

Channel Polarization of Two-dimensional-input Quantum Symmetric Channels

Zhengzhong Yi¹, Zhipeng Liang¹ and Xuan Wang^{1*}

¹Harbin Institute of Technology, Shenzhen. Shenzhen, 518055,
China.

*Corresponding author(s). E-mail(s): wangxuan@cs.hitsz.edu.cn;
Contributing authors: zhengzhongyi@cs.hitsz.edu.cn;
liangzhipenghitsz@163.com;

Abstract

Being attracted by the property of classical polar code, researchers are trying to find its analogue in quantum fields, which is called quantum polar code. The first step and the key to design quantum polar code is to find out for the quantity which can measure the quality of quantum channels, whether there is a polarization phenomenon which is similar to classical channel polarization. Coherent information is believed to be the quantum analogue of classical mutual information and the quantity to measure the capacity of quantum channel. In this paper, we define a class of quantum channels called quantum symmetric channels, and prove that for quantum symmetric channels, under the similar channel combining and splitting process as in the classical channel polarization, the maximum single letter coherent information of the coordinate channels will polarize. That is to say, there is a channel polarization phenomenon in quantum symmetric channels.

Keywords: Quantum symmetric channels, Coherent information, Basis transition probability matrix, Channel polarization

1 Introduction

The potential to solve different problems more efficiently than the state-of-the-art classical computing makes quantum computing attract worldwide attention. To give full play to this potential, quantum computers should have sufficient reliable qubits. However, at present, physical qubits are quite vulnerable, which restricts the development of large-scale fault-tolerant quantum computing and exploiting the advantages of quantum computing. Fortunately, quantum error correcting codes (QECCs) discovered by Shor and Steane provide us with a solution to this problem[1, 2].

Similar to classical error correcting codes (CECCs), QECCs encoding n (which is called **code length**) less reliable physical qubits (with error rate p_0) in a certain way to obtain k ($k < n$) more reliable logic qubits (with error rate $p_L < p_0$ after decoding and recovery). The ratio k/n is called coding rate. The higher it is, the more efficient the QECC is. No matter for CECCs or QECCs, to improve the reliability of the logic bits/qubits, we often need to increase the code length. Good CECCs have constant or increasing coding rate with code length increasing, some[3–6] can even asymptotically achieve the channel capacity which is a quantity measures the upper limit of coding rate. However, for most QECC schemes, the larger the code length n is, the lower the coding rate will be, which will results in excessive physical qubits overhead. This makes reliable large-scale fault-tolerant quantum computing needs millions of physical qubits, which is very difficult to realize for the current technology. For Surface Code[7–24] which is the most promising QECC at present, and the concatenated QECCs[25–27] which is the earliest and also a promising method to realize fault-tolerant quantum computing, their coding rate tends to 0 with the increase of its code length. For quantum low-density parity check (QLDPC) codes[28–37], though their coding rate is constant with code length increasing, whether their coding rate can achieve the channel capacity has not been proven. In some cases, such as hyperbolic codes[33–37], which is a family of QLDPC codes, have a constant coding rate, but their coding rate does not seem to achieve the quantum channel capacity (we measure this capacity by maximum single letter coherent information of the quantum channel, which is explained in Subsection 2.4). For instance, in ref [36], the asymptotic coding rate of 4D-hyperbolic code is 0.18, but the quantum channel capacity of the independent X/Z -flip noise channel considered by the authors with error rate $p = 0.04$ (i.e. a qubit undergoes independently an X error with probability p or a Z error with probability p) is 0.5178, which is rather larger than its asymptotic coding rate.

Classical polar code (CPC) is the only error correcting code whose coding rate has been proven that it can reach the classical channel capacity[6]. The high coding rate has attracted researchers' attention. In the past decade, researchers are trying to apply the channel polarization idea of CPC to quantum channels and find the analogue of CPC in quantum fields, which is called quantum polar code (QPC)[38–45].

The first step and also the key to design QPC is to figure out whether quantum channels will polarize which is similar to classical channel polarization discovered in [6]. Some previous studies[38–43, 45–48] have proved some quantities of classical-quantum channels whose inputs are classical bits and outputs are qubits, such as classical symmetric capacity[38], Bhattacharyya parameter[38, 48], and classical symmetric Holevo information[46] will polarize. Some studies[39, 40, 42, 47, 48] has referred to coherent information, which is a quantum quantity of quantum channels and is believed to be a quantity to measure the channel capacity of pure quantum channels[49–58], but the coherent information of the classical-quantum channels is just the classical mutual information. Based on the polarization of classical-quantum channels, researchers have proposed some quantum polar coding schemes[38–43, 45–48]. Unfortunately, they cannot be applied to quantum computing whose quantum channels are pure quantum channels. In 2019, Dupuis[44] prove that the symmetric coherent information(the coherent information of quantum channel evaluated for Bell-state input[39]) of pure quantum channels will polarize. However, unlike the classical symmetric capacity having been proved that it is the channel capacity of classical symmetric channels, no one has proved that the symmetric coherent information is the maximum single letter coherent information of pure quantum channels.

In this paper, we focus on proving the polarization of pure quantum channels. We first define a class of quantum channels called quantum symmetric channels (QSCs, this term has been used in [59], but in this paper, it has different meaning), and prove some basic properties of them. For QSC, we prove that its maximum single letter coherent information (MSLCI) equals to its symmetric coherent information. Then we prove the MSLCI will polarize in two-dimensional-input QSC under the quantum channel combining and splitting process. Unlike the proof method used by Dupuis[44], our proof uses the basis transition probability matrix proposed by us.

The rest of this paper is organized as follows. Some preliminary knowledge, including coherent information, quantum symmetric channels, quantum channel combining and splitting, will be introduced in Sect. 2. In Sect. 3, we will prove that the combined channel is quantum symmetric channel and all the coordinate channels are two-dimensional-input quasi quantum symmetric channels. In Sect. 4, we will prove the MSLCI of the coordinate channels will polarize. In Sect. 5, we conclude our work.

2 Preliminaries

2.1 Coherent information

Coherent information is proposed by Schumacher and Nielsen to measure the amount of quantum information conveyed in the noisy channel[60]. It is believed to be the analogue of classical mutual information in quantum information theory[61].

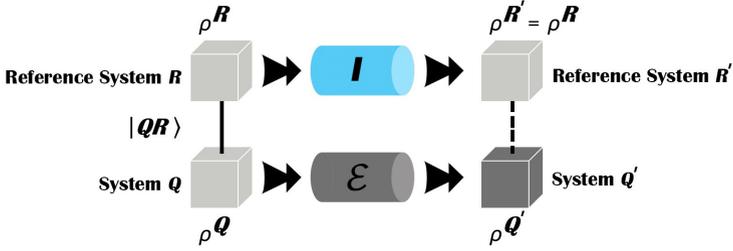


Fig. 1 System Q and its reference system R . System Q is subjected to a quantum channel \mathcal{E} . Notice that the reference system R is only subjected to an identity operator I , namely, $R' = R$. The solid line between system Q and its reference system R indicates Q and R are in a maximally entangled state, which means in a certain basis, the measurement results of Q and R have an one-to-one relationship. Once you get the measurement results of Q , you know the state of R , and vice versa. Hence, we use solid line to represent this “strong” relationship. The dashed line indicates there might be still some entanglement between Q' and R' and the one-to-one relationship might not exist.

As shown in Fig. 1, suppose the state of a quantum system Q is ρ^Q ,

$$\rho^Q = \sum_i p_i |i^Q\rangle \langle i^Q| \quad (1)$$

where $|i^Q\rangle$ is the basis for Q . Suppose Q is subjected to a quantum channel \mathcal{E} which change system Q to Q' and maps the state to $\rho^{Q'}$, namely,

$$\rho^{Q'} = \mathcal{E}(\rho^Q) \quad (2)$$

For system Q , we can always introduce a reference system R which has the same state space as Q to purify Q , namely, map the mixed state ρ^Q to a pure state $|QR\rangle$. The state of system Q and R can be expressed as

$$|QR\rangle = \sum_i \sqrt{p_i} |i^Q\rangle |i^R\rangle \quad (3)$$

where $|i^R\rangle$ is the basis for R , which is the same as $|i^Q\rangle$.

Schumacher defined an intrinsic quantity to Q called entropy exchange S_e [62],

$$S_e \equiv S(RQ') \quad (4)$$

where $S(RQ')$ is the von Neumann entropy of system RQ' .

Coherent information in the process shown in Fig. 1 is defined as

$$I(Q; Q') \equiv S(Q') - S_e = S(Q') - S(RQ') \quad (5)$$

where $S(Q')$ is the von Neumann entropy of system Q' . It's obvious that once Q and \mathcal{E} are given, Q' is determined. Hence, we can also write $I(Q; Q')$ as $I(\rho^Q, \mathcal{E})$.

Assuming the operation elements of \mathcal{E} are $\{E_k\}$, then S_e can be calculated by

$$S_e = S(W) \quad (6)$$

where $W_{ij} = \text{tr} \left(E_i \rho^Q E_j^\dagger \right)$.

It should be emphasized that in this paper, the coherent information which we consider is the single letter coherent information (SLCI). Due to the superadditivity[63] of quantum channel, single letter coherent information is the lower bound of quantum channel capacity. Researchers[64, 65] believes the quantum channel capacity should be more accurately measured by $I(\rho^Q, \mathcal{E}^{\otimes n})$ which is defined by

$$I(\rho^Q, \mathcal{E}^{\otimes n}) \equiv \lim_{n \rightarrow \infty} \frac{1}{n} I(\rho^Q, \mathcal{E}) \quad (7)$$

Whether will $I(\rho^Q, \mathcal{E}^{\otimes n})$ of the coordinate channels polarize has not been proven in this paper.

2.2 Quantum symmetric channels

In classical information theory, there is a class of channels called classical symmetrical channels (CSCs) whose properties have been well-studied, such as binary symmetric channel (BSC). The behavior of a classical channel can be depicted by a transition probability matrix (TPM). Assume the input variable is A , which takes value from $\{a_1, a_2, \dots, a_K\}$, and the output variable is B , which takes value from $\{b_1, b_2, \dots, b_L\}$, then we can write out its TRM as follows.

$$\begin{array}{c} B = b_1 \quad \dots \quad B = b_L \\ \begin{array}{c} A = a_1 \\ A = a_2 \\ \vdots \\ A = a_K \end{array} \left(\begin{array}{ccc} p(B = b_1 | A = a_1) & \dots & p(B = b_L | A = a_1) \\ p(B = b_1 | A = a_2) & \dots & p(B = b_L | A = a_2) \\ \vdots & \ddots & \vdots \\ p(B = b_1 | A = a_K) & \dots & p(B = b_L | A = a_K) \end{array} \right) \end{array} \quad (8)$$

If each row of the TPM is a permutation of the first row, then this channel is symmetric with respect to its input. If each column of the TPM is a permutation of the first column, then this channel is symmetric with respect to its output. If a channel is symmetric with respect to both of its input and output, then this channel is called a symmetric channel. If a channel is symmetric with respect to its input but might not to its output, and its TPM can be divided into several submatrices by column, each of which satisfies that each column of it is a permutation of the first column of it, then this channel is called a quasi symmetric channel.

For some quantum channels, given certain basis of the input space and the output space, we may also find a probability matrix similar to TRM of classical channels. For example, for bit flip channel, if the input state $|Q\rangle$ is $|0\rangle$ (with probability q) or $|1\rangle$ (with probability $1 - q$), then the output state $|Q'\rangle$ will

also take value from $|0\rangle$ or $|1\rangle$, and we can figure out $p(|Q'\rangle = |0\rangle ||Q\rangle = |0\rangle)$, $p(|Q'\rangle = |1\rangle ||Q\rangle = |0\rangle)$, $p(|Q'\rangle = |0\rangle ||Q\rangle = |1\rangle)$, $p(|Q'\rangle = |1\rangle ||Q\rangle = |1\rangle)$. Then we can write out a probability matrix as follows.

$$\begin{array}{cc} & |Q'\rangle = |0\rangle & |Q'\rangle = |1\rangle \\ \begin{array}{l} |Q\rangle = |0\rangle \\ |Q\rangle = |1\rangle \end{array} & \left(\begin{array}{cc} p(|Q'\rangle = |0\rangle ||Q\rangle = |0\rangle) & p(|Q'\rangle = |1\rangle ||Q\rangle = |0\rangle) \\ p(|Q'\rangle = |0\rangle ||Q\rangle = |1\rangle) & p(|Q'\rangle = |1\rangle ||Q\rangle = |1\rangle) \end{array} \right) & \end{array} \quad (9)$$

Here, we name matrix (9) basis transition probability matrix (BTPM), for it shows the transition relationship between the basis of input and output spaces. Different from TPM of classical channels, the above BTPM doesn't seem to fully depicted the behavior of bit flip channel, because quantum mechanics allow the input state to be a superposition state, such as $\frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle$. That is to say, the input state and the output state may take value outside $\{|0\rangle, |1\rangle\}$, which cannot be depicted by BTPM. However, using BTPM (9), given arbitrary input state, we can always determine the output. This is because we can write out the operator elements by matrix (9), which will be proved later. And once the operator elements are determined, the behavior of the quantum channel is determined. Hence, for the quantum channels which have a BTPM, its behavior can be fully depicted by its BTPM.

There is a necessary and sufficient condition for a quantum channel having a BTPM.

Theorem 1 (Necessary and sufficient condition for a quantum channel having a BTPM) *Given a quantum channel \mathcal{E} , it has a BTPM if and only if there is a certain basis $B_{in} = \{|1\rangle, |2\rangle, \dots, |N\rangle\}$ of the input space, any two basis vectors $|i\rangle$ and $|j\rangle$ in B_{in} satisfy that $\mathcal{E}(|i\rangle\langle i|)$ commutes with $\mathcal{E}(|j\rangle\langle j|)$, namely,*

$$[\mathcal{E}(|i\rangle\langle i|), \mathcal{E}(|j\rangle\langle j|)] = \mathcal{E}(|i\rangle\langle i|)\mathcal{E}(|j\rangle\langle j|) - \mathcal{E}(|j\rangle\langle j|)\mathcal{E}(|i\rangle\langle i|) = 0 \quad (10)$$

Proof (1)Sufficiency: If there is a certain basis $B_{in} = \{|1\rangle, |2\rangle, \dots, |N\rangle\}$ of the input space, any two basis vectors $|i\rangle$ and $|j\rangle$ in B_{in} satisfy $[\mathcal{E}(|i\rangle\langle i|), \mathcal{E}(|j\rangle\langle j|)] = 0$, then for all $\mathcal{E}(|i\rangle\langle i|)$, they can be simultaneously diagonalized in a certain basis $B_{out} = \{|1'\rangle, |2'\rangle, \dots, |M'\rangle\}$ of the output space. The result of diagonalization is

$$\mathcal{E}(|i\rangle\langle i|) = \sum_{k=1}^M p_{ik} |k'\rangle\langle k'| \quad (11)$$

It is obvious that p_{ik} forms the BPTM.

