Decoder Ties Do Not Affect the Error Exponent of the Memoryless Binary Symmetric Channel

Ling-Hua Chang*, Po-Ning Chen†, Fady Alajaji‡ and Yunghsiang S. Han§

***Abstract*—The generalized Poor-Verdú error lower bound established in [1] for multihypothesis testing is studied in the classical channel coding context. It is proved that for any sequence of block codes sent over the memoryless binary symmetric channel (BSC), the minimum probability of error (under maximum likelihood decoding) has a relative deviation from the generalized bound that grows at most linearly in blocklength. This result directly implies that for arbitrary codes used over the BSC, decoder ties can only affect the subexponential behavior of the minimum probability of error.**

***Index Terms*—Binary symmetric channel, block codes, error probability bounds, maximum likelihood decoder ties, error exponent, channel reliability function, hypothesis testing.**

I. Introduction

A well-known lower bound on the minimum probability of error P_{e} of multihypothesis testing is the so-called Poor-Verdú bound [2]. The bound was generalized in [3] by tilting, via a parameter $\theta \geq 1$, the posterior hypothesis distribution, with the resulting bound noted to progressively improve with θ except for examples involving the memoryless binary erasure channel (BEC). The closed-form formula of this generalized Poor-Verdú bound, as θ tends to infinity, was recently derived in [1]. An alternative lower bound for P_{e} was established by Verdú and Han in [4]; this bound was subsequently extended and strengthened in [5].

In this paper, we investigate the generalized Poor-Verdú lower bound of [1] in the classical context of the maximum-likelihood (ML) decoding error probability of block codes $\mathcal{C}_n$ with blocklength n and size $|\mathcal{C}_n| = M$ sent over the memoryless binary symmetric channel (BSC) with crossover probability $0 < p < 1/2$. For convenience, we denote this lower bound by b_n (see its expression in (3)). Specifically, for channel inputs uniformly distributed over code $\mathcal{C}_n$, we bound the codes minimum probability of decoding error a_n in terms[1] of b_n as follows:

$$b_n \leq a_n \leq (1 + c\,n)\, b_n, \tag{1}$$

where $c \triangleq (1-p)/p$ is the channel (likelihood ratio) constant and is independent of code $\mathcal{C}_n$. Noting that b_n can be recovered from a_n by disregarding all decoder ties, which occur with probability no larger than $c\,n\,b_n$, we conclude that decoder ties only affect the subexponential behavior of the minimum error probability a_n with respect to an arbitrary sequence of codes $\{\mathcal{C}_n\}_{n\geq 1}$. Indeed, setting $\mathcal{C}_n^{a*} \triangleq \arg\min_{\mathcal{C}_n:|\mathcal{C}_n|=M} a_n(\mathcal{C}_n)$ and $\mathcal{C}_n^{b*} \triangleq \arg\min_{\mathcal{C}_n:|\mathcal{C}_n|=M} b_n(\mathcal{C}_n)$ for codes of blocklength n and size M used over the BSC, (1) implies that:

$$\begin{aligned} b_n(\mathcal{C}_n^{b*}) &\leq b_n(\mathcal{C}_n^{a*}) \leq a_n(\mathcal{C}_n^{a*}) \\ &\leq a_n(\mathcal{C}_n^{b*}) \leq \left(1 + \frac{(1-p)}{p} n\right) b_n(\mathcal{C}_n^{b*}), \end{aligned} \tag{2}$$

which immediately gives that with $M = \lfloor e^{nR} \rfloor$, the BSC reliability function $E(R) \triangleq \limsup_{n\to\infty} -\frac{1}{n}\log a_n(\mathcal{C}_n^{a*})$ satisfies $E(R) = \limsup_{n\to\infty} -\frac{1}{n}\log b_n(\mathcal{C}_n^{b*})$ and can hence be determined via a sequence of codes that minimizes $b_n(\mathcal{C}_n)$ (without considering ties) instead of $a_n(\mathcal{C}_n)$.

The related problem of exactly characterizing the channel reliability function at low rates remains a long-standing open problem; in-depth studies on this focal information-theoretic function and related problems include the classical papers [6]–[9] and texts [10]–[13] and the more recent works [14]–[25] (see also the references therein). In [2], Poor and Verdú conjectured that their original error lower bound for multihypothesis testing, which yields an upper bound on the channel coding reliability function, is tight for all rates and arbitrary channels. The conjecture was disproved in [26], where the bound was shown to be loose for the BEC at low rates. Furthermore, Polyanskiy showed in [17] that the original Poor-Verdú bound [2] coincides with the sphere-packing error exponent bound for discrete memoryless channels (and is hence loose at low rates for this entire class of channels). Our result in (1) which holds for arbitrary sequence of codes $\{\mathcal{C}_n\}_{n\geq 1}$, while not explicitly determining the reliability function for the BSC, provides an alternative approach for studying it.

The rest of the paper is organized as follows. The error bound b_n is analyzed for the channel coding problem over the memoryless BSC in Section II. The proof of the main theorem is provided in detail in Section III. Finally, conclusions are drawn in Section IV.

*Department of Electrical Engineering, Yuan Ze University, Taiwan, R.O.C. (iamjaung@gmail.com).

†Institute of Communications Engineering and Department of Electrical and Computer Engineering, National Yang-Ming Chiao-Tung University, Taiwan, R.O.C. (poningchen@nycu.edu.tw).

‡Department of Mathematics and Statistics, Queen's University, Kingston, ON, Canada (fa@queensu.ca).

§Shenzhen Institute for Advanced Study, University of Electronic Science and Technology of China, Shenzhen, Guangdong, China (yunghsiangh@gmail.com).

The work of Ling-Hua Chang is supported by the Ministry of Science and Technology, Taiwan, R.O.C. (Grant No. MOST 109-2221-E-155 -035 -MY3). The work of Po-Ning Chen is supported by the Ministry of Science and Technology, Taiwan, R.O.C. (Grant No. MOST 110-2221-E-A49-024-MY3). The work of Fady Alajaji is supported by the Natural Sciences and Engineering Research Council of Canada.

[1]Note that a_n and b_n, as well as the notations introduced in Table I, are all functions of the adopted code $\mathcal{C}_n$. For ease of notation, we drop their dependence on $\mathcal{C}_n$ throughout the paper (except in (2) and the discussion related to it).

Throughout the paper, we denote $[M] \triangleq \{1, 2, \ldots, M\}$ for any positive integer M.

II. Analysis of Lower Bound b_n for an Arbitrary Sequence of Binary Codes $\{\mathcal{C}_n\}_{n\geq 1}$

Consider an arbitrary binary code $\mathcal{C}_n \in \{0,1\}^n$ with blocklength n and size $|\mathcal{C}_n| = M$ to be used over the BSC with crossover probability $0 < p < \frac{1}{2}$. It is shown in [1, Eq. (5)] that the generalized Poor-Verdú error lower bound b_n to the minimum probability of decoding error a_n (obtained under maximum-likelihood decoding) is given by

$$b_n = P_{X^n,Y^n}\Big\{(x^n, y^n) \in \mathcal{X}^n \times \mathcal{Y}^n : P_{X^n|Y^n}(x^n|y^n) < \max_{u^n \in \mathcal{C}_n \setminus \{x^n\}} P_{X^n|Y^n}(u^n|y^n)\Big\}, \quad (3)$$

where P_{X^n,Y^n} is the joint input-output distribution that X^n is sent over the BSC (via n uses) and Y^n is received, and $P_{X^n|Y^n}$ is the corresponding posterior conditional distribution of X^n given Y^n. Indeed, by recalling that the (optimal) maximum a posteriori (MAP) estimate of $x^n \in \mathcal{C}_n$ from observing $y^n \in \mathcal{Y}^n$ at the channel output is given by

$$e(y^n) = \arg\max_{x^n \in \mathcal{C}_n} P_{X^n|Y^n}(x^n|y^n), \quad (4)$$

the right-hand-side (RHS) of (3) is nothing but the error probability under a "genie" MAP decoder that correctly resolves ties. We demonstrate that the lower bound b_n in (3), upon scaling it by the affine linear term $(1 + c\,n)$, where $c = (1-p)/p$, becomes an upper bound for a_n, and hence is asymptotically exponentially tight with a_n $\big($i.e., $\limsup_{n\to\infty} \frac{1}{n} \log \frac{a_n}{b_n} = 0\big)$ for arbitrary sequences of block codes sent over the BSC. The exponential tightness result follows directly from the following theorem, which is the main contribution of the paper.

Theorem 1: For any sequence of codes $\{\mathcal{C}_n\}_{n\geq 1}$ of blocklength n and size $|\mathcal{C}_n| = M$ with $\mathcal{C}_n \subseteq \mathcal{X}^n \triangleq \{0,1\}^n$, let a_n denote the minimum probability of decoding error for transmitting $\mathcal{C}_n$ over the BSC with crossover probability $0 < p < 1/2$, under a uniform distribution P_{X^n} over $\mathcal{C}_n$, where X^n is the n-tuple $(X_1, \ldots, X_n)$. Then,

$$b_n \leq a_n \leq \left(1 + \frac{(1-p)}{p} n\right) b_n, \quad (5)$$

where b_n is given in (3).

In Theorem 1, it is implicitly assumed that all M codewords are distinct as the codebook is defined as $\mathcal{C}_n \in \mathcal{X}^n$, with $\mathcal{X}^n$ containing all (distinct) binary sequences of length n. Note that if identical codewords are allowed in $\mathcal{C}_n$, decoder ties may become dominant in the minimum error probability a_n and the right inequality (5) in Theorem 1 no longer holds. Theorem 1 reveals that for any *arbitrary* sequence of block codes $\{\mathcal{C}_n\}_{n\geq 1}$ used over the BSC, the relative deviation, $(a_n - b_n)/b_n$, of the minimum probability of decoding error a_n from b_n is *at most linear* in the blocklength n. It is worth mentioning that this conclusion cannot be applied for the BEC for any code $\mathcal{C}_n$ because decoder ties are the only source of decoding errors on the BEC, which gives $b_n = 0$ since (3) ignores ties.

Overview of the Proof of Theorem 1: Before providing the full proof of Theorem 1 in Section III, we introduce the necessary notation and highlight how we prove (5).

