On the Maximum Size of Block Codes Subject to a Distance Criterion

Ling-Hua Chang*, Po-Ning Chen†, *Senior Member, IEEE*, Vincent Y. F. Tan‡§, *Senior Member, IEEE*, Carol Wang¶, Yunghsiang S. Han‖, *Fellow, IEEE*

***Abstract*—We establish a general formula for the maximum size of finite length block codes with minimum pairwise distance no less than d. The achievability argument involves an iterative construction of a set of radius-d balls, each centered at a codeword. We demonstrate that the number of such balls that cover the entire code space cannot exceed this maximum size. Our approach can be applied to codes $i)$ with elements over arbitrary code alphabets, and $ii)$ under a broad class of distance measures. Our formula indicates that the maximum code size can be fully characterized by the cumulative distribution function of the distance measure evaluated at two independent and identically distributed random codewords. When the two random codewords assume a uniform distribution over the entire code alphabet, our formula recovers and thus naturally generalizes the Gilbert-Varshamov (GV) lower bound. Finally, we extend our study to the asymptotic setting.**

I. Introduction

Given an arbitrary (possibly uncountable) code alphabet $\mathcal{X}$ and a general distance measure (possibly asymmetric or not satisfying the triangle inequality), the determination of the maximal size $M_n^*(d)$ of a block code $\mathcal{C} \subseteq \mathcal{X}^n$ with pairwise minimum distance no less than d and block length $n < \infty$ has been a long-standing problem in information and coding theory. In its applications, one can use $M_n^*(d)$ to obtain an upper bound of the expurgated error exponent [1] and also to characterize the capacity of a graph [2]. Some well-known bounds on $M_n^*(d)$ include the linear programming upper bound [3] and Gilbert-Varshamov (GV) lower bound [3]–[5]. Other famous upper bounds include the Singleton, Plotkin, and Elias bounds [6]. However, these bounds are not tight in general. Since finite-length bounds are usually difficult to obtain, researchers have focused on asymptotic analyses in which blocklength n tends to infinity. One then considers the limit of the code rate $(1/n)\log M_n^*(d)$ subject to a normalized distance constraint $d/n \geq \delta$. Many asymptotic bounds have been derived; see, for example [3], [4], [7]–[13] and the references therein.

*Department of Electrical Engineering, Yuan Ze University, Taiwan, R.O.C.
†Institute of Communications Engineering & Department of Electrical and Computer Engineering, National Chiao Tung University, Taiwan, R.O.C.
‡Department of Electrical and Computer Engineering, National University of Singapore, Singapore
§Department of Mathematics, National University of Singapore, Singapore
¶Work done while in the Department of of Electrical and Computer Engineering, National University of Singapore, Singapore
‖School of Electrical Engineering & Intelligentization, Dongguan University of Technology, China

The work of Ling-Hua Chang is supported by the Ministry of Science and Technology, Taiwan, under grant MOST 107-2218-E-155-012-MY2 and MOST 105-2221-E-009-011-MY3. The work of Po-Ning Chen is supported by the Ministry of Science and Technology, Taiwan, under grant MOST 105-2221-E-009-009-MY3. The work of Carol Wang and Vincent Tan is supported by a Singapore Ministry of Education (MoE) Tier 2 grant (R-263-000-B61-112). The work of Yunghsiang S. Han is support by the National Natural Science Foundation of China (Grant No. 61671007).

A natural question then beckons. Can one derive a "meta-result" concerning the maximum code size subject to a fixed minimum distance $M_n^*(d)$ that recovers some of the above-mentioned bounds as special cases? In [14], using a graph-theoretic framework, Motzkin and Straus derived such a result which implies an exact formula for $M_n^*(d)$ under the condition that $\mathcal{X}$ is finite. See also a related result by Korn [15]. It is then natural to ask if there exists an analogous result for more general code alphabets, e.g., uncountable alphabets. This is precisely the purpose of this paper.

Along this direction, we propose an iterative construction of a set of balls, each centered at a codeword and of a fixed radius d. We then show that the number of such balls that cover the entire code space $\mathcal{X}^n$ cannot exceed the maximum code size. Consequently, we prove that $M_n^*(d)$ for an *arbitrary* code alphabet can be completely determined by the minimum probability (over all distributions) that two i.i.d. random vectors $\hat{X}^n$ and X^n are at distance less than d from each other, i.e.,

$$M_n^*(d) = \frac{1}{\inf_{P_{X^n}} \Pr\left[\min\{\mu(\hat{X}^n, X^n), \mu(X^n, \hat{X}^n)\} < d\right]},$$

where $\mu(\cdot,\cdot)$ is the (possibly asymmetric) distance measure. This formula not only can be used to recover and to naturally generalize the GV bound, but also facilitates the evaluation of the limiting behavior of $(1/n)\log M_n^*(n\delta)$ under the condition that the relative minimum distance is at least δ.

The rest of the paper is organized as follows. The exact formula for $M_n^*(d)$ is presented in Section II. A family of lower bounds to $M_n^*(d)$ is presented in Section III; also included here is the demonstration that the finite length GV lower bound can be recovered from our formula. Extensions to the asymptotic regime are studied in Section IV. Finally, open problems are discussed in Section V.

II. Maximal Code Size Attainable under a Minimum Pairwise Distance

We first introduce the notation used in this paper. An (n, M)-code over $\mathcal{X}^n$ denotes a set of M vectors, each of which belongs to $\mathcal{X}^n$ [16]. A distance measure $\mu(\cdot,\cdot)$ is a real-valued function with domain $\mathcal{X}^n \times \mathcal{X}^n$ which satisfies

$$\mu(u^n, v^n) = \mu_{\min} \triangleq \min_{\hat{x}^n, x^n \in \mathcal{X}^n} \mu(\hat{x}^n, x^n) \text{ if } u^n = v^n. \quad (1)$$

Here, we do not require $\mu(\cdot,\cdot)$ to be symmetric or satisfy the triangle inequality but can be arbitrary as long as it admits its minimum from a point to itself.

