Invited Paper

Quotient Codes and Their Reliability

Mitsuru Hamada[†,††]

This article gives a formula to evaluate the performance of a class of algebraic codes. The class includes quantum codes as well as classical ones. The formula relates a bound on the weight spectrum (distribution), or its generalization, of a code with an upper bound on its decoding error probability.

1. Introduction

The first algebraic quantum error-correcting code (QECC) was invented by Shor [1] in 1995 soon after his discovery of the prime factoring algorithm for quantum computation. The code has been extended to a general class of algebraic QECCs, which are called symplectic codes or stabilizer codes [2)~4)] and have a resemblance with classical linear codes. Due to this resemblance, the design of QECCs, in part, reduces to that of linear codes over finite fields. We discuss this algebraic coding theoretic part of the design of QECCs. In particular, this article gives a formula to evaluate the performance of algebraic codes. The resemblance between classical codes and quantum codes enables us to treat both of them in the proposed formula, which relates a bound on the weight spectrum (distribution), or its generalization, of a code with an upper bound on its decoding error probability. The idea of relating weight spectra with error probability dates back, at least, to 1963 when Gallager's book [5] appeared. The present paper's approach is closer to that of Goppa's [6] and that of Csiszár's [7].

This paper is primarily of expository nature, and most parts (the parts except Section 10) may be viewed as extracted from previous results of the present author's [8),9)]. However, to clarify the exposition, this paper introduces a general framework of codes, for which a name *quotient codes* is coined. As a byproduct, the paper shows that the expurgated exponent of the binary symmetric channel is attainable with a universal scheme of encoding and decoding that are independent of the channel parameter, which resolves a small problem of classical coding.

This paper is organized as follows. Section 2 contains preliminaries, Section 3 introduces the quotient codes, and Section 4 gives the formula on the performance of codes, which is followed by an application to classical coding in Section 5. Section 6 describes symplectic codes, and Section 7 gives an application of the formula to symplectic codes. In Section 8, physical aspects of symplectic codes are explained. In Section 9, we discuss an important class of symplectic codes with applications. Sections 10 and 11, respectively, contain other applications and a summary with remarks.

2. Preliminaries from Classical Theory of Information and Coding

Here we will recall well-known notions and facts in the theory of information and coding [10)~12)].

The *type* of a sequence $y = (y_1, \ldots, y_n) \in \mathcal{Y}^n$ is denoted by P_y, which is defined by

$$\mathsf{P}_y(s) = \frac{|\{i \mid 1 \le i \le n, y_i = s\}|}{n}, \quad s \in \mathcal{Y}.$$

For a fixed type Q, we put $\mathcal{T}_Q^n = \{y \in \mathcal{Y}^n \mid \mathsf{P}_y = Q\}$. The set of all probability distributions on $\mathcal{Y}$ and that of all types of sequences in $\mathcal{Y}^n$ are denoted by $\mathcal{P}(\mathcal{Y})$ and $\mathcal{P}_n(\mathcal{Y})$, respectively. For any $P \in \mathcal{P}(\mathcal{Y})$, we define $P^n \in \mathcal{P}(\mathcal{Y}^n)$ by $P^n(x_1, \ldots, x_n) = P(x_1) \cdots P(x_n)$. Given a random variable $\mathbf{X}$, $P_{\mathbf{X}}$ stands for the probability distribution of $\mathbf{X}$. The expectation operation with respect to a random variable $\mathbf{X}$ taking values in $\mathcal{X}$ is represented by $\mathsf{E}_{\mathbf{X}}$:

$$\mathsf{E}_{\mathbf{X}} f(\mathbf{X}) = \sum_{x \in \mathcal{X}} P_{\mathbf{X}}(x) f(x)$$

where f is a real-valued function on $\mathcal{X}$. The Shannon entropy and the Kullback-Leibler information are denoted by H and D, respectively:

† Research Center for Quantum Information Science, Tamagawa University Research Institute
†† PRESTO, Japan Science and Technology Agency (JST)

$$H(P) = -\sum_{y\in\mathcal{Y}} P(y)\log_d P(y)$$

and

$$D(P||Q) = \sum_{y\in\mathcal{Y}} P(y)\log_d \frac{P(y)}{Q(y)}$$

with the convention $0\log 0 = 0\log(0/0) = 0$. Throughout the paper, the base of logarithms is always $d > 1$. In what follows, we use the basic inequalities

$$\forall Q \in \mathcal{P}_n(\mathcal{Y}),\ |\mathcal{T}^n_Q| \le d^{nH(Q)}, \tag{1}$$

and

$$\sum_{y\in\mathcal{Y}^n:\,\mathsf{P}_y=Q} P^n(x) \le d^{-nD(Q||P)}. \tag{2}$$

Given a set $C \subseteq \mathcal{Y}^n$, we put $M_Q(C) = |\{y \in C \mid \mathsf{P}_y = Q\}|$ for types Q. The list of numbers $(M_Q(C))_{Q\in\mathcal{P}_n(\mathcal{Y})}$ is a generalization of the weight spectrum, and is equivalent to the complete weight enumerator of C (e.g., Ref. 12)). In this paper, we call $(M_Q(C))_{Q\in\mathcal{P}_n(\mathcal{Y})}$ the *spectrum* of C simply.

The symmetric group on $\{1,\dots,n\}$, which is composed of all permutations on $\{1,\dots,n\}$, is denoted by $\mathcal{S}_n$. We define an action of $\mathcal{S}_n$ on $\mathcal{Y}^n$ by

$$\pi([x_1,\dots,x_n]) = [x_{\pi(1)},\dots,x_{\pi(n)}]$$

for any $\pi \in \mathcal{S}_n$ and $[x_1,\dots,x_n] \in \mathcal{Y}^n$, and put

$$\pi(C) = \{\pi(x) \mid x \in C\},\quad \pi \in \mathcal{S}_n,\ C \subseteq \mathcal{Y}^n.$$

For $a = 0, 1$, we abbreviate the n-tuple $(a,\dots,a)$ as a^n.

Additive group codes or *additive codes* over $\mathcal{Y}$ are subgroups of $\mathcal{Y}^n$, where $\mathcal{Y}$ is an additive group and $\mathcal{Y}^n$ denotes the direct sum of n copies of $\mathcal{Y}$. When $\mathcal{Y}$ is a finite field, additive group codes may be linear subspaces of $\mathcal{Y}^n$, which are known as linear codes. By $B \le C$, we mean B is a subgroup of an additive group C. The finite field of d elements is denoted by $\mathbb{F}_d$.

In (algebraic) coding theory, a code usually means a subset of $\mathcal{Y}^n$. When the cardinality of a code $C \subseteq \mathcal{Y}^n$ is $|\mathcal{Y}|^k$, it is called an $[n,k]$ code. Members of a code are called codewords. The common scenario is that the sender encodes a message from $\mathcal{Y}^k$ into a codeword of an $[n,k]$ code C using a one-to-one map, say φ, and sends it through a noisy channel, which is fed with a symbol from $\mathcal{Y}$ and outputs one from $\mathcal{Y}$ at a time, and the receiver decodes the received word back into a message, usually, first into a codeword c' in C by a decoding map ψ, and then into $\varphi^{-1}(c')$. This coding scenario can be summarized schematically as

$$x \in \mathcal{Y}^k \xrightarrow{\varphi} c \in C \xrightarrow{W} y' \in \mathcal{Y}^n \xrightarrow{\psi} c' \in C \xrightarrow{\varphi^{-1}} x' \in \mathcal{Y}^k,$$

where W is a 'stochastic map' that represents the channel.

