Quantum state transmission over partially corrupted quantum information network

Masahito Hayashi$^{1,2,3,4,*}$ and Seunghoan Song$^{4,†}$

1Shenzhen Institute for Quantum Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China
2Guangdong Provincial Key Laboratory of Quantum Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China
3Shenzhen Key Laboratory of Quantum Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China
4Graduate School of Mathematics, Nagoya University, Chikusa-ku, Nagoya 464-8602, Japan

(Received 6 November 2019; accepted 30 June 2020; published 15 July 2020)

Quantum communication over network is particularly appealing in a similar way to classical network communication, but a partially corrupted quantum network has not been sufficiently studied. We discuss quantum state transmission over a partially corrupted quantum network, in which the sender transmits $m_0$ qudits, i.e., $d$-level systems, to the receiver via network, but $m_1$ channels are corrupted. As our result, we show that the optimal transmission rate is at least $(m_0 - 2m_1 + 1) \log d$ under the following two settings. In the first case, the unitaries on intermediate nodes are arbitrary and the corruptions on the $m_1$ channels are individual. In the second case, the unitaries on intermediate nodes are restricted to Clifford operations and the corruptions on the $m_1$ channels are adaptive, i.e., they are caused by an attack with a quantum memory. Further, our code in the second case realizes the noiseless communication even with the single-shot setting and is constructed dependently only on the network topology and the places of the $m_1$ corrupted channels.

DOI: 10.1103/PhysRevResearch.2.033079

I. INTRODUCTION

In the realm of quantum communication, a quantum network can be expected to work a basic infrastructure for quantum communication [1]. A quantum network is composed of edges, i.e., quantum communication channels to transmit quantum states, and nodes, i.e., quantum processors to convert input quantum states to output quantum state. Several coding schemes for a quantum network have been proposed for communication across many players [2–10]. A typical quantum network coding scheme was proposed in Ref. [3], and was already implemented experimentally as a photonic system [11]. Since a quantum network has several communication paths, it enhances robust communication between two players by the diversification of the risk when the quantum communication network is partially corrupted. Although the robust network communication against corruption has been discussed in the classical case by many papers [12–15], this type of research for quantum network are very limited, and only the preceding paper [16] showed the existence of such a network code on the restricted class of quantum networks. Unfortunately, their code construction is not practical due to their asymptotic construction, and the optimality of their transmission rate is not discussed.

This paper derives a simple expression for the optimal value of the reliable transmission rate, which is called the quantum capacity, on a more general class of quantum networks with corruption. In fact, no single-letterized formula is known for the quantum capacity of most of the channels because its existing formula requires a limiting expression [17–20]. Therefore, it is a quite difficult problem to derive the quantum capacity under a certain condition. Also, we propose a reliable quantum network code that realizes a perfect error correction under single transmission without the asymptotic construction when the quantum network satisfies a certain condition. In addition, our code achieves the optimal transmission rate, which is strictly higher than the achievable rate in the preceding paper [16]. While our code construction depends on the network structure and the place of the corruption, it does not depend on the corrupting operations. Whereas conventional network coding considers the optimization of node operations given a directed graph of the network, we consider the quantum capacity when node operations are given as well because it is often quite difficult to control node operations. Specifically, we address the worst-case capacity under a certain condition, which is formulated as follows.

Our model is described as follows. Every quantum channel transmits a $d$-dimensional system by one use of the network. The sender has $m_0$ outgoing quantum channels. Each intermediate node has the same number of incoming quantum channels and outgoing quantum channels. The node applies a fixed unitary across the incoming quantum systems and outputs them to the outgoing quantum channels. Finally, the receiver receives $m_0$ quantum systems via $m_0$ incoming quantum channels. The existing classical networks are composed of noiseless channel by applying error-correcting codes. Hence, existing studies on classical network coding address...
the network composed of noiseless channels unless they are corrupted [12–14]. Therefore, this paper also assumes that the channels are noiseless unless they are corrupted because the errors of these noncorrupted channels can be corrected by quantum error-correcting code. In particular, due to the no-cloning theorem of quantum system, if the dimension of output system is larger than the input system, we cannot avoid the corruption of quantum system. That is, such a node operation is not useful for a quantum network. Hence, each intermediate node is assumed to have the same dimensional output system as the input system. Also, the number of corrupted channels is assumed to be \( m_1 \), and the network is assumed to have no cycle and to be well synchronized, i.e., to have no delayed transmission. Only the sender and the receiver are allowed to optimize their coding operation due to the difficulty of node operation control. Although the paper [16] assumed that the sender and the receiver do not know the places of corrupted channels, the places are assumed to be known to them in our setting.

There are two types of quantum corruptions. The first one is the individual corruption, in which the corruption on each corrupted quantum channel is done individually. If the channel is simply broken due to an accident, the corruption belongs to this type. The other is the adaptive corruption, in which the corruptions on respective corrupted quantum channels are done adaptively. That is, the corruptions are caused by the attack with a quantum memory, and the quantum memory interacts with the corrupted quantum channel on each corruption. When the corruption is done by an adversarial attack, we need to consider the adaptive corruption. This kind of corruption can be written by quantum comb [21,22]. The adaptive corruption is more general than the individual corruption, and the adaptive corruption cannot be reduced to the individual corruption in general. For example, adaptive strategies cannot be reduced to individual strategies in quantum channel discrimination [23].

In this scenario, when all nodes carefully maintain their connected channels, they can find which channels are corrupted. It needs a larger cost to replace the corrupted installed channels by new noiseless channels because these corrupted channels are across physically distant nodes and the replacement of the channels takes longer times and requires many preparations. Also, the intermediate nodes might have physical space limitation so that the node operations are fixed. Hence, as a realistic solution, it is natural to employ the encoding and the decoding on the sender and receiver sides.

In the classical communication over the partially corrupted network, we can show that the capacity, i.e., the maximum transmission rate, is not smaller than \( (m_0 - m_1) \log d \) [13]. In contrast, when our quantum channel has only individual corruptions, we find that the quantum capacity, i.e., the maximum transmission rate of the quantum state, is not smaller than \( (m_0 - 2m_1 + 1) \log d \). This fact is shown by the analysis of coherent information on the quantum network. Further, when the unitaries on intermediate nodes are limited to Clifford operations, the quantum capacity is not smaller than \( (m_0 - 2m_1 + 1) \log d \) even when the corruptions are adaptive while the achievable rate in the preceding paper [16] is \( (m_0 - 2m_1) \log d \).

In this case, our code can be constructed in the single-shot setting by using Clifford operations. This construction depends only on Clifford operations of intermediate nodes and the places of the corrupted channels, and is independent of Eve’s operation to the corrupted channels. Our code construction to achieve the capacity can be intuitively explained in the case of Clifford operations as follows.

We can characterize the \( m_1 \) corruptions by the symplectic structure as follows. When we apply a suitable symplectic matrix, the first corruption can be written as the corruption on the first qudit, i.e., the corrupted computation base and the corrupted Fourier base are given as the first base. However, at the other corruptions, the corrupted computation bases and the corrupted Fourier bases split in general. Hence, in the worst case, totally \( 2m_1 - 1 \) quantum systems are corrupted. Since the symplectic diagonalization extracts noncorrupted bases, it enables us to propose our code construction to achieve the capacity.

The remaining parts of this paper are organized as follows. First, for a comparison, Sec. II discusses the classical network. Next, Sec. III addresses the general unitary network model with the first type of corruption. Then, with the second type of corruption, Sec. IV considers the network model whose node operation is limited to Clifford operations. As its special class, this section also discusses the basis-linear network model, in which the unitary operation is given as a linear operation with respect to the basis. Finally, Sec. V describes a concrete example of a network model.

