DRDO–IISc Programme on
Advanced Research in Mathematical
Engineering

On the Compressibility of non-Abelian-Group Sources in a Distributed Source Coding Setting

(TR-PME-2010-8)

by

V. Lalitha[1], N. Prakash[1], K. Vinodh[1], P. Vijay Kumar[1] and S. S. Pradhan[2]

[1]Department of ECE, Indian Institute of Science, Bangalore
[2] Dept. of EECS, University of Michigan, Ann Arbor, USA.

February 8, 2010

Indian Institute of Science
Bangalore 560 012

On the Compressibility of non-Abelian-Group Sources in a Distributed Source Coding Setting

V. Lalitha, N. Prakash, K. Vinodh, P. Vijay Kumar and S. S. Pradhan

February 8, 2010

ABSTRACT

We consider the problem of compression of a non-Abelian source. This is motivated by the problem of distributed function computation, where it is known that if one is only interested in computing a function of several sources, then one can often improve upon the compression rate required by the Slepian-Wolf bound. Let G be a non-Abelian group having center $\mathcal{Z}(G)$. We show here that it is impossible to compress a source with symbols drawn from G when $\mathcal{Z}(G)$ is trivial if one employs a homomorphic encoder and a typical-set decoder. We provide achievable upper bounds on the minimum rate required to compress a non-Abelian group with non-trivial center. Also, in a two source setting, we provide achievable upper bounds for compression of any non-Abelian group, using a non-homomorphic encoder.

1 Introduction

Distributed source coding is motivated by various applications such as distributed video coding[1] and sensor networks. In such applications, the receiver is often interested in computing a function of the sources rather than in the exact reconstruction of all the sources. Let X_1 and X_2 be two sources having a finite alphabet $\mathcal{X}$ and joint distribution $P_{X_1X_2}$ and let $f(x_1, x_2)$ be the function of interest. Using Slepian-Wolf compression [2], one can achieve lossless compression at a sum rate $H(X_1, X_2)$. Now, consider the case when the function computation can be embedded within a group, i.e., it is possible to associate with every element in the source alphabet, an element in a finite group G such that $f(x_1, x_2) = g \circ h$, where $g, h \in G$ and where multi-

plication is carried out in the group G. The use of homomorphic encoders permits one to compress the function $f(x_1, x_2)$ by compressing the individual sources. A homomorphic encoder is an encoder employing a mapping ϕ where ϕ is a group homomorphism, i.e., $\phi(g \circ h) = \phi(g) \circ \phi(h)$. Thus, by compressing g, h using separate homomorphic encoders ϕ at the two sources and having the receiver compute the product $\phi(g) \circ \phi(h)$, we have in effect achieved distributed compression of the function $f(x_1, x_2)$. For a large class of groups, we can achieve a sum rate, that for certain distributions, improves upon the Slepian-Wolf bound, by using such a homomorphic encoder [3] [4]. We begin a with few examples.

Example 1 *[4] Let the alphabet $\mathcal{X}$ be the set of integers modulo 2, i.e., $\mathcal{X} = \mathbb{Z}_2 = \{0, 1\}$ and let the function that needs to be computed be given by $f(x_1, x_2) = x_1 \oplus x_2$. Assume the sources to have joint distribution given by $P(0,0) = P(1,1) = p/2, P(0,1) = P(1,0) = (1-p)/2,\ 0 < p < 1/2$. In this case, the sum rate required using Slepain-Wolf encoding would equal $H(X_1, X_2) = 1 + h(p)$. But if the compression is done homomorphically, then the required sum rate equals $2h(p)$ which is less than $H(X_1, X_2)$ for $0 < p < \frac{1}{2}$.*

