Abelian Group Codes for Classical and CQ Channel Coding: One-shot and Asymptotic Rate Bounds

James (Chin-Jen) Pang
Marvell Technology, Inc.
Santa Clara, CA, USA
Email: cjpang@umich.edu

S. Sandeep Pradhan
Department of EECS
University of Michigan, USA
Email: pradhanv@umich.edu

Hessam Mahdavifar
Department of ECE
Northeastern University, Boston, USA
Email: h.mahdavifar@northeastern.edu

***Abstract*—We study the *one-shot* channel coding problem over classical and classical-quantum channels, where the underlying codes are constrained to be *group codes*. In the achievability part, we introduce a new distribution that incorporates the encoding homomorphism and the underlying channel law. Using a random coding argument, we characterize the performance in terms of hypothesis testing relative-entropies. In the converse part, we establish bounds by leveraging a hypothesis testing-based approach. Further we apply the one-shot result to the asymptotic use case and establish the *group capacities* for both channels.**

I. Introduction

In this paper, we study channel coding both in the classical and *classical-quantum* (CQ) settings. In both problems, the data to be transmitted reliably are classical, but the channel output of the former is classical in nature, whereas in the latter it is a quantum state characterized by a density operator. The channel coding theorems for classical-quantum channels have been established in [1]–[4].

Due to its algebraic properties, the finite field structure has been adopted in the coding schemes, over the past several decades, to approach information-theoretic performance limits of point-to-point communication [5]–[9]. Later these coding approaches were extended to weaker algebraic structures such as rings and groups [10]–[16]. This is motivated by the two following reasons: a) finite fields exist only for alphabets with a prime power size, and b) for communication under certain constraints, codes with weaker algebraic structures have better properties. For example, when communicating over an additive white Gaussian noise channel with 8-PSK constellation, codes over $\mathbb{Z}_8$, the cyclic group of size 8, are more desirable over binary linear codes because the structure of the code is matched to the structure of the signal set [12]. As another example, construction of polar codes over alphabets of size p^r, for $r > 1$ and p prime, is simpler with a module structure rather than a vector space structure [17]–[19]. Furthermore, Abelian group codes yield better performance in network communication settings such as distributed source coding and interference channels [20]–[22]. Many of the aforementioned works addressed the channel coding problem in the asymptotic regime.

Later researchers considered single-serving scenarios where a given channel is used only once. This approach gives rise to a high level of generality that no assumptions are made on the structure of the channel and the associated capacity is usually referred to as *one-shot* capacity. The one-shot capacity of a classical channel was characterized in terms of min- and max-entropies in [23]. The one-shot classical capacity of a quantum channel is addressed by a hypothesis testing approach in [24] and [25], yielding expressions in terms of the generalized (Rényi) relative entropies and a smooth relative entropy quantity, respectively. These works considered random unstructured codes–those which do not have any group structure–in their achievability approaches.

In this work, we consider performance of structured codes, in particular Abelian group codes, for transmission of classical information over classical and classical-quantum channels in the one-shot regime. This problem has not been studied before. Our derivation is based on the idea of relating the problem of channel coding to hypothesis testing. Here, we use a relative-entropy-type quantity defined in [25] known as hypothesis testing relative entropy, denoted $D_{\mathrm{H}}^{\epsilon}(\cdot\|\cdot)$. We introduce a new hypothesis testing group-based relative entropy that incorporates the underlying subgroup structure of the channel input group alphabet, and derive a tight characterization of the performance of group codes. We use the framework of one-shot quantum typicality developed in [26] for the achievability of CQ channels.

II. Preliminaries

A. Classical and CQ Channel Model

We consider discrete memoryless classical channels used without feedback specified by the tuple $(\mathcal{X}, \mathcal{Y}, W_{Y|X})$, where $\mathcal{X}$ and $\mathcal{Y}$ are the channel input and output alphabets. We also study the case of *classical-quantum channel coding*, where the data to be transmitted reliably are *classical*. Let a (classical) set $\mathcal{X}$ denote the input alphabet. For any input $x \in \mathcal{X}$, the channel produces an output, specified by a density operator ρ_x on a Hilbert space $\mathcal{B}$. We denote a CQ channel by a mapping $\mathcal{N} : x \mapsto \rho_x$ from $\mathcal{X}$ to a set of density operators.

B. Definition of Achievability for Classical Channel Coding

For a group G, a group transmission system with parameters (n, Ω, τ) for reliable communication over a given channel $(\mathcal{X} = G, \mathcal{Y}, W_{Y|X})$ consists of a codebook, an encoding mapping and a decoding mapping. The codebook $\mathbb{C}$ is a shifted subgroup of G^n whose size is equal to Ω and the mappings

This work was supported by the National Science Foundation under grants CCF–2132815 and CCF–1909771.

are defined as

$$\text{Enc} : \{1, 2, \cdots, \Omega\} \to \mathbb{C}\,, \quad \text{Dec} : \mathcal{Y}^n \to \{1, 2, \cdots, \Omega\}\,,$$

such that

$$\sum_{m=1}^{\Omega} \frac{1}{\Omega} \sum_{x \in \mathcal{X}^n} \mathbb{1}_{\{x=\text{Enc}(m)\}} \sum_{y \in \mathcal{Y}^n} \mathbb{1}_{\{m \neq \text{Dec}(y)\}} W^n(y|x) \leqslant \tau\,.$$

A rate R is said to be achievable using group codes if for all $\epsilon > 0$ and for all sufficiently large n, there exists a group transmission system with parameters (n, Ω, τ) such that

$$\frac{1}{n} \log \Omega \geqslant R - \epsilon, \qquad \tau \leqslant \epsilon\,.$$

The group capacity C of the channel is defined as the supremum of the set of all achievable rates using group codes.

C. Definition of Achievability for CQ Channel

Given a classical-quantum channel $\mathcal{N} = \{\rho_x\}_{x \in \mathcal{X}}$ from the classical alphabet $\mathcal{X}$ to the quantum system $\mathcal{B}$, where $\mathcal{X} = G$ is an Abelian group, a group transmission system with parameters (n, Ω, τ) over $\mathcal{N}$ consists of a codebook, an encoding mapping and a decoding positive operator-valued measure (POVM). The codebook $\mathbb{C}$ is a shifted subgroup of G^n whose size is Ω. The encoding mapping is defined as $\text{Enc} : \{1, 2, \cdots, \Omega\} \to \mathbb{C}$. The decoding POVM is a set $\{\Lambda_m\}_{m=1}^{\Omega}$ of operators such that $\Lambda_m \geqslant 0, \forall m$ and $\sum_m \Lambda_m = I$. The probability of obtaining outcome j is $\text{tr}(\Lambda_j \rho)$ if the state is described by some density operator ρ. The group transmission system with parameters (n, Ω, τ) over $\mathcal{N}$ exists if

$$\sum_{m=1}^{\Omega} \frac{1}{\Omega} \sum_{x \in \mathcal{X}^n} \mathbb{1}_{\{x=\text{Enc}(m)\}} [1 - \text{tr}(\Lambda_m \rho_x)] \leqslant \tau\,.$$

Given a channel $\mathcal{N}$, the rate R is said to be achievable using group codes if for all $\epsilon > 0$ and for all sufficiently large n, there exists a group transmission system for reliable communication with parameters (n, Ω, τ) such that

$$\frac{1}{n} \log \Omega \geqslant R - \epsilon, \qquad \tau \leqslant \epsilon\,.$$

The group capacity of the channel $C = C(\mathcal{N})$ is defined as the supremum of all achievable rates using group codes.

D. Groups and Group Codes

All groups referred to in this paper are *Abelian groups*. Given a group $(G, +)$ and a subset H of G, we use $H \leqslant G$ to denote that H is subgroup of G. A *coset* C of a subgroup H is a shift of H by an arbitrary element $a \in G$ (i.e. $C = a + H$ for some $a \in G$). A coset is also referred to as a shifted subgroup. For a prime p dividing the cardinality of G, the *Sylow-p* subgroup of G is the largest subgroup of G whose cardinality is a power of p. Group isomorphism is denoted by $\cong$. Given a group G, a group code $\mathbb{C}$ over G with block length n is coset of a subgroup of G^n.

III. Abelian Group Code Ensemble

In this section, we use a standard characterization of Abelian groups and introduce the ensemble of Abelian group codes used in [15] and this paper.

A. Abelian Groups

For an Abelian group G, let $\mathcal{P}(G)$ denote the set of all distinct primes which divide $|G|$ and for a prime $p \in \mathcal{P}(G)$ let $S_p(G)$ be the corresponding Sylow subgroup of G. It is known that any Abelian group G can be decomposed as a direct sum of its Sylow subgroups in the following manner $G = \bigoplus_{p \in \mathcal{P}(G)} S_p(G)$ Furthermore, each Sylow subgroup $S_p(G)$ can be decomposed into $\mathbb{Z}_{p^r}$ groups as follows: $S_p(G) \cong \bigoplus_{r \in \mathcal{R}_p(G)} \mathbb{Z}_{p^r}^{M_{p,r}}$,where $\mathcal{R}_p(G) \subseteq \mathbb{Z}^+$ and for $r \in \mathcal{R}_p(G)$, $M_{p,r}$ is a positive integer. Thus,

$$G \cong \bigoplus_{p \in \mathcal{P}(G)} \bigoplus_{r \in \mathcal{R}_p(G)} \mathbb{Z}_{p^r}^{M_{p,r}} = \bigoplus_{p \in \mathcal{P}(G)} \bigoplus_{r \in \mathcal{R}_p(G)} \bigoplus_{m=1}^{M_{p,r}} \mathbb{Z}_{p^r}^{(m)} \quad (1)$$

where $\mathbb{Z}_{p^r}^{(m)}$ is called the m^{th} $\mathbb{Z}_{p^r}$ ring of G or the (p, r, m)-th ring of G. We also define two sets, $\mathcal{Q}(G) \subseteq \mathbb{P} \times \mathbb{Z}^+$ by

$$\mathcal{Q}(G) = \{(p, r) \mid p \in \mathcal{P}(G), r \in \mathcal{R}_p(G)\}\ ,$$

and $\mathcal{G}(G) \subseteq \mathbb{P} \times \mathbb{Z}^+ \times \mathbb{Z}^+$ by

$$\mathcal{G}(G) = \{(p, r, m) \mid (p, r) \in \mathcal{Q}(G), m \in \{1, 2, \ldots, M_{p,r}\}\}\ .$$

Hence any element a of the Abelian group G can be regarded as a vector whose components are indexed by $(p, r, m) \in \mathcal{G}(G)$ and whose (p, r, m)-th component $a_{p,r,m}$ takes values from the ring $\mathbb{Z}_{p^r}$.

Example 1: Let $G = \mathbb{Z}_4 \oplus \mathbb{Z}_3 \oplus \mathbb{Z}_9^2$. Then we have $\mathcal{P}(G) = \{2, 3\}$, $S_2(G) = \mathbb{Z}_4$ and $S_3(G) = \mathbb{Z}_3 \oplus \mathbb{Z}_9^2$, $\mathcal{R}_2(G) = \{2\}$, $\mathcal{R}_3(G) = \{1, 2\}$, $M_{2,2} = 1$, $M_{3,1} = 1$, $M_{3,2} = 2$ and $\mathcal{G}(G) = \{(2, 2, 1), (3, 1, 1), (3, 2, 1), (3, 2, 2)\}$. Each element a of G can be represented by a quadruple $(a_{2,2,1}, a_{3,1,1}, a_{3,2,1}, a_{3,2,2})$ where $a_{2,2,1} \in \mathbb{Z}_4$, $a_{3,1,1} \in \mathbb{Z}_3$ and $a_{3,2,1}, a_{3,2,2} \in \mathbb{Z}_9$.