II. CLASSICAL NETWORK MODEL

When all the node operations are invertible and linear on a finite field of order \( q (=d) \), the receiver can find a linear subspace for corrupted information, as discussed in Ref. [13]. The dimension of the subspace is bounded by \( m_1 \). Hence, the capacity is not smaller than \( (m_0 - m_1) \log d \). However, when node operations are not necessarily linear but are invertible, we cannot apply the above discussion. Even in this case, we can show that the capacity is greater than or equal to \( (m_0 - m_1) \log d \) as follows.

Assume that every channel transmits a system whose number of elements is \( d \) by one use of the network and the sender has \( m_0 \) outgoing channels. Node operations are not necessarily linear but are invertible. \( m_1 \) channels are corrupted. Hence, we can assume that \( m_1 \) corruptions are done sequentially. Let \( X_{i-1}', \ldots, X_i' \) and \( X_0, \ldots, X_{m_0} \) be the whole information before and after the \( i \)th corruption, respectively. We denote the input and output information of this network by \( X_0 \) and \( X_{m_0}' \), respectively. Then, \( X_i' \) is written as \( f_i(X_i) \) by using an invertible function \( f_i \). Since the channel capacity of classical communication is given by the maximum mutual information between the input information and the output information, it is sufficient to show that

\[
I(X_0; X_i') = I(X_0; X_i) \geq (m_0 - m_1) \log d \tag{1}
\]

with a certain distribution \( P_{X_0} \) of \( X_0 \).

033079-2
Now, we set the distribution $P_{X_0}$ of $X_0$ to be the uniform distribution, which implies that

$$H(X_0) = m_0 \log d.$$  \hspace{1cm} (2)

From the network structure, we find the relation $H(X_0 | X_{i+1}) \leq \log d$. The chain rule of conditional entropy implies that

$$H(X_0 | X_m) \leq H(X_0, \ldots, X_{m-1} | X_m)$$

$$= \sum_{i=0}^{m_i-1} H(X_i | X_{i+1} X_{i+2} \ldots X_m)$$

$$\leq \sum_{i=0}^{m_i-1} H(X_i | X_{i+1}) \leq m_1 \log d.$$ \hspace{1cm} (3)

The combination of (2) and (3) yields (1).

### III. GENERAL UNITARY NETWORK MODEL

The general unitary network model is described as follows. In this model, we assume that the places of the channels to be corrupted are known. Since our network is composed of unitary operations and partial corruptions, our network model of the adaptive corruption is given as the general form with Fig. 1, whose reason is illustrated in Fig. 2. The input and output systems are the $m_0 \times m_0$-tensor product system $\mathcal{H}^{\otimes m_0}$ of the same system $\mathcal{H}$ of dimension $d$, and $m_0 + 1$ unitaries $U = (U_0, U_1, \ldots, U_{m_1})$ are applied between the input and output systems, which has $m_1$ intervals. Eve can access only the first system on each interval, and has her memory so that the corruption in the $i$th interval is given as the unitary $\tilde{U}_i$ between her memory and the corrupted system, i.e., the first system on the $i$th interval.

Our first result is on the minimum capacity of the general unitary network with individual corruption, in which Eve is assumed to have no memory. Hence, her operation on the $i$th interval can be written as TP-CP maps $\Gamma = (\Gamma_1, \ldots, \Gamma_{m_i})$ as Fig. 3. In this case, the channel between the input and output systems is denoted by $\Lambda(U, \Gamma)$. $C(\Lambda)$ expresses the quantum capacity of a quantum channel $\Lambda$. The following theorem characterizes the quantum capacity under the individual corruption.

**Theorem 1.** The minimum quantum capacity is given as follows:

$$\min C(\Lambda(U, \Gamma)) = (m_0 - 2m_1 + 1) \log d.$$ \hspace{1cm} (4)

Here, the minimum is taken over all channels $\Lambda(U, \Gamma)$ under the individual corruption.
where the maximum is taken over all the input densities on the n-tensor system of the input system of $\Lambda$ \cite{17-20, 25, Theorem 9.10}.

**Lemma 1.** An individual corruption $\Lambda(U, \Gamma)$ satisfies

$$\max_{\rho} I(\rho, \Lambda(U, \Gamma)) \geq (m_0 - 2m_1 + 1) \log d.$$  

(6)

The inequality $\geq$ in Eq. (4) follows from Lemma 1, the additivity property $I_1(\rho^\otimes n, \Lambda^\otimes n) = nI_1(\rho, \Lambda)$, and the above capacity formula (5). In the remaining sections, we will show the existence of a channel $\Lambda(U, \Gamma)$ to satisfy the equality of (4) in the Clifford network model, whose formulation will be given in the next section. Since a Clifford network is a special case of unitary networks, this existence completes the proof of Theorem 1.

**Proof of Lemma 1.** It is sufficient to show the case when $U_0$ is the identity matrix. Let $\rho_{\text{mix}, m_0-1}$ be the completely mixed state on $H^\otimes m_0-1$. We set the initial state to be $|0\rangle\langle 0| \otimes \rho_{\text{mix}, m_0-1}$.

Consider the time after the unitary $U_{i-1}$ is applied but $\Gamma_i$ is not applied yet. At this time, we denote the systems to be attacked and the remaining systems as $A_i$ and $B_i$, respectively. After the application of $\Gamma_i$, we denote the systems to be attacked and the remaining systems as $A_i'$ and $B_i'$, respectively. We consider Steinspring representation $\tilde{U}_i$ of $\Gamma_i$, in which the output of the environment is $E_i$. Figure 4 summarizes the relation among the systems $A_{i-1}', B_{i-1}', A_i, B_i, A_i', B_i', \text{and} E_i$.

Since the state on $A_i' E_i$ is pure, we have $H(A_i') = H(E_i)$. Hence, we have

$$H(A_i' B_i') - H(E_i) = H(A_i') + H(B_i') - H(E_i)$$

$$= H(B_i') = (m_0 - 1) \log d.$$  

(7)

As shown later, for $i = 2, \ldots, m_1$, we have

$$H(A_i' B_i') - H(E_i) \geq H(A_{i-1}' B_{i-1}') - 2 \log d.$$  

(8)

Combining (7) and (8), we have

$$I_1(|0\rangle\langle 0| \otimes \rho_{\text{mix}, m_0-1}, \Lambda(U_0, \Gamma_1, U_1, \ldots, \Gamma_m, U_m))$$

$$= H(A_{m_0}' B_{m_0}') - H(E_1, \ldots, E_{m_1})$$

$$\geq (m_0 - 2m_1 + 1) \log d.$$  

(9)

Now, we show (8). Consider the purification of $\rho_{A_i E_i}$ by using the reference system $R$. Then, $H(R) = H(A_i') = H(E_i)$. Since $H(A_1) + H(A_1') - H(E_i) = H(R) + H(A_1') - H(RA_1') = I(R; A_1') \geq 0$, we have

$$H(E_i) \leq H(A_i) + H(A_i').$$  

(10)

Thus,

$$H(A_i' B_i') - H(E_i)$$

$$\geq H(A_i' B_i') - H(A_i) - H(A_i')$$

$$= H(B_i') + H(A_i') - H(A_i) - H(A_i') - I(A_i'; B_i')$$

$$= H(A_i' B_i') - H(A_i E_i) + I(A_i' E_i; B_i')$$

$$+ H(A_i') - H(A_i) - H(A_i') - I(A_i'; B_i')$$

$$= H(A_i B_i) - H(A_i) + I(E_i; B_i | A_i') - H(A_i)$$

$$\geq H(A_i B_i) - 2H(A_i) = H(A_{i-1}' B_{i-1}') - 2H(A_i)$$

$$\geq H(A_{i-1}' B_{i-1}') - 2 \log d.$$  

(11)

Here, (a) follows from (10), (b) follows from $H(A_i' B_i | E_i) = H(A_i B_i)$ and $H(A_i' E_i) = H(A_i)$, and (c) follows from $I(E_i; B_i | A_i') \geq 0$.

