JOURNAL OF MATHEMATICAL PHYSICS VOLUME 43, NUMBER 9 SEPTEMBER 2002

Lower bound for the quantum capacity of a discrete memoryless quantum channel

Ryutaroh Matsumoto[a] and Tomohiko Uyematsu[b]
Department of Communications and Integrated Systems, Tokyo Institute of Technology, 152-8552 Japan

(Received 24 October 2001; accepted for publication 16 May 2002)

We generalize the random coding argument of stabilizer codes and derive a lower bound on the quantum capacity of an arbitrary discrete memoryless quantum channel. For the depolarizing channel, our lower bound coincides with that obtained by Bennett *et al.* We also slightly improve the quantum Gilbert–Varshamov bound for general stabilizer codes, and establish an analog of the quantum Gilbert–Varshamov bound for linear stabilizer codes. Our proof is restricted to the binary quantum channels, but its extension of to l-adic channels is straight forward. © *2002 American Institute of Physics.* [DOI: 10.1063/1.1497999]

I. INTRODUCTION

The quantum capacity of a quantum channel is the amount of quantum states that can be reliably transmitted through the channel. It is one of the fundamental unsolved problems in the quantum information theory. Except for the quantum erasure channel, we know only lower and upper bounds for the quantum capacity of a quantum channel, and, in addition, a tight lower bound is not known for a general memoryless quantum channel. In this article we shall demonstrate a lower bound on the capacity of a general memoryless quantum channel. A quantum channel is said to be memoryless if the state change of one transmitted quantum system (of the fixed degree of freedom) is statistically independent of the state change of another.

The problem of quantum capacity has attracted great attention, and rapid progress has been made. To be precise, the quantum capacity of a binary memoryless channel Γ is the maximum number $Q(\Gamma)$ such that for any rate $R<Q(\Gamma)$ and any $\epsilon>0$ there exists an $[[n,k]]$ quantum code Q with $k/n \geqslant R$ such that the fidelity between the recovered state and the original state $|\varphi\rangle \in Q$ is at least $1-\epsilon$ for any $|\varphi\rangle$.[1,2] In Refs. 1 and 2, the authors obtained the exact capacity of the quantum erasure channel, and showed lower and upper bounds for that of the quantum depolarizing channel. The same lower bounds for those channels were also obtained in Ref. 3 by using random coding of the stabilizer codes introduced in Refs. 4–6. After that, DiVincenzo *et al.*[7] improved the lower bound for a depolarizing channel by using nonrandom stabilizer codes. The upper bound of the depolarizing channel was improved in Refs. 8–10, and generalized to asymmetric depolarizing channels in Ref. 11. An apparently different definition of the quantum capacity was formalized in Ref. 12, in which an upper bound of a general memoryless quantum channel was established by using the notion of coherent information introduced in Ref. 13. It is informally argued in Ref. 14 that the upper bound in Ref. 12 is achieved by random coding over a general memoryless channel. Barnum *et al.*[15] showed that the definitions of quantum capacity in Refs. 1, 2, and 12 were equivalent.

It is the random coding that is the most commonly used technique in classical information theory to show that a specific rate is achieved by a code in a specific class of codes. For example, Elias showed that the capacity of the binary symmetric channel is achieved by binary linear codes

[a]Electronic address: ryutaroh@rmatsumoto.org; URL: http: //www.rmatsumoto.org
[b]Electronic address: uyematsu@ieee.org; URL: http: //www.it.ss.titech.ac.jp

using random coding[16] (a readable proof of this fact can be found in Sec. 6.2 of Ref. 17). A proof by random coding usually proceeds as follows: one first calculates the average of error probability of all codes in a specific class of codes of the same rate and the same code length, then shows that the average converges to 0 as the code length increases, and finally concludes that there exists at least one sequence of codes of the fixed rate with which the error probability converges to 0.

The technique of random coding is also used in the quantum information theory. Gottesman[3] showed using random coding that the lower bound on the quantum capacity of the depolarizing channel[2] can be achieved by stabilizer codes. However, his proof does not seem to extend easily to the case of general memoryless channel. The aim of the present article is to derive a lower bound on the quantum capacity of a general memoryless channel by using random coding of stabilizer codes. In our argument of random coding, we shall use the fidelity[18,19] as a replacement of error probability in the classical random coding, and use the idea behind the proof of the quantum Gilbert–Varshamov bound for the stabilizer codes.[4] As a byproduct, we also improve the quantum Gilbert–Varshamov bound for stabilizer codes. Our improved bound (Remark 10) is slightly better than the quantum Gilbert–Varshamov bound for general codes.[20]

As a natural consequence of the quantum Gilbert–Varshamov bound[4,20] and the fidelity bound of t-error correcting quantum codes,[21–23] we can also derive lower bounds for the quantum capacity of a general memoryless quantum channel. However, for the depolarizing channel, the derived lower bounds are much smaller than that obtained in Ref. 2. In contrast to this, our lower bound coincides with the bound in Ref. 2 for the depolarizing channel.

It is interesting whether the proposed lower bound is achieved by a subclass of general stabilizer codes. We also show that the random coding of linear stabilizer codes yields the same lower bound on the quantum capacity. As a byproduct we obtain an analog of the quantum Gilbert–Varshamov bound for linear stabilizer codes (Remark 13), which is asymptotically the same as that for general quantum codes.[20]

The quantum channel considered in this article is discrete in the sense that the channel carries finite-dimensional quantum systems, and we do not touch the quantum capacity of a continuous quantum channel recently studied in Refs. 24 and 25. Our proof is restricted to the binary quantum channels for the simplicity of presentation, but its extension to l-adic channels is straightforward for prime l.

This article is organized as follows: In Sec. II we introduce notations and review relevant research results. In Sec. III we derive a lower bound [Eq. (16)] for the quantum capacity of an arbitrary discrete memoryless quantum channel by random coding of stabilizer codes.

II. NOTATIONS AND PRELIMINARIES

In this section we fix notations used in this article, and review known research results that are necessary to establish our results.

A. Quantum channel and its quantum capacity

For a finite-dimensional complex Hilbert space $\mathcal{H}$, let $\mathcal{S}(\mathcal{H})$ be the set of density operators on $\mathcal{H}$ and $\mathcal{L}(\mathcal{H})$ be the set of linear operators on $\mathcal{H}$. The standard description of a quantum channel is the completely positive trace-preserving map (CP map).[26–28] Suppose that we send a state $\rho \in \mathcal{S}(\mathcal{H})$. The statistical ensemble of the received states is described as $\Gamma(\rho)$ by a CP map Γ.

Suppose that we send a state $\rho \in \mathcal{S}(\mathcal{H}^{\otimes n})$ through a quantum channel. The quantum channel is said to be *memoryless* if the received state is described as $\Gamma^{\otimes n}(\rho)$ for all $\rho \in \mathcal{S}(\mathcal{H}^{\otimes n})$ and for some CP map Γ on $\mathcal{L}(\mathcal{H})$.

