Exponential lower bound on the Hilbert–Schmidt norm of quantum error-correcting codes (simply coded codes in a qubit domain)
Let $\mathcal{C}_n$ be the subspace of $\mathcal{R}_n$ consisting of such a subspace $\mathcal{K}_n$ and a TP-CP linear map $\mathcal{R}_n$ is called a code and its performance is evaluated in terms of minimum fidelity [3, 4, 16]

$$F(\mathcal{C}_n, \mathcal{R}_n; A_n) = \inf_{|\psi| \in \mathcal{C}_n} \langle \psi | \mathcal{R}_n; A_n (|\psiangle \langle \psi|) | \psi \rangle.$$ 

Bras $\{ | \rangle$ and kets $\langle | \}$ are assumed normalized. A subspace $\mathcal{C}_n$ alone is also called to code assuming implicitly some recovery operator. Let $F^{*}_{n,k}(A_n)$ denote the supremum of $F(\mathcal{C}_n, \mathcal{R}_n; A_n)$ such that there exists a code $\mathcal{C}_n, \mathcal{R}_n$ with $\log_d \dim \mathcal{C}_n \geq k$. This paper gives an exponential lower bound on $F^{*}_{n,k}(A_n)$, in the case where $\{ A_n \}$ is a slight generalization of the depolarizing channel specified as follows.

Fix an orthonormal basis $\{|0\rangle, \ldots, |d-1\rangle\}$ of $\mathcal{H}$. Put $\mathcal{X} = \{0, \ldots, d-1\}^2$ and $N_{(i,j)} = X^i Z^j$ for $(i,j) \in \mathcal{X}$, where the unitary operators $X, Z \in L(\mathcal{H})$ are defined by

$$X|j\rangle = |(j-1) \mod d\rangle, \quad Z|j\rangle = \omega^{(j^2/2)}|j\rangle$$

with $\omega$ being a primitive $d$-th root of unity [17, 18]. The $\{N_u\}_{u \in \mathcal{X}}$ is a basis of $L(\mathcal{H})$ and a generalization of the Pauli operators (including the identity) in that when $d = 2$, the basis $\{I, X, XZ, Z\}$ is the same as the set of Pauli operators up to a phase factor. For simplicity, we confine ourselves to treating analogs of what are called memoryless channels in classical information theory, i.e., those $\{A_n\}$ such that $A_n = A^{\otimes n}$, $n = 1, 2, \ldots$, for some $A : L(\mathcal{H}) \to L(\mathcal{H})$; such a channel $\{A^{\otimes n}\}$ is referred to as the memoryless channel $A$. In addition, we treat only channels that can be written as $A \sim \{\sqrt{P(u)} N_u \}_{u \in \mathcal{X}}$, where $P$ is a probability distribution on $\mathcal{X}$. This restriction is mainly due to that the codes to be proven to have the desired performance are symplectic (stabilizer, or additive) codes [11, 17, 18, 19, 20, 21], which exploit some algebraic property of the basis $\{N_u\}_{u \in \mathcal{X}}$, and that analysis of code performance naturally turns out to be easy for this class of channels. Analysis for a wider class of channels will be given in future papers.

As is usual in information theory, the classical informational divergence or relative entropy is denoted by $D$ and entropy by $H$ [6, 14]: for probability distributions $P$ and $Q$ on a finite set $\mathcal{X}$, $D(P \| Q) = \sum_{x \in \mathcal{X}} P(x) \log_d P(x) \log_d Q(x)$, and $H(Q) = -\sum_{x \in \mathcal{X}} Q(x) \log_d Q(x)$. This paper’s main result is

**Theorem 1** Let integers $n, k$ and a real number $R$ satisfy $0 \leq k \leq Rn$ and $0 \leq R < 1$ (a typical choice is $k = \lfloor Rn \rfloor$ for an arbitrarily fixed rate $R$). Then, for a memoryless channel $A \sim \{\sqrt{P(u)} N_u \}_{u \in \mathcal{X}}$, we have

$$F^{*}_{n,k}(A^{\otimes n}) \geq 1 - (n + 1)^{2(d-1)} d^{-n} R E(R, P)$$

where

$$E(R, P) = \min_Q \{D(Q \| P) + [1 - H(Q) - R]^+ \},$$

$[x]^+ = \max\{x, 0\}$, and the minimization with respect to $Q$ is taken over all probability distributions on $\mathcal{X}$.

**Remarks:** An immediate consequence of the theorem is that the quantum capacity $[1, 2, 3, 4]$ of $A$ is lower bounded by $1 - H(P)$. To see this, observe that $E(R, P)$ is positive for $R < 1 - H(P)$ due to the basic inequality $D(Q \| P) \geq 0$ where equality occurs if and only if $Q = P$ [6]. The bound $1 - H(P)$ appeared earlier in [9], Sec. 7.16.2.