**IV. CLIFFORD NETWORK MODEL**

**A. Fundamental limit of quantum capacity in Clifford network model**

As pointed out in Ref. [13], in the classical network, when node operations are composed of linear operations, the corruption can be corrected by linear operations in the sender and receiver sides. As a quantum version of a network composed of linear operations, we introduce Clifford network models. In this model, we can expect a simple construction of error correction. Indeed, we can construct a code to achieve the capacity even with the single-shot setting in this case. Also, Theorem 1 can be extended to adaptive corruptions described in Fig. 1. (A concrete example of an adaptive corruption is given in Sec. V.) For this aim, we prepare several notations. Given a prime power $q = p^k$, our Hilbert space $H$ is assumed to be spanned by the computational basis $|i\rangle_{i\in F_p}$, where $F_p$ is the algebraic extension of the finite field $F_p$ with degree $d_q$. That is, the dimension of the Hilbert space $H$ is assumed to be $q$. Then, for $s, t \in F_p$, we define the generalized Pauli operators $X(s)$ and $Z(t)$ as $X(s) := \sum_{x \in F_q} |x + s\rangle \langle x|$ and $Z(t) := \sum_{x \in F_q} |x\rangle \langle x|$, where $\omega := e^{2\pi \sqrt{-1}/p}$. Here, for an element $z \in F_p$, $tr_z$ expresses the element $tr_M_z$ in $F_q$, where $M_z$ denotes the matrix representation of the multiplication map $x \mapsto zx$ with identifying the finite field $F_q$ with the vector space $F_{q^2}$. We define the Fourier basis $|y\rangle_F := |H\rangle_{i\in F_q} \in F_{q^2}$ of the computational basis $|x\rangle_{i\in F_q} \in F_{q^2}$ as

$$|y\rangle_F := \sum_{x \in F_q} \frac{1}{\sqrt{q}} \omega^{yx} |x\rangle.$$

To consider our network model, for vectors $s = (s_1, \ldots, s_n), t = (t_1, \ldots, t_n) \in F_{q^n}$, we define the operators $X(s)$ and $Z(t)$ on the $n$-fold tensor product system $H_{q^n}$ as $X(s) := X(s_1) \otimes \cdots \otimes X(s_n)$ and $Z(t) := Z(t_1) \otimes \cdots \otimes Z(t_n)$. Then, the discrete Weyl operator is defined as $W(s, t) := X(s)Z(t)$. Then, for $(s, t), (s', t') \in F_{q^n}$, we define the skew-symmetric matrix $J$ on $F_{q^n}$ and the
inner product \( \langle (s, t), (s', t') \rangle \in \mathbb{F}_p \) as 
\[
J = \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}
\]
with 
\[
\langle (s, t), (s', t') \rangle := \sum_{i=1}^r (s_i t_i' + t_i s_i').
\]
Then, the commutation relation
\[
W(s, t) W(s', t') = e^{ij(s,t)} J^{(s, t')} W(s', t') W(s, t)
\]
is called a symplectic matrix when \( \langle (s, t), J(s, t') \rangle = \langle g(s, t), J(g(s', t')) \rangle \) for \( (s, t), (s', t') \in \mathbb{F}_q^2 \).

Next, we introduce Clifford group as a subset of the set \( U(H^{2m}) \) of unitaries on \( H^{2m} \). Using the set \( W := \{ cW(s, t) \} 1 \leq (s, t) \in \mathbb{F}_q^2 \) we define the Clifford group \( C \) as
\[
C := \{ U \in U(H^{2m}) | U W U^{-1} = W \}
\]
An element of \( C \) is called a Clifford unitary.

For any element \( U \in C \), there exists a symplectic matrix \( g \) such that
\[
U W(s, t) U^{-1} = cW(g(s, t))
\]
for \( (s, t) \in \mathbb{F}_q^2 \) with a complex number \( c \) satisfying \( |c| = 1 \). Conversely, for any symplectic matrix \( g \), there exists a unitary \( U \) to satisfy (13). A typical construction of such a unitary \( U \) is given in [2.5, Sec. 8.3]. This construction is called metaplectic representation and is denoted by \( U(g) \) in this paper.

Now, the input and output systems are assumed to be \( H^{2m} \), and the unitary \( U \) is to be an element of Clifford group. Such a network is called Clifford network. We choose a symplectic matrix \( g \) as \( U(g) = U \). The minimum capacity of Clifford networks is derived as follows.

Theorem 2. For Clifford network, the minimum quantum capacity is
\[
m_q = 2m_1 + 1 \log q \quad \text{in the adaptive corruption},
\]
ie, the case when the corruption is caused by the attack with a quantum memory.

Since a Clifford network model is a special case of channels mentioned in Theorem 1, Theorem 2 shows the existence of a channel \( (U, \Gamma) \) to satisfy the equality of (4) in Theorem 1. Hence, Theorem 2 implies the remaining part of Theorem 1.

B. Parametric characterization of quantum capacity

To show Theorem 2, we introduce additional parameters to characterize the quantum capacity. For this aim, we describe the behaviors of the errors in the terms of the symplectic structure. The errors are described by vectors in \( \mathbb{F}_q^{2m} \). For this aim, we introduce notations and parameters of the network as follows. Let \( e_j \) be the vector in \( \mathbb{F}_q^{2m} \) that has only \( 1 \) nonzero element \( 1 \) in the \( j \)th entry. Using \( e_1 \) and \( e_m+1 \), we define \( 2m_1 \) vectors \( v_1, \ldots, v_{2m_1} \in \mathbb{F}_q^{2m} \) as
\[
v_i := g_1^{-i} \cdots g_{m_1}^{-i} e_1
\]
and \( \nu_{m_1+i} := g_1^{-i} \cdots g_{m_1}^{-i} e_{m_1+1} \) for \( i = 1, \ldots, m_1 \). Since \( e_1 \) and \( e_{m_1+1} \) describe the directions of errors in the respective interval, all the directions in the linear space \( V \) spanned by \( v_1, \ldots, v_{2m_1} \) are corrupted in this whole network.

In this context, when a matrix \( P \) satisfies \( P^2 = P = \text{Im} P \), it is called a projection onto \( V \). Then, we choose a projection \( P_V \) onto \( V \). Since \( P_{V'} \) is also an antisymmetric matrix, the rank of \( P_{V'} J P_{V'} \) is an even number. The rank of the matrix \( P_{V'} J P_{V'} \) is equal to the rank of matrix \( \langle (v_i, J v_i) \rangle \). Hence, the rank of \( P_{V'} J P_{V'} \) does not depend on the choice of the projection \( P_V \) onto \( V \) while the choice of the projection \( P_{V'} \) onto \( V \) is not unique. With these observations, we define the integers \( m_1 \) and \( m_+ \) as
\[
m_1 = \text{rank } P_{V'} J P_{V'} / 2 \text{ and } m_+ = \dim V - m_1
\]
As shown by using fundamental knowledge of stabilizer codes [26–29][30, Sec. 5.3] in Appendix B, the quantum capacity \( C \) is characterized as follows.