Because the channel input distribution P_{X^n} is uniform over $\mathcal{C}_n$, the code's minimal probability of error a_n is achieved under ML decoding. For the BSC, the ML estimate based on any received n-tuple y^n at the channel output is obtained via the Hamming distances $\{d(x^n, y^n)\}_{x^n \in \mathcal{C}_n, y^n \in \mathcal{Y}^n}$. Define the set of output n-tuples y^n which definitely lead to an ML decoder error when $x^n_{(i)} \in \mathcal{C}_n$ is transmitted as

$$\mathcal{N}_i \triangleq \left\{y^n \in \mathcal{Y}^n : d(x^n_{(i)}, y^n) > \min_{u^n \in \mathcal{C}_n \setminus \{x^n_{(i)}\}} d(u^n, y^n)\right\}, \quad (6)$$

and the set of output n-tuples y^n that induce a decoder tie when transmitting $x^n_{(i)} \in \mathcal{C}_n$ as

$$\mathcal{T}_i \triangleq \left\{y^n \in \mathcal{Y}^n : d(x^n_{(i)}, y^n) = \min_{u^n \in \mathcal{C}_n \setminus \{x^n_{(i)}\}} d(u^n, y^n)\right\}. \quad (7)$$

For the BSC with crossover probability $0 < p < \frac{1}{2}$, we have $P_{Y^n|X^n}(y^n|x^n_{(i)}) = \left(\frac{p}{1-p}\right)^{d(x^n_{(i)}, y^n)}(1-p)^n$. Thus, $d(x^n_{(i)}, y^n) > \min_{u^n \in \mathcal{C}_n \setminus \{x^n_{(i)}\}} d(u^n, y^n)$ if and only if $P_{Y^n|X^n}(y^n|x^n_{(i)}) < \max_{u^n \in \mathcal{C}_n \setminus \{x^n_{(i)}\}} P_{Y^n|X^n}(y^n|u^n)$, and therefore

$$\begin{aligned} b_n &= \sum_{i=1}^{M} P_{X^n}(x^n_{(i)}) P_{Y^n|X^n}(\mathcal{N}_i|x^n_{(i)}) \\ &= \frac{1}{M} \sum_{i\in[M]} P_{Y^n|X^n}(\mathcal{N}_i|x^n_{(i)}). \end{aligned} \quad (8)$$

Similarly, $P_{Y^n|X^n}(y^n|x^n_{(i)}) = \left(\frac{p}{1-p}\right)^{d(x^n_{(i)}, y^n)}(1-p)^n$ implies that the probability of decoder ties, denoted by δ_n, satisfies

$$\begin{aligned} \delta_n &= \sum_{i=1}^{M} P_{X^n}(x^n_{(i)}) P_{Y^n|X^n}\big(\mathcal{T}_i|x^n_{(i)}\big) \\ &= \frac{1}{M} \sum_{i\in[M]} P_{Y^n|X^n}\big(\mathcal{T}_i|x^n_{(i)}\big). \end{aligned} \quad (9)$$

We thus obtain the following relationship:

$$b_n \leq a_n \leq b_n + \delta_n = \left(1 + \frac{\delta_n}{b_n}\right) b_n. \quad (10)$$

Note if $\delta_n = 0$,[2] then (10) is tight and (5) holds trivially; so without loss of generality, we will assume in the proof that $\delta_n > 0$. We then have that

$$\begin{aligned} \frac{\delta_n}{b_n} &= \frac{\sum_{i\in[M]} P_{Y^n|X^n}(\mathcal{T}_i|x^n_{(i)})}{\sum_{i\in[M]} P_{Y^n|X^n}(\mathcal{N}_i|x^n_{(i)})} && (11) \\ &\leq \frac{\sum_{i\in[M]:\mathcal{T}_i\neq\emptyset} P_{Y^n|X^n}(\mathcal{T}_i|x^n_{(i)})}{\sum_{i\in[M]:\mathcal{T}_i\neq\emptyset} P_{Y^n|X^n}(\mathcal{N}_i|x^n_{(i)})} && (12) \\ &\leq \frac{1}{\sum_{i\in[M]:\mathcal{T}_i\neq\emptyset} P_{Y^n|X^n}(\mathcal{N}_i|x^n_{(i)})} \sum_{i\in[M]:\mathcal{T}_i\neq\emptyset} \Bigg(P_{Y^n|X^n}(\mathcal{N}_i|x^n_{(i)}) \\ &\qquad \times \max_{i'\in[M]:\mathcal{T}_{i'}\neq\emptyset} \frac{P_{Y^n|X^n}(\mathcal{T}_{i'}|x^n_{(i')})}{P_{Y^n|X^n}(\mathcal{N}_{i'}|x^n_{(i')})}\Bigg) && (13) \\ &= \max_{i'\in[M]:\mathcal{T}_{i'}\neq\emptyset} \frac{P_{Y^n|X^n}(\mathcal{T}_{i'}|x^n_{(i')})}{P_{Y^n|X^n}(\mathcal{N}_{i'}|x^n_{(i')})}, && (14) \end{aligned}$$

where (12) holds because the assumption of $\delta_n > 0$ guarantees

[2] A straightforward example for which $\delta_n = 0$ is $\mathcal{C}_n$ consisting of only two codewords whose Hamming distance is an odd number.

the existence of at least one non-empty set $\mathcal{T}_i$ for $i \in [M]$. With (10) and (14), the upper bound in (5) follows by proving that

$$\frac{P_{Y^n|X^n}(\mathcal{T}_i|x^n_{(i)})}{P_{Y^n|X^n}(\mathcal{N}_i|x^n_{(i)})} \le \frac{(1-p)}{p}n \quad \text{for non-empty } \mathcal{T}_i. \qquad (15)$$

To achieve this objective, we will construct a number of disjoint covers of $\mathcal{T}_i$ and also construct the same number of disjoint subsets of $\mathcal{N}_i$ such that a one-to-one correspondence between the $\mathcal{T}_i$-covers and the $\mathcal{N}_i$-subsets exists. Since $P_{Y^n|X^n}(\mathcal{T}_i|x^n_{(i)}) > 0$ guarantees the existence of at least one non-empty $\mathcal{T}_i$-cover, a similar derivation to (14) yields that $\frac{P_{Y^n|X^n}(\mathcal{T}_i|x^n_{(i)})}{P_{Y^n|X^n}(\mathcal{N}_i|x^n_{(i)})}$ is upper-bounded by the maximum ratio of the probabilities of the $\mathcal{T}_i$-cover-versus-$\mathcal{N}_i$-subset pairs. The final step (i.e., Proposition 4 in Section III-D) is to enumerate the probabilities of the $\mathcal{T}_i$-cover-versus-$\mathcal{N}_i$-subset pairs and show that it is bounded from above by $\frac{(1-p)}{p}n$. The full details are given in the next section.

III. The Proof of Theorem 1

We divide the proof into four parts. In Section III-A, we obtain a coarse disjoint covering of (non-empty) $\mathcal{T}_i$ and the corresponding disjoint subsets of $\mathcal{N}_i$. In Sections III-B and III-C, we refine the covers of $\mathcal{T}_i$ just obtained by further partitioning each of them in a systematic manner, and the same number of disjoint subsets of $\mathcal{N}_i$ are also constructed. In Section III-D, we enumerate the refined covering sets of $\mathcal{T}_i$ and the corresponding subsets of $\mathcal{N}_i$, which enable us to obtain the desired upper bound for δ_n/b_n. Since we consider the memoryless BSC in this paper, we assume without loss of generality that $x^n_{(1)}$ is the all-zero codeword. We also assume for notational convenience that $i = 1$ and $\mathcal{T}_1 \neq \emptyset$.

For ease of reference, we first summarize in Table I all main symbols used in the proof. We also illustrate in Fig. 1 all sets defined in Table I, based on the code of Example 1 below.

A. A Coarse Disjoint Covering of Non-empty $\mathcal{T}_1$ and the Corresponding Disjoint Subsets of $\mathcal{N}_1$

Before providing a coarse disjoint covering of non-empty $\mathcal{T}_1$ and corresponding disjoint subsets of $\mathcal{N}_1$, we elucidate the idea behind them.

Note from its definition in (7) that $\mathcal{T}_1$ consists of all minimum distance ties when $x^n_{(1)}$ is sent. To obtain disjoint covers of $\mathcal{T}_1$, we first collect all channel outputs y^n that are equidistant from $x^n_{(1)}$ and $x^n_{(2)}$ and we place them in $\mathcal{T}_{2|1}$. We next place into $\mathcal{T}_{3|1}$ those outputs y^n that have *not* been included in $\mathcal{T}_{2|1}$, and that are at equal distance from $x^n_{(1)}$ and $x^n_{(3)}$. We iterate this process sequentially to obtain $\mathcal{T}_{j|1}$ for $j = 4, 5, \ldots, M$ by picking y^n tuples that have not yet been included in all previous collections, and that are equidistant from $x^n_{(1)}$ and $x^n_{(j)}$. This completes the construction of the disjoint covers $\{\mathcal{T}_{j|1}\}_{j=2}^M$ of $\mathcal{T}_1$. Note that for non-empty $\mathcal{T}_1$, we have at least one $\mathcal{T}_{j|1}$ that is non-empty.

The $(M-1)$ disjoint subsets of $\mathcal{N}_i$ are constructed as follows. Suppose $\mathcal{T}_{2|1}$ is non-empty. Given a channel output u^n in $\mathcal{T}_{2|1}$ (that is at equal distance from $x^n_{(1)}$ and $x^n_{(2)}$), we can flip a zero component of u^n to obtain a v^n to fulfill $d(x^n_{(1)}, v^n) - 1 = d(x^n_{(1)}, u^n) = d(x^n_{(2)}, u^n) = d(x^n_{(2)}, v^n) + 1$, implying $d(x^n_{(1)}, v^n) > d(x^n_{(2)}, v^n) \ge \min_{z^n \in \mathcal{C}_n \setminus \{x^n_{(1)}\}} d(z^n, v^n)$. Therefore, it follows from the definition in (6) that $v^n \in \mathcal{N}_1$. Collecting all such v^n from every $u^n \in \mathcal{T}_{2|1}$, we form $\mathcal{N}_{2|1}$. This construction provides an operational connection between $\mathcal{T}_{2|1}$ and $\mathcal{N}_{2|1}$. Iterating this process for $j = 3, 4, \ldots, M$ in this order and deliberately avoiding repeated collections give the desired disjoint subsets of $\mathcal{N}_1$. Here, we force $\mathcal{N}_{j|1} = \emptyset$ whenever $\mathcal{T}_{j|1}$ is an empty set.

The above constructions are formalized in the following definition.

