

Abelian Group Codes for Channel Coding and Source Coding

Aria G. Sahebi and S. Sandeep Pradhan

Department of Electrical Engineering and Computer Science,

University of Michigan, Ann Arbor, MI 48109, USA.

Email: ariaghs@umich.edu, pradhanv@umich.edu

Abstract

In this paper, we study the asymptotic performance of Abelian group codes for the the channel coding problem for arbitrary discrete (finite alphabet) memoryless channels as well as the lossy source coding problem for arbitrary discrete (finite alphabet) memoryless sources. For the channel coding problem, we find the capacity characterized in a single-letter information-theoretic form. This simplifies to the symmetric capacity of the channel when the underlying group is a field. For the source coding problem, we derive the achievable rate-distortion function that is characterized in a single-letter information-theoretic form. When the underlying group is a field, it simplifies to the symmetric rate-distortion function. We give several illustrative examples. Due to the non-symmetric nature of the sources and channels considered, our analysis uses a synergy of information-theoretic and group-theoretic tools.

I. INTRODUCTION

Approaching the information-theoretic performance limits of communication using structured codes has been of great interest for the last several decades [1], [7], [14], [15]. The earlier attempts to design computationally efficient encoding and decoding algorithms for point-to-point communication (both channel coding and source coding) resulted in injection of finite field structures to the coding schemes [12]. In the channel coding problem [24], the channel input alphabets are matched to algebraic structure and encoders are represented by matrices. Similarly in source coding problem [18], the reconstruction alphabets are matched to algebraic structure and decoders are represented by matrices. Later these coding

This work was supported by NSF grant CCF-1116021. This work was presented in part at IEEE International Symposium on Information Theory (ISIT), July 2011, and Allerton conference on communication, control and computing, October 2012.

approaches were extended to weaker algebraic structures such as rings and groups [2], [3], [11], [19], [20]¹. The motivation for this is twofold: a) finite fields exist only for alphabets with size equal to a prime power, 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 [10] (also see [3]), and hence the former have superior error correcting properties. 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 [29], [32], [33]. Subsequently, as interest in network information theory grew, these codes were used to approach the information-theoretic performance limits of certain special cases of multi-terminal communication problems [4], [17], [31], [36], [37]. These limits were obtained earlier using the random coding ensembles in the information theory literature.

In 1979, Korner and Marton, in a significant departure from tradition, showed that for a binary distributed source coding problem, the asymptotic average performance of binary linear code ensembles can be superior to that of the standard random coding ensembles. Although, structured codes were being used in communication mainly for computational complexity reasons, the duo showed that, in contrast, even when computational complexity is not an issue, the use of structured codes leads to superior asymptotic performance limits in multi-terminal communication problems. In the recent past, such gains were shown for a wide class of problems [5], [23], [25], [30]. In our prior work, we developed an inner bound to the optimal rate-distortion region for the distributed source coding problem [23] in which cyclic group codes were used as building blocks in the coding schemes. Similar coding approaches were applied for the interference channel and the broadcast channel in [26], [27]. The motivation for studying Abelian group codes beyond the non-existence of finite fields over arbitrary alphabets is the following. The algebraic structure of the code imposes certain restrictions on the performance. For certain communication problems, linear codes were shown to be not optimal [23], and Abelian group codes exhibit a superior performance. For example, consider a distributed source coding problem with two statistically correlated but individually uniform quaternary sources X and Y that are related via the relation $Y = -X + Z$, where $+$ denotes addition modulo-4 and Z is a hidden quaternary random variable that has a non-uniform distribution and is independent of X . The joint decoder wishes to reconstruct Z losslessly. In this problem, random codes over \mathbb{Z}_4 perform better than random linear codes over the Galois field of size

¹Note that this is an incomplete list. There is a vast body of work on group codes. See [12] for a more complete bibliography.

4. In summary, the main reason for using algebraic structured codes in this context is performance rather than complexity of encoding and decoding. Hence information-theoretic characterizations of asymptotic performance of Abelian group code ensembles for various communication problems and under various decoding constraints became important.

Such performance limits have been characterized in certain special cases. It is well-known that binary linear codes achieve the capacity of binary symmetric channels [16]. More generally, it has also been shown that q -ary linear codes can achieve the capacity of symmetric channels [14] and linear codes can be used to compress a source losslessly down to its entropy [22]. Goblick [1] showed that binary linear codes achieve the rate-distortion function of binary uniform sources with Hamming distortion criterion. Group codes were first studied by Slepian [34] for the Gaussian channel. In [6], the capacity of group codes for certain classes of channels has been computed. Further results on the capacity of group codes were established in [7], [8]. The capacity of group codes over a class of channels exhibiting symmetries with respect to the action of a finite Abelian group has been investigated in [11].

In this work, we focus on two problems. In the first part, we consider the channel coding problem for arbitrary discrete memoryless channels. We assume that the channel input alphabet is equipped with the structure of a finite Abelian group G . We provide an information-theoretic characterization of the capacity of such channels achievable using group codes which are cosets of subgroups of G^n , where n denotes the block length of encoding which is arbitrarily large. This performance limit is equal to the symmetric capacity of the channel when the underlying group is a field; i.e., it is equal to the Shannon mutual information between the channel input and the channel output when the channel input is uniformly distributed. In the general case, additional constraints corresponding to subgroups of the underlying group appear in the characterization, and the achievable rate can be smaller than the symmetric capacity of the channel.

In the second, we consider the lossy source coding problem for arbitrary discrete memoryless sources with single-letter distortion measures and the reconstruction alphabet being equipped with the structure of a finite Abelian group G . We provide an information-theoretic characterization of the rate-distortion function achievable using group codes which are cosets of subgroups of G^n . The performance limit is equal to the symmetric rate-distortion function of the source when the underlying group is a field i.e., the Shannon rate-distortion function with the additional constraint that the reconstruction variable is uniformly distributed. For the general case, as in channel coding, additional constraints corresponding to subgroups of the underlying group appear in the characterization, and this can result in a larger rate compared to the symmetric rate for a given distortion level.

We use joint typicality encoding and decoding [13] for both problems at hand, which makes the analysis more tractable. In this approach we use a synergy of information-theoretic and group-theoretic tools. The traditional approaches have looked at encoding and decoding of structured codes based on either minimum distance or maximum likelihood. However, the approach based on joint typicality does not provide insight into error exponents as compared to traditional ones.

The paper is organized as follows: In Section II, some definitions and basic facts are stated which are used in the paper. In Section III, we give two examples of multi-terminal communication problems where the performance of group code ensembles is strictly superior to that of unstructured code ensembles. In Section IV, we introduce the ensemble of Abelian group codes used in the paper. In Section V, we state the main results of the paper for both the source coding problem as well as the channel coding problem. In Section VI, we prove the converse results for both problems and, in Section VII, we prove the achievability results. We conclude in Section VIII.

II. PRELIMINARIES

1) *Channel Model:* We consider discrete memoryless channels used without feedback. We associate two finite sets \mathcal{X} and \mathcal{Y} with the channel as the channel input and output alphabets. The input-output relation of the channel is characterized by a conditional probability law $W_{Y|X}(y|x)$ for $x \in \mathcal{X}$ and $y \in \mathcal{Y}$. The channel is specified by $(\mathcal{X}, \mathcal{Y}, W_{Y|X})$.

2) *Source Model:* The source is modeled as a discrete-time memoryless random process X with each sample taking values from a finite set \mathcal{X} called alphabet according to the distribution P_X . The reconstruction alphabet is denoted by a finite set \mathcal{U} and the quality of reconstruction is measured by a bounded single-letter distortion function $d : \mathcal{X} \times \mathcal{U} \rightarrow \mathbb{R}^+$. We denote this source by $(\mathcal{X}, \mathcal{U}, P_X, d)$.

3) *Groups:* All groups referred to in this paper are finite *Abelian groups*. Given a group $(G, +)$, a subgroup H of G is denoted by $H \leq G$. 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 $a \in G$). For a subgroup H of G , the number of cosets of H in G is called the *index* of H in G and is denoted by $|G : H|$. The index of H in G is equal to $|G|/|H|$ where $|G|$ and $|H|$ are the cardinality or size of G and H respectively. 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 .

4) *Group Codes:* Given a group G , a group code \mathbb{C} over G with block length n is any subgroup of G^n . A shifted group code over G , $\mathbb{C} + B$ is a translation of a group code \mathbb{C} by a fixed vector $B \in G^n$.

Group codes generalize the notion of linear codes over fields to sources with reconstruction alphabets (and channels with input alphabets) having composite sizes.

5) *Achievability for Channel Coding:* For a group G , a group transmission system with parameters (n, Θ, 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 Θ and the mappings are defined as

$$\text{Enc} : \{1, 2, \dots, \Theta\} \rightarrow \mathbb{C}$$

$$\text{Dec} : \mathcal{Y}^n \rightarrow \{1, 2, \dots, \Theta\}$$

such that

$$\max_{1 \leq m \leq \Theta} \sum_{x \in \mathcal{X}^n} \mathbb{1}_{\{x = \text{Enc}(m)\}} \sum_{y \in \mathcal{Y}^n} \mathbb{1}_{\{m \neq \text{Dec}(y)\}} W^n(y|x) \leq \tau$$

Given a channel $(\mathcal{X} = G, \mathcal{Y}, W_{Y|X})$, 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 for reliable communication with parameters (n, Θ, τ) such that

$$\frac{1}{n} \log \Theta \geq R - \epsilon, \quad \tau \leq \epsilon$$

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

6) *Achievability for Source Coding and the Rate-Distortion Function:* For a group G , a group transmission system with parameters $(n, \Theta, \Delta, \tau)$ for compressing a given source $(\mathcal{X}, \mathcal{U} = G, P_X, d)$ 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 Θ and the mappings are defined as

$$\text{Enc} : \mathcal{X}^n \rightarrow \{1, 2, \dots, \Theta\},$$

$$\text{Dec} : \{1, 2, \dots, \Theta\} \rightarrow \mathbb{C}$$

such that $\mathbb{E}[d(X^n, U^n)] \leq \Delta$, where X^n is the random vector of length n generated by the source, and U^n is the reconstruction vector and is given by $U^n = \text{Dec}(\text{Enc}(X^n))$. In this transmission system, n denotes the block length, $\log \Theta$ denotes the number of “channel uses”, Δ denotes the distortion level and the distortion $d(x, \hat{x})$ between two vectors x and \hat{x} is assumed to be the average of the single-letter

distortion between the components x_i and \hat{x}_i . Let Θ_i denote the size of the range of the i th component of the shifted group code. Given a source $(\mathcal{X}, \mathcal{U} = G, P_X, d)$, a pair of non-negative real numbers (R, D) is said to be achievable using group codes if for every $\epsilon > 0$ and for all sufficiently large numbers n , there exists a group transmission system with parameters (n, Θ, Δ) for compressing the source such that

$$\frac{1}{n} \log \Theta \leq R + \epsilon, \quad \Delta \leq D + \epsilon, \quad \frac{1}{n} \sum_{i=1}^n [\log \Theta_i - H(U_i)] \leq \epsilon.$$

The optimal group rate-distortion function $R^*(D)$ of the source is given by the infimum of the rates R such that (R, D) is achievable using group codes.

The rationale for imposing the third constraint regarding the average entropy of single-letter reconstruction samples is the following. Recall that in channel coding all codewords are used by definition. However in source coding, we may have a situation where only a small fraction of the codewords in the group code is used. In which case, one may have to use some form of entropy coding to remove the redundancy. This leads to a different class of codes called nested group codes. We will not study this class in this paper. To prevent this situation, we impose the third constraint. This also puts the source coding problem on a similar footing as compared to the channel coding problem.

7) *Typicality*: We follow the notion of typicality as found in [13]. Consider two random variables X and Y with joint probability mass function $P_{X,Y}(x, y)$ over $\mathcal{X} \times \mathcal{Y}$. Let n be an integer and ϵ be a positive real number. The sequence pair (x^n, y^n) belonging to $\mathcal{X}^n \times \mathcal{Y}^n$ is said to be jointly ϵ -typical with respect to $P_{X,Y}(x, y)$ if

$$\forall a \in \mathcal{X}, \forall b \in \mathcal{Y} : \left| \frac{1}{n} N(a, b | x^n, y^n) - P_{X,Y}(a, b) \right| \leq \frac{\epsilon}{|\mathcal{X}||\mathcal{Y}|}$$

and none of the pairs (a, b) with $P_{X,Y}(a, b) = 0$ occurs in (x^n, y^n) . Here, $N(a, b | x^n, y^n)$ counts the number of occurrences of the pair (a, b) in the sequence pair (x^n, y^n) . We denote the set of all jointly ϵ -typical sequence pairs in $\mathcal{X}^n \times \mathcal{Y}^n$ by $A_\epsilon^n(X, Y)$.

Given a sequence $x^n \in A_\epsilon^n(X)$, the set of ϵ -typical sequences $A_\epsilon^n(Y | x^n)$ is defined as

$$A_\epsilon^n(Y | x^n) = \{y^n \in \mathcal{Y}^n | (x^n, y^n) \in A_\epsilon^n(X, Y)\}$$

8) *Notation*: In our notation, $O(\epsilon)$ is any function of ϵ such that $\lim_{\epsilon \downarrow 0} O(\epsilon) = 0$, \mathbb{P} is the set of all primes, \mathbb{Z}^+ is the set of positive integers and \mathbb{R}^+ is the set of non-negative reals. Let $|x|^+$ denote $\max\{x, 0\}$. Since we deal with summations over several groups in this paper, when not clear from the context, we indicate the underlying group in each summation; e.g. summation over the group G is denoted by $\sum^{(G)}$. Direct sum of groups is denoted by \oplus and direct product of sets is denoted by \otimes .

III. MOTIVATING EXAMPLES

In this section, we present examples of the use of group codes for two multi-terminal communication problems, namely the distributed source coding problem and the interference channel coding problem. In these problems, one can use group code ensembles to improve upon the asymptotic performance of the standard random coding ensembles used in information theory.

A. Example in Distributed Source Coding

Consider a two-user distributed source coding problem in which the two sources X and Y take values from \mathbb{Z}_4 and a centralized decoder is interested in decoding the sum of the two sources losslessly. The two sources need to be compressed distributively. Furthermore, assume that X is uniformly distributed over \mathbb{Z}_4 and $Y = -X + Z$ where Z is independent from X and is distributed over \mathbb{Z}_4 such that $P_Z(0) = 1 - \tau$ and $P_Z(1) = P_Z(2) = P_Z(3) = \frac{\tau}{3}$ for some $\tau \in (0, 1)$. Let R_1 and R_2 be the rates of the two encoders. Using Slepian-Wolf coding [35], which uses random unstructured codes, one can show that the the following sum rate is achievable

$$R_U = R_1 + R_2 = H(X, Y) = H(X, -X + Z) = H(X) + H(Z) = 2 + h(\tau) + \tau \log 3,$$

where $h(\cdot)$ denotes the binary entropy function. We present a coding scheme based on group codes that is optimal in the sum rate and strictly improves upon the rate achievable using random unstructured codes.