A decoder for an additive code C can be designed as follows. Let J be a set of representatives of cosets in the quotient group $\mathcal{Y}^n/C$. Then, we can specify a decoding map, which makes the code 'J-correcting', as follows. Assume the receiver has obtained a word $y \in \mathcal{Y}^n$. The receiver calculates the unique $c \in C$ such that $c + a = y$ for some $a \in J$, and decodes y into c.

A good code should have small decoding error probability as well as an efficient decoding algorithm. This article mostly deals with the issue of finding codes whose decoding error probabilities, or corresponding quantity if they are quantum codes, are small.

Then, what is a good choice for J? In this paper, we adopt the choice made by Goppa [6]. Namely, we fix some real-valued function γ on $\mathcal{Y}^n$ and choose a word x that maximizes $\gamma(x)$ in each coset lying in $\mathcal{Y}^n/C$ and have J consist of those chosen representatives (to break ties, we use an arbitrarily fixed order, say, a lexicographic order in $\mathcal{Y}^n$). We list three examples of γ with the names of the resulting decoders:

(1) $\gamma(x) = -H(\mathsf{P}_x)$: minimum entropy (syndrome) decoding ☆
(2) $\gamma(x) = -$(Hamming weight of x): minimum Hamming distance decoding
(3) $\gamma(x) = P_n(x)$: maximum likelihood decoding (MLD)

The MLD deserves this name if the channel is additive, namely, if a channel input x is changed into y with probability $P_n(y-x)$, $x, y \in \mathcal{Y}^n$. The MLD minimizes the decoding error probability given a code C, but needs the knowledge on the channel characteristics $P_n(x)$, $x \in \mathcal{Y}^n$. An advantage of the minimum entropy (or Hamming distance) decoding is that it does not depend on the channel characteristics.

3. Coding with Equivalence Classes

Motivated by the structure of algebraic quan-

☆ This is different from the minimum-entropy decoding, i.e., the maximum mutual information (MMI) decoding for constant composition codes [10),13),14)], which is an α-decoding [14)]. Our decoding is a β decoding [6)], which may be called a (wide-sense) syndrome decoding when $\mathcal{Y}$ is a finite field.

tum codes, which will be explained shortly, we will slightly modify the coding scenario above mentioned. Briefly speaking, the idea is to introduce an equivalence relation $\sim$ in $\mathcal{Y}^n$ and have the equivalence classes in $\mathcal{Y}^n/\!\sim$ take over the role of the words in $\mathcal{Y}^n$. In particular, assuming $\mathcal{Y}$ and $B \le \mathcal{Y}^n$ are additive, we use the cosets in $\mathcal{Y}^n/B$ as the equivalence classes.

We call either a quotient group C/B or a pair $(C/B, \{\rho_c\}_{c\in C/B})$ a quotient group code, or simply, a quotient code, where $\rho_c \in \mathcal{P}(c)$ for each $c \in C/B$. The coding scenario is as follows. Assume $B \le C \le \mathcal{Y}^n$, $|B| = |\mathcal{Y}|^b$ and $|C| = |\mathcal{Y}|^{k+b}$. A sender encodes a message from $\mathcal{Y}^k$ into $c \in \widetilde{C} = C/B$ with a prescribed one-to-one map, say φ. Then, a word in c is chosen according to the probability distribution ρ_c, and sent through the channel. The receiver decodes the received word into a member, say c'', of $\widetilde{C}$ and then into a message $\varphi^{-1}(c'')$. This coding scenario can be summarized schematically as

$$x \in \mathcal{Y}^k \overset{\varphi}{\to} c \in C/B \overset{\rho_c}{\to} c' \in C \overset{W}{\to} y' \in \mathcal{Y}^n \overset{\psi}{\to} c'' \in C/B \overset{\varphi^{-1}}{\to} x' \in \mathcal{Y}^k.$$

For quotient codes, we can use essentially the same decoding principle as in the previous section. Namely, after mapping the received word y' into $\widetilde{y} = y' + B$, the map that associates $\widetilde{y} \in \mathcal{Y}^n/B$ with $c'' \in \widetilde{C}$ can be constructed as in Section 2. To be more specific, let us call any coset in $\mathcal{V} = \mathcal{Y}^n/B$ a B-coset. Choose a set of representatives of cosets lying in $\mathcal{V}/\widetilde{C}$, and let $\widetilde{J}$ be the union of those chosen representative B-cosets. We have the following decoding map, which makes the code '$\widetilde{J}$-correcting'[☆]. Assume the receiver has received a word $y' \in \mathcal{Y}^n$, and put $\widetilde{y} = y' + B$. The receiver calculates the unique $c'' \in \widetilde{C}$ such that $c'' + a = \widetilde{y}$ for some B-coset $a \subseteq \widetilde{J}$, and decodes $\widetilde{y}$ into c'', and then into $\varphi^{-1}(c'')$.

We remark such a set $\widetilde{J}$ can be specified alternatively as follows. Let J be a set of representatives of cosets in $\mathcal{Y}^n/C$, and put

$$\widetilde{J} = J + B = \{x + y \mid x \in J,\ y \in B\}. \quad (3)$$

If the choice of J is among those of Goppa listed in Section 2, we also call the resulting decoding scheme minimum entropy decoding, MLD, etc. accordingly as which instance of γ is chosen.

If we use a quotient code and the decoding method as just described on a memoryless additive channel that changes an input symbol x into y with probability $P(y - x)$, $x \in \mathcal{Y}$, then the decoding error probability is $P^n(\widetilde{J}^{\mathrm{c}})$ for any sent 'code-coset' $c \in C/B$ and for any probability distribution $\rho_c \in \mathcal{P}(c)$. Note that as compared to the previous section, the set of correctable errors J has been augmented to $\widetilde{J} = J + B$. Our goal is to find codes with small $P^n(\widetilde{J}^{\mathrm{c}})$.

We remark that the coding schemes in this section fall within the conventional framework of coding in the previous section if either $B = \{0^n\}$ or for any $c \in C/B$, there exists a word $t \in c$ with $\rho_c(t) = 1$. Indeed, for the purpose of transmission of classical data only, one would concentrate the probability of ρ_c, say, on a word $t \in c$ that achieves the minimum decoding error probability in c, for any 'code-coset' $c \in C/B$ [☆☆]. However, it will be turned out that quotient codes are useful, at least, for analyses of channel codes, especially, of quantum codes.

4. Bound on Decoding Error Probability

The following is the basic lemma that relates a bound on the spectrum of a code with its performance.

Lemma 1 Assume subgroups B and C with $B \le C \le \mathcal{Y}^n$ satisfy

$$\frac{M_Q(C \setminus B)}{|\mathcal{T}_Q^n|} \le a_n d^{-nT}, \qquad Q \in \mathcal{P}_n(\mathcal{Y})$$

with some real numbers $d > 1$, $a_n \ge 1$ and T. Then, choosing J as in item (1) of the above list (minimum entropy decoding), we have for any $P_n \in \mathcal{P}(\mathcal{Y}^n)$,

$$\mathsf{E}_{\boldsymbol{\pi}} P_n(\boldsymbol{\pi}(\widetilde{J})^{\mathrm{c}}) \le a_n |\mathcal{P}_n(\mathcal{Y})| \sum_{Q\in\mathcal{P}_n(\mathcal{Y})} P_n(\mathcal{T}_Q^n) d^{-n|T - H(Q)|^+}.$$

where c denotes complement, $\widetilde{J}$ is defined in (3), $|t|^+ = \max\{t, 0\}$, and the random variable $\boldsymbol{\pi}$ is uniformly distributed over $\mathcal{S}_n$.