Example 2 *An example where the function could be embedded in an Abelian group is when the source is interested in the average of N sources, i.e., $f(x_1, \ldots, x_N) = \frac{1}{N} \sum_{i=1}^{N} x_i$. If the source alphabet $\mathcal{X} = \{0, \ldots, M\}$, then the $\sum_{i=1}^{N} x_i$ is equivalent to the addition operation in the group $\mathbb{Z}_q$, where q is a prime greater than MN. For example, with two sources, let $\mathcal{X} = \{0, 1\}$ and $f(x_1, x_2) = (x_1 + x_2)/2$. The function $g(x_1, x_2) = 2f(x_1, x_2)$ can be embedded into $\mathbb{Z}_3$. In distributed compression with homomorphic encoders, the receiver first recovers $g(x_1, x_2)$ and divides it by a factor of 2 to obtain $f(x_1, x_2)$. Consider the same probability distribution on the source alphabet as in Example 1. The sum rate required using homomorphic compression is $2(h(p) + p)$, which, for p such that $h(p) < 1 - 2p$, improves upon the Slepian-Wolf scheme.*

Example 3 *Consider next, the case where the alphabet $\mathcal{X}$ is the set $GL_2(\mathbb{F}_q)$ of 2×2 invertible matrices over the finite field $\mathbb{F}_q$. Then under the binary operation corresponding to matrix multiplication, the set $\mathcal{X} = GL_2(\mathbb{F}_q)$ forms a non-Abelian group of size $(q^2 - 1)(q^2 - q)$. The center of this group is the set of all non-zero scalar multiples of the identity matrix and is hence of size $(q - 1)$. Thus for example, when $q = 3$, the group contains 48 elements*

and the center of the group consists of the elements $I_2,\ 2I_2$ *where* I_2 *is the* (2×2) *identity matrix. Let* A_1, A_2 *be the outputs of two sources whose alphabet* $\mathcal{X} = GL_2(\mathbb{F}_3)$ *and let the function that needs to be computed be given by* $f(A_1, A_2) = A_1A_2$. *Let* $\{C_1, C_2, \ldots, C_{24}\}$ *be the set of coset representatives of* $\{I_2,\ 2I_2\}$ *in* $GL_2(\mathbb{F}_3)$. *The sources are assumed to have a joint distribution given by*

$$P(C_i, C_j) = P(2C_i, 2C_j) = \left(\frac{1}{24^2}\right)\frac{p}{2}, \tag{1}$$

$$P(C_i, 2C_j) = P(2C_i, C_j) = \left(\frac{1}{24^2}\right)\frac{(1-p)}{2}, \tag{2}$$

where $1 \leq i, j \leq 24,\ 0 < p < 1/2$. *Slepian-Wolf encoding in this case would require a sum rate of* $2log_2 24 + 1 + h(p)$. *With homomorphic compression, a sum rate of* $2log_2 24 + 2h(p)$ *is required (see Section 4.3), which is better than the Slepian-Wolf rate.*

The case when the function computation corresponds to a group operation in an Abelian group has recently been addressed by Krithivasan and Pradhan [5]. The properties of non-Abelian groups have been studied in the context of group error correcting codes in [6, 7].

2 System Model

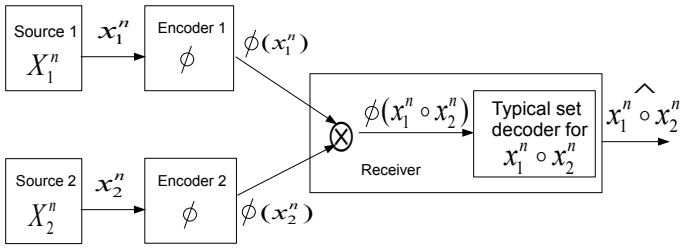

Figure 1: System Model

We consider a distributed compression problem (See Fig 1) involving two correlated but memoryless sources and a receiver that is only interested in computing a function of the two sources, in a lossless manner. We assume further that homomorphic encoders are employed. Though the system model is described here in with only two sources, it can be directly extended to any