Consider a fictitious discrete memoryless channel with input U and output V related via $V = U + Z$, where Z is independent of U . Let \mathbb{C} be a good group channel code for the channel $P_{V|U}$. Using the channel coding result in this paper (see Section VII), the rate of \mathbb{C} can be arbitrarily close to

$$r = \min \left(I(U; V), 2I(U; V|[U]) \right) = \min \left(2 - H(U|V), 2 - 2H(U|[U]V) \right)$$

where for $g \in \mathbb{Z}_4$, $[g] = g + \{0, 2\}$. Note that $H(U|V) = H(Z)$ and

$$H(U|[U]V) \stackrel{(a)}{=} H(V - Z|[Z]V) \stackrel{(b)}{=} H(Z|[Z])$$

where (a) follows since there is a one-to-one correspondence between $([V - Z], V)$ and $([Z], V)$ and (b) follows since V and Z are independent. Therefore, we have

$$r = \min \left(2 - H(Z), 2 - 2H(Z|[Z]) \right) \stackrel{(a)}{=} 2 - H(Z) = 2 - h(\tau) - \tau \log(3),$$

where in (a) we have used the relation, $h(\tau) + \tau \log_2(3) \leq 2h(2\tau/3)$, to show that $I(U; V) < 2I(U; V|[U])$.

The encoding scheme for the distributed source coding problem is as follows: Given a pair of source sequences x^n , and y^n , the X -, Y -encoder send $x^n + \mathbb{C}$, and $y^n + \mathbb{C}$, respectively, to the decoder. In other

words each encoder sends the index of the coset of \mathbb{C} that contains its source word. This implies that, because of the *closure of* \mathbb{C} with respect to group addition, and the commutativity of the group, the decoder has $z^n + \mathbb{C}$, from which it can recover z^n with high probability using the property of the code. Therefore, using group codes, the following sum rate is achievable:

$$R_G = R_1 + R_2 = 2(2 - r) = 2 \max \left(H(Z), 2H(Z|[Z]) \right) = 2H(Z) = 2h(\tau) + 2\tau \log(3) < R_U.$$

It can be shown that this rate is not achievable using random linear codes over Galois field of size 4. The optimality follows from the standard information-theoretic arguments. The algebraic property of the code–closure with respect to addition modulo-4– is exploited in reducing the sum rate from $H(X, Y)$ to $2H(Z)$.

Consider another example with $P_Z(0) = 0.4014$, $P_Z(1) = 0.2035$, $P_Z(2) = 0.3356$ and $P_Z(3) = 0.0595$. Here $H(Z) = 1.7669$, $2H(Z|[Z]) = 1.8712$. The achievable rate for distributed compression using random unstructured codes or random linear codes is $R_U = 3.7669$ and that using random group codes is $R_G = 3.7424$. We do not claim optimality for this example. Although random group codes over \mathbb{Z}_4 are inferior to random unstructured codes in terms of point-to-point compression of the source Z , in the problem of distributed compression of X and Y , they outperform the latter.

B. Example for the Interference Channel

Consider the problem of communication over the following interference channel between three pairs of encoders and decoders. The 3-user interference channel [28, Example 7] has three inputs X_1 , X_2 and X_3 which take values from \mathbb{Z}_4 and has three outputs, which are given by $Y_1 = X_1 + X_2 + X_3 + N_1$, $Y_2 = X_2 + N_2$ and $Y_3 = X_3 + N_3$ where the additions are mod-4 operations and N_1 , N_2 and N_3 are independent random variables distributed according to $P_{N_i}(0) = 1 - \delta_i$, $P_{N_i}(1) = P_{N_i}(2) = P_{N_i}(3) = \frac{\delta_i}{3}$ for $i = 1, 2, 3$. For simplicity, we assume $\delta_2 = \delta_3 = \delta$. The input X_1 is costed according to $w(0) = 0$, $w(1) = w(2) = w(3) = 1$ but X_2 and X_3 are not costed. Let τ be the cost constraint on X_1 . Let us assume for simplicity $\delta_1, \delta < \frac{1}{4}$ and $\tau < \frac{3}{4}$. Observe that the channel inputs of the second and the third transmitters interfere with that of the first. There is no interference for the second and third receivers.

Consider the following optimal coding scheme based on group codes. Let \mathbb{C} be a good group channel code for the channel with input X_2 and output Y_2 . The same code is employed for communication between X_3 and Y_3 . Using the arguments used in distributed source coding, it follows that the following rates are achievable for transmitters 2 and 3:

$$R_2 = R_3 = 2 - h(\delta) - \delta \log(3).$$

Note that $X_2 + X_3$ interferes with X_1 at Y_1 , and the decoder wishes to decode the interference first. Using the *closure of \mathbb{C}* with respect to group addition, note that the rate of $X_2^n + X_3^n$ is $2 - h(\delta) - \delta \log(3)$. The interference can be decoded reliably if the effective noise given by $X_1 + N_1$ satisfies: $P(X_1 + N_1 \neq 0) \leq P(N_2 \neq 0) = P(N_3 \neq 0)$. This condition implies that $\delta_1 + \tau - \frac{4\delta_1\tau}{3} \leq \delta$. After decoding the interference, the effective channel becomes $(Y_1 - X_2 - X_3) = X_1 + N_1$, and the first transmitter can communicate at the following rate, given by the capacity of this effective channel,

$$R_1 = C^* \triangleq \sup_{P_{X_1}: \mathbb{E}w(X_1) \leq \tau} I(X_1; Y_1 | X_2 + X_3).$$

In summary, the rate triple $(C^*, 2 - h(\delta) - \delta \log(3), 2 - h(\delta) - \delta \log(3))$ is achievable using group codes. It can be shown that [28, Lemma 8] if

$$C^* + 2(2 - h(\delta) - \delta \log 3) > 2 - h(\delta_1) - \delta_1 \log 3,$$

then the above triple of rates cannot be achieved using either random unstructured codes or random linear codes over the Galois field of size 4. The following example of δ_1 , τ and δ satisfies all the conditions: $\delta_1 = \tau = \frac{3}{4} - \frac{\sqrt{30}}{8}$, $\delta = \frac{1}{8}$. Note that group codes outperform random codes in this case because there is a match between the structure of the group code and the structure of the channel.

IV. 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 the 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 [21, Theorem 3.3.1] that any Abelian group G can be decomposed as a direct sum of its Sylow subgroups in the following manner

$$G = \bigoplus_{p \in \mathcal{P}(G)} S_p(G) \tag{1}$$

Furthermore, each Sylow subgroup $S_p(G)$ can be decomposed into \mathbb{Z}_{p^r} groups as follows:

$$S_p(G) \cong \bigoplus_{r \in \mathcal{R}_p(G)} \mathbb{Z}_{p^r}^{M_{p,r}} \tag{2}$$

where $\mathcal{R}_p(G) \subseteq \mathbb{Z}^+$ and for $r \in \mathcal{R}_p(G)$, $M_{p,r}$ is a positive integer. Note that $\mathbb{Z}_{p^r}^{M_{p,r}}$ is defined as the direct sum of the ring \mathbb{Z}_{p^r} with itself for $M_{p,r}$ times. Combining Equations (1) and (2), we can represent any Abelian group as follows:

$$G \cong \bigoplus_{p \in \mathcal{P}(G)} \bigoplus_{r \in \mathcal{R}_p(G)} \mathbb{Z}_{p^r}^{M_{p,r}} = \bigoplus_{p \in \mathcal{P}(G)} \bigoplus_{r \in \mathcal{R}_p(G)} \bigoplus_{m=1}^{M_{p,r}} \mathbb{Z}_{p^r}^{(m)} \quad (3)$$

where $\mathbb{Z}_{p^r}^{(m)}$ is called the m^{th} \mathbb{Z}_{p^r} ring of G or the $(p, r, m)^{\text{th}}$ ring of G . Equivalently, this can be written as follows

$$G \cong \bigoplus_{(p,r) \in \mathcal{Q}(G)} \mathbb{Z}_{p^r}^{M_{p,r}} = \bigoplus_{(p,r,m) \in \mathcal{G}(G)} \mathbb{Z}_{p^r}^{(m)}$$

where $\mathcal{Q}(G) \subseteq \mathbb{P} \times \mathbb{Z}^+$ is defined as:

$$\mathcal{Q}(G) = \{(p, r) | p \in \mathcal{P}(G), r \in \mathcal{R}_p(G)\}, \quad (4)$$

and $\mathcal{G}(G) \subseteq \mathbb{P} \times \mathbb{Z}^+ \times \mathbb{Z}^+$ is defined as:

$$\mathcal{G}(G) = \{(p, r, m) \in \mathbb{P} \times \mathbb{Z}^+ \times \mathbb{Z}^+ | (p, r) \in \mathcal{Q}(G), m \in \{1, 2, \dots, M_{p,r}\}\}$$

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

With a slight abuse of notation, we represent an element a of G as

$$a = \bigoplus_{(p,r,m) \in \mathcal{G}(G)} a_{p,r,m}$$

Furthermore, for two elements $a, b \in G$, we have

$$a + b = \bigoplus_{(p,r,m) \in \mathcal{G}(G)} a_{p,r,m} +_{p^r} b_{p,r,m}$$

where $+$ denotes the group operation and $+_{p^r}$ denotes addition mod- p^r . More generally, let a, b, \dots, z be any number of elements of G . Then we have

$$a + b + \dots + z = \bigoplus_{(p,r,m) \in \mathcal{G}(G)} (a_{p,r,m} +_{p^r} b_{p,r,m} +_{p^r} \dots +_{p^r} z_{p,r,m}) \quad (5)$$

This can equivalently be written as

$$[a + b + \dots + z]_{p,r,m} = a_{p,r,m} +_{p^r} b_{p,r,m} +_{p^r} \dots +_{p^r} z_{p,r,m}$$

where $[\cdot]_{p,r,m}$ denotes the $(p, r, m)^{\text{th}}$ component of it's argument.

Let $\mathbb{I}_{G:p,r,m} \in G$ be a generator for the group which is isomorphic to the $(p, r, m)^{\text{th}}$ ring of G . Then we have

$$a = \sum_{(p,r,m) \in \mathcal{G}(G)}^{(G)} a_{p,r,m} \mathbb{I}_{G:p,r,m} \quad (6)$$

where the summations are done with respect to the group operation and the multiplication $a_{p,r,m} \mathbb{I}_{G:p,r,m}$ is by definition the summation (with respect to the group operation) of $\mathbb{I}_{G:p,r,m}$ to itself for $a_{p,r,m}$ times.

In other words, $a_{p,r,m} \mathbb{I}_{G:p,r,m}$ is the short hand notation for

$$a_{p,r,m} \mathbb{I}_{G:p,r,m} = \sum_{i \in \{1, \dots, a_{p,r,m}\}}^{(G)} \mathbb{I}_{G:p,r,m}$$

where the summation is the group operation.

Example: 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 $a_{3,2,1}, a_{3,2,2} \in \mathbb{Z}_9$. Finally, we have $\mathbb{I}_{G:2,2,1} = (1, 0, 0, 0)$, $\mathbb{I}_{G:3,1,1} = (0, 1, 0, 0)$, $\mathbb{I}_{G:3,2,1} = (0, 0, 1, 0)$, $\mathbb{I}_{G:3,2,2} = (0, 0, 0, 1)$ so that Equation (6) holds.

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 subgroups of G^n . We use the following fact to characterize all subgroups of G^n :

Lemma IV.1. *For an Abelian group \tilde{G} , let $\phi : J \rightarrow \tilde{G}$ be a homomorphism from some Abelian group J to \tilde{G} . Then $\phi(J) \leq \tilde{G}$; i.e., the image of the homomorphism is a subgroup of \tilde{G} . Moreover, for any subgroup \tilde{H} of \tilde{G} there exists a corresponding Abelian group J and a homomorphism $\phi : J \rightarrow \tilde{G}$ such that $\tilde{H} = \phi(J)$.*

Proof: The first part of the lemma is proved in [9, Theorem 12-1]. For the second part, Let J be isomorphic to \tilde{H} and let ϕ be the identity mapping (more rigorously, let ϕ be the isomorphism between J and \tilde{H}). ■

In order to use the above lemma to construct the ensemble of subgroups of G^n , we need to identify all groups J from which there exist non-trivial homomorphisms to G^n . Then the above lemma implies that for each such J and for each homomorphism $\phi : J \rightarrow G^n$, the image of the homomorphism is a group code over G of length n and for each group code $\mathbb{C} \leq G^n$, there exists a group J and a homomorphism such that \mathbb{C} is the image of the homomorphism. This ensemble corresponds to the ensemble of linear codes characterized by their generator matrix when the underlying group is a field of prime size. Note that as in the case of standard ensembles of linear codes, the correspondence between this ensemble and the set of Abelian group codes over G of length n may not be one-to-one.

