Exponential lower bound on the highest fidelity achievable by quantum error-correcting codes

Mitsuru Hamada
Quantum Computation and Information Project (ERATO)
Japan Science and Technology Corporation, 5-28-3, Hongo, Bunkyo-ku, Tokyo 113-0032, Japan

(Dated: Sep. 21, 2001; Received, Phys. Rev.: Oct. 12, 2001; Published, Phys. Rev. A, 65, 052305, Apr. 15, 2002)

On a class of quantum channels which includes the depolarizing channel, the highest fidelity of quantum error-correcting codes of length n and rate R is proven to be lower bounded by $1 - \exp[-nE(R) + o(n)]$ for some function $E(R)$. The $E(R)$ is positive below some threshold $R_0$, which implies $R_0$ is a lower bound on the quantum capacity.

PACS numbers: 03.67.Lx, 03.67.Hk, 89.70.+c

I. INTRODUCTION

Quantum error-correcting codes (simply called codes in this paper) are deemed indispensable for quantum computation as schemes that protect quantum states from decoherence. An information theoretic problem relevant to such codes is one of determining the quantum capacity of a channel, which is far from settled \[6, 7, 8, 9\]. This paper treats a problem closely related to the quantum capacity. The corresponding problem in classical information theory is that of determining the highest fidelity of a channel, which is far from settled \[1, 2, 3, 4, 5\].

This paper employs an analog of minimum entropy decoding \[14\] together with the method of types from classical information theory \[3, 4, 8, 9\]. Note that Shannon’s coding theorem directly follows from (1) and (2).

Motivated by this classical issue, this paper presents an exponential lower bound on the highest possible fidelity of a code used on a class of quantum channels, which includes the depolarizing channel often discussed in the literature \[3, 4, 8, 9\]. This work was inspired by the recent result of Matsumoto and Uyematsu \[11\], who used an algebraic fact due to Calderbank et al. \[12\] \[Eq. (9) below\] to deduce a lower bound on the quantum capacity. This work’s approach resembles theirs in that both bounds are exponential in the data transmission rate of the code on channel $W$. The $E^*(R, W)$ actually equals $E_\ell(R, W)$ for relatively large rates $R$, and complete determination of $E^*(R, W)$ is one of the central issues in classical information theory \[3, 4, 8, 9\].

II. EXPONENTIAL BOUND ON FIDELITY

We follow the standard formalism of quantum information theory which assumes all possible quantum operations and state changes, including the effects of quantum channels, are described in terms of completely positive (CP) linear maps \[1, 2, 10\]. In this paper, only trace-preserving completely positive (TPCP) linear maps are treated. Given a Hilbert space $H$ of finite dimension, let $L(H)$ denote the set of linear operators on $H$. In general, every CP linear map $\mathcal{M} : L(H) \rightarrow L(H)$ has an operator-sum representation $\mathcal{M}(\rho) = \sum_{i \in \mathcal{I}} M_i \rho M_i^\dagger$ for some $M_i \in L(H), \ i \in \mathcal{I}$. When $\mathcal{M}$ is specified by a set of operators $\{M_i\}_{i \in \mathcal{I}}$, which is not unique, in this way, we write $\mathcal{M} \sim \{M_i\}_{i \in \mathcal{I}}$. Hereafter, $H$ denotes an arbitrarily fixed Hilbert space.
where the unitary operators $\{A_n : L(H^{\otimes n}) \rightarrow L(H^{\otimes n})\}$. We want a large subspace $C = C_n \subseteq H^{\otimes n}$ in which every state vector remains almost unchanged after the effect of a channel followed by the action of some suitable recovery process. The recovery process is again described as a TPCP linear map $R_n : L(H^{\otimes n}) \rightarrow L(H^{\otimes n})$. A pair $(C_n, R_n)$ consisting of such a subspace $C_n$ and a TPCP linear map $R_n$ is called a code and its performance is evaluated in terms of minimum fidelity \[ F(C_n, R_n, A_n) = \min_{|\psi\rangle \in C_n} \langle \psi | R_n A_n \langle \psi | \rangle \langle \psi |, \] where $R_n A_n$ denotes the composition of $A_n$ and $R_n$. Throughout, bras $\langle \cdot |$ and kets $| \cdot \rangle$ are assumed normalized. A subspace $C_n$ alone is also called a code assuming implicitly some recovery operator. Let $\{C_n, R_n\}$ be some memoryless channels in classical information theory, i.e., those $R_n$’s satisfying $R_n = R_n \circ R_n$. As is usual in information theory, the classical informational divergence or relative entropy is defined as $D(Q || P) = \sum_{x \in A} Q(x) \log \frac{Q(x)}{P(x)}$, and the minimization with respect to $Q$ takes over all probability distributions on $A$.