(2)Necessity: Assume quantum channel \mathcal{E} has a BTPM whose elements are A_{ik} ($1 \leq i \leq N, 1 \leq k \leq M$), and the corresponding basis for the input and output space are $B_{in} = \{|1\rangle, |2\rangle, \dots, |N\rangle\}$ and $B_{out} = \{|1'\rangle, |2'\rangle, \dots, |M'\rangle\}$, respectively, then $\mathcal{E}(|i\rangle\langle i|)$ can be expressed as

$$\mathcal{E}(|i\rangle\langle i|) = \sum_{k=1}^M A_{ik} |k'\rangle\langle k'| \quad (12)$$

which means that all $\mathcal{E}(|i\rangle\langle i|)$ can be simultaneously diagonalized in $B_{out} = \{|1'\rangle, |2'\rangle, \dots, |M'\rangle\}$. Hence, any two basis vectors $|i\rangle$ and $|j\rangle$ in B_{in} satisfy $[\mathcal{E}(|i\rangle\langle i|), \mathcal{E}(|j\rangle\langle j|)] = 0$. The proof is completed. \square

Next, we are going to prove that one can derive the channel operation elements by BTPM.

Theorem 2 (Derive the channel operation elements from BTPM) *For a quantum channel which has BTPM, its BTPM determine a set of quantum operations.*

Proof Assume quantum channel \mathcal{E} has a BTPM A , for arbitrary input $\rho = \sum_{i=1}^N q_i |i\rangle\langle i|$, the corresponding output $\mathcal{E}(\rho)$ is

$$\mathcal{E}(\rho) = \sum_{i=1}^N q_i \mathcal{E}(|i\rangle\langle i|) = \sum_{i=1}^N q_i \sum_{k=1}^M A_{ik} |k'\rangle\langle k'| = \sum_{k=1}^M E_k \left(\sum_{i=1}^N q_i |i\rangle\langle i| \right) E_k^\dagger \quad (13)$$

where $\{E_k\}$ are the operation elements of channel \mathcal{E} , and $E_k |i\rangle = \sqrt{A_{ik}} |k'\rangle$, according to which one can easily write out the matrix representation of E_k . This completes the proof. \square

According to the above proof, one can see that the number of independent operation elements equals to the number of dimensions of the output space.

Similar to classical symmetric channels, we can define quantum symmetric channels by BTPM.

Definition 1 (Quantum symmetric channels) For the quantum channels which have BTPM, if each row of the BTPM is a permutation of the first row, then this quantum channel is symmetric with respect to its input. If each column of the BTPM is a permutation of the first column, then this quantum channel is symmetric with respect to its output. If a quantum channel is symmetric with respect to both of its input and output, then this channel is called quantum symmetric channel (QSC). If a channel is symmetric with respect to its input but might not to its output, and its BTPM can be divided into several submatrices by column, each of which satisfies that each column of it is a permutation of the first column of it, then this channel is called a quantum quasi symmetric channel (QQSC). Actually, a QSC can be regarded as a special QQSC.

Theorem 3 (Operation elements of two-dimensional-input QQSC) *For a two-dimensional-input QQSC whose output space is M -dimensional, there is always a set of operation elements $\{E_k\}$, $1 \leq k \leq M$, which satisfies*

$$E_k |0\rangle = \sqrt{p_k} |k'\rangle, E_k |1\rangle = \sqrt{p_k} |\pi(k)'\rangle \quad (14)$$

where π is a certain permutation, $\{|k'\rangle\}$ is a basis of the output space, $\sum_{k=1}^M p_k = 1$. Notice that $|0\rangle$ and $|1\rangle$ are only basis vectors of the input space, they are not necessary to be the computational basis vectors $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

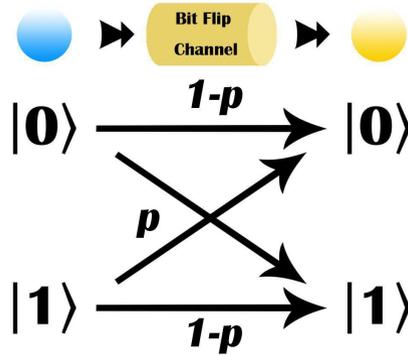


Fig. 2 Bit flip channel

Proof We only need to determine the matrix representation of each E_k and prove that $\sum_k E_k^\dagger E_k = I$. Notice that the matrix representation of E_k can be calculated by

$$E_{k_{ji}} = \langle j | E_k | i - 1 \rangle \quad (15)$$

Here, the ket vector is $|i - 1\rangle$ rather than $|i\rangle$. This is because we use $|0\rangle$ and $|1\rangle$ to represent the input basis rather than $|1\rangle$ and $|2\rangle$, so the index of the column of the matrix starts from 1. Through Eq. (14) and Eq. (15), it's easy to obtain

$$E_{k_{j1}} = \sqrt{p_k} \delta_{jk} \quad (16)$$

$$E_{k_{j2}} = \sqrt{p_k} \delta_{j\pi(k)} \quad (17)$$

where δ is the Kronecker Delta. Hence,

$$\sum_k E_k^\dagger E_k = \begin{pmatrix} \sum_{k=1}^M p_k & 0 \\ 0 & \sum_{k=1}^M p_k \end{pmatrix} = I \quad (18)$$

which completes the proof. \square

2.3 Two examples of QSC

Bit flip channel and phase flip channel are two typical QSCs, as shown in Fig. 2 and Fig. 3, respectively.

Bit flip channel flips $|0\rangle$ and $|1\rangle$ with the same probability p , and phase flip channel flips $|+\rangle$ and $|-\rangle$ with the same probability p . It's easy to write out the operation elements[61] for them. The operation elements for bit flip channel is

$$E_0 = \sqrt{p}I, E_1 = \sqrt{1-p}X \quad (19)$$

where X is the pauli X operator. And the operation elements for phase flip channel is

$$\widetilde{E}_0 = \sqrt{p}I, \widetilde{E}_1 = \sqrt{1-p}Z \quad (20)$$

where Z is the pauli Z operator.

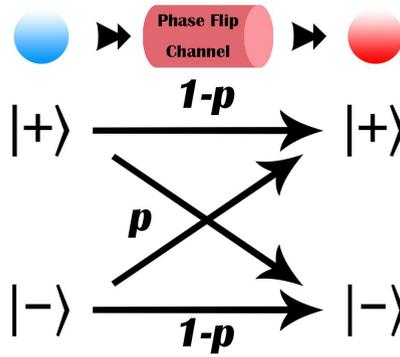


Fig. 3 Phase flip channel

2.4 Symmetric coherent information and the MSLCI of two-dimensional-input QQSC

The symmetric coherent information was first proposed in [40], which is similar to the definition of symmetric capacity used by Arikan[6].

Definition 2 (Symmetric coherent information) For a quantum channel \mathcal{E} , the number of whose input qubits is n , its input state can be represented by $\rho = \sum_{i=1}^{2^n} q_i |i\rangle \langle i|$, its symmetric coherent information I_U is defined as the coherent information $I(\rho, \mathcal{E})$ when $q_1 = q_2 = \dots = q_{2^n} = 1/2^n$, namely,

$$I_U \equiv I \left(\rho = \sum_{i=1}^{2^n} \frac{1}{2^n} |i\rangle \langle i|, \mathcal{E} \right) \quad (21)$$

Arikan has proved that for a classical symmetric channel (actually, the classical symmetric channel mentioned by Arikan in the Part A of Sec. I of [6] means a classical binary quasi symmetric channel), the symmetric capacity is its Shannon capacity. However, up to the present, none of the previous studies[38–48] has proved that the symmetric coherent information of a pure quantum channel is its MSLCI. Next, we will prove this theorem for two-dimensional-input QQSC.

Theorem 4 (The MSLCI of two-dimensional-input QQSC) *The MSLCI of two-dimensional-input QQSC is its symmetric coherent information.*

Proof Assume the input state of a two-dimensional-input QQSC \mathcal{E} is $\rho = q|0\rangle \langle 0| + (1 - q)|1\rangle \langle 1|$. According to Theorem 3, there is a set of operation elements $\{E_k\}$, $1 \leq k \leq M$. By Eq. (5) and Eq. (6), the coherent information of \mathcal{E} is

$$I(\rho, \mathcal{E}) = S(\mathcal{E}(\rho)) - S_e = S(\mathcal{E}(\rho)) - S(W) \quad (22)$$

Using Eq. (15), one can easily obtain

$$\begin{aligned} W_{ij} &= \text{tr} \left(E_i \rho E_j^\dagger \right) = \text{tr} \left(E_i (q |0\rangle \langle 0| + (1-q) |1\rangle \langle 1|) E_j^\dagger \right) \\ &= q \times \text{tr} \left(\sqrt{p_i p_j} |i'\rangle \langle j'| \right) + (1-q) \times \text{tr} \left(\sqrt{p_i p_j} |\pi(i)'\rangle \langle \pi(j)'\rangle \right) \\ &= p_i \delta_{ij} \end{aligned} \quad (23)$$

where δ is the Kronecker Delta and π is a certain permutation.

Hence, $S(W) = H(p_i)$, where $H(p_i)$ is the Shannon entropy of the probability distribution $\{p_1, \dots, p_M\}$. It's obviously that $S(W)$ has nothing to do with q .

Next, we analyze the first term $S(\mathcal{E}(\rho))$. Using Eq. (14), we get

$$\begin{aligned} S(\mathcal{E}(\rho)) &= S \left(\sum_{k=1}^M E_k \rho E_k^\dagger \right) \\ &= S \left(\sum_{k=1}^M q p_k |k'\rangle \langle k'| + \sum_{k=1}^M (1-q) p_k |\pi(k)'\rangle \langle \pi(k)'\rangle \right) \\ &= S \left(\sum_{k=1}^M q p_k |k'\rangle \langle k'| + \sum_{m=1}^M (1-q) p_{\pi(m)} |m'\rangle \langle m'| \right) \\ &= S \left(\sum_{k=1}^M q p_k |k'\rangle \langle k'| + \sum_{k=1}^M (1-q) p_{\pi(k)} |k'\rangle \langle k'| \right) \end{aligned} \quad (24)$$

The third equality in Eq. (24) holds because $\pi(\pi(k)) = k$. If one let $\pi(k) = m$, then $\pi(m) = k$. The last equality is obtained simply by renaming m .

Notice that von Neumann entropy has a property which states that when ρ_i have support on orthogonal subspaces, the following equation holds.

$$S \left(\sum_i p_i \rho_i \right) = \sum_i p_i S(\rho_i) + H(p_i) \quad (25)$$

Using Eq. (25), we can further simplify Eq. (24).

$$\begin{aligned} S(\mathcal{E}(\rho)) &= \sum_{k=1}^M \left[q p_k + (1-q) p_{\pi(k)} \right] S(|k'\rangle \langle k'|) + H \left(q p_k + (1-q) p_{\pi(k)} \right) \\ &= H \left(q p_k + (1-q) p_{\pi(k)} \right) \end{aligned} \quad (26)$$

Taking the derivative with respect to q , we obtain

$$\begin{aligned} \frac{dS(\mathcal{E}(\rho))}{dq} &= - \sum_{k=1}^M \left\{ \left(p_k - p_{\pi(k)} \right) \log_2 \left[\left(p_k - p_{\pi(k)} \right) q + p_{\pi(k)} \right] + \frac{\left(p_k - p_{\pi(k)} \right)}{\ln 2} \right\} \\ &= - \sum_{k=1}^M t_k \end{aligned} \quad (27)$$

where $t_k = - \left(p_k - p_{\pi(k)} \right) \log_2 \left[\left(p_k - p_{\pi(k)} \right) q + p_{\pi(k)} \right] + \left(p_k - p_{\pi(k)} \right) / \ln 2$. Notice that there are M terms in the summation sign, which can be divided into $M/2$ pairs, each of which can be represented by

$$y_k = t_k + t_{\pi(k)} \quad (28)$$

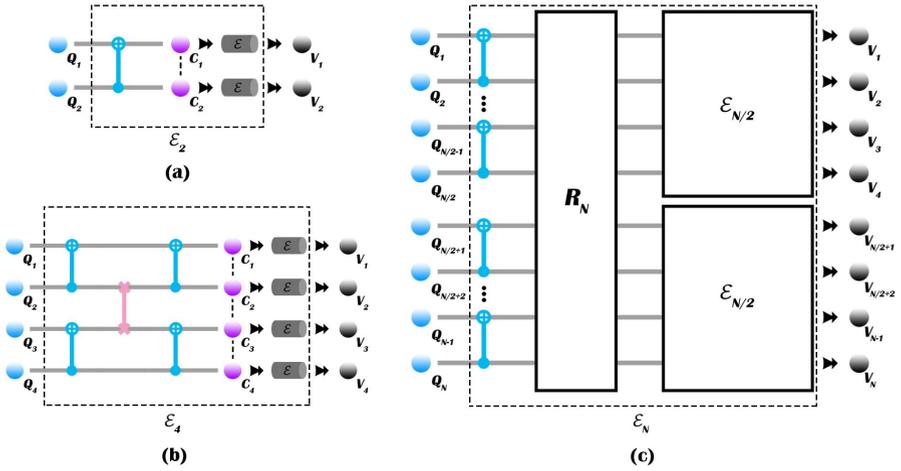


Fig. 4 (a) Two primal channel \mathcal{E} combines to form channel \mathcal{E}_2 . (b) Two primal \mathcal{E}_2 combines to form channel \mathcal{E}_4 . (c) Two primal $\mathcal{E}_{N/2}$ combines to form channel \mathcal{E}_N , R_N is the reverse shuffle operator[6]. The blue gates are quantum CNOT gates and the pink gates are quantum SWAP gates.

It's easy to prove that for all y_k , when $q \in [0, \frac{1}{2})$, $y_k < 0$, when $q \in (\frac{1}{2}, 1]$, $y_k > 0$, and when $q = \frac{1}{2}$, $y_k = 0$. Hence, when $q \in [0, \frac{1}{2})$, $\frac{dS(\mathcal{E}(\rho))}{dq} > 0$, when $q \in (\frac{1}{2}, 1]$, $\frac{dS(\mathcal{E}(\rho))}{dq} < 0$, and $q = \frac{1}{2}$, $\frac{dS(\mathcal{E}(\rho))}{dq} = 0$. Therefore, $q = \frac{1}{2}$ is the maximum point of $S(\mathcal{E}(\rho))$, which completes the proof. \square

2.5 Quantum channel combining and splitting

2.5.1 Quantum channel combining

Quantum channel combining and splitting are two steps to polarize quantum channels. The quantum channel combining is similar to classical channel combining. Assume the primal channel is $\mathcal{E} : \rho^Q \rightarrow \rho^V$, which maps the state of a qubit to another state. We denote the input by Q , and the output by V . As shown in Fig. 4, we use the same recursive manner as in classical channel combining to combine N primal quantum channels. The difference is that we replace XOR gates in classical channel combining by quantum CNOT gates, and we use quantum SWAP gates to realize the reverse shuffle operator[6]. The channel combining process produces the channel $\mathcal{E}_N : \rho^{Q_1 \dots Q_N} \rightarrow \rho^{V_1 \dots V_N}$, where the subscript i ($1 \leq i \leq N$) means the i th qubit. This paper follows the Arikan's rule to denote a row vector, namely, we use a_1^i as a shorthand for denoting (a_1, \dots, a_i) , and the notation 0_1^N is used to denote the all-zero vector. According to this rule, \mathcal{E}_N can be rewritten as $\mathcal{E}_N : \rho^{Q_1^N} \rightarrow \rho^{V_1^N}$.