Let \tilde{G} and J be two Abelian groups with decompositions:

$$\begin{aligned}\tilde{G} &= \bigoplus_{(p,r,m) \in \mathcal{G}(\tilde{G})} \mathbb{Z}_{p^r}^{(m)} \\ J &= \bigoplus_{(q,s,l) \in \mathcal{G}(J)} \mathbb{Z}_{q^s}^{(l)}\end{aligned}$$

and let ϕ be a homomorphism from J to \tilde{G} . For $(q, s, l) \in \mathcal{G}(J)$ and $(p, r, m) \in \mathcal{G}(\tilde{G})$, let

$$g_{(q,s,l) \rightarrow (p,r,m)} = [\phi(\mathbb{I}_{J;q,s,l})]_{p,r,m}$$

where $\mathbb{I}_{J;q,s,l} \in J$ is the standard generator for the $(q, s, l)^{\text{th}}$ ring of J and $[\phi(\mathbb{I}_{J;q,s,l})]_{p,r,m}$ is the $(p, r, m)^{\text{th}}$ component of $\phi(\mathbb{I}_{J;q,s,l}) \in \tilde{G}$. For $a = \bigoplus_{(q,s,l) \in \mathcal{G}(J)} a_{q,s,l} \in J$, let $b = \phi(a)$ and write $b = \bigoplus_{(p,r,m) \in \mathcal{G}(\tilde{G})} b_{p,r,m}$. Note that as in Equation (6), we can write:

$$\begin{aligned}a &= \underbrace{\sum_{(q,s,l) \in \mathcal{G}(J)}^{(J)}}_{(q,s,l) \in \mathcal{G}(J)} a_{q,s,l} \mathbb{I}_{J;q,s,l} \\ &= \underbrace{\sum_{(q,s,l) \in \mathcal{G}(J)}^{(J)}}_{(q,s,l) \in \mathcal{G}(J)} \underbrace{\sum_{i \in \{1, \dots, a_{q,s,l}\}}^{(J)}}_{i \in \{1, \dots, a_{q,s,l}\}} \mathbb{I}_{J;q,s,l}\end{aligned}$$

where the summations are the group summations. We have

$$\begin{aligned}
b_{p,r,m} &= [\phi(a)]_{p,r,m} \\
&= \left[\phi \left(\sum_{(q,s,l) \in \mathcal{G}(J)} \sum_{i \in \{1, \dots, a_{q,s,l}\}} \mathbb{I}_{J:q,s,l} \right) \right]_{p,r,m} \\
&\stackrel{(a)}{=} \left[\sum_{(q,s,l) \in \mathcal{G}(J)} \sum_{i \in \{1, \dots, a_{q,s,l}\}} \phi(\mathbb{I}_{J:q,s,l}) \right]_{p,r,m} \\
&\stackrel{(b)}{=} \sum_{(q,s,l) \in \mathcal{G}(J)} \sum_{i \in \{1, \dots, a_{q,s,l}\}} [\phi(\mathbb{I}_{J:q,s,l})]_{p,r,m} \\
&\stackrel{(c)}{=} \sum_{(q,s,l) \in \mathcal{G}(J)} a_{q,s,l} [\phi(\mathbb{I}_{J:q,s,l})]_{p,r,m} \\
&= \sum_{(q,s,l) \in \mathcal{G}(J)} a_{q,s,l} g_{(q,s,l) \rightarrow (p,r,m)}
\end{aligned}$$

Note that (a) follows since ϕ is a homomorphism; (b) follows from Equation (5); and (c) follows by using $a_{q,s,l} [\phi(\mathbb{I}_{J:q,s,l})]_{p,r,m}$ as the short hand notation for the summation of $[\phi(\mathbb{I}_{J:q,s,l})]_{p,r,m}$ to itself for $a_{q,s,l}$ times.

Note that $g_{(q,s,l) \rightarrow (p,r,m)}$ represents the effect of the $(q, s, l)^{\text{th}}$ component of a on the $(p, r, m)^{\text{th}}$ component of b dictated by the homomorphism. This means that the homomorphism ϕ can be represented by

$$\phi(a) = \bigoplus_{(p,r,m) \in \mathcal{G}(\tilde{G})} \sum_{(q,s,l) \in \mathcal{G}(J)} a_{q,s,l} g_{(q,s,l) \rightarrow (p,r,m)} \quad (7)$$

where $a_{q,s,l} g_{(q,s,l) \rightarrow (p,r,m)}$ is the short-hand notation for the mod- p^r addition of $g_{(q,s,l) \rightarrow (p,r,m)}$ to itself for $a_{q,s,l}$ times. We have the following lemma on $g_{(q,s,l) \rightarrow (p,r,m)}$:

Lemma IV.2. *For a homomorphism described by (7), we have*

$$\begin{aligned}
g_{(q,s,l) \rightarrow (p,r,m)} &= 0 && \text{If } p \neq q \\
g_{(q,s,l) \rightarrow (p,r,m)} &\in p^{r-s} \mathbb{Z}_{p^r} && \text{If } p = q, r \geq s
\end{aligned}$$

Moreover, any mapping described by (7) and satisfying these conditions is a homomorphism.

Proof: The proof is provided in Appendix IX-A. ■

This lemma implies that in order to construct a subgroup of \tilde{G} , we only need to consider homomorphisms from an Abelian group J to \tilde{G} such that

$$\mathcal{P}(J) \subseteq \mathcal{P}(\tilde{G})$$

since if for some $(q, s, l) \in \mathcal{G}(J)$, $q \notin \mathcal{P}(\tilde{G})$ then $\phi(a)$ would not depend on $a_{q,s,l}$. For $p \in \mathcal{P}(\tilde{G})$, define

$$r_p = \max \mathcal{R}_p(G) \quad (8)$$

We show that we can restrict ourselves to J 's such that for all $(q, s, l) \in \mathcal{G}(J)$, $s \leq r_q$. Let $(p, r, m) \in \mathcal{G}(\tilde{G})$ be such that $p = q$. Since $g_{(q,s,l) \rightarrow (p,r,m)} \in \mathbb{Z}_{p^r}$ and $r \leq r_q$, we have

$$\begin{aligned} (a_{q,s,l} g_{(q,s,l) \rightarrow (p,r,m)}) \pmod{p^r} &= ((a_{q,s,l}) \pmod{p^r}) g_{(q,s,l) \rightarrow (p,r,m)} \pmod{p^r} \\ &= ((a_{q,s,l}) \pmod{p^{r_q}}) g_{(q,s,l) \rightarrow (p,r,m)} \pmod{p^r} \end{aligned}$$

This implies that for all $a \in J$ and all $(q, s, l) \in \mathcal{G}(J)$, in the expression for the $(p, r, m)^{\text{th}}$ component of $\phi(a)$ with $p = q$, $a_{q,s,l}$ appears as $(a_{q,s,l}) \pmod{p^{r_q}}$. Therefore, it suffices for $a_{q,s,l}$ to take values from $\mathbb{Z}_{p^{r_q}}$ and this happens if $s \leq r_q$.

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

$$G^n \cong \bigoplus_{p \in \mathcal{P}(G)} \bigoplus_{r \in \mathcal{R}_p} \mathbb{Z}_{p^r}^{nM_{p,r}} = \bigoplus_{p \in \mathcal{P}(G)} \bigoplus_{r \in \mathcal{R}_p} \bigoplus_{m=1}^{nM_{p,r}} \mathbb{Z}_{p^r}^{(m)} = \bigoplus_{(p,r,m) \in \mathcal{G}(G^n)} \mathbb{Z}_{p^r}^{(m)} \quad (9)$$

Define J as

$$J = \bigoplus_{q \in \mathcal{P}(G)} \bigoplus_{s=1}^{r_q} \mathbb{Z}_{q^s}^{k_{q,s}} = \bigoplus_{q \in \mathcal{P}(G)} \bigoplus_{s=1}^{r_q} \bigoplus_{l=1}^{k_{q,s}} \mathbb{Z}_{q^s}^{(l)} = \bigoplus_{(q,s,l) \in \mathcal{G}(J)} \mathbb{Z}_{q^s}^{(l)} \quad (10)$$

for some positive integers $k_{q,s}$.

Example: Let $G = \mathbb{Z}_8 \oplus \mathbb{Z}_9 \oplus \mathbb{Z}_5$. Then we have

$$J = \mathbb{Z}_2^{k_{2,1}} \oplus \mathbb{Z}_4^{k_{2,2}} \oplus \mathbb{Z}_8^{k_{2,3}} \oplus \mathbb{Z}_3^{k_{3,1}} \oplus \mathbb{Z}_9^{k_{3,2}} \oplus \mathbb{Z}_5^{k_{5,1}}$$

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 $q \in \mathcal{P}(G)$ and $s = 1, \dots, r_q$ so that we can write

$$J = \bigoplus_{q \in \mathcal{P}(G)} \bigoplus_{s=1}^{r_q} \bigoplus_{l=1}^{k w_{q,s}} \mathbb{Z}_{q^s}^{(l)} \quad (11)$$

for some constants $w_{q,s}$ adding up to one. Note that

$$\mathcal{G}(J) = \{(q, s, l) : q \in \mathcal{P}(G), 1 \leq s \leq r_q, 1 \leq l \leq kw_{q,s}\}.$$

Define

$$\mathcal{S}(G) = \{(p, s) | p \in \mathcal{P}(G), 1 \leq s \leq r_p\}. \quad (12)$$

Note that $\mathcal{S}(J) = \mathcal{Q}(J) = \mathcal{S}(G)$.

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

$$\phi(a) = \bigoplus_{(p,r,m) \in \mathcal{G}(G^n)} \sum_{(q,s,l) \in \mathcal{G}(J)}^{(\mathbb{Z}_{p^r})} a_{q,s,l} g_{(q,s,l) \rightarrow (p,r,m)} \quad (13)$$

for $a \in J$ where $g_{(q,s,l) \rightarrow (p,r,m)} = 0$ if $p \neq q$, $g_{(q,s,l) \rightarrow (p,r,m)}$ is a uniform random variable over \mathbb{Z}_{p^r} if $p = q, r \leq s$, and $g_{(q,s,l) \rightarrow (p,r,m)}$ is a uniform random variable over $p^{r-s}\mathbb{Z}_{p^r}$ if $p = q, r \geq s$. The corresponding shifted group code with parameters (n, k, w) is defined by

$$\mathbb{C} = \{\phi(a) + B | a \in J\} \quad (14)$$

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

$$R = \frac{1}{n} \log |J| = \frac{k}{n} \sum_{q \in \mathcal{P}(G)} \sum_{s=1}^{r_q} sw_{q,s} \log q \quad (15)$$

Remark IV.3. *An alternate approach to constructing Abelian group codes is to consider kernels of homomorphisms (the kernel ensemble). To construct the ensemble of Abelian group codes in this manner, let ϕ be a homomorphism from J into G^n such that for $a \in G^n$,*

$$\phi(a) = \bigoplus_{(q,s,l) \in \mathcal{G}(J)} \sum_{(p,r,m) \in \mathcal{G}(G^n)}^{(\mathbb{Z}_{q^s})} a_{p,r,m} g_{(p,r,m) \rightarrow (q,s,l)}$$

where $g_{(p,r,m) \rightarrow (q,s,l)} = 0$ if $q \neq p$, $g_{(p,r,m) \rightarrow (q,s,l)}$ is a uniform random variable over \mathbb{Z}_{q^s} if $q = p, s \leq r$, and $g_{(p,r,m) \rightarrow (q,s,l)}$ is a uniform random variable over $p^{s-r}\mathbb{Z}_{q^s}$ if $q = p, s \geq r$. The code is given by $\mathbb{C} = \{a \in G^n | \phi(a) = c\}$ where c is a uniform random variable over J .

In this paper, we use the image ensemble for both the channel and the source coding problem; however, similar results can be derived using the kernel ensemble as well.

V. MAIN RESULTS

In this section, we provide an information-theoretic characterization of the optimal rate-distortion function of a given source and the capacity of a given channel using group codes when the underlying group is an arbitrary finite Abelian group represented by Equation (3). We define two subgroups of G and then state two theorems using these subgroups, and finally provide an interpretation of the results and these subgroups with two examples.

A. Definitions

Consider four vectors $\hat{\theta}$, w , η and b . The components $\hat{\theta}_{p,s}$ and $w_{p,s}$ of $\hat{\theta}$ and w , respectively, are indexed by $(p, s) \in \mathcal{S}(G)$, and satisfy: $0 \leq \hat{\theta}_{p,s} \leq s$,

$$\sum_{(p,s) \in \mathcal{S}(G)} w_{p,s} = 1, \quad w_{p,s} \geq 0.$$

Let $\mathbf{0}$ denote the all-zero vector, and \mathbf{s} denote the vector whose components satisfy $\mathbf{s}_{p,s} = s$ for all $(p, s) \in \mathcal{S}(G)$.

The components $\eta_{p,r,m,s}$ of η are indexed by (p, r, m, s) , for every $(p, r, m) \in \mathcal{G}(G)$ and every $s \in \{1, \dots, r_p\}$, and satisfy $0 \leq \eta_{p,r,m,s} \leq r - |r - s|^+$. The components $b_{p,r,m}$ of b are indexed by $(p, r, m) \in \mathcal{P}(G)$ and satisfy $b_{p,r,m} \in \mathbb{Z}_{p^r}$. Let α be a probability distribution on the set of all η and b , where $\alpha_{\eta,b}$ denote the probability assigned to a particular η and b .

For $\hat{\theta}$, and η , define

$$\boldsymbol{\theta}(\eta) = \left(\min_{\substack{1 \leq s \leq r_p \\ w_{p,s} \neq 0}} |r - s|^+ + \eta_{p,r,m,s} \right)_{(p,r,m) \in \mathcal{P}(G)}.$$

Define $H_\eta \leq G$ and $H_{\eta+\hat{\theta}} \leq H_\eta$ as

$$H_\eta = \bigoplus_{(p,r,m) \in \mathcal{G}(G)} p^{\boldsymbol{\theta}(\eta)_{p,r,m}} \mathbb{Z}_{p^r}^{(m)} \quad (16)$$

$$H_{\eta+\hat{\theta}} = \bigoplus_{(p,r,m) \in \mathcal{G}(G)} p^{\boldsymbol{\theta}(\eta+\hat{\theta})_{p,r,m}} \mathbb{Z}_{p^r}^{(m)}, \quad (17)$$

where $\eta + \hat{\theta}$ denote a vector obtained by adding the components with index (p, s) in $\mathcal{S}(G)$.

For a given $\hat{\theta}$ and w , define

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

B. Main Results

The following theorem is the first main result of this paper.

Consider a channel $(\mathcal{X} = G, \mathcal{Y}, W_{Y|X})$. For every η and b , let $X_{\eta,b}$ be a random variable distributed uniformly over $H_\eta + b$. Let $[X_{\eta,b}]_{\hat{\theta}} = X_{\eta,b} + H_{\eta+\hat{\theta}}$ which takes values from the cosets of $H_{\eta+\hat{\theta}}$ in $H_\eta + b$.

Theorem V.1. For a channel $(\mathcal{X} = G, \mathcal{Y}, W_{Y|X})$, the group capacity is given by

$$C = \sup_{\alpha, w} \min_{\hat{\theta} \neq \mathbf{s}} \frac{1}{1 - \omega_{\hat{\theta}}} \sum_{\eta, b} \alpha_{\eta, b} I(X_{\eta, b}; Y | [X_{\eta, b}]_{\hat{\theta}}). \quad (18)$$

Proof: The proof is provided in Section VII-A2. ■

The following theorem is the second main result of this paper.

Consider a source $(\mathcal{X}, \mathcal{U} = G, P_X, d)$, and distortion level D . For every η and b , let $U_{\eta, b}$ be a reconstruction random variable distributed uniformly over $H_\eta + b$ and is statistically correlated with the source X . Let us denote this collection of random variables as \mathbf{U} . Let $[U_{\eta, b}]_{\hat{\theta}} = U_{\eta, b} + H_{\eta+\hat{\theta}}$.

