Lecture 6: Quantum error correction and quantum capacity

Mark M. Wilde*

The quantum capacity theorem is one of the most important theorems in quantum Shannon theory. It is a fundamentally "quantum" theorem in that it demonstrates that a fundamentally quantum information quantity, the coherent information, is an achievable rate for quantum communication over a quantum channel. The fact that the coherent information does not have a strong analog in classical Shannon theory truly separates the quantum and classical theories of information.

The no-cloning theorem provides the intuition behind quantum error correction. The goal of any quantum communication protocol is for Alice to establish quantum correlations with the receiver Bob. We know well now that every quantum channel has an isometric extension, so that we can think of another receiver, the environment Eve, who is at a second output port of a larger unitary evolution. Were Eve able to learn anything about the quantum information that Alice is attempting to transmit to Bob, then Bob could not be retrieving this information—otherwise, they would violate the no-cloning theorem. Thus, Alice should figure out some subspace of the channel input where she can place her quantum information such that only Bob has access to it, while Eve does not. That the dimensionality of this subspace is exponential in the coherent information is perhaps then unsurprising in light of the above no-cloning reasoning. The coherent information is an entropy difference $H(B) - H(E)$—a measure of the amount of quantum correlations that Alice can establish with Bob less the amount that Eve can gain.

Perhaps the most surprising result in quantum Shannon theory is that it is possible to "superactivate" the quantum capacity. That is, suppose that two channels on their own have zero capacity for transmitting quantum information (for the phenomenon to occur, these channels are specific channels). Then it is possible for the joint channel (the tensor product of the individual channels) to have a non-zero quantum capacity, in spite of them being individually useless for quantum data transmission. This latter result implies that we are rather distant from having a complete quantum theory of information, in spite of the many successes reviewed in this book.

We structure this lecture as follows. We first overview the information processing task relevant for quantum communication. Next, we discuss the no-cloning intuition for quantum capacity in some more detail, presenting the specific example of a quantum erasure channel. We follow with a brief introduction to the theory of quantum error correction and demonstrate a proof of the direct part of the quantum capacity theorem for quantum stabilizer codes used for protecting quantum data sent over many independent uses of a Pauli channel.

*Mark M. Wilde is with the Department of Physics and Astronomy and the Center for Computation and Technology, Louisiana State University, Baton Rouge, Louisiana 70803. Much of the material is from the book preprint "From Classical to Quantum Shannon Theory" available as `arXiv:1106.1445`.

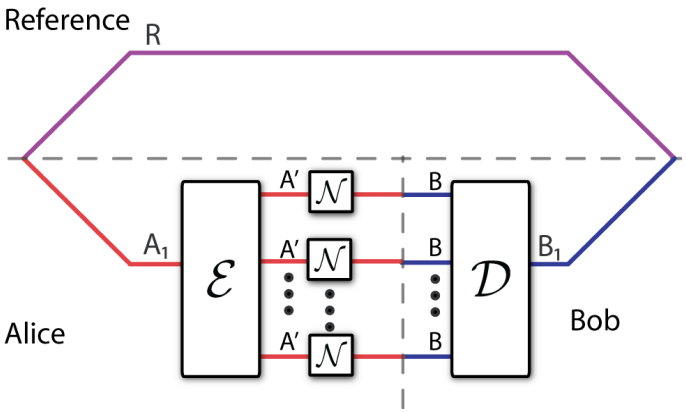

Figure 1: The information processing task for entanglement transmission. Alice is trying to preserve the entanglement with some inaccessible reference system by encoding her system and transmitting the encoded quantum data over many independent uses of a noisy quantum channel. Bob performs a decoding of the systems he receives, and the state at the end of the protocol is close to the original state shared between Alice and the reference if the protocol is any good for entanglement transmission.