Fidelity is a measure of closeness between two quantum states. The fidelity F between a pure state $|\varphi\rangle \in \mathcal{H}$ and a state $\rho \in \mathcal{S}(\mathcal{H})$ is defined by $\langle\varphi|\rho|\varphi\rangle$.[18,19] We have $0 \leqslant F \leqslant 1$ and two states are closer if the fidelity between them is larger.

Let H_2 be the two-dimensional complex Hilbert space. Unless otherwise stated we consider the binary memoryless quantum channel, that is, when we send $\rho \in \mathcal{S}(H_2^{\otimes n})$ we receive $\Gamma^{\otimes n}(\rho)$, where Γ is a CP map on $\mathcal{L}(H_2)$. We shall identify a binary memoryless channel with a CP map on $\mathcal{L}(H_2)$.

A binary $[[n,k]]$ quantum code Q is a 2^k-dimensional subspace of $H_2^{\otimes n}$. The rate of an $[[n,k]]$ quantum code is k/n. The quantum capacity of a binary memoryless channel Γ is the maximum number $Q(\Gamma)$ such that for any rate $R<Q(\Gamma)$ and any $\epsilon>0$ there exists an $[[n,k]]$ quantum code Q with $k/n \geqslant R$ such that the fidelity between the recovered state and the original state $|\varphi\rangle \in Q$ is at least $1-\epsilon$ for any $|\varphi\rangle$.[1,2]

B. Fidelity bound of the quantum error correction

In this subsection we review Preskill's lower bound on the fidelity of quantum error correction in terms of the set of uncorrectable errors of a quantum code. Let

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

and $\mathcal{E}=\{w_1 \otimes \cdots \otimes w_n\}$, where w_i is either I, σ_x, σ_z or $\sigma_x\sigma_z$. For a quantum code Q and a fixed error correction process for Q, an operator $M \in \mathcal{E}$ is said to be correctable if the error correction process of Q recovers $M|\varphi\rangle$ to $|\varphi\rangle$ for all $|\varphi\rangle \in Q$. An operator M is uncorrectable if it is not correctable. Let $\mathcal{E}_{\mathrm{unc}} \subset \mathcal{E}$ be the set of uncorrectable errors of a quantum code $Q \subset H_2^{\otimes n}$. Suppose that we send a pure state $|\varphi\rangle \in Q$ through a binary memoryless channel described by a CP map Γ on $\mathcal{L}(H_2)$. By a unitary representation of a CP map,[29] there exists a finite-dimensional Hilbert space H_{env}, a pure state $|0_{\mathrm{env}}\rangle \in H_{\mathrm{env}}$ and a unitary operator U on $H_2^{\otimes n} \otimes H_{\mathrm{env}}$ such that

$$\Gamma(\rho) = \mathrm{Tr}_{H_{\mathrm{env}}}(U(\rho \otimes |0_{\mathrm{env}}\rangle\langle 0_{\mathrm{env}}|)U^*) \tag{1}$$

for all $\rho \in \mathcal{S}(H_2^{\otimes n})$, where $\mathrm{Tr}_{H_{\mathrm{env}}}$ is the partial trace over H_{env}. Since $\mathcal{E}$ is a basis of $\mathcal{L}(H_2^{\otimes n})$ we can write U in Eq. (1) as

$$U = \sum_{M \in \mathcal{E}} M \otimes L_M,$$

where L_M is a linear operator on H_{env}. Preskill proved the following theorem in Sec. 7.4 of Ref. 23.

Theorem 1: *Let Q and $\mathcal{E}_{\mathrm{unc}}$ be as above. When we send a pure state $|\varphi\rangle \in Q$, the fidelity between $|\varphi\rangle$ and the recovered state is not less than*

$$1 - \left\| \sum_{M \in \mathcal{E}_{\mathrm{unc}}} M|\varphi\rangle \otimes L_M |0_{\mathrm{env}}\rangle \right\|^2,$$

where $\|\cdot\|$ denotes the norm of a vector.

C. Stabilizer codes and their error correction process

In this subsection we review stabilizer quantum codes introduced in Refs. 4–6. Let $E = \{\pm w_1 \otimes \cdots \otimes w_n\}$, where w_i is either I, σ_x, σ_z or $\sigma_x\sigma_z$, S is a commutative subgroup of E, and

$$S' = \{M \in E : \forall N \in S, \quad MN = NM\}.$$

A stabilizer code Q is defined as a simultaneous eigenspace of all matrices in S. If S' has 2^{n+k+1} elements, then $\dim Q = 2^k$. The set of simultaneous eigenspaces of S is equal to $\{MQ: \ M \in \mathcal{E}\}$, where $MQ = \{M|\varphi\rangle : |\varphi\rangle \in Q\}$.

We shall describe the error correction process of a stabilizer code. Suppose that we send a pure state $|\varphi\rangle \in Q$ and received $\rho \in \mathcal{S}(H_2^{\otimes n})$. We measure an observable of $H_2^{\otimes n}$ whose eigenspaces are the same as those of S. Then the received state ρ is projected to a state ρ' that is an ensemble of pure states in some eigenspace Q' of S. For $M = \pm w_1 \otimes \cdots \otimes w_n \in E$ we define the weight $w(M)$ of M by $\sharp\{i : w_i \neq I\}$, where $\sharp$ denotes the number of elements in a set. Let

$M \in \mathcal{E}$ such that $MQ = Q'$ and that, if $MQ = M'Q$ for $M' \in \mathcal{E}$, then $w(M) \leqslant w(M')$. We recover ρ' to $M^{-1}\rho'(M^{-1})^*$. With this error correction process the set of uncorrectable errors is contained in

$$\{M \in \mathcal{E}: \text{ there exists } M' \in \mathcal{E} \text{ such that } w(M') \leqslant w(M),$$
$$M'Q = MQ, \text{ and } MS \neq \pm M'S\} = \{M \in \mathcal{E}: \text{ there exists } M' \in \mathcal{E} \text{ such that } w(M') \leqslant w(M), \tag{2}$$
$$M'S' = MS', \text{ and } MS \neq \pm M'S\}.$$

Hamada[30] showed the following theorem based on Theorem 1.

Theorem 2: *Notations as in Theorem 1. Let Q be a stabilizer code with the decoding process described above. Then there exists a subspace $Q' \subset Q$ such that* $\dim Q' = \dim Q/2$ *and that for all pure state $|\varphi\rangle \in Q'$, the fidelity between $|\varphi\rangle$ and the recovered state is not less than*

$$1 - 2\sum_{M \in \mathcal{E}_{\text{unc}}} \| L_M |0_{\text{env}}\rangle \|^2. \tag{3}$$

Observe that the information rates of Q and Q' in Theorem 2 differ by $\log 2/n$, which becomes negligible as $n \to \infty$. We call a subspace Q' as a subcode of Q as in the classical coding theory. We shall consider the subcode Q' of a stabilizer code Q in the discussion of Sec. III.