Theorem V.2. For a source $(\mathcal{X}, \mathcal{U} = G, P_X, d)$, group rate-distortion function is given by

$$R^*(D) = \inf_{\mathbf{U}} \inf_{\alpha, w} \max_{\hat{\theta} \neq \mathbf{0}} \frac{1}{\omega_{\hat{\theta}}} \sum_{\eta, b} \alpha_{\eta, b} I([U_{\eta, b}]_{\hat{\theta}}; X), \quad (19)$$

where infimum is over all α , w and \mathbf{U} , such that

$$D \geq \sum_{\eta, b} \alpha_{\eta, b} \mathbb{E}[d(X, U_{\eta, b})]$$

Proof: The proof is provided in Section VII-B2. ■

C. Interpretation of the Result

In this section, we give some intuition about the result and the quantities defined above using several examples. At a high level, $w_{p,s}$ characterizes the normalized weight given to the \mathbb{Z}_{p^s} component of the input group J in constructing the homomorphism from J to G^n , and $\hat{\theta}$ indexes subgroups of J . η characterizes the the collection of input distributions used on the channel. $\frac{1}{(1-\omega_{\hat{\theta}})} I(X_{\eta, b}; Y | [X_{\eta, b}]_{\hat{\theta}})$ in channel coding and $\frac{1}{\omega_{\hat{\theta}}} I([U_{\eta, b}]_{\hat{\theta}}; X)$ in source coding denote the rate constraints imposed by the subgroup $H_{\eta+\hat{\theta}}$. Due to the algebraic structure of the code, in the ensemble, two random codewords corresponding to two distinct indexes are statistically dependent, unless G is a finite field. For the channel coding problem, when a random codeword corresponding to a given message index is transmitted over the channel, consider the event that all components of the difference between the codeword transmitted and a random codeword corresponding to another message index belong to a proper subgroup $H_{\eta+\hat{\theta}}$ of G .

Then the probability that the latter is decoded instead of the former is higher than the case when no algebraic structure on the code is enforced. For the source coding problem, when the code is chosen randomly, consider the event that all components of their difference belong to a proper subgroup $H_{\eta+\hat{\theta}}$ of G . Then if one of them is a poor representation of a given source sequence, so is the other with a probability that is higher than the case when no algebraic structure on the code is enforced. This means that the code size has to be larger so that with high probability one can find a good representation of the source.

Example: We start with the simple example where $G = \mathbb{Z}_8$. In this case, we have $\mathcal{P}(G) = \{2\}$, $r_2 = 3$, $\mathcal{S}(G) = \{(2, 1), (2, 2), (2, 3)\}$, and $\mathcal{Q}(G) = \{(2, 3)\}$. Let η be the all-zero vector and $b = 0$. For vectors w , $\hat{\theta}$ and θ defined as above, we have $w = (w_{2,1}, w_{2,2}, w_{2,3})$, $\hat{\theta} = (\hat{\theta}_{2,1}, \hat{\theta}_{2,2}, \hat{\theta}_{2,3})$ and $\theta = \theta_{2,3,1}$. Recall that the ensemble of Abelian group codes used in the random coding argument consists of the set of all homomorphisms from some $J = \mathbb{Z}_2^{kw_{2,1}} \oplus \mathbb{Z}_4^{kw_{2,2}} \oplus \mathbb{Z}_8^{kw_{2,3}}$, and hence the vector of weights w determines the input group of the homomorphism. Any vector $\hat{\theta} = (\hat{\theta}_{2,1}, \hat{\theta}_{2,2}, \hat{\theta}_{2,3})$ with $0 \leq \hat{\theta}_{2,1} \leq 1$, $0 \leq \hat{\theta}_{2,2} \leq 2$ and $0 \leq \hat{\theta}_{2,3} \leq 3$ corresponds to a subgroup $K_{\hat{\theta}}$ of the input group J given by

$$K_{\hat{\theta}} = 2^{\hat{\theta}_{2,1}} \mathbb{Z}_2^{kw_{2,1}} \oplus 2^{\hat{\theta}_{2,2}} \mathbb{Z}_4^{kw_{2,2}} \oplus 2^{\hat{\theta}_{2,3}} \mathbb{Z}_8^{kw_{2,3}}$$

Similarly, any $\theta(\eta + \hat{\theta}) = \theta_{2,3,1}$ corresponds to a subgroup $H_{\eta+\hat{\theta}} = 2^{\theta_{2,3,1}} \mathbb{Z}_8$ of the group G .

Example: Next, we consider the case where $G = \mathbb{Z}_4 \oplus \mathbb{Z}_3$. In this case, we have $\mathcal{P}(G) = \{2, 3\}$, $r_2 = 2$, $r_3 = 1$, $\mathcal{S}(G) = \{(2, 1), (2, 2), (3, 1)\}$, and $\mathcal{Q}(G) = \{(2, 2), (3, 1)\}$. For vectors w , $\hat{\theta}$ and θ defined as before, we have $w = (w_{2,1}, w_{2,2}, w_{3,1})$, $\hat{\theta} = (\hat{\theta}_{2,1}, \hat{\theta}_{2,2}, \hat{\theta}_{3,1})$ and $\theta = (\theta_{2,2,1}, \theta_{3,1,1})$. Let η be the all-zero vector and $b = 0$. The ensemble of Abelian group codes consists of the set of all homomorphisms from some $J = \mathbb{Z}_2^{kw_{2,1}} \oplus \mathbb{Z}_4^{kw_{2,2}} \oplus \mathbb{Z}_3^{kw_{3,1}}$. Any vector $\hat{\theta} = (\hat{\theta}_{2,1}, \hat{\theta}_{2,2}, \hat{\theta}_{3,1})$ with $0 \leq \hat{\theta}_{2,1} \leq 1$, $0 \leq \hat{\theta}_{2,2} \leq 2$ and $0 \leq \hat{\theta}_{3,1} \leq 1$ corresponds to a subgroup $K_{\hat{\theta}}$ of the input group J given by

$$K_{\hat{\theta}} = 2^{\hat{\theta}_{2,1}} \mathbb{Z}_2^{kw_{2,1}} \oplus 2^{\hat{\theta}_{2,2}} \mathbb{Z}_4^{kw_{2,2}} \oplus 3^{\hat{\theta}_{3,1}} \mathbb{Z}_3^{kw_{3,1}}$$

Similarly, any $\theta(\eta + \hat{\theta}) = (\theta_{2,2,1}, \theta_{3,1,1})$ corresponds to a subgroup $H_{\eta+\hat{\theta}} = 2^{\theta_{2,2,1}} \mathbb{Z}_4 \oplus 3^{\theta_{3,1,1}} \mathbb{Z}_3$ of the group G .

VI. PROOF OF CONVERSE

Consider an arbitrary shifted group code \mathbb{C} with parameters (n, k, w) . We assume that the associated homomorphism is a one-to-one mapping. We can express the code compactly as follows:

$$\mathbb{C} = \left\{ \bigoplus_{i=1}^n \left[\bigoplus_{(p,r,m)=\mathcal{G}(G)} \sum_{s=1}^{r_p} a_{p,s} g_{(p,s) \rightarrow (r,m)}^{(i)} + B^{(i)} \right] : a_{p,s} \in \mathbb{Z}_{p^s}^{kw_{p,s}}, \forall (p,s) \in \mathcal{S}(G) \right\}.$$

For every pair of vectors (η, b) as defined in Section V-A, define

$$\Gamma_{\eta, b} = \left\{ i \in [1, n] : g_{(p,s) \rightarrow (r,m)}^{(i)} \in p^{\eta_{p,r,m,s} + |r-s|^+} \mathbb{Z}_{p^r}^{kw_{p,s}} \setminus p^{\eta_{p,r,m,s} + 1 + |r-s|^+} \mathbb{Z}_{p^r}^{kw_{p,s}}, B^{(i)} = b, \forall (p, r, m, s) \right\}$$

Let $\hat{\theta}$ be an arbitrary vector whose components $\hat{\theta}_{p,s}$ are indexed by $(p, s) \in \mathcal{S}(G)$ and satisfy $0 \leq \hat{\theta}_{p,s} \leq s$. Construct a one-to-one correspondence $a_{p,s} \leftrightarrow (\tilde{a}_{p,s}, \hat{a}_{p,s})$ where $\tilde{a}_{p,s} \in p^{\hat{\theta}_{p,s}} \mathbb{Z}_{p^s}^{kw_{p,s}}$ and $\hat{a}_{p,s} \in \mathbb{Z}_{p^{\hat{\theta}_{p,s}}}^{kw_{p,s}}$.

A. Channel Coding

For an arbitrary $\hat{a}_{p,s} \in \mathbb{Z}_{p^{\hat{\theta}_{p,s}}}^{kw_{p,s}}$, consider the following subcode of \mathbb{C} :

$$\mathbb{C}_1(\hat{\theta}, \hat{a}) = \left\{ \bigoplus_{i=1}^n \left[\bigoplus_{(p,r,m) \in \mathcal{G}(G)} \sum_{s=1}^{r_p} (\hat{a}_{p,s} + \tilde{a}_{p,s}) g_{(p,s) \rightarrow (r,m)}^{(i)} + B^{(i)} \right] : \tilde{a}_{p,s} \in p^{\hat{\theta}_{p,s}} \mathbb{Z}_{p^s}^{kw_{p,s}}, \forall (p, s) \in \mathcal{S}(G) \right\}.$$

The rate of the code $\mathbb{C}_1(\hat{\theta}, \hat{a})$ is given by $(1 - \omega_{\hat{\theta}}) \frac{k}{n} \sum_{(p,s) \in \mathcal{S}(G)} sw_{p,s} \log p$.

For a given channel $(\mathcal{X} = G, \mathcal{Y}, W_{Y|X})$, suppose that rate R is achievable using group codes. Consider an arbitrary $\epsilon > 0$. This implies that there exists a shifted group code \mathbb{C} with parameters (n, k, w) that yields a maximal error probability τ such that $\tau \leq \epsilon$ and $\frac{k}{n} \sum_{(p,s) \in \mathcal{S}(G)} sw_{p,s} \log p \geq R - \epsilon$. Since the maximal error probability of the code is τ , the average error probability is no greater than τ . Using a uniform distribution on a , we let X_i denote the random channel input at the i th channel use induced by this code.

For the subcode $\mathbb{C}_1(\hat{\theta}, \hat{a})$, the average error probability is no greater than τ . Using the fact that \tilde{a} is uniformly distributed over its range, for $i \in \Gamma_{\eta, b}$, in the code $\mathbb{C}_1(\hat{\theta}, \hat{a})$, the channel input $X_i(\hat{\theta}, \hat{a})$ at the i th channel use has the following distribution

$$P(X_i(\hat{\theta}, \hat{a}) = \beta) = \prod_{(p,r,m) \in \mathcal{G}(G)} p^{-|r - \boldsymbol{\theta}(\eta + \hat{\theta})_{(p,r,m)}|^+} = \frac{1}{|H_{\eta + \hat{\theta}}|},$$

if $\beta_{(p,r,m)} \in \sum_{s=1}^{r_p} \hat{a}_{p,s} g_{(p,s) \rightarrow (r,m)}^{(i)} + b_{p,r,m} + p^{\boldsymbol{\theta}(\eta + \hat{\theta})_{(p,r,m)}} \mathbb{Z}_{p^r}$ for all $(p, r, m) \in \mathcal{G}(G)$, and $P(X_i(\hat{\theta}, \hat{a}) = \beta) = 0$ otherwise. Using Fano's inequality, and the standard information theoretic arguments, we have for every $\hat{\theta}$ with $0 \leq \hat{\theta}_{p,s} \leq s$, and \hat{a} with $\hat{a}_{p,s} \in \mathbb{Z}_{p^{\hat{\theta}_{p,s}}}^{kw_{p,s}}$

$$(1 - \omega_{\hat{\theta}})(R - \epsilon)(1 - \tau) - \frac{1}{n} \leq \frac{1}{n} \sum_{i=1}^n I(X_i(\hat{\theta}, \hat{a}); Y_i)$$

Hence, averaging with uniform distribution on \hat{a} , we get for all $\hat{\theta} \neq \mathbf{s}$,

$$(1 - \omega_{\hat{\theta}})(R - \epsilon)(1 - \tau) - \frac{1}{n} \leq \frac{1}{n} \sum_{i=1}^n \sum_{\hat{a}} P(\hat{a}) I(X_i(\hat{\theta}, \hat{a}); Y_i) \quad (20)$$

$$\stackrel{(a)}{=} \sum_{\eta, b} \sum_{i \in \Gamma(\eta, b)} \frac{1}{n} \sum_{\hat{a}} P(\hat{a}) I(X_i; Y_i | X_i \in \hat{\mathbf{a}}\mathbf{g}^{(i)} + b + H_{\eta+\hat{\theta}}) \quad (21)$$

$$\stackrel{(b)}{=} \sum_{\eta, b} \sum_{i \in \Gamma(\eta, b)} \frac{1}{n} I(X_i; Y_i | X_i \in \hat{\mathbf{A}}\mathbf{g}^{(i)} + b + H_{\eta+\hat{\theta}}) \quad (22)$$

$$\stackrel{(c)}{=} \sum_{\eta, b} \frac{|\Gamma(\eta, b)|}{n} I(X_{\eta, b}; Y | [X_{\eta, b}]_{\hat{\theta}}) \quad (23)$$

where in (a) we have expressed $\bigoplus_{(p,r,m) \in \mathcal{G}(G)} \sum_{s=1}^{r_p} \hat{a}_{p,s} g_{(p,s) \rightarrow (r,m)}^{(i)} + b_{p,r,m}$ as $\hat{\mathbf{a}}\mathbf{g}^{(i)} + b$, in (b) $\hat{\mathbf{A}}$ denotes the random variable corresponding to \hat{a} , in (c) we have used the fact that $\hat{\mathbf{A}}\mathbf{g}^{(i)} + b + H_{\eta+\hat{\theta}}$ is uniform over the set of cosets of $H_{\eta+\hat{\theta}} + b$ in $H_{\eta} + b$. Hence the converse follows.

B. Source Coding

Let us express \mathbb{C} as

$$\mathbb{C} = \left\{ \bigoplus_{\eta, b} \bigoplus_{i \in \Gamma(\eta, b)} \left[\bigoplus_{(p,r,m) \in \mathcal{G}(G)} \sum_{s=1}^{r_p} a_{p,s} g_{(p,s) \rightarrow (r,m)}^{(i)} + b \right] : a_{p,s} \in \mathbb{Z}_{p^s}^{kw_{p,s}}, \forall (p, s) \in \mathcal{S}(G) \right\}.$$

Consider the following code:

$$\begin{aligned} \mathbb{C}_2(\hat{\theta}) &= \left\{ \bigoplus_{\eta, b} \bigoplus_{i \in \Gamma(\eta, b)} \left[\bigoplus_{(p,r,m) \in \mathcal{G}(G)} \sum_{s=1}^{r_p} a_{p,s} g_{(p,s) \rightarrow (r,m)}^{(i)} + b + H_{\eta+\hat{\theta}} \right] : a_{p,s} \in \mathbb{Z}_{p^s}^{kw_{p,s}}, \forall (p, s) \in \mathcal{S}(G) \right\} \\ &\stackrel{(a)}{=} \left\{ \bigoplus_{\eta, b} \bigoplus_{i \in \Gamma(\eta, b)} \left[\bigoplus_{(p,r,m) \in \mathcal{G}(G)} \sum_{s=1}^{r_p} \hat{a}_{p,s} g_{(p,s) \rightarrow (r,m)}^{(i)} + b + H_{\eta+\hat{\theta}} \right] : \hat{a}_{p,s} \in \mathbb{Z}_{p^{\hat{\theta}_{p,s}}}^{kw_{p,s}}, \forall (p, s) \in \mathcal{S}(G) \right\} \end{aligned}$$

where (a) follows from the fact that for $i \in \Gamma(\eta, b)$, we have

$$\left\{ \bigoplus_{(p,r,m) \in \mathcal{G}(G)} \sum_{s=1}^{r_p} \tilde{a}_{p,s} g_{(p,s) \rightarrow (r,m)}^{(i)} : \tilde{a}_{p,s} \in p^{\hat{\theta}_{p,s}} \mathbb{Z}_{p^s}^{kw_{p,s}}, \forall (p, s) \in \mathcal{S}(G) \right\} = H_{\eta+\hat{\theta}}.$$

The rate of the code $\mathbb{C}_2(\hat{\theta})$ is given by $\omega_{\hat{\theta}} \frac{k}{n} \sum_{(p,s) \in \mathcal{S}(G)} sw_{p,s} \log p$.