In the following section, we introduce the ensemble of Abelian group codes which we use in the paper.

B. The Image Ensemble

Recall that for a positive integer n, an Abelian group code of length n over the group G is a coset of a subgroup of G^n. Our ensemble of codes consists of *all* Abelian group codes over G; i.e., we consider all cosets of subgroups of G^n. The following lemma ([15, Lemma 1]) effectively characterizes all subgroups of G^n:

Lemma 1: For a group $\tilde{G}$, let $\phi : J \to \tilde{G}$ be a homomorphism from some group J to $\tilde{G}$. Then $\phi(J) \leqslant \tilde{G}$. Moreover, for any subgroup $\tilde{H}$ of $\tilde{G}$ there exists a corresponding group J and a homomorphism $\phi : J \to \tilde{G}$ such that $\tilde{H} = \phi(J)$.

Definition 1: Let G be an Abelian group. For $p \in \mathcal{P}(G)$, define $r_p = \max \mathcal{R}_p(G)$, and $\mathcal{S}(G) = \{(p, s) \mid p \in \mathcal{P}(G), 1 \leqslant s \leqslant r_p\}$.

It is shown in [15] that we only need to consider homomorphisms from an Abelian group J to $\tilde{G}$ such that $\mathcal{P}(J) \subseteq \mathcal{P}(\tilde{G})$, and $s \leqslant r_q = \max \mathcal{R}_q(\tilde{G})$ for all $(q, s, l) \in \mathcal{G}(J)$.

To construct Abelian group codes of length n over G, let $\tilde{G} = G^n$. We have

$$G^n \cong \bigoplus_{p \in \mathcal{P}(G)} \bigoplus_{r \in \mathcal{R}_p} \mathbb{Z}_{p^r}^{nM_{p,r}} = \bigoplus_{p \in \mathcal{P}(G)} \bigoplus_{r \in \mathcal{R}_p} \bigoplus_{m=1}^{nM_{p,r}} \mathbb{Z}_{p^r}^{(m)}\,. \quad (2)$$

Define J as

$$J = \bigoplus_{q\in\mathcal{P}(G)} \bigoplus_{s=1}^{r_q} \mathbb{Z}_{q^s}^{k_{q,s}} = \bigoplus_{q\in\mathcal{P}(G)} \bigoplus_{s=1}^{r_q} \bigoplus_{l=1}^{k_{q,s}} \mathbb{Z}_{q^s}^{(l)} \tag{3}$$

for some positive integers $k_{q,s}$. Define $k = \sum_{q\in\mathcal{P}(G)} \sum_{s=1}^{r_q} k_{q,s}$ and $w_{q,s} = \frac{k_{q,s}}{k}$ for $(q,s) \in \mathcal{S}(G)$.

Definition 2: The ensemble of Abelian group encoders consists of all mappings $\phi : J \to G^n$ of the form

$$\phi(a) = \bigoplus_{(p,r,m)\in\mathcal{G}(G^n)} \sum\nolimits_{(q,s,l)\in\mathcal{G}(J)} a_{q,s,l} g_{(q,s,l)\to(p,r,m)} \tag{4}$$

for $a \in J$, where $a_{q,s,l} g_{(q,s,l)\to(p,r,m)}$ is the short-hand notation for the mod-p^r addition of $g_{(q,s,l)\to(p,r,m)}$ to itself for $a_{q,s,l}$ times, the sum is over $\mathbb{Z}_{p^r}$, and

$$g_{(q,s,l)\to(p,r,m)} \begin{cases} = 0 & \text{if } p \neq q \\ \sim \text{Unif}(\mathbb{Z}_{p^r}) & \text{if } p = q, r \leqslant s \\ \sim \text{Unif}(p^{r-s}\mathbb{Z}_{p^r}) & \text{if } p = q, r \geqslant s \end{cases}$$

The corresponding group code is defined by

$$\mathbb{C} = \{\phi(a) + V | a \in J\}, \tag{5}$$

where V is a uniform random variable over G^n.
The rate of this code is given by

$$R = \frac{1}{n} \log |J| = \frac{k}{n} \sum_{q\in\mathcal{P}(G)} \sum_{s=1}^{r_q} s w_{q,s} \log q. \tag{6}$$

C. The $H_{\hat{\theta}}$ coset

For an Abelian group G defined in (1), denote a vector $\hat{\theta}$ whose components are non-negative integer-valued and indexed by $(p,s) \in \mathcal{S}(G)$ by $(\hat{\theta}_{p,s})_{(p,s)\in\mathcal{S}(G)}$, where $0 \leqslant \hat{\theta}_{p,s} \leqslant s$. Let $\mathbf{s}$ denote the vector whose components satisfy $\mathbf{s}_{(p,s)} = s$ for all $(p,s) \in \mathcal{S}(G)$. Let $\Theta = \Theta(G)$ be the set of vectors $\hat{\theta}$ indexed by $(p,s) \in \mathcal{S}(G)$ such that $0 \leqslant \hat{\theta}_{p,s} \leqslant s$ and $\hat{\theta} \neq \mathbf{s}$, and denote its size by $M \triangleq |\Theta|$. For $\hat{\theta} = (\hat{\theta}_{p,s})_{(p,s)\in\mathcal{S}(G)}$, define a vector $\boldsymbol{\theta}(\hat{\theta})$ indexed by $(p,r) \in \mathcal{Q}(G)$ and

$$\left(\boldsymbol{\theta}(\hat{\theta})\right)_{p,r} = \min_{\substack{(p,s)\in\mathcal{S}(G) \\ w_{p,s}\neq 0}} |r-s|^+ + \hat{\theta}_{p,s} \ .$$

Let $H_{\hat{\theta}}$ be a subgroup of G defined as

$$H_{\hat{\theta}} = \bigoplus_{(p,r,m)\in\mathcal{G}(G)} p^{\boldsymbol{\theta}(\hat{\theta})_{p,r}} \mathbb{Z}_{p^r}^{(m)}. \tag{7}$$

For $a \in J$ and $\hat{\theta} = (\hat{\theta}_{p,s})_{(p,s)\in\mathcal{S}(G)}$, let $T_{\hat{\theta}}(a)$ denote the set of vectors $\tilde{a} \in J$ such that

$$\tilde{a}_{p,s} - a_{p,s} \in p^{\hat{\theta}_{p,s}} \mathbb{Z}_{p^s}^{k_{p,s}} \backslash p^{\hat{\theta}_{p,s}+1} \mathbb{Z}_{p^s}^{k_{p,s}}, \forall (p,s) \in \mathcal{S}(G) \ .$$

Then we have $\left|T_{\hat{\theta}}(a)\right| = \prod_{(p,s)\in\mathcal{S}(G)} p^{(s-\hat{\theta}_{p,s})k_{p,s}}$ for all $a \in J$. Therefore, we may write $\left|T_{\hat{\theta}}(a)\right| = \left|T_{\hat{\theta}}\right|$ without any ambiguity. Let $\omega_{\hat{\theta}}$ be defined by

$$\omega_{\hat{\theta}} = \frac{\sum_{(p,s)\in\mathcal{S}(G)} \hat{\theta}_{p,s} w_{p,s} \log p}{\sum_{(p,s)\in\mathcal{S}(G)} s w_{p,s} \log p}, \tag{8}$$

we show in Appendix Section B the following result:

$$\log \left|T_{\hat{\theta}}\right| = (1 - \omega_{\hat{\theta}}) nR \ . \tag{9}$$

For any $a \in J$, $\left\{T_{\hat{\theta}}(a)\right\}_{\hat{\theta}}$ is a collection of disjoint sets whose union is $\cup_{\hat{\theta}} T_{\hat{\theta}}(a) = J$. Hence $\sum_{\hat{\theta}} \left|T_{\hat{\theta}}\right| = \sum_{\hat{\theta}} \left|T_{\hat{\theta}}(a)\right| = |J|$. Exploiting equation (9), we have that $\sum_{\hat{\theta}} 2^{(1-\omega_{\hat{\theta}})nR} = |J|$, or equivalently, $\sum_{\hat{\theta}} 2^{(1-\omega_{\hat{\theta}})} = 1$.

Definition 3: For an element $x \in G$ there is a one-to-one mapping $x \leftrightarrow ([x]_{\hat{\theta}}, \overline{x}_{\hat{\theta}})$, where $[x]_{\hat{\theta}}$ is the representative of the coset of $H_{\hat{\theta}}$ which x belongs to, and $\overline{x}_{\hat{\theta}} \in H_{\hat{\theta}}$, such that $x = [x]_{\hat{\theta}} + \overline{x}_{\hat{\theta}}$.

We use the following notations for the conditional distributions of the codeword and channel output given the *coset* information.

Definition 4: Let $H = H_{\hat{\theta}}$ be a subgroup of G and $x_r \in G$. Let X be distributed according to $P_X \equiv \text{Unif}(\mathcal{X})$, the uniform distribution over $\mathcal{X} = G$, and $W = (\mathcal{X}, \mathcal{Y}, W_{Y|X})$ be a classical channel. Then, for a representative $[x_r]_{\hat{\theta}}$ of a coset of H in G, define $P_{[X]}([x_r]) \triangleq \Pr([X] = [x_r]) = \frac{|H|}{|G|}, P_{X|[X]}(x \,|\, [x_r]) \triangleq \Pr(X = x \mid [X] = [x_r]) = \frac{1}{|H|}$ if $x \in [x_r], 0$ otherwise, and $P_{Y|[X]}(y \,|\, [x_r]) \triangleq \Pr(Y = y \mid [X] = [x_r]) = \sum_{x\in[x_r]+H} P_{X|[X]}(x|[x_r]) W_{Y|X}(y|x) = \sum_{x\in[x_r]+H} \frac{1}{|H|} W_{Y|X}(y|x)$, where we write $[x_r]$ and $[x_r]_{\hat{\theta}}$ interchangeably when the dependency of $\hat{\theta}$ is clear from the context.

IV. One-shot Classical Group Coding

Given a channel $W = (\mathcal{X} = G, \mathcal{Y}, W_{Y|X})$, let the joint distribution P_{XY} be $P_{XY} = P_X \cdot W_{Y|X}$ with P_X being the uniform distribution over $\mathcal{X}$, and P_Y be the marginal distribution of P_{XY} over $\mathcal{Y}$. Let $H = H_{\hat{\theta}}$ be a subgroup of G defined in (7) and $[X] = [X]_{\hat{\theta}}$ for some $\hat{\theta}$. Define

$$I_H^{\epsilon,\hat{\theta}}(\overline{X}; [X]Y) \triangleq D_H^{\epsilon}(P_{XY} \| P_{[X]} P_{X|[X]} P_{Y|[X]}). \tag{10}$$

A. Achievability

We have the first main result of this section.