2.5.2 Quantum channel splitting

Having combining N quantum channels \mathcal{E} to \mathcal{E}_N , the next step to polarize quantum channels is splitting \mathcal{E}_N to N quantum coordinate channels $\mathcal{E}_N^{(i)} : \rho^{Q_i} \rightarrow \rho^{V_1^N, R_1^{i-1}}$, namely, $\mathcal{E}_N^{(i)} : \rho^{Q_i} \rightarrow \rho^{V_1 \cdots V_N, R_1 \cdots R_{i-1}}$, where R_i is the reference system of Q_i , and $1 \leq i \leq N$. The quantum coordinate channels we define is a little bit different from the classical coordinate channels. If we follow the classical definition, the quantum coordinate channels should be $\mathcal{E}_N^{(i)} : \rho^{Q_i} \rightarrow \rho^{V_1^N, Q_1^{i-1}}$. However, due to quantum no-cloning theorem, $\rho^{Q_1^{i-1}}$ and $\rho^{V_1^{i-1}}$ cannot appear at the same side. Moreover, according to the manner of Eq. (3) we introduce the reference systems, R_i and Q_i are in maximally entangled states, which means that the state of R_i is the same as Q_i . Hence, $\mathcal{E}_N^{(i)} : \rho^{Q_i} \rightarrow \rho^{V_1^N, R_1^{i-1}}$ is a more reasonable definition.

According to Theorem 4, for a two-dimensional-input QSC, when the input state is a completely mixed state, its SLCI takes maximum.

3 Symmetry of the quantum combined channel and coordinate channels

In Sect. 2.5, quantum combined channel \mathcal{E}_N and coordinate channels $\{\mathcal{E}_N^{(i)}\}$ have been defined. The main goal of this section is to prove that if the primal channel \mathcal{E} is a two-dimensional-input QSC with two-dimensional output, then \mathcal{E}_N is a QSC and $\{\mathcal{E}_N^{(i)}\}$ are two-dimensional-input QQSCs. We will refer to the proof method which Arıkan used to prove that if the primal binary-input discrete memoryless channel $W : \mathcal{X} \rightarrow \mathcal{Y}$ is symmetric with input alphabet $\mathcal{X} = \{0, 1\}$ and output alphabet $\mathcal{Y} = \{0', 1'\}$, classical combined channel $W_N : \mathcal{X}^N \rightarrow \mathcal{Y}^N$ and classical coordinate channels $W_N^{(i)} : \mathcal{X} \rightarrow \mathcal{Y}^N \times \mathcal{X}^{i-1}$, $1 \leq i \leq N$, are symmetric.

3.1 Symmetry of the quantum combined channel \mathcal{E}_N

If \mathcal{E} is a two-dimensional-input QSC with two-dimensional output, the BTM of the channel \mathcal{E} can be expressed as

$$\begin{array}{l} |0\rangle \begin{pmatrix} \langle 0'| & \langle 1'| \\ Pr(|0'\rangle|0\rangle & Pr(|1'\rangle|0\rangle) \end{pmatrix} \\ |1\rangle \begin{pmatrix} \langle 0'| & \langle 1'| \\ Pr(|0'\rangle|1\rangle & Pr(|1'\rangle|1\rangle) \end{pmatrix} \end{array} \quad (29)$$

where $Pr(|0'\rangle|0\rangle) = Pr(|1'\rangle|1\rangle)$ and $Pr(|1'\rangle|0\rangle) = Pr(|0'\rangle|1\rangle)$. According to Theorem 2, we can derive a set of quantum operations $\{E_0, E_1\}$ of this channel \mathcal{E} , which satisfy $E_0|0\rangle = \sqrt{Pr(|0'\rangle|0\rangle)}|0'\rangle$, $E_1|0\rangle = \sqrt{Pr(|1'\rangle|0\rangle)}|0'\rangle$, $E_0|1\rangle = \sqrt{Pr(|1'\rangle|1\rangle)}|1'\rangle$ and $E_1|1\rangle = \sqrt{Pr(|0'\rangle|1\rangle)}|0'\rangle$.

Definition 3 (*N -copy channel $\mathcal{E}^{\otimes N}$ of the primal QSC \mathcal{E}*) We define a N -copy channel $\mathcal{E}^{\otimes N} : \rho^{Q_1} \otimes \dots \otimes \rho^{Q_N} \rightarrow \rho^{V_1} \otimes \dots \otimes \rho^{V_N}$ which is simply composed by N independent copies of the primal $\mathcal{E} : \rho^Q \rightarrow \rho^V$. The operation elements $\{F_k\}$ of $\mathcal{E}^{\otimes N}$ is

$$F_k = E_{b_1}^1 \otimes E_{b_2}^2 \otimes \dots \otimes E_{b_N}^N \quad (30)$$

where the subscript $b_j \in \{0, 1\}$, $1 \leq j \leq N$. The superscript i of $E_{b_j}^i$ means the operation element $E_{b_j}^i$ only acts on the i th input state ρ^{Q_i} , and the subscript k ($0 \leq k \leq 2^N - 1$) of operation elements F_k is the decimal number of the binary sequence $b_1 b_2 \dots b_N$.

Assume that N uncorrelated pure input states of the channel $\mathcal{E}^{\otimes N}$ is $|Q_1^N\rangle = |Q_1\rangle \otimes \dots \otimes |Q_N\rangle$, we have

$$\begin{aligned} F_k |Q_1^N\rangle &= E_{b_1}^1 \otimes \dots \otimes E_{b_N}^N (|Q_1\rangle \otimes \dots \otimes |Q_N\rangle) \\ &= \prod_{i=1}^N \sqrt{Pr(|V_i\rangle || Q_i\rangle)} (|V_1\rangle \otimes \dots \otimes |V_N\rangle) \\ &= \sqrt{Pr_N(|V_1^N\rangle || Q_1^N\rangle)} |V_1^N\rangle \end{aligned} \quad (31)$$

where we let

$$|V_1\rangle \otimes \dots \otimes |V_N\rangle = |V_1^N\rangle \quad (32)$$

and

$$Pr_N(|V_1^N\rangle || Q_1^N\rangle) = \prod_{i=1}^N Pr(|V_i\rangle || Q_i\rangle) \quad (33)$$

for all $V_1^N \in \mathcal{Y}^N$, $Q_1^N \in \mathcal{X}^N$. \mathcal{X}^N is the N -power extension alphabet of \mathcal{X} and \mathcal{Y}^N is the N -power extension alphabet of \mathcal{Y} . Eq. (33) means $Pr_N(|V_1^N\rangle || Q_1^N\rangle)$ is the transition probability when the input state of $\mathcal{E}^{\otimes N}$ is $|Q_1^N\rangle$ and the output state of $\mathcal{E}^{\otimes N}$ is $|V_1^N\rangle$.

In Fig. 4, one can see that $\mathcal{E}^{\otimes N}$ is just the last layer of \mathcal{E}_N , which is to say, if the recursive combining circuits are omitted, \mathcal{E}_N will become $\mathcal{E}^{\otimes N}$. Intuitively, it seems that the BTPM of \mathcal{E}_N should have some connections with that of $\mathcal{E}^{\otimes N}$. Next, we prove this intuition.

Proposition 5 (*The BTPM of quantum combined channel \mathcal{E}_N*) If each input state of the channel \mathcal{E}_N is uncorrelated, the basis transition probabilities of the channel \mathcal{E}_N can be obtained by the following equation

$$Pr_N(|V_1^N\rangle || Q_1^N\rangle) = \prod_{i=1}^N Pr(|V_i\rangle || C_i\rangle) \quad (34)$$

for all $C_i \in \mathcal{X}$, $V_i \in \mathcal{Y}$, $V_1^N \in \mathcal{Y}^N$, $Q_1^N \in \mathcal{X}^N$, where $|Q_1^N\rangle$ and $|V_1^N\rangle$ are the input basis vector and the output basis vector of channel \mathcal{E}_N respectively. $|C_i\rangle$ and $|V_i\rangle$ are the i th input basis vector and the i th output basis vector of the channel $\mathcal{E}^{\otimes N}$ respectively, as shown in Fig. 4.

Proof Assume that each uncorrelated input state ρ^{Q_i} of the quantum combined channel \mathcal{E}_N is $\rho^{Q_i} = q|0\rangle\langle 0| + (1-q)|1\rangle\langle 1|$. Then we have

$$\begin{aligned}\rho^{Q_1^N} &= \rho^{Q_1} \otimes \dots \otimes \rho^{Q_N} \\ &= (q|0\rangle\langle 0| + (1-q)|1\rangle\langle 1|)^{\otimes N} \\ &= \sum_{Q_1^N \in \mathcal{X}^N} Pr(|Q_1^N\rangle\langle Q_1^N|) |Q_1^N\rangle\langle Q_1^N|\end{aligned}\quad (35)$$

where alphabet $\mathcal{X} = \{0, 1\}$ and \mathcal{X}^N is the N -power extension alphabet of \mathcal{X} , $Pr(|Q_1^N\rangle\langle Q_1^N|)$ is the probability of $|Q_1^N\rangle\langle Q_1^N|$. Since each input state ρ^{Q_i} is uncorrelated with other input states, we have

$$Pr(|Q_1^N\rangle\langle Q_1^N|) = \prod_{i=1}^N Pr(|Q_i\rangle\langle Q_i|) \quad (36)$$

Since the process $|Q_1^N\rangle \rightarrow |C_1^N\rangle$ which can be seen as an encoding process only includes quantum CNOT gates and quantum SWAP gates, this process must be unitary. As shown in Fig. 4, we use unitary operator U_N to denote this encoding process, and obtain

$$\begin{aligned}\rho^{C_1^N} &= U_N \rho^{Q_1^N} U_N^\dagger \\ &= U_N \left(\sum_{Q_1^N \in \mathcal{X}^N} Pr(|Q_1^N\rangle\langle Q_1^N|) |Q_1^N\rangle\langle Q_1^N| \right) U_N^\dagger \\ &= \sum_{Q_1^N \in \mathcal{X}^N} Pr(|Q_1^N\rangle\langle Q_1^N|) |Q_1^N G_N\rangle\langle Q_1^N G_N| \\ &= \sum_{C_1^N \in \mathcal{X}^N} Pr(|Q_1^N\rangle\langle Q_1^N|) |C_1^N\rangle\langle C_1^N|\end{aligned}\quad (37)$$

where $C_1^N = Q_1^N G_N$, and G_N is generator matrix[6].

CNOT gates can produce entanglement between its two input qubits while the control qubit is in a superposition state in the computational basis. However, since the input states $|Q_i\rangle$ will only take value from $|0\rangle$ or $|1\rangle$, CNOT gates will not produce entanglement[66, 67]. Besides, SWAP gates will not produce entanglement between its two inputs. Hence, all $|C_i\rangle$ are uncorrelated. By Eq. (33), we have

$$Pr_N(|V_1^N\rangle\langle V_1^N|) = \prod_{i=1}^N Pr(|V_i\rangle\langle V_i|) \quad (38)$$

Since $|C_1^N\rangle = |Q_1^N G_N\rangle$, once Q_1^N is determined, C_1^N will be determined. Thus,

$$\begin{aligned}Pr_N(|V_1^N\rangle\langle V_1^N|) &= Pr_N(|V_1^N\rangle\langle V_1^N|) \\ &= Pr_N(|V_1^N\rangle\langle C_1^N|) \\ &= \prod_{i=1}^N Pr(|V_i\rangle\langle C_i|)\end{aligned}\quad (39)$$

which completes the proof. \square

Let $\mathcal{E} : \rho^Q \rightarrow \rho^V$ is a two-dimensional-input QSC with two-dimensional output. By definition, there is a permutation π_1 on \mathcal{Y} such that 1) $\pi_1^{-1} = \pi_1$ and 2) $Pr(|V\rangle || 1\rangle) = Pr(|\pi_1(V)\rangle || 0\rangle)$ for all $V \in \mathcal{Y} = \{0', 1'\}$. Let π_0 be the identity permutation on \mathcal{Y} . Using the compact notation mentioned by Arikan, we denote $\pi_Q(V)$ by $Q \cdot V$, for all $Q \in \mathcal{X} = \{0, 1\}$ and $V \in \mathcal{Y} = \{0', 1'\}$.

Observe that $Pr(|V\rangle || Q \oplus a) = Pr(|a \cdot V\rangle || Q)$ for all $a, Q \in \mathcal{X} = \{0, 1\}$ and $V \in \mathcal{Y} = \{0', 1'\}$. It's easy to verify that $Pr(|V\rangle || Q \oplus a) = Pr((Q \oplus a) \cdot V || 0) = Pr(|Q \cdot (a \cdot V)\rangle || 0)$ and $Pr(|V\rangle || Q \oplus a) = Pr(|Q \cdot V\rangle || a)$ since \oplus is commutative operation on \mathcal{X} .

For $Q_1^N \in \mathcal{X}^N$, $V_1^N \in \mathcal{Y}^N$, let

$$Q_1^N \cdot V_1^N \triangleq (Q_1 \cdot V_1, \dots, Q_N \cdot V_N) \quad (40)$$

Next, we will prove the quantum combined channel \mathcal{E}_N is symmetric.

Theorem 6 (the quantum combined channel \mathcal{E}_N is a QSC) *If the primal channel \mathcal{E} is a two-dimensional-input QSC with two-dimensional output, then the quantum combined channel \mathcal{E}_N is QSC in the sense that*

$$Pr_N(|V_1^N\rangle || Q_1^N) = Pr_N(|a_1^N G_N \cdot V_1^N\rangle || Q_1^N \oplus a_1^N) \quad (41)$$

for all $Q_1^N, a_1^N \in \mathcal{X}^N$ and $V_1^N \in \mathcal{Y}^N$.

The Eq. (41) means arbitrary row of the BTPM of \mathcal{E}_N is a permutation of the first row, and arbitrary column of the BTPM of \mathcal{E}_N is a permutation of the first column.

Proof By Proposition 5, we have

$$\begin{aligned} Pr_N(|V_1^N\rangle || Q_1^N) &= \prod_{i=1}^N Pr(|V_i\rangle || C_i) \\ &= \prod_{i=1}^N Pr(|C_i \cdot V_i\rangle || 0) \\ &= Pr_N(|C_1^N \cdot V_1^N\rangle || 0_1^N) \end{aligned} \quad (42)$$

Let $b_1^N = a_1^N G_N$, we have

$$\begin{aligned} Pr_N(|b_1^N \cdot V_1^N\rangle || Q_1^N \oplus a_1^N) &= Pr_N(|(C_1^N \oplus b_1^N) \cdot (b_1^N \cdot V_1^N)\rangle || 0_1^N) \\ &= Pr_N(|C_1^N \cdot V_1^N\rangle || 0_1^N) \\ &= Pr_N(|V_1^N\rangle || Q_1^N) \end{aligned} \quad (43)$$

which completes the proof. \square

3.2 Symmetry of the quantum coordinate channels

$$\{\mathcal{E}_N^{(i)} : 0 \leq i \leq N\}$$

In this part, we will prove that if the primal channel \mathcal{E} is a two-dimensional-input QSC with two-dimensional output, the coordinate channels $\{\mathcal{E}_N^{(i)} : 0 \leq i \leq N\}$ are QQSCs. The key of the proof is to find out the BTPMs of $\{\mathcal{E}_N^{(i)} : 0 \leq i \leq N\}$, and prove their arbitrary row is a permutation of another row.