An (n,M,d)-code $\mathcal{C}$ denotes an (n,M)-code with the minimum pairwise distance among codewords at least d, i.e.,

$$\min_{\hat{x}^n,\, x^n\in\mathcal{C} \text{ and } \hat{x}^n\neq x^n} \mu(\hat{x}^n,x^n)\geq d. \tag{2}$$

The maximal code size $M_n^*(d)$ subject to a pairwise minimum distance d is given by

$$M_n^*(d)\triangleq\max\left\{M\in\mathbb{N}:\exists\,(n,M,d)\text{-code}\right\},$$

where $\mathbb{N}$ is the set of positive integers. For convenience, a code that satisfies (2) is referred to as a *distance-d code*. Throughout this paper, $\hat{X}^n$ and X^n denote two independent random variables with a common distribution P_{X^n} over $\mathcal{X}^n$.

We now present a general formula for the maximum size $M_n^*(d)$ of distance-d codes over an arbitrary code alphabet $\mathcal{X}$ (not necessarily countable) and general distance measure $\mu(\cdot,\cdot)$.

Theorem 1: Fix an arbitrary code alphabet $\mathcal{X}$ and a distance measure $\mu(\cdot,\cdot)$ that satisfies (1). For all $n\geq 1$ and $d>\mu_{\min}$, we have

$$M_n^*(d)=\frac{1}{\inf\limits_{P_{X^n}}\Pr\big[\min\{\mu(\hat{X}^n,X^n),\mu(X^n,\hat{X}^n)\}<d\big]}. \tag{3}$$

Proof: We first prove the validity of (3) under the assumption that $M_n^*(d)<\infty$. Its extension to $M_n^*(d)=\infty$ will be done next.

Subject to the condition that $M_n^*(d)$ is finite, the equality in (3) can be proved in two steps. We first show that for every distribution P_{X^n} over $\mathcal{X}^n$, the following inequality holds:

$$\Pr\left[\tilde{\mu}(\hat{X}^n,X^n)<d\right]\geq\frac{1}{M_n^*(d)}, \tag{4}$$

where for convenience, we denote $\tilde{\mu}(\hat{x}^n,x^n)\triangleq\min\{\mu(\hat{x}^n,x^n),\mu(x^n,\hat{x}^n)\}$; hence,

$$\inf_{P_{X^n}}\Pr\left[\tilde{\mu}(\hat{X}^n,X^n)<d\right]\geq\frac{1}{M_n^*(d)}. \tag{5}$$

The proof is then completed by exhibiting a distribution $P_{X^{n*}}$ that results in equality in (5); consequently, given that $M_n^*(d)$ is finite, the infimum in (3) can be replaced by a minimum.

1) Achievability (*Validation of* (4) *under finite* $M_n^*(d)$): Fix a distribution P_{X^n} over $\mathcal{X}^n$ and an arbitrarily small $\epsilon>0$. Let

$$a_1\triangleq\inf_{x^n\in\mathcal{X}^n}\Pr[X^n\in\mathcal{B}(x^n)],$$

where $\mathcal{B}(x^n)\triangleq\{\hat{x}^n\in\mathcal{X}^n:\tilde{\mu}(\hat{x}^n,x^n)<d\}$. Find an element u_1^n in $\mathcal{X}^n$ such that $p_1\triangleq\Pr[X^n\in\mathcal{B}(u_1^n)]<a_1+\epsilon$. Note that the existence of u_1^n is guaranteed by the definition of the infimum. Let

$$a_2\triangleq\inf_{x^n\in\mathcal{X}^n\setminus\mathcal{B}(u_1^n)}\Pr[X^n\in\mathcal{B}(x^n)\setminus\mathcal{B}(u_1^n)].$$

Find an element u_2^n in $\mathcal{X}^n\setminus\mathcal{B}(u_1^n)$ such that $p_2\triangleq\Pr[X^n\in\mathcal{B}(u_2^n)\setminus\mathcal{B}(u_1^n)]<a_2+\epsilon$. We repeat this procedure to obtain

$$a_i\triangleq\inf_{x^n\in\mathcal{X}^n\setminus\cup_{j=1}^{i-1}\mathcal{B}(u_j^n)}\Pr[X^n\in\mathcal{B}(x^n)\setminus\cup_{j=1}^{i-1}\mathcal{B}(u_j^n)]$$

and an u_i^n in $\mathcal{X}^n\setminus\cup_{j=1}^{i-1}\mathcal{B}(u_j^n)$ with $p_i\triangleq\Pr[X^n\in\mathcal{B}(u_i^n)\setminus\cup_{j=1}^{i-1}\mathcal{B}(u_j^n)]<a_i+\epsilon$ for $i=3,4,\ldots,k$ until $\cup_{j=1}^{k}\mathcal{B}(u_j^n)$ covers the entire $\mathcal{X}^n$, i.e., $\mathcal{X}^n\setminus\cup_{j=1}^{k}\mathcal{B}(u_j^n)=\emptyset$ but $\mathcal{X}^n\setminus\cup_{j=1}^{k-1}\mathcal{B}(u_j^n)\neq\emptyset$. Two observations are made: $i)$ $\{u_1^n,u_2^n,\ldots,u_k^n\}$ is a distance-d code and hence by the definition of $M_n^*(d)$ and its assumed finiteness, $k\leq M_n^*(d)$ is a finite integer so the above procedure is repeated at most $M_n^*(d)$ times; $ii)$ $\sum_{j=1}^{k}p_j=1$. Denoting $\mathcal{D}_i\triangleq\mathcal{B}(u_i^n)\setminus\cup_{j=1}^{i-1}\mathcal{B}(u_j^n)$ and noting $\mathcal{D}_i\cap\mathcal{D}_j=\emptyset$ for $i\neq j$ and $\mathcal{X}^n=\cup_{j=1}^{k}\mathcal{D}_j$ (i.e., $\{\mathcal{D}_j\}_{j=1}^{k}$ is a partition of $\mathcal{X}^n$), we can derive the following chain of inequalities:

$$\begin{aligned}
&\Pr[\tilde{\mu}(\hat{X}^n,X^n)<d]\\
&=\int_{\mathcal{X}^n}\int_{\mathcal{X}^n}\mathbf{1}\{\tilde{\mu}(\hat{x}^n,x^n)<d\}\,\mathrm{d}P_{X^n}(\hat{x}^n)\,\mathrm{d}P_{X^n}(x^n)\\
&=\sum_{j=1}^{k}\int_{\mathcal{D}_j}\int_{\mathcal{X}^n}\mathbf{1}\{\tilde{\mu}(\hat{x}^n,x^n)<d\}\,\mathrm{d}P_{X^n}(\hat{x}^n)\,\mathrm{d}P_{X^n}(x^n)\\
&=\sum_{j=1}^{k}\int_{\mathcal{D}_j}\int_{\mathcal{B}(x^n)}\mathrm{d}P_{X^n}(\hat{x}^n)\,\mathrm{d}P_{X^n}(x^n)\\
&\geq\sum_{j=1}^{k}\int_{\mathcal{D}_j}a_j\,\mathrm{d}P_{X^n}(x^n) &(6)\\
&=\sum_{j=1}^{k}a_jp_j &(7)\\
&>\sum_{j=1}^{k}(p_j-\epsilon)p_j &(8)\\
&=\left(\sum_{j=1}^{k}p_j^2\right)-\epsilon &(9)\\
&\geq\frac{1}{k}-\epsilon &(10)\\
&\geq\frac{1}{M_n^*(d)}-\epsilon, &(11)
\end{aligned}$$

where $\mathbf{1}\{\cdot\}$ is the set indicator function; (6) holds because

$$\begin{aligned}
&\inf_{x^n\in\mathcal{D}_j}\int_{\mathcal{B}(x^n)}\mathrm{d}P_{X^n}(\hat{x}^n)\\
&=\inf_{x^n\in\mathcal{B}(u_j^n)\setminus\cup_{\ell=1}^{j-1}\mathcal{B}(u_\ell^n)}\Pr[X^n\in\mathcal{B}(x^n)]\\
&\geq\inf_{x^n\in\mathcal{B}(u_j^n)\setminus\cup_{\ell=1}^{j-1}\mathcal{B}(u_\ell^n)}\Pr[X^n\in\mathcal{B}(x^n)\setminus\cup_{\ell=1}^{j-1}\mathcal{B}(u_\ell^n)]\\
&\geq\inf_{x^n\in\mathcal{X}^n\setminus\cup_{\ell=1}^{j-1}\mathcal{B}(u_\ell^n)}\Pr\left[X^n\in\mathcal{B}(x^n)\setminus\cup_{\ell=1}^{j-1}\mathcal{B}(u_\ell^n)\right]\\
&=a_j;
\end{aligned}$$

(7) follows from the definition of p_j; (8) holds since $p_j<a_j+\epsilon$; (9) applies since $\sum_{j=1}^{k}p_j=1$; (10) is a consequence of the Cauchy-Schwarz inequality;[1] and the last inequality in (11) follows from $k\leq M_n^*(d)$.

[1]The Cauchy-Schwarz inequality can be used to assert that $1=\left(\sum_{j=1}^{k}1\cdot p_j\right)^2\leq\left(\sum_{j=1}^{k}1^2\right)\left(\sum_{j=1}^{k}p_j^2\right)=k\sum_{j=1}^{k}p_j^2$.

The proof of (4) is completed by noting that the above derivations hold for arbitrarily small ϵ.

2) Converse (*Equality of* (5) *under finite* $M_n^*(d)$): Let $P_{X^{n*}}$ be the uniform distribution over a distance-d code $\mathcal{C}^*$ that achieves $M_n^*(d)$. We then have

$$\Pr[\tilde{\mu}(\hat{X}^{n*}, X^{n*}) < d] = \sum_{x^n \in \mathcal{C}^*} \left[P_{X^{n*}}(x^n)\right]^2 = \sum_{x^n \in \mathcal{C}^*} \frac{1}{|\mathcal{C}^*|^2} = \frac{1}{M_n^*(d)}, \quad (12)$$

where $|\mathcal{C}^*|$ denotes the cardinality of $\mathcal{C}^*$.

The above two steps complete the proof of

$$\min_{P_{X^n}} \Pr\left[\tilde{\mu}(\hat{X}^n, X^n) < d\right] = \frac{1}{M_n^*(d)}.$$

subject to finite $M_n^*(d)$.

When $M_n^*(d) = \infty$, again, let $\mathcal{C}^*$ denote an infinite distance-d code that achieves $M_n^*(d)$. Then, any finite subset $\mathcal{S}$ of $\mathcal{C}^*$ is a distance-d code. Using a derivation similar to that leading to (12) gives that

$$\Pr[\tilde{\mu}(\hat{X}^{n\circ}, X^{n\circ}) < d] = \frac{1}{|\mathcal{S}|},$$

where $P_{X^{n\circ}}$ is the uniform distribution over $\mathcal{S}$. As $|\mathcal{S}|$ can be made arbitrarily large,

$$\inf_{P_{X^n}} \Pr\left[\tilde{\mu}(\hat{X}^n, X^n) < d\right] = 0.$$

This completes the proof. ■

Some remarks concerning Theorem 1 are in order. First, the theorem can be applied to an arbitrary code alphabet and any distance measure satisfying (1). Its generality thus extends the study of the maximal code size of distance-d codes from the conventional finite code alphabets and the Hamming distance to, for example, $\mathcal{X} = [0, 1)$ and the Euclidean distance (cf. Example 1).

Secondly, the crux of the proof of Theorem 1 is the observation that the entire space $\mathcal{X}^n$ can be covered by k "open" balls of radius d with $k \le M_n^*(d)$,[2] where the radius is defined via the distance $\tilde{\mu}(\cdot,\cdot)$. In addition, the selection of the center u_i^n of the next ball $\mathcal{B}(u_i^n)$ is chosen such that $p_i = \Pr[X^n \in \mathcal{B}(u_i^n) \setminus \cup_{j=1}^{i-1}\mathcal{B}(u_j^n)]$ is ϵ-close to its minimum possible value and therefore k can be made as large as possible, ideally as close to $M_n^*(d)$ as possible.