D. Symplectic geometry

In this subsection we review the symplectic geometric interpretation of stabilizer codes introduced in Refs. 4 and 5. A symplectic geometry is a linear space with a nondegenerate symplectic form.[31] Let $\mathbf{F}_2$ be the finite field with two elements. For $\vec{a} = (a_1, \ldots, a_n) \in \mathbf{F}_2^n$ and $\vec{b} = (b_1, \ldots, b_n) \in \mathbf{F}_2^n$, we define $(\vec{a}|\vec{b})$ by $(a_1, \ldots, a_n, b_1, \ldots, b_n) \in \mathbf{F}_2^{2n}$ and

$$f(\pm \sigma_x^{a_1}\sigma_z^{b_1} \otimes \cdots \otimes \sigma_x^{a_n}\sigma_z^{b_n}) = (\vec{a}|\vec{b}).$$

We also define the standard symplectic form of $(\vec{a}|\vec{b})$ and $(\vec{a}'|\vec{b}') \in \mathbf{F}_2^{2n}$ by

$$\langle \vec{a}, \vec{b}' \rangle - \langle \vec{a}', \vec{b} \rangle, \tag{4}$$

where $\langle \cdot, \cdot \rangle$ denotes the standard inner product in $\mathbf{F}_2^n$. For a subspace $C \subseteq \mathbf{F}_2^{2n}$ we denote by $C^\perp$ the orthogonal space of C with respect to (4). For a subgroup $S \subseteq E$, S is commutative if and only if $f(S) \subseteq (f(S))^\perp$, and $S' = f^{-1}((f(S))^\perp)$.

E. Linear stabilizer codes and unitary geometry

Calderbank *et al.*[5] related stabilizer codes to classical error-correcting codes and unitary geometry, which is a linear space with a nondegenerate hermitian form.[31] Let ω be a primitive element in $\mathbf{F}_4$, and define

$$g(\pm \sigma_x^{a_1}\sigma_z^{b_1} \otimes \cdots \otimes \sigma_x^{a_n}\sigma_z^{b_n}) = \omega\vec{a} + \omega^2\vec{b} \in \mathbf{F}_4^n.$$

For vectors $\vec{x}$, $\vec{y} \in \mathbf{F}_4^n$ we define an $\mathbf{F}_2$-bilinear map

$$\langle \vec{x}^2, \vec{y} \rangle - \langle \vec{x}, \vec{y}^2 \rangle, \tag{5}$$

where $\langle , \rangle$ denotes the standard inner product in $\mathbf{F}_4^n$ and $\vec{x}^2 = (x_1^2, \ldots, x_n^2)$. For an $\mathbf{F}_2$-linear subspace C of $\mathbf{F}_4^n$, let $C^\perp$ denotes the orthogonal space of C with respect to (5). For a subgroup $S \subseteq E$, S is commutative if and only if $g(S) \subseteq (g(S))^\perp$, and $S' = g^{-1}((g(S))^\perp)$.

For $\vec{x}, \vec{y} \in \mathbf{F}_4^n$, we define the standard Hermitian form of $\vec{x}, \vec{y}$ by

$$\tau(\vec{x},\vec{y}) = \langle \vec{x}^2, \vec{y} \rangle, \tag{6}$$

which is used only in Sec. III D. If C is an $\mathbf{F}_4$-linear subspace of $\mathbf{F}_4^n$, $C^\perp$ is equal to the orthogonal space of C with respect to (6). A stabilizer code constructed from an $\mathbf{F}_2$-linear self-orthogonal space $C \subset \mathbf{F}_4^n$ is said to be linear if C is $\mathbf{F}_4$-linear. This connection between binary stabilizer codes and classical codes over $\mathbf{F}_4$ is generalized to nonbinary case in Refs. 32–34.

III. LOWER BOUND ON THE QUANTUM CAPACITY

As described in the Introduction, we have to calculate the average of fidelity over all the stabilizer codes, and show that the average converges to 1. Strictly speaking, we shall use the subcode of a stabilizer code introduced in Theorem 2. This section is organized as follows: In Sec. III A we introduce a definition of the distance of a quantum channel from the identity channel. In Sec. III B we calculate the average of fidelity over subcodes of general stabilizer codes of the fixed rate. In Sec. III C we deduce a sufficient condition for the rate to let the average of fidelity converge to 1. In Sec. III D we indicate that a small modification of the argument in Secs. III B and III C shows that the same lower bound on the capacity is obtained from the random coding of linear stabilizer codes.

A. A definition of distance of a quantum channel from the identity channel

We give a lower bound on the quantum capacity in terms of the distance of a quantum channel from the identity channel. Let us first review a definition of the distance of a quantum channel from the identity channel introduced in Ref. 22, then show some properties of the definition. Let Γ be a CP map on $\mathcal{L}(H_2)$.

Definition 3: Suppose that there exist a four-dimensional space H_E, $|e_0\rangle \in H_E$, and a unitary operator U on $H_2 \otimes H_E$ such that

$$\Gamma(\rho) = \mathrm{Tr}_E[U(\rho \otimes |e_0\rangle\langle e_0|)U^*] \tag{7}$$

for all $\rho \in \mathcal{S}(H_2)$. Write U as

$$U = I \otimes L_I + \sigma_x \otimes L_x + \sigma_z \otimes L_z + \sigma_x\sigma_z \otimes L_{xz},$$

where L_I, L_x, L_z, and L_{xz} are linear operators on H_E. Then the distance $p(\Gamma)$ and $q(\Gamma)$ of the channel Γ from the identity channel are defined by

$$q(\Gamma) = \| L_I |e_0\rangle \|^2,$$

$$p(\Gamma) = \| L_x |e_0\rangle \|^2 + \| L_z |e_0\rangle \|^2 + \| L_{xz} |e_0\rangle \|^2.$$

It is not clear whether the values of $p(\Gamma)$ and $q(\Gamma)$ are uniquely determined by Γ alone, that is, whether they are independent of choice of U and $|e_0\rangle$ in Eq. (7). In order to answer this question in Corollary 5, we shall represent $p(\Gamma)$ and $q(\Gamma)$ using the operator-sum representation of Γ induced by U and $|e_0\rangle$.