Corollary 1 Assume subgroups B and C with $B \le C \le \mathcal{Y}^n$ have $M_Q(C \setminus B)$ bounded as in Lemma 1. Then, with J and $\widetilde{J}$ as in the lemma, we have for any $P \in \mathcal{P}(\mathcal{Y})$,

[☆] Conversely, a quotient code C/B being $\widetilde{J}$-correcting means $\widetilde{J}$ can be chosen in this way.

[☆☆] More quantitatively, with the codelength n and the size $|\mathcal{M}|$ of the set of messages $\mathcal{M}$ fixed, the minimum of the decoding error probability of a conventional code cannot be improved by a quotient code $(C/B, \{\rho_c\}_{c\in C/B})$ (or a more general coding scheme in which a code, i.e., the image of an encoder, has form $\{\rho_c \in \mathcal{P}(\mathcal{Y}^n) \mid c \in \mathcal{M}\}$).

$$P^n(\widetilde{J}^{\mathrm{c}}) \le a_n|\mathcal{P}_n(\mathcal{Y})|^2 d^{-nE^{(n)}(P,T)}$$

where

$$E^{(n)}(P,T) = \min_{Q\in\mathcal{P}_n(\mathcal{Y})}[D(Q||P) + |T - H(Q)|^+].$$

A proof of Lemma 1 is given in the appendix though it is almost the same as the proof of Theorem 1 of Ref. 9), which gives an upper bound on the 'error probability' of CSS codes.

Proof of Corollary 1. Clearly, $\mathsf{E}_{\boldsymbol{\pi}} P^n(\boldsymbol{\pi}(\widetilde{J})^{\mathrm{c}}) = P^n(\widetilde{J}^{\mathrm{c}})$. Then, inserting the estimate of $P^n(\mathcal{T}_Q^n)$ in Eq. (2) into the bound on $\mathsf{E}_{\boldsymbol{\pi}} P^n(\boldsymbol{\pi}(\widetilde{J})^{\mathrm{c}})$ in the lemma, we have

$$P^n(\widetilde{J}^{\mathrm{c}}) \le a_n|\mathcal{P}_n(\mathcal{Y})| \sum_{Q\in\mathcal{P}_n(\mathcal{Y})} d^{-n[D(Q||P)+|T-H(Q)|^+]}$$

and hence, the corollary.

In general, the spectra of specific codes are hard to calculate. However, calculating the average of the spectra over an ensemble is sometimes easy, in which case, the next lemma is useful to obtain bounds on the spectra of deterministic codes.

Lemma 2 Suppose $\mathcal{Y}$ is a finite set (not necessarily a group), and A is a family of subsets of $\mathcal{Y}^n$. Let $\mathbf{S}$ be a random variable taking values in A. Then, there exists a subset $S \in \mathsf{A}$ such that

$$\forall Q \in \mathcal{P}_n(\mathcal{Y}), \quad \mathsf{M}_Q(S) \le |\mathcal{P}_n(\mathcal{Y})|\mathsf{E}_{\mathbf{S}}\mathsf{M}_Q(\mathbf{S}).$$

The easy proof of Lemma 2 of Ref. 9) applies directly to this lemma.

5. Applications of Lemma 1 to Classical Coding

Fix a set $B \in \mathcal{Y}^n$, and suppose A is an ensemble of subsets of $\mathcal{Y}^n$ that contain B. If there exists a constant V such that $|\{C \in \mathsf{A} \mid x \in C \setminus B\}| = V$ for any word $x \in \mathcal{Y}^n \setminus B$, the ensemble A is said to be *balanced except* B. An ensemble balanced except $B = \{0^n\}$ is simply called *balanced*. Then, by Lemma 2, we have the next lemma.

Lemma 3 Suppose an ensemble A of $[n,k]$ additive codes over $\mathcal{Y}$ is balanced. Then,

$$\frac{M_Q(C\setminus\{0^n\})}{|\mathcal{T}_Q^n|} \le |\mathcal{P}_n(\mathcal{Y})||\mathcal{Y}|^{-n(1-R)}, \quad Q \in \mathcal{P}_n(\mathcal{Y})$$

for some code $C \in \mathsf{A}$ of rate $R = k/n$.

A generalization of this lemma will appear below as Lemma 6 with a proof. If we apply Corollary 1, putting $d = |\mathcal{Y}|$ and $T = 1 - R$, to the code in Lemma 3, we have the next theorem.

Theorem 1 Suppose an ensemble A of $[n,k]$ additive codes over $\mathcal{Y}$ is balanced. Then, there exists a J-correcting code in A such that

$$P^n(J^{\mathrm{c}}) \le |\mathcal{P}_n(\mathcal{Y})|^3 d^{-nE_{\mathrm{r}}(P,R)}$$

for any $P \in \mathcal{P}(\mathcal{Y})$, where $R = k/n$ and

$$E_{\mathrm{r}}(P,R) = \min_{Q\in\mathcal{P}(\mathcal{Y})}[D(Q||P) + |1 - R - H(Q)|^+].$$

The function E_{r} is known as the random coding exponent for the additive memoryless channel[10] W characterized by P via $W(y|x) = P(y-x)$ [or $W(y|x) = P(x-y)$].

An example of ensembles balanced except an arbitrarily fixed subspace B is the set of all $[n,k]$ linear codes containing B over a finite field $\mathcal{Y}$. To see this, we only need to notice that given any pair of words from $\mathcal{Y}^n \setminus B$, say, x and y, we have a one-to-one linear map on $\mathcal{Y}^n$ that sends x to y.

6. Symplectic Codes

Throughout, H is a Hilbert space of $\dim \mathsf{H} = d$. A scenario of quantum error correction is that provided n primitive quantum systems, each represented by H, are available, a d^k-dimensional subspace of the n-th tensor power $\mathsf{H}^{\otimes n}$ of H, $0 < k < n$, is protected against quantum noise so as to be used for k-quantum-digit computation.

The $2n$-dimensional linear space $\mathbb{F}_d^{2n}$ over $\mathbb{F}_d$ equipped with the standard symplectic form

$$\mathsf{f}_{\mathrm{sp}}((x_1,z_1,\ldots,x_n,z_n),(x_1',z_1',\ldots,x_n',z_n')) = \sum_i x_i z_i' - z_i x_i'$$

plays a crucial role in algebraic QECCs. We can define the dual $L^{\perp_{\mathrm{sp}}}$ of L by $L^{\perp_{\mathrm{sp}}} = \{y \in \mathbb{F}_d^{2n} \mid \forall x \in L, \mathsf{f}_{\mathrm{sp}}(x,y) = 0\}$. Let us call a subspace L with $L^{\perp_{\mathrm{sp}}} \subseteq L$ an f_{sp}-dual-containing code or a *dual-containing code* (with respect to the symplectic form f_{sp}). Then, we have a quantum code whose performance is closely related to that of the classical code L. The code is called a *symplectic (quantum) code* with parity check set $(g_1,\ldots,g_{n-k})$, where $g_1,\ldots,g_{n-k} \in \mathbb{F}_d^{2n}$ form a basis of L, or a symplectic code with stabilizer N_L. Here, $N : u \mapsto N_u$ is Weyl's projective representation[15] of $\mathbb{F}_d^{2n}$ (see Section 8), and $N_J = \{N_y \mid y \in J\}$.

Suppose $\mathsf{A}' = \mathsf{A}'_{n,k}$ is the ensemble of

$[2n, n+k]$ $\mathsf{f}_{\rm sp}$-dual-containing codes over $\mathbb{F}_d$. We can regard them $[n, (n+k)/2]$ additive codes over $\mathcal{Y} = \mathbb{F}_d^2$ if we pair up the coordinates of any word $(x_1, z_1, \ldots, x_n, z_n)$ to have $((x_1, z_1), \ldots, (x_n, z_n)) \in \mathcal{Y}^n$. We can associate with an $[n, (n+k)/2]$ $\mathsf{f}_{\rm sp}$-dual-containing code a set of d^k-dimensional subspaces of $\mathsf{H}^{\otimes n}$, which can be used for quantum error correction[2)~4)]. Namely, we have the next lemma, which is a slight reformulation of the original one[2),3)].