Theorem 1: Let ϵ and $\left\{\epsilon_{\hat{\theta}}\right\}$ be given with $\epsilon_{\hat{\theta}} > 0$ for all $\hat{\theta}$ and $\sum_{\hat{\theta}} \epsilon_{\hat{\theta}} \leqslant \epsilon$. Then there exists a $(1, |J|, \epsilon')$-code such that

$$\epsilon' \leqslant \epsilon + \sum_{\hat{\theta}\neq\mathbf{s}} \exp_2 \left\{(1-\omega_{\hat{\theta}})R - I_H^{\epsilon_{\hat{\theta}}}(\overline{X}_{\hat{\theta}}; [X]_{\hat{\theta}} Y)\right\},$$

where the rate R is given in Equation (6).

Proof: Let the ensemble of homomorphisms ϕ from J to G and the group code $\mathbb{C} = \{\phi(a) + V | a \in J\}$ be given as in Definition 2 with $n = 1$. Given a channel $W = (\mathcal{X} = G, \mathcal{Y}, W_{Y|X})$, consider a decision region $A_\epsilon \subset \mathcal{X} \times \mathcal{Y}$, which will be constructed explicitly later, such that

$$P_{XY}(A_\epsilon) = \sum_{(x,y)\in A_\epsilon} P_X(x) W_{Y|X}(y|x) \geqslant 1 - \epsilon, \tag{11}$$

where P_X is uniform over G.

To find an achievable rate, we use a random coding argument in which the random encoder is characterized by the random homomorphism ϕ and a random vector V uniformly distributed over G. Given a message $u \in J$, the encoder maps it

to $x = \phi(u) + V$ and x is then fed to the channel. At the receiver, after receiving the channel output $y \in \mathcal{Y}$, the decoder looks for a unique $\tilde{u} \in J$ such that $(\phi(\tilde{u}) + V, y) \in A_\epsilon$. If the decoder does not find such $\tilde{u}$ or if such $\tilde{u}$ is not unique, it declares error. Thus, the error event can be characterized by the union of two events: $E(u) = E_1(u) \cup E_2(u)$ where $E_1(u)$ is the event that $(\phi(u) + V, y) \notin A_\epsilon$ and $E_2(u)$ is the event that there exists a $\tilde{u} \neq u$ such that $(\phi(\tilde{u}) + V, y) \in A_\epsilon$. We can provide an upper bound on the probability of the error event as $\Pr(E(u)) \leqslant \Pr(E_1(u)) + \Pr(E_2(u) \cap (E_1(u))^c)$.

In Appendix C, we show that $\Pr(E(u)) \leqslant \epsilon + \sum_{\hat{\theta} \neq \mathbf{s}} \left|T_{\hat{\theta}}(u)\right| \exp_2\left\{-I_H^{\epsilon_{\hat{\theta}}}(\overline{X}; [X]Y)\right\}$. The average probability of error of the group transmission scheme can be upper bounded by $\Pr(\text{error}) = \sum_{u \in J} \frac{1}{|J|} \Pr(E(u)) \leqslant \epsilon + \sum_{\hat{\theta} \neq \mathbf{s}} \left|T_{\hat{\theta}}(u)\right| \exp_2\left\{-I_H^{\epsilon_{\hat{\theta}}}(\overline{X}; [X]Y)\right\}$. Exploiting equation (9), we get the desired result in terms of the rate R of the code.

B. Converse

Toward the converse, we have the following theorem whose proof is given in Appendix D

Theorem 2: Assume that a group transmission system with parameters $(1, |J|, \epsilon)$ exists over a channel $(\mathcal{X} = G, \mathcal{Y}, W_{Y|X})$, and that the group J takes the form as in equation (3). Then the rate of the code, $R = \log |J|$, is bounded as:

$$R \leqslant \min_{\hat{\theta} \neq \mathbf{s}} \frac{1}{1 - \omega_{\hat{\theta}}} I_H^{\epsilon}(\overline{X}_{\hat{\theta}}; [X]_{\hat{\theta}} Y).$$

V. One-shot Classical-Quantum Group Coding

A. Achievability

Consider arbitrary output and input Abelian groups G and J, a CQ channel $\{\rho_x\}_{x \in G}$, and a subgroup $H_{\hat{\theta}}$ of G indexed by $\hat{\theta}$. We assume the uniform distribution on the input of the CQ channel yielding the input-output joint state as

$$\rho^{AB} = \sum_{x \in G} \frac{1}{|G|} |x\rangle\langle x|^A \otimes \rho_x^B, \tag{12}$$

where $\mathcal{A}$ denotes the input space and $\mathcal{B}$ denotes the output space. Define the transversal, the set of coset representatives of $H_{\hat{\theta}}$ in G, as $R_{\hat{\theta}}$. We drop $\hat{\theta}$ from the subscript, when it is clear from the context.

For any $H_{\hat{\theta}}$, using the one-to-one mapping $x \leftrightarrow ([x]_{\hat{\theta}}, \overline{x}_{\hat{\theta}})$, the joint state can be viewed as follows[1]

$$\rho^{AB} = \rho^{[A]\bar{A}B} \triangleq \sum_{[x], \overline{x}} \frac{1}{|G|} |[x], \overline{x}\rangle\langle[x], \overline{x}|^{[A]\bar{A}} \otimes \rho_{[x], \overline{x}}^B,$$

and we let $P_{[X]}([x]) = \frac{|H|}{|G|}$ for all $[x]$ and $P_{\overline{X}}(\overline{x}) = \frac{1}{|H|}$ for all $\overline{x}$. Here $[\mathcal{A}]$ denotes the space associated with $R_{\hat{\theta}}$ and $\bar{\mathcal{A}}$ that associated with $H_{\hat{\theta}}$. Define

$$\rho_{[x]}^B \triangleq \sum_{\overline{x}} P_{\overline{X}}(\overline{x}) \rho_{[x], \overline{x}}^B \ , \quad \rho^{\bar{A}} \triangleq \sum_{\overline{x}} P_{\overline{X}}(\overline{x}) |\overline{x}\rangle\langle\overline{x}|^{\bar{A}}$$
$$\rho^{[A]B} \triangleq \sum_{[x]} P_{[X]}([x]) |[x]\rangle\langle[x]|^{[A]} \otimes \rho_{[x]}^B \ .$$

[1]Recall that the classical variable (channel input) X with alphabet $\mathcal{X}$ is stored in a quantum register with Hilbert space also denoted as $\mathcal{X}$.

Also define the hypothesis testing mutual information:

$$I_H^{\epsilon, \hat{\theta}}(\overline{X}; [X]Y) \triangleq D_H^{\epsilon}(\rho^{[A]\bar{A}B} || \rho^{\bar{A}} \rho^{[A]B}) \ .$$

With these definitions, we are ready to state the main result of this section.

Theorem 3: Let ϵ and $\left\{\epsilon_{\hat{\theta}}\right\}$ be given with $\epsilon_{\hat{\theta}} > 0$ for all $\hat{\theta}$ and $\sum_{\hat{\theta}} \epsilon_{\hat{\theta}} \leqslant \epsilon$. Then there exists a (J, ϵ')-code such that

$$\epsilon' \leqslant \epsilon + \sum_{\hat{\theta} \neq \mathbf{s}} \exp_2\left\{(1 - \omega_{\hat{\theta}}) R - I_H^{\epsilon_{\hat{\theta}}}(\overline{X}_{\hat{\theta}}; [X]_{\hat{\theta}} Y)\right\},$$

where the rate R is given in Equation (6).

Proof: Let $\epsilon > 0$, and consider the optimizing POVM $(\Pi'')_{[X]}$ in $[A]\bar{A}B$ arising in the definition of $I_H^{\epsilon, \hat{\theta}}(\overline{X}; [X]Y)$. The POVM satisifies:

$$(\Pi'')_{[X]} = \sum_{[x], \overline{x}} |[x], \overline{x}\rangle\langle[x], \overline{x}|^{[A]\bar{A}} (\Pi'')_{[X];[x], \overline{x}}^B$$
$$\operatorname{tr}\left[(\Pi'')\rho^{AB}\right] \geqslant 1 - \epsilon \ , \operatorname{tr}\left[(\Pi'')\rho^{\bar{A}} \otimes \rho^{[A]B}\right] \leqslant 2^{-I_H^{\epsilon, \hat{\theta}}(\overline{X}; [X]Y)}. \tag{13}$$

We use the approach given in [26]. Define $\widehat{\mathcal{B}} \triangleq \mathcal{B} \otimes \mathbb{C}^2$. By Fact 2 in [26], there are orthogonal projections $\Pi_{[X];x}^{\widehat{B}}$ in $\widehat{\mathcal{B}}$ that give the same measurement probability on states $\sigma^B \otimes |0\rangle\langle 0|$ that POVM elements $(\Pi'')_{[X];x}^B$ give on states σ^B. Let $W_{[X];x}$ denote the orthogonal complement of the support of $\Pi_{[X];x}^{\widehat{B}}$ in $\widehat{\mathcal{B}}$.

Step 1: Consider a new Hilbert space $\mathcal{L}$ that is used only as a quantum register to store classical values, and define the extended output space

$$\mathcal{B}' \triangleq (\mathcal{B} \otimes \mathbb{C}^2) \oplus \bigoplus_{(p,r,m,k) \in \mathcal{G}^*(G)} (\mathcal{B} \otimes \mathbb{C}^2 \otimes \mathcal{L}^{(p,r,m,k)}).$$

where $\mathcal{L}^{(p,r,m,k)}$ is isomorphic to $\mathcal{L}$.

Remark 1: Define $\zeta(G) = \sum_{(p,r,m) \in \mathcal{G}(G)} r$, the sum of prime powers in the prime factorization of $|G|$. Also define the set $\mathcal{G}^*(G) \triangleq \{(p, r, m, k) : (p, r, m) \in \mathcal{G}(G), 1 \leqslant k \leqslant r\}$. We use a vector representation for the elements of G as follows. First we note that every element of a of $\mathbb{Z}_{p^r}$ can be represented uniquely as a r-length vector $(a_1, a_2, \ldots, a_r)$ such that a_i takes value in the transversal of $p^{r-i}\mathbb{Z}_{p^r}/p^{r-i+1}\mathbb{Z}_{p^r}$. A shorter vector $(a_1, a_2, \ldots, a_k)$ can represent any element in the subgroup $p^{r-k}\mathbb{Z}_{p^r}$. This is extended to all of G using the direct sum operation. We denote this mapping as $\nu(a)$. Hence every element of G is represented uniquely as a vector of length $\zeta(G)$. This is also extended to all subgroups of G. Furthermore, we use $(p, r, m, k) \in \mathcal{G}^*(G)$ to index the elements of a vector. The set formed by the (p, r, m, k)-th element of the $\nu(G)$ is denoted as $\mathcal{X}_{(p,r,m,k)}$. This also denotes the Hilbert space associated with the quantum register that stores the corresponding classical values, where a direct sum is replaced with a tensor product.