Another direct consequence of the theorem is

$$\lim_{n \to \infty} \frac{1}{n} \min_{u \in \mathcal{X}} \sqrt[n]{\prod_{x \in \mathcal{X}} P(x)^{N(u,x)} A^{\otimes n}(u \otimes \bar{u})} \geq E(R, P) \geq \frac{1}{n} \min_{u \in \mathcal{X}} \prod_{x \in \mathcal{X}} P(x)^{N(u,x)} A^{\otimes n}(u \otimes \bar{u}),$$

where $N(u,x) = x^i z^j$ for $(i,j) \in \mathcal{X}$, which resembles (1). In fact, we can see that $E(R, P)$ is closely related to $E_\mathcal{C}(R, W)$ in (1) as follows. A specific form of $E_\mathcal{C}$ is $E_\mathcal{C}(R, W) = \max_p E_\mathcal{C}(R, p, W)$, where $E_\mathcal{C}(R, p, W)$ is defined by $W(x) = \frac{1}{n} \prod_{x \in \mathcal{X}} P(x)^{N(u,x)} A^{\otimes n}(u \otimes \bar{u})$. Rewriting $E_\mathcal{C}(R, p, W)$ into the other well-known form (see [6], pp. 168, 192–193, and [5, 7]), we have another form of $E$:

$$E(R, P) = \max_{0 \leq \delta \leq 1} \delta (R - 1) - (\delta + 1) \log_d \sum_{u \in \mathcal{X}} P(u)^{\frac{1}{\delta}}.$$ 

Furthermore, putting $\hat{P}_\delta(u) = P(u)^{\frac{1}{1-\delta}} \sum_{u \in \mathcal{X}} P(u)^{\frac{1}{\delta}}$, we obtain

$$E(R, P) = \begin{cases} R - 1 - 2 \log_d \sum_u P(u)^{\frac{1}{\delta}} & \text{if } 0 \leq R < R_1, \\ D(\hat{P}_\delta \| P) & \text{if } R_1 \leq R < R_0, \\ 0 & \text{if } R_0 \leq R, \end{cases}$$

where $\delta$ is a $\delta$ with $R_0 = R$; see FIG. 1.

To prove the theorem, we use a lemma on codes for quantum channels. We can regard the index of $N_{(i,j)} = X^i Z^j$, $(i,j) \in \mathcal{X}$, as a pair of elements from the field $F = \mathbb{F}_d = \mathbb{Z} / d\mathbb{Z}$, the finite field consisting of $d$ elements.
From these, we obtain a basis \( N_n = \{ N_x \mid x \in (F^2)^n \} \) of \( L(H^{\otimes n}) \), where \( N_x = N_{x_1} \otimes \cdots \otimes N_{x_n} \) for \( x = (x_1, \ldots, x_n) \in (F^2)^n \). We write \( N_f \) for \( \{ N_x \in N_n \mid x \in J \} \) where \( J \subseteq (F^2)^n \). The index of a basis element

\[
(\{ u_1, v_1 \}, \ldots, \{ u_n, v_n \} ) \in (F^2)^n,
\]

can be regarded as

\[
x = (u_1, v_1, \ldots, u_n, v_n) \in F^{2n}.
\]

We can equip the vector space \( F^{2n} \) over \( F \) a symplectic paring (bilinear form, or inner product) defined by

\[
(x, y)_{KP} = \sum_{i=1}^{n} u_i v_i' - v_i u_i'
\]

for the above \( x \) and \( y = (u_1', v_1', \ldots, u_n', v_n') \in F^{2n} [22, 23] \). Given a subspace \( L \subseteq F^{2n} \), let

\[
L^\perp = \{ x \in F^{2n} \mid \forall y \in L, (x, y)_{KP} = 0 \}.
\]

**Lemma 1** \([11]\) 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

\[
\{ y - x \mid x \in J, y \in J \} \subseteq (L^\perp \setminus L)^c,
\]

where the superscript \( C \) denotes complement. Then, there exist \( d^k \)-dimensional \( N_J \)-correcting codes.

The codes in the lemma have the form \( \{ \psi \in H^{\otimes n} \mid \forall M \in N_L, \psi = \sum_{M} \tau(M) \psi \} \) with some scalars \( \tau(M), M \in N_L \). A precise definition of \( N_J \)-correcting codes can be found in Sec. III of [16] 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 \([17, 18, 20, 21]\).

Now, for a memoryless channel \( A = \{ \{ P(y) \} \}_{y \in X} \) and an \( N_J \)-correcting code \( C \subseteq H^{\otimes n} \), write

\[
F(C) = \sup_{N_n} F(C, R_n, A^{\otimes n})
\]

where \( R_n \) ranges over all TPCP linear maps on \( L(H^{\otimes n}) \). Then, clearly,

\[
1 - F(C) \leq \sum_{x \notin J} P^0(x), \tag{5}
\]

where we have written \( P^0(x_1, \ldots, x_n) \) for \( P(x_1) \ldots P(x_n) \).