Definition 1: Define for $j \in [M] \setminus \{1\}$,

$$\begin{cases} \mathcal{T}_{j|1} \triangleq \Big\{ y^n \in \mathcal{Y}^n : d(x^n_{(1)}, y^n) = d(x^n_{(j)}, y^n) \\ \qquad\qquad < \min_{r \in [j-1]\setminus\{1\}} d(x^n_{(r)}, y^n) \Big\}; & (16a) \\ \mathcal{N}_{j|1} \triangleq \Big\{ y^n \in \mathcal{Y}^n : \\ \qquad d(x^n_{(1)}, y^n) - 1 = d(x^n_{(j)}, y^n) + 1 \\ \qquad \neq d(x^n_{(r)}, y^n) + 1 \text{ for } r \in [j-1] \setminus \{1\} \Big\}. & (16b) \end{cases}$$

To better understand the terms just introduced, we provide the following example.

Example 1: Suppose $M = 3$ and $\mathcal{C}_4 = \{x^4_{(1)}, x^4_{(2)}, x^4_{(3)}\} = \{0000, 1100, 0110\}$. Then, $\mathcal{T}_1 = \{0100, 1000, 0101, 1001, 1010, 1011, 0010, 0011\}$ and $\mathcal{N}_1 = \{1100, 0110, 0111, 1101, 1110, 1111\}$. Furthermore, we have $\mathcal{T}_{2|1} = \{0100, 1000, 0101, 1001, 1010, 1011, 0110, 0111\}$ and $\mathcal{T}_{3|1} = \{0010, 0011\}$. Note that the last two elements in $\mathcal{T}_{2|1}$ satisfy both $d(x^n_{(1)}, y^n) = d(x^n_{(2)}, y^n)$ and $d(x^n_{(1)}, y^n) > d(x^n_{(3)}, y^n)$, and hence they result in *ties* but not in *minimum distance ties* as required for $\mathcal{T}_1$ in (7), indicating that $\mathcal{T}_{2|1} \cup \mathcal{T}_{3|1}$ is a proper covering of $\mathcal{T}_1$ as shown in Fig. 1. On the other hand, we have $\mathcal{N}_{2|1} = \{1100, 1101, 1110, 1111\}$ and $\mathcal{N}_{3|1} = \{0110, 0111\}$, showing that they are disjoint subsets of $\mathcal{N}_1$. □

The observations we made from Example 1 are proved in the next proposition.

Proposition 1: For nonempty $\mathcal{T}_1$, the following two properties hold.

i) The collection $\{\mathcal{T}_{j|1}\}_{j\in[M]\setminus\{1\}}$ forms a disjoint covering of $\mathcal{T}_1$.

ii) $\{\mathcal{N}_{j|1}\}_{j\in[M]\setminus\{1\}}$ is a collection of disjoint subsets of $\mathcal{N}_i$.

Proof: The strict inequality in (16a) and the non-equality condition in (16b) guarantee no multiple inclusions of an element from the previous collections; therefore, $\{\mathcal{T}_{j|1}\}_{j\in[M]\setminus\{1\}}$ are disjoint and so are $\{\mathcal{N}_{j|1}\}_{j\in[M]\setminus\{1\}}$. Now for any $y^n \in \mathcal{T}_1$, we have $d(x^n_{(1)}, y^n) = d(x^n_{(m)}, y^n)$ for some $m \neq 1$; therefore, this y^n must be collected in $\mathcal{T}_{j|1}$ for some $j \le m$, confirming that $\{\mathcal{T}_{j|1}\}_{j\in[M]\setminus\{1\}}$ forms a covering of $\mathcal{T}_1$. Next, for any $y^n \in \mathcal{N}_{j|1}$, we have $d(x^n_{(1)}, y^n) - 1 = d(x^n_{(j)}, y^n) + 1 \ge \min_{u^n \in \mathcal{C}_n \setminus \{x^n_{(1)}\}} d(u^n, y^n) + 1$, leading to $d(x^n_{(1)}, y^n) > d(x^n_{(j)}, y^n) \ge \min_{u^n \in \mathcal{C}_n \setminus \{x^n_{(1)}\}} d(u^n, y^n)$; hence, this y^n must be contained in $\mathcal{N}_1$, confirming that $\{\mathcal{N}_{j|1}\}_{j\in[M]\setminus\{1\}}$ are subsets of $\mathcal{N}_i$. ■

From Proposition 1, we have that

$$\frac{P_{Y^n|X^n}(\mathcal{T}_1|x^n_{(1)})}{P_{Y^n|X^n}(\mathcal{N}_1\big|x^n_{(1)})} \le \frac{P_{Y^n|X^n}\big(\bigcup_{j\in[M]\setminus\{1\}} \mathcal{T}_{j|1}\big|x^n_{(1)}\big)}{P_{Y^n|X^n}\big(\bigcup_{j\in[M]\setminus\{1\}} \mathcal{N}_{j|1}\big|x^n_{(1)}\big)}$$

TABLE I

SUMMARY OF ALL MAIN SYMBOLS USED IN THE PROOF.

Symbol	Description	Definition
$[M]$	A shorthand for $\{1, 2, \ldots, M\}$	
$\mathcal{C}_n$	The code $\{x_1^{(n)}, x_2^{(n)}, \ldots, x_M^{(n)}\}$ with $x_1^{(n)}$ being the all-zero codeword	
$d(u^n, v^n \vert \mathcal{S})$	The Hamming distance between the portions of u^n and v^n with indices in $\mathcal{S}$	
All terms below are functions of $\mathcal{C}_n$ (this dependence is not explicitly shown to simplify notation)		
$\mathcal{N}_j$	The set of channel outputs y^n that lead to an ML decoder error when $x_{(i)}^n$ is sent	(6)
$\mathcal{T}_j$	The set of channel outputs y^n that induce a decoder tie when $x_{(i)}^n$ is sent	(7)
$\mathcal{T}_{j\vert 1}$	The set of channel outputs y^n that are at equal distance from $x_{(1)}^n$ and $x_{(j)}^n$ and that are not included in $\mathcal{T}_{i\vert 1}$ for $2 \le i \le j-1$	(16a)
$\mathcal{N}_{j\vert 1}$	The set of channel outputs y^n that satisfy $d(x_{(1)}^n, y^n) - 1 = d(x_{(j)}^n, y^n) + 1$ and that are not included in $\mathcal{N}_{i\vert 1}$ for $2 \le i \le j-1$	(16b)
$\mathcal{S}_j$	The set of indices for which the components of $x_{(j)}^n$ equal one	
ℓ_j	The size of $\mathcal{S}_j$, i.e., $\vert\mathcal{S}_j\vert$	
$\mathcal{S}_{r;\lambda_r}$	It is equal to $\mathcal{S}_r$ if $\lambda_r = 1$, and $\mathcal{S}_r^{\mathrm{c}}$ if $\lambda_r = 0$ (only used in (19) to define $\mathcal{S}_j^{(m)}$)	
$\mathcal{S}_j^{(m)}$	The subset of $\mathcal{S}_j$ defined according to whether each index in $\mathcal{S}_j$ is in each of $\mathcal{S}_2, \ldots, \mathcal{S}_{j-2}$	(19)
$\mathscr{S}_j^{(m)}$	The union of $\mathcal{S}_j^{(1)}, \mathcal{S}_j^{(2)}, \ldots, \mathcal{S}_j^{(m)}$	(20)
$\ell_j^{(m)}$	The size of $\mathscr{S}_j^{(m)}$, i.e., $\vert\mathscr{S}_j^{(m)}\vert$	
$\sigma(\cdot)$	The mapping from $\{0, 1, \ldots, \ell_j - 1\}$ to $[2^{j-2}]$ for partitioning $\mathcal{T}_{j\vert 1}$ into ℓ_j subsets $\{\mathcal{T}_{j\vert 1}(k)\}_{0 \le k < \ell_j}$	(24)
$\mathcal{T}_{j\vert 1}(k)$	The kth partition of $\mathcal{T}_{j\vert 1}$ for $k = 0, 1, \ldots, \ell_j - 1$	(25a)
$\mathcal{N}_{j\vert 1}(k)$	The kth subset of $\mathcal{N}_{j\vert 1}$ for $k = 0, 1, \ldots, \ell_j - 1$	(25b)
$\mathcal{U}_{j\vert 1}(k)$	The group of representative elements in $\mathcal{T}_{j\vert 1}(k)$ for defining the partitions of $\mathcal{T}_{j\vert 1}(k)$	
$\mathcal{T}_{j\vert 1}(u^n; k)$	The partition of $\mathcal{T}_{j\vert 1}(k)$ associated with $u^n \in \mathcal{U}_{j\vert 1}(k)$	(30a)
$\mathcal{N}_{j\vert 1}(u^n; k)$	The subset of $\mathcal{N}_{j\vert 1}(k)$ associated with $u^n \in \mathcal{U}_{j\vert 1}(k)$	(30b)

$$= \frac{\sum_{j\in[M]\setminus\{1\}} P_{Y^n|X^n}\left(\mathcal{T}_{j|1}\middle|x_{(1)}^n\right)}{\sum_{j\in[M]\setminus\{1\}} P_{Y^n|X^n}\left(\mathcal{N}_{j|1}\middle|x_{(1)}^n\right)}, \quad (17)$$

which implies, using the same method to derive (14), that

$$\frac{P_{Y^n|X^n}(\mathcal{T}_1|x_{(1)}^n)}{P_{Y^n|X^n}(\mathcal{N}_1|x_{(1)}^n)} \le \max_{j\in[M]\setminus\{1\}:\mathcal{T}_{j|1}\neq\emptyset} \frac{P_{Y^n|X^n}(\mathcal{T}_{j|1}|x_{(1)}^n)}{P_{Y^n|X^n}(\mathcal{N}_{j|1}|x_{(1)}^n)} \quad (18)$$

for non-empty $\mathcal{T}_1$.

In the next section, we continue decomposing non-empty $\mathcal{T}_{j|1}$ and its corresponding $\mathcal{N}_{j|1}$.

B. A Partition of Non-empty $\mathcal{T}_{j|1}$ and the Corresponding Disjoint Subsets of $\mathcal{N}_{j|1}$

For the enumeration analysis in Section III-D, further decompositions of $\mathcal{T}_{j|1}$ and $\mathcal{N}_{j|1}$ are needed in order to facilitate the identification of which portions of $x_{(r)}^n$ are ones and which portions of $x_{(r)}^n$ are zeros for *every* $r \in [j]$. Let $\mathcal{S}_j$ denote the set of indices for which the (bit) components of $x_{(j)}^n$ equal one.