Lemma 2. The capacity \( C \) is lower bounded as
\[
log q \leq C \leq \log q \quad \text{in the adaptive case, i.e., the case when the corruption is caused by the attack with a quantum memory.}
\]
Proof of Lemma 2. To consider this problem, we employ a stabilizer code [26–29][30, Sec. 5.3] after applying the unitary \( U^{-1} U \) to the received system. Define \( V_1 := \text{ker} P_{V'} J P_{V'} \cap V = V \cap V_1^⊥ \), where \( V_1^⊥ := \{ x \in \mathbb{F}_q^{2m} | (x, y_J) = 0 \} \). Due to symplectic diagonalization, we can choose \( m_1 \)-independent vectors \( w_1, \ldots, w_m \in V \subset \mathbb{F}_q^{2m} \) and other \( m_+ \)-independent vectors \( w_1', \ldots, w_m' \in V \subset \mathbb{F}_q^{2m} \) such that \( \langle w_i, J w_i \rangle = \delta_{i,j} \), \( \langle w_i, J w_j \rangle = 0 \) for \( i, j = 1, \ldots, m_m \). Hence, by symplectic diagonalization, see Appendix A, we define \( V_2 \) and \( V_3 \) as the spaces spanned by \( w_1, \ldots, w_m \) and \( w_1', \ldots, w_m' \), respectively. Then, we define the subspace \( N := V_1 + V_2 \) with dimension \( m_1 \). We find that \( N \) is self-orthogonal, i.e., \( N \cap N^⊥ \), where \( N^⊥ := \{ x \in \mathbb{F}_q^{2m} | (x, y) = 0 \} \). Define the projection \( \sigma : \mathbb{F}_q^{2m} \rightarrow \mathbb{F}_q^{2m}/N^⊥ \). Since \( N_1 \cap N_3 = \{ 0 \} \), the map \( \sigma \) is injective on \( V_1 \). That is, there is an error map \( \tau \) from \( \mathbb{F}_q^{2m}/N^⊥ \) to \( \mathbb{F}_q^{2m} \) such that \( N_3 \subset \text{Im } \tau \), i.e., \( \tau(\sigma(v)) = v \) for any \( v \in V_3 \). When we apply the stabilizer code with the correcting set \( \text{Im } \tau \), the set of correctable errors is \( \text{Im } \tau + N \), which includes \( V \). (The detail of this error-correcting code is given in Appendix B.) That is, any error in \( V \) can be corrected. In this code, the logarithm of the dimension of the transmitted space is \( \frac{2m_1}{2m} \). This discussion shows Lemma 2. Our code construction depends only on \( g_0, \ldots, g_{m_1} \). That is, it is independent of the remaining unitaries \( U, \ldots, U_m \) of Eve’s corruption.

The lower bound of Lemma 2 is achieved by the following lemma.

Lemma 2. When Eve changes the state on the corrupted edge to the completely mixed state, the capacity \( C \) equals \( (m_0 - m_+) \log q \).

Proof of Lemma 2. Now, we consider the case when Eve changes the state on the corrupted edge to the completely mixed state. It is equivalent to the case when all errors \( \sigma(\{ x \in \mathbb{F}_q^{2m} | (x, y) = 0 \} \) happen with equal probability. Then, we focus on the spectral decomposition \( P_{V'} \langle y \in N \rangle / \langle y \in N \rangle \) of the common eigenspace of \( \{ W(x) \}_{x \in N} \) as \( W(x) = \sum_{y \in \mathbb{F}_q^{2m}/N^⊥} \langle y \rangle_{x} P_{V'} \langle y \rangle \). The coherence between different eigenvectors in operators in \( N \) is collapsed. The rank of \( P_{V'} \langle y \rangle \) does not depend on \( y \) \( \langle y \rangle \in \mathbb{F}_q^{2m}/N^⊥ \) and equals \( q^{m_0 - m_+} \). Hence, the application of Lemma 8 to the subspace \( \langle y \rangle \) given in the proof of Lemma 2 with \( E = V_3 \) shows Lemma 3. That is, it is impossible to transmit a space larger than \( \text{Im } P_{V'} \langle y \rangle \).

C. Range of network parameters and basis-linear network model

As described in Lemmas 2 and 3, the capacity is characterized by rank \( P_{V'} J P_{V'} \) and \( \dim V \). The range of these network parameters can be summarized as the following lemma.
Lemma 4. The following conditions are equivalent for two integers $l_i$ and $l_{ss}$.

1. $2m_i - 1 > l_i > l_{ss} > 1$ and $m_0 > l_{ss}$.
2. There exists a sequence of Clifford unitaries $U(g_0), \ldots, U(g_{m_0})$ such that $2l_i = \text{rank } P_{V_i}^T J P_V$ and $l_i + l_{ss} = \dim V$.

Applying Lemma 4 to Lemmas 2 and 3 for the case of $m_{ss} = l_{ss} = 2m_i - 1$, we obtain Theorem 2. Lemma 4 will be shown in Sec. IV D. Although the relation (2)⇒(1) of Lemma 4 can be directly shown, the proof of the opposite direction needs the detailed network construction stated in condition (2). For this aim, we introduce a special class of Clifford networks, called basis-linear networks.

In basis-linear networks, we assume that each Clifford unitary $U_i$ is characterized as the basis exchange caused by an invertible matrix $\tilde{g}_i$ on $\mathbb{F}_2^{m_0}$, which is similar to the case of CSS (Calderbank-Shor-Steane) code [31,32]. That is, the Clifford unitary $U_i$ is given as the unitary $\tilde{U}(\tilde{g}_i)$ defined by $\tilde{U}(\tilde{g}_i)\mathbf{x} = [\tilde{g}_i]\mathbf{x}$. Its action on the Fourier basis $\{|y\}^F_{\mathbb{F}_2^{m_0}}$ is characterized as $\tilde{U}(\tilde{g}_i)|y\rangle_F = \langle \tilde{g}_i |y\rangle_F$, where $\langle \tilde{g}_i |y\rangle_F$ is defined as the transpose $(\tilde{g}_i^{-1})^T$ of the inverse matrix $\tilde{g}_i^{-1}$ [16, Appendix A]. Hence, we have

$$\tilde{U}(\tilde{g}_i) = U\left[\begin{array}{cc} \tilde{g}_i & 0 \\
0 & \tilde{g}_i^T_F \end{array}\right].$$

Let $\tilde{e}_i$ be the vector in $\mathbb{F}_2^{m_0}$ that has only one nonzero element 1 in the $i$th entry. By using the vector $\tilde{e}_i = (1, 0, \ldots, 0) \in \mathbb{F}_2^{m_0}$, the vectors $v_1, \ldots, v_{2m_0}$ are written as $v_i = (\tilde{e}_i, 0)$ and $v_{m_0+i} = (0, \tilde{e}_i)$ with $\tilde{e}_i := \tilde{g}_0^{-1} \tilde{g}_{i-1}^{-1} \tilde{e}_1$ and $\tilde{v}_i := \tilde{g}_0^{-1} \tilde{g}_{i-1}^{-1} \tilde{v}_1$ for $i = 1, \ldots, m_1$. We define the matrices $V_i$ and $V_i^\top$ as $(v_1, \ldots, v_{m_0})$ and $(\tilde{v}_1, \ldots, \tilde{v}_{m_0})$. Then, we have

$$m_i = \text{rank} (V_i^\top)^T V_i, \quad m_{ss} = \text{rank} V + \text{rank} V^\top - m_{ss}. \quad (15)$$

Introducing new parameters $(l_1, l_2, l_3)$ satisfying $l_i = l_1 + l_2 - l_3$, we can characterize the possible range of these parameters under this submodel as follows.