**Proof of Theorem 1.** We employ the method of type \([6, 13, 14]\), on which a few basic facts to be used are collected here. For \( x = (x_1, \ldots, x_n) \in \mathcal{X}^n \), define a probability distribution \( P_x \) on \( \mathcal{X} \) by

\[
P_x(u) = \frac{\{ i \mid 1 \leq i \leq n, x_i = u \}}{n}, \quad u \in \mathcal{X},
\]

which is called the type (empirical distribution) of \( x \). With \( \mathcal{X} \) fixed, the set of all possible types of sequences from \( \mathcal{X}^n \) is denoted by \( Q_n(\mathcal{X}) \) or simply \( Q_n \). For a \( Q \in Q_n \), \( T^Q_0 \) is defined as \( \{ x \in \mathcal{X}^n \mid P_x = Q \} \). In what follows, we use

\[
|Q_n| \leq (n+1)^{|\mathcal{X}|-1}, \quad \forall Q \in Q_n, \quad |T^Q_0| \leq e^{dH(Q)}. \tag{6}
\]

Note that if \( x \in \mathcal{X}^n \) has type \( Q \), then \( P^0(x) = \prod_{a \in \mathcal{X}} P(a)^n(Q(a)) \times \exp \{ -|n| H(Q) + |Q(a)||P(a)| \} \).

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 error decoder known in the classical information theory literature \([13]\). 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

\[
A = \{ L \subseteq F^{2n} \mid L \text{ linear, } L \subseteq L^\perp, \dim L = n-k \}
\]

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

\[
\mathcal{F} = \frac{1}{|A|} \sum_{L \in A} F(C(L)).
\]

We will show that \( \mathcal{F} \) is bounded from below by \( B_{H'}(b + 1)^{(d^2-1)d^{-n}E(R_P)} \), which establishes the theorem. Such a method for a proof is called random coding \([6, 10, 12]\).

The \( \{ 0, 1 \} \)-valued indicator function \( I(T) \) equals 1 if and only if the statement \( T \) is true and equals 0 otherwise. From (5), we have

\[
1 - \mathcal{F} \leq \frac{1}{|A|} \sum_{L \in A} \sum_{x \notin J(L)} P^0(x)
\]

\[
= \frac{1}{|A|} \sum_{L \in A} \sum_{x \in F^{2n}} P^0(x) \mathbb{1}[x \notin J(L)]
\]

\[
= \sum_{x \in F^{2n}} P^0(x) \frac{|B(x)|}{|A|}, \tag{7}
\]

where we have put

\[
B(x) = \{ L \in A \mid x \notin J(L) \}, \quad x \in F^{2n}.
\]

The fraction \( |B(x)|/|A| \) is trivially bounded as

\[
\frac{|B(x)|}{|A|} \leq 2, \quad x \in F^{2n}. \tag{8}
\]

We use the next inequality \([10, 11]\). Let

\[
A(x) = \{ L \in A \mid x \notin L \setminus J \}.
\]
Then, $|A(0)| = 0$ and

$$\frac{|A(x)|}{|A|} \leq \frac{1}{d^{n-k}}, \quad x \in F^{2^n}, \ x \neq 0.$$  \hspace{1cm} (9)

Since $B(x) \subseteq \{ L \in A \mid \exists y \in F^{2^n}, H(P_y) \leq H(P_x), y - x \in L \setminus L \}$ from the design of $J(L)$ specified above (cf. [12]),

$$|B(x)| \leq \sum_{y \in F^{2^n} : H(P_y) \leq H(P_x), \ y \neq x} |A(y - x)| \leq \sum_{y \in F^{2^n} : H(P_y) \leq H(P_x), \ y \neq x} |A|d^{-(n-k)}, \hspace{1cm} (10)$$

$$1 - \mathcal{F} \leq \sum_{x \in F^{2^n}} P^{(x)} \min \left\{ \sum_{y \in F^{2^n} : H(Q') \leq H(Q)} \frac{|T(y)|}{P^{(1-h)}} , \ 1 \right\} \leq \sum_{Q \in \mathcal{Q}_n} \left( \prod_{x \in X} P(a)^{n(Q(x))} \right) \min \left\{ \sum_{Q' \in \mathcal{Q}_n} \frac{|T(y)|}{P^{(1-h)}} , \ 1 \right\} \leq \sum_{Q \in \mathcal{Q}_n} \exp[nD(Q||P)] \sum_{Q' \in \mathcal{Q}_n} \exp[n] \max_{H(Q') \leq H(Q)} \exp[n] \leq \sum_{Q \in \mathcal{Q}_n} \exp[n] \max_{H(Q) \leq H(Q')} \exp[n] \leq (n+1)^{(d^2-1)} \exp[n] \mathbb{E}(R, P),$$

which is the desired bound. Q.E.D.

This author conjectures that the bound in (4) is not tight in view of the existence of the Shor-Smolins codes [3].

The author would like to thank R. Matsumoto for valuable information on algebraic matters, discussions, comments, and especially, pointing out an error in the author's unpublished proof of a fact used in the earlier manuscript, use of which was circumvented in the present paper. M. Hayashi and K. Matsumoto for helpful comments, and H. Inoue for support.

---

* Electronic address: mitsuru@ieee.org