### Remarks

An immediate consequence of the theorem is that the quantum capacity of $A \subseteq H^{\otimes n}$ is lower bounded by $1 - H(P)$. To see this, observe that $E(R, P)$ is positive for $P' < 1 - H(P)$ due to the basic inequality $D(Q||P) \geq 0$ where equality occurs if and only if $Q = P'$. The bound $1 - H(P)$ appeared earlier in [6], Sec. 7.16.2.

Another direct consequence of the theorem is

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log d \{ 1 - F_{n, R_n} (A^{\otimes n}) \} \geq E(R, P),$$

which resembles (\ref{e:4}). In fact, we can see that $E(R, P)$ is closely related to $E_r(\bar{R}, W)$ in (\ref{e:3}) as follows. A specific form of $E_r$ is $E_r(\bar{R}, W) = \max_{v} E_r(\bar{R}, p, W)$, where

$$E_r(\bar{R}, p, W) = \min_{v} D(V||W[p]) + |I_0(p, v) - R|^{+}.$$

see (\ref{e:6}, \ref{e:3}) for detail. For consistency, we assume all logarithms appearing in the definitions of the rate of a code \( R \) and functions $I$ are to be base $d$. The function $E(R, P)$ coincides with $E_r(R + 1, p, W)$ with $p$ being the uniform probability distribution on $X$ and $W$ being the channel defined by $W(v|u) = P(v-u)$, $u, v \in X = \mathbb{F}_d^2$, where $\mathbb{F}_d$ denotes the finite field consisting of $d$ elements. Rewriting $E_r(\bar{R}, W)$ into the other well-known form (see (\ref{e:6}), pp. 168, 192–193, and (\ref{e:5}, \ref{e:3})), we have another form of $E$: $E(R, P) = \max_{0 \leq \delta \leq 1} -\delta (R-1) - (\delta + 1) \log_d \sum_{u \in X} P(u) \frac{1}{\delta}$. Furthermore, putting $\tilde{P}_d(u) = P(u) \frac{1}{\sum_{v \in X} P(v) \frac{1}{\delta}}$, $u \in X$, and $R_\delta = 1 - H(\tilde{P}_d)$, we obtain

$$E(R, P) = \left\{ \begin{array}{ll} -R + 1 - 2 \log_d \sum_{u} P(u) \frac{1}{\delta} & \text{if } 0 \leq R < R_1, \\ 0 & \text{if } R_1 \leq R < R_0, \\ \frac{1}{\delta} & \text{if } R_0 \leq R, \end{array} \right.$$ 

where $\delta^*$ is a $\delta$ with $R_\delta = \bar{R}$; see FIG. I.

### III. QUANTUM ERROR-CORRECTING CODES

To prove the theorem, we use a lemma on codes for quantum channels. We can regard the index of $N_{(i,j)} = X^t Z^t$, $(i,j) \in X$, as a pair of elements from the field $\mathbb{F} = \mathbb{F}_d = \mathbb{Z}/d\mathbb{Z}$. From these, we obtain a basis $N_i = \{N_x | x \in (\mathbb{F}^2)^n\}$ of $L(H^{\otimes n})$, where $N_x = N_{x_1} \otimes \cdots \otimes N_{x_n}$ for $x = (x_1, \ldots, x_n) \in (\mathbb{F}^2)^n$. We write $N_j$ for $\{N_x \in N_j | x \in J\}$ where $J \subseteq (\mathbb{F}^2)^n$. The index of a basis element $((u_1, v_1), \ldots, (u_n, v_n)) \in (\mathbb{F}^2)^n$
can be regarded as the plain $2n$-dimensional vector
\[ x = (u_1, v_1, \ldots, u_n, v_n) \in F^{2n}. \]

We can equip the vector space $F^{2n}$ over $F$ with a symplectic paring (bilinear form, or inner product) defined by
\[ (x, y)_{sp} = \sum_{i=1}^{n} u_i v'_i - v_i u'_i \]
for the above $x$ and $y = (u'_1, v'_1, \ldots, u'_n, v'_n) \in F^{2n}$.

Given a subspace $L \subseteq F^{2n}$, let
\[ L^\perp = \{ x \in F^{2n} \mid \forall y \in L, (x, y)_{sp} = 0 \}. \]

**Lemma 1** \([2]\) Let a subspace $L \subseteq F^{2n}$ satisfy $L \subseteq L^\perp$ and $\dim L = n-k$. Choose a set $J \subseteq F^{2n}$, not necessarily linear, such that
\[ \{ y - x \mid x \in J, y \in L \} \subseteq (L^\perp \setminus L)^c, \]
where the superscript $C$ denotes complement. Then, there exist $d^k$-dimensional $N_j$-correcting codes.