Lemma 5. The following conditions are equivalent for three integers $l_1, l_2,$ and $l_3$.

(a) $m_1 \geq l_1 \geq l_2 \geq l_3 \geq 1$, and $m_0 \geq l_1 + l_2 - l_3$.

(b) There exists a sequence of invertible matrices $\tilde{g}_0, \ldots, \tilde{g}_{m_0}$ over finite field $\mathbb{F}_2$ such that rank $V = l_1$, rank $V^\top = l_2$, and rank $(V^\top)^T V = l_3$.

Lemma 5 will be shown in Sec. IV D as well as Lemma 4. Our proof of (1)⇒(2) of Lemma 4 will be given by using (a)⇒(b) of Lemma 5.

D. Proofs of Lemmas 4 and 5

Our proofs of Lemmas 4 and 5 are composed of the parameter constraint part and the network construction part. The parameter constraint part proves the relations (2)⇒(1) of Lemma 4 and (b)⇒(a) of Lemma 5. The network construction part proves (a)⇒(b) of Lemma 5. Lastly, we show the relation (1)⇒(2) of Lemma 4 using Lemma 5.

1. Parameter constraint part

The relations (2)⇒(1) of Lemma 4 and (b)⇒(a) of Lemma 5 show the constraint for the range of parameters $l_i$ and $l_{ss}$ to characterize our network. Before starting our proof of these relations, we prepare a more general discussion. Remind that $m_i$ and $m_{ss}$ are defined as $m_i = \text{rank } P_{V_i}^T J P_V$ and $m_{ss} = \dim V - m_i$. Since the rank of the submatrix $(|u_i, v_j\rangle)_{i=1,m_i+1}$ is 2, the rank of $(|u_i, v_j\rangle)_{i=1,m_i+1}^{2m_0}$ is at least 2. As the rank of $(|u_i, v_j\rangle)_{i=1,m_i+1}^{2m_0}$ equals the rank of $P_{V_i}^T J P_V$, the inequality

$$m_i \geq 1$$

holds. Thus, the inequality $\dim V \geq \text{rank } P_{V_i}^T J P_V$ implies

$$m_i \geq m_{ss}. \quad (17)$$

So, since $2m_i \geq m_i + m_{ss}$, we have

$$m_{ss} \leq 2m_i - 1. \quad (18)$$

Now we prove (2)⇒(1) of Lemma 4. The inequality $l_{ss} \geq l_i$ follows from (17), and the inequality $l_i \geq 1$ follows from (16). Since rank $P_{V_i}^T J P_V = 2l_i$, there exist $l_{ss}$ independent vectors $x_1, \ldots, x_{l_{ss}} \in V$ such that $(x_1, x_2) = 0$ for $i, j = 1, \ldots, l_{ss}$. Since the number of such vectors is upper bounded by $m_0$, we have $m_0 \geq l_{ss}$. The relation $2m_i - 1 \geq l_{ss}$ follows from (18). Hence, we have the relation (2)⇒(1) of Lemma 4.

Next, we show (b)⇒(a) of Lemma 5. Since the relations $m_1 \geq l_1 \geq l_2 \geq l_3$ are trivial, it is sufficient to show $m_0 \geq l_1 + l_2 - l_3$ and $l_1 \geq 1$. Since the matrix $(V_i^\top)^T V_i$ is not zero, we have $l_1 \geq 1$. Hence, the relations (15) and (18) imply the relation $m_0 \geq l_1 + l_2 - l_3$ as

$$m_0 \geq 2m_i - 1 \geq m_{ss} = \text{rank } V + \text{rank } V^\top - m_{ss},$$

$$\text{rank } V + \text{rank } V^\top - \text{rank } (V_{i+1})^T V_i = l_1 + l_2 - l_3.$$

Thus, we obtain the relation (b)⇒(a) of Lemma 5.

2. Network construction part

Now, we show (a)⇒(b) of Lemma 5. That is, we concretely construct our network based on the parameters $l_1, l_2,$ and $l_3$. We assume that $l_1 \geq l_2$. Otherwise, we can exchange the computation basis and the Fourier basis. We choose $g_{m_0}$ to be the identity matrix. For $i = 1, \ldots, m_1$, we choose $\tilde{g}_{i-1}$ to be $A_i^\top A_{i-1}$, where the matrix $A_i$ is defined as follows. The constructed matrices $V, V_i, (V_i^\top)^T V_i$ are described in Fig. 5.

$A_0$ is the identity matrix. For $i = 1, \ldots, l_1$, we define $A_1$ as the transposition between the first entry and the $i$th entry. For $i = l_1 + 1, \ldots, m_1$, we define $A_i$ as the identity matrix. For $i = 1, \ldots, l_2 - l_3$, we define $A_{i+1}$ in the following way. For the first, $(l_3 + 2i - 1)$th, and $(l_3 + 2i)$th entries, it is defined as $A_{i+1} = A_{i+1} = 1$. For other indices $i$, $j$, the matrix component $(A_{i+1})_{i,j}$ is defined as $\delta_{i,j}$.

For $i = 1, \ldots, l_2 - l_3$, we define $A_{i+1}$ in the following way. For the first and $(2l_2 - 1 + i)$th entries, it is defined as (1 0 0). For other indices $i$, $j$, the matrix component $(A_{i+1})_{i,j}$ is defined as $\delta_{i,j}$.

Then, for $i = 1, \ldots, l_3$, we have $\tilde{v}_i = \tilde{e}_i$. For $i = 1, \ldots, l_2 - l_3$, we have $\tilde{v}_{i+1} = \tilde{e}_i + \tilde{e}_{i+2l_2-1}$. For $i = 1, \ldots, l_2 - l_3$, we have $\tilde{v}_{i+1} = \tilde{e}_i + \tilde{e}_{2l_2-1+i}$. For $i = l_1 + 1, \ldots, m_1$, we have $\tilde{v}_i = \tilde{e}_i$.

The matrix $[A_{i+1}]_{i,j}$ is characterized for $i = 1, \ldots, l_2 - l_3$ as follows. For the first, $(l_3 + 2i - 1)$th, and $(l_3 + 2i)$th
entries, it is given as \(\begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}\). For other indices \(j, j'\), 
\([A_{l_2+i}]_{\ell, j'}\) is given as \(\delta_{j, j'}\).

The matrix \([A_{l_2+i}]_{\ell, j'}\) is characterized for \(i = 1, \ldots, l_1 - l_2\) as follows. For the first and \((2l_2 - l_3 + i)\)th entries, it is given as \(\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}\). For other indices \(j, j'\), the matrix component \([A_{l_2+i}]_{\ell, j'}\) is given as \(\delta_{j, j'}\). Then, for \(i = 1, \ldots, l_3\), we have \(\bar{v}_{i} = \bar{e}_{i}\). For \(i = 1, \ldots, l_3\), \(l_1 - l_2\), we have \(\bar{v}_{l_1+i} = \bar{e}_{i} - \bar{e}_{l_1+i}\). For \(i = 1, \ldots, l_3\), \(l_1 - l_2\), we have \(\bar{v}_{l_1+i} = \bar{e}_{i}\). For \(i = 1, \ldots, m_1\), we have \(\bar{v}_{i} = \bar{e}_{1}\). Therefore, we have rank \(\bar{V} = l_1\) and rank \(\bar{V}' = l_2\).