For a given source $(\mathcal{X}, \mathcal{U} = G, p_X, d)$, suppose the pair of rate and distortion (R, D) is achievable using group codes. Consider an arbitrary $\epsilon > 0$. This implies that there exists a shifted group code \mathbb{C} with parameters (n, k, w) that yields a distortion Δ such that $\Delta \leq D + \epsilon$, $\frac{k}{n} \sum_{(p,s) \in \mathcal{S}(G)} sw_{p,s} \log p \leq R + \epsilon$, and $\frac{1}{n} \sum_{i=1}^n [\log \Theta_i - H(U_i)] \leq \epsilon$. Let $U^n = \text{Dec}(\text{Enc}(X^n))$, induced by the code \mathbb{C} . Note that for $i \in \Gamma(\eta, b)$, the i th sample U_i takes values in $H_{\eta} + b$. Let $[U_i]_{\hat{\theta}} = U_i + H_{\eta+\hat{\theta}}$ denote the unique coset of $H_{\eta+\hat{\theta}}$ in

$H_\eta + b$ that contains U_i . Observe that $[U_i]_{\hat{\theta}}$ is a function of U_i for $i \in [1, n]$. Let $[U^n]_{\hat{\theta}} = \bigoplus_{i \in [1, n]} [U_i]_{\hat{\theta}}$. Note that $[U^n]_{\hat{\theta}} \in \mathbb{C}_2(\hat{\theta})$. Hence we get the following conditions for the rate of the code for all $\hat{\theta} \neq \mathbf{0}$:

$$\omega_{\hat{\theta}}(R + \epsilon) \geq \frac{1}{n} H([U^n]_{\hat{\theta}}) = \frac{1}{n} I([U^n]_{\hat{\theta}}; X^n) \geq \frac{1}{n} \sum_{i=1}^n I([U_i]_{\hat{\theta}}; X_i) \quad (24)$$

$$= \sum_{\eta, b} \frac{|\Gamma(\eta, b)|}{n} \sum_{i \in \Gamma(\eta, b)} \frac{1}{|\Gamma(\eta, b)|} I([U_i]_{\hat{\theta}}; X_i) \quad (25)$$

$$\stackrel{(a)}{=} \sum_{\eta, b} \frac{|\Gamma(\eta, b)|}{n} \sum_{i \in \Gamma(\eta, b)} \frac{1}{|\Gamma(\eta, b)|} I(P_X, \hat{P}_{\hat{\theta}, i}) \quad (26)$$

$$\stackrel{(b)}{\geq} \sum_{\eta, b} \frac{|\Gamma(\eta, b)|}{n} I \left(P_X, \sum_{i \in \Gamma(\eta, b)} \frac{1}{|\Gamma(\eta, b)|} \hat{P}_{\hat{\theta}, i} \right), \quad (27)$$

$$\stackrel{(c)}{\geq} \sum_{\eta, b} \frac{|\Gamma(\eta, b)|}{n} I([U_{\eta, b}]_{\hat{\theta}}; X). \quad (28)$$

Here (a) follows by denoting the conditional probability $P([U_i]_{\hat{\theta}} = u | X_i = x)$ as $\hat{P}_{\hat{\theta}, i}(u|x)$, and writing explicitly the mutual information as a function of the source distribution and the conditional distribution of the corresponding function of the reconstruction given the source. (b) follows from convexity of mutual information. In (c) $[U_{\eta, b}]_{\hat{\theta}}$ denote the random variable which is related to X through the conditional probability distribution $\frac{1}{|\Gamma(\eta, b)|} \sum_{i \in \Gamma(\eta, b)} \hat{P}_{\hat{\theta}, i}$.

Note that for $\hat{\theta} = \mathbf{s}$ we have $[U_{\eta, b}]_{\mathbf{s}} = U_{\eta, b}$. Hence we have the following conditions regarding the distortion of the code:

$$\begin{aligned} D + \epsilon &\geq \Delta \geq \frac{1}{n} \sum_{i=1}^n \mathbb{E}(d(X_i, U_i)) = \sum_{\eta, b} \frac{|\Gamma(\eta, b)|}{n} \sum_{i \in \Gamma(\eta, b)} \frac{1}{|\Gamma(\eta, b)|} \sum_{x \in \mathcal{X}} P_X(x) \sum_{u \in H_{\eta+b}} \hat{P}_{\mathbf{s}, i}(u|x) d(x, u) \\ &= \sum_{\eta, b} \frac{|\Gamma(\eta, b)|}{n} \sum_{x \in \mathcal{X}} P_X(x) \sum_{u \in H_{\eta+b}} P(U_{\eta, b} = u | X = x) d(x, u) \end{aligned} \quad (29)$$

$$= \sum_{\eta, b} \frac{|\Gamma(\eta, b)|}{n} \mathbb{E} d(X, U_{\eta, b}). \quad (30)$$

Finally, we ensure that $U_{\eta, b}$ is distributed nearly uniformly over its range. For every η and b , let $\bar{U}_{\eta, b}$ be uniformly distributed over $H_\eta + b$. Now using the relation between the variational distance and

information divergence [13, p.58], we have

$$\sum_{\eta,b} \frac{|\Gamma(\eta, b)|}{2n \ln 2} d_v^2(U_{\eta,b}, \bar{U}_{\eta,b}) \leq \sum_{\eta,b} \frac{|\Gamma(\eta, b)|}{n} D(U_{\eta,b} \| \bar{U}_{\eta,b}) \quad (31)$$

$$= \sum_{\eta,b} \frac{|\Gamma(\eta, b)|}{n} D \left(\frac{1}{|\Gamma(\eta, b)|} \sum_{i \in \Gamma(\eta,b)} U_i \middle\| \bar{U}_{\eta,b} \right) \quad (32)$$

$$\leq \sum_{\eta,b} \frac{|\Gamma(\eta, b)|}{n} \frac{1}{|\Gamma(\eta, b)|} \sum_{i \in \Gamma(\eta,b)} D(U_i \| \bar{U}_{\eta,b}) \quad (33)$$

$$= \sum_{\eta,b} \frac{1}{n} \sum_{i \in \Gamma(\eta,b)} D(U_i \| \bar{U}_i) = \frac{1}{n} \sum_{i=1}^n [\log \Theta_i - H(U_i)] \leq \epsilon. \quad (34)$$

The converse follows from the continuity of mutual information.

VII. ACHIEVABILITY

A. Channel Coding

We give proof of only the essential elements for conciseness. Fix a triple of code parameters (n, k, w) . Let R denote the rate of the code. First we consider the special case where for every channel use, the input can take all possible values in G . This corresponds to the choice: $\alpha_{\eta^*, b} = 1$, where η^* is the all-zero vector and b is arbitrary, and hence $H_{\eta^*} = G$. Generalization to arbitrary probability distributions is relatively straightforward.

1) *Encoding and Decoding*: Following the analysis of Section IV-B, we construct the ensemble of group codes of length n over G as the image of all homomorphisms ϕ from some Abelian group J into G^n where J and G^n are as in Equations (11) and (9), respectively. The random homomorphism ϕ is described in Equation (13).

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 B uniformly distributed over G^n . Given a message $a \in J$, the encoder maps it to $x = \phi(a) + B$ and x is then fed to the channel. At the receiver, after receiving the channel output $y \in \mathcal{Y}^n$, the decoder looks for a unique $\tilde{a} \in J$ such that $\phi(\tilde{a}) + B$ is jointly typical with y with respect to the distribution $P_X W_{Y|X}$ where P_X is uniform over G . If the decoder does not find such \tilde{a} or if such \tilde{a} is not unique, it declares error.

2) *Error Analysis*: Let a , x and y be the message, the channel input and the channel output, respectively. The error event can be characterized by the union of two events: $E(a) = E_1(a) \cup E_2(a)$ where $E_1(a)$ is the event that $\phi(a) + B$ is not jointly typical with y and $E_2(a)$ is the event that there exists a

$\tilde{a} \neq a$ such that $\phi(\tilde{a}) + B$ is jointly typical with y . We can provide an upper bound on the probability of the error event as $P(E(a)) \leq P(E_1(a)) + P(E_2(a) \cap (E_1(a))^c)$. Using the standard approach, one can show that $P(E_1(a)) \rightarrow 0$ as $n \rightarrow \infty$. The probability of the error event $E_2(a) \cap (E_1(a))^c$ can be written as

$$P_{avg}(E_2(a) \cap (E_1(a))^c) = \sum_{x \in G^n} \mathbb{1}_{\{\phi(a)+B=x\}} \sum_{y \in A_\epsilon^n(Y|x)} W_{Y|X}^n(y|x) \mathbb{1}_{\{\exists \tilde{a} \in J: \tilde{a} \neq a, \phi(\tilde{a})+B \in A_\epsilon^n(X|y)\}}$$

The expected value of this probability over the ensemble is given by $\mathbb{E}\{P_{avg}(E_2(a) \cap (E_1(a))^c)\} = P_{err}$ where

$$P_{err} = \sum_{x \in G^n} \sum_{y \in A_\epsilon^n(Y|x)} W_{Y|X}^n(y|x) P(\phi(a) + B = x, \exists \tilde{a} \in J : \tilde{a} \neq a, \phi(\tilde{a}) + B \in A_\epsilon^n(X|y))$$

Using the union bound, we have

$$P_{err} \leq \sum_{x \in G^n} \sum_{y \in A_\epsilon^n(Y|x)} \sum_{\substack{\tilde{a} \in J \\ \tilde{a} \neq a}} \sum_{\tilde{x} \in A_\epsilon^n(X|y)} W_{Y|X}^n(y|x) P(\phi(a) + B = x, \phi(\tilde{a}) + B = \tilde{x})$$

We need the following lemmas to proceed.

Lemma VII.1. For $a, \tilde{a} \in J$, $x, \tilde{x} \in G^n$ and for $(p, s) \in \mathcal{Q}(J) = \mathcal{S}(G)$, let $\hat{\theta}_{p,s} \in \{0, 1, \dots, s\}$ be such that

$$\tilde{a}_{p,s} - a_{p,s} \in p^{\hat{\theta}_{p,s}} \mathbb{Z}_{p^s}^{kw_{p,s}} \setminus p^{\hat{\theta}_{p,s}+1} \mathbb{Z}_{p^s}^{kw_{p,s}}$$

Then,

$$P(\phi(a) + B = x, \phi(\tilde{a}) + B = \tilde{x}) = \begin{cases} \frac{1}{|G|^n} \frac{1}{|H_{\eta^*+\hat{\theta}}|^n} & \text{If } \tilde{x} - x \in H_{\eta^*+\hat{\theta}}^n \\ 0 & \text{Otherwise} \end{cases}$$

Proof: The proof is provided in Appendix IX-B ■

Lemma VII.2. Let X be a random variable taking values from the group G and for a subgroup H of G , define $[X] = X + H$. For $y \in A_\epsilon^n(Y)$ and $x \in A_\epsilon^n(X|y)$, let $z = [x] = x + H^n$. Then we have

$$(x + H^n) \cap A_\epsilon^n(X|y) = A_\epsilon^n(X|zy)$$

and

$$(1 - \epsilon) 2^{n[H(X|Y[X]) - O(\epsilon)]} \leq |(x + H^n) \cap A_\epsilon^n(X|y)| \leq 2^{n[H(X|Y[X]) + O(\epsilon)]}$$

Proof: The proof is provided in Appendix IX-C ■

For $a \in J$, and for $\hat{\theta}$ with $\hat{\theta}_{p,s} \in \{0, 1, \dots, s\}$ for all $(p, s) \in \mathcal{S}(G)$, let

$$T_{\hat{\theta}}(a) = \{\tilde{a} \in J | \tilde{a}_{p,s} - a_{p,s} \in p^{\hat{\theta}_{p,s}} \mathbb{Z}_{p^s}^{kw_{p,s}} \setminus p^{\hat{\theta}_{p,s}+1} \mathbb{Z}_{p^s}^{kw_{p,s}}, \forall (p, s) \in \mathcal{S}(G)\}$$

It follows that for all $a \in J$

$$|T_{\hat{\theta}}(a)| = \prod_{(p,s) \in \mathcal{S}(G)} p^{(s-\hat{\theta}_{p,s})kw_{p,s}},$$

and

$$P_{err} \leq \sum_{x \in G^n} \sum_{y \in A_e^n(Y|x)} \sum_{\hat{\theta} \neq \mathbf{s}} \sum_{\tilde{a} \in T_{\hat{\theta}}(a)} \sum_{\tilde{x} \in A_e^n(X|y)} W_{Y|X}^n(y|x) P(\phi(a) + B = x, \phi(\tilde{a}) + B = \tilde{x})$$

Using Lemmas VII.1, and VII.2, we have

$$\begin{aligned} P_{err} &\leq \sum_{\hat{\theta} \neq \mathbf{s}} \sum_{x \in G^n} \sum_{y \in A_e^n(Y|x)} \sum_{\tilde{a} \in T_{\hat{\theta}}(a)} \sum_{\substack{\tilde{x} \in A_e^n(X|y) \\ \tilde{x} \in x + H_{\eta^* + \hat{\theta}}^n}} W_{Y|X}^n(y|x) \frac{1}{|G|^n} \frac{1}{|H_{\eta^* + \hat{\theta}}|^n} \\ &\leq \sum_{\hat{\theta} \neq \mathbf{s}} \sum_{x \in G^n} \sum_{y \in A_e^n(Y|x)} \sum_{\tilde{a} \in T_{\hat{\theta}}(a)} W_{Y|X}^n(y|x) 2^{n[H(X_{\eta^*,b}|Y, [X_{\eta^*,b}]_{\hat{\theta}}) + O(\epsilon)]} \frac{1}{|G|^n} \frac{1}{|H_{\eta^* + \hat{\theta}}|^n} \\ &\leq \sum_{\hat{\theta} \neq \mathbf{s}} \left(\prod_{(p,s) \in \mathcal{S}(G)} p^{(s-\hat{\theta}_{p,s})kw_{p,s}} \right) 2^{n[H(X_{\eta^*,b}|Y, [X_{\eta^*,b}]_{\hat{\theta}}) + O(\epsilon)]} \frac{1}{|H_{\eta^* + \hat{\theta}}|^n} \\ &= \sum_{\hat{\theta} \neq \mathbf{s}} \exp_2 \left\{ n \left[\frac{k}{n} \sum_{(p,s) \in \mathcal{S}(G)} (s - \hat{\theta}_{p,s}) w_{p,s} \log p + H(X_{\eta^*,b}|Y, [X_{\eta^*,b}]_{\hat{\theta}}) - \log |H_{\eta^* + \hat{\theta}}| + O(\epsilon) \right] \right\} \end{aligned}$$