Thirdly, as noted by Korn [15], when the code alphabet $\mathcal{X}^n$ is finite, the optimization problem $\inf_{P_{X^n}} \Pr[\tilde{\mu}(\hat{X}^n, X^n) < d]$ corresponds exactly to the minimization of the quadratic form $\mathbf{p}\mathbb{A}\mathbf{p}^{\mathsf{T}}$, where $\mathbf{p}$ is the row vector formed by listing the probability masses of P_{X^n} and $\mathbb{A}$ is the corresponding $|\mathcal{X}|^n \times |\mathcal{X}|^n$ matrix with entries given by $\mathbf{1}\{\tilde{\mu}(\hat{x}^n, x^n) < d\}$. This quadratic optimization problem was considered by Korn [15] in his study of the maximization of Gallager's lower bound for the zero-error capacity of discrete memoryless channels (DMCs) [17]. The same solution can also be found in Motzkin and Straus' work [14], where the order of the maximal complete graph contained in a *finite graph* is considered. Here, instead of iteratively removing one codeword from any two codewords within distance d until the size of the set of candidate codewords is reduced to $M_n^*(d)$ as suggested by Korn's technique in [15], we define a "proper" notion of progress to iteratively add codewords to a distance-d code. Specifically, we select a representative vector u_i^n in some "ϵ-neighborhood" defined as $\{u^n \in \mathcal{X}^n : \Pr[X^n \in \mathcal{B}(u^n) \setminus \cup_{j=1}^{i-1}\mathcal{B}(u_j^n)] < \inf_{x^n\in\mathcal{X}^n} \Pr[X^n \in \mathcal{B}(x^n) \setminus \cup_{j=1}^{i-1}\mathcal{B}(u_j^n)] + \epsilon\}$ for a given distribution P_{X^n}. This selection is repeated until the entire code alphabet can be covered by the union of radius-d balls centered at u_i's. The assumed finiteness of $M_n^*(d)$ ensures that the iterative selection will terminate. Note that the proof of Theorem 1 is not restricted to code alphabets that are finite (cf. [14], [15]). In addition to being applicable to general arbitrary code alphabets, it provides a different perspective of the general formula in (3).

[2]Here "open" means a strict inequality is used to define the ball.

Lastly, we recall that finding the maximal distance-d code size is equivalent to obtaining the zero-error capacity [18]. Consequently, Theorem 1 can be used to establish a general formula for the zero-error capacity for *arbitrary* channels as summarized below. This result complements the general formula for the (vanishing error) capacity of arbitrary channels considered by Verdú and Han in [19].

Definition 1 (Zero-error capacity): Let Ω_n be the maximum code size that can be transmitted error-free (i.e., with exactly zero error probability) over the channel $P_{Y^n|X^n}$. Then, the zero-error capacity for a sequence of channels $\{P_{Y^n|X^n}\}_{n=1}^{\infty}$ is defined as

$$C_0 \triangleq \sup_{n\ge1} \frac{1}{n} \log \Omega_n.$$

Corollary 1 (General zero-error capacity): The zero-error capacity for an arbitrary sequence of channels $\{P_{Y^n|X^n}\}_{n=1}^{\infty}$ (not necessarily with countable alphabets) can be expressed as

$$C_0 = \sup_{n\ge1} \; -\frac{1}{n} \log \inf_{P_{X^n}} \Pr[\mu(\hat{X}^n, X^n) = 0], \quad (13)$$

where

$$\mu(\hat{x}^n, x^n) \triangleq \begin{cases} 1, & (\exists\ \mathcal{T} \subset \mathcal{Y}^n)\ \Pr(Y^n \in \mathcal{T} | X^n = \hat{x}^n) \\ & \quad = \Pr(Y^n \notin \mathcal{T} | X^n = x^n) = 1; \\ 0, & \text{otherwise}. \end{cases} \quad (14)$$

When we particularize the infimum in (13) to product distributions, we obtain that for DMCs with finite channel input alphabet $\mathcal{X}$ and finite channel output alphabet $\mathcal{Y}$ [15],

$$C_0 \ge -\frac{1}{n} \log \inf_{P_X} \Pr\left[\sum_{y\in\mathcal{Y}} P_{Y|X}(y|\hat{X}) P_{Y|X}(y|X) > 0\right].$$

Note that when $\mathcal{Y}$ is finite, (14) implies that $\mu(\hat{x}^n, x^n) = 0$ if and only if there exists an y^n such that the channel $P_{Y^n|X^n}$ maps both $\hat{x}^n$ and x^n to y^n with positive probabilities, i.e., $\hat{x}^n$ and x^n are *confusable* [18].

III. Implications of the Distance Spectrum Formula for $M_n^*(d)$

In this section, we further explore the implications of the theoretical result presented in the previous section. Specifically, we show that the GV lower bound for discrete alphabets can be recovered from (3) by letting P_{X^n} be a uniform distribution over $\mathcal{X}^n$. An example in which the alphabet $\mathcal{X}$ is continuous and hence uncountable is also provided.

An immediate consequence of Theorem 1 is that a family of lower bounds to $M_n^*(d)$ can be obtained by evaluating $L_{X^n}(d) \triangleq 1/\Pr[\min\{\mu(\hat{X}^n, X^n), \mu(X^n, \hat{X}^n)\} < d]$ for different distributions P_{X^n}. This implies even if we do not use an optimal distribution P_{X^n}, we may still be able to obtain good lower bounds to the optimal code size. In addition, the converse proof of Theorem 1 shows that $M_n^*(d)$ can actually be achieved using a distribution which is *uniform* over an appropriate subset of $\mathcal{X}^n$ (that is, over an optimal code). Thus, $L_{X^n}(d)$ based on uniform P_{X^n} is an important family of lower bounds to $M_n^*(d)$.

In particular, the Gilbert-Varshamov (GV) lower bound [4] can be recovered with a uniform distribution over all possible codewords. As an example, consider a finite code alphabet $\mathcal{X}$ with $|\mathcal{X}| = Q$ and the Hamming distance measure $\mu(\cdot,\cdot)$. Let the components of $X^n = (X_1\ X_2\ \ldots\ X_n)$ be i.i.d. and uniform over $\mathcal{X}$. This choice yields exactly the GV lower bound $G_n(d)$ [4]:

$$\begin{aligned} M_n^*(d) \geq L_{X^n}(d) &= \frac{1}{\Pr[\sum_{i=1}^n \mu(\hat{X}_i, X_i) < d]} \\ &\geq \frac{Q^n}{\sum_{i=0}^{d-1}\binom{n}{i}(Q-1)^i} \triangleq G_n(d). \end{aligned} \tag{15}$$

The same observation has been stated by Kolesnik and Krachkovsky in [20, pp. 1446].