Theorem 7 (the quantum coordinate channels $\{\mathcal{E}_N^{(i)} : 0 \leq i \leq N\}$ are QQSCs) *If the primal channel \mathcal{E} is a two-dimensional-input QSC with two-dimensional output, and the input state $\rho^{Q_i} = q|0\rangle\langle 0| + (1-q)|1\rangle\langle 1|$, then the arbitrary quantum coordinate channel $\mathcal{E}_N^{(i)} : \rho^{Q_i} \rightarrow \rho^{V_1^N, R_1^{i-1}}$, $1 \leq i \leq N$, is QQSC. The density operator $\rho^{V_1^N, R_1^{i-1}}$ of the joint system V_1^N, R_1^{i-1} can be written as*

$$\rho^{V_1^N, R_1^{i-1}} = \sum_{m=0}^{2^N-1} \left[q Pr_N^{(i)}(|m'\rangle||0\rangle) |m'\rangle\langle m'| + (1-q) Pr_N^{(i)}(|m'\rangle||1\rangle) |m'\rangle\langle m'| \right] \quad (44)$$

where $|m'\rangle = \sum_{\substack{Q_1^{i-1} \\ =R_1^{i-1} \\ \in \mathcal{X}^{i-1}}} \sqrt{Pr(|Q_1^{i-1}\rangle\langle Q_1^{i-1}|)} |(Q_1^{i-1}, 0, 0_{i+1}^N) G_N \cdot V_1^N, R_1^{i-1}\rangle$, $0 \leq m \leq 2^N - 1$, form a set of basis $\{|m'\rangle\}_{m=0, \dots, 2^N-1}$ which contains 2^N basis vectors.

And the basis transition probabilities are

$$\begin{aligned} & Pr_N^{(i)}(|m'\rangle||Q_i\rangle) \\ &= Pr_N^{(i)} \left(\sum_{\substack{Q_1^{i-1} \\ =R_1^{i-1} \\ \in \mathcal{X}^{i-1}}} \sqrt{Pr(|Q_1^{i-1}\rangle\langle Q_1^{i-1}|)} |(Q_1^{i-1}, 0, 0_{i+1}^N) G_N \cdot V_1^N, R_1^{i-1}\rangle ||Q_i\rangle \right) \\ &= \sum_{Q_{i+1}^N \in \mathcal{X}^{N-i}} Pr(|Q_{i+1}^N\rangle\langle Q_{i+1}^N|) Pr_N(|V_1^N\rangle||0_1^{i-1}, Q_i, Q_{i+1}^N\rangle) \\ &= Pr_N^{(i)} \left(\sum_{Q_1^{i-1}=R_1^{i-1} \in \mathcal{X}^{i-1}} \sqrt{Pr(|Q_1^{i-1}\rangle\langle Q_1^{i-1}|)} \right. \\ & \quad \left. |(a_1^{i-1}, 1, a_{i+1}^N \oplus Q_1^{i-1}, 0, 0_{i+1}^N) G_N \cdot V_1^N, R_1^{i-1}\rangle ||Q_i \oplus 1\rangle \right) \end{aligned} \quad (45)$$

for all $V_1^N \in \mathcal{Y}^N$, $Q_i \in \mathcal{X}$, $(a_1^{i-1}, 1, a_{i+1}^N)$, $(Q_1^{i-1}, Q_i, Q_{i+1}^N) \in \mathcal{X}^N$, $N = 2^n$, $n \geq 0$, $1 \leq i \leq N$, which means arbitrary row of the BTPM of $\mathcal{E}_N^{(i)}$ is a permutation of another row.

The proof of Theorem 7 is given in Appendix B.

Since $\mathcal{E}_N^{(i)}$ is two-dimensional-input QQSC, according to Theorem 4, the MSLCI of $\mathcal{E}_N^{(i)}$ is equal to its symmetric coherent information, namely, the SLCI of $\mathcal{E}_N^{(i)}$ takes the maximum when the input state is $\rho^{Q_i} = \frac{1}{2} |0\rangle \langle 0| + \frac{1}{2} |1\rangle \langle 1|$, therefore Eq. (44) and Eq. (45) are reduced to

$$\rho^{V_1^N, R_1^{i-1}} = \frac{1}{2} \sum_{m=0}^{2^N-1} Pr_N^{(i)}(|m\rangle || 0\rangle) |m\rangle \langle m| + \frac{1}{2} \sum_{m=0}^{2^N-1} Pr_N^{(i)}(|m\rangle || 1\rangle) |m\rangle \langle m| \quad (46)$$

and

$$\begin{aligned} & Pr_N^{(i)}(|m\rangle || Q_i) \\ &= Pr_N^{(i)} \left(\sum_{Q_1^{i-1}=R_1^{i-1} \in \mathcal{X}^{i-1}} \frac{1}{2^{\frac{i-1}{2}}} |(Q_1^{i-1}, 0, 0_{i+1}^N) G_N \cdot V_1^N, R_1^{i-1}\rangle || Q_i \right) \\ &= \frac{2^{i-1}}{2^{N-1}} \sum_{Q_{i+1}^N \in \mathcal{X}^{N-i}} Pr_N(|V_1^N\rangle || 0_1^{i-1}, 0, Q_{i+1}^N) \\ &= Pr_N^{(i)} \left(\sum_{\substack{Q_1^{i-1} \\ =R_1^{i-1} \\ \in \mathcal{X}^{i-1}}} \frac{1}{2^{\frac{i-1}{2}}} |(a_1^{i-1}, 1, a_{i+1}^N \oplus Q_1^{i-1}, 0, 0_{i+1}^N) G_N \cdot V_1^N, R_1^{i-1}\rangle || Q_i \oplus 1 \right) \end{aligned} \quad (47)$$

where $|m\rangle = \sum_{Q_1^{i-1}=R_1^{i-1} \in \mathcal{X}^{i-1}} \frac{1}{2^{\frac{i-1}{2}}} |(Q_1^{i-1}, 0, 0_{i+1}^N) G_N \cdot V_1^N, R_1^{i-1}\rangle$, $0 \leq m \leq 2^N - 1$.

4 Polarization of two-dimensional-input QSC

The goal of this section is to prove the MSLCI of coordinate channels $\{\mathcal{E}_N^{(i)}\}$ will polarize.

One can see that the quantum combined channel \mathcal{E}_N corresponds to a classical combined channel W_N , which is obtained by simply replace the quantum circuits in Fig. 4 to a classical ones, and the primal channel \mathcal{E} to a classical channel W . Our proof in this section makes use of the connection between \mathcal{E}_N and W_N .

If the BTPM of the primal QSC \mathcal{E} and the TPM of classical primal BSC W are the same, first of all, we prove that the BTPM of the quantum combined channel \mathcal{E}_N and the TPM of classical combined channel W_N are the same; secondly, we prove the BTPM of quantum coordinate channel $\mathcal{E}_N^{(i)}$ can be derived from the TPM of classical coordinate channel $W_N^{(i)}$ which reveals the relationship between the BTPM of $\mathcal{E}_N^{(i)}$ and the TPM of $W_N^{(i)}$; finally we use

this relationship to prove that the MSLCI of $\mathcal{E}_N^{(i)}$ numerically equals to the Shannon capacity of $W_N^{(i)}$. Since the Shannon capacity of $\{W_N^{(i)}\}$ will polarize, the MSLCI of $\{\mathcal{E}_N^{(i)}\}$ will polarize as well. Moreover, due to the MSLCI of the primal channel \mathcal{E} being equal to the Shannon capacity of the classical primal channel W , the polarization rate of $\{\mathcal{E}_N^{(i)}\}$ equals to the MSLCI of \mathcal{E} , which is referred to Arikan's method[6].

Proposition 8 (Relationship between the BTM of \mathcal{E}_N and the TPM of W_N) Assume that the BTM of a two-dimensional-input QSC with two-dimensional output \mathcal{E} is

$$\begin{array}{l} |0\rangle \\ |1\rangle \end{array} \begin{pmatrix} \begin{array}{c} |0'\rangle \\ |1'\rangle \end{array} \\ \begin{array}{c} \Pr(|0'\rangle||0\rangle) \\ \Pr(|0'\rangle||1\rangle) \end{array} \end{pmatrix} \begin{array}{c} \begin{array}{c} |1'\rangle \\ |1\rangle \end{array} \\ \begin{array}{c} \Pr(|1'\rangle||0\rangle) \\ \Pr(|1'\rangle||1\rangle) \end{array} \end{array} \quad (48)$$

where $\Pr(|0'\rangle||0\rangle) = \Pr(|1'\rangle||1\rangle) = W(0'0) = W(1'1)$ and $\Pr(|1'\rangle||0\rangle) = \Pr(|0'\rangle||1\rangle) = W(1'0) = W(0'1)$. Then the BTM of quantum combined channel \mathcal{E}_N and the TPM of classical combined channel W_N are the same, that is to say

$$\Pr_N(|V_1^N\rangle||Q_1^N\rangle) = W_N(y_1^N|u_1^N) \quad (49)$$

for all $V_1^N = y_1^N \in \mathcal{Y}^N$ and $Q_1^N = u_1^N \in \mathcal{X}^N$, where $y, V \in \mathcal{Y} = \{0', 1'\}$ and $u, Q \in \mathcal{X} = \{0, 1\}$.

Proof By Proposition 5, we have

$$\begin{aligned} \Pr_N(|V_1^N\rangle||Q_1^N\rangle) &= \Pr_N(|V_1^N\rangle||Q_1^N G_N\rangle) \\ &= \Pr_N(|V_1^N\rangle||C_1^N\rangle) \\ &= \prod_{i=1}^N \Pr(|V_i\rangle||C_i\rangle) \end{aligned} \quad (50)$$

According to Arikan's method[6], we have

$$\begin{aligned} W_N(y_1^N|u_1^N) &= W_N(y_1^N|u_1^N G_N) \\ &= W_N(y_1^N|x_1^N) \\ &= \prod_{i=1}^N W(V_i|x_i) \end{aligned} \quad (51)$$

where $u_1^N G_N = x_1^N$. Since $V_1^N = y_1^N$ and $Q_1^N = u_1^N$, then we have $Q_1^N G_N = u_1^N G_N = C_1^N = x_1^N$. Thus, we have $\Pr(|V_i\rangle||C_i\rangle) = W(y_i|x_i)$, and obtain

$$\prod_{i=1}^N \Pr(|V_i\rangle||C_i\rangle) = \prod_{i=1}^N W(V_i|x_i) \quad (52)$$

which completes the proof. \square

Proposition 9 (Relationship between the BTM of $\mathcal{E}_N^{(i)}$ and the TPM of $W_N^{(i)}$) According to Eq. (46) and Eq. (47), when the input state ρ^{Q_i} of the channel $\mathcal{E}_N^{(i)}$ is $\rho^{Q_i} = \frac{1}{2} |0\rangle\langle 0| + \frac{1}{2} |1\rangle\langle 1|$, the output state $|m\rangle$ of the channel $\mathcal{E}_N^{(i)}$ is

$$|m\rangle = \sum_{Q_1^{i-1}=R_1^{i-1} \in \mathcal{X}^{i-1}} \frac{1}{2^{\frac{i-1}{2}}} |(Q_1^{i-1}, 0, 0_{i+1}^N) G_N \cdot V_1^N, R_1^{i-1}\rangle \quad (53)$$

and the basis transition probabilities are

$$\begin{aligned} Pr_N^{(i)}(|m\rangle || Q_i) &= \\ Pr_N^{(i)} &\left(\sum_{Q_1^{i-1}=R_1^{i-1} \in \mathcal{X}^{i-1}} \frac{1}{2^{\frac{i-1}{2}}} |(Q_1^{i-1}, 0, 0_{i+1}^N) G_N \cdot V_1^N, R_1^{i-1}\rangle || Q_i \right) \\ &= \frac{2^{i-1}}{2^{N-1}} \sum_{Q_1^{i-1} \in \mathcal{X}^{i-1}} Pr_N(|V_1^N\rangle || 0_1^{i-1}, Q_i, Q_{i+1}^N) \end{aligned} \quad (54)$$

We can derive $Pr_N^{(i)}(|m\rangle || Q_i)$ from the TPM of classical coordinate channels $W_N^{(i)}$

$$\begin{aligned} Pr_N^{(i)}(|m\rangle || Q_i) &= \\ &= \sum_{u_1^{i-1} \in \mathcal{X}^{i-1}} \frac{1}{2^{N-1}} \sum_{u_{i+1}^N \in \mathcal{X}^{N-i}} W_N((u_1^{i-1}, 0, 0_{i+1}^N) G_N \cdot y_1^N | u_1^{i-1}, u_i, u_{i+1}^N) \\ &= \frac{2^{i-1}}{2^{N-1}} \sum_{u_{i+1}^N \in \mathcal{X}^{N-i}} W_N(y_1^N | 0_1^{i-1}, u_i, u_{i+1}^N) \\ &= 2^{i-1} W_N^{(i)}(y_1^N, 0_1^{i-1} | u_i) \end{aligned} \quad (55)$$

for all $V_1^N = y_1^N \in \mathcal{Y}^N$ and $Q_1^N = u_1^N \in \mathcal{X}^N$, where $y, V \in \mathcal{Y} = \{0', 1'\}$ and $u, Q \in \mathcal{X} = \{0, 1\}$.

The proof of Proposition 9 is given in Appendix C.

The Proposition 9 means that arbitrary column of the BTM of each $\mathcal{E}_N^{(i)}$ is the sum of some 2^{i-1} columns of the TPM whose corresponding elements are all equal, hence the TPM of each $W_N^{(i)}$ has 2^{N+i-1} columns while the BTM of each $\mathcal{E}_N^{(i)}$ has 2^N columns.

Theorem 10 (the polarization of quantum coordinate channels $\{\mathcal{E}_N^{(i)}\}$) If the BTM of the primal QSC \mathcal{E} and the TPM of classical primal BSC W are the same, the MSLCI $I(\rho^{Q_i}, \mathcal{E}_N^{(i)})$ of the quantum coordinate channel $\mathcal{E}_N^{(i)}$ is numerically equal to the Shannon capacity $I(W_N^{(i)})$ of classical coordinate channel $W_N^{(i)}$, namely,

$$\begin{aligned} I(\rho^{Q_i}, \mathcal{E}_N^{(i)}) &= S(\rho^{V_1^N, R_1^{i-1}}) - S(\rho^{V_1^N, R_1^i}) = I(W_N^{(i)}) \\ &= H(y_1^N u_1^{i-1}) - H(y_1^N u_1^i) + H(u_i) \end{aligned} \quad (56)$$

where $p(u_i = 0) = p(u_i = 1) = \frac{1}{2}$ and the density operator $\rho^{Q_i} = \frac{1}{2} |0\rangle \langle 0| + \frac{1}{2} |1\rangle \langle 1|$ is the input of the quantum coordinate channel $\mathcal{E}_N^{(i)}$, which is also the i th input of the quantum combined channel \mathcal{E}_N . $S(\cdot)$ is von Neumann entropy, and $H(\cdot)$ is Shannon entropy. Since classical coordinate channels $\{W_N^{(i)}\}$ polarize, quantum coordinate channels $\{\mathcal{E}_N^{(i)}\}$ polarize as well.