Recall that $R = \frac{k}{n} \sum_{(p,s) \in \mathcal{S}(G)} s w_{p,s} \log p$. Noting that P_{err} is the probability of error for message a averaged over the ensemble, in order for the maximal error probability to go to zero, we require the exponent of all the terms to be negative; or equivalently, for all $\hat{\theta} \neq \mathbf{s}$,

$$R \frac{\sum_{(p,s) \in \mathcal{S}(G)} (s - \hat{\theta}_{p,s}) w_{p,s} \log p}{\sum_{(p,s) \in \mathcal{S}(G)} s w_{p,s} \log q} < \log |H_{\eta^* + \hat{\theta}}| - H(X_{\eta^*,b}|Y, [X_{\eta^*,b}]_{\hat{\theta}}) - O(\epsilon).$$

Therefore, the achievability conditions are

$$R \leq \frac{1}{1 - \omega_{\hat{\theta}}} I(X_{\eta^*,b}; Y | [X_{\eta^*,b}]_{\hat{\theta}})$$

for all $\hat{\theta} \neq \mathbf{s}$. This means that the following rate is achievable

$$R = \min_{\hat{\theta} \neq \mathbf{s}} \frac{1}{1 - \omega_{\hat{\theta}}} I(X_{\eta^*,b}; Y | [X_{\eta^*,b}]_{\hat{\theta}}).$$

For a general α , we use an extended random coding argument. We construct the ensemble of group codes of length n over G as the image of all homomorphisms ϕ from some Abelian group J into $\bigoplus_{\eta,b} (H_{\eta} + b)^{n\alpha_{\eta,b}}$. The random homomorphism ϕ is described in Equation (13). Given a message

$a \in J$, the encoder maps it to $x = \phi(a) + B$ and x is then fed to the channel. At the receiver, after receiving the channel output $y \in \mathcal{Y}^n$, the decoder looks for a unique $\tilde{a} \in J$ such that for every (η, b) , the appropriate set of $n\alpha_{\eta,b}$ samples of $\phi(\tilde{a}) + B$ is jointly typical with the corresponding set of $n\alpha_{\eta,b}$ samples of y with respect to the distribution $P_{\eta,b}W_{Y|X}$ where $P_{\eta,b}$ is uniform over $H_\eta + b$. If the decoder does not find such \tilde{a} or if such \tilde{a} is not unique, it declares error. Using a similar analysis, one gets the desired achievability result.

3) \mathbb{Z}_4 : Evaluating the expression for the group capacity for codes over \mathbb{Z}_4 , we get three non-redundant terms as follows:

$$R < \sup_{\alpha_1, \alpha_2, w_0} \min\{T_1, T_2, T_3\}$$

where supremum is over all α_1 and α_2 such that $0 \leq \alpha_1, \alpha_2, \alpha_1 + \alpha_2 \leq 1$, and $0 \leq w_0 \leq 1$ and $T_1 = \alpha_1 I_4 + \alpha_2 I_2 + (1 - \alpha_1 - \alpha_2) I'_2$, $T_2 = (1 + w_0) \left[\frac{\alpha_1}{2} (I_2 + I'_2) + \alpha_2 I_2 + (1 - \alpha_1 - \alpha_2) I'_2 \right]$ and $T_3 = \frac{(1+w_0)\alpha_1}{w_0} (I_2 + I'_2)$, $I_4 = I(X; Y)$, $I_2 = I(X; Y|X \in 2\mathbb{Z}_4)$ and $I'_2 = I(X; Y|X \notin 2\mathbb{Z}_4)$. This can be solved to obtain the following

$$C = \max\{\min\{I_4, (I_2 + I'_2)\}, I_2, I'_2\}.$$

B. Source Coding

Fix a triple of code parameters (n, k, w) . Let R denote the rate of the code. We consider the special case where $\alpha_{\eta^*, b} = 1$, where η^* is the all-zero vector, and hence $H_{\eta^*} = G$. Generalization to arbitrary probability distributions is straightforward. Fix a conditional distribution $P_{U|X}$ on G such that U is uniform on G .

1) *Encoding and Decoding*: To find an achievable rate for a distortion level D , we use a random coding argument as in channel coding. Consider a random shifted group code as $\mathbb{C} = \{\phi(a) + B : a \in J\}$, where ϕ and B are uniformly distributed over their range and are independent of each other. Given the source output sequence $x \in \mathcal{X}^n$, the random encoder looks for a codeword $u \in \mathbb{C}$ such that u is jointly typical with x with respect to p_{XU} . If it finds at least one such u , it encodes x to u (if it finds more than one such u it picks one of them at random). Otherwise, it declares error. The decoder outputs u as the source reconstruction.

2) *Error Analysis*: Let $x = (x_1, \dots, x_n)$ and $u = (u_1, \dots, u_n)$ be the source output and the decoder output, respectively. Note that if the encoder declares no error then since x and u are jointly typical, $(d(x_i, u_i))_{i=1, \dots, n}$ is typical with respect to the distribution of $d(X, U)$. Therefore for large n , $\frac{1}{n} d(x, u) = \frac{1}{n} \sum_{i=1}^n d(x_i, u_i) \approx \mathbb{E}\{d(X, U)\} \leq D$. It remains to show that the rate can be as small as

given in the theorem while keeping the probability of encoding error small.

Given the source output $x \in \mathcal{X}^n$, define

$$\alpha(x) = \sum_{u \in A_\epsilon^n(U|x)} \mathbb{1}_{\{u \in \mathcal{C}\}} = \sum_{u \in A_\epsilon^n(U|x)} \sum_{a \in J} \mathbb{1}_{\{\phi(a) + B = u\}}$$

An encoding error occurs if and only if $\alpha(x) = 0$. We use the following Chebyshev's inequality to show that under certain conditions the probability of error can be made arbitrarily small:

$$P(\alpha(x) = 0) \leq \frac{\text{var}\{\alpha(x)\}}{\mathbb{E}\{\alpha(x)\}^2}$$

We have

$$\begin{aligned} \mathbb{E}\{\alpha(x)\} &= \sum_{u \in A_\epsilon^n(U|x)} \sum_{a \in J} P(\phi(a) + B = u) \\ &= \frac{|A_\epsilon^n(U|x)| \cdot |J|}{|G|^n} \end{aligned}$$

and

$$\begin{aligned} \mathbb{E}\{\alpha(x)^2\} &= \mathbb{E} \left\{ \sum_{u, \tilde{u} \in A_\epsilon^n(U|x)} \sum_{a, \tilde{a} \in J} \mathbb{1}_{\{\phi(a) + B = u, \phi(\tilde{a}) + B = \tilde{u}\}} \right\} \\ &= \sum_{u, \tilde{u} \in A_\epsilon^n(U|x)} \sum_{a, \tilde{a} \in J} P(\{\phi(a) + B = u, \phi(\tilde{a}) + B = \tilde{u}\}) \\ &= \sum_{\hat{\theta}} \sum_{a \in J} \sum_{u \in A_\epsilon^n(U|x)} \sum_{\tilde{a} \in T_{\hat{\theta}}(a)} \sum_{\substack{\tilde{u} \in A_\epsilon^n(U|x) \\ \tilde{u} - u \in H_{\eta^* + \hat{\theta}}^n}} \frac{1}{|G|^n} \cdot \frac{1}{|H_{\eta^* + \hat{\theta}}|^n} \end{aligned}$$

Note that the term corresponding to $\hat{\theta} = \mathbf{0}$ is bounded from above by $\mathbb{E}\{\alpha(x)\}^2$. Using Lemma VII.2, we have

$$\left| A_\epsilon^n(U|x) \cap \left(u + H_{\eta^* + \hat{\theta}}^n \right) \right| \leq 2^{n[H(U_{\eta^*, b}|[U_{\eta^*, b}]_{\hat{\theta}}, X) + O(\epsilon)]}$$

Therefore,

$$\begin{aligned} \text{var}\{\alpha\} &= \mathbb{E}\{\alpha(x)^2\} - \mathbb{E}\{\alpha(x)\}^2 \\ &\leq \sum_{\hat{\theta} \neq \mathbf{0}} |J| \cdot |A_\epsilon^n(U|x)| \left(\prod_{(p,s) \in \mathcal{S}(G)} p^{(s - \hat{\theta}_{p,s})kw_{q,s}} \right) \frac{2^{n[H(U_{\eta^*, b}|[U_{\eta^*, b}]_{\hat{\theta}}, X) + O(\epsilon)]}}{|G|^n \cdot |H_{\eta^* + \hat{\theta}}|^n} \end{aligned}$$

Therefore,

$$P(\alpha(x) = 0) \leq \frac{\text{var}\{\alpha(x)\}}{\mathbb{E}\{\alpha(x)\}^2} \leq \sum_{\hat{\theta} \neq \mathbf{0}} \left(\prod_{(p,s) \in \mathcal{S}(G)} p^{(s-\hat{\theta}_{p,s})kw_{p,s}} \right) \frac{2^{-n[H(U_{\eta^*,b}|X) - H(U_{\eta^*,b}|[U_{\eta^*,b}]_{\hat{\theta}}X) - O(\epsilon)]|G|^n}}{|J| \cdot |H_{\eta^*+\hat{\theta}}|^n}$$

Note that $H(U_{\eta^*,b}|X) - H(U_{\eta^*,b}|[U_{\eta^*,b}]_{\hat{\theta}}, X) = H([U_{\eta^*,b}]_{\hat{\theta}}|X)$ and

$$|J| = \prod_{(p,s) \in \mathcal{S}(G)} p^{ksw_{p,s}}$$

Therefore,

$$P(\alpha(x) = 0) \leq \sum_{\hat{\theta} \neq \mathbf{0}} \exp_2 \left\{ -n \left[H([U_{\eta^*,b}]_{\hat{\theta}}|X) - \log |H_{\eta^*} : H_{\eta^*+\hat{\theta}}| + \frac{k}{n} \sum_{(p,s) \in \mathcal{S}(G)} \hat{\theta}_{p,s} w_{p,s} \log p - O(\epsilon) \right] \right\}$$

In order for the probability of error to go to zero as n increases, we require the exponent of all the terms to be negative; or equivalently,

$$R = \max_{\hat{\theta} \neq \mathbf{0}} \frac{1}{\omega_{\hat{\theta}}} I([U_{\eta^*,b}]_{\hat{\theta}}; X)$$

is achievable. Using an extension of the random coding argument to general $\alpha_{\eta,b}$, similar to that considered in channel coding, we get the achievability of the theorem.

VIII. CONCLUSION

We derived the achievable set of rates using Abelian group codes for arbitrary discrete memoryless channels. In the case of linear codes, it simplifies to the symmetric capacity of the channel i.e., the Shannon capacity with the additional constraint that the channel input distribution is uniformly distributed. For the case where the underlying group is not a field, we observe that several subgroups of the group appear in the achievable rate and this causes the rate to be smaller than the symmetric capacity of the channel in general.

We derived a similar result for the source coding problem; i.e., the achievable rate-distortion function using Abelian group codes for arbitrary discrete memoryless sources. When the underlying group is a field, these group codes are linear codes, and this function is equivalent to the symmetric rate-distortion function i.e., the Shannon rate-distortion function with the additional constraint that the reconstruction random variable is uniformly distributed. We showed that when the underlying group is not a field, due to the algebraic structure of the code, certain subgroups of the group appear in the rate-distortion function and cause a larger rate for a given distortion level.

IX. APPENDIX

A. Proof of Lemma IV.2

We first prove that for a homomorphism $\phi, g_{(q,s,l) \rightarrow (p,r,m)}$ satisfies the above conditions. First assume $p \neq q$. Note that the only nonzero component of $\mathbb{I}_{J:q,s,l}$ takes values from \mathbb{Z}_{q^s} and therefore

$$q^s \mathbb{I}_{J:q,s,l} = \sum_{i=1, \dots, q^s}^{(J)} \mathbb{I}_{J:q,s,l} = 0$$

Note that since ϕ is a homomorphism, we have $\phi(q^s \mathbb{I}_{J:q,s,l}) = 0$. On the other hand,

$$\begin{aligned} \phi(q^s \mathbb{I}_{J:q,s,l}) &= \phi\left(\sum_{i=1, \dots, q^s}^{(J)} \mathbb{I}_{J:q,s,l}\right) \\ &= \sum_{i=1, \dots, q^s}^{(\tilde{G})} \phi(\mathbb{I}_{J:q,s,l}) \\ &= \bigoplus_{(p,r,m) \in \mathcal{G}(\tilde{G})} \left[\sum_{i=1, \dots, q^s}^{(\tilde{G})} \phi(\mathbb{I}_{J:q,s,l}) \right]_{p,r,m} \\ &= \bigoplus_{(p,r,m) \in \mathcal{G}(\tilde{G})} \sum_{i=1, \dots, q^s}^{(\mathbb{Z}_{p^r})} [\phi(\mathbb{I}_{J:q,s,l})]_{p,r,m} \\ &= \bigoplus_{(p,r,m) \in \mathcal{G}(\tilde{G})} q^s [\phi(\mathbb{I}_{J:q,s,l})]_{p,r,m} \\ &= \bigoplus_{(p,r,m) \in \mathcal{G}(\tilde{G})} q^s g_{(q,s,l) \rightarrow (p,r,m)} \end{aligned}$$

Therefore, we have $q^s g_{(q,s,l) \rightarrow (p,r,m)} = 0 \pmod{p^r}$ or equivalently $q^s g_{(q,s,l) \rightarrow (p,r,m)} = Cp^r$ for some integer C . Since $p \neq q$, this implies $p^r | g_{(q,s,l) \rightarrow (p,r,m)}$ and since $g_{(q,s,l) \rightarrow (p,r,m)}$ takes value from \mathbb{Z}_{p^r} , we have $g_{(q,s,l) \rightarrow (p,r,m)} = 0$.