Next, two examples are given, where the corresponding GV lower bound $G_n(d)$ are obtained.

Example 1: Here we derive lower bounds to $M_2^*(d)$ for Euclidean distance $\mu(\cdot,\cdot)$ and a bounded code alphabet $\mathcal{X} = [0,1)$. Taking P_{X^2} to be the uniform distribution over $\mathcal{X}^2$ and letting $Z_i \triangleq (\hat{X}_i - X_i)^2$ yields that for $d > 0$,

$$\begin{aligned} \Pr[\mu(\hat{X}^2, X^2) < d] &= \Pr\left[Z_1 + Z_2 < d^2\right] \\ &= \int_0^1 \int_0^{d^2 - z_1} f_Z(z_1)\, f_Z(z_2) \mathrm{d}z_2\, \mathrm{d}z_1, \end{aligned}$$

where $f_Z(z) = \left(\frac{1}{\sqrt{z}} - 1\right)\mathbf{1}\{0 \leq z < 1\}$, which implies

$$M_2^*(d) \geq \lceil G_2(d) \rceil = \lceil L_{X^2}(d) \rceil = \begin{cases} 3, & d = \frac{1}{2}; \\ 2, & d = 1. \end{cases} \tag{16}$$

Via a procedure suggested by the proof of Theorem 1, we can actually obtain

$$M_2^*(d) \geq \begin{cases} 8, & d = \frac{1}{2}; \\ 2, & d = 1. \end{cases}$$

This indicates that there is room for improving the generalised $G_n(d)$ (i.e., $L_{X^n}(d)$ with respect to uniform P_{X^n} over $\mathcal{X}^2$) and the codeword selection procedure in the proof of Theorem 1 could be further explored for finding a better lower bound.

Example 2: In this example, we demonstrate a case that $M_n^*(d)$ can be exactly determined. Let the distance measure be given by $\mu(\hat{x}^n, x^n) = |\kappa_n(\hat{x}^n) - \kappa_n(x^n)|$, where $\hat{x}^n$ and x^n are in $\{0,1\}^n$, and $\kappa_n(x^n) \triangleq x_n 2^{n-1} + x_{n-1}2^{n-2} + \ldots + x_2 2^1 + x_1$ is the binary representation of $x^n = (x_1\ x_2\ \ldots\ x_n)$. In other words, $\mu(\hat{x}^n, x^n)$ is the absolute difference between two decimal numbers $\kappa_n(\hat{x}^n)$ and $\kappa_n(x^n)$, and is a *separable distance measure* [21, Def. 1].

Since $\kappa_n(x^n)$ is an integer in $\{0, 1, 2, \ldots, 2^n - 1\}$, it can be easily seen that for $d > 0$,

$$M_n^*(d) = \left\lceil \frac{2^n}{\lceil d \rceil} \right\rceil,$$

where $\lceil\cdot\rceil$ is the ceiling function. Notably, one of the uniform X^n's that results in $L_{X^n}(d) = M_n^*(d)$ has support $\{0, \lceil d\rceil, 2\lceil d\rceil, \ldots, (M_n^*(d)-1)\lceil d\rceil\}$, and there are exactly $\lceil d\rceil$ optimizers that can achieve $M_n^*(d)$. We then recall that (15) has illustrated that $G_n(d)$ can be regarded as a special case of $L_{X^n}(d)$ with uniform X^n over the entire $\mathcal{X}^n$. As such, we derive

$$G_n(d) = \begin{cases} \dfrac{2^{2n}}{(3\lceil d\rceil - 1)\lceil d\rceil + (2\lceil d\rceil - 1)(2^n - 2\lceil d\rceil)}, \\ \qquad\qquad \text{for } 0 < \lceil d\rceil \leq 2^{n-1}; \\ \dfrac{2^{2n}}{2^{2n} + (\lceil d\rceil - 2^n)(2^n - \lceil d\rceil + 1)}, \\ \qquad\qquad \text{for } 2^{n-1} < \lceil d\rceil \leq 2^n - 1; \\ 1, \qquad\qquad \text{for } \lceil d\rceil > 2^n - 1, \end{cases}$$

showing that $G_n(d)$ is strictly less than $M_n^*(d)$ except when $\lceil d\rceil = 1$ and $\lceil d\rceil \geq 2^n$. This result confirms that the finite length GV lower bound is not tight in general.

We close this example by noting that an upper bound $U_n(d)$ for $M_n^*(d)$ can also be provided based on Theorem 1. If there exists a function $U_n(d)$ such that

$$U_n(d) \geq \frac{1}{\Pr\left[\min\{\mu(\hat{X}^n, X^n), \mu(X^n, \hat{X}^n)\} < d\right]}$$

for all P_{X^n}'s, then

$$\begin{aligned} U_n(d) &\geq \frac{1}{\inf_{P_{X^n}} \Pr\left[\min\{\mu(\hat{X}^n, X^n), \mu(X^n, \hat{X}^n)\} < d\right]} \\ &= M_n^*(d). \end{aligned}$$

Now setting $j = j(n,d) \triangleq 2^n/\lceil d\rceil$, we derive

$$\begin{aligned} &\Pr\left\{\left|\frac{\kappa_n(\hat{X}^n)}{2^n} - \frac{\kappa_n(X^n)}{2^n}\right| < \frac{\lceil d\rceil}{2^n}\right\} \\ &\geq \sum_{i=0}^{\lceil j\rceil - 1} \Pr\left\{\frac{i}{\lceil j\rceil} \leq \frac{\kappa_n(\hat{X}^n)}{2^n} < \frac{i+1}{\lceil j\rceil}\right. \\ &\qquad\qquad \left.\text{and } \frac{i}{\lceil j\rceil} \leq \frac{\kappa_n(X^n)}{2^n} < \frac{i+1}{\lceil j\rceil}\right\} \\ &= \sum_{i=0}^{\lceil j\rceil - 1}\left(\Pr\left\{\frac{i}{\lceil j\rceil} \leq \frac{\kappa_n(X^n)}{2^n} < \frac{i+1}{\lceil j\rceil}\right\}\right)^2 \\ &\geq \frac{1}{\lceil j\rceil}, \end{aligned} \tag{17}$$

where (17) again follows from the Cauchy-Schwarz inequality. This gives an upper bound coinciding with $M_n^*(d)$

$$U_n(d) = \lceil j\rceil = \left\lceil \frac{2^n}{\lceil d\rceil} \right\rceil = M_n^*(d).$$