Now as an example, if we decompose $\mathcal{S}_3$ into $\mathcal{S}_2^{\mathrm{c}} \bigcap \mathcal{S}_3$ and $\mathcal{S}_2 \bigcap \mathcal{S}_3$, then we are certain that the portions of $x_{(2)}^n$ with indices in $\mathcal{S}_2^{\mathrm{c}} \bigcap \mathcal{S}_3$ are zeros, and those with indices in $\mathcal{S}_2 \bigcap \mathcal{S}_3$ are ones. Furthermore, when considering the portions of $x_{(4)}^n$ that are ones, $\mathcal{S}_4$ can be decomposed into $\mathcal{S}_2^{\mathrm{c}} \bigcap \mathcal{S}_3^{\mathrm{c}} \bigcap \mathcal{S}_4$, $\mathcal{S}_2^{\mathrm{c}} \bigcap \mathcal{S}_3 \bigcap \mathcal{S}_4$, $\mathcal{S}_2 \bigcap \mathcal{S}_3^{\mathrm{c}} \bigcap \mathcal{S}_4$ and $\mathcal{S}_2 \bigcap \mathcal{S}_3 \bigcap \mathcal{S}_4$, and the values of $x_{(2)}^n$ and $x_{(3)}^n$ are known exactly when considering their portions with indices in any of these four sets. In general, we shall partition S_j into 2^{j-2} subsets based on $\mathcal{S}_2, \mathcal{S}_3, \ldots, \mathcal{S}_{j-1}$ and their respective complements. As such, $\mathcal{S}_4$ is partitioned into $2^{j-2} = 4$ subsets (here $j = 4$). For convenience, we use the positive integer $m \triangleq 1 + \sum_{r=2}^{j-1} \lambda_r \cdot 2^{r-2}$, where $1 \le m \le 2^{j-2}$, to enumerate the 2^{j-2} joint intersections, where $\lambda_r = 0$ implies $\mathcal{S}_r^{\mathrm{c}}$ is involved in the joint intersections, while $\lambda_r = 1$ implies $\mathcal{S}_r$ is taken instead. Thus, with $j = 4$, the four sets $\mathcal{S}_2^{\mathrm{c}} \bigcap \mathcal{S}_3^{\mathrm{c}} \bigcap \mathcal{S}_4$, $\mathcal{S}_2 \bigcap \mathcal{S}_3^{\mathrm{c}} \bigcap \mathcal{S}_4$, $\mathcal{S}_2^{\mathrm{c}} \bigcap \mathcal{S}_3 \bigcap \mathcal{S}_4$ and $\mathcal{S}_2 \bigcap \mathcal{S}_3 \bigcap \mathcal{S}_4$ are respectively indexed by $m = 1$, 2, 3 and 4, which correspond to $(\lambda_2, \lambda_3) = (0, 0)$, $(1, 0)$, $(0, 1)$ and $(1, 1)$, respectively.

For $j \in [M] \setminus \{1\}$, partition $\mathcal{S}_j$ into 2^{j-2} subsets according to whether each index in $\mathcal{S}_j$ is in $\mathcal{S}_2, \ldots, \mathcal{S}_{j-2}$ or not as follows:

$$\mathcal{S}_j^{(m)} \triangleq \left(\bigcap_{r=2}^{j-2} \mathcal{S}_{r;\lambda_r}\right) \bigcap \mathcal{S}_j$$

$$\text{for } 1 \le m = 1 + \sum_{r=2}^{j-1} \lambda_r \cdot 2^{r-2} \le 2^{j-2}, \quad (19)$$

where $\mathcal{S}_{r;1} \triangleq \mathcal{S}_r$ and $\mathcal{S}_{r;0} \triangleq \mathcal{S}_r^{\mathrm{c}}$, and each $\lambda_r \in \{0, 1\}$. Define

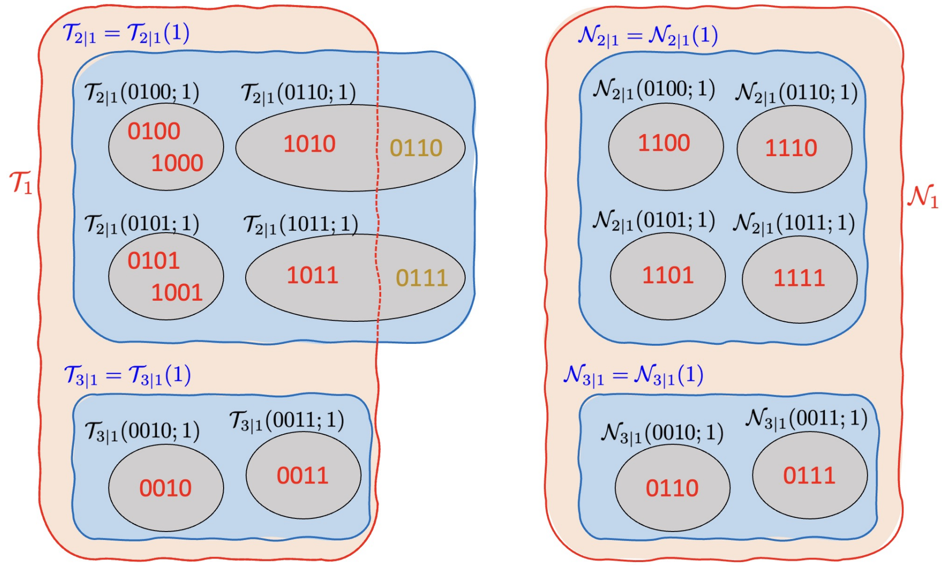

Fig. 1. An illustration of the sets defined in Table I, based on the setting in Example 1, where $\mathcal{T}_{2|1}(0) = \mathcal{T}_{3|1}(0) = \mathcal{N}_{2|1}(0) = \mathcal{N}_{3|1}(0) = \emptyset$, $\mathcal{U}_{2|1}(1) = \{0100, 0101, 0110, 1011\}$ and $\mathcal{U}_{3|1}(1) = \{0010, 0011\}$.

incrementally $\mathscr{S}_j^{(0)} \triangleq \emptyset$ and

$$\mathscr{S}_j^{(m)} \triangleq \bigcup_{q=1}^{m} \mathcal{S}_j^{(q)}, \quad m \in [2^{j-2}]. \tag{20}$$

Let $\ell_j \triangleq |\mathcal{S}_j|$ and $\ell_j^{(m)} \triangleq |\mathscr{S}_j^{(m)}|$ denote the sizes of $\mathcal{S}_j$ and $\mathscr{S}_j^{(m)}$, respectively. Then, as mentioned at the beginning of this section, for all $r \in [j]$, the components of $x_{(r)}^n$ with indices in $\mathcal{S}_j^{(m)}$ can now be unambiguously identified and are all equal to λ_r. As a result, with $x_{(1)}^n$ being the all-zero codewords,

$$d(x_{(1)}^n, x_{(r)}^n | \mathcal{S}_j^{(m)}) = \begin{cases} |\mathcal{S}_j^{(m)}|, & \lambda_r = 1; \\ 0, & \lambda_r = 0, \end{cases} \tag{21}$$

where $d(u^n, v^n|\mathcal{S})$ denotes the Hamming distance between the portions of u^n and v^n with indices in $\mathcal{S}$, and by convention, we set $d(u^n, v^n|\mathcal{S}) = 0$ when $\mathcal{S} = \emptyset$. We will see later in the proof of Proposition 4 that (21) facilitates our evaluation of $d(x_{(r)}^n, y^n)$ for channel output y^n.

We illustrate the sets and quantities just introduced in the following example.

Example 2: Suppose $\mathcal{C}_6 = \{x_{(1)}^6, x_{(2)}^6, x_{(3)}^6\} = \{000000, 111100, 001111\}$. Then, from (16a) and (16b), we obtain $\mathcal{T}_{3|1} = \{001010, 001001, 000110, 000101, 000011, 010011, 100011\}$ and $\mathcal{N}_{3|1} = \{000111, 001011, 001101, 001110, 101011, 011011, 100111, 010111, 111101, 111110\}$. Next, it can be seen that $\mathcal{S}_2 = \{1, 2, 3, 4\}$, $\mathcal{S}_3 = \{3, 4, 5, 6\}$ and $\ell_2 = \ell_3 = 4$. In addition, by varying $m = 1 + \lambda_2$ for $\lambda_2 \in \{0, 1\}$, $\mathcal{S}_3$ can be partitioned into $2^{3-1} = 2$ sets, which are:

$$\mathcal{S}_3^{(m)} = \begin{cases} \mathcal{S}_{2;0} \bigcap \mathcal{S}_3 = \{5, 6\}, & m = 1; \\ \mathcal{S}_{2;1} \bigcap \mathcal{S}_3 = \{3, 4\}, & m = 2. \end{cases} \tag{22}$$

Hence,

$$\mathscr{S}_3^{(m)} = \begin{cases} \mathcal{S}_3^{(1)} = \{5, 6\}, & m = 1; \\ \mathcal{S}_3^{(1)} \bigcup \mathcal{S}_3^{(2)} = \{3, 4, 5, 6\}, & m = 2, \end{cases} \tag{23}$$

and $\ell_3^{(1)} = |\mathscr{S}_3^{(1)}| = 2$ and $\ell_3^{(2)} = |\mathscr{S}_3^{(2)}| = 4$. □

We are now ready to describe how we partition $\mathcal{T}_{j|1}$ and construct the corresponding disjoint subsets of $\mathcal{N}_{j|1}$. Recall from Section III-A that we can flip a zero component of u^n in $\mathcal{T}_{j|1}$ to recover a v^n in $\mathcal{N}_{j|1}$. This observation indicates that the number of zero components (equivalently, the number of one components) of $u^n \in \mathcal{T}_{j|1}$ with indices in $\mathscr{S}_j^{(m)}$ can be used as a factor to relate each partition of $\mathcal{T}_{j|1}$ to its corresponding subset of $\mathcal{N}_{j|1}$. As $x_{(1)}^n$ is assumed all-zero, this factor can be parameterized via $d(x_{(1)}^n, u^n|\mathscr{S}_j^{(m)}) = k$ for $0 \le k < \ell_j^{(m)}$.