Proposition 4: Let $|e_0\rangle, \ldots, |e_3\rangle$ be an orthonormal basis of H_E, and

$$A_i = \langle e_i | U | e_0 \rangle.$$

By Eq. (8.10) of Ref. 35,

$$\Gamma(\rho) = \sum_{i=0}^{3} A_i \rho A_i^*$$

for all $\rho \in \mathcal{S}(H_2)$. Write A_i as

$$A_i = a_{i,I} I + a_{i,x}\sigma_x + a_{i,z}\sigma_z + a_{i,xz}\sigma_x\sigma_z .$$

Then we have another representations of $p(\Gamma)$ *and* $q(\Gamma)$ *as*

$$p(\Gamma) = \sum_{i=0}^{3} |a_{i,x}|^2 + |a_{i,z}|^2 + |a_{i,xz}|^2,$$

$$q(\Gamma) = \sum_{i=0}^{3} |a_{i,I}|^2.$$

Proof: By definition of A_i,

$$A_i = \langle e_i|L_I|e_0\rangle I + \langle e_i|L_x|e_0\rangle\sigma_x + \langle e_i|L_z|e_0\rangle\sigma_z + \langle e_i|L_{xz}|e_0\rangle\sigma_x\sigma_z .$$

Therefore,

$$a_{i,I} = \langle e_i|L_I|e_0\rangle, \quad a_{i,x} = \langle e_i|L_x|e_0\rangle, \quad a_{i,z} = \langle e_i|L_z|e_0\rangle, \quad a_{i,xz} = \langle e_i|L_{xz}|e_0\rangle.$$

Since $|e_0\rangle,\ldots,|e_3\rangle$ are an orthonormal basis, we have

$$q(\Gamma) = \| L_I|e_0\rangle \|^2 = \sum_{i=0}^{3} |a_{i,I}|^2.$$

The equality of $p(\Gamma)$ can be shown in a similar way. ■

Corollary 5: The values of $p(\Gamma)$ *and* $q(\Gamma)$ *do not depend on choice of* U *and* $|e_0\rangle$ *in Eq. (7).*

Proof: Let

$$\Gamma(\rho) = \sum_{i=0}^{3} B_i \rho B_i^*$$

be another operator-sum representation of Γ. By Theorem 8.2 of Ref. 35, there exists a 4×4 unitary matrix V such that

$$\begin{pmatrix} B_0 \\ \vdots \\ B_3 \end{pmatrix} = V \begin{pmatrix} A_0 \\ \vdots \\ A_3 \end{pmatrix}.$$

Write B_i as

$$B_i = b_{i,I} I + b_{i,x}\sigma_x + b_{i,z}\sigma_z + b_{i,xz}\sigma_x\sigma_z ,$$

and define $\vec{a}_I = (a_{0,I},\ldots,a_{3,I})^T$, $\vec{b}_I = (b_{0,I},\ldots,b_{3,I})^T$. Since I, σ_x, σ_z, and $\sigma_x\sigma_z$ are linearly independent, we have $\vec{b}_I = V\vec{a}_I$. Since V is unitary, $\|\vec{b}_I\| = \|\vec{a}_I\|$, which shows that $q(\Gamma)$ does not depend on choice of representation. The independence of $p(\Gamma)$ can be shown in a similar way. ■

The following corollary drastically simplifies the formula for the lower bound in Eq. (16).

Corollary 6: $p(\Gamma) + q(\Gamma) = 1$.

Proof: Notations as in Proposition 4. We have

$$I = \sum_{i=0}^{3} A_i^* A_i .$$

Taking trace on the both side, we have

$$\mathrm{Tr}[I]=\sum_{i=0}^{3}\mathrm{Tr}[A_i^* A_i]$$

$$=\sum_{i=0}^{3}\mathrm{Tr}[(a_{i,I}I+a_{i,x}\sigma_x+a_{i,z}\sigma_z+a_{i,xz}\sigma_x\sigma_z)^*\times(a_{i,I}I+a_{i,x}\sigma_x+a_{i,z}\sigma_z+a_{i,xz}\sigma_x\sigma_z)]$$

$$=\mathrm{Tr}[I]\sum_{i=0}^{3}|a_{i,I}|^2+|a_{i,x}|^2+|a_{i,z}|^2+|a_{i,xz}|^2=\mathrm{Tr}[I](p(\Gamma)+q(\Gamma)).$$

■

B. Average of the fidelity over all the stabilizer codes

In this subsection we shall prove that the average fidelity of the subcodes of all $[[n,\lfloor Rn\rfloor]]$ stabilizer codes converges to 1 as $n\to\infty$ under the conditions (8) and (9). The proof is proceeded as follows:

(1) For every stabilizer code, there exists its subcode whose fidelity of error correction is lower bounded by Eq. (3). The lower bound (3) is expressed as a sum indexed by error operators. The average of the sum will be divided according to the weight of error operators in Eq. (10).
(2) It is easy to see the part indexed by operators of larger weights converges to 0 as $n\to\infty$.
(3) We shall show that the other part indexed by operators of smaller weights converges to 0 by the fact that most of stabilizer codes can correct an error of small weight, which will be rigorously proved in Eq. (13) from Lemma 9.

Let δ and R be real numbers such that

$$\lim_{n\to\infty}\sum_{i=\lfloor\delta n\rfloor+1}^{n}\binom{n}{i}p(\Gamma)^i q(\Gamma)^{n-i}=0, \tag{8}$$

$$1-\lim_{n\to\infty}\frac{\log_2\left[\sum_{i=1}^{\lfloor\delta n\rfloor}\binom{n}{i}p(\Gamma)^i q(\Gamma)^{n-i}\Sigma_{j=0}^{i}\binom{n}{j}3^j\right]}{n}>R, \tag{9}$$

where $\lfloor x\rfloor$ denotes the largest integer $\leqslant x$.

Let

$$A_n=\{C\subset\mathbf{F}_2^{2n}:\ C \text{ is linear},\quad \dim C=n-\lfloor Rn\rfloor, C\subseteq C^{\perp}\}.$$

Recall that we can construct an $[[n,\lfloor Rn\rfloor]]$ stabilizer code from every $C\in A_n$. Note that A_n is not empty because there exists a self-orthogonal subspace of dimension n in $\mathbf{F}_2^{2n}$.[31] This subsection is devoted to show the following.

Proposition 7: If R satisfies Eq. (9), then there exists a sequence of subcodes of stabilizer codes whose rates are greater than or equal to R and whose fidelity converges to 1 as $n\to\infty$.

Since the information rates of the subcode Q' and Q in Theorem 2 are asymptotically the same as $n\to\infty$, it is sufficient to show that the average of the fidelity bound (3) of Q' over all the stabilizer codes in A_n converges to 1 as $n\to\infty$.