Lemma 4 Suppose a subspace $L \in \mathsf{A}'_{n,k}$ and a set J of representatives of cosets of L in $\mathbb{F}_d^{2n}$ are given. Then, we have a d^k-dimensional subspace of $\mathsf{H}^{\otimes n}$ that works as an $N_{\widetilde{J}}$-correcting code with a suitable recovery operator, where $\widetilde{J} = J + L^{\perp_{\rm sp}} = \{x + y \mid x \in J, y \in L^{\perp_{\rm sp}}\}$.

For a proof, see Ref. 3) or, e.g., Refs. 16), 17). Roughly speaking, given a set of operators $\mathcal{E}$, a quantum code being $\mathcal{E}$-correcting or a code corrects 'errors' in $\mathcal{E}$ means that it recovers any state in the code subspace perfectly after the state suffers 'errors' belonging to $\mathcal{E}$[18)]. The precise definition of $\mathcal{E}$-correcting is not requisite for evaluating the performance of quantum codes. Indeed, the next fact is enough to treat symplectic codes: If we properly define the performance measure, called fidelity, of symplectic codes, it equals the probability $P_{\mathcal{A}}(\widetilde{J})$, where $P_{\mathcal{A}} \in \mathcal{P}(\mathcal{Y}^n)$ is associated with the considered quantum channel (completely positive map) $\mathcal{A}$ in a definite manner[9),17),19)]. It might be said that the structure of quotient codes were inherent in quantum error-correcting codes and some codes used in quantum cryptography (Sections 9 and 10.3).

7. Application of Lemma 1 to Quantum Coding

Noticing A' in the previous section is balanced in the sense of Section 5, we obtain the next theorem either from Theorem 1 or from Lemma 3 and Corollary 1 with $a_n = |\mathcal{P}_n(\mathbb{F}_d^2)|$ and $T = 1 - k/n$.

Theorem 2 For any prime or power of a prime d and any integers n, k with $0 \le k \le n$, there exists a dual-containing $[n, (n+k)/2]$ code C over $\mathbb{F}_d^2$ with respect to the symplectic form $\mathsf{f}_{\rm sp}$ and a set of coset representatives for $\mathbb{F}_d^{2n}/C$ such that for any $P \in \mathcal{P}_n(\mathbb{F}_d^2)$,

$$P^n(J^{\rm c}) \le |\mathcal{P}_n(\mathbb{F}_d^2)|^3 d^{-nE_{\rm q}(P,R)}$$

where $R = k/n$ and

$$E_{\rm q}(P, R) = \min_{Q \in \mathcal{P}(\mathbb{F}_d^2)} [D(Q||P) + |1 - R - H(Q)|^+].$$

Here, we emphasize H and D are defined with logarithms of base d. This recovers the bound on the fidelity of symplectic quantum codes in Refs. 8), 19). In fact, this is better than those in Refs. 8), 19), in that the choice of L in Theorem 2 does not depend on the channel parameter P[☆]. Such a property is referred to as universality in information theory.

8. A Bit of Physics

We have seen in Section 6 that a symplectic quantum code can be characterized by the corresponding $\mathsf{f}_{\rm sp}$-dual-containing code L. Note that instead of specifying a dual-containing code we can specify a self-orthogonal code, i.e., a subspace S with $S \subseteq S^{\perp_{\rm sp}}$. This is because $L^{\perp_{\rm sp}}$ is uniquely determined from L and vice versa due to the property $(L^{\perp_{\rm sp}})^{\perp_{\rm sp}} = L$. As compared with Refs. 8), 19), the roles of L and $L^{\perp_{\rm sp}}$ are interchanged in this article in order to emphasize the next respect. Given a subspace L with $L^{\perp_{\rm sp}} \subseteq L$, the performance of a symplectic quantum code with stabilizer N_L is closely related to that of the classical code L, not $L^{\perp_{\rm sp}}$.

On the other hand, the self-orthogonal subspace $L^{\perp_{\rm sp}}$ has a direct physical meaning. Namely, the symplectic quantum codes corresponding to a dual-containing code L are defined as simultaneous eigenspaces of N_u, $u \in L^{\perp_{\rm sp}}$, where $N : u \mapsto N_u$ is Weyl's projective representation of $\mathbb{F}_d^{2n}$. Specifically, let d be prime, and X and Z be a pair of unitary operators on H of dimension d satisfying

$$XZ = \omega ZX, \tag{4}$$

for a primitive d-th root of unity ω in the field of complex numbers (such as $e^{{\rm i}2\pi/d}$). Let $X^{(x_1,\ldots,x_n)}$ denote $X^{x_1} \otimes \cdots \otimes X^{x_n}$, etc. Then, N is defined by

$$N_{(x_1,z_1,\ldots,x_n,z_n)} = X^{(x_1,\ldots,x_n)} Z^{(z_1,\ldots,z_n)}.$$

What is important is the commutation relation

$$N_u N_v = \omega^{\mathsf{f}_{\rm sp}(u,v)} N_v N_u \tag{5}$$

☆ There was an attempt[20)] to derive a similar bound from a result on classical constant composition codes and the MMI decoding[10),11)]. However, Ref. 20) did not give a quantum operation (completely positive map) to decode the symplectic code, whereas we have such an operation that generalizes a (wide-sense) syndrome decoding for classical codes[17),19)].

which follows from Eq. (4). Observe that N_u and N_v commute if and only if $\mathsf{f}_{\rm sp}(u,v)=0$. Hence, specifying a set of commuting operators in $\{N_u \mid u \in \mathbb{F}_d^{2n}\}$ is equivalent to specifying a self-orthogonal subspace of $\mathbb{F}_d^{2n}$. From the engineering point of view, the role of $N_{L^{\perp_{\rm sp}}}$ (more precisely, $N_{g_1},\ldots,N_{g_{n-k}}$ for a basis $(g_1,\ldots,g_{n-k})$ of $L^{\perp_{\rm sp}}$) is a syndrome measurement. More details may be found in Refs. 2), 3), or Appendix A of Ref. 17).

9. Calderbank-Shor-Steane Codes

9.1 Symmetric CSS Codes

We have explained the scenario of using symplectic codes for protection of quantum states. It is known that a class of symplectic codes are also useful for quantum key distribution (QKD). In particular, Shor and Preskill [21] argued that the security of the famous Bennett-Brassard 1984 (BB84) QKD protocol could be proved by evaluating the fidelity of quantum error-correcting codes underlying the protocol. The codes are called Calderbank-Shor-Steane (CSS) codes.

First, we consider a class of CSS codes of simple structure. Given a classical code $C \subseteq \mathbb{F}_d^n$ with $C^\perp \subseteq C$, where $C^\perp$ is the dual of C with respect to the bilinear form $\sum_i x_i y_i$, a simultaneous eigenspace of the commuting operators $X^x Z^z$, $x,z \in C^\perp$, is called a CSS code. Given a set of coset representatives Γ for $\mathbb{F}_d^n/C$, with a suitable decoding (recovery) operator, the CSS code can correct errors $X^x Z^z$ for $x \in \Gamma' = \Gamma + C^\perp$ and $z \in \Gamma'$, so that the 'decoding error probability' (one minus fidelity) of the quantum code is upper-bounded by ☆

$$1-\Pr[\mathbf{X}^n \in \Gamma' \text{ and } \mathbf{Z}^n \in \Gamma'] \le \Pr[\mathbf{X}^n \in \Gamma'^{\rm c}] + \Pr[\mathbf{Z}^n \in \Gamma'^{\rm c}]. \quad (6)$$

9.2 Applications of Lemma 1 to CSS Quantum Coding

We can show that the ensemble of dual-containing codes with respect to the bilinear form $\sum_i x_i y_i$ has a good balance, though it may not be completely balanced in the sense of Section 5, which leads to the next lemma and theorem [9]. The case of $d \ge 3$ is more tractable [9].