Next, assume $p = q$ and $r \geq s$. Note that same as above, we have $\phi(q^s \mathbb{I}_{J:q,s,l}) = 0$ and

$$\phi(q^s \mathbb{I}_{J:q,s,l}) = \bigoplus_{(p,r,m) \in \mathcal{G}(\tilde{G})} q^s g_{(q,s,l) \rightarrow (p,r,m)}$$

and therefore, $q^s g_{(q,s,l) \rightarrow (p,r,m)} = 0 \pmod{p^r}$. Since $g_{(q,s,l) \rightarrow (p,r,m)}$ takes values from \mathbb{Z}_{p^r} and $p = q$, this implies $p^{r-s} | g_{(q,s,l) \rightarrow (p,r,m)}$ or equivalently $g_{(q,s,l) \rightarrow (p,r,m)} \in p^{r-s} \mathbb{Z}_{p^r}$.

Next we show that any mapping described by (7) satisfying the conditions of the lemma is a homomorphism. For two elements $a, b \in J$ and for $(p, r, m) \in \mathcal{G}(\tilde{G})$ we have

$$\begin{aligned}
[\phi(a + b)]_{p,r,m} &= \left[\phi \left(\bigoplus_{(q,s,l) \in \mathcal{G}(J)} (a_{q,s,l} +_{q^s} b_{q,s,l}) \right) \right]_{p,r,m} \\
&= \left[\phi \left(\bigwedge_{(q,s,l) \in \mathcal{G}(J)}^{(J)} (a_{q,s,l} +_{q^s} b_{q,s,l}) \mathbb{I}_{J:q,s,l} \right) \right]_{p,r,m} \\
&= \left[\phi \left(\bigwedge_{(q,s,l) \in \mathcal{G}(J)}^{(J)} \bigwedge_{i=1, \dots, a_{q,s,l} +_{q^s} b_{q,s,l}}^{(J)} \mathbb{I}_{J:q,s,l} \right) \right]_{p,r,m} \\
&= \left[\bigwedge_{(q,s,l) \in \mathcal{G}(J)}^{(\tilde{G})} \bigwedge_{i=1, \dots, a_{q,s,l} +_{q^s} b_{q,s,l}}^{(\tilde{G})} \phi(\mathbb{I}_{J:q,s,l}) \right]_{p,r,m} \\
&= \bigwedge_{(q,s,l) \in \mathcal{G}(J)}^{(\mathbb{Z}_{p^r})} \bigwedge_{i=1, \dots, a_{q,s,l} +_{q^s} b_{q,s,l}}^{(\mathbb{Z}_{p^r})} [\phi(\mathbb{I}_{J:q,s,l})]_{p,r,m} \\
&= \bigwedge_{(q,s,l) \in \mathcal{G}(J)}^{(\mathbb{Z}_{p^r})} \bigwedge_{i=1, \dots, a_{q,s,l} +_{q^s} b_{q,s,l}}^{(\mathbb{Z}_{p^r})} g_{(q,s,l) \rightarrow (p,r,m)} \tag{35}
\end{aligned}$$

On the other hand, we have

$$\begin{aligned}
[\phi(a) + \phi(b)]_{p,r,m} &= [\phi(a)]_{p,r,m} +_{p^r} [\phi(b)]_{p,r,m} \\
&= \left(\bigwedge_{(q,s,l) \in \mathcal{G}(J)}^{(\mathbb{Z}_{p^r})} a_{q,s,l} g_{(q,s,l) \rightarrow (p,r,m)} \right) +_{p^r} \left(\bigwedge_{(q,s,l) \in \mathcal{G}(J)}^{(\mathbb{Z}_{p^r})} b_{q,s,l} g_{(q,s,l) \rightarrow (p,r,m)} \right) \\
&= \left(\bigwedge_{(q,s,l) \in \mathcal{G}(J)}^{(\mathbb{Z}_{p^r})} \bigwedge_{i=1, \dots, a_{q,s,l}}^{(\mathbb{Z}_{p^r})} g_{(q,s,l) \rightarrow (p,r,m)} \right) +_{p^r} \left(\bigwedge_{(q,s,l) \in \mathcal{G}(J)}^{(\mathbb{Z}_{p^r})} \bigwedge_{i=1, \dots, b_{q,s,l}}^{(\mathbb{Z}_{p^r})} g_{(q,s,l) \rightarrow (p,r,m)} \right) \\
&= \bigwedge_{(q,s,l) \in \mathcal{G}(J)}^{(\mathbb{Z}_{p^r})} \bigwedge_{i=1, \dots, a_{q,s,l} +_{q^s} b_{q,s,l}}^{(\mathbb{Z}_{p^r})} g_{(q,s,l) \rightarrow (p,r,m)} \tag{36}
\end{aligned}$$

where the addition in $a_{q,s,l} +_{q^s} b_{q,s,l}$ is the integer addition.

In order to show that ϕ is a homomorphism, it suffices to show that under the conditions of the lemma, Equations (35) and (36) are equivalent. We show that for a fixed $(q, s, l) \in \mathcal{G}(J)$, if the conditions of the

lemma are satisfied, then

$$\sum_{i=1, \dots, a_{q,s,l} + b_{q,s,l}}^{\binom{\mathbb{Z}_{p^r}}{}} g_{(q,s,l) \rightarrow (p,r,m)} = \sum_{i=1, \dots, a_{q,s,l} + q^s b_{q,s,l}}^{\binom{\mathbb{Z}_{p^r}}{}} g_{(q,s,l) \rightarrow (p,r,m)} \quad (37)$$

Note that if $p \neq q$, then both summations are zero. Note that we have

$$\sum_{i=1, \dots, a_{q,s,l} + b_{q,s,l}}^{\binom{\mathbb{Z}_{p^r}}{}} g_{(q,s,l) \rightarrow (p,r,m)} = \sum_{i=1, \dots, (a_{q,s,l} + b_{q,s,l}) \pmod{p^r}}^{\binom{\mathbb{Z}_{p^r}}{}} g_{(q,s,l) \rightarrow (p,r,m)}$$

and

$$\sum_{i=1, \dots, a_{q,s,l} + q^s b_{q,s,l}}^{\binom{\mathbb{Z}_{p^r}}{}} g_{(q,s,l) \rightarrow (p,r,m)} = \sum_{i=1, \dots, (a_{q,s,l} + q^s b_{q,s,l}) \pmod{p^r}}^{\binom{\mathbb{Z}_{p^r}}{}} g_{(q,s,l) \rightarrow (p,r,m)}$$

If $p = q$ and $r \leq s$, then we have $(a_{q,s,l} + q^s b_{q,s,l}) \pmod{p^r} = (a_{q,s,l} + b_{q,s,l}) \pmod{p^r}$ and hence it follows that Equation (37) is satisfied. If $p = q$ and $r \geq s$, since $g_{(q,s,l) \rightarrow (p,r,m)} \in p^{r-s} \mathbb{Z}_{p^r}$ we have

$$\sum_{i=1, \dots, a_{q,s,l} + b_{q,s,l}}^{\binom{\mathbb{Z}_{p^r}}{}} g_{(q,s,l) \rightarrow (p,r,m)} = \sum_{i=1, \dots, (a_{q,s,l} + b_{q,s,l}) \pmod{p^s}}^{\binom{\mathbb{Z}_{p^r}}{}} g_{(q,s,l) \rightarrow (p,r,m)}$$

and hence it follows that Equation (37) is satisfied.

B. Proof of Lemma VII.1

Note that since $g_{(q,s,l) \rightarrow (p,r,m)}$'s and B are uniformly distributed, in order to find the desired joint probability, we need to count the number of choices for $g_{(q,s,l) \rightarrow (p,r,m)}$'s and B such that for $(p, r, m) \in \mathcal{G}(G^n)$,

$$\left(\sum_{(q,s,l) \in \mathcal{G}(J)}^{\binom{\mathbb{Z}_{p^r}}{}} a_{q,s,l} g_{(q,s,l) \rightarrow (p,r,m)} \right) +_{p^r} B_{p,r,m} = u_{p,r,m}$$

$$\left(\sum_{(q,s,l) \in \mathcal{G}(J)}^{\binom{\mathbb{Z}_{p^r}}{}} \tilde{a}_{q,s,l} g_{(q,s,l) \rightarrow (p,r,m)} \right) +_{p^r} B_{p,r,m} = \tilde{u}_{p,r,m}$$

and divide this number by the total number of choices which is equal to

$$|G|^n \cdot \prod_{(p,r,m) \in \mathcal{G}(G^n)} \prod_{\substack{(q,s,l) \in \mathcal{G}(J) \\ q=p}} p^{\min(r,s)} = |G|^n \cdot \left[\prod_{(p,r,m) \in \mathcal{G}(G)} \prod_{\substack{(q,s,l) \in \mathcal{G}(J) \\ q=p}} p^{\min(r,s)} \right]^n$$

where the term $p^{\min(r,s)}$ appears since the number of choices for $g_{(q,s,l) \rightarrow (p,r,m)}$ is p^r if $p = q, r \leq s$ and is equal to p^s if $p = q, r \geq s$. Since B can take values arbitrarily from G^n , the number of choices for the above set of conditions is equal to the number of choices for $g_{(q,s,l) \rightarrow (p,r,m)}$'s such that,

$$\left(\sum_{(q,s,l) \in \mathcal{G}(J)}^{\binom{\mathbb{Z}_{p^r}}{}} (\tilde{a}_{q,s,l} - a_{q,s,l}) g_{(q,s,l) \rightarrow (p,r,m)} \right) = \tilde{u}_{p,r,m} - u_{p,r,m}$$

Let $\hat{\theta}_{p,s,l} \in \{0, 1, \dots, s\}$ be such that $\tilde{a}_{p,s,l} - a_{p,s,l} \in p^{\hat{\theta}_{p,s,l}} \mathbb{Z}_{p^s} \setminus p^{\hat{\theta}_{p,s,l}+1} \mathbb{Z}_{p^s}$. Note

$$\hat{\theta}_{p,s} = \min_{\{1 \leq l \leq kw_{p,s}\}} \hat{\theta}_{p,s,l}.$$

Note that for all $(q, s, l) \in \mathcal{G}(J)$, $(\tilde{a}_{q,s,l} - a_{q,s,l}) g_{(q,s,l) \rightarrow (p,r,m)} \in p^{\hat{\theta}_{p,s,l}} \mathbb{Z}_{p^r}$. Therefore we require $\tilde{u}_{p,r,m} - u_{p,r,m} \in p^{\hat{\theta}_{p,s,l}} \mathbb{Z}_{p^r}$ and therefore we require $\tilde{u} - u \in H_{\eta^* + \hat{\theta}}$ or otherwise the probability would be zero.

For fixed $p \in \mathcal{P}(G)$ and $r \in \mathcal{R}_p(G)$, let $(q^*, s^*, l^*) \in \mathcal{G}(J)$ be such that $q^* = p$ and

$$\boldsymbol{\theta}(\eta^* + \hat{\theta})_{p,r,m} = |r - s^*|^+ + \hat{\theta}_{q^*, s^*, l^*}.$$

For fixed $(p, r, m) \in \mathcal{G}(G^n)$, and for $(q, s, l) \neq (q^*, s^*, l^*)$, choose $g_{(q,s,l) \rightarrow (p,r,m)}$ arbitrarily from its domain. The number of choices for this is equal to

$$\left[\prod_{(p,r,m) \in \mathcal{G}(G)} \prod_{\substack{(q,s,l) \in \mathcal{G}(J) \\ q=p \\ (q,s,l) \neq (q^*, s^*, l^*)}} p^{\min(r,s)} \right]^n$$

For each $(p, r, m) \in \mathcal{G}(G^n)$, we need to have

$$(\tilde{a}_{q^*, s^*, l^*} - a_{q^*, s^*, l^*}) g_{(q^*, s^*, l^*) \rightarrow (p,r,m)} = \tilde{u}_{p,r,m} - u_{p,r,m} - \left(\sum_{\substack{(q,s,l) \in \mathcal{G}(J) \\ (q,s,l) \neq (q^*, s^*, l^*)}}^{\binom{\mathbb{Z}_{p^r}}{}} (\tilde{a}_{q,s,l} - a_{q,s,l}) g_{(q,s,l) \rightarrow (p,r,m)} \right)$$

Note that the right hand side is included in $p^{\hat{\theta}_{p,r,m}} \mathbb{Z}_{p^r}$ and $(\tilde{a}_{q^*, s^*, l^*} - a_{q^*, s^*, l^*})$ is included in $p^{\hat{\theta}_{q^*, s^*, l^*}} \mathbb{Z}_{(q^*)^{(s^*)}}$.

We need to count the number of solutions for $g_{(q^*, s^*, l^*) \rightarrow (p,r,m)}$ in $p^{|r-s^*|^+} \mathbb{Z}_{p^r}$. Using Lemma IX.1, we can show that the number of solutions is equal to $p^{\hat{\theta}_{q^*, s^*, l^*}}$. The total number of solutions for ϕ is equal to

$$\left[\left(\prod_{(p,r,m) \in \mathcal{G}} \prod_{\substack{(q,s,l) \in \mathcal{G}(J) \\ q=p \\ (q,s,l) \neq (q^*, s^*, l^*)}} p^{\min(r,s)} \right) \cdot p^{\hat{\theta}_{q^*, s^*, l^*}} \right]^n$$

Hence we have

$$\begin{aligned}
P(\phi(a) + B = u, \phi(\tilde{a}) + B = \tilde{u}) &= \frac{\left[\prod_{(p,r,m) \in \mathcal{G}(G)} \left(p^{\hat{\theta}_{q^*,s^*,l^*}} \cdot \prod_{\substack{(q,s,l) \in \mathcal{G}(J) \\ q=p \\ (q,s,l) \neq (q^*,s^*,l^*)}} p^{\min(r,s)} \right) \right]^n}{\left[\prod_{(p,r,m) \in \mathcal{G}(G)} \prod_{\substack{(q,s,l) \in \mathcal{G}(J) \\ q=p}} p^{\min(r,s)} \right]^n} \\
&= \left[\prod_{(p,r,m) \in \mathcal{G}(G)} \prod_{\substack{(q,s,l) \in \mathcal{G}(J) \\ q=p \\ (q,s,l) = (q^*,s^*,l^*)}} \frac{p^{\hat{\theta}_{q^*,s^*,l^*}}}{p^{\min(r,s)}} \right]^n
\end{aligned}$$

Note that for $(q, s, l) = (q^*, s^*, l^*)$ we have

$$\min(r, s) = \min(r, s^*) = r - |r - s^*|^+ = r - (\theta_{p,r,m} - \hat{\theta}_{q^*,s^*,l^*})$$

Therefore, the above probability is equal to

$$\begin{aligned}
\left[\prod_{(p,r,m) \in \mathcal{G}} \prod_{\substack{(q,s,l) \in \mathcal{G}(J) \\ q=p \\ (q,s,l) = (q^*,s^*,l^*)}} \frac{p^{\hat{\theta}_{q^*,s^*,l^*}}}{p^{r - (\theta_{p,r,m} - \hat{\theta}_{q^*,s^*,l^*})}} \right]^n &= \left[\prod_{(p,r,m) \in \mathcal{G}} \prod_{\substack{(q,s,l) \in \mathcal{G}(J) \\ q=p \\ (q,s,l) = (q^*,s^*,l^*)}} \frac{1}{p^{r - \theta_{p,r,m}}} \right]^n \\
&= \left[\prod_{(p,r,m) \in \mathcal{G}} \frac{p^{\theta_{p,r,m}}}{p^r} \right]^n = \frac{1}{|H_{\eta^* + \hat{\theta}}|^n}
\end{aligned}$$

Since the dither B is uniform, we conclude that

$$P \begin{pmatrix} \phi(u) + B = x \\ \phi(\tilde{u}) + B = \tilde{x} \end{pmatrix} = \frac{1}{|G|^n} \frac{1}{|H_{\eta^* + \hat{\theta}}|^n}$$

C. Proof of Lemma VII.2

First, we show that $(x + H^n) \cap A_\epsilon^n(X|y)$ is contained in $A_\epsilon^n(X|zy)$. Since z is a function of x , we have $(x, z, y) \in A_\epsilon^n(X, [X], Y)$. For $x' \in (x + H^n) \cap A_\epsilon^n(X|y)$, we have $[x'] = x' + H^n = x + H^n = z$ and $(x', z, y) = (x', [x'], y) \in A_\epsilon^n(X, [X], Y)$. Therefore, $x' \in A_\epsilon^n(X|zy)$ and hence,

$$(x + H^n) \cap A_\epsilon^n(X|y) \subseteq A_\epsilon^n(X|zy)$$

Conversely, for $x' \in A_\epsilon^n(X|zy)$, since $(x, z) \in A_\epsilon^n(X, [X])$ where $[X]$ is a function of X , we have $[x'] = z$. This implies $x' \in z + H^n = x + H^n$. Clearly, we also have $x' \in A_\epsilon^n(X|y)$. The claim on the size of the set follows since $(z, y) \in A_\epsilon^n([X]Y)$.