Also, when \(j = 2, \ldots, l_3\) or \(j' = 2, \ldots, l_3\), we have \((\bar{V})^\top \bar{V})_{j, j'} = \delta_{j, j'}\). When \(j = 1, l_3 + 1, \ldots, m_1\) and \(j' = 1, l_3 + 1, \ldots, m_1\), we have \((\bar{V})^\top \bar{V})_{j, j'} = 1\). Hence, \(\text{rank}(\bar{V})^\top \bar{V} = l_3\). Thus, we obtain the relation \((a) \Rightarrow (b)\) of Lemma 5.

3. Proof of \((1) \Rightarrow (2)\) of Lemma 4

We choose \((l_1, l_2, l_3) := (1, l_1 + l_{\omega n} - 1, l_3)\). Then, \((l_1, l_2, l_3)\) satisfies condition \((a)\) of Lemma 4. We also have \((a) \Rightarrow (b)\) from Sec. IV D 2. From (15), the condition \((b)\) of Lemma 4 implies the existence of Clifford network with \((m_0, m_\omega) := (l_1, l_1 + l_2 - l_3) = (l_1, l_\omega)\), which implies \((1) \Rightarrow (2)\) of Lemma 5. Therefore, our proof of Lemmas 4 and 5 is completed.

V. CONCRETE EXAMPLE FOR ADAPTIVE CORRUPTION

To clarify the usefulness of our result, we discuss a concrete example for the worst case with \(m_0 = 4\) and \(m_1 = 2\) in the framework of Fig. 1. Also, we show a concrete example for adaptive corruption in this case. Consider the case given in Lemma 5 with \(d = 2\), \(m_0 = 4\), and \(m_1 = 2\). That is, we focus on the case given in Sec. IV D 2 with \(l_1 = l_2 = 2\) and \(l_3 = 1\). This discussion is also helpful to understand the construction given in Sec. IV D 2.

As shown in Fig. 6, we set \(U_0 = U_2 = I\) and \(U_1 = \text{CNOT}_{1 \rightarrow 2} \text{CNOT}_{3 \rightarrow 1}\). In this case, the capacity is 1, and only the fourth qubit is noiseless. Hence, the use of the fourth qubit is the best error correction to achieve capacity.

As a typical example of Eve’s corruption, this figure illustrates the following adaptive corruption on a network. Eve corrupts each channel between the sender and the intermediate node and between the intermediate node and the receiver. Eve has one qubit system as her memory \(E\). Eve applies the swap gate on \(H_1 \otimes E\) at the first corruption and the CNOT gate on \(E \otimes H_1\), where \(H_1\) is the first qubit system.
Finally to see that this case realizes the worst capacity, we show the relations $l_1 = l_2 = 2$ and $l_3 = 1$. We notice that
\begin{equation}
A_0 = A_1 = I, \quad A_2 = \begin{pmatrix} 1 & 0 & 1 \\
1 & 1 & 0 \\
0 & 0 & 1 \end{pmatrix}.
\end{equation}
\begin{equation}
\tilde{g}_0 = \tilde{g}_2 = I, \text{ and the first three entries of } \tilde{g}_1 = A_1A_2^{-1} \text{ are}
\begin{pmatrix} 1 & 0 & 1 \\
1 & 1 & 1 \\
0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\
1 & 1 & 0 \\
0 & 0 & 1 \end{pmatrix}.
\end{equation}
Since $\tilde{g}_1$ and $\tilde{g}_1^{-1}$ are
\begin{equation}
\tilde{g}_1 = \begin{pmatrix} 1 & 0 & 1 \\
1 & 1 & 0 \\
0 & 0 & 1 \end{pmatrix}, \quad \tilde{g}_1^{-1} = \begin{pmatrix} 1 & 0 & 1 \\
1 & 1 & 0 \\
0 & 0 & 1 \end{pmatrix},
\end{equation}
we have
\begin{equation}
\tilde{V} = \begin{pmatrix} 1 & 1 \\
0 & 1 \\
0 & 0 \end{pmatrix}, \quad \tilde{V}^\prime = \begin{pmatrix} 1 & 0 \\
0 & 1 \\
0 & 0 \end{pmatrix}, \quad (\tilde{V}^\prime)^\top \tilde{V} = \begin{pmatrix} 1 & 1 \\
1 & 1 \end{pmatrix},
\end{equation}
which implies $m_{ss} = \text{rank } \tilde{V} + \text{rank } \tilde{V}^\prime - \text{rank } (\tilde{V}^\prime)^\top \tilde{V} = 3$.
Since $m_0 - 2m_1 + 1 = m_0 - m_{ss} = 1$, the capacity of this network is the minimum capacity. The optimal code for this case is to transmit the state only on the fourth channel of the network regardless of Eve’s operations $U_1, U_2$.

VI. DISCUSSION

We have shown that the quantum capacity is not smaller than $(m_0 - 2m_1 + 1) \log d$ when the sender has $m_0$ outgoing channels, the receiver has $m_0$ incoming channels, each intermediate node applies invertible unitary, only $m_1$ channels are corrupted in our quantum network model, and other noncorrupted channels are noiseless. Our result holds with the following two cases. In the first case, the unitaries on intermediate nodes are arbitrary and the corruptions on the $m_1$ channels are individual. In the second case, the unitaries on intermediate nodes are restricted to Clifford operations and the corruptions on the $m_1$ channels are adaptive, i.e., they are caused by the attack allowed to have a quantum memory. Further, our code in the second case realizes the noiseless communication even with the single-shot setting, and depends only on the node operations, the network topology, and the places of the $m_1$ corrupted channels. That is, it is independent of Eve’s operation on the $m_1$ corrupted channels. This code utilizes the following structure of this model. The error in the first corrupted channel can be concentrated to one quantum system. However, the errors of the computation basis and the Fourier basis in another corrupted channel split into two quantum systems in general. Hence, $2m_1 - 1$ quantum systems are corrupted in the worst case. The first case has been shown by the analysis of the coherent information, and symplectic structure including symplectic diagonalization on the discrete system plays a key role in the second case. It is an interesting remaining problem to derive the quantum capacity when the operations on intermediate nodes are arbitrary unitaries and the corruptions on the $m_1$ channels are adaptive.

ACKNOWLEDGMENTS

M.H. is grateful to Professor M. Owari and Dr. G. Kato for helpful comments. M.H. was supported in part by Guangdong Provincial Key Laboratory (Grant No. 2019B121203002), Japan Society for the Promotion of Science (JSPS) Grant-in-Aid for Scientific Research (A) Grants No. 17H01280, (B) No. 16K0017, and Kayamori Foundation of Informational Science Advancement. S.S. is supported by Lotte Foundation and JSPS Grant-in-Aid for JSPS Fellows No. JP20J11484.

APPENDIX A: SYMPLECTIC DIAGONALIZATION

For the choice of vectors $w_1, \ldots, w_m$, and $w'_1, \ldots, w'_m$, we summarize fundamental knowledge for symplectic diagonalization in the finite-dimensional system. Assume that $\mathcal{V}$ is a finite-dimensional vector space over a finite field $\mathbb{F}_q$. We consider a bilinear form $Q$ from $\mathcal{V} \times \mathcal{V}$ to $\mathbb{F}_q$. In the following discussion, $\mathcal{V}$ is not necessarily nondegenerate. Given an element $v \in \mathcal{V}$, $Q(v, \cdot)$ can be regarded as an element of the dual space $\mathcal{V}^*$ of $\mathcal{V}$. In this sense, $Q$ can be regarded as a linear map from $\mathcal{V}$ to $\mathcal{V}^*$. A bilinear form $Q$ from $\mathcal{V} \times \mathcal{V}$ to $\mathbb{F}_q$ is called antisymmetric when $Q(v_1, v_2) = -Q(v_2, v_1)$ and $Q(v_1, v_1) = 0$ for $v_1, v_2 \in \mathcal{V}$.