1 The Information Processing Task

We begin the technical development in this lecture by describing the information processing task for quantum communication (we define an $(n, Q - \delta, \epsilon)$ quantum communication code). First, there are several protocols that we can consider for quantum communication, but perhaps the strongest definition of quantum capacity corresponds to a task known as *entanglement transmission*. Suppose that Alice shares entanglement with a reference system to which she does not have access. Then their goal is to devise a quantum coding scheme such that Alice can transfer this entanglement to Bob. To this end, suppose that Alice and the reference share an arbitrary state $|\varphi\rangle^{RA_1}$. Alice then performs some encoder on system A_1 to prepare it for input to many instances of a quantum channel $\mathcal{N}^{A'\to B}$. The resulting state is as follows:

$$\mathcal{E}^{A_1\to A'^n}(\varphi^{RA_1}). \tag{1}$$

Alice transmits the systems A'^n through many independent uses of the channel, resulting in the following state:

$$\mathcal{N}^{A'^n\to B^n}(\mathcal{E}^{A_1\to A'^n}(\varphi^{RA_1})), \tag{2}$$

where $\mathcal{N}^{A'^n\to B^n} \equiv (\mathcal{N}^{A'\to B})^{\otimes n}$. After Bob receives the systems B^n from the channel outputs, he performs some decoding map $\mathcal{D}^{B^n\to B_1}$, where B_1 is some system of the same dimension as A_1. The final state after Bob decodes is as follows:

$$\omega^{RB_1} \equiv \mathcal{D}^{B^n\to B_1}(\mathcal{N}^{A'^n\to B^n}(\mathcal{E}^{A_1\to A'^n}(\varphi^{RA_1}))). \tag{3}$$

Figure 1 depicts all of the above steps.

If the protocol is good for quantum communication, then the following condition should hold for all states $|\varphi\rangle^{RA_1}$:

$$\left\|\varphi^{RA_1} - \omega^{RB_1}\right\|_1 \leq \epsilon. \tag{4}$$

The rate Q of this scheme is equal to the number of qubits transmitted per channel use:

$$Q \equiv \frac{1}{n} \log d_{A_1} + \delta, \tag{5}$$

where d_{A_1} is the dimension of the A_1 register and δ is an arbitrarily small positive number. We say that a rate Q is achievable if there exists an $(n, Q - \delta, \epsilon)$ quantum communication code for all $\epsilon, \delta > 0$ and sufficiently large n.

The above notion of quantum communication encompasses other quantum information processing tasks such as mixed state transmission, pure state transmission, and entanglement generation. Alice can transmit any mixed or pure state if she can preserve the entanglement with a reference system. Also, she can generate entanglement with Bob if she can preserve entanglement with a reference system—she just needs to create an entangled state locally and apply the above protocol to one system of the entangled state.

2 The No-Cloning Theorem and Quantum Communication

We first discuss quantum communication over a quantum erasure channel before stating and proving the quantum capacity theorem. Consider the quantum erasure channel that gives Alice's input state to Bob with probability $1 - \epsilon$ and an erasure flag to Bob with probability ϵ:

$$\rho \to (1 - \epsilon)\rho + \epsilon |e\rangle\langle e|, \tag{6}$$

where $\langle e|\rho|e\rangle = 0$ for all inputs ρ. An isometric extension of this channel is as follows:

$$|\psi\rangle^{RA} \to \sqrt{1 - \epsilon}|\psi\rangle^{RB}|e\rangle^{E} + \sqrt{\epsilon}|\psi\rangle^{RE}|e\rangle^{B}, \tag{7}$$

so that the channel now has the other interpretation that Eve gets the state with probability ϵ while giving her the erasure flag with probability $1 - \epsilon$.

Now suppose that the erasure parameter is set to $1/2$. In such a scenario, the channel to Eve is the *same* as the channel to Bob, namely, both have the channel $\rho \to 1/2(\rho + |e\rangle\langle e|)$. We can argue that the quantum capacity of such a channel should be zero, by invoking the no-cloning theorem. More specifically, suppose there is a scheme (an encoder and decoder as given in Figure 1) for Alice and Bob to communicate quantum information reliably at a non-zero rate over such a channel. If so, Eve could simply use the same decoder that Bob does, and she should also be able to obtain the quantum information that Alice is sending. But the ability for both Bob and Eve to decode the quantum information that Alice is transmitting violates the no-cloning theorem. Thus, the quantum capacity of such a channel should vanish.