□

IV. Extensions to the Asymptotic Regime

We now extend the result in Theorems 1 to the asymptotic regime in which the length n of the code goes to infinity. In what follows, $\log$ denotes the natural logarithm. A distance spectrum formula for the largest code rate $R = \log(M)/n$ subject to a normalized minimum distance $\delta = d/n$ can be obtained on the basis of Theorem 1 in a straightforward manner:

$$\begin{aligned} R_n^*(\delta) &\triangleq \frac{1}{n}\log M_n^*(n\delta) \\ &= \sup_{P_{X^n}}\left(-\frac{1}{n}\log \Pr\left[\frac{1}{n}\mu(\hat{X}^n, X^n) < \delta\right]\right). \quad (18) \end{aligned}$$

The formula of $R_n^*(\delta)$ in (18) provides a quantitative characterization of the largest code rate attainable for an $(n, M, n\delta)$-code, based on which a first-order expression for the largest asymptotic code rate attainable for a sequence of $(n, M, n\delta)$-codes can be obtained when the normalized distance measure is uniformly bounded.

Theorem 2: (Largest Asymptotic Code Rate) Fix an arbitrary code alphabet $\mathcal{X}$ and a (sequence of) general distance measures $\mu(\cdot,\cdot)$ that satisfy the condition mentioned in Theorem 1 and also satisfy

$$\sup_{n\geq 1}\max_{\hat{x}^n, x^n\in\mathcal{X}^n}\frac{1}{n}\mu(\hat{x}^n, x^n) < \infty. \quad (19)$$

Then,

$$\limsup_{n\to\infty} R_n^*(\delta) = \limsup_{n\to\infty}\sup_{P_{X^n}} J_{X^n}(\delta)$$

and

$$\liminf_{n\to\infty} R_n^*(\delta) = \liminf_{n\to\infty}\sup_{P_{X^n}} J_{X^n}(\delta),$$

where

$$J_{X^n}(\delta) \triangleq \inf_{a\leq\delta}\sup_{\theta\in\mathbb{R}}\left\{a\theta - \frac{1}{n}\log\mathbb{E}\left[e^{\theta\mu(\hat{X}^n, X^n)}\right]\right\}. \quad (20)$$

Proof: The proof can be found in Appendix A. In particular, an upper bound on the second-order term of $R_n^*(\delta)$ is also provided (cf. Lemma 2). ■

The above theorem indicates that $R_n^*(\delta)$ and $\sup_{P_{X^n}} J_{X^n}(\delta)$ are asymptotically close. In fact, the proof in Appendix A shows that

$$R_n^*(\delta) \geq \sup_{P_{X^n}} J_{X^n}(\delta) \quad (21)$$

for every n. However, the proof of the upper bound (which shows that the second-order term is $O(1/\sqrt{n})$) is significantly more involved and requires delicate twistings of probability distributions [22]. Using a large deviations technique, we can slightly improve (21) by the addition of a logarithmic term. For example, when $\mathcal{X}$ is binary and $\mu(\cdot,\cdot)$ is the Hamming distance measure,

$$R_n^*(\delta) \geq D\left(\delta\,\middle\|\,\frac{1}{2}\right) + \frac{\log n}{2n} + \Theta\left(\frac{1}{n}\right) \quad \text{as } n\to\infty. \quad (22)$$

Although Jiang and Vardy [23, Thm. 1] have shown, by using a graph-theoretic framework, that the achievable second-order term in (22) is at least $(\log n)/n$, which is slightly stronger than the term $(\log n)/(2n)$, Eq. (22) provides some additional insight into the suboptimality of choosing $\hat{X}^n$ and X^n with i.i.d. components.

V. Conclusion and Future Work

In this paper, we developed an exact formula for the maximal size of distance-d codes for arbitrary alphabets and general distance measures. The implications of the established formula were discussed. The extension to the asymptotic regime was also explored. Some natural directions for future work include:

- Understanding the structure of optimal or even "good" distributions P_{X^n} to give lower bounds on the optimal code size. For example, based on our numerical experiments, we know that the optimal distribution may not be unique. Studying the binary Hamming distance for small block lengths suggests that there may be an optimizer whose marginals are uniform on each coordinate.
- Finding $i)$ a similar formula of the minimum code size subject to a covering radius constraint (cf. [24]) and $ii)$ a formula of maximal code size under a minimum multi-wise distance constraint (cf. [25]). The latter would constitute a generalization of Turán's Theorem.

Appendix A
Proof of Theorem 2

The theorem can be verified via the following two lemmas. The first lemma shows that for arbitrary distance measures, $R_n^*(\delta)$ is lower-bounded by $\sup_{P_{X^n}} J_{X^n}(\delta)$. The second lemma proves that $R_n^*(\delta)$ is upper bounded by $\sup_{X^n} J_{X^n}(\delta) + \Theta(\frac{1}{\sqrt{n}})$ when the normalized distance measure is uniformly bounded. Then, the two lemmas imply Theorem 2.

Lemma 1: Fix an arbitrary code alphabet and an arbitrary distance measure that satisfies (1). Then,

$$R_n^*(\delta) \geq \sup_{P_{X^n}} J_{X^n}(\delta). \quad (23)$$

Proof: This is a consequence of two observations that for $\theta > 0$,

$$\begin{aligned} \Pr\left[\frac{1}{n}\mu(\hat{X}^n, X^n) < \delta\right] &= \Pr[Y > 0] = \Pr[e^{\theta Y} > 1] \\ &\leq \mathbb{E}[e^{\theta Y}] \triangleq M_Y(\theta), \quad (24) \end{aligned}$$

where $Y \triangleq n\delta - \mu(\hat{X}^n, X^n)$, and that for $\delta < \frac{1}{n}\mathbb{E}[\mu(\hat{X}^n, X^n)]$,

$$\begin{aligned} \Pr[Y > 0] &\leq \inf_{\theta>0} M_Y(\theta) = \inf_{\theta\in\mathbb{R}} M_Y(\theta) = \inf_{\theta\in\mathbb{R}} M_Y(-\theta) \\ &= \exp\left\{-n\sup_{\theta\in\mathbb{R}}\left(\delta\theta - \frac{1}{n}\log\mathbb{E}\left[e^{\theta\mu(\hat{X}^n, X^n)}\right]\right)\right\} \\ &= \exp\{-n\cdot J_{X^n}(\delta)\}. \end{aligned}$$

■

Lemma 2: Fix an arbitrary code alphabet and an arbitrary distance measure that satisfies both (1) and (19). Then,

$$R_n^*(\delta) \leq \sup_{X^n} J_{X^n}(\delta) + \Theta(\tfrac{1}{\sqrt{n}}) \quad (25)$$

for those δ satisfying $\sup_{P_{X^n}} J_{X^n}(\delta) > 0$.