Proof For $W_N^{(i)}$, its Shannon capacity is $I(W_N^{(i)}) = H(y_1^N u_1^{i-1}) - H(y_1^N u_1^i) + H(u_i)$. To calculate the Shannon capacity of $W_N^{(i)}$, we should first calculate $H(y_1^N u_1^{i-1})$ which is the Shannon entropy of the output $y_1^N u_1^{i-1}$ of $W_N^{(i)}$

$$H(y_1^N u_1^{i-1}) = \sum_{y_1^N u_1^{i-1} \in \mathcal{Y}^N \times \mathcal{X}^{i-1}} -p(y_1^N, u_1^{i-1}) \log_2 p(y_1^N, u_1^{i-1}) \quad (57)$$

where $p(y_1^N, u_1^{i-1}) = \frac{1}{2} W_N^{(i)}(y_1^N, u_1^{i-1} | u_i = 0) + \frac{1}{2} W_N^{(i)}(y_1^N, u_1^{i-1} | u_i = 1)$. Notice that by Proposition 9, we have $W_N^{(i)}(y_1^N, 0_1^{i-1} | u_i) = W_N^{(i)}((u_1^{i-1}, 0, 0_{i+1}^N) G_N \cdot y_1^N, 0_1^{i-1} \oplus u_1^{i-1} | u_i)$ for all $u_1^{i-1} \in \mathcal{X}^{i-1}$, hence

$$H(y_1^N u_1^{i-1}) = 2^{i-1} \sum_{y_1^N \in \mathcal{Y}^N} -p(y_1^N, 0_1^{i-1}) \log_2 p(y_1^N, 0_1^{i-1}) \quad (58)$$

and

$$\sum_{y_1^N u_1^{i-1} \in \mathcal{Y}^N \times \mathcal{X}^{i-1}} p(y_1^N, u_1^{i-1}) = 2^{i-1} \sum_{y_1^N \in \mathcal{Y}^N} p(y_1^N, 0_1^{i-1}) = 1 \quad (59)$$

Now, we calculate $S(\rho^{V_1^N, R_1^{i-1}})$, the von Neumann entropy of the output state $\rho^{V_1^N, R_1^{i-1}}$ of $\mathcal{E}_N^{(i)}$. By Eq. (46), we have

$$S(\rho^{V_1^N, R_1^{i-1}}) = - \sum_m p(|m\rangle) \log_2 p(|m\rangle) \quad (60)$$

where $p(|m\rangle) = \frac{1}{2} Pr_N^{(i)}(|m\rangle || 0) + \frac{1}{2} Pr_N^{(i)}(|m\rangle || 1)$. By Proposition 9, we have

$$\begin{aligned} Pr_N^{(i)}(|m\rangle || Q_i) &= Pr_N^{(i)} \left(\sum_{\substack{Q_1^{i-1} \\ = R_1^{i-1} \\ \in \mathcal{X}^{i-1}}} \frac{1}{2^{\frac{i-1}{2}}} |(Q_1^{i-1}, 0, 0_{i+1}^N) G_N \cdot V_1^N, R_1^{i-1}\rangle || Q_i \right) \\ &= 2^{i-1} W_N^{(i)}(y_1^N, 0_1^{i-1} | u_i) \end{aligned} \quad (61)$$

for all $y_1^N = V_1^N \in \mathcal{Y}^N$ and $Q_1^N = u_1^N \in \mathcal{X}^N$. Thus $S(\rho^{V_1^N, R_1^{i-1}})$ can be rewritten as

$$\begin{aligned} S(\rho^{V_1^N, R_1^{i-1}}) &= - \sum_{y_1^N \in \mathcal{Y}^N} 2^{i-1} p(y_1^N, 0_1^{i-1}) \log_2 \left[2^{i-1} p(y_1^N, 0_1^{i-1}) \right] \\ &= -2^{i-1} \sum_{y_1^N \in \mathcal{Y}^N} p(y_1^N, 0_1^{i-1}) \log_2 p(y_1^N, 0_1^{i-1}) - (i-1) \times 2^{i-1} \sum_{y_1^N \in \mathcal{Y}^N} p(y_1^N, 0_1^{i-1}) \\ &= H(y_1^N u_1^{i-1}) - (i-1) \end{aligned} \tag{62}$$

Using the same method, we have $S(\rho^{V_1^N, R_1^i}) = H(y_1^N u_1^i) - i$. Thus $I(\rho^{Q_i}, \mathcal{E}_N^{(i)}) = S(\rho^{V_1^N, R_1^{i-1}}) - S(\rho^{V_1^N, R_1^i}) = H(y_1^N u_1^{i-1}) - H(y_1^N u_1^i) + 1$. Notice that $H(u_i) = H(\frac{1}{2}) = 1$, thus we have

$$I(\rho^{Q_i}, \mathcal{E}_N^{(i)}) = H(y_1^N u_1^{i-1}) - H(y_1^N u_1^i) + H(u_i) = I(W_N^{(i)}) \tag{63}$$

which completes the proof. \square

5 Conclusion

The core of this paper is to prove that there is a polarization phenomenon in quantum channels similar to classical channel polarization. To prove this, we first define BTPM, and show how to use BTPM to determine a set of operation elements of a quantum channel. Then we use BTPM to define QSC and QQSC, and prove that the MSLCI of two-dimensional-input QQSC is its symmetric coherent information, which was not proved before our work. After this, we introduce the quantum channel combining and splitting, and obtain the quantum combined channel \mathcal{E}_N and coordinate channels $\{\mathcal{E}_N^{(i)}\}$. It has been proved in Sect. 3 that if the primal channel \mathcal{E} is a two-dimensional-input QSC, then \mathcal{E}_N is a two-dimensional-input QSC and $\{\mathcal{E}_N^{(i)}\}$ are two-dimensional-input QQSCs. Based on the above work, we prove that the MSLCI of the coordinate channels will polarize – some of them tend to 1 while the others tend to 0 with the increase of N , and the ratio of the former to N equals to the MSLCI of the primal channel \mathcal{E} , which completes the proof that there is a polarization phenomenon in quantum channels.

However, whether we can make use of this polarization phenomenon of quantum channels to design a quantum error correcting code which can achieve the MSLCI of QSC is still unknown.

Declarations

- Funding
Not applicable.
- Competing interests
The authors have no competing interests to declare that are relevant to the content of this article.

- Ethics approval
Not applicable.
- Consent to participate
Not applicable.
- Consent for publication
Not applicable.
- Availability of data and materials
Data sharing not applicable to this article as no datasets were generated or analysed during the current study.
- Code availability
Not applicable.
- Authors' contributions
All authors conceived the work, analysed the results and wrote the manuscript.

Appendix A A Particular Rule

Before proving Theorem 7, we make a particular rule which will be used in the second step of the proof.

This rule is used to label the operator elements of a channel through a one-to-one relationship between operator elements and output states. First, we fixed the input state $|Q_1^N\rangle$ of the quantum combined channel \mathcal{E}_N to $|0_1^N\rangle$, then arbitrary operator element $F_k \in \{F_k\}_{k=0, \dots, 2^N-1}$ of the N -copy channel $\mathcal{E}^{\otimes N}$ uniquely corresponds to a output state $|V_1^N\rangle$, $V_1^N \in \mathcal{Y}^N$, namely,

$$F_k |0_1^N G_N\rangle = \sqrt{Pr_N(|V_1^N\rangle ||0_1^N\rangle)} |V_1^N\rangle \quad (\text{A1})$$

By Definition 3, we have

$$F_k = E_{b_1}^1 \otimes E_{b_2}^2 \otimes \dots \otimes E_{b_N}^N \quad (\text{A2})$$

the subscript k of F_k is the decimal number of the binary sequence $b_1 b_2 \dots b_N$.

To further understanding this rule, we take 2-copy channel $\mathcal{E}^{\otimes 2}$ for example, and primal channel \mathcal{E} is Bit flip channel whose operator elements are $\{E_0 = \sqrt{p}X, E_1 = \sqrt{1-p}I\}$. It's easy to obtain that four operator elements of $\mathcal{E}^{\otimes 2}$ are $F_0 = pX^1 \otimes X^2$, $F_1 = \sqrt{p(1-p)}X^1 \otimes I^2$, $F_2 = \sqrt{p(1-p)}I^1 \otimes X^2$ and $F_3 = (1-p)I^1 \otimes I^2$, respectively. Assume that the input state of primal channel \mathcal{E} will only take value from $|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ or $|1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Then the input space $\{|Q_1^2\rangle\}$ of the quantum combined channel \mathcal{E}_2 must be $\{|Q_1^2\rangle\} = \{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$, and the output space $\{|V_1^2\rangle\}$ of the quantum combined channel \mathcal{E}_2 must be $\{|V_1^2\rangle\} = \{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$, which means different operator element F_k , $0 \leq k \leq 3$, will map the input space $\{|Q_1^2\rangle\}$ to the same output space $\{|V_1^2\rangle\}$. Thus we fixed the input state to $|00\rangle$, and a one-to-one relationship between operator element F_k ($0 \leq k \leq 3$) and output state $|V_1^2\rangle$ of

the channel $\mathcal{E}^{\otimes 2}$ is established, namely, F_0 corresponds to $|11\rangle$, F_1 corresponds to $|10\rangle$, F_2 corresponds to $|01\rangle$ and F_3 corresponds to $|00\rangle$.

By using Theorem 6 and Eq. (A1), we have

$$F_k |Q_1^N G_N\rangle = \sqrt{Pr_N(|Q_1^N G_N \cdot V_1^N\rangle | |Q_1^N\rangle)} |Q_1^N G_N \cdot V_1^N\rangle \quad (\text{A3})$$

for all $Q_1^N \in \mathcal{X}^N$ and $V_1^N \in \mathcal{Y}^N$.

Appendix B Proof of Theorem 7

In this section, we prove Theorem 7 that the quantum coordinate channels $\{\mathcal{E}_N^{(i)}\}$ are QQSCs. At the second step of the proof, we use the particular rule that we make in Appendix A.

Proof In subsection 2.5, we define quantum coordinate channel $\mathcal{E}_N^{(i)}$, $1 \leq i \leq N$, whose input is ρ^{Q_i} and output is $\rho^{V_1^N, R_1^{i-1}}$.

1. The first step of the proof: obtain the general form of density operator $\rho^{V_1^N, R_1^{i-1}}$ of quantum joint system V_1^N, R_1^{i-1} .

Assume that each input state ρ^{Q_i} of the quantum combined channel \mathcal{E}_N is $\rho^{Q_i} = q|0\rangle\langle 0| + (1-q)|1\rangle\langle 1|$. Then we have

$$\begin{aligned} \rho^{Q_1^N} &= \rho^{Q_1} \otimes \dots \otimes \rho^{Q_N} \\ &= (q|0\rangle\langle 0| + (1-q)|1\rangle\langle 1|)^{\otimes N} \\ &= \sum_{Q_1^N \in \mathcal{X}^N} Pr(|Q_1^N\rangle\langle Q_1^N|) |Q_1^N\rangle\langle Q_1^N| \end{aligned} \quad (\text{B4})$$

where $Pr(|Q_1^N\rangle\langle Q_1^N|) = \prod_{i=1}^N Pr(|Q_i\rangle\langle Q_i|)$, alphabet $\mathcal{X} = \{0, 1\}$ and \mathcal{X}^N is the N-power extension alphabet of \mathcal{X} . Introduce reference system $\rho^{R_1^N} = \rho^{R_1} \otimes \dots \otimes \rho^{R_N}$ to purify $\rho^{Q_1^N}$, where $\rho^{R_1} = \dots = \rho^{R_N} = \rho^{Q_1} = \dots = \rho^{Q_N}$. We have

$$|\varphi_{Q_1^N, R_1^N}\rangle = \sum_{Q_1^N = R_1^N \in \mathcal{X}^N} \sqrt{Pr(|Q_1^N\rangle\langle Q_1^N|)} |Q_1^N, R_1^N\rangle \quad (\text{B5})$$

Then the density operator $\rho^{Q_1^N, R_1^N}$ of the joint system Q_1^N, R_1^N is

$$\begin{aligned} &\rho^{Q_1^N, R_1^N} \\ &= |\varphi_{Q_1^N, R_1^N}\rangle\langle \varphi_{Q_1^N, R_1^N}| \\ &= \sum_{\substack{Q_1^N = R_1^N \in \mathcal{X}^N \\ \tilde{Q}_1^N = \tilde{R}_1^N \in \mathcal{X}^N}} \sqrt{Pr(|Q_1^N\rangle\langle Q_1^N|)} |Q_1^N, R_1^N\rangle \sqrt{Pr(|\tilde{Q}_1^N\rangle\langle \tilde{Q}_1^N|)} |\tilde{Q}_1^N, \tilde{R}_1^N\rangle \end{aligned} \quad (\text{B6})$$

We use a unitary operator U_N which only acts on system Q_1^N to represent the encoding process $|Q_1^N\rangle \rightarrow |C_1^N\rangle$, and we have

$$\begin{aligned}
& \rho^{C_1^N, R_1^N} \\
&= U_N \rho^{Q_1^N R_1^N} U_N^\dagger \\
&= U_N \left(\sum_{\substack{Q_1^N = R_1^N \in \mathcal{X}^N \\ \tilde{Q}_1^N = \tilde{R}_1^N \in \mathcal{X}^N}} \sqrt{Pr(|Q_1^N\rangle\langle Q_1^N|)} |Q_1^N, R_1^N\rangle \sqrt{Pr(|\tilde{Q}_1^N\rangle\langle \tilde{Q}_1^N|)} \langle \tilde{Q}_1^N, \tilde{R}_1^N| \right) U_N^\dagger \\
&= \sum_{\substack{Q_1^N = R_1^N \in \mathcal{X}^N \\ \tilde{Q}_1^N = \tilde{R}_1^N \in \mathcal{X}^N}} \sqrt{Pr(|Q_1^N\rangle\langle Q_1^N|)} |Q_1^N G_N, R_1^N\rangle \sqrt{Pr(|\tilde{Q}_1^N\rangle\langle \tilde{Q}_1^N|)} \langle \tilde{Q}_1^N G_N, \tilde{R}_1^N| \\
&= \sum_{R_1^N \in \mathcal{X}^N} \sqrt{Pr(|Q_1^N\rangle\langle Q_1^N|)} |C_1^N, R_1^N\rangle \sum_{\tilde{R}_1^N \in \mathcal{X}^N} \sqrt{Pr(|\tilde{Q}_1^N\rangle\langle \tilde{Q}_1^N|)} \langle \tilde{C}_1^N, \tilde{R}_1^N|
\end{aligned} \tag{B7}$$

where $C_1^N = Q_1^N G_N$, $\tilde{C}_1^N = \tilde{Q}_1^N G_N$ and G_N is generator matrix.