Exercise 1 *Prove that the quantum capacity of an amplitude damping channel vanishes if its damping parameter is equal to* $1/2$.

The no-cloning theorem plays a more general role in the analysis of quantum communication over quantum channels. In the construction of a quantum code, we are trying to find a "no-cloning" subspace of the input Hilbert space that is protected from Eve. If Eve is able to obtain any of the quantum information in this subspace, then this information cannot be going to Bob by the same no-cloning argument featured in the previous paragraph.

3 Stabilizer Codes and the Hashing Bound

We now describe a well-known class of quantum error-correcting codes known as the stabilizer codes, and we prove that a randomly chosen stabilizer code achieves a quantum communication rate known as the hashing bound of a Pauli channel (the hashing bound is equal to the coherent information of a Pauli channel when sending one share of a Bell state through it). The proof of this theorem is different from our proof above that the coherent information rate is achievable, and we consider it instructive to see this other approach for the special case of stabilizer codes used for protecting quantum information sent over many independent instances of a Pauli channel. Before delving into the proof, we first briefly introduce the simple repetition code and the more general stabilizer quantum codes.

3.1 The Qubit Repetition Code

The simplest quantum error correction code is the repetition code, which encodes one qubit $|\psi\rangle \equiv \alpha|0\rangle + \beta|1\rangle$ into three physical qubits as follows:

$$\alpha|0\rangle + \beta|1\rangle \rightarrow \alpha|000\rangle + \beta|111\rangle. \tag{8}$$

A simple way to perform this encoding is to attach two ancilla qubits in the state $|0\rangle$ to the original qubit and perform a CNOT gate from the first qubit to the second and from the first to the last. This encoding illustrates one of the fundamental principles of quantum error correction: the quantum information is spread across the correlations between the three physical qubits after the encoding takes place. (Of course, this was also the case for the codes we constructed in the direct part of the quantum capacity theorem.)

The above encoding will protect the encoded qubit against an artificial noise where either the first, second, or third qubit is subjected to a bit flip (and no other errors occur). For example, if a bit flip occurs on the second qubit, the encoded state changes as follows:

$$X_2(\alpha|000\rangle + \beta|111\rangle) = \alpha|010\rangle + \beta|101\rangle, \tag{9}$$

where the notation X_2 indicates that a Pauli operator X acts on the second qubit. The procedure for the receiver to recover from such an error is to perform collective measurements on all three qubits that learn only about the error and nothing about the encoded quantum data. In this case, the receiver can perform a measurement of the operators Z_1Z_2 and Z_2Z_3 to learn only about the error, so that the coherent superposition is preserved. One can easily verify that Z_1Z_2 and Z_2Z_3 are as follows:

$$Z_1Z_2 \equiv Z \otimes Z \otimes I = [(|00\rangle\langle 00| + |11\rangle\langle 11|) - (|01\rangle\langle 01| + |10\rangle\langle 10|)] \otimes I, \tag{10}$$

$$Z_2Z_3 \equiv I \otimes Z \otimes Z = I \otimes [(|00\rangle\langle 00| + |11\rangle\langle 11|) - (|01\rangle\langle 01| + |10\rangle\langle 10|)], \tag{11}$$

revealing that these measurements return a $+1$ if the parity of the basis states is even and -1 if the parity is odd. So, for our example error in (9), the syndrome measurements will return -1 for Z_1Z_2 and -1 for Z_2Z_3, which the receiver can use to identify the error that occurs. He can then perform the bit flip operator X_2 to invert the action of the error. One can verify that the following

syndrome table identifies which type of error occurs:

Measurement Result	Error
$+1, +1$	I
$+1, -1$	X_3
$-1, +1$	X_1
$-1, -1$	X_2

. (12)

Thus, if the only errors that occur are either no error or a single-qubit bit-flip error, then it is possible to perfectly correct these. If errors besides these ones occur, then it is not possible to correct them with this code.