We extend the space corresponding to the alphabet of (p, r, m, k)-th element of $\nu(G)$ as follows. For every $(p, r, m, k) \in \mathcal{G}^*(G)$ define

$$\mathcal{X}'_{(p,r,m,k)} = \mathcal{X}_{(p,r,m,k)} \otimes \mathcal{L}^{(p,r,m,k)}.$$

This leads to the following extensions. For each subgroup $p^{r-s}\mathbb{Z}_{p^r}^{(m)}$, $s = 1, 2, \ldots, r$, the extended space is

$$p^{r-s}\mathbb{Z}_{p^r}^{(m)} \otimes \bigotimes_{k=1}^{s} \mathcal{L}^{(p,r,m,k)},$$

and for the transversal of $\mathbb{Z}_{p^r}^{(m)}/p^{r-s}\mathbb{Z}_{p^r}^{(m)}$, the extended input space is given by

$$(\mathbb{Z}_{p^r}^{(m)}/p^{r-s}\mathbb{Z}_{p^r}^{(m)}) \otimes \bigotimes_{k=s+1}^{r} \mathcal{L}^{(p,r,m,k)}.$$

Step 2: Let $0 \leqslant \delta \leqslant 1/10$. Consider a vector $\boldsymbol{\ell}$ indexed by (p, r, m, k) for $(p, r, m, k) \in \mathcal{G}^*(G)$, where $\ell_{(p,r,m,k)} \in \mathcal{L}^{(p,r,m,k)}$, and is a basis element of $\mathcal{L}^{(p,r,m,k)}$. Let $\boldsymbol{\ell}_{\hat{\theta}}$ denote the sub-vector corresponding to the subgroup $H_{\hat{\theta}}$, and the complementary sub-vector $[\boldsymbol{\ell}]_{\hat{\theta}}$ corresponding to the transversal $T_{\hat{\theta}}$. For the trivial case $H_{\hat{\theta}} = \{0\}$, we have $[\boldsymbol{\ell}] = \boldsymbol{\ell}$. Recall that in this case, $[X] = X$.

For any subgroup $H_{\hat{\theta}}$ of G, define the *tilting* map $\mathcal{T}_{[X];[\boldsymbol{\ell}],\delta} : \widehat{\mathcal{B}} \to \mathcal{B}'$ defined as

$$\mathcal{T}_{[X];[\boldsymbol{\ell}],\delta} : |h\rangle \mapsto \frac{1}{\sqrt{1+\zeta(G)\delta^2}}\big(|h\rangle + \sum_{(p,r,m,k)\in\mathcal{G}^*(G)/\mathcal{G}^*(H)} \delta|h\rangle|l_{(p,r,m,k)}\rangle\big) .$$

Note that we are tilting only along the direction of transversal of $H_{\hat{\theta}}$ in G. Define a state

$$(\rho')^{B'}_{x,\boldsymbol{\ell},\delta} \triangleq \mathcal{T}_{X;\boldsymbol{\ell},\delta}\left(\rho_x^B \otimes |0\rangle\langle 0|^{\mathbb{C}^2}.\right)$$

Consider the classical–quantum state

$$(\rho')^{A'B'} \triangleq |\mathcal{L}|^{-\zeta(G)} \sum_{x,\boldsymbol{\ell}} P_X(x)|x,\boldsymbol{\ell}\rangle\langle x,\boldsymbol{\ell}|^{A'} \otimes (\rho')^{B'}_{x,\boldsymbol{\ell},\delta} . \quad (14)$$

It can be shown that

$$\left\|(\rho')^{B'}_{x,\boldsymbol{\ell},\delta} - \rho_x^B \otimes |0\rangle\langle 0|^{\mathbb{C}^2}\right\|_1 \leqslant 2\zeta(G)\delta^2. \quad (15)$$

Define the tilted space

$$W'_{[X];x,\boldsymbol{\ell},\delta} \triangleq \mathcal{T}_{[X];[\boldsymbol{\ell}],\delta}(W_{[X];x})$$

residing in $\mathcal{B}'$. Define the subspace

$$W'_{x,\boldsymbol{\ell},\delta} \triangleq \bigoplus_{\hat{\theta}\in\Theta} W'_{[X];x,\boldsymbol{\ell},\delta}, \quad (16)$$

and $(\Pi')^{B'}_{W'} = (\Pi')^{B'}_{W'_{x,\boldsymbol{\ell},\delta}}$ the orthogonal projection in $\mathcal{B}'$ onto $W'_{x,\boldsymbol{\ell},\delta}$. Let $\Pi^{B'}_{\widehat{B}}$ be the orthogonal projection in $\mathcal{B}'$ onto $\widehat{\mathcal{B}}$. Define POVM elements

$$(\Pi')^{B'}_{x,\boldsymbol{\ell},\delta} \triangleq \left(\mathbf{1}^{B'} - (\Pi')^{B'}_{W'}\right)\Pi^{B'}_{\widehat{B}}\left(\mathbf{1}^{B'} - (\Pi')^{B'}_{W'}\right) , \quad (17)$$

$$(\Pi')^{A'B'} \triangleq \sum_{x,\boldsymbol{\ell}} |x,\boldsymbol{\ell}\rangle\langle x,\boldsymbol{\ell}|^{A'} \otimes (\Pi')^{B'}_{x,\boldsymbol{\ell},\delta} . \quad (18)$$

Define the following states for any subgroup $H_{\hat{\theta}}$:

$$(\rho')^{B'}_{[x],[\boldsymbol{\ell}]} \triangleq \frac{1}{|\mathcal{L}|^{\zeta(H)}} \sum_{\overline{x},\overline{\boldsymbol{\ell}}} P(\overline{x})(\rho')^{B'}_{x,\boldsymbol{\ell}}.$$

$$(\rho')^{B'} \triangleq \frac{1}{|\mathcal{L}|^{\zeta(G)}} \sum_{x,\boldsymbol{\ell}} P(x)(\rho')^{B'}_{x,\boldsymbol{\ell}}.$$

We make the following observations using the arguments given in [26]:

$$(\rho')^{B'}_{[x],[\boldsymbol{\ell}]} = \frac{1+(\zeta(G)-\zeta(H))\delta^2}{1+\zeta(G)\delta^2}\mathcal{T}_{[X];[\boldsymbol{\ell}]}(\rho^B_{[x]} \otimes |0\rangle\langle 0|) + N_{[X];[x],[\boldsymbol{\ell}]},$$

for some operators satisfying

$$\|N_{[X];[x],[\boldsymbol{\ell}]}\|_\infty \leqslant \frac{2\zeta(G)\delta}{\sqrt{|\mathcal{L}|}}. \quad (19)$$

Furthermore, using Holder's inequality we have

$$\|(\Pi')^{B'}_{x,\boldsymbol{\ell}}\|_1 \leqslant 2|\mathcal{B}|. \quad (20)$$

We perform random coding and perform error analysis and show the desired result. The details are given in Appendix E.

B. Converse

Toward the converse, we have the following theorem whose proof is given in Appendix F.

Theorem 4: Assume that a group transmission system with parameters $(1, |J|, \epsilon)$ exists over a classical-quatum channel $\mathcal{N} = \{\rho_x^B\}_{x\in\mathcal{X}}$, and that the group J takes the form as in equation (3). Then the rate of the code, $R = \log|J|$, is bounded as:

$$R \leqslant \min_{\hat{\theta}\neq\mathbf{s}} \frac{1}{1-\omega_{\hat{\theta}}} I_H^{\epsilon,\hat{\theta}}(\overline{X}; [X]Y).$$

VI. Group Coding in the Asymptotic Regime

We leverage the one-shot results for classical channel group transmission and show the following capacity result.

Theorem 5:

$$C = \min_{\hat{\theta}\neq\mathbf{s}} \frac{1}{1-\omega_{\hat{\theta}}} I(X; Y \mid [X]_{\hat{\theta}}).$$

Proof: The proof is provided in Appendix Section G. ■

Similarly, we provide a characterization of the capacity for the CQ channel:

Theorem 6:

$$C = \min_{\hat{\theta}\neq\mathbf{s}} \frac{1}{1-\omega_{\hat{\theta}}} I(X; Y \mid [X]_{\hat{\theta}}),$$

where

$$I(X; Y \mid [X]_{\hat{\theta}}) \triangleq D(\rho^{AB} || \rho^{\bar{A}_{\hat{\theta}}} \rho^{[A]_{\hat{\theta}}B}) .$$

Proof: The proof is provided in Appendix Section G. ■

REFERENCES

[1] A. S. Holevo, "The capacity of the quantum channel with general signal states," *IEEE Transactions on Information Theory*, vol. 44, no. 1, pp. 269–273, 1998.

[2] B. Schumacher and M. D. Westmoreland, "Sending classical information via noisy quantum channels," *Physical Review A*, vol. 56, no. 1, p. 131, 1997.

[3] M. Hayashi and H. Nagaoka, "General formulas for capacity of classical-quantum channels," *IEEE Transactions on Information Theory*, vol. 49, no. 7, pp. 1753–1768, 2003.

[4] D. Kretschmann and R. F. Werner, "Quantum channels with memory," *Physical Review A*, vol. 72, no. 6, p. 062323, 2005.

[5] R. Ahlswede and J. Gemma, "Bounds on algebraic code capacities for noisy channels. i," *Information and Control*, vol. 19, no. 2, pp. 124–145, 1971.

[6] ——, "Bounds on algebraic code capacities for noisy channels. ii," *Information and Control*, vol. 19, no. 2, pp. 146–158, 1971.

[7] T. J. Goblick, "Coding for a discrete information source with a distortion measure," Ph.D. dissertation, Massachusetts Institute of Technology, 1963.

[8] R. Dobrusin, "Asymptotic bounds of the probability of error for the transmission of messages over a discrete memoryless channel with a symmetric transition probability matrix," *Teor. Veroyatnost. i Primenen*, vol. 7, pp. 283–311, 1962.

[9] R. Dobrushin, "Asymptotic optimality of group and systematic codes for some channels," *Theory of Probability & Its Applications*, vol. 8, no. 1, pp. 47–60, 1963.

[10] G. D. Forney and M. D. Trott, "The dynamics of group codes: state spaces, trellis diagrams, and canonical encoders," *IEEE Transactions on Information Theory*, vol. 39, no. 5, pp. 1491–1513, 1993.

[11] G. Como and F. Fagnani, "The capacity of finite abelian group codes over symmetric memoryless channels," *IEEE Transactions on Information Theory*, vol. 55, no. 5, pp. 2037–2054, 2009.

[12] H.-A. Loeliger, "Signal sets matched to groups," *IEEE Transactions on Information Theory*, vol. 37, no. 6, pp. 1675–1682, 1991.

[13] H.-A. Loeliger and T. Mittelholzer, "Convolutional codes over groups," *IEEE Transactions on Information Theory*, vol. 42, no. 6, pp. 1660–1686, 1996.

[14] R. Garello and S. Benedetto, "Multilevel construction of block and trellis group codes," *IEEE Transactions on Information Theory*, vol. 41, no. 5, pp. 1257–1264, 1995.

[15] A. G. Sahebi and S. S. Pradhan, "Abelian group codes for channel coding and source coding," *IEEE Transactions on Information Theory*, vol. 61, no. 5, pp. 2399–2414, 2015.

[16] D. Slepian, "Group codes for the gaussian channel," *Bell System Technical Journal*, vol. 47, no. 4, pp. 575–602, 1968.

[17] A. G. Sahebi and S. S. Pradhan, "Multilevel polarization of polar codes over arbitrary discrete memoryless channels," in *2011 49th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE, 2011, pp. 1718–1725.