The codes in the lemma have the form $\{ \psi \in H^{\otimes n} \mid \forall M \in N_L, M \psi = \tau(M)\psi \}$ with some scalars $\tau(M)$, $M \in N_L$. A precise definition of $N_j$-correcting codes can be found in Sec. III of \([17]\) and the above lemma has been verified with Theorem III.2 therein. Most constructions of quantum error-correcting codes rely on this lemma, which is valid even if $d$ is a prime other than two \([13, 19, 21, 22]\).

Now, for a memoryless channel $A \sim \{ \sqrt{P(u)N_u} \}_{u \in X}$, and an $N_j$-correcting code $C \subseteq H^{\otimes n}$, write
\[ F(C) = \sup_{R_n} F(C, R_n A^{\otimes n}) \]
where $R_n$ ranges over all TPCP linear maps on $L(H^{\otimes n})$.

Then, since a recovery operator $R_n$ can be constructed explicitly so as to correct all errors in $N_j$, as in the proof of Theorem III.2 of \([17]\), we have
\[ 1 - F(C) \leq \sum_{x \notin J} P^n(x), \]
where we have written $P^n(x_1 \ldots x_n)$ for $P(x_1) \ldots P(x_n)$.

**IV. PROOF OF THEOREM 1**

We employ the method of types \([6, 14, 13]\), on which a few basic facts to be used are collected here. For $x = (x_1, \ldots, x_n) \in X^n$, define a probability distribution $P_x$ on $X$ by
\[ P_x(u) = |\{ i \mid 1 \leq i \leq n, x_i = u \}|/n, \quad u \in X, \]
which is called the type (empirical distribution) of $x$. With $X$ fixed, the set of all possible types of sequences from $X^n$ is denoted by $Q_n(X)$ or simply by $Q_n$. For a type $Q \in Q_n$, $T_Q^n$ is defined as $\{ x \in X^n \mid P_x = Q \}$. In what follows, we use
\[ |Q_n| \leq (n + 1)|X|^{-1}, \quad \forall Q \in Q_n, |T_Q^n| \leq d^{nH(Q)}, \]

Note that if $x \in X^n$ has type $Q$, then $P^n(x) = \prod_{u \in X} P(a)^{nQ(a)} = \exp\{ -n[H(Q) + D(Q||P)] \}$.

We apply Lemma 1 choosing $J$ as follows. Assume $\dim L = n - k$. Then, $\dim L^\perp = n + k$. From each of the $d^{n-k}$ cosets of $L^\perp$ in $F^{2n}$, select a vector that minimizes $H(P_x)$, i.e., a vector $x$ satisfying $H(P_x) \leq H(P_y)$ for any $y$ in the coset. This selection uses the idea of the minimum entropy decoder known in the classical information theory literature \([14]\). Let $J_0(L)$ denote the set of the $d^{n-k}$ selected vectors. If we take $J$ in Lemma 1 as $J(L) = \{ z + w \mid z \in J_0(L), w \in L \}$, the condition in the lemma is clearly satisfied. Let
\[ A = \{ L \subseteq F^{2n} \mid \text{L linear, } L \subseteq L^\perp, \dim L = n - k \}, \]
and for each $L \in A$, let $C(L)$ be an $N_j(L)$-correcting code existence of which is ensured by Lemma 1. Put
\[ T = \frac{1}{|A|} \sum_{L \in A} F(C(L)). \]

We will show that $T$ is bounded from below by $1 - \sum_{x \notin J(L)} P^n(x)$, which establishes the theorem. Such a method for a proof is called random coding \([6, 11, 13]\).

The $\{ 0, 1 \}$-valued indicator function $[T]$ equals 1 if and only if the statement $T$ is true and equals 0 otherwise. From (1), we have
\[ 1 - T \leq \frac{1}{|A|} \sum_{L \in \mathcal{A}} \sum_{x \notin J(L)} P^n(x) \]
\[ = \frac{1}{|A|} \sum_{L \in \mathcal{A}} \sum_{x \in F^{2n}} P^n(x) [x \notin J(L)] \]
\[ = \sum_{x \in F^{2n}} P^n(x) \frac{B(x)}{|A|}, \]
where we have put
\[ B(x) = \{ L \in \mathcal{A} \mid x \notin J(L) \}, \quad x \in F^{2n}. \]
The fraction $|B(x)|/|A|$ is trivially bounded as
\[
\frac{|B(x)|}{|A|} \leq 1, \quad x \in \mathbb{F}^{2^n}.
\] (8)

We use the next inequality \[11, 12\]. Let
\[A(x) = \{L \in A \mid x \in L^\perp \setminus L\}.
\]
Then, $|A(0)| = 0$ and
\[
\frac{|A(x)|}{|A|} \leq \frac{1}{d^{n-k}}, \quad x \in \mathbb{F}^{2^n}, \ x \neq 0.
\] (9)