3.2 Stabilizer Codes

We can generalize the main idea behind the above qubit repetition code to formulate the class of quantum stabilizer codes. These stabilizer codes then generalize the classical theory of linear error correction to the quantum case.

In the repetition code, observe that the encoded state in (8) is a $+1$-eigenstate of the operators Z_1Z_2 and Z_2Z_3, i.e., it holds that

$$Z_1Z_2(\alpha|000\rangle + \beta|111\rangle) = \alpha|000\rangle + \beta|111\rangle = Z_2Z_3(\alpha|000\rangle + \beta|111\rangle). \tag{13}$$

We say that the operators Z_1Z_2 and Z_2Z_3 stabilize the encoded state. The stabilizing operators form a group under multiplication because we obtain another stabilizing operator if we multiply two of them: one can check that the operator Z_1Z_3 stabilizes the encoded state and that $Z_1Z_3 = (Z_1Z_2)(Z_2Z_3)$. Also, the two operators Z_1Z_2 and Z_2Z_3 commute, implying that the encoded state is in the simultaneous eigenspace of these operators, and that it is possible to measure the operators Z_1Z_2 and Z_2Z_3 in any order, in order to learn about errors that occur.

We now describe the theory of quantum stabilizer codes. Recall that the Pauli matrices for one qubit are I, X, Y, and Z, whose action on the computational basis is as follows:

$$I|0\rangle = |0\rangle, \qquad I|1\rangle = |1\rangle, \tag{14}$$
$$X|0\rangle = |1\rangle, \qquad X|1\rangle = |0\rangle, \tag{15}$$
$$Y|0\rangle = i|1\rangle, \qquad Y|1\rangle = -i|0\rangle, \tag{16}$$
$$Z|0\rangle = |0\rangle, \qquad Z|1\rangle = -|1\rangle. \tag{17}$$

The X operator is known as the "bit-flip" operator, Z as the "phase-flip" operator, and Y as the "bit and phase flip" operator. The Pauli group $\mathcal{G}_n$ acting on n qubits consists of n-fold tensor products of these operators along with the phase factors ± 1 and $\pm i$:

$$\mathcal{G}_n \equiv \{\pm 1, \pm i\} \otimes \{I, X, Y, Z\}^{\otimes n}. \tag{18}$$

The inclusion of the phase factors, along with the relations $Y = iXZ$, $Z = iYX$, and $X = iZY$ and the fact that any one of X, Y, and Z anticommutes with the other two ensures that the set $\mathcal{G}_n$ is closed under multiplication. It is useful in the theory of quantum error correction to consider the Pauli group quotiented out by its center: $\mathcal{G}_n/\{\pm 1, \pm i\}$, essentially because global phases are not physically observable. This reduced version of the Pauli group has 4^n elements.

Let $\mathcal{S}$ be an abelian subgroup of the Pauli group $\mathcal{G}_n$. Any such subgroup $\mathcal{S}$ has size 2^{n-k} for some integer k such that $0 \leq k \leq n$. This subgroup $\mathcal{S}$ can be generated by a set of size $n-k$, so that $\mathcal{S} = \langle S_1, \ldots, S_{n-k} \rangle$. A state $|\psi\rangle$ is stabilized by the subgroup $\mathcal{S}$ if

$$S|\psi\rangle = |\psi\rangle \qquad \forall S \in \mathcal{S}. \tag{19}$$

The 2^k-dimensional subspace of the full 2^n-dimensional space for the n qubits that is stabilized by $\mathcal{S}$ is known as the codespace, or equivalently, an $[n, k]$ stabilizer code that encodes k logical qubits into n physical qubits. The decoding operation that the receiver performs is analogous to that for the repetition code—he just measures the $n-k$ operators constituting some generating set of $\mathcal{S}$ and performs a recovery operation based on the results of these measurements.