[18] E. Şaşoğlu, E. Telatar, and E. Arikan, "Polarization for arbitrary discrete memoryless channels," in *2009 IEEE Information Theory Workshop*. IEEE, 2009, pp. 144–148.

[19] W. Park and A. Barg, "Polar codes for q-ary channels, $q = 2^r$," *IEEE Transactions on Information Theory*, vol. 59, no. 2, pp. 955–969, 2012.

[20] D. Krithivasan and S. S. Pradhan, "Distributed source coding using abelian group codes: A new achievable rate-distortion region," *IEEE Transactions on Information Theory*, vol. 57, no. 3, pp. 1495–1519, 2011.

[21] A. Padakandla, A. G. Sahebi, and S. S. Pradhan, "An achievable rate region for the three-user interference channel based on coset codes," *IEEE Transactions on Information Theory*, vol. 62, no. 3, pp. 1250–1279, 2016.

[22] S.-N. Hong and G. Caire, "On interference networks over finite fields," *IEEE Transactions on Information Theory*, vol. 60, no. 8, pp. 4902–4921, 2014.

[23] R. Renner, S. Wolf, and J. Wullschleger, "The single-serving channel capacity," in *2006 IEEE International Symposium on Information Theory*. IEEE, 2006, pp. 1424–1427.

[24] M. Mosonyi and N. Datta, "Generalized relative entropies and the capacity of classical-quantum channels," *Journal of Mathematical physics*, vol. 50, no. 7, 2009.

[25] L. Wang and R. Renner, "One-shot classical-quantum capacity and hypothesis testing," *Physical Review Letters*, vol. 108, no. 20, p. 200501, 2012.

[26] P. Sen, "Unions, intersections and a one-shot quantum joint typicality lemma," *Sādhanā*, vol. 46, no. 1, p. 57, 2021.

[27] T. Ogawa and H. Nagaoka, "Strong converse and stein's lemma in quantum hypothesis testing," *IEEE Transactions on Information Theory*, vol. 46, no. 7, pp. 2428–2433, 2000.

APPENDIX

A. Hypothesis testing relative entropy

We recall the definition the *hypothesis testing relative entropy* $D_{\mathrm{H}}^{\epsilon}(\rho\|\sigma)$ from [25]. Let ρ and σ be two possible states of a system, and Q a positive operator with $0 \leqslant Q \leqslant I$. The hypothesis testing relative entropy $D_{\mathrm{H}}^{\epsilon}(\rho\|\sigma)$ is defined by

$$D_{\mathrm{H}}^{\epsilon}(\rho\|\sigma) \triangleq -\log_2 \inf_{\substack{Q:0\leqslant Q\leqslant I,\\ \mathrm{tr}[Q\rho]\geqslant 1-\epsilon}} \mathrm{tr}[Q\sigma]. \tag{21}$$

For two distributions P, Q on a classical alphabet $\mathcal{X}$, the hypothesis testing relative entropy $D_H^{\epsilon}(P\|Q)$ can be similarly defined by

$$D_H^{\epsilon}(P\|Q) = -\log_2 \inf_{A:P(A)\geqslant 1-\epsilon} Q(A),$$

where $A \subset \mathcal{X}$ is sometimes called the *decision region*.

B. More on the $H_{\hat{\theta}}$ coset

We show the derivation for (9) as follows:

$$\begin{aligned}
\log\left|T_{\hat{\theta}}\right| &= \sum_{(p,s)\in\mathcal{S}(G)} (s-\hat{\theta}_{p,s})k_{p,s}\log p \\
&= k\cdot\left[\frac{\sum_{(p,s)\in\mathcal{S}(G)}(s-\hat{\theta}_{p,s})\omega_{p,s}\log p}{\sum_{(p,s)\in\mathcal{S}(G)} s\omega_{p,s}\log p}\right] \\
&\quad\times \sum_{(p,s)\in\mathcal{S}(G)} s\omega_{p,s}\log p \\
&= (1-\omega_{\hat{\theta}})k \sum_{(p,s)\in\mathcal{S}(G)} s\omega_{p,s}\log p = (1-\omega_{\hat{\theta}})nR.
\end{aligned} \tag{22}$$

We recall the following result from [15].

Lemma 2: For $a, \tilde{a} \in J$, $x, \tilde{x} \in G^n$ and for $(p,s) \in \mathcal{Q}(J) = \mathcal{S}(G)$, let $\hat{\theta}_{p,s} \in \{0,1,\cdots,s\}$ be such that $\tilde{a}_{p,s} - a_{p,s} \in p^{\hat{\theta}_{p,s}}\mathbb{Z}_{p^s}^{k_{p,s}} \backslash p^{\hat{\theta}_{p,s}+1}\mathbb{Z}_{p^s}^{k_{p,s}}$, i.e., $\tilde{a} \in T_{\hat{\theta}}(a)$. Consider a random homomorphism ϕ and a dither V with distribution specified as in Definition 2. Then,

$$\Pr\begin{pmatrix}\phi(a)+V=x,\\ \phi(\tilde{a})+V=\tilde{x}\end{pmatrix} = \begin{cases} \frac{1}{|G|^n}\frac{1}{|H_{\hat{\theta}}|^n} & \text{if } \tilde{x}-x \in H_{\hat{\theta}}^n \\ 0 & \text{otherwise}\end{cases}$$

C. Proof of Achievability for Classical Channel

Denote by $P_{err}(u)$ the probability of the event $E_2(u) \cap (E_1(u))^c$, averaged over the randomness of ϕ, V.

$$\begin{aligned}
P_{err}(u) &\triangleq \mathbb{E}\left[\Pr(E_2(u)\cap(E_1(u))^c)\right] \\
&= \sum_{(x,y)\in A_\epsilon} W(y|x)\Pr\begin{pmatrix}\phi(u)+V=x,\exists\tilde{u}\in J:\\ \tilde{u}\neq u,(\phi(\tilde{u})+V,y)\in A_\epsilon\end{pmatrix} \\
&\leqslant \sum_{(x,y)\in A_\epsilon}\sum_{\substack{\tilde{u}\in J\\ \tilde{u}\neq u}}\sum_{\substack{\tilde{x}\in G\\ (\tilde{x},y)\in A_\epsilon}} W(y|x)\Pr\begin{pmatrix}\phi(u)+V=x,\\ \phi(\tilde{u})+V=\tilde{x}\end{pmatrix} \\
&= \sum_{\hat{\theta}\neq\mathbf{s}}\sum_{(x,y)\in A_\epsilon}\sum_{\tilde{u}\in T_{\hat{\theta}}(u)}\sum_{\substack{\tilde{x}\in G\\ (\tilde{x},y)\in A_\epsilon}} W(y|x)\Pr\begin{pmatrix}\phi(u)+V=x,\\ \phi(\tilde{u})+V=\tilde{x}\end{pmatrix} \\
&= \sum_{\hat{\theta}\neq\mathbf{s}} P_{err}(u,\hat{\theta}),
\end{aligned} \tag{23}$$

where $P_{err}(u,\hat{\theta}) \triangleq \sum_{\tilde{u}\in T_{\hat{\theta}}(u)} P_{err}(u,\tilde{u})$ and for $\tilde{u} \in T_{\hat{\theta}}(u)$,

$$P_{err}(u,\tilde{u}) \triangleq \sum_{(x,y)\in A_\epsilon}\sum_{\substack{\tilde{x}\in G\\ (\tilde{x},y)\in A_\epsilon}} W(y|x)\Pr\begin{pmatrix}\phi(u)+V=x,\\ \phi(\tilde{u})+V=\tilde{x}\end{pmatrix}, \tag{24}$$

and $\mathbf{s}$ denote the vector whose compnents satisfy $\mathbf{s}_{(p,s)} = s$ for all $(p,s) \in \mathcal{S}(G)$. The term $\Pr(\phi(u)+V=x,\ \phi(\tilde{u})+V=\tilde{x})$ in (24) can be found using Lemma 2. Let $x \in [x_r]$ be a shorthand for $[x] = [x_r]$, or equivalently, $x \in [x_r]+H$. Hence

$$\begin{aligned}
P_{err}(u,\tilde{u}) &= \sum_{(x,y)\in A_\epsilon}\sum_{\tilde{x}\in x+H_{\hat{\theta}}} W(y|x)\mathbb{I}_{A_\epsilon}(\tilde{x},y)\frac{1}{|G|}\frac{1}{|H_{\hat{\theta}}|} \\
&= \sum_{x\in\mathcal{X}}\sum_{\tilde{x}\in x+H_{\hat{\theta}}}\sum_{\substack{y:(x,y)\in A_\epsilon\\ (\tilde{x},y)\in A_\epsilon}} W(y|x)\frac{1}{|G|}\frac{1}{|H_{\hat{\theta}}|} \\
&\leqslant \sum_{x\in\mathcal{X}}\sum_{\tilde{x}\in x+H_{\hat{\theta}}}\sum_{y:(\tilde{x},y)\in A_\epsilon} W(y|x)\frac{1}{|G|}\frac{1}{|H_{\hat{\theta}}|} \\
&= \sum_{[x_r]}\frac{|H_{\hat{\theta}}|}{|G|}\sum_{x\in[x_r]}\sum_{\tilde{x}\in[x_r]}\sum_{y:(\tilde{x},y)\in A_\epsilon} W(y|x)\frac{1}{|H_{\hat{\theta}}|^2} \\
&= \sum_{[x_r]} P([x_r])\sum_{\tilde{x}\in[x_r]}\sum_{y:(\tilde{x},y)\in A_\epsilon} P(\tilde{x}\,|\,[x_r])P(y\,|\,[x_r]).
\end{aligned} \tag{25}$$

Let a set of parameters $\{\epsilon_{\hat{\theta}}\}_{\hat{\theta}\neq\mathbf{s}}$ be given such that $\epsilon_{\hat{\theta}} > 0$ for each $\hat{\theta}$ and that $\sum_{\hat{\theta}}\epsilon_{\hat{\theta}} = \epsilon$. Let the set $A_{\epsilon_{\hat{\theta}}}^*$ be a maximizer of the right-hand side of (10) for $I_H^{\epsilon_{\hat{\theta}},\hat{\theta}}(\overline{X};[X]Y)$, i.e., $P_{XY}(A_{\epsilon_{\hat{\theta}}}^*) \geqslant 1-\epsilon_{\hat{\theta}}$ and

$$\begin{aligned}
&I_H^{\epsilon_{\hat{\theta}},\hat{\theta}}(\overline{X};[X]Y) = -\log_2 \\
&\quad\Big[\sum_{[x_r]} P([x_r])\sum_{x\in[x_r]} P(x\,|\,[x_r])\sum_{y:(x,y)\in A_{\epsilon_{\hat{\theta}}}^*} P(y\,|\,[x_r])\Big].
\end{aligned}$$

Now we set explicitly $A_\epsilon = \cap_{\hat{\theta}} A_{\epsilon_{\hat{\theta}}}^*$. The probability of the event $E_1(u)$ can be bounded as:

$$\begin{aligned}
\Pr(E_1(u)) &= \Pr\left((\phi(u)+V,Y)\notin A_\epsilon\right) = \Pr((X,Y)\notin A_\epsilon) \\
&= \Pr\left((X,Y)\in\cup_{\hat{\theta}}(A_{\epsilon_{\hat{\theta}}}^*)^C\right) \\
&\leqslant \sum_{\hat{\theta}}\Pr\left((X,Y)\in(A_{\epsilon_{\hat{\theta}}}^*)^C\right) \leqslant \sum_{\hat{\theta}}\epsilon_{\hat{\theta}} = \epsilon.
\end{aligned}$$

