.; Omura, J.K. *Principles of Digital Communication and Coding*; McGraw-Hill: New York, NY, USA, 1979.
7. Csiszár, I.; Körner, J. *Information Theory: Coding Theorems for Discrete Memoryless Systems*; Academic Press: New York, NY, USA, 1981.
8. Blahut, R. *Principles and Practice of Information Theory*; Addison-Wesley Longman Publishing Co., Inc.: Albany, NY, USA, 1988.
9. Barg, A.; McGregor, A. Distance distribution of binary codes and the error probability of decoding. *IEEE Trans. Inf. Theory* **2005**, *51*, 4237–4246. [CrossRef]
10. Haroutunian, E.A.; Haroutunian, M.E.; Harutyunyan, A.N. Reliability Criteria in Information Theory and in Statistical Hypothesis Testing. In *Foundations and Trends in Communications and Information Theory*; Now Publishers Inc.: Delft, The Netherlands 2007; Volume 4, pp. 97–263.
11. Dalai, M. Lower bounds on the probability of error for classical and classical-quantum channels. *IEEE Trans. Inf. Theory* **2013**, *59*, 8027–8056. [CrossRef]
12. Burnashev, M.V. On the BSC reliability function: Expanding the region where it is known exactly. *Probl. Inf. Transm.* **2015**, *51*, 307–325. [CrossRef]
13. Csiszár, I. Joint source-channel error exponent. *Probl. Control. Inf. Theory* **1980**, *9*, 315–328.
14. Zhong, Y.; Alajaji, F.; Campbell, L. On the joint source-channel coding error exponent for discrete memoryless systems. *IEEE Trans. Inf. Theory* **2006**, *52*, 1450–1468. [CrossRef]

15. Alajaji, F.; Phamdo, N.; Fuja, T. Channel codes that exploit the residual redundancy in CELP-encoded speech. *IEEE Trans. Speech Audio Process.* **1996**, *4*, 325–336. [[CrossRef](#)]
16. Xu, W.; Hagenauer, J.; Hollmann, J. Joint source-channel decoding using the residual redundancy in compressed images. In Proceedings of the Proceedings of the International Conference on Communications, Washington, DC, USA, 25–28 February 1996; Volume 1, pp. 142–148.
17. Hagenauer, J. Source-controlled channel decoding. *IEEE Trans. Commun.* **1995**, *43*, 2449–2457. [[CrossRef](#)]
18. Goertz, N. *Joint Source-Channel Coding of Discrete-Time Signals with Continuous Amplitudes*; World Scientific: Singapore, 2007.
19. Duhamel, P.; Kieffer, M. *Joint Source-Channel Decoding: A Cross-Layer Perspective with Applications in Video Broadcasting*; Academic Press: Cambridge, MA, USA, 2009.
20. Fresia, M.; Pérez-Cruz, F.; Poor, H.V.; Verdú, S. Joint source and channel coding. *IEEE Signal Process. Mag.* **2010**, *27*, 104–113. [[CrossRef](#)]
21. Alajaji, F.; Chen, P.N. *An Introduction to Single-User Information Theory*; Springer: Berlin/Heidelberg, Germany, 2018.
22. Chang, L.H.; Chen, P.N.; Alajaji, F.; Han, Y.S. The asymptotic generalized Poor-Verdú bound achieves the BSC error exponent at zero rate. In Proceedings of the IEEE International Symposium on Information Theory, Los Angeles, CA, USA, 21–26 June 2020.
23. Chen, P.N.; Alajaji, F. A generalized Poor-Verdú error bound for multihypothesis testings. *IEEE Trans. Inf. Theory* **2012**, *58*, 311–316. [[CrossRef](#)]
24. Poor, H.V.; Verdú, S. A lower bound on the probability of error in multihypothesis testing. *IEEE Trans. Inf. Theory* **1995**, *41*, 1992–1994. [[CrossRef](#)]
25. Chang, L.H.; Chen, P.N.; Alajaji, F.; Han, Y.S. Tightness of the asymptotic generalized Poor-Verdú error bound for the memoryless symmetric channel. *arXiv* **2020**, arXiv: 2007.04080v1.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.