We can define logical operations on the quantum information encoded inside an $[n, k]$ stabilizer code. These are operations that manipulate the quantum information inside of the codespace without taking the encoded information outside the codespace. These logical operations are part of the normalizer of $\mathcal{S}$, defined as

$$N(\mathcal{S}) \equiv \left\{ U \in \mathbb{U}(2^n) : U\mathcal{S}U^\dagger = \mathcal{S} \right\}, \tag{20}$$

where $\mathbb{U}(2^n)$ denotes the unitary group for n qubits. We can easily see that any $U \in N(\mathcal{S})$ does not take a state $|\psi\rangle$ in the codespace outside of it. First, for all $U \in N(\mathcal{S})$, it follows that $U^\dagger \in N(\mathcal{S})$, so that for all $S \in \mathcal{S}$, we have

$$SU|\psi\rangle = UU^\dagger SU|\psi\rangle = US_U|\psi\rangle = U|\psi\rangle, \tag{21}$$

where $S_U = U^\dagger SU$ and $S_U \in \mathcal{S}$ from the definition of the normalizer. From the above, we conclude that the state $U|\psi\rangle$ is in the codespace since it is stabilized by all $S \in \mathcal{S}$: $SU|\psi\rangle = U|\psi\rangle$. It also follows that $\mathcal{S} \subseteq N(\mathcal{S})$ because $\mathcal{S}$ is abelian, implying that

$$S_1 S_2 S_1^\dagger = S_2 S_1 S_1^\dagger = S_2 \quad \forall S_1, S_2 \in \mathcal{S}. \tag{22}$$

In quantum error correction, we are concerned with correcting a fixed set of errors $\mathcal{E} \subseteq \mathcal{G}_n$ such that each element of $\mathcal{E}$ acts on the n physical qubits. In doing so, we might not be able to correct all of the errors in a set $\mathcal{E}$ if there exists a pair $E_1, E_2 \in \mathcal{E}$ such that

$$E_1^\dagger E_2 \in N(\mathcal{S}). \tag{23}$$

Consider that for all $S \in \mathcal{S}$, we have

$$E_1^\dagger E_2 S = (-1)^{g(S,E_1)+g(S,E_2)} S E_1^\dagger E_2, \tag{24}$$

where we define $g(P, Q)$ by $PQ = (-1)^{g(P,Q)}QP$ for all $P, Q \in \mathcal{G}_n$. The above relation then implies the following one for all $S \in \mathcal{S}$:

$$E_1^\dagger E_2 S \left(E_1^\dagger E_2 \right)^\dagger = (-1)^{g(S,E_1)+g(S,E_2)} S. \tag{25}$$

Since we assumed that $E_1^\dagger E_2 \in N(\mathcal{S})$, the only way that the above relation can be true for all $S \in \mathcal{S}$ is if $g(S, E_1) = g(S, E_2)$. Thus, during the error correction procedure, Bob will measure a set $\{S_j\}$ of generators, and since the outcome of a measurement of S_j on $E|\psi\rangle$ is $g(S, E)$, the errors E_1 and E_2 will be assigned the same syndrome. Since they have the same syndrome, the receiver will have to reverse these errors with the same recovery operation, and this is only possible if $E_1|\psi\rangle = E_2|\psi\rangle$ for all states $|\psi\rangle$ in the codespace. This latter condition is only true if $E_1^\dagger E_2 \in \mathcal{S}$, leading us to the error correcting conditions for quantum stabilizer codes:

Theorem 2 *It is possible to correct a set of errors $\mathcal{E}$ with a quantum stabilizer code if every pair $E_1, E_2 \in \mathcal{E}$ satisfies*

$$E_1^\dagger E_2 \notin N(\mathcal{S})/\mathcal{S}. \tag{26}$$

A simple way to satisfy the error-correcting conditions is just to demand that every pair of errors in $\mathcal{E}$ be such that $E_1^\dagger E_2 \notin N(\mathcal{S})$. In such a case, each error is assigned a unique syndrome, and codes along with an error set satisfying this property are known as non-degenerate codes. Codes with a corresponding error set not satisfying this are known as degenerate codes.