Since $A_\epsilon \subset A_{\epsilon_{\hat{\theta}}}^*$, the term $P_{err}(u,\tilde{u})$ in (24) can then be upper bounded by

$$\begin{aligned}
&\sum_{[x_r]} P([x_r])\sum_{\tilde{x}\in[x_r]}\sum_{y:(\tilde{x},y)\in A_{\epsilon_{\hat{\theta}}}^*} P(\tilde{x}|\,[x_r])P(y|\,[x_r]) \\
&= \exp_2\{-I_H^{\epsilon_{\hat{\theta}},\hat{\theta}}(\overline{X};[X]Y)\},
\end{aligned}$$

which leads to the following bound,

$$P_{err}(u) \leqslant \sum_{\hat{\theta}\neq\mathbf{s}}\left|T_{\hat{\theta}}(u)\right|\exp_2\{-I_H^{\epsilon_{\hat{\theta}},\hat{\theta}}(\overline{X};[X]Y)\}. \tag{26}$$

Therefore, we have

$$\Pr(E(u)) \leqslant \epsilon + \sum_{\hat{\theta}\neq\mathbf{s}}\left|T_{\hat{\theta}}(u)\right|\exp_2\{-I_H^{\epsilon_{\hat{\theta}},\hat{\theta}}(\overline{X};[X]Y)\}.$$

The average probability of error of the group transmission scheme can be upper bounded by

$$\begin{aligned}\Pr(\text{error}) &= \sum_{u\in J}\frac{1}{|J|}\Pr(E(u))\\ &\leqslant \epsilon + \sum_{\hat{\theta}\neq \mathbf{s}} \left|T_{\hat{\theta}}(u)\right| \exp_2\{-I_H^{\epsilon_{\hat{\theta}},\hat{\theta}}(\overline{X};[X]Y)\}.\end{aligned}$$

Example 2: Let $J=\mathbb{Z}_4, G=\mathbb{Z}_8$. In this example, we have $\mathcal{G}(G)=\{(2,3,1)\}$ and $\mathcal{G}(J)=\{(2,2,1)\}$, $k_{2,1}=0, k_{2,2}=1, k_{2,3}=0$, and the term $g_{(2,2,1)\to(2,3,1)}$ is a uniform random variable over $2\mathbb{Z}_8$. For simplicity, we write $u=u_{2,2,1}\in J$ and $g=g_{(2,2,1)\to(2,3,1)}\in\mathbb{Z}_8$. Then $r_2=\max\mathcal{R}_2(G)=3$ and the set $\mathcal{Q}(J)=\mathcal{S}(G)=\{(2,1),(2,2),(2,3)\}$, and $\mathbf{s}_{(2,1)}=1, \mathbf{s}_{(2,2)}=2, \mathbf{s}_{(2,3)}=3$. For distinct $u,\tilde{u}\in J$, the vector $\hat{\theta}=(\hat{\theta}_{2,1},\hat{\theta}_{2,2},\hat{\theta}_{2,3})$ for which $\tilde{u}\in T_{\hat{\theta}}(u)$ must have $\hat{\theta}_{2,1}=1, 0\leqslant\hat{\theta}_{2,2}<2, \hat{\theta}_{2,3}=3$. Thus $P_{err}(u)\leqslant P_{err}(u,(1,0,3))+P_{err}(u,(1,1,3))$.

The set $\mathcal{Q}(G)=\{(2,3)\}$, so $\boldsymbol{\theta}(\hat{\theta})=\boldsymbol{\theta}(\hat{\theta})_{(2,3)}$ and

$$\begin{aligned}\boldsymbol{\theta}(\hat{\theta})_{(2,3)} &= \min_{\substack{(2,s)\in\mathcal{S}(G)\\ w_{2,s}\neq 0}}\left\{|3-s|^+ + \hat{\theta}_{2,s}\right\}\\ &= |3-2|^+ + \hat{\theta}_{2,2} = 1+\hat{\theta}_{2,2}.\end{aligned}$$

Case 1: $\hat{\theta}_{2,2}=0$, $\boldsymbol{\theta}(\hat{\theta})_{(2,3)}=1$

For $\tilde{u}\in T_{\hat{\theta}}(u)$, $\tilde{u}-u\in Z_4\backslash 2\mathbb{Z}_4$, and $H_{\hat{\theta}}=2\mathbb{Z}_8$. Let $A^*_{\epsilon/2,\hat{\theta}}$ be a maximizer for $D_H^{\epsilon/2}(P_{XY}\|P_{[X]_{\hat{\theta}}}P_{X|[X]_{\hat{\theta}}}P_{Y|[X]_{\hat{\theta}}})$. Thus we have

$$P_{err}(u,(1,0,3))\leqslant \left|T_{(1,0,3)}(u)\right|\exp_2\{-I_H^{\epsilon/2}(\overline{X};[X]Y)\}.$$

Case 2: $\hat{\theta}_{2,2}=1$, $\boldsymbol{\theta}(\hat{\theta})_{(2,3)}=2$ In this case, $H_{\hat{\theta}}=4\mathbb{Z}_8$, and we have

$$P_{err}(u,(1,1,3))\leqslant \left|T_{(1,1,3)}(u)\right|\exp_2\{-I_H^{\epsilon_{\hat{\theta}},\hat{\theta}}(\overline{X};[X]Y)\}.$$

Therefore the error probability for a message u is

$$\begin{aligned}&\Pr(E(u))\leqslant \Pr(E_1(u))+\Pr(E_2(u)\cap(E_1(u))^c)\\ &\leqslant \epsilon + \left|T_{(1,0,3)}(u)\right|\exp_2\{-I_H^{\epsilon_{\hat{\theta}},\hat{\theta}}(\overline{X};[X]Y)\}_{\hat{\theta}=(1,0,3)}\\ &+ \left|T_{(1,1,3)}(u)\right|\exp_2\{-I_H^{\epsilon_{\hat{\theta}},\hat{\theta}}(\overline{X};[X]Y)\}_{\hat{\theta}=(1,1,3)},\end{aligned}$$

where we consider the decision region $A_\epsilon=\cap_{\hat{\theta}}A^*_{\epsilon/2,\hat{\theta}}$.

D. Proof of Converse for Classical Channel

Based on Lemma 1, for each group code $\mathbb{C}\leqslant G$, there exists a group J and a homomorphism such that $\mathbb{C}$ is the image of the homomorphism. Assume now that a group transmission system with parameters $(1,|J|,\epsilon)$ exists over a channel $(\mathcal{X}=G,\mathcal{Y},W_{Y|X})$, and that the group J takes the form in equation (3). Assume that the homomorphism ϕ for the group code $\mathbb{C}$ is a one-to-one mapping. We have:

$$\mathbb{C}=\Big\{\bigoplus_{(p,r,m)\in\mathcal{G}(G)}\sum_{s=1}^{r_p}u_{p,s}g_{(p,s)\to(r,m)}+V : u\in J\Big\}.$$

Let $\hat{\theta}$ be a vector indexed by $(p,s)\in\mathcal{S}(G)$ with $0\leqslant\hat{\theta}_{p,s}\leqslant s$. For an message $u\in J$, construct a one-to-one correspondence between $u_{p,s}\in\mathbb{Z}_{p^s}^{k_{p,s}}$ and the tuple $(\tilde{u}_{p,s},\hat{u}_{p,s})$ where $\tilde{u}_{p,s}\in p^{\hat{\theta}_{p,s}}\mathbb{Z}_{p^s}^{k_{p,s}}$ and $\hat{u}_{p,s}\in\mathbb{Z}_{p^{\hat{\theta}_{p,s}}}^{k_{p,s}}$. Let U denote the random message of the group transmission system of the code. Let $\hat{U}$ denote the part of the random message such that $\hat{U}_{p,s}\in\mathbb{Z}_{p^{\hat{\theta}_{p,s}}}^{k_{p,s}}$, for all $(p,s)\in\mathcal{S}(G)$. Consider the subcode of $\mathbb{C}$:

$$\begin{aligned}\mathbb{C}_1(\hat{\theta},\hat{u})=\Big\{&\bigoplus_{(p,r,m)\in\mathcal{G}(G)}\sum_{s=1}^{r_p}(\tilde{u}_{p,s}+\hat{u}_{p,s})g_{(p,s)\to(r,m)}+V :\\ &\tilde{u}_{p,s}\in p^{\hat{\theta}_{p,s}}\mathbb{Z}_{p^s}^{k_{p,s}},\forall(p,s)\in\mathcal{S}(G)\Big\}.\end{aligned}$$

Let $x=\phi(u)+V$ be the channel input and $H_{\hat{\theta}}$ be given as in (7). Then $\mathbb{C}_1(\hat{\theta},\hat{u})=x+H_{\hat{\theta}}$. That is, there is a one-to-one correspondence between $\hat{U}$ and $[X]_{\hat{\theta}}$. Also, $\left|T_{\hat{\theta}}(u)\right|=\left|\mathbb{C}_1(\hat{\theta},\hat{u})\right|=\left|H_{\hat{\theta}}\right|$ for all $u\in J$. Let $x\in[x_r]$ be a shorthand for $[x]=[x_r]$, or equivalently, $x\in[x_r]+H$, where we omit the $\hat{\theta}$ subscript when it is clear from the context.

Define a one-to-one correspondence between x and the tuple $(\tilde{x}_{\hat{\theta}},[x]_{\hat{\theta}})$ where $\tilde{x}_{\hat{\theta}}=\phi(\tilde{u})$. Consider a genie-aided receiver which gets access to $\hat{U}$ and performs maximum likelihood decoding. Equivalently, this receiver has access to the coset information $[X]_{\hat{\theta}}$ of X and can be written as $\mathcal{D}^{ga}:([x]_{\hat{\theta}},y)\mapsto x'\in\mathcal{X}$. Clearly the average probability of error for this decoder must be not greater than ϵ. Let $X'\in\mathcal{X}$ be the output of $\mathcal{D}^{ga}$. For every $\hat{\theta}$ with $0\leqslant\hat{\theta}_{p,s}\leqslant s$, $\hat{\theta}\neq\mathbf{s}$, the average probability of error for this decoder is

$$\begin{aligned}&\sum_{\hat{u}}\sum_{x,x'}\Pr(\hat{u})P_{XX'|\hat{U}}(x,x'\mid\hat{u})\mathbf{1}_{\{x'\neq x\}}\\ &=\sum_{x,x'}P_{XX'}(x,x')\mathbf{1}_{\{x'\neq x\}}\leqslant\epsilon,\end{aligned}$$

where

$$P_{XX'|\hat{U}}(x,x'|\hat{u})\triangleq P(x|[x])\sum_{y:\mathcal{D}^{ga}([x],y)=x'}W(y|x).$$