The channel $\mathcal{E}^{\otimes N}$, whose operator elements are $\{F_k\}_{k=0, \dots, 2^N-1}$, follows the encoding process $|Q_1^N\rangle \rightarrow |C_1^N\rangle$. Then the density operator $\rho^{V_1^N, R_1^N}$ of the output of the channel $\mathcal{E}^{\otimes N}$ is

$$\begin{aligned}
\rho^{V_1^N, R_1^N} &= \sum_{k=0}^{2^N-1} F_k \rho^{C_1^N, R_1^N} F_k^\dagger \\
&= \sum_{k=0}^{2^N-1} F_k \sum_{Q_1^N = R_1^N \in \mathcal{X}^N} \sqrt{Pr(|Q_1^N\rangle\langle Q_1^N|)} |Q_1^N G_N, R_1^N\rangle \\
&\quad \times \sum_{\tilde{Q}_1^N = \tilde{R}_1^N \in \mathcal{X}^N} \sqrt{Pr(|\tilde{Q}_1^N\rangle\langle \tilde{Q}_1^N|)} \langle \tilde{Q}_1^N G_N, \tilde{R}_1^N| F_k^\dagger
\end{aligned} \tag{B8}$$

Notice that the channel $\mathcal{E}^{\otimes N}$ is the last layer of the channel \mathcal{E}_N , so the density operator $\rho^{V_1^N, R_1^N}$ is also the output of the channel \mathcal{E}_N . Then we perform partial

trace over the system R_i^N and obtain

$$\begin{aligned}
& \rho_{Q_1^N, R_1^N}^{V_1^N, R_1^N} \\
&= \text{tr}_{R_i^N} \left[\sum_{k=0}^{2^N-1} F_k \sum_{Q_1^N=R_1^N \in \mathcal{X}^N} \sqrt{\text{Pr}(|Q_1^N\rangle\langle Q_1^N|)} |Q_1^N G_N, R_1^N\rangle \right. \\
&\times \left. \sum_{\tilde{Q}_1^N=\tilde{R}_1^N \in \mathcal{X}^N} \sqrt{\text{Pr}(|\tilde{Q}_1^N\rangle\langle \tilde{Q}_1^N|)} \langle \tilde{Q}_1^N G_N, \tilde{R}_1^N | F_k^\dagger \right] \\
&= \text{tr}_{R_i^N} \left[\sum_{k=0}^{2^N-1} F_k \sum_{Q_i^N=R_i^N \in \mathcal{X}^{N-i+1}} \right. \\
&\times \sum_{Q_1^{i-1}=R_1^{i-1} \in \mathcal{X}^{i-1}} \sqrt{\text{Pr}(|Q_1^{i-1}\rangle\langle Q_1^{i-1}|)} \text{Pr}(|Q_i^N\rangle\langle Q_i^N|) |Q_1^N G_N, R_1^{i-1}, R_i^N\rangle \\
&\times \left. \sum_{\tilde{Q}_1^{i-1}=\tilde{R}_1^{i-1} \in \mathcal{X}^{i-1}} \sqrt{\text{Pr}(|\tilde{Q}_1^{i-1}\rangle\langle \tilde{Q}_1^{i-1}|)} \text{Pr}(|Q_i^N\rangle\langle Q_i^N|) \langle \tilde{Q}_1^N G_N, \tilde{R}_1^{i-1}, R_i^N | F_k^\dagger \right] \\
&= \sum_{k=0}^{2^N-1} F_k \left[\sum_{Q_i^N \in \mathcal{X}^{N-i+1}} \text{Pr}(|Q_i^N\rangle\langle Q_i^N|) \right. \\
&\times \sum_{Q_1^{i-1}=R_1^{i-1} \in \mathcal{X}^{i-1}} \sqrt{\text{Pr}(|Q_1^{i-1}\rangle\langle Q_1^{i-1}|)} |Q_1^N G_N, R_1^{i-1}\rangle \\
&\times \left. \sum_{\tilde{Q}_1^{i-1}=\tilde{R}_1^{i-1} \in \mathcal{X}^{i-1}} \sqrt{\text{Pr}(|\tilde{Q}_1^{i-1}\rangle\langle \tilde{Q}_1^{i-1}|)} \langle \tilde{Q}_1^N G_N, \tilde{R}_1^{i-1} | F_k^\dagger \right] \tag{B9}
\end{aligned}$$

Eq. (B9) guarantees that

$$\sum_{Q_1^{i-1}=R_1^{i-1} \in \mathcal{X}^{i-1}} \sqrt{\text{Pr}(|Q_1^{i-1}\rangle\langle Q_1^{i-1}|)} |Q_1^N G_N, R_1^{i-1}\rangle \tag{B10}$$

must be a unit vector, since it is easy to verify $\sum_{Q_1^{i-1} \in \mathcal{X}^{i-1}} \text{Pr}(|Q_1^{i-1}\rangle\langle Q_1^{i-1}|) = 1$. Divide the Eq. (B9) into two parts: $Q_i = 0$ and $Q_i = 1$, we have

$$\rho_{V_1^N, R_1^N}^{V_1^N, R_1^N} = \rho_{V_1^N, R_1^{i-1}}^{(0)} + \rho_{V_1^N, R_1^{i-1}}^{(1)} \tag{B11}$$

where

$$\begin{aligned} \rho_{V_1^N, R_1^{i-1}}^{(0)} &= q \sum_{k=0}^{2^N-1} F_k \left[\sum_{Q_{i+1}^N \in \mathcal{X}^{N-i}} Pr \left(|Q_{i+1}^N\rangle \langle Q_{i+1}^N| \right) \right. \\ &\quad \times \sum_{Q_1^{i-1}=R_1^{i-1} \in \mathcal{X}^{i-1}} \sqrt{Pr \left(|Q_1^{i-1}\rangle \langle Q_1^{i-1}| \right)} |(Q_1^{i-1}, 0, Q_{i+1}^N)G_N, R_1^{i-1}\rangle \\ &\quad \times \left. \sum_{\tilde{Q}_1^{i-1}=\tilde{R}_1^{i-1} \in \mathcal{X}^{i-1}} \sqrt{Pr \left(|\tilde{Q}_1^{i-1}\rangle \langle \tilde{Q}_1^{i-1}| \right)} \langle (\tilde{Q}_1^{i-1}, 0, Q_{i+1}^N)G_N, \tilde{R}_1^{i-1} | \right] F_k^\dagger \end{aligned} \quad (\text{B12})$$

and

$$\begin{aligned} \rho_{V_1^N, R_1^{i-1}}^{(1)} &= (1-q) \sum_{k=0}^{2^N-1} F_k \left[\sum_{Q_{i+1}^N \in \mathcal{X}^{N-i}} Pr \left(|Q_{i+1}^N\rangle \langle Q_{i+1}^N| \right) \right. \\ &\quad \times \sum_{Q_1^{i-1}=R_1^{i-1} \in \mathcal{X}^{i-1}} \sqrt{Pr \left(|Q_1^{i-1}\rangle \langle Q_1^{i-1}| \right)} |(Q_1^{i-1}, 1, Q_{i+1}^N)G_N, R_1^{i-1}\rangle \\ &\quad \times \left. \sum_{\tilde{Q}_1^{i-1}=\tilde{R}_1^{i-1} \in \mathcal{X}^{i-1}} \sqrt{Pr \left(|\tilde{Q}_1^{i-1}\rangle \langle \tilde{Q}_1^{i-1}| \right)} \langle (\tilde{Q}_1^{i-1}, 1, Q_{i+1}^N)G_N, \tilde{R}_1^{i-1} | \right] F_k^\dagger \end{aligned} \quad (\text{B13})$$

For $\rho_{V_1^N, R_1^{i-1}}^{(0)}$ and $\rho_{V_1^N, R_1^{i-1}}^{(1)}$, we exchange summation order, and obtain

$$\begin{aligned} \rho_{V_1^N, R_1^{i-1}}^{(0)} &= q \sum_{Q_{i+1}^N \in \mathcal{X}^{N-i}} Pr \left(|Q_{i+1}^N\rangle \langle Q_{i+1}^N| \right) \\ &\quad \times \sum_{k=0}^{2^N-1} F_k \left[\sum_{Q_1^{i-1}=R_1^{i-1} \in \mathcal{X}^{i-1}} \sqrt{Pr \left(|Q_1^{i-1}\rangle \langle Q_1^{i-1}| \right)} |(Q_1^{i-1}, 0, Q_{i+1}^N)G_N, R_1^{i-1}\rangle \right. \\ &\quad \times \left. \sum_{\tilde{Q}_1^{i-1}=\tilde{R}_1^{i-1} \in \mathcal{X}^{i-1}} \sqrt{Pr \left(|\tilde{Q}_1^{i-1}\rangle \langle \tilde{Q}_1^{i-1}| \right)} \langle (\tilde{Q}_1^{i-1}, 0, Q_{i+1}^N)G_N, \tilde{R}_1^{i-1} | \right] F_k^\dagger \end{aligned} \quad (\text{B14})$$

and

$$\begin{aligned}
& \rho_{V_1^N, R_1^{i-1}}^{(1)} \\
&= (1-q) \sum_{Q_{i+1}^N \in \mathcal{X}^{N-i}} Pr \left(|Q_{i+1}^N\rangle \langle Q_{i+1}^N| \right) \\
&\times \sum_{k=0}^{2^N-1} F_k \left[\sum_{Q_1^{i-1}=R_1^{i-1} \in \mathcal{X}^{i-1}} \sqrt{Pr \left(|Q_1^{i-1}\rangle \langle Q_1^{i-1}| \right)} |(Q_1^{i-1}, 1, Q_{i+1}^N)G_N, R_1^{i-1}\rangle \right. \\
&\times \left. \sum_{\tilde{Q}_1^{i-1}=\tilde{R}_1^{i-1} \in \mathcal{X}^{i-1}} \sqrt{Pr \left(|\tilde{Q}_1^{i-1}\rangle \langle \tilde{Q}_1^{i-1}| \right)} \langle (\tilde{Q}_1^{i-1}, 1, Q_{i+1}^N)G_N, \tilde{R}_1^{i-1}| \right] F_k^\dagger
\end{aligned} \tag{B15}$$

2. The second step of the proof: prove that the density operator $\rho_{V_1^N, R_1^{i-1}}^{(1)}$ can be diagonalized with respect to a set of basis $\{|m'\rangle\}_{m=0, \dots, 2^N-1}$.

We will prove that density operators $\rho_{V_1^N, R_1^{i-1}}^{(0)}$ and $\rho_{V_1^N, R_1^{i-1}}^{(1)}$ can be diagonalized with respect to a same set of basis $\{|m'\rangle\}_{m=0, \dots, 2^N-1}$, namely,

$$\rho_{V_1^N, R_1^{i-1}}^{(0)} = q \sum_{m=0}^{2^N-1} Pr_N^{(i)} \left(|m'\rangle ||0\rangle \right) |m'\rangle \langle m'| \tag{B16}$$

$$\rho_{V_1^N, R_1^{i-1}}^{(1)} = (1-q) \sum_{m=0}^{2^N-1} Pr_N^{(i)} \left(|m'\rangle ||1\rangle \right) |m'\rangle \langle m'| \tag{B17}$$

We consider $\rho_{V_1^N, R_1^{i-1}}^{(0)}$ only, since the proof method of $\rho_{V_1^N, R_1^{i-1}}^{(1)}$ is the same as that of $\rho_{V_1^N, R_1^{i-1}}^{(0)}$. We first prove that the vector $|m'\rangle$ can be written as

$$|m'\rangle = \sum_{Q_1^{i-1}=R_1^{i-1} \in \mathcal{X}^{i-1}} \sqrt{Pr \left(|Q_1^{i-1}\rangle \langle Q_1^{i-1}| \right)} |(Q_1^{i-1}, 0, 0_{i+1}^N)G_N \cdot V_1^N, R_1^{i-1}\rangle \tag{B18}$$

Since for all $Q_{i+1}^N \in \mathcal{X}^{N-i}$, operation elements $\{F_k\}_{k=0, \dots, 2^N-1}$ will map vector

$$\sum_{Q_1^{i-1}=R_1^{i-1} \in \mathcal{X}^{i-1}} \sqrt{Pr \left(|Q_1^{i-1}\rangle \langle Q_1^{i-1}| \right)} |(Q_1^{i-1}, 0, Q_{i+1}^N)G_N, R_1^{i-1}\rangle \tag{B19}$$

to a same set of orthogonal basis $\{|m'\rangle\}_{m=0, \dots, 2^N-1}$, which contains 2^N basis vectors. Thus, without losing generality, we can let $Q_{i+1}^N = 0_{i+1}^N$. Using Eq. (A1), Eq.

(A3) and Theorem 6, we have

$$\begin{aligned}
F_k & \sum_{Q_1^{i-1}=R_1^{i-1} \in \mathcal{X}^{i-1}} \sqrt{\Pr(|Q_1^{i-1}\rangle \langle Q_1^{i-1}|)} |(Q_1^{i-1}, 0, Q_{i+1}^N)G_N, R_1^{i-1}\rangle \\
& = \sum_{Q_1^{i-1}=R_1^{i-1} \in \mathcal{X}^{i-1}} \sqrt{\Pr_N(|(Q_1^{i-1}, 0, 0_{i+1}^N)G_N \cdot V_1^N\rangle ||Q_1^{i-1}, 0, 0_{i+1}^N\rangle)} \\
& \times \sqrt{\Pr(|Q_1^{i-1}\rangle \langle Q_1^{i-1}|)} |(Q_1^{i-1}, 0, 0_{i+1}^N)G_N \cdot V_1^N, R_1^{i-1}\rangle \\
& = \sqrt{\Pr_N(|V_1^N\rangle ||0_1^N\rangle)} \\
& \times \sum_{Q_1^{i-1}=R_1^{i-1} \in \mathcal{X}^{i-1}} \sqrt{\Pr(|Q_1^{i-1}\rangle \langle Q_1^{i-1}|)} |(Q_1^{i-1}, 0, 0_{i+1}^N)G_N \cdot V_1^N, R_1^{i-1}\rangle \\
& = \sqrt{\Pr_N(|V_1^N\rangle ||0_1^N\rangle)} |m'\rangle
\end{aligned} \tag{B20}$$

which proves the Eq. (B18).