D. Useful Lemma

Lemma IX.1. *Let p be a prime and s, r a positive integer such that $s \leq r$. For $a \in \mathbb{Z}_{p^s}$ and $b \in \mathbb{Z}_{p^r}$, let $0 \leq \hat{\theta} \leq s$ and $\hat{\theta} \leq \theta \leq r$ be such that*

$$\begin{aligned} a &\in p^{\hat{\theta}}\mathbb{Z}_{p^s} \setminus p^{\hat{\theta}+1}\mathbb{Z}_{p^s} \\ b &\in p^{\theta}\mathbb{Z}_{p^r} \end{aligned}$$

Write $a = p^{\hat{\theta}}\alpha$ for some invertible element $\alpha \in \mathbb{Z}_{p^r}$ and $b = p^{\theta}\beta$ for some $\beta \in \beta \in \{0, 1, \dots, p^{r-\theta} - 1\}$. Then, the set of solutions to the equation $ax \pmod{p^r} = b$ is

$$\left\{ p^{\theta-\hat{\theta}}\alpha^{-1}\beta + i\alpha^{-1}p^{r-\hat{\theta}} \mid i = 0, 1, \dots, p^{\hat{\theta}} - 1 \right\}$$

Proof: Note that the representation of b as $b = p^{\theta}\beta$ is not unique and for any $\tilde{\beta}$ of the form $\tilde{\beta} = \beta + ip^{r-\theta}$ for $i = 0, 1, \dots, p^{\theta} - 1$, b can be written as $p^{\theta}\tilde{\beta}$. Also, the representation of a as $a = p^{\hat{\theta}}\alpha$ is not unique and for any $\tilde{\alpha} = \alpha + ip^{r-\hat{\theta}}$ for $i = 0, 1, \dots, p^{\hat{\theta}} - 1$, we have $a = p^{\hat{\theta}}\tilde{\alpha}$. The set of solutions to $ax = b$ is identical to the set of solutions to $p^{\hat{\theta}}x = p^{\theta}\alpha^{-1}\beta$. The set of solutions to the latter is

$$\left\{ p^{\theta-\hat{\theta}}\alpha^{-1}\beta + i\alpha^{-1}p^{r-\hat{\theta}} \mid i = 0, 1, \dots, p^{\hat{\theta}} - 1 \right\}$$

It remains to show that this set of solutions is independent of the choice of α and β . First, we show that the set of solutions is independent of the choice of β . For $\tilde{\beta} = \beta + jp^{r-\theta}$ for some $j \in \{0, 1, \dots, p^{\theta} - 1\}$, we have

$$\begin{aligned} &\left\{ p^{\theta-\hat{\theta}}\alpha^{-1}\tilde{\beta} + i\alpha^{-1}p^{r-\hat{\theta}} \mid i = 0, 1, \dots, p^{\hat{\theta}} - 1 \right\} \\ &= \left\{ p^{\theta-\hat{\theta}}\alpha^{-1}(\beta + jp^{r-\theta}) + i\alpha^{-1}p^{r-\hat{\theta}} \mid i = 0, 1, \dots, p^{\hat{\theta}} - 1 \right\} \\ &= \left\{ p^{\theta-\hat{\theta}}\alpha^{-1}\beta + (i+j)\alpha^{-1}p^{r-\hat{\theta}} \mid i = 0, 1, \dots, p^{\hat{\theta}} - 1 \right\} \\ &\stackrel{(a)}{=} \left\{ p^{\theta-\hat{\theta}}\alpha^{-1}\beta + i\alpha^{-1}p^{r-\hat{\theta}} \mid i = 0, 1, \dots, p^{\hat{\theta}} - 1 \right\} \end{aligned}$$

where (a) follows since the set $p^{r-\hat{\theta}}\{0, 1, \dots, p^{\hat{\theta}} - 1\}$ is a subgroup of \mathbb{Z}_{p^r} and $jp^{r-\hat{\theta}}$ lies in this set.

Next, we show that the set of solutions is independent of the choice of α . For $\tilde{\alpha} = \alpha + jp^{r-\hat{\theta}}$ for some $j \in \{0, 1, \dots, p^{\hat{\theta}} - 1\}$, we have

$$\tilde{\alpha} \left(\alpha^{-1} - \alpha^{-1}jp^{r-\hat{\theta}}\tilde{\alpha}^{-1} \right) = 1$$

Therefore, it follows that the unique inverse of $\tilde{\alpha}$ satisfies $\alpha^{-1} - \tilde{\alpha}^{-1} \in \alpha^{-1}p^{r-\hat{\theta}}\mathbb{Z}_p$. Assume $\tilde{\alpha}^{-1} = \alpha^{-1} + k\alpha^{-1}p^{r-\hat{\theta}}$. We have,

$$\begin{aligned} & \left\{ p^{\theta-\hat{\theta}}\tilde{\alpha}^{-1}\beta + i\tilde{\alpha}^{-1}p^{r-\hat{\theta}} \mid i = 0, 1, \dots, p^{\hat{\theta}} - 1 \right\} \\ &= \left\{ p^{\theta-\hat{\theta}} \left(\alpha^{-1} + k\alpha^{-1}p^{r-\hat{\theta}} \right) \beta + i \left(\alpha^{-1} + k\alpha^{-1}p^{r-\hat{\theta}} \right) p^{r-\hat{\theta}} \mid i = 0, 1, \dots, p^{\hat{\theta}} - 1 \right\} \\ &= \left\{ p^{\theta-\hat{\theta}}\alpha^{-1}\beta + \left(i + ikp^{r-\hat{\theta}} + k\beta p^{\theta-\hat{\theta}} \right) \alpha^{-1}p^{r-\hat{\theta}} \mid i = 0, 1, \dots, p^{\hat{\theta}} - 1 \right\} \\ &\stackrel{(a)}{=} \left\{ p^{\theta-\hat{\theta}}\alpha^{-1}\beta + i\alpha^{-1}p^{r-\hat{\theta}} \mid i = 0, 1, \dots, p^{\hat{\theta}} - 1 \right\} \end{aligned}$$

where same as above, (a) follows since the set $p^{r-\hat{\theta}}\{0, 1, \dots, p^{\hat{\theta}} - 1\}$ is a subgroup of \mathbb{Z}_p and $(ikp^{r-\hat{\theta}} + k\beta p^{\theta-\hat{\theta}})p^{r-\hat{\theta}}$ lies in this set. \blacksquare

REFERENCES

- [1] T. J. Goblick, Jr., "Coding for a discrete information source with a distortion measure", *Ph.D. dissertation*, Dept. Electr. Eng., MIT, Cambridge, MA, 1962.
- [2] H. A. Loeliger and T. Mittelholzer, "Convolutional codes over groups", *IEEE Trans. Inform. Theory*, vol. 42, no. 6, pp. 1660–1686, November 1996.
- [3] H. A. Loeliger, "Signal sets matched to groups", *IEEE Trans. Inform. Theory*, vol. 37, no. 6, pp. 1675–1682, November 1991.
- [4] I. Csiszar, "Linear codes for sources and source networks: Error exponents, universal coding," *IEEE Trans. Inform. Theory*, vol. IT- 28, pp. 585–592, July 1982.
- [5] S. Shridharan, A. Jafarian, S. Vishwanath, and S. A. Jafar, "Lattice Coding for K User Gaussian Interference Channels", *IEEE Transactions on Information Theory*, April 2010. .
- [6] R. Ahlswede. Group codes do not achieve Shannons's channel capacity for general discrete channels. *The annals of Mathematical Statistics*, 42(1):224–240, Feb. 1971.
- [7] R. Ahlswede and J. Gemma. Bounds on algebraic code capacities for noisy channels I. *Information and Control*, 19(2):124–145, 1971.
- [8] R. Ahlswede and J. Gemma. Bounds on algebraic code capacities for noisy channels II. *Information and Control*, 19(2):146–158, 1971.
- [9] N. J. Bloch. *Abstract Algebra With Applications*. Prentice-Hall, Inc, Englewood Cliffs, New Jersey, 1987.
- [10] G. Como. Group codes outperform binary coset codes on non-binary symmetric memoryless channels. *IEEE Trans. Information Theory*, 56(9):4321–4334, Sept. 2010.
- [11] G. Como and F. Fagnani. The capacity of finite abelian group codes over symmetric memoryless channels. *IEEE Transactions on Information Theory*, 55(5):2037–2054, 2009.
- [12] J. H. Conway and N. J. A. Sloane. *Sphere Packings, Lattices and Groups*. Springer-Verlag, New York, 1988.
- [13] I. Csiszar and J. Korner. *Information theory: Coding theorems for Discrete memoryless Systems*. 1981.
- [14] R. L. Dobrushin. Asymptotic optimality of group and systematic codes for some channels. *Theor. Probab. Appl.*, 8:47–59, 1963.

- [15] R. L. 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*, pages 283–311, 1962.
- [16] P. Elias. Coding for noisy channels. *IRE Conv. Record*, part. 4:37–46, 1955.
- [17] U. Erez and S. tenBrink. A close-to-capacity dirty paper coding scheme. *IEEE Trans. Inform. Theory*, 51:3417–3432, October 2005.
- [18] S. Litsyn G. Cohen, I. Honkala and A. Lobstein. *Covering Codes*. Elsevier-North-Holland, Amsterdam, 1997.
- [19] R. Garello and S. Benedetto. Multilevel construction of block and trellis group codes. *IEEE Trans. Inform. Theory*, 41:1257–1264, Sep. 1995.
- [20] G. D. Forney Jr and M. Trott. The dynamics of group codes: State spaces, trellis diagrams, and canonical encoders. *IEEE Transactions on Information Theory*, 39(9):1491–1513, 1993.
- [21] M. Hall Jr. *The Theory of Groups*. The Macmillan Company, New York, 1959.
- [22] J. Korner and K. Marton. How to encode the modulo-two sum of binary sources. *IEEE Transactions on Information Theory*, IT-25:219–221, Mar. 1979.
- [23] D. Krithivasan and S. S. Pradhan. Distributed source coding using abelian group codes. 2011. *IEEE Transactions on Information Theory*(57)1495-1519.
- [24] F. J. MacWilliams and N. J. A. Sloane. *The theory of error-correcting codes*. Elsevier-North-Holland, 1977.
- [25] B. A. Nazer and M. Gastpar. Computation over multiple-access channels. *IEEE Transactions on Information Theory*, 53(10):3498–3516, Oct. 2007.
- [26] A. Padakandla and S.S. Pradhan. A new coding theorem for three user discrete memoryless broadcast channel. 2012. Online: <http://128.84.158.119/abs/1207.3146v2>.
- [27] A. Padakandla, A.G. Sahebi, and S.S. Pradhan. A new achievable rate region for the 3-user discrete memoryless interference channel. In *Information Theory Proceedings (ISIT), 2012 IEEE International Symposium on*, pages 2256–2260, 2012.
- [28] A. Padakandla, A.G. Sahebi, and S.S. Pradhan. An achievable rate region for the 3-user interference channel based on coset codes. 2014. Online: <http://arxiv.org/abs/1403.4583>.
- [29] W. Park and A. Barg. Polar codes for q -ary channels, $q = 2^r$. 2012. Online: <http://arxiv.org/abs/1107.4965>.
- [30] T. Philosof and R. Zamir. On the loss of single-letter characterization: The dirty multiple access channel. *IEEE Transactions on Information Theory*, 55:2442–2454, June 2009.
- [31] S. S. Pradhan and K. Ramchandran. Distributed source coding using syndromes (DISCUS): Design and construction. *IEEE Transactions on Information Theory*, 49(3):626–643, 2003.
- [32] A. G. Sahebi and S. S. Pradhan. Multilevel Polarization of Polar Codes Over Arbitrary Discrete Memoryless Channels. *Proc. 49th Allerton Conference on Communication, Control and Computing*, Sept. 2011.
- [33] E. Sasoglu, E. Telatar, and E. Arikan. Polarization for arbitrary discrete memoryless channels. *IEEE Information Theory Workshop*, Dec. 2009. Lausanne, Switzerland.
- [34] D. Slepian. Group codes for for the Gaussian channel. *Bell Syst. Tech. Journal*, 1968.
- [35] D. Slepian and J. K. Wolf. A coding theorem for multiple access channels with correlated sources. *bell Syst. tech. J.*, 52:1037–1076, 1973.
- [36] A. D. Wyner. Recent results in the Shannon theory. *IEEE Trans. Inform. Theory*, IT-20:2–10, January 1974.
- [37] R. Zamir, S. Shamai, and U. Erez. Nested linear/lattice codes for structured multiterminal binning. *IEEE Trans. Inform. Theory*, 48:1250–1276, June 2002.