Proof: Given that P_{X^n} is the optimizer of $\sup_{X^n} J_{X^n}(\delta)$ and following the notations used in the proof of Lemma 1, we define the twisted distribution of Y as $\mathrm{d}P_{Y^{(\theta)}}(y) \triangleq e^{\theta y}\,\mathrm{d}P_Y(y)/M_Y(\theta)$. Then,

$$\begin{aligned}\Pr[Y>0] &= \int_0^\infty \mathrm{d}P_Y(y) = \int_0^\infty M_Y(\theta^*)e^{-\theta^* y}\,\mathrm{d}P_{Y^{(\theta^*)}} \\ &= M_Y(\theta^*)\int_0^\infty e^{-\theta^* y}\,\mathrm{d}P_{Y^{(\theta^*)}}(y), \qquad (26)\end{aligned}$$

where θ^* is the minimizer of $\inf_{\theta\in\mathbb{R}} M_Y(\theta)$. Let W be a nonnegative random variable with distribution $\mathrm{d}P_W(y) \triangleq \mathrm{d}P_{Y^{(\theta^*)}}(y)/\Pr[Y^{(\theta)}>0]$. Then, (26) can be rewritten as

$$\begin{aligned}\Pr[Y>0] &= M_Y(\theta^*)\cdot\Pr[Y^{(\theta^*)}>0]\int_0^\infty e^{-\theta^* y}\,\mathrm{d}P_W(y) \\ &= M_Y(\theta^*)\cdot\Pr[Y^{(\theta^*)}>0]\cdot\mathbb{E}\big[e^{-\theta^* W}\big].\end{aligned}$$

Using the fact that $\mathbb{E}[Y^{(\theta^*)}]=0$ [26, Thm. 9.2], we obtain

$$\begin{aligned}\frac{1}{\Pr[Y^{(\theta^*)}>0]} &\le \frac{4\,\mathbb{E}[(Y^{(\theta^*)})^4]}{\mathbb{E}^2[(Y^{(\theta^*)})^2]} \\ &= 4\left(\frac{1}{n}\frac{\varphi^{(4)}_{X^n}(-\theta^*)}{(\varphi''_{X^n}(-\theta^*))^2}+3\right) \qquad (27)\end{aligned}$$

where $\varphi_{X^n}(\theta) \triangleq \frac{1}{n}\log\mathbb{E}[e^{\theta\mu(\hat{X}^n,X^n)}]$. Using Jensen's inequality, i.e., $\mathbb{E}[e^{-\theta^* W}] \ge e^{-\theta^*\cdot\mathbb{E}[W]}$, and

$$\begin{aligned}\mathbb{E}[W] &= \int_0^\infty y\frac{\mathrm{d}P_{Y^{(\theta^*)}}(y)}{\Pr[Y^{(\theta^*)}>0]} \\ &\le \frac{1}{\Pr[Y^{(\theta^*)}>0]}\int_{-\infty}^\infty |y|\,\mathrm{d}P_{Y^{(\theta^*)}}(y) \\ &\le \frac{1}{\Pr[Y^{(\theta^*)}>0]}\sqrt{\mathbb{E}[(Y^{(\theta^*)})^2]} \\ &= \frac{1}{\Pr[Y^{(\theta^*)}>0]}\sqrt{n\cdot\varphi''_{X^n}(-\theta^*)},\end{aligned}$$

we conclude from all the above derivations that

$$\Pr[Y>0] \ge e^{-n\cdot J_{X^n}(\delta)} \times\frac{\exp\left\{4\theta^*\left(\frac{1}{n}\frac{\varphi^{(4)}_{X^n}(-\theta^*)}{(\varphi''_{X^n}(-\theta^*))^2}+3\right)\sqrt{n\cdot\varphi''_{X^n}(-\theta^*)}\right\}}{4\left(\frac{1}{n}\frac{\varphi^{(4)}_{X^n}(-\theta^*)}{(\varphi''_{X^n}(-\theta^*))^2}+3\right)}.$$

We completes the proof of (25) by remarking that with probability one, $(1/n)\mu(\hat{X}^n,X^n)$ is not only bounded, but uniformly upper bounded in the block length n, and so are its moments and cumulants. Since a twisted random variable generated from $(1/n)\mu(\hat{X}^n,X^n)$ must have the same support as $(1/n)\mu(\hat{X}^n,X^n)$, its twisted moments as well as twist cumulants are also uniformly bounded. Accordingly, $\varphi^{(4)}_{X^n}(-\theta^*)=O(1)$ and $\varphi''_{X^n}(-\theta^*)=O(1)$, based on which (25) implies $R_n^*(\delta)\le\sup_{X^n} J_{X^n}(\delta)+\Theta(\frac{1}{\sqrt{n}})$. ∎

Acknowledgements

A special acknowledgement is given to Dr. Mladen Kovačević, who brought [20] to the authors' attention. The authors sincerely thank the anonymous reviewers for their constructive feedback to improve the quality of the paper.

References

[1] R. E. Blahut, *Principles and Practice of Information Theory*. Addison Wesley, 1987.

[2] L. Lovasz, "On the Shannon capacity of a graph," *IEEE Trans. Inf. Theory*, vol. 25, no. 1, pp. 1–7, 1979.

[3] R. J. McEliece, E. R. Rodemich, H. Rumsey, Jr., and L. R. Welch, "New upper bounds on the rate of a code via the Delsarte-MacWilliams inequalities," *IEEE Trans. Inf. Theory*, vol. 23, no. 2, pp. 157–166, March 1977.