number of sources when the encoding functions are homomorphisms. The source alphabet is assumed to be a finite non-Abelian group G. Let $X_i^n = \{X_i(1) \dots X_i(n)\}$ denote the random variables corresponding to an $n-$length output sequence of the i^{th} source, $i = 1, 2$. We assume $(X_1(i), X_2(i)), 1 \leq i \leq n$ to be jointly distributed according to P_{X_1,X_2}. Also, for a fixed i, $X_i(j)$ is assumed to be independent of $X_i(k)$, $j \neq k$. We shall denote the realization of the random variables X_i^n by x_i^n. The function that needs to be computed at the receiver is $y(i) = x_1(i) \circ x_2(i), 1 \leq i \leq n$, in which the multiplication is carried out in the group G. The corresponding random variables, $\{Y(i)\}$, are independent and identically distributed as

$$P_Y(y) = \sum_{x_1,x_2:x_1 \circ x_2 = y} P_{X_1X_2}(x_1, x_2). \tag{3}$$

Encoder : Since we restrict our attention to homomorphic encoders, encoding of both sources is carried out using the group homomorphism $\phi^{(n)}$,

$$\phi^{(n)} : G_1 \times \ldots \times G_n \longrightarrow H^{(n)}, \tag{4}$$

where $G_i = G, \ \forall i$ and $H^{(n)}$ denotes the range of the homomorphism. For example, $H^{(n)}$ could be G^k, where $k = \lceil \alpha n \rceil, 0 \leq \alpha \leq 1$ where α may be viewed as a crude measure of the amount of compression taking place. Let $\phi^{(n)}(x_1^n)$ and $\phi^{(n)}(x_2^n)$ denote the output of the two encoders.

Receiver : Since the function of interest corresponds to the multiplication of two elements in the non-Abelian group, as noted in the introduction, the first step taken by the receiver is to multiply the outputs of the two encoders to obtain

$$\phi^{(n)}(x_1^n) \circ \phi^{(n)}(x_2^n) = \phi^{(n)}(x_1^n \circ x_2^n) = \phi^{(n)}(y^n), \tag{5}$$

where (5) follows since $\phi^{(n)}$ is a homomorphism. The decoder declares $\hat{y}^n$ as the output if there exists a unique $\hat{y}^n$ such that $\hat{y}^n \in A_\epsilon^{(n)}(Y)$ and $\phi^{(n)}(\hat{y}^n) = \phi^{(n)}(y^n)$, where $A_\epsilon^{(n)}(Y)$ denotes the $\epsilon-$strongly typical set, defined as follows.

$$A_\epsilon^{(n)}(Y) = \left\{ y^n \in G^n : \left| \frac{N(g|y^n)}{n} - P_Y(g) \right| \leq \epsilon, \ \forall g \in G \right\}, \tag{6}$$

where $N(g|y^n)$ denotes the number of occurrences of $g \in G$ in y^n. If no such $\hat{y}^n$ is found, the decoder declares an error. Throughout this paper, we will

make the assumption of a typical set decoder as described above.

Error event : For large n, $y^n \in A_\epsilon^{(n)}(Y)$ with high probability. Thus, error occurs if there exists a sequence $\tilde{y}^n (\neq y^n) \in G^n$ such that $\tilde{y}^n \in A_\epsilon^{(n)}(Y)$ and $\phi^{(n)}(\tilde{y}^n) = \phi^{(n)}(y^n)$. Equivalently, an error occurs if there exists a sequence $a^n \in \mathrm{Ker}(\phi^{(n)})$ such that $y^n \circ a^n \in A_\epsilon^{(n)}(Y)$. Let $P_e^{(n)}$ denote the probability of error averaged over all the source sequences.