Irrespective of the construction of disjoint subsets of $\mathcal{N}_{3|1}$, one may improperly infer from Example 2 that $\mathcal{T}_{3|1}$ can be subdivided into ℓ_3 partitions according to $d(x_{(1)}^6, u^6|\mathscr{S}_3^{(1)}) = k$ for each $0 \le k < \ell_3^{(1)}$, and then according to $d(x_{(1)}^6, u^6|\mathscr{S}_3^{(1)}) = \ell_3^{(1)}$ and $d(x_{(1)}^6, u^6|\mathscr{S}_3^{(2)}) = k$ for $\ell_3^{(1)} \le k < \ell_3^{(2)} = \ell_3$. However, the above setup could have two u^6 tuples, in respectively two different partitions of $\mathcal{T}_{3|1}$, recover the same v^6, leading to two *non-disjoint* subsets of $\mathcal{N}_{3|1}$. For example, flipping the last bit of 000110 that belongs to the partition constrained by $d(x_{(1)}^6, 000110|\mathscr{S}_3^{(1)}) = 1$, and flipping the 4th bit of 000011 that is included in the partition constrained by $d(x_{(1)}^6, 000011|\mathscr{S}_3^{(1)}) = \ell_3^{(1)}$ and $d(x_{(1)}^6, 000011|\mathscr{S}_3^{(2)}) = 2$ yield identical tuples given by $v^6 = 000111$; hence, the two partitions, indexed respectively by $k = 1$ and $k = 2$, recover two non-disjoint subsets of $\mathcal{N}_{3|1}$. To avoid repetitive constructions of the same v^6 from distinct partitions of $\mathcal{T}_{3|1}$, we note that multiple constructions of the same v^6 could happen only when the flipped zero component of u^6 is the only zero component in $\mathscr{S}_3^{(1)}$, i.e. $d(x_{(1)}^6, u^n|\mathscr{S}_3^{(1)}) = \ell_3^{(1)} - 1$. A solution is to place all u^6 tuples that result in multiple constructions of the same v^6 in one partition, based on which for $k \ge 2$, we refine the constraint of the kth partition as $\ell_3^{(1)} - 1 \le d(x_{(1)}^6, u^6|\mathscr{S}_3^{(1)}) \le$

$d(x_{(1)}^6, u^6|\mathscr{S}_3^{(2)}) = k$. In this manner, 000110 and 000011 are both included in the partition indexed by $k = 2$.

As a generalization, we constrain the kth partition of $\mathcal{T}_{j|1}$ by $\ell_j^{(m-1)} - 1 \le d(x_{(1)}^n, u^n|\mathscr{S}_j^{(m-1)}) \le d(x_{(1)}^n, u^n|\mathscr{S}_j^{(m)}) = k$ for $\ell_j^{(m-1)} - 1 \le k < \ell_j^{(m)} - 1$. After flipping a zero component of u^n in the kth partition of $\mathcal{T}_{j|1}$, the resulting v^n that belongs to the kth subset of $\mathcal{N}_{j|1}$ satisfies $\ell_j^{(m-1)} = d(x_{(1)}^n, v^n|\mathscr{S}_j^{(m-1)}) \le d(x_{(1)}^n, v^n|\mathscr{S}_j^{(m)}) = k+1$. To simplify our set constructions in the following definition, we define the mapping from the partition index k to the number m satisfying $\ell_j^{(m-1)} - 1 \le k < \ell_j^{(m)} - 1$, which designates the set $\mathscr{S}_j^{(m)}$ the flipped zero component of u^n is located in, as follows:

$$\sigma(k) \triangleq \begin{cases} m, & \text{for } \ell_j^{(m-1)} - 1 \le k < \ell_j^{(m)} - 1; \\ \min\left\{m : \ell_j^{(m)} = \ell_j\right\}, & \text{for } k = \ell_j - 1. \end{cases} \tag{24}$$

Definition 2: Define for $k = 0, 1, \ldots, \ell_j - 1$,

$$\begin{cases} \mathcal{T}_{j|1}(k) \triangleq \Big\{y^n \in \mathcal{T}_{j|1} : \\ \qquad \ell_j^{(m-1)} - 1 \le d\big(x_{(1)}^n, y^n \big| \mathscr{S}_j^{(m-1)}\big) \\ \qquad\qquad \le d\big(x_{(1)}^n, y^n \big| \mathscr{S}_j^{(m)}\big) = k\Big\}; & \text{(25a)} \\ \mathcal{N}_{j|1}(k) \triangleq \Big\{y^n \in \mathcal{N}_{j|1} : \\ \qquad \ell_j^{(m-1)} = d\big(x_{(1)}^n, y^n \big| \mathscr{S}_j^{(m-1)}\big) \\ \qquad\qquad \le d\big(x_{(1)}^n, y^n \big| \mathscr{S}_j^{(m)}\big) = k+1\Big\}, & \text{(25b)} \end{cases}$$

where $m = \sigma(k)$ is given in (24).

An example to illustrate the $\mathcal{T}_{j|1}$-partitions and $\mathcal{N}_{j|1}$-subsets is given below.

Example 3: Using the setting of Example 2, we show how we partition $\mathcal{T}_{3|1}$ according to $\mathscr{S}_3^{(1)}$ and $\mathscr{S}_3^{(2)}$ and construct the corresponding disjoint subsets of $\mathcal{N}_{3|1}$. From (25a) and (25b), we can obtain the partition $\{\mathcal{T}_{3|1}(k)\}_{0\le k<\ell_3}$ and disjoint subsets $\{\mathcal{N}_{3|1}(k)\}_{0\le k<\ell_3}$ as follows:

$$\mathcal{T}_{3|1}(k) = \begin{cases} \emptyset, & k = 0, 1, 3; \\ \mathcal{T}_{3|1}, & k = 2, \end{cases} \tag{26}$$

and

$$\mathcal{N}_{3|1}(k) = \begin{cases} \emptyset, & k = 0, 1, 3; \\ \mathcal{N}_{3|1}, & k = 2, \end{cases} \tag{27}$$

as a result of the mapping

$$\sigma(k) = \begin{cases} 1, & \ell_3^{(0)} - 1 \le k < \ell_3^{(1)} - 1 \text{ (equiv. } k = 0); \\ 2, & \ell_3^{(1)} - 1 \le k < \ell_3^{(2)} - 1 \text{ (equiv. } k = 1, 2); \\ 2, & k = \ell_3 - 1 = 3. \end{cases} \tag{28}$$

□

With the above definition, we next verify the partitions of non-empty $\mathcal{T}_{j|1}$ and the corresponding disjoint subsets of $\mathcal{N}_{j|1}$.

Proposition 2: For non-empty $\mathcal{T}_{j|1}$, the following two properties hold.

i) $\{\mathcal{T}_{j|1}(k)\}_{0\le k<\ell_j}$ *forms a partition of* $\mathcal{T}_{j|1}$*;*

ii) $\{\mathcal{N}_{j|1}(k)\}_{0\le k<\ell_j}$ *is a collection of disjoint subsets of* $\mathcal{N}_{j|1}$.

Proof: It can be seen from the definitions of $\{\mathcal{T}_{j|1}(k)\}_{0\le k<\ell_j}$ and $\{\mathcal{N}_{j|1}(k)\}_{0\le k<\ell_j}$ that they are collections of mutually disjoint subsets of $\mathcal{T}_{j|1}$ and $\mathcal{N}_{j|1}$, respectively. It remains to show that $\mathcal{T}_{j|1} = \bigcup_{0\le k<\ell_j} \mathcal{T}_{j|1}(k)$. Recall that $\mathscr{S}_j^{(m)}$ is a subset of $\mathcal{S}_j$ and every element y^n in $\mathcal{T}_{j|1}$ must satisfy $\ell_j > d(x_{(1)}^n, y^n|\mathcal{S}_j) = d(x_{(j)}^n, y^n|\mathcal{S}_j) = \frac{\ell_j}{2} \ge d(x_{(1)}^n, y^n|\mathscr{S}_j^{(m)})$; hence, no element in $\mathcal{T}_{j|1}$ can fulfill $d(x_{(1)}^n, y^n|\mathscr{S}_j^{(m)}) = \ell_j$. This confirms that in defining $\mathcal{T}_{j|1}(k)$ in (25a), we can exclude the case of $k = \ell_j$. Since every element in $\mathcal{T}_{j|1}$ must satisfy the two constraints in $\mathcal{T}_{j|1}(k)$ for exactly one $0 \le k < \ell_j$, $\{\mathcal{T}_{j|1}(k)\}_{0\le k<\ell_j}$ forms a partition of $\mathcal{T}_{j|1}$. ■

By applying a similar technique that leads to (14) and (18), Proposition 2 results in the following inequality:

$$\frac{P_{Y^n|X^n}(\mathcal{T}_{j|1}|x_{(1)}^n)}{P_{Y^n|X^n}(\mathcal{N}_{j|1}|x_{(1)}^n)} \le \max_{0\le k<\ell_j : \mathcal{T}_{j|1}(k)\ne\emptyset} \frac{P_{Y^n|X^n}(\mathcal{T}_{j|1}(k)|x_{(1)}^n)}{P_{Y^n|X^n}(\mathcal{N}_{j|1}(k)|x_{(1)}^n)} \tag{29}$$

for non-empty $\mathcal{T}_{j|1}$. We further decompose non-empty $\mathcal{T}_{j|1}(k)$ and its corresponding $\mathcal{N}_{j|1}(k)$ in the next section.

C. A Fine Partition of $\mathcal{T}_{j|1}(k)$ and the Corresponding Disjoint Subsets of $\mathcal{N}_{j|1}(k)$

The final decomposition of $\mathcal{T}_{j|1}(k)$ and $\mathcal{N}_{j|1}(k)$ is a little involved. We elucidate its underlying concept via an example before formally presenting it. The idea is to further partition $\mathcal{T}_{j|1}(k)$ using a group of representative elements in $\mathcal{T}_{j|1}(k)$ and construct the corresponding subsets of $\mathcal{N}_{j|1}(k)$ based on the same group of representative elements.