Let $|0_{\mathrm{env}}\rangle=|e_0\rangle^{\otimes n}$, and for $M=\sigma_{i_1}\otimes\cdots\otimes\sigma_{i_n}\in\mathcal{E}$ let

$$L_M=L_{i_1}\otimes\cdots\otimes L_{i_n},$$

where $\sigma_I=I$ and $|e_0\rangle$, L_I, L_x, L_z, and L_{xz} are as defined in Definition 3. For $C\in A_n$ we denote the set of uncorrectable errors of C in $\mathcal{E}$ by $\mathcal{E}_{\mathrm{unc}}(C)$. The average of the fidelity bound (3) of Q' over all the stabilizer codes in A_n is not less than

$$\frac{1}{\sharp A_n}\sum_{C\in A_n}\left(1-2\sum_{M\in\mathcal{E}_{\mathrm{unc}}(C)}\|L_M|0_{\mathrm{env}}\rangle\|^2\right)$$
$$=1-\frac{2}{\sharp A_n}\sum_{C\in A_n}\left(\sum_{\substack{M\in\mathcal{E}_{\mathrm{unc}}(C)\\1\leqslant w(M)\leqslant\lfloor\delta n\rfloor}}\|L_M|0_{\mathrm{env}}\rangle\|^2+\sum_{\substack{M\in\mathcal{E}_{\mathrm{unc}}(C)\\w(M)>\lfloor\delta n\rfloor}}\|L_M|0_{\mathrm{env}}\rangle\|^2\right)$$
$$\geqslant 1-\left(\frac{2}{\sharp A_n}\sum_{C\in A_n}\sum_{\substack{M\in\mathcal{E}_{\mathrm{unc}}(C)\\1\leqslant w(M)\leqslant\lfloor\delta n\rfloor}}\|L_M|0_{\mathrm{env}}\rangle\|^2\right)-2\sum_{\substack{M\in\mathcal{E}\\w(M)>\lfloor\delta n\rfloor}}\|L_M|0_{\mathrm{env}}\rangle\|^2. \tag{10}$$

By the same argument as Ref. 22, one can show that

$$\sum_{\substack{M\in\mathcal{E}\\w(M)>\lfloor\delta n\rfloor}}\|L_M|0_{\mathrm{env}}\rangle\|\leqslant\sum_{i=\lfloor\delta n\rfloor+1}^{n}\binom{n}{i}p(\Gamma)^i q(\Gamma)^{n-i},$$

which converges to 0 as $n\to\infty$ by the condition (8).

We shall calculate an upper bound for the second term in Eq. (10). For $M\in\mathcal{E}$ we define

$$B_n(M)=\{C\in A_n : M\in\mathcal{E}_{\mathrm{unc}}(C)\}.$$

It follows that

$$\frac{1}{\sharp A_n}\sum_{C\in A_n}\sum_{\substack{M\in\mathcal{E}_{\mathrm{unc}}(C)\\1\leqslant w(M)\leqslant\lfloor\delta n\rfloor}}\|L_M|0_{\mathrm{env}}\rangle\|^2\leqslant\frac{1}{\sharp A_n}\sum_{\substack{M\in\mathcal{E}\\1\leqslant w(M)\leqslant\lfloor\delta n\rfloor}}\sharp B_n(M)\|L_M|0_{\mathrm{env}}\rangle\|^2. \tag{11}$$

Note that we omitted the factor 2 from Eq. (10) for simplicity, because we shall show that the right-hand side of Eq. (11) converges to 0 and factor 2 is negligible.

We shall give an upper bound for $\sharp B_n(M)$. To estimate $\sharp B_n(M)$ we shall introduce Lemma 9. In the proof of Lemma 9 we use the Witt theorem, so we review it.

Theorem 8 (Witt): *Let K be a field, V_1, V_2 finite-dimensional K-linear spaces, and τ_1, τ_2 symplectic forms on V_1, V_2, respectively. An injective linear map $T:V_1\to V_2$ is said to be an* isometry *if*

$$\tau_1(x,y)=\tau_2(Tx,Ty).$$

Let W_1 be a subspace of V_1. If there exists a bijective isometry from V_1 to V_2 and an isometry $T_{W_1}:W_1\to V_2$, then there exists an isometry $T_{V_1}:V_1\to V_2$ such that the restriction of T_{V_1} to W_1 is equal to T_{W_1}. The same result also holds when τ_1,τ_2 are Hermitian forms.

Proof: See Sec. 20 of Ref. 31. ■

Lemma 9: For $M\in E-\{\pm I\}$, let $A_n(M)=\{C\in A_n : f(M)\in C^{\perp}\backslash C\}$. We have

$$\sharp A_n(M)\leqslant(1/2^{n-\lfloor Rn\rfloor})\frac{1-2^{-2\lfloor Rn\rfloor}}{1-2^{-2n}}\sharp A_n<\sharp A_n/2^{n-\lfloor Rn\rfloor}.$$

Proof: Let $\mathrm{Sp}_n(\mathbf{F}_2)$ be the group of bijective linear maps on $\mathbf{F}_2^{2n}$ preserving the symplectic form (4). For every pair of spaces C_1, $C_2\in A_n$, every bijective linear map from C_1 to C_2 is an isometry. Consequently, there exists $\sigma\in\mathrm{Sp}_n(\mathbf{F}_2)$ such that $\sigma C_1=C_2$ by the Witt theorem. A similar argument shows that there exists $\sigma'\in\mathrm{Sp}_n(\mathbf{F}_2)$ such that $\sigma'(\vec{a}|\vec{b})=(\vec{a}'|\vec{b}')$ for every pair of nonzero vectors $(\vec{a}|\vec{b})$, $(\vec{a}'|\vec{b}')\in\mathbf{F}_2^{2n}$.

It follows that

$$\begin{aligned}\sharp A_n(M) &= \sharp\{C \in A_n : f(M) \in C^{\perp}\backslash C\} \\ &= \sharp\{\alpha C_1 : f(M) \in (\alpha C_1)^{\perp}\backslash \alpha C_1, \alpha \in \mathrm{Sp}_n(\mathbf{F}_2)\} \\ &= \sharp\{\alpha C_1 : \beta(f(M)) \in (\alpha C_1)^{\perp}\backslash \alpha C_1, \alpha \in \mathrm{Sp}_n(\mathbf{F}_2)\},\end{aligned}$$

where C_1 (resp. β) is an arbitrary fixed element in A_n [resp. $\mathrm{Sp}_n(\mathbf{F}_2)$]. Therefore $\sharp A_n(M)$ is the same among every nonzero $f(M)$.