Rate: The rate of compression of each source in bits per symbol is given by

$$R^{(n)} = \frac{\log_2 |\mathrm{Im}(\phi^{(n)})|}{n}. \tag{7}$$

Achievability : A compression rate R is said to be achievable under the above encoder and decoder, if $\forall \delta > 0$, there exists a sequence of maps $\{\phi^{(n)}\}$ such that for sufficiently large n, $R^{(n)} < R + \delta$ and $\lim_{n\to\infty} P_e^{(n)} = 0$.

Remark 1 *Although no side-information is considered in the receiver described above, any side-information s (generated according to some distribution P_S) available at the receiver can be taken into account by using the conditional typical set $A_\epsilon^{(n)}(Y|s)$ in place of $A_\epsilon^{(n)}(Y)$ while decoding.*

3 Compression of Abelian groups

Although the focus of the paper is on the compression of non-Abelian groups, we present a summary of some known results on the homomorphic compression of Abelian groups in this section as we shall subsequently have the need of these results.

Theorem 1 *[8]* [Primary decomposition theorem] *Every finite Abelian group G is the direct product of primary cyclic groups, i.e.,*

$$G \;=\; \mathbb{Z}_{p_1^{r1}} \times \mathbb{Z}_{p_2^{r2}} \times \ldots \times \mathbb{Z}_{p_m^{rm}}, \tag{8}$$

where p_i is prime, $r_i \geq 1,\ 1 \leq i \leq m$.

The compression of Abelian group sources via an approach that makes use of the above primary decomposition theorem is studied in [3]. This paper identifies achievable rates, given below, for the homomorphic compression of the primary cyclic groups $\mathbb{Z}_{p^r}$.

Theorem 2 *[3] Suppose Y is random variable over the group $\mathbb{Z}_{p^r}$. Let P_Y be the distribution of Y such that $P_Y(y) > 0$ for at least one symbol*

$y \in \mathbb{Z}_{p^r} \backslash p\mathbb{Z}_{p^r}$. Then there exists a sequence of homomorphic encoders $\{\phi^{(n)} : \mathbb{Z}_{p^r}^n \to \mathbb{Z}_{p^r}^{k(n)}\}$ that achieves the rate

$$R = \lim_{n\to\infty} \frac{k(n)}{n} log_2 p^r \geq \max_{0\leq i<r} \left(\frac{r}{r-i}\right)(H(Y) - H([Y]_i)), \tag{9}$$

where $[Y]_i$ is a random variable taking values over the set of all distinct cosets $C_{ij} := j + p^i\mathbb{Z}_{p^r}$ of $p^i\mathbb{Z}_{p^r}$ in $\mathbb{Z}_{p^r}$ and where $[Y]_i = C_{i\ell}$ if $Y \in \ell + p^i\mathbb{Z}_{p^r}$ with probability distribution that is induced by the distribution of Y. For example, if the group is $\mathbb{Z}_4$, then $[Y]_1$ is a binary random variable defined on the cosets of $2\mathbb{Z}_4$ in $\mathbb{Z}_4$, i.e, defined on the set $\{\{0,2\},\{1,3\}\}$ with associated symbol probabilities $(P_Y(0) + P_Y(2), P_Y(1) + P_Y(3))$. Note that if $r = 1$, $R = H(Y)$ is achievable.

For a general Abelian group as given in (8), at the encoder, a sequence $y^n \in G^n$ is decomposed into its direct product components $(y_1^n, \ldots, y_k^n)$, $y_i^n \in \mathbb{Z}_{p_i^{r_i}}, 1 \leq i \leq m$. Let Y_i denote the random variable corresponding to the i^{th} component. Then the distribution P_Y can be factored as $P_Y = \prod_{i=1}^m P_{Y_i|Y_1\ldots Y_{i-1}}$. The compression takes place in m stages. In the i^{th} stage, the component y_i^n is encoded using the homomorphism $\phi_i : \mathbb{Z}_{p_i^{r_i}}^n \to \mathbb{Z}_{p_i^{r_i}}^{k(n)}$, while assuming that all the previous components have been decoded successfully at the receiver. Thus the quantity $s = (\hat{y}_1^n, \ldots, \hat{y}_{i-1}^n)$ represents the side-information available to the decoder at the i^{th} stage. The decoder searches for a unique $\hat{y}_i^n$ such that $\phi_i(\hat{y}_i^n) = \phi_i(y_i^n)$ and $\hat{y}_i^n \in A_\epsilon^{(n)}(Y_i|s)$. In the situation, the rate achievable in the i^{th} stage is given by [3]