3.3 The Hashing Bound

We now provide a proof that the hashing bound for a Pauli channel (coherent information when sending one share of a Bell state through a Pauli channel) is an achievable rate for quantum communication. Of course, our proof of the direct part of the quantum capacity theorem already suffices as a proof of this statement, but we think it is instructive to provide a proof of this statement using the theory of stabilizer codes. The main idea of the proof is to choose a stabilizer code randomly from the set of all stabilizer codes and show that such a code can correct the typical errors issued by a tensor-product Pauli channel.

Theorem 3 (Hashing Bound) *There exists a stabilizer quantum error-correcting code that achieves the hashing bound $R = 1 - H(\mathbf{p})$ for a Pauli channel of the following form:*

$$\rho \to p_I \rho + p_X X\rho X + p_Y Y\rho Y + p_Z Z\rho Z, \tag{27}$$

where $\mathbf{p} = (p_I, p_X, p_Y, p_Z)$ and $H(\mathbf{p})$ is the entropy of this probability vector.

Proof. We consider a decoder that corrects only the typical errors. That is, consider defining the typical error set as follows:

$$T_\delta^{\mathbf{p}^n} \equiv \left\{ a^n : \left| -\frac{1}{n} \log_2(\Pr\{E_{a^n}\}) - H(\mathbf{p}) \right| \leq \delta \right\}, \tag{28}$$

where a^n is some sequence consisting of letters corresponding to the Pauli operators $\{I, X, Y, Z\}$ and $\Pr\{E_{a^n}\}$ is the probability that an IID Pauli channel issues some tensor-product error $E_{a^n} \equiv E_{a_1} \otimes \cdots \otimes E_{a_n}$. This typical set consists of the likely errors in the sense that

$$\sum_{a^n \in T_\delta^{\mathbf{p}^n}} \Pr\{E_{a^n}\} \geq 1 - \epsilon, \tag{29}$$

for all $\epsilon > 0$ and sufficiently large n. The error-correcting conditions for a stabilizer code in this case are that $\{E_{a^n} : a^n \in T_\delta^{\mathbf{p}^n}\}$ is a correctable set of errors if

$$E_{a^n}^\dagger E_{b^n} \notin N(\mathcal{S})/\mathcal{S}, \tag{30}$$

for all error pairs E_{a^n} and E_{b^n} such that $a^n, b^n \in T_\delta^{\mathbf{p}^n}$. Also, we consider the expectation of the error probability under a random choice of a stabilizer code. We proceed as follows:

$$\mathbb{E}_{\mathcal{S}}\{p_e\} = \mathbb{E}_{\mathcal{S}}\left\{\sum_{a^n} \Pr\{E_{a^n}\}\mathcal{I}(E_{a^n} \text{ is uncorrectable under } \mathcal{S})\right\} \tag{31}$$

$$\leq \mathbb{E}_{\mathcal{S}}\left\{\sum_{a^n \in T_\delta^{\mathbf{p}^n}} \Pr\{E_{a^n}\}\mathcal{I}(E_{a^n} \text{ is uncorrectable under } \mathcal{S})\right\} + \epsilon \tag{32}$$

$$= \sum_{a^n \in T_\delta^{\mathbf{p}^n}} \Pr\{E_{a^n}\}\mathbb{E}_{\mathcal{S}}\{\mathcal{I}(E_{a^n} \text{ is uncorrectable under } \mathcal{S})\} + \epsilon \tag{33}$$

$$= \sum_{a^n \in T_\delta^{\mathbf{p}^n}} \Pr\{E_{a^n}\} \Pr_{\mathcal{S}}\{E_{a^n} \text{ is uncorrectable under } \mathcal{S}\} + \epsilon \tag{34}$$