Pick an arbitrary element from $\mathcal{T}_{3|1}(2)$ in Example 3 as the first representative element, say $u^6 = 001010$. We collect all outputs y^6 in $\mathcal{T}_{3|1}(2)$ such that its components with indices outside $\mathscr{S}_3^{(\sigma(2))}$ are exact duplications of the components of u^6 at the same positions, and place them in $\mathcal{T}_{3|1}(u^6; 2)$. In other words, we require $d\big(u^6, y^6\big|\big(\mathscr{S}_3^{(2)}\big)^{\mathrm{c}}\big) = 0$. With $\big(\mathscr{S}_3^{(2)}\big)^{\mathrm{c}} = \{1, 2\}$, we have $\mathcal{T}_{3|1}(u^6; 2) = \mathcal{T}_{3|1}(001010; 2) = \{000011, 001010, 001001, 000110, 000101\}$, where the first two bits of each tuple in $\mathcal{T}_{3|1}(u^6; 2)$ must be equal to the first two bits of $u^6 = 001010$. Analogously, $\mathcal{N}_{3|1}(u^6; 2)$ collects all elements in $\mathcal{N}_{3|1}(2)$ satisfying $d\big(u^6, y^6\big|\big(\mathscr{S}_3^{(2)}\big)^{\mathrm{c}}\big) = 0$, and is given by $\mathcal{N}_{3|1}(001010; 2) = \{000111, 001011, 001101, 001110\}$.

We can further pick another element 100011 in $\mathcal{T}_{3|1} \setminus \mathcal{T}_{3|1}(001010; 2)$ as the second representative to construct $\mathcal{T}_{3|1}(100011; 2) = \{100011\}$ and the corresponding $\mathcal{N}_{3|1}(100011; 2) = \{101011, 100111\}$, where the first two bits of elements in the two sets must equal 10. Continuing this process to construct $\mathcal{T}_{3|1}(010011; 2) = \{010011\}$ and $\mathcal{N}_{3|1}(010011; 2) = \{011011, 010111\}$, we can see that all elements in $\mathcal{T}_{3|1}(2)$ have been exhausted. Thus, $\mathcal{U}_{3|1}(2) = \{001010, 100011, 010011\}$ is exactly the required group of representatives.

We formalize the above set constructions in the following definition and proposition, whose proof is omitted, being a direct consequence of the construction process.

Definition 3: Define for $u^n \in \mathcal{T}_{j|1}(k)$ with $m = \sigma(k)$,

$$\begin{cases} \mathcal{T}_{j|1}(u^n;k) \triangleq \big\{y^n \in \mathcal{T}_{j|1}(k) : \\ \qquad d\big(u^n, y^n \big| (\mathscr{S}_j^{(m)})^{\mathrm{c}}\big) = 0\big\}; & \text{(30a)} \\ \mathcal{N}_{j|1}(u^n;k) \triangleq \big\{y^n \in \mathcal{N}_{j|1}(k) : \\ \qquad d\big(u^n, y^n \big| (\mathscr{S}_j^{(m)})^{\mathrm{c}}\big) = 0\big\}. & \text{(30b)} \end{cases}$$

Proposition 3: For non-empty $\mathcal{T}_{j|1}(k)$, there exists a group of representative $\mathcal{U}_{j|1}(k) \subseteq \mathcal{T}_{j|1}(k)$ such that the following two properties hold.

i) *$\big\{\mathcal{T}_{j|1}(u^n;k)\big\}_{u^n \in \mathcal{U}_{j|1}(k)}$ forms a (non-empty) partition of $\mathcal{T}_{j|1}(k)$;*

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

Again, by applying a similar technique to derive (14), Proposition 3 yields that for non-empty $\mathcal{T}_{j|1}(k)$,

$$\begin{aligned} &\frac{P_{Y^n|X^n}(\mathcal{T}_{j|1}(k)|x^n_{(1)})}{P_{Y^n|X^n}(\mathcal{N}_{j|1}(k)|x^n_{(1)})} \\ &\le \max_{u^n \in \mathcal{U}_{j|1}(k)} \frac{P_{Y^n|X^n}(\mathcal{T}_{j|1}(u^n;k)|x^n_{(1)})}{P_{Y^n|X^n}(\mathcal{N}_{j|1}(u^n;k)|x^n_{(1)})}. \qquad (31) \end{aligned}$$

What remains to confirm is that $\frac{(1-p)}{p}n$ is an upper bound on $\frac{P_{Y^n|X^n}(\mathcal{T}_{j|1}(u^n;k)|x^n_{(1)})}{P_{Y^n|X^n}(\mathcal{N}_{j|1}(u^n;k)|x^n_{(1)})}$; this will be shown in the next section.

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

The constraints of $\mathcal{T}_{j|1}(u^n;k)$ in (30a) and $\mathcal{N}_{j|1}(u^n;k)$ in (30b) indicate that when dealing with $\frac{P_{Y^n|X^n}(\mathcal{T}_{j|1}(u^n;k)|x^n_{(1)})}{P_{Y^n|X^n}(\mathcal{N}_{j|1}(u^n;k)|x^n_{(1)})}$, we only need to consider those bits with indices in $\mathscr{S}_j^{(m)}$ with $m = \sigma(k)$ 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. Since elements in $\mathcal{T}_{j|1}(u^n;k)$ with indices in $\mathscr{S}_j^{(\sigma(k))}$ have exactly k ones, and those in $\mathcal{N}_{j|1}(u^n;k)$ with indices in $\mathscr{S}_j^{(\sigma(k))}$ have exactly $k+1$ ones, we can immediately infer that

$$\begin{aligned} &\frac{P_{Y^n|X^n}(\mathcal{T}_{j|1}(u^n;k)|x^n_{(1)})}{P_{Y^n|X^n}(\mathcal{N}_{j|1}(u^n;k)|x^n_{(1)})} \\ &= \frac{p^k(1-p)^{n-k}}{p^{k+1}(1-p)^{n-(k+1)}} \cdot \frac{|\mathcal{T}_{j|1}(u^n;k)|}{|\mathcal{N}_{j|1}(u^n;k)|} \qquad (32) \\ &= \frac{(1-p)}{p} \cdot \frac{|\mathcal{T}_{j|1}(u^n;k)|}{|\mathcal{N}_{j|1}(u^n;k)|}. \qquad (33) \end{aligned}$$

The desired upper bound can thus be established by proving that $\frac{|\mathcal{T}_{j|1}(u^n;k)|}{|\mathcal{N}_{j|1}(u^n;k)|} \le n$, as shown in the next proposition.

Proposition 4: For non-empty $\mathcal{T}_{j|1}(u^n;k)$, we have

$$\frac{P_{Y^n|X^n}(\mathcal{T}_{j|1}(u^n;k)|x^n_{(1)})}{P_{Y^n|X^n}(\mathcal{N}_{j|1}(u^n;k)|x^n_{(1)})} \le \frac{(1-p)}{p}n. \qquad (34)$$

Proof: Recall from (16a), (25a) and (30a) that $y^n \in \mathcal{T}_{j|1}(u^n;k)$ with $m = \sigma(k)$ if and only if

$$\begin{cases} d(x^n_{(1)}, y^n) = d(x^n_{(j)}, y^n); & \text{(35a)} \\ d(x^n_{(1)}, y^n) < \min_{r \in [j-1]\setminus\{1\}} d(x^n_{(r)}, y^n); & \text{(35b)} \\ \ell_j^{(m-1)} - 1 \le d\big(x^n_{(1)}, y^n \big| \mathscr{S}_j^{(m-1)}\big) \\ \qquad \le d\big(x^n_{(1)}, y^n \big| \mathscr{S}_j^{(m)}\big) = k; & \text{(35c)} \\ d\big(u^n, y^n \big| (\mathscr{S}_j^{(m)})^{\mathrm{c}}\big) = 0. & \text{(35d)} \end{cases}$$

Thus, we can enumerate the number of elements in $\mathcal{T}_{j|1}(u^n;k)$ by counting the number of channel outputs y^n fulfilling the above four conditions.

We then examine the number of y^n satisfying (35c) and (35d). Nothing that these y^n have either $\ell_j^{(m-1)} - 1$ ones or $\ell_j^{(m-1)}$ ones with indices in $\mathscr{S}_j^{(m-1)}$, we know there are

$$\begin{aligned} &\binom{\ell_j^{(m-1)}}{\ell_j^{(m-1)} - 1}\binom{\ell_j^{(m)} - \ell_j^{(m-1)}}{k - (\ell_j^{(m-1)} - 1)} \\ &\qquad + \binom{\ell_j^{(m-1)}}{\ell_j^{(m-1)}}\binom{\ell_j^{(m)} - \ell_j^{(m-1)}}{k - \ell_j^{(m-1)}} \qquad (36) \end{aligned}$$

of y^n tuples satisfying (35c) and (35d).[3] Considering the additional two conditions in (35a) and (35b), we get that the number of elements in $\mathcal{T}_{j|1}(u^n;k)$ is upper-bounded by (36).

On the other hand, from (16b), (25b), (30b) and $\mathcal{N}_{j|1}(u^n;k) \subseteq \mathcal{N}_{j|1}(k) \subseteq \mathcal{N}_{j|1}$, we obtain that $w^n \in \mathcal{N}_{j|1}(u^n;k)$ if and only if

$$\begin{cases} d(x^n_{(1)}, w^n) - 1 = d(x^n_{(j)}, w^n) + 1; & \text{(37a)} \\ d(x^n_{(1)}, w^n) - 1 \ne d(x^n_{(r)}, w^n) + 1 \\ \qquad \text{for } r \in [j-1]\setminus\{1\}; & \text{(37b)} \\ \ell_j^{(m-1)} = d\big(x^n_{(1)}, w^n \big| \mathscr{S}_j^{(m-1)}\big) \\ \qquad \le d\big(x^n_{(1)}, w^n \big| \mathscr{S}_j^{(m)}\big) = k+1; & \text{(37c)} \\ d\big(u^n, w^n \big| (\mathscr{S}_j^{(m)})^{\mathrm{c}}\big) = 0. & \text{(37d)} \end{cases}$$

We then claim that any w^n satisfying (37c) and (37d) should automatically validate (37a) and (37b). Note that the validity of the claim, which we prove in Appendix A, immediately implies that the number of elements in $\mathcal{N}_{j|1}(u^n;k)$ can be determined by (37c) and (37d), and hence

$$|\mathcal{N}_{j|1}(u^n;k)| = \binom{\ell_j^{(m)} - \ell_j^{(m-1)}}{k + 1 - \ell_j^{(m-1)}}. \qquad (38)$$

Under this claim, we complete the proof of the proposition using (33), (36) and (38) as follows:

$$\begin{aligned} &\frac{P_{Y^n|X^n}\big(\mathcal{T}_{j|1}(u^n;k)|x^n_{(1)}\big)}{P_{Y^n|X^n}\big(\mathcal{N}_{j|1}(u^n;k)|x^n_{(1)}\big)} \\ &\le \frac{(1-p)}{p} \end{aligned}$$