$$R_i \geq \max_{0\leq j<r} \left(\frac{r}{r-j}\right)(H(Y_i|S) - H([Y_i]_j)|S), \tag{10}$$

where S denotes the random variable corresponding to the side information. The rate of compression for the group G in (8) in terms of the compression rates of the component groups is given by $R = \sum_{i=1}^m R_i$.

4 Compression of non-Abelian groups

Unlike in the case of Abelian groups, it is not at all clear that any homomorphic compression is possible in the case of a general non-Abelian group. We begin with an investigation into the possibility of compression under

a fixed homomorphic encoder ϕ. We obtain two necessary conditions for compression to be possible:

1. the map ϕ when restricted to any finite set of input co-ordinates must represent an isomorphism of groups and

2. the compression rate must satisfy the lower bound:

$$R \geq \log_2 \frac{|G|}{|\mathcal{Z}(G)|}. \tag{11}$$

It follows from the second condition that homomorphic source compression is not possible in the case of non-Abelian groups G having a trivial center $\mathcal{Z}(G)$. We say compression is not possible if the necessary condition for achievable rate is $R \geq \log_2|G|$. For the case when the group does possess a non-trivial center $\mathcal{Z}(G)$, we divide the discussion into two cases. In the first case, the group G is assumed to be the direct product $G = P \times Q$ of an Abelian group P and a non-Abelian group Q. Here we show that compression is indeed possible using a homomorphic encoder. In the second case, when such a decomposition is not possible, we show that all is not lost. We show how one can achieve compression by making use of an encoder that is "almost" homomorphic. Both these compression methods, though will be discussed only in the context of two sources, can be extended to any finite number of sources. Finally, we give a third compression strategy applicable only for a two source setting, but which will potentially achieve much better rates than both the previous strategies.

We begin by describing certain properties of the homomorphic encoder.

4.1 Properties of the encoder

Consider the homomorphic encoder described in (4). Let $\phi_i^{(n)}$ be the restriction of ϕ to the i^{th} copy of G, i.e.

$$\phi_i^{(n)} = \phi^{(n)}|_{1\times\ldots\times G_i\times\ldots\times 1} \colon G_i \twoheadrightarrow H_i, \tag{12}$$

with $H_i = \mathrm{Im}(\phi_i^{(n)})$. Then for any $y^n \in G^n$, we have

$$\phi^{(n)}(y^n) = \prod_{i=1}^{n} \phi_i^{(n)}(y(i)). \tag{13}$$

Property 1 *$\forall i, j, i \neq j,\ H_i, H_j$ commute element-wise; i.e. $h_i \circ h_j = h_j \circ h_i,\ h_i \in H_i,\ h_j \in H_j$.*

This follows from the fact that the groups $1 \times \ldots \times G_i \times \ldots \times 1$ and $1 \times \ldots \times G_j \times \ldots \times 1$ commute element-wise and hence the same holds for their homomorphic images. The property also implies that $H_k \lhd \prod_{i=1}^{n} H_i,\ \forall k$.

Property 2 *$Im(\phi^{(n)}) = \prod_{i=1}^{n} H_{\pi(i)}$ where π is any permutation of $\{1, 2, \ldots, n\}$.*

The fact follows from (13) and Property 1.