[4] T. K. Moon, *Error Correction Coding: Mathematical Methods and Algorithms*. Wiley, 2005.

[5] J. H. van Lint, *Introduction to Coding Theory*, 2nd ed. New York: Springer Verlag, 1992.

[6] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*. Amsterdam, New York, and North-Holland: Oxford University Press, 1983.

[7] S. N. Litsyn and M. A. Tsfasman, "A note on lower bounds," *IEEE Trans. Inf. Theory*, vol. IT-32, no. 5, pp. 705–706, September 1986.

[8] T. Ericson and V. A. Zinoviev, "An improvement of the Gilbert bound for constant weight codes," *IEEE Trans. Inf. Theory*, vol. IT-33, pp. 721–723, September 1987.

[9] S. G. Vladut, "An exhaustion bound for algebraic-geometric modular codes," *Probl. Inform. Transm.*, vol. 23, pp. 22–34, September 1987.

[10] M. Svanstrom, "A lower bound for ternary constant weight codes," *IEEE Trans. Inf. Theory*, vol. 43, no. 5, pp. 1630–1632, September 1997.

[11] H. Stichtenoth and C. Xing, "Excellent nonlinear codes from algebraic function fields," *IEEE Trans. Inf. Theory*, vol. 51, no. 11, pp. 4044–4046, October 2005.

[12] P. Gaborit and G. Zemor, "Asymptotic improvement of the Gilbert-Varshamov bound for linear codes," *IEEE Trans. Inf. Theory*, vol. 54, no. 9, pp. 3865–3872, August 2008.

[13] A. Bassa, P. Beelen, A. Garcia, and H. Stichtenoth, "An improvement of the Gilbert-Varshamov bound over nonprime fields," *IEEE Trans. Inf. Theory*, vol. 60, no. 7, pp. 3859–3861, April 2014.

[14] T. S. Motzkin and E. G. Straus, "Maxima for graphs and a new proof of a theorem of Turan," *Canad. J. Math*, vol. 17, no. 4, pp. 533–540, 1965.

[15] I. Korn, "On the lower bound of zero-error capacity," *IEEE Trans. Inf. Theory*, vol. 40, no. 4, pp. 509–510, May 1968.

[16] S. Lin and D. J. Costello, Jr., *Error Control Coding*, 2nd ed. Upper Saddle River, NJ, USA: Prentice-Hall, 2004.

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

[18] C. E. Shannon, "The zero error capacity of a noisy channel," *IEEE Trans. Inf. Theory*, vol. 2, no. 3, pp. 8–19, September 1956.

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

[20] V. D. Kolesnik and V. Y. Krachkovsky, "Lower bounds on achievable rates for limited bitshift correcting codes," *IEEE Trans. Inf. Theory*, vol. 40, no. 5, pp. 1443–1458, September 1994.

[21] P.-N. Chen, T.-Y. Lee, and Y. S. Han, "Distance-spectrum formulas on the largest minimum distance of block codes," *IEEE Trans. Inf. Theory*, vol. 46, no. 3, pp. 869–885, May 2000.

[22] P.-N. Chen, "Generalization of Gartner-Ellis theorem," *IEEE Trans. Inf. Theory*, vol. 46, no. 7, pp. 2752–2760, 2000.

[23] T. Jiang and A. Vardy, "Asymptotic improvement of the Gilbert-Varshamov bound on the size of binary codes," *IEEE Trans. Inf. Theory*, vol. 50, no. 8, pp. 1655–1664, November 2004.

[24] P.-N. Chen and Y. S. Han, "Asymptotic minimum covering radius of block codes," *SIAM J. Discrete Math.*, vol. 14, no. 4, pp. 549–564, 2001.

[25] H.-Y. Lin, S. M. Moser, and P.-N. Chen, "The r-wise Hamming distance and its operational interpretation for block codes," in *52th Annual Conference on Information Sciences and Systems (CISS)*, Princeton, NJ, USA, Mar 2018.

[26] G. van der Geer and J. H. van Lint, *Large Deviation Techniques in Decision, Simulation, and Estimation*. New York, NY, USA: Wiley, 1990.

Ling-Hua Chang is currently an assistant professor in the Department of Electrical Engineering (group A) at Yuan Ze University, Taiwan. She received the B. S. and Ph.D. degrees in Electrical Engineering from National Chiao

Tung University, Taiwan, respectively in 2010 and 2016. 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. From 2012–2015, he was the associate chief director of Microelectronics and Information Systems Research Center, 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. His research interests generally lie in information and coding theory, large deviation theory, distributed detection and sensor networks.

Vincent Y. F. Tan (S'07-M'11-SM'15) was born in Singapore in 1981. He is currently a Dean's Chair Associate Professor in the Department of Electrical and Computer Engineering and the Department of Mathematics at the National University of Singapore (NUS). He received the B.A. and M.Eng. degrees in Electrical and Information Sciences from Cambridge University in 2005 and the Ph.D. degree in Electrical Engineering and Computer Science (EECS) from the Massachusetts Institute of Technology (MIT) in 2011. His research interests include information theory, machine learning, and statistical signal processing.

Dr. Tan received the MIT EECS Jin-Au Kong outstanding doctoral thesis prize in 2011, the NUS Young Investigator Award in 2014, the NUS Engineering Young Researcher Award in 2018, and the Singapore National Research Foundation (NRF) Fellowship (Class of 2018). He is also an IEEE Information Theory Society Distinguished Lecturer for 2018/9. He has authored a research monograph on *"Asymptotic Estimates in Information Theory with Non-Vanishing Error Probabilities"* in the Foundations and Trends in Communications and Information Theory Series (NOW Publishers). He is currently serving as an Associate Editor of the IEEE Transactions on Signal Processing.

Carol Wang received her Bachelor's degree from the California Institute of Technology in 2010 and her Ph.D. in 2015 from Carnegie Mellon University, supported by a National Science Foundation graduate research fellowship. Following her Ph.D., she was a research fellow in the Department of Electrical and Computer Engineering at the National University of Singapore. Dr. Wang works primarily in algebraic coding theory, and is always interested in extending coding-theoretic techniques to new coding models.

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. Now he is with School of Electrical Engineering & Intelligentization at Dongguan University of Technology, 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.