The first equality follows by definition—$\mathcal{I}$ is an indicator function equal to one if E_{a^n} is uncorrectable under $\mathcal{S}$ and equal to zero otherwise. The first inequality follows from (29)—we correct only the typical errors because the atypical error set has negligible probability mass. The second equality follows by exchanging the expectation and the sum. The third equality follows because the expectation of an indicator function is the probability that the event it selects occurs. Continuing, we have

$$= \sum_{a^n \in T_\delta^{\mathbf{p}^n}} \Pr\{E_{a^n}\} \Pr_{\mathcal{S}}\left\{\exists E_{b^n} : b^n \in T_\delta^{\mathbf{p}^n},\ b^n \neq a^n,\ E_{a^n}^\dagger E_{b^n} \in N(\mathcal{S})/\mathcal{S}\right\} \tag{35}$$

$$\leq \sum_{a^n \in T_\delta^{A^n}} \Pr\{E_{a^n}\} \Pr_{\mathcal{S}}\left\{\exists E_{b^n} : b^n \in T_\delta^{\mathbf{p}^n},\ b^n \neq a^n,\ E_{a^n}^\dagger E_{b^n} \in N(\mathcal{S})\right\} \tag{36}$$

$$= \sum_{a^n \in T_\delta^{\mathbf{p}^n}} \Pr\{E_{a^n}\} \Pr_{\mathcal{S}}\left\{\bigcup_{b^n \in T_\delta^{\mathbf{p}^n},\ b^n \neq a^n} E_{a^n}^\dagger E_{b^n} \in N(\mathcal{S})\right\} \tag{37}$$

$$\leq \sum_{a^n, b^n \in T_\delta^{\mathbf{p}^n},\ b^n \neq a^n} \Pr\{E_{a^n}\} \Pr_{\mathcal{S}}\left\{E_{a^n}^\dagger E_{b^n} \in N(\mathcal{S})\right\} \tag{38}$$

$$\leq \sum_{a^n, b^n \in T_\delta^{\mathbf{p}^n},\ b^n \neq a^n} \Pr\{E_{a^n}\} 2^{-(n-k)} \tag{39}$$

$$\leq 2^{2n[H(\mathbf{p})+\delta]} 2^{-n[H(\mathbf{p})+\delta]} 2^{-(n-k)} \tag{40}$$

$$= 2^{-n[1-H(\mathbf{p})-k/n-3\delta]}. \tag{41}$$

The first equality follows from the error-correcting conditions for a quantum stabilizer code, where $N(\mathcal{S})$ is the normalizer of $\mathcal{S}$. The first inequality follows by ignoring any potential degeneracy in the code—we consider an error uncorrectable if it lies in the normalizer $N(\mathcal{S})$ and the probability can only be larger because $N(\mathcal{S})/\mathcal{S} \subseteq N(\mathcal{S})$. The second equality follows by realizing that the probabilities for the existence criterion and the union of events are equivalent. The second inequality follows by applying the union bound. The third inequality follows from the fact that the probability for a fixed operator $E_{a^n}^\dagger E_{b^n}$ not equal to the identity commuting with the stabilizer operators of a

random stabilizer can be upper bounded as follows:

$$\Pr_{\mathcal{S}}\Big\{ E_{a^n}^{\dagger} E_{b^n} \in N(\mathcal{S}) \Big\} = \frac{2^{n+k}-1}{2^{2n}-1} \leq 2^{-(n-k)}. \tag{42}$$

The random choice of a stabilizer code is equivalent to fixing operators $Z_1, \ldots, Z_{n-k}$ and performing a uniformly random Clifford unitary U. The probability that a fixed operator commutes with $UZ_1U^{\dagger}, \ldots, UZ_{n-k}U^{\dagger}$ is then just the number of non-identity operators in the normalizer $(2^{n+k}-1)$ divided by the total number of non-identity operators $(2^{2n}-1)$. After applying the above bound, we then exploit the following typicality bounds:

$$\forall a^n \in T_{\delta}^{\mathbf{p}^n} : \Pr\{E_{a^n}\} \leq 2^{-n[H(\mathbf{p})+\delta]}, \tag{43}$$

$$\left| T_{\delta}^{\mathbf{p}^n} \right| \leq 2^{n[H(\mathbf{p})+\delta]}. \tag{44}$$

We conclude that as long as the rate $k/n = 1 - H(\mathbf{p}) - 4\delta$, the expectation of the error probability becomes arbitrarily small, so that there exists at least one choice of a stabilizer code with the same bound on the error probability. ■