**Lemma 6.** Consider an antisymmetric bilinear form $Q$.

Then, the rank of $Q$ is an even number $2k$. There exists a basis $w_1, w_2, w'_1, \ldots, w'_k \in \mathcal{V}$ such that $Q(w'_i, w'_j) = \delta_{ij}$ and $Q(w_i, w_j) = Q(w'_i, w'_j) = 0$ for $i, j = 1, \ldots, k$.

This statement can be generalized as follows.

**Lemma 7.** Consider an antisymmetric bilinear form $Q$.

Then, the rank of $Q$ is an even number $2k$. Assume that $l + l'$ linearly independent vectors $w_1, \ldots, w_l, w'_1, \ldots, w'_l \in \mathcal{V}$ with $l \leq l' \leq k$ satisfy $Q(w'_i, w'_j) = \delta_{ij}$ and $Q(w_i, w_j) = Q(w'_i, w'_j) = 0$ for $i, j = 1, \ldots, l$ and $i', j' = 1, \ldots, l'$.

There exist $2k - l - l'$ vectors $w_{l+1}, \ldots, w_k, w'_{l+1}, \ldots, w'_{k} \in \mathcal{V}$ such that $Q(w_i, w_{j}) = \delta_{ij}$ and $Q(w_i, w_{j}) = Q(w'_i, w'_j) = 0$ for $i, j = 1, \ldots, k$.

**Proof of Lemma 7.** First, we show that there exists a vector $w_{l+1} \in \mathcal{V}$ such that $Q(w'_i, w_{l+1}) = \delta_{i,l+1}$ and $Q(w_i, w_{l+1}) = Q(w'_i, w_{l+1}) = 0$ for $i = 1, \ldots, l$ and $l' = 1, \ldots, l'$. Since the number of constraints is $2l$ and is smaller than the dimension of the dual space of $\mathcal{V}$, we can choose such a vector. Since any linear combination of $w_{l+1}, \ldots, w_k, w'_{l+1}, \ldots, w'_{k}$ does not satisfy this condition, $w_{l+1}$ is linearly independent of $w_1, \ldots, w_l, w'_1, \ldots, w'_k$.

When $l' > l$, it is sufficient to choose one vector $w_{l+1}$ to satisfy $Q(w_{l+1}, w_i) = Q(w_{l+1}, w'_i) = 0$ for $i = 1, \ldots, l'$. Since the number of constraints is $l + l'$ and is smaller than the dimension, we can choose such a vector. Since any linear combination of $w_1, \ldots, w_l, w'_1, \ldots, w'_l$ does not satisfy this condition, $w_{l+1}$ is linearly independent of $w_1, \ldots, w_l, w'_1, \ldots, w'_l$.

When $l' = l$, we exchange vectors $w_{l+1}, \ldots, w_k$ and $w'_1, \ldots, w'_{l'}$ and repeat the above procedure. Otherwise, we repeat the above procedure. Therefore, we can choose the desired vectors inductively.
APPENDIX B: STABILIZER CODE

1. Construction of stabilizer code

First, we summarize the fundamental knowledge for the stabilizer code on the system $\mathcal{H}^{\otimes m}$. Here, we assume that the Weyl operator $W(x)$ is defined in the same way as Sec. IV. Although it was formulated in Refs. [26–29], the following discussion is based on the notation in [30, Sec. 5.3]. When a subspace $N \subset \mathbb{F}_q^{2 m}$ satisfies the condition $N \subset N^\perp := \{ x \in \mathbb{F}_q^{2m} | \langle x, y \rangle = 0 \text{ for any } y \in N \}$, it is called self-orthogonal. We consider the spectral decomposition $P_{y}|N\rangle\langle y| = \sum_{y \in \mathbb{F}_q^{2m}/N^\perp} c_{y|x} P_{y}|N\rangle\langle y|$. Here, $c_{y|x}$ is the eigenvalue and $P_{y}|N\rangle\langle y|$ is the projection to the common eigenspace.

Given a self-orthogonal subspace $N \subset \mathbb{F}_q^{2 m}$, we choose a map $\tau$ from $\mathbb{F}_q^{2 m}/N^\perp$ to $\mathbb{F}_q^{2 m}/N^\perp$ such that $\tau(y) = [y] \in \mathbb{F}_q^{2 m}/N^\perp$. Then, the stabilizer code with the correcting set $\mathrm{Im} \tau$ is generated as follows. The sender sets the initial state in the subspace $\mathrm{Im} P_{y}|N\rangle\langle y|$. After receiving the system, the receiver applies the projective measurement $P_{y}|N\rangle\langle y|\mathbb{F}_q^{2 m}/N^\perp$. Then, when we observe $[y] \in \mathbb{F}_q^{2 m}/N^\perp$, the resultant state belongs to $\mathrm{Im} P_{y}|N\rangle\langle y|$. We apply the unitary $W(-\tau(y))$ so that the state is transferred from $\mathrm{Im} P_{y}|N\rangle\langle y|$ to $\mathrm{Im} P_{\tau(y)}|N\rangle\langle y|$. When the error is $W(\tau(y))$, it can be corrected in this error correction. Also, when the error belongs to $\{W(x) | x \in N\}$, it does not change the subspace $\mathrm{Im} P_{y}|N\rangle\langle y|$. Hence, when the error belongs to $\{W(\tau(y)) + x) | [y] \in \mathbb{F}_q^{2 m}/N^\perp, x \in N\}$, it can be corrected.

2. Capacity

We consider another space $E \subset \mathbb{F}_q^{2 m}$ such that $E \cap N^\perp = \{0\}$ and $E \subset E^\perp$. We define the projection $\sigma : \mathbb{F}_q^{2 m} \rightarrow \mathbb{F}_q^{2 m}/N^\perp$. The map $\sigma$ is injective on $E$. Hence, we choose $\tau$ to satisfy $E \subset \mathrm{Im} \tau$.

Now, we consider the case when all errors in $N^\perp \subset E$ happen with equal probability. To address this case, we focus on the noisy channel $\rho \mapsto \Lambda(\rho) := \sum_{x \in \mathbb{F}_q^{2 m}/N\perp} W(x) \rho W(x)^{\dagger}$. In this case, an error $x$ in $E$ makes the state from $\mathrm{Im} P_{y}|N\rangle\langle y|$ to $\mathrm{Im} P_{x}|N\rangle\langle y|$. Then, as shown later, we have

$$\Lambda(\rho) = \frac{1}{|E|} \sum_{[y] \in \mathbb{F}_q^{2 m}/N^\perp} W(x) P_{y}|N\rangle\langle y| W(x)^{\dagger}. \quad (B1)$$