Property 3 *If $\mathcal{I}, \mathcal{J} \subseteq \{1, \ldots, n\}$ such that $\mathcal{I}, \mathcal{J} \neq \Phi$ (the null set) and $\mathcal{I} \cap \mathcal{J} = \Phi$, then*

$$\prod_{i \in \mathcal{I}} H_i \bigcap \prod_{j \in \mathcal{J}} H_j = \mathcal{Z}\left(\prod_{i \in \mathcal{I}} H_i\right) \bigcap \mathcal{Z}\left(\prod_{j \in \mathcal{J}} H_j\right). \tag{14}$$

This follows from Property 1.

4.2 Necessary conditions for compressibility

We begin with a useful lemma regarding typical sets. We omit the proof due to lack of space.

Lemma 3 *Consider a random variable U on the group G, distributed according to P_U. Let $u^n \in A_{\epsilon-\delta}^{(n)}(U),\ 0 \leq \delta < \epsilon$, and $a^n \in G^n$ such that $wt(a^n) = r$, where $wt(a^n)$ is the number of non identity components in a^n. If r is such that $lim_{n\to\infty} \frac{r}{n} = 0$, then for large n, $u^n \circ a^n \in A_{\epsilon}^{(n)}(U)$.*

Theorem 4 *Consider the subgroup K of G^n, where $K = 1 \times \ldots \times G_{i_1} \times \ldots \times G_{i_2} \times \ldots \times G_{i_r} \times \ldots \times 1$, with $1 \leq i_1 < i_2 < \ldots < i_r \leq n$. If r is such that $lim_{n\to\infty} \frac{r}{n} = 0$, then for large n, the restriction of $\phi^{(n)}$ to K must necessarily be an isomorphism for compression to be possible under a typical set decoder assuming a homomorphic encoder.*

Proof: The proof follows directly by method of contradiction and an application of Lemma 3. Without loss of generality, assume $K = G_1 \times \ldots \times G_r$. Let $H_K \triangleq \mathrm{Im}(\phi^{(n)}|_K) = \prod_{j=1}^{r} H_{i_j}$. Assume $\phi^{(n)}|_K$ is not an isomorphism; which implies $\exists b^r (\neq 1_G^r) \in G^r$ such that $\phi^{(n)}|_K(b^r) = 1_{H_K}$. Consider the sequence $a^n = [b^r \ \ 1_G^{n-r}]$. Clearly, $a^n \in \mathrm{Ker}(\phi^{(n)})$. Let $A_{\epsilon}^{(n)}(Y)$ be the typical set corresponding to the source distribution, and hence the typical set used for decoding. Fix a δ such that $0 \leq \delta < \epsilon$. The average probability of error

can be lower bounded as

$$P_e^{(n)} \geq \sum_{y^n \in A_{\epsilon-\delta}^{(n)}(Y)} P_{Y^n}(y^n) P_{e|y^n}^{(n)}. \tag{15}$$

Now, since $\mathrm{wt}(a^n) = r$ such that $\lim_{n\to\infty} \frac{r}{n} = 0$, by Lemma 3, $y^n \in A_{\epsilon-\delta}^{(n)}(Y) \Rightarrow y^n \circ a^n \in A_{\epsilon}^{(n)}(Y)$. Thus both y^n and $y^n \circ a^n$ are typical. Also since $\phi^{(n)}(y^n) = \phi^{(n)}(y^n \circ a^n)$, $P_{e|y^n}^{(n)} = 1$, under a typical set decoder. Thus the bound on error probability becomes

$$P_e^{(n)} \geq P\left(A_{\epsilon-\delta}^{(n)}(Y)\right) \to 1, \text{ as } n \to \infty. \tag{16}$$

which implies compression is impossible if $\phi^{(n)}|_K$ is not an isomorphism. □

By setting $r = 1$ in Theorem (4), we get $H_i \cong G,\ 1 \leq i \leq n$, for large n. We will make use of this fact below, where we establish the necessity of second condition given in (11).