Lemma 5 Assume $d=2$ and $n \ge 2$ is even. Then, for any $\overline{\kappa}$ with $n/2 \le \overline{\kappa} \le n$, there exists a $\overline{\kappa}$-dimensional subspace C of $\mathbb{F}_d^n$ such that $\{1^n\} \subseteq C^\perp \subseteq C$ and for any $Q \in \mathcal{P}_n(\mathbb{F}_d)$,

$$\frac{M_Q(C \setminus \{0^n,1^n\})}{|\mathcal{T}_Q^n|} \le |\mathcal{P}_n(\mathbb{F}_d)| d^{\overline{\kappa}-n+1}.$$

Theorem 3 Assume $d=2$. Let a number $0 \le R \le 1$ be given. There exists a sequence of pairs (C_n,Γ_n), $n=2,4,6,\ldots$, each consisting of a subspace $C_n \subseteq \mathbb{F}_d^n$ with $\{1^n\} \subseteq C_n^\perp \subseteq C_n$ and $2\dim_{\mathbb{F}_d} C_n - n \ge nR$, and a set of coset representatives Γ_n of $\mathbb{F}_d^n/C_n$, such that for any probability distribution P on $\mathbb{F}_d$,

$$P(\widetilde{\Gamma}_n^{\rm c}) \le 4(n+1)^3 d^{-nE_{\rm r}((1+R)/2,P)}$$

where $\widetilde{\Gamma}_n = \mathsf{span}\, 1^n + \Gamma_n = \Gamma_n \cup \{1^n + x \mid x \in \Gamma_n\}$, and $E_{\rm r}$ is defined in Theorem 1.

Proof. Put $B=\{0^n,1^n\}$, $\mathcal{Y}=\mathbb{F}_d$, $a_n = d^2|\mathcal{P}_n(\mathcal{Y})|$ and $T=(1-R)/2$. Then, Theorem 3 follows from Corollary 1 applied to the codes in Lemma 5.

This lemma and Eq. (6) ensure the existence of a CSS code whose 'decoding error probability' is upper-bounded by $8(n+1)^3 d^{-nE(R,P_{\mathbf{X}},P_{\mathbf{Z}})}$ where

$$E(R,P_{\mathbf{X}},P_{\mathbf{Z}}) = \min\{E_{\rm r}((1+R)/2,P_{\mathbf{X}}), E_{\rm r}((1+R)/2,P_{\mathbf{Z}})\},$$

and $P_{\mathbf{XZ}} = P_{\mathcal{A}_1}$ is associated with the considered memoryless quantum channel $\mathcal{A}_1$ in the manner described in Refs. 9), 17), 19). The resulting achievable rate is not so large as $1-H(P_{\mathbf{XZ}})$ obtained from Theorem 2.

10. Other Applications

10.1 Ensemble of Quotient Codes

The ensemble average of $\mathsf{M}_Q(C \setminus B)$ over those (B,C) with $B=C^\perp$ and $B \le C$ was evaluated to prove the existence of codes of 'balanced' spectra in Lemma 5. We treat other ensembles in this section.

Lemma 6 Let $\mathcal{Y}$ be a finite set and integers k, b with $0<k\le n-b$ be given. Suppose A is a family of pairs of subsets (B,C) of $\mathcal{Y}^n$ with $B' \subseteq B \subseteq C \subseteq \mathcal{Y}^n$, $|B|=|\mathcal{Y}|^b$, $|C|=|\mathcal{Y}|^{k+b}$ for some $B' \subseteq \mathcal{Y}^n$ such that the number $V = |\{(B,C) \in \mathsf{A} \mid x \in C \setminus B\}|$ does not depend on $x \in \mathcal{Y}^n \setminus B'$. Then, there exists a pair $(B,C) \in \mathsf{A}$ such that for any $Q \in \mathcal{P}_n(\mathcal{Y})$,

$$\frac{M_Q(C \setminus B)}{|\mathcal{T}_Q^n|} \le |\mathcal{P}_n(\mathcal{Y})| |\mathcal{Y}|^{k+b-n}.$$

Corollary 2 Let a b-dimensional subspace B of $\mathcal{Y}^n$, where $\mathcal{Y}$ is any finite field, and an integer k, $0<k\le n-b$ be given. Let A' be the family of all $(k+b)$-dimensional subspaces

☆ The $\mathbf{X}^n$ stands for $(\mathbf{X}_1,\ldots,\mathbf{X}_n)$, etc., and the probability distribution of $(\mathbf{X}_1,\mathbf{Z}_1,\ldots,\mathbf{X}_n,\mathbf{Z}_n)$, is $P_{\mathcal{A}}$ mentioned in Section 6.

C with $B \le C \le \mathcal{Y}^n$. Then, the bound on $M_Q(C \setminus B)$ in the lemma is fulfilled by some subspace $C \in \mathsf{A}'$.

Proof. Put $d = |\mathcal{Y}|$. Counting the pairs $(x, (B, C))$ such that $x \in C \setminus B$ and $(B, C) \in \mathsf{A}$ in two ways, we have $(d^n - d^b)V \le (d^n - |B'|)V \le |\mathsf{A}|(d^{k+b} - d^b)$, from which $\mathsf{E}_{\mathbf{B},\mathbf{C}} M_Q(\mathbf{C} \setminus \mathbf{B}) \le |\mathcal{T}_Q^n| d^{k+b-n}$ follows, where $(\mathbf{B}, \mathbf{C})$ is a random variable uniformly distributed over A, and hence, the desired estimate follows by Lemma 2. The corollary follows from the last remark in Section 5.

Theorem 4 Let an additive group $\mathcal{Y}$ of order d and integers k, b with $0 < k \le n - b$ be given. Suppose we have a quotient code C/B, i.e., a pair (B, C) with $B \le C \le \mathcal{Y}^n$, $|B| = d^b$, and $|C| = d^{k+b}$ such that

$$\frac{M_Q(C \setminus B)}{|\mathcal{T}_Q^n|} \le a_n d^{k-n}, \quad Q \in \mathcal{P}_n(\mathcal{Y}).$$

Then, if $a_n \ge 1$, its decoding error probability with the minimum entropy decoding on additive memoryless channel P is upper-bounded by

$$a_n |\mathcal{P}_n(\mathcal{Y})|^2 d^{-nE_{\mathrm{r}}(P,R)}$$

where $R = k/n$ and $E_{\mathrm{r}}(P, R)$ is defined in Theorem 1. The same code C/B have its decoding error probability upper-bounded by

$$|\mathcal{P}_n(\mathcal{Y})| d^{b - nE_{\mathrm{ex}}(P, R + [\log_d a_n]/n)}$$

with the MLD, where $R = k/n$,

$$E_{\mathrm{ex}}(P, R) = \min_{Q \in \mathcal{P}(\mathcal{Y}):\ H(Q) \ge 1 - R} \Big[1 - R - H(Q) - \sum_{s \in \mathcal{Y}} Q(s) \log \sum_{a \in \mathcal{Y}} \sqrt{P(a)P(a+s)} \Big].$$

Recall that the decoding error probability is given by $P^n(\widetilde{J}^{\mathrm{c}})$ as in Section 3. We remark that the code in Corollary 2 satisfies the premise, and hence, both bounds in the theorem with $a_n = d^b |\mathcal{P}_n(\mathcal{Y})|$ or $a_n = d^b (n+1)^{d-1}$.