Consider a strategy to distinguish $P_{XX'}$ and $P_{\hat{U}}(P_{X|\hat{U}}\otimes P_{X'|\hat{U}})$ as follows. The strategy guesses $P_{XX'}$ if it sees $X=X'$, and guesses $P_{\hat{U}}(P_{X|\hat{U}}\otimes P_{X'|\hat{U}})$ otherwise. When $P_{XX'}$ is the true underlying distribution, the type-I error probability is exactly the probability that $X\neq X'$ computed from $P_{XX'}$, namely, the average probability of a decoding error, and is thus not larger than ϵ. When $P_{\hat{U}}(P_{X|\hat{U}}\otimes P_{X'|\hat{U}})$ is the true underlying distribution, the probability of type-II error (misdetection) is

$$\begin{aligned}&\sum_{\hat{u}}P_{\hat{U}}(\hat{u})\sum_{x,x'}P_{X|\hat{U}}(x\mid\hat{u})P_{X'|\hat{U}}(x'\mid\hat{u})\mathbf{1}_{\{x'=x\}}\\ &=\sum_{[x_r]}P([x_r])\sum_{x,x'}P(x\mid[x_r])P(x'\mid[x_r])\mathbf{1}_{\{x'=x\}} \qquad (27)\\ &=\sum_{[x_r]}\frac{|H|}{|G|}\sum_{x}P(x\mid[x_r])P(x\mid[x_r])\\ &=\sum_{[x_r]}\frac{|H|}{|G|}\sum_{x\in[x_r]}\frac{1}{|H|}P(x\mid[x_r])=\frac{1}{|H|},\end{aligned}$$

where (27) follows from the one-to-one mapping between $\hat{U}$ and $[X]_{\hat{\theta}}$. Thus,

$$\begin{aligned}
&D_H^{\epsilon,\hat{\theta}}(P_{XY}\|P_{[X]}P_{X|[X]}P_{Y|[X]})\\
\geqslant&D_H^{\epsilon,\hat{\theta}}(P_{XX'}\|P_{[X]}P_{X|[X]}P_{X'|[X]})\\
=&D_H^{\epsilon}(P_{XX'}\|P_{\hat{U}}(P_{X|\hat{U}}\otimes P_{X'|\hat{U}}))\\
\geqslant&-\log_2\frac{1}{|H|}=\log_2|H|=\log_2\left|T_{\hat{\theta}}(u)\right|\\
=&(1-\omega_{\hat{\theta}})k\sum_{(p,s)\in\mathcal{S}(G)}s\omega_{p,s}\log p., \qquad (28)
\end{aligned}$$

where the first inequality follows from the DPI. Equivalently,

$$I_H^{\epsilon,\hat{\theta}}(\overline{X};[X]Y)\geqslant(1-\omega_{\hat{\theta}})R\,,$$

which yields Theorem 2.

E. Proof of Achievability for CQ Channel

Remark 2: We make a general remark here about an arbitrary subgroup H of G. We note that $[x]=x \mod H$. Using the distributive property of $\mod$ operation, i.e., $[x+y]=[[x]+y]$, we note that

$$[x]=[\Phi(u)+\beta]=[[\Phi([u])+\Phi(\overline{u})]+\beta]=[[\Phi([u])]+\beta],$$

$$\overline{x}=x-[x]=\Phi(\overline{u})+\overline{\Phi([u])+\beta}.$$

Hence $[x]$ depends only on $[u]$, where as $\overline{x}$ depends on the entirety of u in general.

Step 3: Probability of Error Analysis: We construct a random code as follows. We generate the random homomorphism as stated in the previous section. In addition, we generate for every $u\in J$, a random vector $\boldsymbol{\ell}$ independently and uniformly such that $[\boldsymbol{\ell}]$ depends only on $[u]$ and $\bar{\boldsymbol{\ell}}$ depends on the entire u. (Need to show the existence of such a PMF). We start by computing the average probability of error for a fixed u using Hayashi-Nagaoka inequality [3] as follows:

$$\begin{aligned}
&\mathbb{E}_C(P(\text{Error}|u))=\mathbb{E}_C\mathrm{tr}\left[\left(\mathbf{1}^{B'}-\Lambda_u^{B'}\right)(\rho')^{B'}_{u,\delta}\right]\\
&\leqslant2\mathbb{E}_C\mathrm{tr}\left[\left(\mathbf{1}^{B'}-(\Pi')^{B'}_{(x,\boldsymbol{\ell})(u)}(\rho')^{B'}_{(x,\boldsymbol{\ell})(u),\delta}\right)\right]\\
&+4\sum_{\hat{\theta}\in\Theta}\sum_{u'\in T_{\hat{\theta}}(u)}\mathbb{E}_C\mathrm{tr}\Big[(\Pi')^{B'}_{([x],[\boldsymbol{\ell}])([u']),(\overline{x},\bar{\boldsymbol{\ell}})(u')}\\
&\quad\times(\rho')^{B'}_{([x],[\boldsymbol{\ell}])([u]),(\overline{x},\bar{\boldsymbol{\ell}})(u),\delta}\Big]
\end{aligned}$$

We work on the two terms in the right hand side. The first term can be simplified as follows:

$$\begin{aligned}
&\mathbb{E}_C\mathrm{tr}\left[\left(\mathbf{1}^{B'}-(\Pi')^{B'}_{([x],[\boldsymbol{\ell}])([u]),(\overline{x},\bar{\boldsymbol{\ell}})(u)}\right)(\rho')^{B'}_{([x],[\boldsymbol{\ell}])([u]),(\overline{x},\bar{\boldsymbol{\ell}})(u),\delta}\right]\\
&=\frac{1}{|\mathcal{L}|^{\zeta(G)}}\sum_{x,\boldsymbol{\ell}}P(x)\mathrm{tr}[(\rho')^{B'}_{x,\boldsymbol{\ell}}-(\Pi')^{B'}_{x,\boldsymbol{\ell}}(\rho')^{B'}_{x,\boldsymbol{\ell}}]\\
&\overset{(a)}{\leqslant}\frac{4}{|\mathcal{L}|^{\zeta(G)}}\sum_{x,\boldsymbol{\ell}}P(x)\mathrm{tr}[(I-(\Pi)^{B'}_{\hat{B}})(\rho')^{B'}_{x,\boldsymbol{\ell}}+(\Pi')^{B'}_{W'_{x,\boldsymbol{\ell}}}(\rho')^{B'}_{x,\boldsymbol{\ell}}]\\
&\overset{(b)}{\leqslant}16\zeta(G)\delta^2+\frac{4}{|\mathcal{L}|^{\zeta(G)}}\sum_{x,\boldsymbol{\ell}}P(x)\mathrm{tr}[(\Pi')^{B'}_{W'_{x,\boldsymbol{\ell}}}(\rho^B_x\otimes|0\rangle\langle0|)]\\
&\overset{(c)}{\leqslant}16\zeta(G)\delta^2+\frac{4(1+\zeta(G)\delta^2)}{|\mathcal{L}|^{\zeta(G)}\delta^2}\sum_{x,\boldsymbol{\ell}}P(x)\\
&\qquad\times\sum_{\hat{\theta}\in\Theta}\left[1-\mathrm{tr}[(\Pi)^{\hat{B}}_{[X];x}(\rho^B_x\otimes|0\rangle\langle0|)]\right]\\
&=16\zeta(G)\delta^2+\frac{4(1+\zeta(G)\delta^2)}{|\mathcal{L}|^{\zeta(G)}\delta^2}\sum_{x,\boldsymbol{\ell}}P(x)\\
&\qquad\times\sum_{\hat{\theta}\in\Theta}[1-\mathrm{tr}[(\Pi'')^{B}_{[X];x}\rho^B_x]\\
&\overset{(d)}{\leqslant}16\zeta(G)\delta^2+\frac{4(1+\zeta(G)\delta^2)|\Theta|\epsilon}{\delta^2},
\end{aligned}$$

where we provide the following arguments. (a) follows from Fact 3 of [26] and (b) from (15). (c) follows from Proposition 2 [26] by using $l=|\Theta|$, and $\alpha=\frac{\delta^2}{1+\zeta(G)\delta^2}$. (d) follows from (13).

Next we look at the second term as follows. Note that

$$T_{\hat{\theta}}(u)=\{u':[u']=[u],\overline{u'}\neq\overline{u}\}.$$

For any $u'\in T_{\hat{\theta}}(u)$, we have

$$\begin{aligned}
&\mathbb{E}_C\mathrm{tr}\left[\left((\Pi')^{B'}_{([x],[\boldsymbol{\ell}])([u]),(\overline{x},\bar{\boldsymbol{\ell}})(u')}\right)(\rho')^{B'}_{([x],[\boldsymbol{\ell}])([u]),(\overline{x},\bar{\boldsymbol{\ell}})(u),\delta}\right]\\
&=\frac{1}{|\mathcal{L}|^{\zeta(G)+\zeta(H)}}\sum_{[x],[\boldsymbol{\ell}],\overline{x},\bar{\boldsymbol{\ell}}}\sum_{\overline{x}',\bar{\boldsymbol{\ell}}'}P([x])P(\overline{x})P(\overline{x'})\\
&\quad\times\mathrm{tr}[(\Pi')^{B'}_{[x],[\boldsymbol{\ell}],\overline{x'},\bar{\boldsymbol{\ell}}'}(\rho')^{B'}_{[x],[\boldsymbol{\ell}],\overline{x},\bar{\boldsymbol{\ell}},\delta}]\\
&=\frac{1}{|\mathcal{L}|^{\zeta(G)}}\sum_{[x],[\boldsymbol{\ell}],\overline{x'},\bar{\boldsymbol{\ell}}'}P([x])P(\overline{x'})\mathrm{tr}[(\Pi')^{B'}_{[x],[\boldsymbol{\ell}],\overline{x'},\bar{\boldsymbol{\ell}}'}(\rho')^{B'}_{[x],[\boldsymbol{\ell}]}]\\
&\overset{(a)}{\leqslant}\sum_{[x],[\boldsymbol{\ell}],\overline{x'},\bar{\boldsymbol{\ell}}'}\frac{P([x])P(\overline{x'})}{|\mathcal{L}|^{\zeta(G)}}\\
&\quad\times\mathrm{tr}[(\Pi')^{B'}_{[x],[\boldsymbol{\ell}],\overline{x'},\bar{\boldsymbol{\ell}}'}\mathcal{T}_{[X];[\boldsymbol{\ell}]}(\rho^B_{[x]}\otimes|0\rangle\langle0|)]+\frac{4\zeta(G)\delta|B|}{\sqrt{|\mathcal{L}|}}\\
&\overset{(b)}{\leqslant}\sum_{[x],[\boldsymbol{\ell}],\overline{x'},\bar{\boldsymbol{\ell}}'}\frac{P([x])P(\overline{x'})}{|\mathcal{L}|^{\zeta(G)}}\mathrm{tr}[(I-(\Pi')^{B'}_{W'_{[x],[\boldsymbol{\ell}],\overline{x'},\bar{\boldsymbol{\ell}}'}})\\
&\quad\times\mathcal{T}_{[X];[\boldsymbol{\ell}]}(\rho^B_{[x]}\otimes|0\rangle\langle0|)]+\frac{4\zeta(G)\delta|B|}{\sqrt{|\mathcal{L}|}}\\
&\overset{(c)}{\leqslant}\sum_{[x],[\boldsymbol{\ell}],\overline{x'},\bar{\boldsymbol{\ell}}'}\frac{P([x])P(\overline{x'})}{|\mathcal{L}|^{\zeta(G)}}\mathrm{tr}[(I-(\Pi')^{B'}_{W'_{[X];[x],[\boldsymbol{\ell}],\overline{x'},\bar{\boldsymbol{\iota}}'}})\\
&\quad\times\mathcal{T}_{[X];[\boldsymbol{\ell}]}(\rho^B_{[x]}\otimes|0\rangle\langle0|)]+\frac{4\zeta(G)\delta|B|}{\sqrt{|\mathcal{L}|}}
\end{aligned}$$