Since $B(x) \subseteq \{L \in A \mid \exists y \in \mathbb{F}^{2^n}, H(P_y) \leq H(P_x), y - x \in L^\perp \setminus L\}$ from the design of $J(L)$ specified above (cf. \[3\]),
\[
|B(x)| \leq \sum_{y \in \mathbb{F}^{2^n} : H(P_y) \leq H(P_x), y \neq x} |A(y - x)| \leq \sum_{y \in \mathbb{F}^{2^n} : H(P_y) \leq H(P_x), y \neq x} |A|d^{-n+k},
\] (10)
where we have used \[\ref{R1}\] for the latter inequality. Combining \(\ref{R3}\), \(\ref{R2}\) and \(\ref{R4}\), we can proceed as follows with the aid of the basic inequalities in \(\ref{R1}\) and the inequality $\min\{a + b, 1\} \leq \min\{a, 1\} + \min\{b, 1\}$ for $a, b \geq 0$.

\[1 - F \leq \sum_{x \in \mathbb{F}^{2^n}} P^n(x) \min_{y \in \mathbb{F}^{2^n} : H(P_y) \leq H(P_x), y \neq x} \sum_{Q' \in Q_n : H(Q') \leq H(Q)} \exp_d[-nD(Q||P)] \sum_{Q' \in Q_n : H(Q') \leq H(Q)} \max_{Q' \in Q_n : H(Q') \leq H(Q)} \exp_d[-n|1 - R - H(Q')|^+] \exp_d[-n|1 - R - H(Q')|^+] \leq (n + 1)^{2(2^d-1)} \exp_d[-nE(R, P)],
\]
which is the promised bound.

V. CONCLUDING REMARK

This author conjectures that the bound in \[\ref{R5}\] is not tight in view of the existence of the Shor-Smolin codes \[\ref{R3}\].

Acknowledgments

The author would like to thank R. Matsumoto for valuable information on algebraic matters, discussions, comments, and especially, pointing out an error in the earlier manuscript, M. Hayashi and K. Matsumoto for helpful comments, and H. Imai for support.

[1] B. Schumacher, Phys. Rev. A 54, 2614 (1996), quant-ph/9604023.
[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. P. DiVincenzo, P. W. Shor, and J. A. Smolin, Phys. Rev. A 57, 830 (1998), quant-ph/9706061.
[4] H. Barnum, E. Knill, and M. A. Nielsen, IEEE Trans. Inf. Theory 46, 1317 (2000), quant-ph/9809010.
[5] P. W. Shor, Phys. Rev. A 52, R2493 (1995).
[6] R. G. Gallager, Information Theory and Reliable Communication (John Weily & Sons, NY, 1968).
[7] I. Csiszár and J. Körner, Information Theory: Coding Theorems for Discrete Memoryless Systems (Academic,
[8] R. G. Gallager, IEEE Trans. Inf. Theory 11, 3 (1965).
[9] S. Litsyn, IEEE Trans. Inf. Theory 45, 385 (1999).
[10] J. Preskill, Lecture notes for physics 229: Quantum information and computation (1998). URL http://www.theory.caltech.edu/people/preskill/ph229
[11] R. Matsumoto and T. Uyematsu, e-Print quant-ph/0105151, LANL (2001).
[12] A. R. Calderbank, E. M. Rains, P. W. Shor, and N. J. A. Sloane, Phys. Rev. Lett. 78, 405 (1997), quant-ph/9605005.
[13] V. D. Goppa, Problems of Information Transmission 10, 89 (1974).
[14] I. Csiszár and J. Körner, IEEE Trans. Inf. Theory 27, 5 (1981).
[15] I. Csiszár, IEEE Trans. Inf. Theory 44, 2505 (1998).
[16] M.-D. Choi, Linear Algebra and Its Applications 10, 285 (1975).
[17] E. Knill and R. Laflamme, Phys. Rev. A 55, 900 (1997), quant-ph/9604034.
[18] E. Knill, e-Print quant-ph/9608048, LANL (1996).
[19] E. Knill, e-Print quant-ph/9608049, LANL (1996).
[20] D. Gottesman, Phys. Rev. A 54, 1862 (1996), quant-ph/9604038.
[21] E. M. Rains, IEEE Trans. Inf. Theory 45, 1827 (1999), quant-ph/9703048.
[22] A. Ashikhmin and E. Knill, IEEE Trans. Inf. Theory 47, 3065 (2001), quant-ph/0005008.
[23] E. Artin, Geometric Algebra (Interscience Publisher, New York, 1957).
[24] M. Aschbacher, Finite Group Theory (Cambridge University Press, Cambridge, UK, 2000), 2nd ed.