Observe Eq. (B20), there is a one-to-one relationship between F_k and V_1^N , thus sum over all F_k is sum over all V_1^N and Eq. (B14) can be rewritten as

$$\begin{aligned}
\rho_{V_1^N, R_1^{i-1}}^{(0)} & = q \sum_{Q_{i+1}^N \in \mathcal{X}^{N-i}} \Pr(|Q_{i+1}^N\rangle \langle Q_{i+1}^N|) \sum_{V_1^N \in \mathcal{Y}^N} \\
& \times \sum_{Q_1^{i-1}=R_1^{i-1} \in \mathcal{X}^{i-1}} \Pr_N(|(Q_1^{i-1}, 0, 0_{i+1}^N)G_N \cdot V_1^N\rangle ||Q_1^{i-1}, 0, Q_{i+1}^N\rangle) \\
& \times \sqrt{\Pr(|Q_1^{i-1}\rangle \langle Q_1^{i-1}|)} |(Q_1^{i-1}, 0, 0_{i+1}^N)G_N \cdot V_1^N, R_1^{i-1}\rangle \\
& \times \sum_{\tilde{Q}_1^{i-1}=\tilde{R}_1^{i-1} \in \mathcal{X}^{i-1}} \sqrt{\Pr(|\tilde{Q}_1^{i-1}\rangle \langle \tilde{Q}_1^{i-1}|)} \langle (\tilde{Q}_1^{i-1}, 0, 0_{i+1}^N)G_N \cdot V_1^N, \tilde{R}_1^{i-1} | \\
& = q \sum_{Q_{i+1}^N \in \mathcal{X}^{N-i}} \Pr(|Q_{i+1}^N\rangle \langle Q_{i+1}^N|) \sum_{V_1^N \in \mathcal{Y}^N} \Pr_N(|V_1^N\rangle ||0_1^{i-1}, 0, Q_{i+1}^N\rangle) \\
& \times \sum_{Q_1^{i-1}=R_1^{i-1} \in \mathcal{X}^{i-1}} \sqrt{\Pr(|Q_1^{i-1}\rangle \langle Q_1^{i-1}|)} |(Q_1^{i-1}, 0, 0_{i+1}^N)G_N \cdot V_1^N, R_1^{i-1}\rangle \\
& \times \sum_{\tilde{Q}_1^{i-1}=\tilde{R}_1^{i-1} \in \mathcal{X}^{i-1}} \sqrt{\Pr(|\tilde{Q}_1^{i-1}\rangle \langle \tilde{Q}_1^{i-1}|)} \langle (\tilde{Q}_1^{i-1}, 0, 0_{i+1}^N)G_N \cdot V_1^N, \tilde{R}_1^{i-1} |
\end{aligned} \tag{B21}$$

Here we use the fact that $\Pr_N(|V_1^N\rangle ||Q_1^N\rangle) = \Pr_N(|a_1^N G_N \cdot V_1^N\rangle ||Q_1^N \oplus a_1^N\rangle)$ which is according to Theorem 6, so let $a_1^N = Q_1^{i-1}, 0, 0_{i+1}^N$, we have

$$\Pr_N(|V_1^N\rangle ||0_1^{i-1}, 0, Q_{i+1}^N\rangle) = \Pr_N(|(Q_1^{i-1}, 0, 0_{i+1}^N)G_N \cdot V_1^N\rangle ||Q_1^{i-1}, 0, Q_{i+1}^N\rangle) \tag{B22}$$

For Eq. (B21), we exchange summation order and obtain

$$\begin{aligned}
& \rho_{V_1^N, R_1^{i-1}}^{(0)} \\
&= q \sum_{V_1^N \in \mathcal{Y}^N} \left[\sum_{Q_{i+1}^N \in \mathcal{X}^{N-i}} Pr \left(|Q_{i+1}^N\rangle \langle Q_{i+1}^N| \right) Pr_N \left(|V_1^N\rangle \| 0_1^{i-1}, 0, Q_{i+1}^N \right) \right. \\
&\times \sum_{Q_1^{i-1} = R_1^{i-1} \in \mathcal{X}^{i-1}} \sqrt{Pr \left(|Q_1^{i-1}\rangle \langle Q_1^{i-1}| \right)} |(Q_1^{i-1}, 0, 0_{i+1}^N) G_N \cdot V_1^N, R_1^{i-1}\rangle \\
&\times \left. \sum_{\tilde{Q}_1^{i-1} = \tilde{R}_1^{i-1} \in \mathcal{X}^{i-1}} \sqrt{Pr \left(|\tilde{Q}_1^{i-1}\rangle \langle \tilde{Q}_1^{i-1}| \right)} \langle (\tilde{Q}_1^{i-1}, 0, 0_{i+1}^N) G_N \cdot V_1^N, \tilde{R}_1^{i-1} | \right] \\
&= q \sum_{m=0}^{2^N-1} Pr_N^{(i)} \left(|m'\rangle \| 0 \right) |m'\rangle \langle m'|
\end{aligned} \tag{B23}$$

where

$$|m'\rangle = \sum_{Q_1^{i-1} = R_1^{i-1} \in \mathcal{X}^{i-1}} \sqrt{Pr \left(|Q_1^{i-1}\rangle \langle Q_1^{i-1}| \right)} |(Q_1^{i-1}, 0, 0_{i+1}^N) G_N \cdot V_1^N, R_1^{i-1}\rangle \tag{B24}$$

and

$$Pr_N^{(i)} \left(|m'\rangle \| 0 \right) = \sum_{Q_{i+1}^N \in \mathcal{X}^{N-i}} Pr \left(|Q_{i+1}^N\rangle \langle Q_{i+1}^N| \right) Pr_N \left(|V_1^N\rangle \| 0_1^{i-1}, 0, Q_{i+1}^N \right) \tag{B25}$$

$Pr_N^{(i)} \left(|m'\rangle \| 0 \right)$ is the transition probability which means the probability of the input state $|0\rangle \langle 0|$ changing into $|m'\rangle \langle m'|$. Using Eq. (B22), then Eq. (B25) can be rewritten as

$$\begin{aligned}
Pr_N^{(i)} \left(|m'\rangle \| 0 \right) &= \sum_{Q_1^{i-1} \in \mathcal{X}^{i-1}} \frac{1}{2^{i-1}} \sum_{Q_{i+1}^N \in \mathcal{X}^{N-i}} Pr \left(|Q_{i+1}^N\rangle \langle Q_{i+1}^N| \right) \\
&\times Pr_N \left(\left| (Q_1^{i-1}, 0, 0_{i+1}^N) G_N \cdot V_1^N \right\| |Q_1^{i-1}, 0, Q_{i+1}^N\rangle \right)
\end{aligned} \tag{B26}$$

Using the same method, Eq. (B17) can be easily proved, and we have

$$\begin{aligned}
Pr_N^{(i)} \left(|m'\rangle \| 1 \right) &= \sum_{Q_1^{i-1} \in \mathcal{X}^{i-1}} \frac{1}{2^{i-1}} \sum_{Q_{i+1}^N \in \mathcal{X}^{N-i}} Pr \left(|Q_{i+1}^N\rangle \langle Q_{i+1}^N| \right) \\
&\times Pr_N \left(\left| (Q_1^{i-1}, 0, 0_{i+1}^N) G_N \cdot V_1^N \right\| |Q_1^{i-1}, 1, Q_{i+1}^N\rangle \right) \\
&= \sum_{Q_{i+1}^N \in \mathcal{X}^{N-i}} Pr \left(|Q_{i+1}^N\rangle \langle Q_{i+1}^N| \right) Pr_N \left(|V_1^N\rangle \| 0_1^{i-1}, 1, Q_{i+1}^N \right)
\end{aligned} \tag{B27}$$

Thus, the basis transition probabilities can be uniformly expressed as

$$\begin{aligned}
 Pr_N^{(i)}(|m'\rangle||Q_i\rangle) &= \sum_{Q_1^{i-1} \in \mathcal{X}^{i-1}} \frac{1}{2^{i-1}} \sum_{Q_{i+1}^N \in \mathcal{X}^{N-i}} Pr(|Q_{i+1}^N\rangle\langle Q_{i+1}^N|) \\
 &\quad \times Pr_N(|(Q_1^{i-1}, 0, 0_{i+1}^N)\rangle G_N \cdot V_1^N ||Q_1^{i-1}, Q_i, Q_{i+1}^N\rangle) \\
 &= \sum_{Q_{i+1}^N \in \mathcal{X}^{N-i}} Pr(|Q_{i+1}^N\rangle\langle Q_{i+1}^N|) Pr_N(|V_1^N\rangle||0_1^{i-1}, Q_i, Q_{i+1}^N\rangle)
 \end{aligned} \tag{B28}$$

3. The third step of the proof: use Arikan's method to prove the basis transition probability matrix is symmetric.

Next, we will prove that the basis transition probability matrix is symmetric. We will refer to the proof method which Arikan used to prove that classical coordinate channels $\{W_N^{(i)}\}$ are symmetric if the primal binary-input discrete memoryless channel W is symmetric.

By Theorem 6, we have

$$\begin{aligned}
 &Pr_N(|(Q_1^{i-1}, 0, 0_{i+1}^N)\rangle G_N \cdot V_1^N ||Q_1^{i-1}, Q_i, Q_{i+1}^N\rangle) \\
 &= Pr_N(|(a_1^{i-1}, 1, a_{i+1}^N)G_N \cdot (Q_1^{i-1}, 0, 0_{i+1}^N)G_N \cdot V_1^N | \\
 &\quad |(Q_1^{i-1}, Q_i, Q_{i+1}^N) \oplus (a_1^{i-1}, 1, a_{i+1}^N)\rangle)
 \end{aligned} \tag{B29}$$

for arbitrary $(a_1^{i-1}, 1, a_{i+1}^N) \in \mathcal{X}^N$, thus the Eq. (B28) can be rewritten as

$$\begin{aligned}
 &Pr_N^{(i)}(|m'\rangle||Q_i\rangle) \\
 &= \sum_{Q_1^{i-1} \in \mathcal{X}^{i-1}} \frac{1}{2^{i-1}} \sum_{Q_{i+1}^N \in \mathcal{X}^{N-i}} Pr(|Q_{i+1}^N\rangle\langle Q_{i+1}^N|) \\
 &\quad \times Pr_N(|(Q_1^{i-1}, 0, 0_{i+1}^N)G_N \cdot V_1^N ||Q_1^{i-1}, Q_i, Q_{i+1}^N\rangle) \\
 &= \sum_{Q_1^{i-1} \in \mathcal{X}^{i-1}} \frac{1}{2^{i-1}} \sum_{Q_{i+1}^N \in \mathcal{X}^{N-i}} Pr(|Q_{i+1}^N\rangle\langle Q_{i+1}^N|) \\
 &\quad \times Pr_N(|(a_1^{i-1}, 1, a_{i+1}^N)G_N \cdot (Q_1^{i-1}, 0, 0_{i+1}^N)G_N \cdot V_1^N | \\
 &\quad |(Q_1^{i-1}, Q_i, Q_{i+1}^N) \oplus (a_1^{i-1}, 1, a_{i+1}^N)\rangle) \\
 &= \sum_{Q_1^{i-1} \in \mathcal{X}^{i-1}} \frac{1}{2^{i-1}} \sum_{Q_{i+1}^N \in \mathcal{X}^{N-i}} Pr(|Q_{i+1}^N\rangle\langle Q_{i+1}^N|) \\
 &\quad \times Pr_N(|(a_1^{i-1}, 1, a_{i+1}^N \oplus Q_1^{i-1}, 0, 0_{i+1}^N)G_N \cdot V_1^N | \\
 &\quad |(Q_1^{i-1}, Q_i, Q_{i+1}^N) \oplus (a_1^{i-1}, 1, a_{i+1}^N)\rangle) \\
 &= Pr_N^{(i)}\left(\sum_{Q_1^{i-1}=R_1^{i-1} \in \mathcal{X}^{i-1}} \sqrt{Pr(|Q_1^{i-1}\rangle\langle Q_1^{i-1}|)} \right. \\
 &\quad \left. \times |(a_1^{i-1}, 1, a_{i+1}^N \oplus Q_1^{i-1}, 0, 0_{i+1}^N)G_N \cdot V_1^N, R_1^{i-1} \oplus a_1^{i-1} ||Q_i \oplus 1\rangle)\right)
 \end{aligned} \tag{B30}$$

Substitute Eq. (B18) into $Pr_N^{(i)}(|m'\rangle||Q_i\rangle)$, and connect with Eq. (B30), we have

$$\begin{aligned}
& Pr_N^{(i)}(|m'\rangle||Q_i\rangle) \\
&= Pr_N^{(i)} \left(\sum_{\substack{Q_1^{i-1} \\ =R_1^{i-1} \\ \in \mathcal{X}^{i-1}}} \sqrt{Pr(|Q_1^{i-1}\rangle\langle Q_1^{i-1}|)} |(Q_1^{i-1}, 0, 0_{i+1}^N) G_N \cdot V_1^N, R_1^{i-1}\rangle||Q_i\rangle \right) \\
&= Pr_N^{(i)} \left(\sum_{Q_1^{i-1}=R_1^{i-1} \in \mathcal{X}^{i-1}} \sqrt{Pr(|Q_1^{i-1}\rangle\langle Q_1^{i-1}|)} \right. \\
&\quad \left. \times |(a_1^{i-1}, 1, a_{i+1}^N \oplus Q_1^{i-1}, 0, 0_{i+1}^N) G_N \cdot V_1^N, R_1^{i-1} \oplus a_1^{i-1}\rangle||Q_i \oplus 1\rangle \right) \tag{B31}
\end{aligned}$$

Here we take $a_1^N = (a_1^{i-1}, 1, a_{i+1}^N)$, and the proof is completed. The Eq. (B31) means arbitrary row of the BTPM of the quantum coordinate channel $\mathcal{E}_N^{(i)}$ is a permutation of another row. \square

Appendix C Proof of Proposition 9

In this section, we prove Proposition 9 that we can derive $Pr_N^{(i)}(|m\rangle||Q_i\rangle)$ from the TPM of classical coordinate channels $W_N^{(i)}$.

Proof According to Arikan's theorem[6], the transition probabilities of classical coordinate channels $\{W_N^{(i)}\}$ are

$$\begin{aligned}
& W_N^{(i)}(y_1^N, u_1^{i-1} | u_i) \\
&= \sum_{u_{i+1}^N \in \mathcal{X}^{N-i}} \frac{1}{2^{N-1}} W_N(y_1^N | u_1^N) \\
&= \sum_{u_{i+1}^N \in \mathcal{X}^{N-i}} \frac{1}{2^{N-1}} W_N(a_1^N G_N \cdot y_1^N | u_1^N \oplus a_1^N) \\
&= W_N^{(i)}(a_1^N G_N \cdot y_1^N, u_1^{i-1} \oplus a_1^{i-1} | u_i \oplus a_i) \tag{C32}
\end{aligned}$$

and Arikan has proved that classical combined channel W_N and classical coordinate channels $\{W_N^{(i)}\}$ are all symmetric, which satisfies

$$W_N(y_1^N | 0_1^{i-1}, u_i, u_{i+1}^N) = W_N\left(\left(u_1^{i-1}, 0, 0_{i+1}^N\right) G_N \cdot y_1^N | u_1^{i-1}, u_i, u_{i+1}^N\right) \tag{C33}$$

and

$$W_N^{(i)}(y_1^N, 0_1^{i-1} | u_i) = W_N^{(i)}\left(\left(u_1^{i-1}, 0, 0_{i+1}^N\right) G_N \cdot y_1^N, 0_1^{i-1} \oplus u_1^{i-1} | u_i\right) \tag{C34}$$

for all $u_1^{i-1} \in \mathcal{X}^{i-1}$.

Using Theorem 6, Proposition 8 and Eq (C33), we have

$$\begin{aligned} Pr_N \left(|V_1^N\rangle |0_1^{i-1}, Q_i, Q_{i+1}^N\rangle \right) &= W_N \left(y_1^N | 0_1^{i-1}, u_i, u_{i+1}^N \right) \\ &= W_N \left((u_1^{i-1}, 0, 0_{i+1}^N) G_N \cdot y_1^N | u_1^{i-1}, u_i, u_{i+1}^N \right) \\ &= Pr_N \left(|(Q_1^{i-1}, 0, 0_{i+1}^N) G_N \cdot V_1^N\rangle |Q_1^{i-1}, Q_i, Q_{i+1}^N\rangle \right) \end{aligned} \quad (\text{C35})$$

for all $V_1^N = y_1^N \in \mathcal{Y}^N$ and $Q_1^N = u_1^N \in \mathcal{X}^N$.