Since $\sharp(C^{\perp}\backslash C) = 2^{n+\lfloor Rn\rfloor} - 2^{n-\lfloor Rn\rfloor}$, there are $(2^{n+\lfloor Rn\rfloor} - 2^{n-\lfloor Rn\rfloor})\sharp A_n$ pairs of $((\vec{a}|\vec{b}), C)$ such that $(\vec{a}|\vec{b}) \in C^{\perp}\backslash C$ and $C \in A_n$. Thus if $M \neq \pm I$, then

$$\sharp A_n(M) \leqslant \frac{2^{n+\lfloor Rn\rfloor} - 2^{n-\lfloor Rn\rfloor}}{2^{2n}-1}\sharp A_n = (1/2^{n-\lfloor Rn\rfloor})\frac{1-2^{-2\lfloor Rn\rfloor}}{1-2^{-2n}}\sharp A_n < \sharp A_n/2^{n-\lfloor Rn\rfloor}.$$

■

Remark 10: From Lemma 9 we can improve the quantum Gilbert–Varshamov bound slightly. There exists an $[[n,k,d]]$ *stabilizer code if*

$$\frac{1-2^{-2k}}{1-2^{-2n}} \cdot \frac{1}{2^{n-k}} \sum_{i=1}^{d-1} 3^i \binom{n}{i} < 1. \tag{12}$$

The proof is as follows: For each error $M \in \mathcal{E}$, $A_n(M)$ *is equal to the set of stabilizer codes unable to detect* M *as an error. Therefore, by replacing* $\lfloor Rn\rfloor$ *with* k *in Lemma 9, we see that if Eq. (12) holds, then there is at least one stabilizer code* C *is able to detect all the errors* M *with* $w(M) < d$, *which means that the minimum distance of* C *is at least* d. *The idea behind this proof already appeared in the original proof of the quantum Gilbert–Varshamov bound for stabilizer codes.*[4] *Observe that our bound is slightly better than the quantum Gilbert–Varshamov bound for general codes,*[20] *which implies that an* $[[n,k,d]]$ *quantum code exists if*

$$\frac{1}{2^{n-k}} \sum_{i=0}^{d-1} 3^i \binom{n}{i} < 1.$$

By Eq. (2), $M \in \mathcal{E}$ belongs to $\mathcal{E}_{\mathrm{unc}}(C)$ only if there exists $M' \in \mathcal{E}$ such that $w(M') \leqslant w(M)$, $Mf^{-1}(C^{\perp}) = M'f^{-1}(C^{\perp})$, and $Mf^{-1}(C) \neq M'f^{-1}(C)$. A space $C \in A_n$ belongs to $B_n(M)$ only if there exists $M' \in \mathcal{E}$ such that $w(M') \leqslant w(M)$ and $M^{-1}M' \in f^{-1}(C^{\perp}\backslash C)$. The last condition is equivalent to $C \in A_n(M^{-1}M')$. Since there are

$$\sum_{j=0}^{w(M)} \binom{n}{j} 3^j$$

operators $M' \in \mathcal{E}$ such that $w(M') \leqslant w(M)$, it follows that

$$\begin{aligned}\sharp B_n(M) &\leqslant \sum_{\substack{M' \in \mathcal{E} \\ w(M') \leqslant w(M)}} \sharp A_n(M^{-1}M') \\ &\leqslant \sum_{\substack{M' \in \mathcal{E} \\ w(M') \leqslant w(M)}} \frac{\sharp A_n}{2^{n-[Rn]}} \leqslant \frac{\sharp A_n}{2^{n-[Rn]}} \sum_{j=0}^{w(M)} \binom{n}{j} 3^j.\end{aligned} \tag{13}$$

An upper bound for Eq. (11) is derived as follows:

$$\frac{1}{\sharp A_n}\sum_{\substack{M\in\mathcal{E}\\ 1\leqslant w(M)\leqslant\lfloor\delta n\rfloor}}\sharp B_n(M)\|L_M|0_{\mathrm{env}}\rangle\|^2$$
$$\leqslant\frac{1}{\sharp A_n}\sum_{\substack{M\in\mathcal{E}\\ 1\leqslant w(M)\leqslant\lfloor\delta n\rfloor}}\frac{\sharp A_n}{2^{n-[Rn]}}\sum_{j=0}^{w(M)}\binom{n}{j}3^j\|L_M|0_{\mathrm{env}}\rangle\|^2$$
$$=\frac{1}{2^{n-[Rn]}}\sum_{\substack{M\in\mathcal{E}\\ 1\leqslant w(M)\leqslant\lfloor\delta n\rfloor}}\sum_{j=0}^{w(M)}\binom{n}{j}3^j\|L_M|0_{\mathrm{env}}\rangle\|^2. \tag{14}$$

For an integer $0\leqslant i\leqslant n$, by the same argument as Ref. 22, one can show that

$$\sum_{\substack{M\in\mathcal{E}\\ w(M)=i}}\|L_M|0_{\mathrm{env}}\rangle\|^2=\binom{n}{i}p(\Gamma)^i q(\Gamma)^{n-i}.$$

Therefore Eq. (14) is equal to

$$\frac{1}{2^{n-[Rn]}}\sum_{i=1}^{\lfloor\delta n\rfloor}\binom{n}{i}p(\Gamma)^i q(\Gamma)^{n-i}\sum_{j=0}^{i}\binom{n}{j}3^j,$$

which converges to 0 as $n\to\infty$ by the condition (9).

C. Achievable rate by general stabilizer codes

In the previous subsection, we have shown that if the rate R satisfies Eq. (9), then there exists at least one sequence of subcodes of stabilizer codes of the rate R such that the average of fidelity converges to 1. In this subsection we shall simplify Eqs. (8) and (9) with which we can easily compute a lower bound on the capacity of the channel Γ.

We shall deduce a sufficient condition for δ to satisfy Eq. (8). By Appendix A of Ref. 36, for $0\leqslant\epsilon<\lambda\leqslant 1$ we have

$$\sum_{i=\lambda n}^{n}\binom{n}{i}\epsilon^i(1-\epsilon)^{n-i}\leqslant 2^{-nD(\lambda\|\epsilon)},$$

where $D(\lambda\|\epsilon)$ is the classical relative entropy defined by

$$\lambda\log_2\frac{\lambda}{\epsilon}+(1-\lambda)\log_2\frac{1-\lambda}{1-\epsilon}.$$

Since $p(\Gamma)+q(\Gamma)=1$ by Corollary 6, the condition (8) holds if

$$\delta>p(\Gamma). \tag{15}$$

The term inside of $\log_2$ in Eq. (9) can be bounded as follows:

$$\sum_{i=1}^{\lfloor \delta n \rfloor} \binom{n}{i} p(\Gamma)^i q(\Gamma)^{n-i} \sum_{j=0}^{i} \binom{n}{j} 3^j \leq \sum_{i=0}^{n} \binom{n}{i} p(\Gamma)^i q(\Gamma)^{n-i} \sum_{j=0}^{\lfloor \delta n \rfloor} \binom{n}{j} 3^j$$

$$= [p(\Gamma)+q(\Gamma)]^n \sum_{j=0}^{\lfloor \delta n \rfloor} \binom{n}{j} 3^j = \sum_{j=0}^{\lfloor \delta n \rfloor} \binom{n}{j} 3^j \quad \text{(by Corollary 6)}$$

$$\leq (\delta n+1) \binom{n}{\lfloor \delta n \rfloor} 3^{\delta n}$$

$$\leq (\delta n+1) \exp_2[nH_e(\delta)] 3^{\delta n} \quad \text{(by Appendix B of Ref. 36)}$$

$$= (\delta n+1) \exp_2\{n[H_e(\delta)+\delta \log_2 3]\},$$

where H_e is the binary entropy function.