$$
\begin{aligned}
&\overset{(d)}{=}\frac{1}{|\mathcal{L}|^{\zeta(G)}}\sum_{[x],[\boldsymbol{\ell}],\overline{x}',\bar{\boldsymbol{\ell}}'}P([x])P(\overline{x}')\mathrm{tr}[(I-(\Pi)^{\hat{B}}_{W_{[X];[x],\overline{x}'}})\\
&\quad\times(\rho^B_{[x]}\otimes|0\rangle\langle 0|)]+\frac{4\zeta(G)\delta|B|}{\sqrt{|\mathcal{L}|}}\\
&=\sum_{[x],[\boldsymbol{\ell}],\overline{x}',\bar{\boldsymbol{\ell}}'}\frac{P([x])P(\overline{x}')}{|\mathcal{L}|^{\zeta(G)}}\mathrm{tr}[(\Pi'')^B_{[X];[x],\overline{x}'}(\rho^B_{[x]})]+\frac{4\zeta(G)\delta|B|}{\sqrt{|\mathcal{L}|}}\\
&\overset{(e)}{\leqslant}2\left(2^{-I_H^{\epsilon,\hat{\theta}}(\overline{X};[X],B)}\right),
\end{aligned}
$$

for large enough $\mathcal{L}$, where (a) follows from (19), (20), (b) from using (17), ((c) from using (16), (d) from the fact that $\mathcal{T}_{[X];[\boldsymbol{\ell}]}$ is an isometry, and hence

$$
\begin{aligned}
&\mathrm{tr}[(\Pi')^{B'}_{W'_{[X];[x],[\boldsymbol{\ell}],\overline{x}',\bar{\boldsymbol{\ell}}'}}\mathcal{T}_{[X];[\boldsymbol{\ell}]}(\rho^B_{[x]}\otimes|0\rangle\langle 0|)]\\
&\qquad=\mathrm{tr}[(\Pi)^{\hat{B}}_{W_{[X];[x],\overline{x}'}}(\rho^B_{[x]}\otimes|0\rangle\langle 0|)],
\end{aligned}
$$

and (e) from (13). Combining the three terms we obtain the average probability of error for a fixed u as

$$
\begin{aligned}
\mathbb{E}_C(P(\text{Error}|u))\leqslant 16\zeta(G)\delta^2+\frac{4(1+\zeta(G)\delta^2)|\Theta|\epsilon}{\delta^2}\\
+8\sum_{\hat{\theta}\in\Theta}|T_{\hat{\theta}}|2^{-I_H^{\epsilon,\hat{\theta}}(\overline{X},[X];B)}.
\end{aligned}
$$

for large enough $\mathcal{L}$.

F. Proof for Converse for CQ Channel

Based on Lemma 1, for each group code $\mathbb{C}\leqslant G$, there exists a group J and a homomorphism such that $\mathbb{C}$ is the image of the homomorphism. Assume now that a group transmission system with parameters $(1,|J|,\epsilon)$ exists over a classical-quantum channel $\mathcal{N}=\left\{\rho^B_x\right\}_{x\in\mathcal{X}}$, where $\mathcal{X}=G$ is an Abelian group, and $J=\bigoplus_{p\in\mathcal{P}(G)}\bigoplus_{s=1}^{r_p}\mathbb{Z}_{p^s}^{k_{p,s}}$. Assume that the homomorphism ϕ for the group code $\mathbb{C}$ is a one-to-one mapping.

Let $\hat{\theta}$ be a vector indexed by $(p,s)\in\mathcal{S}(G)$ with $0\leqslant\hat{\theta}_{p,s}\leqslant s$. For an message $u\in J$, construct a one-to-one correspondence between $u_{p,s}\in\mathbb{Z}_{p^s}^{k_{p,s}}$ and the tuple $(\tilde{u}_{p,s},\hat{u}_{p,s})$ where $\tilde{u}_{p,s}\in p^{\hat{\theta}_{p,s}}\mathbb{Z}_{p^s}^{k_{p,s}}$ and $\hat{u}_{p,s}\in\mathbb{Z}_{p^{\hat{\theta}_{p,s}}}^{k_{p,s}}$. Consider the subcode $\mathbb{C}_1(\hat{\theta},\hat{u})$ of $\mathbb{C}$ as defined in the classical channel case. Let $x=\phi(u)+V$ be the channel input and $H_{\hat{\theta}}$ be given as in (7). Then $\mathbb{C}_1(\hat{\theta},\hat{u})=[x]_{\hat{\theta}}=x+H_{\hat{\theta}}$. That is, there is an one-to-one correspondence between $\hat{U}$ and $[X]_{\hat{\theta}}$. Also, $\left|T_{\hat{\theta}}(u)\right|=\left|\mathbb{C}_1(\hat{\theta},\hat{u})\right|=\left|H_{\hat{\theta}}\right|$ for all $u\in J$.

Define a one-to-one correspondence between x and the tuple $(\tilde{x}_{\hat{\theta}},[x]_{\hat{\theta}})$ where $\tilde{x}_{\hat{\theta}}=\phi(\tilde{u})$. Consider a genie-aided receiver which gets access to $\hat{U}$ and denote it by $\mathcal{D}^{ga}$. Equivalently, this receiver has access to the coset information $[X]_{\hat{\theta}}$ of X and can be realized by a family of POVMs $\left\{E_x^{[x]}\right\}$. Clearly the average probability of error for this decoder must be not greater than ϵ. Let $X'\in\mathcal{X}$ be the output of $\mathcal{D}^{ga}$. For every $\hat{\theta}\neq\mathbf{s}$, the average probability of error for this decoder is

$$
\sum_{\hat{u}}\sum_{x,x'}\Pr(\hat{u})P_{XX'|\hat{U}}(x,x'\mid\hat{u})\mathbf{1}_{\{x'\neq x\}}\leqslant\epsilon,
$$

where $P_{XX'|\hat{U}}(x,x'|\hat{u})\triangleq P_{X|[X_r]}(x|[x])\mathrm{tr}\left[E_{x'}^{[x]}\rho_x\right]$.

Note that the decoding POVM can be viewed as a CPM. This CPM maps ρ^{AB} to the (classical) state $P_{XX'}$ denoting the joint distribution of the transmitted codeword X and the decoder's guess X'. Similarly, it maps $\rho^{\bar{A}}\rho^{[A]B}$ to $P_{\hat{U}}(P_{X|\hat{U}}\otimes P_{X'|\hat{U}})$. Hence, it follows from the DPI for $D^{\epsilon}_{\mathrm{H}}(\rho\|\sigma)$ that

$$
D^{\epsilon}_{\mathrm{H}}(P_{XX'}\|P_{\hat{U}}(P_{X|\hat{U}}\otimes P_{X'|\hat{U}}))\leqslant D^{\epsilon}_{\mathrm{H}}(\rho^{AB}\|\rho^{\bar{A}}\rho^{[A]B})\ .
$$

Consider the strategy to distinguish $P_{XX'}$ and $P_{\hat{U}}(P_{X|\hat{U}}\otimes P_{X'|\hat{U}})$ as given in the classical channel case, under which the type-I error probability is not larger than ϵ, and the probability of type-II error (misdetection) is $\frac{1}{|H|}$. That is,

$$
\begin{aligned}
D^{\epsilon}_{\mathrm{H}}(\rho^{AB}\|\rho^{\bar{A}}\rho^{[A]B})&\geqslant D^{\epsilon}_{\mathrm{H}}(P_{XX'}\|P_{\hat{U}}(P_{X|\hat{U}}\otimes P_{X'|\hat{U}}))\\
&=\log_2|H|=\log_2\left|T_{\hat{\theta}}(u)\right|\ .
\end{aligned}
\tag{29}
$$

Using (9), we may rewrite (29) compactly as $I_H^{\epsilon,\hat{\theta}}(\overline{X};[X]Y)\geqslant(1-\omega_{\hat{\theta}})R$.

G. Proofs for Section VI

Proof of Theorem 5:

Achievability– We use the n independent copies of the channel, and make the observation that $\mathcal{S}(G)=\mathcal{S}(G^n)$ for all $n\geqslant 1$. Using this and Theorem 1, we see that there exists a $(n,|J|,\epsilon')$-code such that

$$
\epsilon'\leqslant\epsilon+\sum_{\hat{\theta}\neq\mathbf{s}}\exp_2\left\{(1-\omega_{\hat{\theta}})nR-I_H^{\epsilon_{\hat{\theta}}}(\overline{X}^n_{\hat{\theta}};[X]^n_{\hat{\theta}}Y^n)\right\},
$$

where the rate R is given in Equation (6), and the joint distribution of the input and the output of the channel is given by

$$
P(X^n=x^n,Y^n=y^n)=\prod_{i=1}^{n}P_X(x_i)W_{Y|X}(y_i|x_i),
$$

where $P_X(x)=\frac{1}{|G|}$. Then we have the random vectors will have the following distributions. For $x_r^n=(x_{r,1},x_{r,2},\ldots,x_{r,n})\in G^n$, $[x_r^n]$ denotes the coset representative of $x_r^n+H^n$ in G^n, and the product conditional distribution $P^n_{Y|[X]}$ is defined as

$$
\begin{aligned}
P^n_{Y|[X]}(y^n\mid[x_r^n])&\triangleq\prod_{i=1}^{n}P_{Y|[X]}(y_i\mid[x_{r,i}])\\
&=\sum_{x^n\in[x_r^n]+H^n}P^n_{X|[X]}(x^n|[x_r^n])W^n_{Y|X}(y^n|x^n),
\end{aligned}
$$

where $P_{[X]}$ and $P_{X|[X]}$ are given in Definition 4 and used in the one-shot case.

$$
P^n_{[X]}=\frac{|H|^n}{|G|^n},P^n_{X|[X]}(x^n\mid[x_r^n])=\begin{cases}\frac{1}{|H|^n} & \text{if } x^n\in[x_r]^n,\\ 0 & \text{otherwise.}\end{cases}
$$

Now since all distributions are in a product form, we can use AEP [25] for hypothesis testing relative entropy as: for all $\epsilon>0$, and all $\hat{\theta}$,

$$
\lim_{n\to\infty}\frac{1}{n}I_H^{\epsilon,\hat{\theta}}(\overline{X}^n_{\hat{\theta}};[X]^n_{\hat{\theta}}Y^n)=I(\overline{X}_{\hat{\theta}};[X]_{\hat{\theta}},Y)=I(\overline{X}_{\hat{\theta}};Y|[X]_{\hat{\theta}}).
$$

This gives the desired achievability result. Theorem 5 *Converse–* The converse follows from [15].

Proof of Theorem 6:

The achievability follows by exploiting the product nature of the input distribution and the quantum Stein's lemma [27, Theorem 2]. We skip the details due to lack of space. We skip the proof of the converse due to lack of space. The complete details will be provided in the longer version of the paper. ■