Substitute Eq. (C35) and Eq. (C34) into Eq. (47), we have

$$\begin{aligned} Pr_N^{(i)}(|m\rangle |Q_i\rangle) &= \sum_{u_1^{i-1} \in \mathcal{X}^{i-1}} \frac{1}{2^{N-1}} \\ &\times \sum_{u_{i+1}^N \in \mathcal{X}^{N-i}} W_N \left((u_1^{i-1}, 0, 0_{i+1}^N) G_N \cdot y_1^N | u_1^{i-1}, u_i, u_{i+1}^N \right) \\ &= \frac{2^{i-1}}{2^{N-1}} \sum_{u_{i+1}^N \in \mathcal{X}^{N-i}} W_N \left(y_1^N | 0_1^{i-1}, u_i, u_{i+1}^N \right) \\ &= 2^{i-1} W_N^{(i)} \left(y_1^N, 0_1^{i-1} | u_i \right) \end{aligned} \quad (\text{C36})$$

The Eq. (C36) means we can derive $Pr_N^{(i)}(|m\rangle |Q_i\rangle)$ from the TPM of classical coordinate channels $W_N^{(i)}$, which completes the proof. \square

References

- [1] Shor, P.W.: Scheme for reducing decoherence in quantum computer memory. *Phys. Rev. A* **52**, 2493–2496 (1995). <https://doi.org/10.1103/PhysRevA.52.R2493>
- [2] Steane, A.M.: Error correcting codes in quantum theory. *Phys. Rev. Lett.* **77**, 793–797 (1996). <https://doi.org/10.1103/PhysRevLett.77.793>
- [3] Gallager, R.: Low-density parity-check codes. *IRE Transactions on Information Theory* **8**(1), 21–28 (1962). <https://doi.org/10.1109/TIT.1962.1057683>
- [4] MacKay, D.J.C.: Good error-correcting codes based on very sparse matrices. *IEEE Transactions on Information Theory* **45**(2), 399–431 (1999). <https://doi.org/10.1109/18.748992>
- [5] MacKay, D.J., Neal, R.M.: Near shannon limit performance of low density parity check codes. *Electronics letters* **32**(18), 1645 (1996)
- [6] Arikan, E.: Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels. *IEEE Transactions on Information Theory* **55**(7), 3051–3073 (2009). <https://doi.org/10.1109/TIT.2009.2021379>

- [7] Bravyi, S.B., Kitaev, A.Y.: Quantum codes on a lattice with boundary. arXiv preprint quant-ph/9811052 (1998)
- [8] Stephens, A.M.: Fault-tolerant thresholds for quantum error correction with the surface code. *Phys. Rev. A* **89**, 022321 (2014). <https://doi.org/10.1103/PhysRevA.89.022321>
- [9] Bullock, S.S., Brennen, G.K.: Qudit surface codes and gauge theory with finite cyclic groups. *Journal of Physics A: Mathematical and Theoretical* **40**(13), 3481 (2007)
- [10] Zémor, G.: On cayley graphs, surface codes, and the limits of homological coding for quantum error correction. In: *International Conference on Coding and Cryptology*, pp. 259–273 (2009). Springer
- [11] Wang, D.S., Fowler, A.G., Stephens, A.M., Hollenberg, L.C.L.: Threshold error rates for the toric and surface codes. arXiv preprint arXiv:0905.0531 (2009)
- [12] Fowler, A.G., Stephens, A.M., Groszkowski, P.: High-threshold universal quantum computation on the surface code. *Phys. Rev. A* **80**, 052312 (2009). <https://doi.org/10.1103/PhysRevA.80.052312>
- [13] Bravyi, S., Duclos-Cianci, G., Poulin, D., Suchara, M.: Subsystem surface codes with three-qubit check operators. arXiv preprint arXiv:1207.1443 (2012)
- [14] Ghosh, J., Fowler, A.G., Geller, M.R.: Surface code with decoherence: An analysis of three superconducting architectures. *Physical Review A* **86**(6), 062318 (2012)
- [15] Fowler, A.G.: Proof of finite surface code threshold for matching. *Phys. Rev. Lett.* **109**, 180502 (2012). <https://doi.org/10.1103/PhysRevLett.109.180502>
- [16] Wootton, J.R., Loss, D.: High threshold error correction for the surface code. *Phys. Rev. Lett.* **109**, 160503 (2012). <https://doi.org/10.1103/PhysRevLett.109.160503>
- [17] Fowler, A.G., Whiteside, A.C., Hollenberg, L.C.L.: Towards practical classical processing for the surface code: Timing analysis. *Phys. Rev. A* **86**, 042313 (2012). <https://doi.org/10.1103/PhysRevA.86.042313>
- [18] Fowler, A.G., Mariantoni, M., Martinis, J.M., Cleland, A.N.: Surface codes: Towards practical large-scale quantum computation. *Phys. Rev. A* **86**, 032324 (2012). <https://doi.org/10.1103/PhysRevA.86.032324>

- [19] Fowler, A.G.: Optimal complexity correction of correlated errors in the surface code. arXiv preprint arXiv:1310.0863 (2013)
- [20] Barends, R., Kelly, J., Megrant, A., Veitia, A., Sank, D., Jeffrey, E., White, T.C., Mutus, J., Fowler, A.G., Campbell, B., *et al.*: Superconducting quantum circuits at the surface code threshold for fault tolerance. *Nature* **508**(7497), 500–503 (2014)
- [21] Hill, C.D., Peretz, E., Hile, S.J., House, M.G., Fuechsle, M., Rogge, S., Simmons, M.Y., Hollenberg, L.C.: A surface code quantum computer in silicon. *Science advances* **1**(9), 1500707 (2015)
- [22] Delfosse, N., Iyer, P., Poulin, D.: A linear-time benchmarking tool for generalized surface codes. arXiv preprint arXiv:1611.04256 (2016)
- [23] Versluis, R., Poletto, S., Khammassi, N., Tarasinski, B., Haider, N., Michalak, D.J., Bruno, A., Bertels, K., DiCarlo, L.: Scalable quantum circuit and control for a superconducting surface code. *Phys. Rev. Applied* **8**, 034021 (2017). <https://doi.org/10.1103/PhysRevApplied.8.034021>
- [24] Huang, C., Ni, X., Zhang, F., Newman, M., Ding, D., Gao, X., Wang, T., Zhao, H.-H., Wu, F., Zhang, G., *et al.*: Alibaba cloud quantum development platform: Surface code simulations with crosstalk. arXiv preprint arXiv:2002.08918 (2020)
- [25] Aharonov, D., Ben-Or, M.: Fault-tolerant quantum computation with constant error rate. *SIAM Journal on Computing* **38**(4), 1207–1282 (2008). <https://doi.org/10.1137/S0097539799359385>
- [26] Knill, E., Laflamme, R.: Concatenated quantum codes. arXiv preprint quant-ph/9608012 (1996)
- [27] Knill, E.: Quantum computing with realistically noisy devices. *Nature* **434**(7029), 39–44 (2005)
- [28] Gottesman, D.: Fault-tolerant quantum computation with constant overhead. arXiv preprint arXiv:1310.2984 (2013)
- [29] Tillich, J.-P., Zémor, G.: Quantum ldpc codes with positive rate and minimum distance proportional to the square root of the blocklength. *IEEE Transactions on Information Theory* **60**(2), 1193–1202 (2013)
- [30] Freedman, M.H., Hastings, M.B.: Quantum systems on non- k -hyperfinite complexes: A generalization of classical statistical mechanics on expander graphs. arXiv preprint arXiv:1301.1363 (2013)

- [31] Guth, L., Lubotzky, A.: Quantum error correcting codes and 4-dimensional arithmetic hyperbolic manifolds. *Journal of Mathematical Physics* **55**(8), 082202 (2014)
- [32] Kovalev, A.A., Pryadko, L.P.: Fault tolerance of quantum low-density parity check codes with sublinear distance scaling. *Physical Review A* **87**(2), 020304 (2013)
- [33] Hastings, M.B.: Decoding in hyperbolic spaces: Ldpc codes with linear rate and efficient error correction. arXiv preprint arXiv:1312.2546 (2013)
- [34] Breuckmann, N.P., Terhal, B.M.: Constructions and noise threshold of hyperbolic surface codes. *IEEE transactions on Information Theory* **62**(6), 3731–3744 (2016)
- [35] Breuckmann, N.P., Vuillot, C., Campbell, E., Krishna, A., Terhal, B.M.: Hyperbolic and semi-hyperbolic surface codes for quantum storage. *Quantum Science and Technology* **2**(3), 035007 (2017)
- [36] Breuckmann, N.P., Londe, V.: Single-shot decoding of linear rate ldpc quantum codes with high performance. *IEEE Transactions on Information Theory* **68**(1), 272–286 (2021)
- [37] Grospellier, A., Grouès, L., Krishna, A., Leverrier, A.: Combining hard and soft decoders for hypergraph product codes. *Quantum* **5**, 432 (2021)
- [38] Guo, Y., Lee, M.H., Zeng, G.: Polar quantum channel coding with optical multi-qubit entangling gates for capacity-achieving channels. *Quantum information processing* **12**(4), 1659–1676 (2013)
- [39] Renes, J.M., Dupuis, F., Renner, R.: Efficient polar coding of quantum information. *Physical Review Letters* **109**(5), 050504 (2012)
- [40] Wilde, M.M., Guha, S.: Polar codes for degradable quantum channels. *IEEE Transactions on Information Theory* **59**(7), 4718–4729 (2013)
- [41] Hirche, C.: Polar codes in quantum information theory. arXiv preprint arXiv:1501.03737 (2015)
- [42] Renes, J.M., Sutter, D., Dupuis, F., Renner, R.: Efficient quantum polar codes requiring no preshared entanglement. *IEEE Transactions on Information Theory* **61**(11), 6395–6414 (2015). <https://doi.org/10.1109/TIT.2015.2468084>
- [43] Hirche, C., Morgan, C., Wilde, M.M.: Polar codes in network quantum information theory. *IEEE Transactions on Information Theory* **62**(2), 915–924 (2016). <https://doi.org/10.1109/TIT.2016.2514319>

- [44] Dupuis, F., Goswami, A., Mhalla, M., Savin, V.: Purely quantum polar codes. In: 2019 IEEE Information Theory Workshop (ITW), pp. 1–5 (2019). <https://doi.org/10.1109/ITW44776.2019.8989387>
- [45] Babar, Z., Kaykac Egilmez, Z.B., Xiang, L., Chandra, D., Maunder, R.G., Ng, S.X., Hanzo, L.: Polar codes and their quantum-domain counterparts. *IEEE Communications Surveys Tutorials* **22**(1), 123–155 (2020). <https://doi.org/10.1109/COMST.2019.2937923>
- [46] Wilde, M.M., Guha, S.: Polar codes for classical-quantum channels. *IEEE Transactions on Information Theory* **59**(2), 1175–1187 (2013). <https://doi.org/10.1109/TIT.2012.2218792>
- [47] Wilde, M.M., Renes, J.M.: Quantum polar codes for arbitrary channels. In: 2012 IEEE International Symposium on Information Theory Proceedings, pp. 334–338 (2012). <https://doi.org/10.1109/ISIT.2012.6284203>
- [48] Goswami, A., Mhalla, M., Savin, V.: Quantum polarization of qudit channels. arXiv preprint arXiv:2101.10194 (2021)
- [49] Ramakrishnan, N., Iten, R., Scholz, V.B., Berta, M.: Computing quantum channel capacities. *IEEE Transactions on Information Theory* **67**(2), 946–960 (2021). <https://doi.org/10.1109/TIT.2020.3034471>
- [50] Gyongyosi, L., Imre, S., Nguyen, H.V.: A survey on quantum channel capacities. *IEEE Communications Surveys Tutorials* **20**(2), 1149–1205 (2018). <https://doi.org/10.1109/COMST.2017.2786748>
- [51] Holevo, A.S.: Quantum channel capacities. *Quantum Electronics* **50**(5), 440 (2020)
- [52] Smith, G.: Quantum channel capacities. In: 2010 IEEE Information Theory Workshop, pp. 1–5 (2010). <https://doi.org/10.1109/CIG.2010.5592851>
- [53] Holevo, A.S., Shirokov, M.E.: Mutual and coherent information for infinite-dimensional quantum channels. *Problems of information transmission* **46**(3), 201–218 (2010)
- [54] Bennett, C.H., Shor, P.W.: Quantum channel capacities. *Science* **303**(5665), 1784–1787 (2004)
- [55] Barnum, H., Nielsen, M.A., Schumacher, B.: Information transmission through a noisy quantum channel. *Phys. Rev. A* **57**, 4153–4175 (1998). <https://doi.org/10.1103/PhysRevA.57.4153>

- [56] Lloyd, S.: Capacity of the noisy quantum channel. *Phys. Rev. A* **55**, 1613–1622 (1997). <https://doi.org/10.1103/PhysRevA.55.1613>
- [57] Kretschmann, D., Werner, R.F.: Tema con variazioni: quantum channel capacity. *New Journal of Physics* **6**(1), 26 (2004)
- [58] Shor, P.W.: Capacities of quantum channels and how to find them. arXiv preprint quant-ph/0304102 (2003)
- [59] Javidian, M.A., Aggarwal, V., Bao, F., Jacob, Z.: Quantum entropic causal inference. arXiv preprint arXiv:2102.11764 (2021)
- [60] Schumacher, B., Nielsen, M.A.: Quantum data processing and error correction. *Phys. Rev. A* **54**, 2629–2635 (1996). <https://doi.org/10.1103/PhysRevA.54.2629>
- [61] Nielsen, M.A., Chuang, I.: *Quantum computation and quantum information*. American Association of Physics Teachers (2002)
- [62] Schumacher, B.: Sending entanglement through noisy quantum channels. *Phys. Rev. A* **54**, 2614–2628 (1996). <https://doi.org/10.1103/PhysRevA.54.2614>
- [63] Hastings, M.B.: Superadditivity of communication capacity using entangled inputs. *Nature Physics* **5**(4), 255–257 (2009)
- [64] Cubitt, T., Elkouss, D., Matthews, W., Ozols, M., Pérez-García, D., Strelchuk, S.: Unbounded number of channel uses may be required to detect quantum capacity. *Nature communications* **6**(1), 1–4 (2015)
- [65] Smith, G., Yard, J.: Quantum communication with zero-capacity channels. *Science* **321**(5897), 1812–1815 (2008)
- [66] Baumgratz, T., Cramer, M., Plenio, M.B.: Quantifying coherence. *Phys. Rev. Lett.* **113**, 140401 (2014). <https://doi.org/10.1103/PhysRevLett.113.140401>
- [67] Streltsov, A., Singh, U., Dhar, H.S., Bera, M.N., Adesso, G.: Measuring quantum coherence with entanglement. *Phys. Rev. Lett.* **115**, 020403 (2015). <https://doi.org/10.1103/PhysRevLett.115.020403>