10.2 Implication on Classical Coding

With b small, say, $b = 0, 1$, the combination of the two bounds in Theorem 4 gives the asymptotically best bound among those known for additive memoryless channels (for the rates below the critical rate, the random coding exponent $E_{\mathrm{r}}(P, R)$ is optimum, and the expurgated exponent $E_{\mathrm{ex}}(P, R)$ improves this for high rates). Note that whereas $E_{\mathrm{r}}(P, R)$ is attainable by universal systems of encoding and decoding, we have used the MLD in Theorem 4 to prove the attainability of $E_{\mathrm{ex}}(P, R)$, and this is true for any derivations of $E_{\mathrm{ex}}(P, R)$ known (to the present author). A natural question arises (Cf. Csiszár and Körner [14]): For which class of channels can we find a good code and its decoding that attain the expurgated exponent and do not depend on the channel characteristics?

Theorem 4, together with Corollary 2 or Lemma 5, ensures the existence of such a system of encoding and decoding that works for all binary symmetric channels. In fact, for $\mathcal{Y} = \{0, 1\}$ and $B = \{0^n, 1^n\}$, the minimum entropy decoding and the MLD (together with the minimum Hamming distance decoding) described in Section 3 happen to be the same.

10.3 Asymmetric CSS Codes

A general d-ary CSS code is a simultaneous commuting operators $X^x Z^z$, $x \in C_2$, $z \in C_1^\perp$ for some $C_1, C_2 \subseteq \mathbb{F}_d^n$. For d prime, these operators really commute if (and only if) $C_2 \subseteq C_1$ as can be easily checked with Eq. (5). Then, by the general principle of symplectic codes in Lemma 4, it is easy to see that if the quotient codes C_1/C_2 and $C_2^\perp/C_1^\perp$ are both good, then the CSS codes are good. This was stated by Mayers [22], where the goodness of a quotient code was measured by minimal distance ☆. In the framework of quotient codes, this can be said more clearly: If C_1/C_2 is $\widetilde{\Gamma}_1$-correcting and $C_2^\perp/C_1^\perp$ is $\widetilde{\Gamma}_2$-correcting, then the CSS code is $N_{\widetilde{J}}$-correcting with

$$\widetilde{J} = \{[x, z] \mid x \in \widetilde{\Gamma}_2 \text{ and } z \in \widetilde{\Gamma}_1\}$$

where

$$[(x_1, \ldots, x_n), (z_1, \ldots, z_n)] = (x_1, z_1, \ldots, x_n, z_n).$$

In applications of CSS codes to quantum cryptography, either the quotient code C_1/C_2 or $C_2^\perp/C_1^\perp$ should be efficiently decodable [21),22)]. Then, it was proposed to use known classical codes as C_1 and a randomly chosen subspace of C_1 as C_2 [22]. If we knew which specific subspace were good as C_2, we would use it rather than random ones as C_2. Moreover, the deterministic choice of C_2 saves randomness and public communication.

Then, a natural question arises: After a (good) code C_1 over a finite field $\mathcal{Y}$ is arbitrarily fixed, how good $C_2^\perp/C_1^\perp$ can be by a proper

☆ Related metrics for quotient spaces are proposed in Appendix A.3.

choice of $C_2 \leq C_1$? Corollaries 1 and 2 answer this question. Namely, putting $B = C_1^{\perp}$, $C = C_2^{\perp}$, $b = (1 - R_1)n$, and $k = (R_1 - R_2)n$, we conclude from these corollaries the existence of a $\widetilde{\Gamma}_2$-correcting quotient code $C_2^{\perp}/C_1^{\perp}$ with $P^n(\widetilde{\Gamma}_2) \leq |\mathcal{P}_n(\mathcal{Y})|^3 d^{-nE_{\mathrm{r}}(1-R_2)}$ for any $P \in \mathcal{P}(\mathcal{Y})$.

If we use the random selection of C_2 instead as in Refs. 22), 23), the resulting bound is $\mathsf{E}_{C_2} P^n(\widetilde{\Gamma}_2) \leq |\mathcal{P}_n(\mathcal{Y})|^2 d^{-nE_{\mathrm{r}}(1-R_2)}$ by Corollary 3 in the next subsection. Observe the bound is better by the factor of $|\mathcal{P}_n(\mathcal{Y})|$ because the process of choosing a deterministic good code (Lemma 2), where the factor stems from, is not necessary in this case.

Note the argument in this section is applicable to the case where $\mathcal{Y} = \mathbb{F}_d^m$ for some $m \geq 1$ (we define the duals of codes as before viewing $\mathcal{Y}^n$ as $\mathbb{F}_d^{nm}$), which recovers the result of Appendix B of Ref. 9), though the decoding complexity is sacrificed slightly for the generality of the argument.

10.4 Random Codes

Expecting possible applications in the future, this subsection gives general forms of Lemma 1 and Corollary 1, which are applicable to random codes. The proof of Lemma 1 in the appendix and that of Corollary 1 can be accommodated to this case only by applying $\mathsf{E}_{\mathbf{B},\mathbf{C}}$ where appropriate.

Lemma 7 Assume a pair of random variables $(\mathbf{B}, \mathbf{C})$ that take values in a set $\{(B, C) \mid B \leq C \leq \mathcal{Y}^n\}$ satisfy

$$\mathsf{E}_{\mathbf{B},\mathbf{C}} \frac{M_Q(\mathbf{C} \setminus \mathbf{B})}{|\mathcal{T}_Q^n|} \leq a_n d^{-nT}, \quad Q \in \mathcal{P}_n(\mathcal{Y})$$

with some real numbers $d > 1$, $a_n \geq 1$ and T. Then, choosing J_C as in item (1) of the list in Section 2 (minimum entropy decoding), we have for any $P_n \in \mathcal{P}(\mathcal{Y}^n)$,

$$\begin{aligned}&\mathsf{E}_{\mathbf{B},\mathbf{C}} \mathsf{E}_{\boldsymbol{\pi}} P_n(\boldsymbol{\pi}(\widetilde{J}_{\mathbf{B},\mathbf{C}})^{\mathrm{c}}) \\ &\quad \leq a_n |\mathcal{P}_n(\mathcal{Y})| \sum_{Q \in \mathcal{P}_n(\mathcal{Y})} P_n(\mathcal{T}_Q^n) d^{-n|T - H(Q)|^+}\end{aligned}$$

where $\widetilde{J}_{B,C} = J_C + B = \{x + y \mid x \in J_C, y \in B\}$.

Corollary 3 Assume a pair of random variables $(\mathbf{B}, \mathbf{C})$ fulfills the condition of Lemma 7. Then, for the above choice of J_C and for any $P \in \mathcal{P}(\mathcal{Y})$, we have

$$\mathsf{E}_{\mathbf{B},\mathbf{C}} P^n(\widetilde{J}_{\mathbf{B},\mathbf{C}^{\mathrm{c}}}) \leq a_n |\mathcal{P}_n(\mathcal{Y})|^2 d^{-nE^{(n)}(P,T)}$$

where $E^{(n)}(P, T)$ is given in Corollary 1.

11. Summary and Remarks

This article gave a formula to evaluate the performance of an algebraic code in terms of its (weight) spectrum.

The achievable rate $1 - H(P)$ for quantum channels resulting from Theorem 2 is not the best [24),25)]. The optimum rate achievable by general quantum codes is now known (see, e.g., Ref. 26) and the references therein, especially, Refs. 27), 28) for details). The problem of determining the optimum attainable exponent is left open [17),29)].