[3]To unify the expression, when $m = 1$, in which case $\ell_j^{(0)} = 0$, we assign $\binom{0}{-1} = 0$ and $\binom{0}{0} = 1$ in (36). Similarly, when $k = \ell_j^{(m-1)} - 1$, we set $\binom{\ell_j^{(m)} - \ell_j^{(m-1)}}{k - \ell_j^{(m-1)}} = \binom{\ell_j^{(m)} - \ell_j^{(m-1)}}{-1} = 0$.

$$\cdot \frac{\binom{\ell_j^{(m-1)}}{\ell_j^{(m-1)}-1}\binom{\ell_j^{(m)}-\ell_j^{(m-1)}}{k-(\ell_j^{(m-1)}-1)}+\binom{\ell_j^{(m-1)}}{\ell_j^{(m-1)}}\binom{\ell_j^{(m)}-\ell_j^{(m-1)}}{k-\ell_j^{(m-1)}}}{\binom{\ell_j^{(m)}-\ell_j^{(m-1)}}{k+1-\ell_j^{(m-1)}}} \quad (39)$$

$$= \frac{(1-p)}{p}\left(\ell_j^{(m-1)} + \frac{k+1-\ell_j^{(m-1)}}{\ell_j^{(m)}-k}\right) \quad (40)$$

$$\leq \frac{(1-p)}{p}\left(\ell_j^{(m-1)} + \frac{\ell_j^{(m)}-\ell_j^{(m-1)}}{1}\right) \quad (41)$$

$$\leq \frac{(1-p)}{p} n, \quad (42)$$

where (41) holds because $\ell_j^{(m-1)}-1 \leq k \leq \ell_j^{(m)}-1$ by (24), and (42) follows from $\ell_j^{(m)} \leq \ell_j \leq n$. ■

Using (18), (29), (31) and Proposition 4, we obtain

$$\frac{P_{Y^n|X^n}(\mathcal{T}_1|x^n_{(1)})}{P_{Y^n|X^n}(\mathcal{N}_1|x^n_{(1)})} \leq \frac{(1-p)}{p} n. \quad (43)$$

We close this section by remarking that the same inequality as (43), i.e.,

$$\frac{P_{Y^n|X^n}(\mathcal{T}_i|x^n_{(i)})}{P_{Y^n|X^n}(\mathcal{N}_i|x^n_{(i)})} \leq \frac{(1-p)}{p} n, \quad (44)$$

can be analogously established for all $i \in [M]$ with $\mathcal{T}_i \neq \emptyset$. Consequently, (14) implies

$$\frac{\delta_n}{b_n} \leq \max_{i\in[M]:\mathcal{T}_i\neq\emptyset} \frac{P_{Y^n|X^n}(\mathcal{T}_i|x^n_{(i)})}{P_{Y^n|X^n}(\mathcal{N}_i|x^n_{(i)})} \leq \frac{(1-p)}{p} n. \quad (45)$$

IV. Conclusion

In this paper, the generalized Poor-Verdú error lower bound of [1] was considered in the classical channel coding context over the BSC. We proved that the bound is exponentially tight in blocklength as a direct consequence of a key inequality, showing that for any block code with distinct codewords used over the BSC, the relative deviation of the code's minimum probability of error from the lower bound grows at most linearly in blocklength.

Even though the exact determination of the reliability function of the BSC at low rates remains a daunting open problem, our results offer potentially a new perspective or tool for subsequent studies. Other future work includes investigating sharp bounds for codes with small-to-moderate blocklengths (e.g., see [5], [27], [28]) used over symmetric channels. As our counting analysis for the binary symmetric channel relies heavily on the equivalence between ML decoding and minimum Hamming distance decoding, which does not hold for non-symmetric channels, extending our results to general channels may require more sophisticated enumerating techniques.

Acknowledgment

The authors sincerely thank the Associate Editor and the reviewers for their insightful and constructive comments which were quite instrumental in improving the manuscript.

Appendix A
The Proof of (37c) and (37d) implying (37a) and (37b)

We validate the claim 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 $d\big(x^n_{(1)}, u^n \big| \mathscr{S}_j^{(m-1)}\big)$ equals $\ell_j^{(m-1)}$ or $\ell_j^{(m-1)}-1$ as follows.

i) $d(x^n_{(1)}, u^n|\mathscr{S}_j^{(m-1)}) = \ell_j^{(m-1)}$: Since in this case, u^n has no zero components with indices in $\mathscr{S}_j^{(m-1)}$, we flip a zero component of u^n with its index in $\mathscr{S}_j^{(m)} \setminus \mathscr{S}_j^{(m-1)} = \mathcal{S}_j^{(m)}$ to construct a v^n such that

$$d(x^n_{(1)}, v^n) = d(x^n_{(1)}, u^n) + 1 \quad (46)$$

and

$$d(x^n_{(j)}, v^n) = d(x^n_{(j)}, u^n) - 1, \quad (47)$$

where the existence of such v^n is guaranteed by $k \leq \ell_j^{(m)}-1$. Then, v^n must fulfill (37a), (37c) and (37d) (with w^n replaced by v^n) as u^n satisfies (35a), (35c) and (35d). We next prove v^n also fulfills (37b) by contradiction. Suppose there exists a $r \in [j-1] \setminus \{1\}$ satisfying

$$d(x^n_{(1)}, v^n) - 1 = d(x^n_{(r)}, v^n) + 1. \quad (48)$$

We then recall from (21) that $d(x^n_{(1)}, x^n_{(r)}|\mathcal{S}_j^{(m)})$ is either 0 or $|\mathcal{S}_j^{(m)}|$. Thus, (48) can be disproved by differentiating two cases: 1) $d(x^n_{(1)}, x^n_{(r)}|\mathcal{S}_j^{(m)}) = 0$, and 2) $d(x^n_{(1)}, x^n_{(r)}|\mathcal{S}_j^{(m)}) = |\mathcal{S}_j^{(m)}|$.

In case 1), v^n that is obtained by flipping a zero component of u^n with index in $\mathcal{S}_j^{(m)}$ must satisfy $d(x^n_{(1)}, v^n) = d(x^n_{(1)}, u^n) + 1$ and $d(x^n_{(r)}, v^n) = d(x^n_{(r)}, u^n) + 1$. Then, (48) implies $d(x^n_{(1)}, u^n) - 1 = d(x^n_{(r)}, u^n) + 1$. A contradiction to the fact that u^n satisfies (35b) is obtained. In case 2), the flipping manipulation on u^n results in $d(x^n_{(1)}, v^n) = d(x^n_{(1)}, u^n) + 1$ and $d(x^n_{(r)}, v^n) = d(x^n_{(r)}, u^n) - 1$. Therefore, (48) implies $d(x^n_{(1)}, u^n) = d(x^n_{(r)}, u^n)$, which again contradicts (35b). Accordingly, v^n must also fulfill (37b); hence, $v^n \in \mathcal{N}_{j|1}(u^n;k)$.

With this auxiliary v^n, we are ready to prove that every w^n satisfying (37c) and (37d) also validates (37a) and (37b). This can be done by showing $d(x^n_{(r)}, w^n) = d(x^n_{(r)}, v^n)$ for every $r \in [j]$, which can be verified as follows:

$$\begin{aligned} &d(x^n_{(r)}, w^n) \\ &= d\big(x^n_{(r)}, w^n \big| \mathscr{S}_j^{(m-1)}\big) + d\big(x^n_{(r)}, w^n \big| \mathcal{S}_j^{(m)}\big) \\ &\quad + d\big(x^n_{(r)}, w^n \big| (\mathscr{S}_j^{(m)})^{\mathrm{c}}\big) && (49) \\ &= d\big(x^n_{(r)}, v^n \big| \mathscr{S}_j^{(m-1)}\big) + d\big(x^n_{(r)}, v^n \big| \mathcal{S}_j^{(m)}\big) \\ &\quad + d\big(x^n_{(r)}, v^n \big| (\mathscr{S}_j^{(m)})^{\mathrm{c}}\big) && (50) \\ &= d(x^n_{(r)}, v^n), && (51) \end{aligned}$$

where the substitution in the first term of (50) holds because both v^n and w^n satisfy (37c), implying all components of v^n and w^n with indices in $\mathscr{S}_j^{(m-1)}$ are equal to one; the substitution in the 2nd term of (50) holds because when considering only those portions with indices in $\mathcal{S}_j^{(m)}$, $x^n_{(r)}$ are either all ones or all zeros according to (21), and both w^n and v^n have exactly $k+1-\ell_j^{(m-1)}$ ones according to (37c); and the substitution in the 3rd term of (50) is valid since both v^n and w^n satisfy (37d).

ii) $d(x^n_{(1)}, u^n|\mathscr{S}_j^{(m-1)}) = \ell_j^{(m-1)} - 1$: Now we let v^n be equal to u^n in all positions but one in $\mathscr{S}_j^{(m-1)}$ such that $d(x^n_{(1)}, v^n|\mathscr{S}_j^{(m-1)}) = \ell_j^{(m-1)}$. Then, v^n must fulfill (37a), (37c) and (37d) as u^n satisfies (35a), (35c) and (35d). With the components of $x^n_{(r)}$ with respect to $\mathcal{S}_j^{(m)}$ being either all zeros or all ones, the same contradiction argument after (48) can disprove the validity of (48) for this v^n and for any $r \in [j-1] \setminus \{1\}$. Therefore, v^n also fulfills (37b), implying $v^n \in \mathcal{N}_{j|1}(u^n; k)$. With this auxiliary v^n, we can again verify (51) via the same derivation in (51). The claim that w^n satisfying (37c) and (37d) validates (37a) and (37b) is thus confirmed.

References

[1] L.-H. Chang, P.-N. Chen, F. Alajaji, and Y. S. Han, "The asymptotic generalized Poor-Verdú bound achieves the BSC error exponent at zero rate," in *IEEE Int. Symp. Inf. Theory (ISIT)*, Los Angeles, CA, USA, June 21–26, 2020.

[2] H. V. Poor and S. Verdú, "A lower bound on the probability of error in multihypothesis testing," *IEEE Trans. Inf. Theory*, vol. 41, no. 6, pp. 1992–1994, November 1995.

[3] P.-N. Chen and F. Alajaji, "A generalized Poor-Verdú error bound for multihypothesis testings," *IEEE Trans. Inf. Theory*, vol. 58, no. 1, pp. 311–316, January 2012.

[4] S. Verdú and T. S. Han, "A general formula for channel capacity," *IEEE Trans. Inf. Theory*, vol. 40, no. 4, pp. 1147–1157, July 1994.