To derive (B1), we prepare several notations. Let $k$ and $l$ be the dimensions of $E$ and $N$, respectively. From the definition, we have $k \leq l$. Let $w_0^i, \ldots, w_j^i$ be a basis of $E$, which automatically satisfies the condition that $\langle w_i^j, w_j^i \rangle = 0$ for $i, j = 1, \ldots, k$. Since the $k$ nonzero linear maps $x \mapsto (x, w_i^j), \ldots, x \mapsto (x, w_j^i)$ on $N$ are linearly independent, considering the dual basis of the above linear maps, we can choose $l$ vectors $w_i^j, w_0^j$ of $N$ such that $\langle w_i^j, w_j^i \rangle = \delta_{ij}$, $\langle w_0^j, w_i^j \rangle = 0$ for $i, j = 1, \ldots, l$ and $j = 1, \ldots, k$. According to Lemma 7, we choose vectors $w_0^1, \ldots, w_0^l, w_0^{l+1}, \ldots, w_0^{2 m}$ in $\mathbb{F}_q^{2 m}$. Then, we choose a symplectic matrix $g$ such that $g_{ij} = w_i^j$ and $g_{i+j} = w_i$ for $i = 1, \ldots, m$, where $e_i$ is defined in Sec. IV. Hence, $U(g)|cW(x) | x \in E, |c| = 1|U(g)^{-1}$

equals $\{cW_0(0, t) | t \in \mathbb{F}_q^l, |c| = 1\}$, where $H_A := H_{\otimes l}$ and $W_A(x)$ is the discrete Weyl operator on $H_A$. With a $k$-dimensional subspace $E \subset \mathbb{F}_q^l$, $U(g)|cW(x) | x \in E, |c| = 1|U(g)^{-1}$

equals $\{cW_0(s, 0) | s \in E, |c| = 1\}$. We define the system $H_B := H_{\otimes m-1}$ so that $H_{\otimes m} = H_A \otimes H_B$, and denote the discrete Weyl operator on $H_B$ by $W_0(x)$. In this notation, $P_{\tau(y)}|N\rangle\langle y|$ is the same dimensional system as $H_B$.

Now, instead of the original system $H_{\otimes m}$, we focus on the system $H_A \otimes H_B$ by applying the unitary $U(g)$. The noisy channel $A$ is unitarily equivalent to $A_A \otimes I_B$, where $I_B$ is the noisecase channel on $H_B$ and $A_A$ is given as

$$\sum_{s \in \mathbb{F}_q^l} \frac{1}{d_k} W_A(s,0) W_A(0, t) \rho W_A(0, t)^\dagger W_A(s,0)^\dagger = \sum_{s \in \mathbb{F}_q^l} \frac{1}{d_k} \sum_{t \in \mathbb{F}_q^l} W_A(s,0) |t\rangle |\rho(t)| |t\rangle |W_A(s,0)^\dagger. \quad (B2)$$

That is, the relation $U(g)^{\dagger} U(g) \rho U(g)^{\dagger} U(g) = (A_A \otimes I_B)(\rho)$ holds for any state $\rho$. The symbol $[y] \in \mathbb{F}_q^{2 m}/N^\perp$ in Eq. (B1) corresponds to $t \in \mathbb{F}_q^l$ in Eq. (B2), and $P_{\tau(y)}|N\rangle\langle y|$ in Eq. (B1) corresponds to $|t\rangle |I_B$. Hence, $U(g)(A_A \otimes I_B) U(g)^{\dagger} U(g)^{\dagger} U(g) = (A_A \otimes I_B)$ equals the right-hand side of (B1), which shows (B1).

Since the channel $A_A$ in $H_A$ is entanglement breaking, it is impossible to transport a noisecase system from $H_B$, i.e., $\mathrm{Im} P_{\tau(y)}|N\rangle\langle y|$. Since the above stabilizer code protects the space $\mathrm{Im} P_{\tau(y)}|N\rangle\langle y|$, the maximum dimension of the correctable space is $d = \dim N$.

Now, we consider the case of $n$ times use of this channel, i.e., $A_{\otimes n}$, which is unitarily equivalent to $(A_A \otimes I_B)^{\otimes n} = A_A^{\otimes n} \otimes I_B^{\otimes n}$. Since $A_A$ is an entanglement-breaking channel, using Lemma 9 given below, we have

$$\max_{\rho} I(\rho, A_{\otimes n}) = \max_{\rho} I(\rho, A_A^{\otimes n} \otimes I_B^{\otimes n}) = \max_{\rho} I(\tau, \rho^{\otimes n}) = n \max_{\tau} I(\tau, \rho^{\otimes n}) = n(m_0 - \dim N) \log q,$$

where $I(\rho, A)$ is the coherent information. Hence, we obtain the following lemma.

Lemma 9. When all errors in $E \subset N^\perp$ happen with equal probability, the capacity $C$ equals $(m_0 - \dim N) \log q$.

3. Lemma for calculation of capacity

Lemma 9. When a channel $A_A$ is entanglement breaking, a channel $A_B$ satisfies the condition

$$\max_{\rho} I(\rho, A_A \otimes A_B) = \max_{\rho} I(\tau, A_B). \quad (B3)$$

Proof of Lemma 9. Let $A$ and $B$ ($A'$ and $B'$) be the input (output) systems of $A_A$ and $A_B$, respectively. We choose a state $\rho_A B$ on $A B$. Let $C$ be the reference system of the state $\rho_{ABC}^{AB}$ so that $\rho_{ABC}^{AB}$ is the purification of $\rho_{ABC}^{AB}$. Let $\rho'$ be the output system on the whole system of $A', B'$, and $C$.

Since $A_A$ is entanglement breaking, it is written as $A_A(\sigma) = \sum_{\rho_{ABC}^{AB}} \mathrm{Tr} M_{ABC} \rho_{ABC}^{AB}$, where $\{M_A\}$ is a PVM and $\mathrm{rank} M_A = 1$. Hence, $A_A(\rho_{ABC}^{AB})$ is written as $\sum_{\rho_{ABC}^{AB}} \mathrm{Tr} M_{ABC} \rho_{ABC}^{AB} \otimes \rho_{PABC}^{ABC}$, where $\mathrm{rank} \rho_{PABC}^{ABC} = 1$. Then, we denote $A_A(\rho_{ABC}^{AB}) = \rho_{PABC}^{ABC}$. The coherent information $I(\rho_{ABC}^{AB} | A_B \otimes A_B)$ equals $D(\rho_{ABC}^{AB} || \rho_{A'B'C}^{ABC} \otimes I_C)$.
which is evaluated as
\[
D(\rho'_{AB'C} \| \rho'_{AB} \otimes I_C) = D \left( \sum_a P_{A=a} \rho'_{B|A=a} \otimes \rho_{A=a} \right) \\
\leq D \left( \sum_a P_{A=a} \rho'_{B|A=a} \otimes |a\rangle\langle a| \otimes I_C \right) \\
\leq \sum_a P_{A=a} D(\rho'_{B|A=a} \| \rho'_{B|A=a} \otimes I_C). \quad (B4)
\]

The inequality follows from the information processing inequality for the map $|a\rangle\langle a| \mapsto \rho_{A=a}$.

Since rank $\rho_{B'C|A=a} = 1$, the state $\rho_{B'C|A=a}$ is a purification of $\rho_{B|A=a}$. Thus, $D(\rho'_{B'C|A=a} \| \rho'_{B|A=a} \otimes I_C)$ equals the coherent information $I_c(\rho'_{B|A=a}, \Lambda_B)$. Hence, we have
\[
I_c(\rho_{AB}, \Lambda_A \otimes \Lambda_B) \leq \sum_a P_{A=a} I_c(\rho_{B|A=a}, \Lambda_B), \quad (B5)
\]
which implies that
\[
\max_{\rho} I_c(\rho, \Lambda_A \otimes \Lambda_B) \leq \max_{\rho} I_c(\tau, \Lambda_B). \quad (B6)
\]

Next, we show the converse inequality
\[
\max_{\rho} I_c(\rho, \Lambda_A \otimes \Lambda_B) \geq \max_{\rho} I_c(\tau, \Lambda_B). \quad (B7)
\]
For any state $\tau$ on the system $B$, define $\rho_{AB} = \rho_A \otimes \tau$ where $\rho_A$ is a pure state. Then, we have
\[
I_c(\rho_{AB}, \Lambda_A \otimes \Lambda_B) = I_c(\tau, \Lambda_B). \quad (B8)
\]
Therefore, we obtain (B7). 