Acknowledgments Useful discussions with Masahito Hayashi are acknowledged. A part of this work was performed while the author was with the ERATO Quantum Computation and Information Project, JST, Japan. He is grateful to the project director Hiroshi Imai and the staff for their continuing support.

References

1) Shor, P.W.: Scheme for reducing decoherence in quantum computer memory, *Phys. Rev. A*, Vol.52, pp.R2493–2496 (1995).
2) Calderbank, A.R., Rains, E.M., Shor, P.W. and Sloane, N.J.A.: Quantum error correction and orthogonal geometry, *Phys. Rev. Lett.*, Vol.78, pp.405–408 (Jan. 1997).
3) Calderbank, A.R., Rains, E.M., Shor, P.W. and Sloane, N.J.A.: Quantum error correction via codes over GF(4), *IEEE Trans. Inf. Theory*, Vol.44, pp.1369–1387 (July 1998).
4) Gottesman, D.: Class of quantum error-correcting codes saturating the quantum Hamming bound, *Phys.Rev.A*, Vol.54, pp.1862–1868 (Sep. 1996).
5) Gallager, R.G.: *Low-density parity-check codes*, Cambridge, MA, MIT Press (1963).
6) Goppa, V.D.: Universal coding for symmetric channels, *Problems of Information Transmission*, Vol.11, pp.11–17 (Jan.–Mar. 1975).
7) Csiszár, I.: Linear codes for sources and source networks: Error exponents, universal coding, *IEEE Trans. Inf. Theory*, Vol.IT-28, pp.585–592 (July 1982).
8) Hamada, M.: Exponential lower bound on the highest fidelity achievable by quantum error-correcting codes, *Phys. Rev. A*, Vol.65, pp.052305–1–4 (Apr. 2002).
9) Hamada, M.: Reliability of Calderbank-Shor-Steane codes and security of quantum key distribution, *J. Phys. A: Math. Gen.*, Vol.37, pp.8303–8328 (2004).
10) Csiszár, I. and Körner, J.: *Information Theory: Coding Theorems for Discrete Memoryless*

Systems, NY, Academic (1981).
11) Csiszár, I.: The method of types, *IEEE Trans. Inf. Theory*, Vol.IT-44, pp.2505–2523 (Oct. 1998).
12) MacWilliams, F.J. and Sloane, N.J.A.: *The Theory of Error-Correcting Codes*, NY, North-Holland (1977).
13) Goppa, V.D.: Nonprobabilistic mutual information without memory, *Problems of Control and Information Theory*, Vol.4, No.2, pp.97–102 (English translation: pp.1–6) (1975).
14) Csiszár, I. and Körner, J.: Graph decomposition: A new key to coding theorems, *IEEE Trans. Inf. Theory*, Vol.IT-27, pp.5–12 (Jan. 1981).
15) Weyl, H.: *Gruppentheorie und Quantenmechanik*, Leipzig, Verlag von S. Hirzel in Leipzig (1928). English translation of the second ed. (1931): *The Theory of Groups and Quantum Mechanics*, Dover (1950).
16) Ashikhmin, A. and Knill, E.: Nonbinary quantum stabilizer codes, *IEEE Trans. Inf. Theory*, Vol.47, pp.3065–3072 (Nov. 2001).
17) Hamada, M.: Notes on the fidelity of symplectic quantum error-correcting codes, *International Journal of Quantum Information*, Vol.1, No.4, pp.443–463 (2003).
18) Knill, E. and Laflamme, R.: Theory of quantum error-correcting codes, *Phys. Rev. A*, Vol.55, pp.900–911 (Feb. 1997).
19) Hamada, M.: Lower bounds on the quantum capacity and highest error exponent of general memoryless channels, *IEEE Trans. Inf. Theory*, Vol.48, pp.2547–2557 (Sep. 2002).
20) Barg, A.: A low-rate bound on the reliability of a quantum discrete memoryless channel, *IEEE Trans. Inf. Theory*, Vol.48, pp.3096–3100 (2002).
21) Shor, P. and Preskill, J.: Simple proof of security of the BB84 quantum key distribution protocol, *Phys. Rev. Lett.*, Vol.85, pp.441–444 (July 2000).
22) Mayers, D.: Shor and Preskill's and Mayers's security proof for the BB84 quantum key distribution protocol, *The European Physical Journal D*, Vol.18, pp.161–170 (2002).
23) Watanabe, S., Matsumoto, R. and Uyematsu, T.: Noise tolerance of the BB84 protocol with random privacy amplification, e-Print quant-ph/0412070, LANL (2004).
24) Shor, P.W. and Smolin, J.A.: Quantum error-correcting codes need not completely reveal the error syndrome, e-Print quant-ph/9604006, LANL (1996).
25) Hamada, M.: Information rates achievable with algebraic codes on quantum discrete memoryless channels, to appear in *IEEE Trans. Inf. Theory*, e-Print, quant-ph/0207113, LANL (2002).
26) Devetak, I. and Winter, A.: Distillation of secret key and entanglement from quantum states, e-Print qunat-ph/0306078, LANL (2003).
27) Barnum, H., Knill, E. and Nielsen, M.A.: On quantum fidelities and channel capacities, *IEEE Trans. Inf. Theory*, Vol.46, pp.1317–1329 (July 2000).
28) Horodecki, M., Horodecki, P. and Horodecki, R.: Unified approach to quantum capacities: Towards quantum noisy coding theorem, *Phys. Rev. Lett.*, Vol.85, pp.433–436 (July 2000).
29) Kretschmann, D. and Werner, R.F.: Tema con variazioni: quantum channel capacity, *New J. Phys.*, Vol.6, No.26, pp.1–33 (2004).
30) Gabidulin, E.M.: Combinatorial metrics in coding theory, *2nd International Symposium on Information Theory* (Armenia, USSR, Sep. 1971), Petrov, B.N. and Csaki, F. (Eds.), Budapest, Akademiai Kiado (1973).
31) Hamada, M.: A note on combinatorial metrics for error-correcting codes, *IEICE Technical Report*, Vol.IT99-31, pp.97–102 (July 1999).

(Received February 9, 2005)
(Accepted July 4, 2005)
(Online version of this article can be found in the IPSJ Digital Courier, Vol.1, pp.450–460.)

Appendix

A.1 Proof of Lemma 1

In the proof, $\mathcal{P}_n(\mathcal{Y})$ is abbreviated as $\mathcal{P}_n$. We will show that $G = \mathsf{E}_{\boldsymbol{\pi}} P_n(\boldsymbol{\pi}(\widetilde{J})^{\mathrm{c}})$ is bounded above by the claimed quantity.