From Proposition 7, Eq. (15), and the observations above, we see that the capacity of the channel Γ is at least

$$1-\{H_e[p(\Gamma)]+p(\Gamma)\log_2 3\}. \tag{16}$$

Note that the same lower bound on the capacity can also be obtained by the method of Bennett *et al.*,[2] though they stated their result only for the depolarizing channel. However, Bennett *et al.*[2] did not address the achievability of the bound (16) with stabilizer codes, which is the main focus of this article.

We shall compare our bound on the capacity [Eq. (16)] with the conventional bound for a general memoryless channel derived from the quantum Gilbert–Varshamov bound[4,20] and the fidelity bounds for t-error-correcting codes.[21,22] Suppose that we have a sequence of $\lfloor \delta n_i \rfloor$-error-correcting quantum codes of length n_i with $\lim_{i\to\infty} n_i = \infty$. The condition (15) is sufficient in order that the fidelity of error correction by $\lfloor \delta n_i \rfloor$-error-correcting codes converges to 1 as $i \to \infty$. By the quantum Gilbert–Varshamov bound the derived lower bound on the capacity is

$$1-\{H_e[2p(\Gamma)]+2p(\Gamma)\log_2 3\},$$

which is always smaller than Eq. (16).

When the channel Γ is the depolarizing channel of the fidelity parameter f, $p(\Gamma)=1-f$ and $q(\Gamma)=f$. The proposed lower bound [Eq. (16)] for the capacity is

$$1-[H_e(1-f)+(1-f)\log_2 3],$$

which coincides with the lower bound given in Ref. 2. It is not clear to the authors whether our lower bound can be improved by the method in Ref. 7.

Our analysis for the quantum capacity can be generalized to the capacity of an l-adic channel using the l-adic stabilizer codes[37,38] in a straightforward manner when l is prime. The quantum Gilbert–Varshamov bound for l-adic stabilizer codes can also be proved by Lemma 9.

D. Achievable rate by linear stabilizer codes

In this subsection we shall show that the achievable rate (16) by subcodes of general stabilizer codes can also be achieved by those of linear stabilizer codes, which shows the asymptotic optimality of linear stabilizer codes among general ones. As a byproduct we establish an analog of Gilbert–Varshamov bound for linear stabilizer codes.

Let

$$A_n' = \{C \subset \mathbf{F}_4^n : C \text{ is } \mathbf{F}_4\text{-linear}, \ \dim_{\mathbf{F}_4} C = \lfloor (n-Rn)/2 \rfloor, \quad C \subseteq C^{\perp}\}.$$

Recall that we can construct an $[[n,n-2\lfloor(n-Rn)/2\rfloor]]$ linear stabilizer code from every $C \in A_n'$. Note that A_n' is not empty because there exists a self-orthogonal subspace of dimension $\lfloor n/2 \rfloor$ in $\mathbf{F}_4^n$ (see Proposition 2.3.2 in Ref. 39). For $M \in \mathcal{E}$, define

$$A_n'(M)=\{C \in A_n' : g(M) \in C^{\perp} \backslash C\},$$

$$B_n'(M)=\{C \in A_n' : M \text{ is uncorrectable by } C\}.$$

By these definitions of $A_n'(M)$ and $B_n'(M)$, all the arguments except Lemma 9 in the previous subsections can be used for showing that the rate (16) is achieved by subcodes of linear stabilizer codes. In this subsection we prove an upper bound (Lemma 11) for $\sharp A_n'(M)$ that can be used as a substitute for Lemma 9.

Lemma 11: Define τ by Eq. (6). The number of nonzero vectors $\vec{x} \in \mathbf{F}_4^n$ such that $\tau(\vec{x},\vec{x})=0$ is $2^{2n-1}+(-1)^n 2^{n-1}-1$.

Proof: See the proof of Proposition 2.3.3 in Ref. 39. ■

Lemma 12: Let $u=\lfloor(n-Rn)/2\rfloor$. For $M \in E\backslash\{\pm I\}$

$$\sharp A_n'(M) \leqslant \frac{4^{n-u}-4^u}{\min\{2^{2n-1}+(-1)^n 2^{n-1}-1, 2^{2n-1}-(-1)^n 2^{n-1}\}} \sharp A_n' \leqslant \frac{4^{n-u}-4^u}{2^{2n-1}-2^{n-1}-1} \sharp A_n'.$$

Proof: Let $\mathrm{GU}_n(\mathbf{F}_4)$ be the group of bijective linear maps on $\mathbf{F}_4^n$ that preserve the value of the Hermitian form τ. For every pair of spaces $C_1, C_2 \in A_n'$, every bijective linear map from C_1 to C_2 is an isometry. Thus there exists $\sigma \in \mathrm{GU}_n(\mathbf{F}_4)$ such that $\sigma C_1 = C_2$ by the Witt theorem. For a pair of nonzero vectors $\vec{x},\vec{y} \in \mathbf{F}_4^n$ with $\tau(\vec{x},\vec{x})=\tau(\vec{y},\vec{y})=0$, a similar argument shows that there exists $\sigma \in \mathrm{GU}_n(\mathbf{F}_4)$ such that $\sigma\vec{x}=\vec{y}$.

We want to show that for a pair of vectors $\vec{x},\vec{y} \in \mathbf{F}_4^n$ with $\tau(\vec{x},\vec{x}) \neq 0$ and $\tau(\vec{y},\vec{y}) \neq 0$, there exists $\sigma \in \mathrm{GU}_n(\mathbf{F}_4)$ such that $\sigma\vec{x}=\vec{y}$. Since τ is a Hermitian form, $\tau(\vec{x},\vec{x}) \in \mathbf{F}_2$. Therefore $\tau(\vec{x},\vec{x})=\tau(\vec{y},\vec{y})=1$, and there exists $\sigma \in \mathrm{GU}_n(\mathbf{F}_4)$ such that $\sigma\vec{x}=\vec{y}$ by the Witt theorem.