[5] G. Vazquez-Vilar, A. T. Campo, A. G. i Fábregas, and A. Martinez, "Bayesian M-ary hypothesis testing: The meta-converse and Verdú-Han bounds are tight," *IEEE Trans. Inf. Theory*, vol. 62, no. 5, pp. 2324–2333, May 2016.

[6] R. G. Gallager, "A simple derivation of the coding theorem and some applications," *IEEE Trans. Inf. Theory*, vol. 11, no. 1, pp. 3–18, January 1965.

[7] C. E. Shannon, R. G. Gallager, and E. R. Berlekamp, "Lower bounds to error probability for coding on discrete memoryless channels - i," *Inf. Contr.*, vol. 10, pp. 65–103, January 1967.

[8] ——, "Lower bounds to error probability for coding on discrete memoryless channels - ii," *Inf. Contr.*, vol. 10, pp. 522–552, May 1967.

[9] R. J. McEliece and J. K. Omura, "An improved upper bound on the block coding error exponent for binary-input discrete memoryless channels," *IEEE Trans. Inf. Theory*, vol. 23, no. 5, pp. 611–613, September 1977.

[10] R. G. Gallager, *Information Theory and Reliable Communication*. NY: Wiley, 1968.

[11] A. J. Viterbi and J. K. Omura, *Principles of Digital Communication and Coding*. NY: McGraw-Hill, 1979.

[12] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. NY: Academic Press, 1981.

[13] R. Blahut, *Principles and Practice of Information Theory*. A. Wesley, MA, 1988.

[14] S. Litsyn, "New upper bounds on error exponents," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 385–398, March 1999.

[15] A. Barg and A. McGregor, "Distance distribution of binary codes and the error probability of decoding," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4237–4246, December 2005.

[16] E. A. Haroutunian, M. E. Haroutunian, and A. N. Harutyunyan, "Reliability criteria in information theory and in statistical hypothesis testing," *Foundations and Trends in Communications and Information Theory, Now Publishers Inc.*, vol. 4, pp. 97–263, January 2007.

[17] Y. Polyanskiy, "Saddle point in the minimax converse for channel coding," *IEEE Trans. Inf. Theory*, vol. 59, no. 5, pp. 2576–2595, May 2013.

[18] M. Dalai, "Lower bounds on the probability of error for classical and classical-quantum channels," *IEEE Trans. Inf. Theory*, vol. 59, no. 12, pp. 8027–8056, December 2013.

[19] J. Scarlett, A. Martinez, and A. G. i Fábregas, "Mismatched decoding: Error exponents, second-order rates and saddlepoint approximations," *IEEE Trans. Inf. Theory*, vol. 60, no. 5, pp. 2647–2666, May 2014.

[20] J. Scarlett, I. Peng, N. Merhav, A. Martinez, and A. G. i Fábregas, "Expurgated random-coding ensembles: Exponents, refinements, and connections," *IEEE Trans. Inf. Theory*, vol. 60, no. 8, pp. 4449–4462, August 2014.

[21] Y. Altŭg and A. B. Wagner, "Refinement of the random coding bound," *IEEE Trans. Inf. Theory*, vol. 60, no. 10, pp. 6005–6023, October 2014.

[22] M. V. Burnashev, "On the BSC reliability function: Expanding the region where it is known exactly," *Probl. Inf. Trans.*, vol. 51, no. 4, pp. 307–325, January 2015.

[23] A. Tandon, V. Y. F. Tan, and L. R. Varshney, "The bee-identification problem: Bounds on the error exponent," *IEEE Trans. Commun.*, vol. 67, no. 11, pp. 7405–7416, November 2019.

[24] ——, "On the bee-identification error exponent with absentee bees," *arXiv preprint arXiv:1910.10333*, October 2019.

[25] N. Merhav, "A Lagrange-dual lower bound to the error exponent of the typical random code," *IEEE Trans. Inf. Theory*, vol. 66, no. 6, pp. 3456–3464, June 2020.

[26] F. Alajaji, P.-N. Chen, and Z. Rached, "A note on the Poor-Verdú conjecture for the channel reliability function," *IEEE Trans. Inf. Theory*, vol. 48, no. 1, pp. 309–313, January 2002.

[27] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2307–2359, May 2010.

[28] P.-N. Chen, H.-Y. Lin, and S. M. Moser, "Optimal ultrasmall block-codes for binary discrete memoryless channels," *IEEE Trans. Inf. Theory*, vol. 59, no. 11, pp. 7346–7378, November 2013.

Ling-Hua Chang is currently an assistant professor in the Department of Electrical Engineering (group A) at the Yuan Ze University, Taiwan. She received the B.S. and Ph.D. degrees both in Electrical Engineering from National Chiao Tung University, Taiwan, in 2010 and 2016, respectively. Her research interests include signal processing, linear algebra, and information theory.

Po-Ning Chen (S'93–M'95–SM'01) received the B.S. and M.S. degrees in electrical engineering from National Tsing-Hua University, Taiwan, in 1985 and 1987, respectively, and the Ph.D. degree in electrical engineering from University of Maryland, College Park, in 1994.

From 1985 to 1987, he was with Image Processing Laboratory in National Tsing-Hua University, where he worked on the recognition of Chinese characters. During 1989, he was with Star Tech. Inc., where he focused on the development of finger-print recognition systems. After the reception of Ph.D. degree in 1994, he jointed Wan Ta Technology Inc. as a vice general manager, conducting several projects on Point-of-Sale systems. In 1995, he became a research staff in Advanced Technology Center, Computer and Communication Laboratory, Industrial Technology Research Institute in Taiwan, where he led a project on Java-based Network Managements. Since 1996, he has been an Associate Professor in Department of Communications Engineering at National Chiao Tung University (NCTU), Taiwan, and was promoted to a full professor in 2001. He was elected to be the Chair of IEEE Communications Society Taipei Chapter in 2006 and 2007, during which IEEE ComSoc Taipei Chapter won the 2007 IEEE ComSoc Chapter Achievement Awards (CAA) and 2007 IEEE ComSoc Chapter of the Year (CoY). He has served as the chairman of Department of Communications Engineering, NCTU, during 2007–2009. During 2012–2015, he was the associate chief director of Microelectronics and Information Systems Research Center, NCTU. In 2017, he became the associate dean of College of Electrical and Computer Engineering, NCTU.

Dr. Chen received the annual Research Awards from National Science Council, Taiwan, from 1996–2000, and received the 2000 Young Scholar Paper Award from Academia Sinica, Taiwan. His Experimental Handouts for the course of Communication Networks Laboratory have been awarded as the Annual Best Teaching Materials for Communications Education by Ministry of Education, Taiwan, in 1998. He has been selected as the Outstanding Tutor Teacher of NCTU in 2002, 2013, and 2014. He was also the recipient of Distinguished Teaching Award from College of Electrical and Computer Engineering, NCTU, Taiwan, in 2003 and 2014, and the Outstanding Teaching Award of NCTu, in 2020. His research interests generally lie in information and coding theory, large deviation theory, distributed detection and sensor networks.

Fady Alajaji (S90M94SM00) received the B.E. degree with distinction from the American University of Beirut, Lebanon, and the M.Sc. and Ph.D. degrees from the University of Maryland, College Park, all in electrical engineering, in 1988, 1990 and 1994, respectively. He held a postdoctoral appointment in 1994 at the Institute for Systems Research, University of Maryland.

In 1995, he joined the Department of Mathematics and Statistics at Queens University, Kingston, Ontario, where he is currently a Professor of Mathematics and Engineering. Since 1997, he has also been cross-appointed in the Department of Electrical and Computer Engineering at the same university. In 2013-2014, he served as acting head of the Department of Mathematics and Statistics. From 2003 to 2008 and in 2018-2019, he served as chair of the Queens Mathematics and Engineering program. His research interests include information theory, digital communications, error control coding, data compression, joint source-channel coding, network epidemics, generative-adversarial networks, and data privacy and fairness in machine learning.

Dr. Alajaji is an Associate Editor for Shannon Theory for the IEEE Transactions on Information Theory. He served as Area Editor for Source-Channel Coding and Signal Processing (2008-2015) and as Editor for Source and Source-Channel Coding (2003-2012) for the IEEE Transactions on Communications. He served as organizer and Technical Program Committee member of several international conferences and workshops. He received the Premiers Research Excellence Award from the Province of Ontario.

Yunghsiang S. Han (S'90-M'93-SM'08-F'11) was born in Taipei, Taiwan, 1962. He received B.Sc. and M.Sc. degrees in electrical engineering from the National Tsing Hua University, Hsinchu, Taiwan, in 1984 and 1986, respectively, and a Ph.D. degree from the School of Computer and Information Science, Syracuse University, Syracuse, NY, in 1993. He was from 1986 to 1988 a lecturer at Ming-Hsin Engineering College, Hsinchu, Taiwan. He was a teaching assistant from 1989 to 1992, and a research associate in the School of Computer and Information Science, Syracuse University from 1992 to 1993. He was, from 1993 to 1997, an Associate Professor in the Department of Electronic Engineering at Hua Fan College of Humanities and Technology, Taipei Hsien, Taiwan. He was with the Department of Computer Science and Information Engineering at National Chi Nan University, Nantou, Taiwan from 1997 to 2004. He was promoted to Professor in 1998. He was a visiting scholar in the Department of Electrical Engineering at University of Hawaii at Manoa, HI from June to October 2001, the SUPRIA visiting research scholar in the Department of Electrical Engineering and Computer Science and CASE center at Syracuse University, NY from September 2002 to January 2004 and July 2012 to June 2013, and the visiting scholar in the Department of Electrical and Computer Engineering at University of Texas at Austin, TX from August 2008 to June 2009. He was with the Graduate Institute of Communication Engineering at National Taipei University, Taipei, Taiwan from August 2004 to July 2010. From August 2010 to January 2017, he was with the Department of Electrical Engineering at National Taiwan University of Science and Technology as Chair Professor. From February 2017 to February 2021, he was with School of Electrical Engineering & Intelligentization at Dongguan University of Technology, China. Now he is with the Shenzhen Institute for Advanced Study, University of Electronic Science and Technology of China. He is also a Chair Professor at National Taipei University from February 2015. His research interests are in error-control coding, wireless networks, and security.

Dr. Han was a winner of the 1994 Syracuse University Doctoral Prize and a Fellow of IEEE. One of his papers won the prestigious 2013 ACM CCS Test-of-Time Award in cybersecurity.