Imagine we list up all words in $\pi(C \setminus B)$ for all $\pi \in \mathcal{S}_n$ allowing duplication. Clearly, the number of appearances of any fixed word $y \in \mathcal{Y}^n$ in the list only depends on its type $\mathsf{P}_y \in \mathcal{P}_n$. Namely, for any $Q \in \mathcal{P}_n$, there exists a constant, say L_Q, such that $|\{\pi \in \mathcal{S}_n \mid y \in \pi(C \setminus B)\}| = L_Q$ for any word y with $\mathsf{P}_y = Q$. Then, counting the number of words of a fixed type Q in the list in two ways, we have $|\mathcal{T}_Q^n| L_Q = |\mathcal{S}_n| M_Q(C \setminus B)$. Hence, for any type $Q \in \mathcal{P}_n(\mathcal{Y})$

$$\frac{L_Q}{|\mathcal{S}_n|} = \frac{M_Q(C \setminus B)}{|\mathcal{T}_Q^n|} \leq a_n d^{-nT}$$

by assumption. This implies that for $y \in \mathcal{Y}^n$, we have

$$\frac{|\mathsf{A}_y(C \setminus B)|}{|\mathcal{S}_n|} \leq a_n d^{-nT} \tag{7}$$

where

$$\mathsf{A}_y(C \setminus B) = \{\pi \in \mathcal{S}_n \mid y \in \pi(C \setminus B)\}.$$

We have

$$G = \frac{1}{|\mathcal{S}_n|} \sum_{\pi \in \mathcal{S}_n} \sum_{x \notin \pi(\widetilde{J})} P_n(x)$$
$$= \sum_{x \in \mathcal{Y}^n} P_n(x) \frac{|\{\pi \in \mathcal{S}_n \mid x \notin \pi(\widetilde{J})\}|}{|\mathcal{S}_n|}. \quad (8)$$

Since $x \notin \pi(\widetilde{J})$ occurs only if there exists a word $u \in \mathcal{Y}^n$ such that $H(\mathsf{P}_u) \le H(\mathsf{P}_x)$ and $u - x \in \pi(C \setminus B)$ from the design of $\widetilde{J}$ specified above (minimum entropy decoding), it follows

$$|\{\pi \in \mathcal{S}_n \mid x \notin \pi(\widetilde{J})\}|/|\mathcal{S}_n|$$
$$\le \sum_{u \in \mathcal{Y}^n : H(\mathsf{P}_u) \le H(\mathsf{P}_x)} |\mathsf{A}_{u-x}(C \setminus B)|/|\mathcal{S}_n|$$
$$\le \sum_{u \in \mathcal{Y}^n : H(\mathsf{P}_u) \le H(\mathsf{P}_x)} a_n d^{-nT}$$
$$= \sum_{Q' \in \mathcal{P}_n : H(Q') \le H(\mathsf{P}_x)} a_n |\mathcal{T}^n_{Q'}| d^{-nT}$$
$$\le \sum_{Q' \in \mathcal{P}_n : H(Q') \le H(\mathsf{P}_x)} a_n d^{nH(Q') - nT} \quad (9)$$

where we have used Eq. (7) for the second inequality, and Eq. (1) for the last inequality. Then, using the inequalities $\min\{at, 1\} \le a \min\{t, 1\}$ and $\min\{s + t, 1\} \le \min\{s, 1\} + \min\{t, 1\}$ for $a \ge 1, s, t \ge 0$, we can proceed from Eq. (8) as follows, which completes the proof:

$$G \le \sum_{x \in \mathcal{Y}^n} P_n(x) \min\Big\{ \sum_{Q' \in \mathcal{P}_n : H(Q') \le H(\mathsf{P}_x)} a_n d^{nH(Q') - nT},\ 1 \Big\}$$
$$\le a_n \sum_{Q \in \mathcal{P}_n} P_n(\mathcal{T}^n_Q) \min\Big\{ \sum_{Q' \in \mathcal{P}_n : H(Q') \le H(Q)} d^{nH(Q') - nT}, 1 \Big\}$$
$$\le a_n \sum_{Q \in \mathcal{P}_n} P_n(\mathcal{T}^n_Q) \sum_{Q' \in \mathcal{P}_n : H(Q') \le H(Q)} \min\big\{ d^{-n[T - H(Q')]},\ 1 \big\}$$
$$\le a_n |\mathcal{P}_n| \sum_{Q \in \mathcal{P}_n} P_n(\mathcal{T}^n_Q) \max_{Q' \in \mathcal{P}(\mathbb{F}_d) : H(Q') \le H(Q)} d^{-n|T - H(Q')|^+}$$
$$= a_n |\mathcal{P}_n| \sum_{Q \in \mathcal{P}_n} P_n(\mathcal{T}^n_Q) d^{-n|T - H(Q)|^+}.$$

A.2 Proof of Theorem 4

The first bound with the random coding exponent in theorem immediately follows from Corollary 1. We proceed to proving the second.

By the choice of J (MLD),

$$P^n(\widetilde{J}^{\mathrm{c}}) \le \sum_{\widetilde{x} \in \mathcal{Y}^n / B} |B| \sum_{\widetilde{y} \in C/B \setminus B} \sqrt{\max_{x \in \widetilde{x}} P^n(x) \max_{y \in \widetilde{y}} P^n(x+y)}$$

This can be seen by noticing that when the additive error falls in a coset $\widetilde{x} \in \mathcal{Y}^n/B$, the decoding error occurs only if there is a coset $\widetilde{z} \in \mathcal{Y}^n/B$ with $\max_{z \in \widetilde{z}} P^n(z) \ge \max_{x \in \widetilde{x}} P^n(x)$ and $\widetilde{z} - \widetilde{x} \in C/B \setminus B$. Hence, we have

$$P^n(\widetilde{J}^{\mathrm{c}})$$
$$\le |B| \sum_{\widetilde{y} \in C/B \setminus B} \sum_{\widetilde{x} \in \mathcal{Y}^n/B} \sum_{y \in \widetilde{y}} \sum_{x \in \widetilde{x}} \sqrt{P^n(x) P^n(x+y)}$$
$$\le |B| \sum_{y \in C \setminus B} \sum_{x \in \mathcal{Y}^n} \sqrt{P^n(x) P^n(x+y)}$$
$$= |B| \sum_{y \in C \setminus B} d^{n \sum_{s \in \mathcal{Y}} \mathsf{P}_y(s) \log \sum_{a \in \mathcal{Y}} \sqrt{P(a)P(a+s)}}$$
$$= |B| \sum_{Q \in \mathcal{P}_n(\mathcal{Y})} M_Q(C \setminus B) d^{n \sum_{s \in \mathcal{Y}} Q(s) \log \sum_{a \in \mathcal{Y}} \sqrt{P(a)P(a+s)}}.$$

Note that the premise in the theorem implies

$$M_Q(C \setminus B) \le d^{-n[1 - R - H(Q) - (\log a_n)/n]}, \quad Q \in \mathcal{P}_n(\mathcal{Y}),$$

by Eq. (1), and hence, $M_Q(C \setminus B) = 0$ if $1 - R - H(Q) - (\log a_n)/n > 0$ (incidentally, this implies that the code satisfies the Gilbert-Varshamov bound asymptotically). Thus, we have the second bound in the theorem.

A.3 Metrics for Quotient Spaces

We have treated spaces of the form $\mathcal{V} = \mathcal{Z}/B$, where $B \le \mathcal{Z}$ are finite additive groups. In this appendix, a natural way to derive metrics on $\mathcal{V}$ from those on $\mathcal{Z}$, such as the Hamming metric on $\mathcal{Z} = \mathcal{Y}^n$, is introduced. Given a nonnegative function W on $\mathcal{Z}$, a function D on $\mathcal{Z} \times \mathcal{Z}$ defined by $\mathsf{D}(x, y) = \mathsf{W}(y - x)$ is a metric if W satisfies (i) triangle inequality $\mathsf{W}(x + y) \ge \mathsf{W}(x) + \mathsf{W}(y)$, $x, y \in \mathcal{Z}$, (ii) $\mathsf{W}(x) = 0$ if and only if x is zero, and (iii) $\mathsf{W}(x) = \mathsf{W}(-x)$.

Lemma 8 Given a function W on $\mathcal{Z}$, define $\mathsf{W}_B(\widetilde{x}) = \min_{x \in \widetilde{x}} \mathsf{W}_B(x)$ for $\widetilde{x} \in \mathcal{Z}/B$. Then, whichever of properties (i), (ii) and (iii) W has, W_B inherits the same properties from W.

The easy proof is omitted. A broad class of metrics that include metrics of this type are known as combinatorial metrics [30] or combinatorial metrics with cost [31].