A similar argument to the proof of Lemma 9 shows that for $M \in E\backslash\{\pm I\}$ we have

$$\sharp A_n'(M) \leqslant \frac{4^{n-u}-4^u}{2^{2n-1}+(-1)^n 2^{n-1}-1} \sharp A_n' \quad \text{if} \quad \tau(g(M),g(M))=0,$$

$$\sharp A_n'(M) \leqslant \frac{4^{n-u}-4^u}{2^{2n-1}-(-1)^n 2^{n-1}} \sharp A_n' \quad \text{if} \quad \tau(g(M),g(M)) \neq 0.$$

■

Remark 13: From Lemma 11 we can show that there exists an $[[n,k,d]]$ linear stabilizer code if k is even and

$$\frac{2(1-2^{-2k})}{1-2^{-n}-2^{-2n+1}} \cdot \frac{1}{2^{n-k}} \sum_{i=1}^{d-1} 3^i \binom{n}{i} < 1,$$

which is asymptotically the same as the quantum Gilbert–Varshamov bound for general quantum codes.[20]

Remark 14: The connection between stabilizer codes and classical codes over $\mathbf{F}_4$ was generalized to nonbinary case in Refs. 32–34. The argument in this subsection can be extended to linear l-adic stabilizer codes for a prime l with the following exception: In the proof of Lemma 11, there does not always exist $\sigma \in \mathrm{GU}_n(\mathbf{F}_{l^2})$ such that $\sigma\vec{x}=\vec{y}$ for a pair of vectors $\vec{x},\vec{y} \in \mathbf{F}_{l^2}^n$ with $\tau(\vec{x},\vec{x}) \neq 0$ and $\tau(\vec{y},\vec{y}) \neq 0$. However, there always exists $\sigma \in \mathcal{U}$ such that $\sigma\vec{x}=\vec{y}$, where $\mathcal{U}$ is the group generated by $\mathrm{GU}_n(\mathbf{F}_{l^2})$ and nonzero scalar multiples of the identity map on $\mathbf{F}_{l^2}^n$.

ACKNOWLEDGMENTS

We would like to thank Dr. Mitsuru Hamada for drawing our attention to the random coding of stabilizer codes and for providing detailed comments on this article, and Dr. Masahito Hayashi, Dr. Keiji Matsumoto, and Dr. Tomohiro Ogawa for helpful discussions.

[1] C. H. Bennett, D. P. DiVincenzo, and J. A. Smolin, Phys. Rev. Lett. **78**, 3217 (1997); quant-ph/9701015.
[2] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, Phys. Rev. A **54**, 3824 (1996); quant-ph/9604024.
[3] D. Gottesman, Ph.D. thesis, California Institute of Technology, 1997; quant-ph/9705052.
[4] A. R. Calderbank, E. M. Rains, P. W. Shor, and N. J. A. Sloane, Phys. Rev. Lett. **78**, 405 (1997); quant-ph/9605005.
[5] A. R. Calderbank, E. M. Rains, P. W. Shor, and N. J. A. Sloane, IEEE Trans. Inf. Theory **44**, 1369 (1998); quant-ph/9608006.
[6] D. Gottesman, Phys. Rev. A **54**, 1862 (1996); quant-ph/9604038.
[7] D. P. DiVincenzo, P. W. Shor, and J. A. Smolin, Phys. Rev. A **57**, 830 (1998); quant-ph/9706061.
[8] D. Bruß, D. P. DiVincenzo, A. Ekert, C. A. Fuchs, C. Macchiavello, and J. A. Smolin, Phys. Rev. A **57**, 2368 (1998); quant-ph/9705038.
[9] E. M. Rains, quant-ph/9707002.
[10] V. Vedral and M. B. Plenio, Phys. Rev. A **57**, 1619 (1998); quant-ph/9707035.
[11] N. J. Cerf, Phys. Rev. Lett. **84**, 4497 (2000); quant-ph/9803058.
[12] H. Barnum, M. A. Nielsen, and B. Schumacher, Phys. Rev. A **57**, 4153 (1998); quant-ph/9702049.
[13] B. Schumacher and M. A. Nielsen, Phys. Rev. A **54**, 2629 (1996); quant-ph/9604022.
[14] S. Lloyd, Phys. Rev. A **55**, 1613 (1997); quant-ph/9604015.
[15] H. Barnum, E. Knill, and M. A. Nielsen, IEEE Trans. Inf. Theory **46**, 1317 (2000); quant-ph/9809010.
[16] P. Elias, in *IRE Convention Record, Part 4* (1955), pp. 37–46.
[17] R. G. Gallager, *Information Theory and Reliable Communication* (Wiley, New York, 1968).
[18] R. Jozsa and B. Schumacher, J. Mod. Opt. **41**, 2343 (1994).
[19] A. Uhlmann, Rep. Math. Phys. **9**, 273 (1976).
[20] A. Ekert and C. Macchiavello, Phys. Rev. Lett. **77**, 2585 (1996); quant-ph/9602022.
[21] E. Knill and R. Laflamme, Phys. Rev. A **55**, 900 (1997); quant-ph/9604034.
[22] R. Matsumoto, Phys. Rev. A **64**, 022314 (2001); quant-ph/0011047.
[23] J. Preskill, *Lecture Notes for Physics 229: Quantum information and computation* (1998), URL http://www.theory.caltech.edu/people/preskill/ph229
[24] D. Gottesman, A. Kitaev, and J. Preskill, Phys. Rev. A **64**, 012310 (2001); quant-ph/0008040.
[25] J. Harrington and J. Preskill, Phys. Rev. A **64**, 062301 (2001); quant-ph/0105058.
[26] A. S. Holevo, Probl. Peredachi Inf., **8**, 62 (1972) [in Russian; English translation in Probl. Inf. Transm. **8**, 47 (1974)].
[27] A. S. Holevo, Rep. Math. Phys. **12**, 273 (1977).
[28] B. Schumacher, Phys. Rev. A **54**, 2614 (1996); quant-ph/9604023.
[29] K. Kraus, *States, Effects, and Operations*, Vol. 190 of *Lecture Notes in Physics* (Springer-Verlag, Berlin, 1983).
[30] M. Hamada, IEEE Trans. Inf. Theory **48** (2002); quant-ph/0112103.
[31] M. Aschbacher, *Finite Group Theory*, Vol. 10 of *Cambridge Studies in Advanced Mathematics* 2nd ed. (Cambridge University Press, Cambridge, 2000).
[32] A. Ashikhmin and E. Knill, IEEE Trans. Inf. Theory **47**, 3065 (2001); quant-ph/0005008.
[33] J. Bierbrauer and Y. Edel, J. Combinatorial Designs **8**, 174 (2000).
[34] R. Matsumoto and T. Uyeniatsu, IEICE Trans. Fundamentals **E83-A**, 1878 (2000); quant-ph/9911011.
[35] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).
[36] W. W. Peterson and E. J. Weldon, Jr., *Error-Correcting Codes*, 2nd ed. (MIT, Cambridge, 1972).
[37] E. Knill, quant-ph/9608048.
[38] E. M. Rains, IEEE Trans. Inf. Theory **45**, 1827 (1999); quant-ph/9703048.
[39] P. Kleidman and M. Liebeck, *The Subgroup Structure of the Finite Classical Groups*, Vol. 129 of *London Mathematical Society Lecture Notes Series* (Cambridge University Press, Cambridge, 1990).