Theorem 5 *For a group G with center $\mathcal{Z}(G)$, compression rates less than $log_2 \frac{|G|}{|\mathcal{Z}(G)|}$ cannot be achieved using homomorphic encoders. Specifically, if $\mathcal{Z}(G) = \{1_G\}$, no compression is possible.*

Proof: Consider the homomorphic encoder given in (4). From (7) and property 2, the rate corresponding to the encoder $\phi^{(n)}$ in (4) is given by

$$R^{(n)} = \frac{\log_2 |\prod_{i=1}^n H_i|}{n}. \tag{17}$$

Using the fact that for any two finite groups A and B $|AB| = (|A|\,|B|)/(|A \cap B|)$, the cardinality of the image, $\prod_{i=1}^n H_i$, of $\phi^{(n)}$ can be lower bounded by,

$$\left|\prod_{i=1}^n H_i\right| = \frac{|H_1|\,|\prod_{i=2}^n H_i|}{|H_1 \cap \prod_{i=2}^n H_i|} \tag{18}$$

$$= \frac{|H_1|\,|\prod_{i=2}^n H_i|}{|\mathcal{Z}(H_1) \cap \mathcal{Z}(\prod_{i=2}^n H_i)|} \tag{19}$$

$$\geq \frac{|H_1|\,|\prod_{i=2}^{n} H_i|}{|\mathcal{Z}(H_1)|} \tag{20}$$

$$= \frac{|G|\,|\prod_{i=2}^{n} H_i|}{|\mathcal{Z}(G)|} \tag{21}$$

$$\geq \frac{|G|^n}{|\mathcal{Z}(G)|^{n-1}}\ , \tag{22}$$

where (19) follows from Property 3, (21) follows since, for compression to be possible, Theorem 4 implies that $H_i \cong G,\ \forall i$. Combining equation (17) and (22), we get

$$R^{(n)} \geq \log_2 \frac{|G|}{|\mathcal{Z}(G)|} + \frac{1}{n}\log_2|\mathcal{Z}(G)|, \tag{23}$$

from which, using the definition of achievability of rate R, it can be shown that no rate less than $\log\frac{|G|}{|\mathcal{Z}(G)|}$ is achievable. □

Theorem (5) rules out homomorphic compression of many of the commonly known non-Abelian groups such as the dihedral-group, D_{2n}, for n odd, the symmetric group S_n, for $n \geq 3$, and the alternating group, A_n, for $n \geq 4$, and all non-abelian simple groups, since all of the above groups have trivial centers. See [8] for a discussion on these groups.

4.3 Achievable rates for non-Abelian groups

We separate the discussion on the compression of non-Abelian groups into three cases. The first two cases are applicable in distributed source coding setting with any number of sources, whenever the center $\mathcal{Z}(G)$ of G is non-trivial. We then describe a third case, which will give a compression strategy for any non-Abelian group, applicable to the two-source distributed source-coding setting.

Case 1 : Here in this case, we consider the compression of non-Abelian group G that can be decomposed as a direct product of groups, at least one of which is Abelian. Suppose G is decomposable as

$$G \cong B_1 \times \ldots \times B_s \times A_1 \times \ldots \times A_t, \tag{24}$$

where $B_i, 1 \leq i \leq s$, are the non-Abelian components and $A_j, 1 \leq j \leq t$, are the Abelian components. Then

$$G^n \cong B_1^n \times \ldots \times B_s^n \times A_1^n \times \ldots \times A_t^n, \tag{25}$$

Note that from Theorem 1, we can assume without loss of generality that $A_j \cong \mathbb{Z}_{p^r}$, for some p, r. Also note that $\mathcal{Z}(G)$ is non-trivial due to the presence of the Abelian components. Now, suppose that $y^n (\in G^n) = (b_1^n, \ldots, b_s^n, a_1^n, \ldots, a_t^n)$ needs to be compressed. The proposed scheme is simple: the non-Abelian components are transmitted uncompressed, while the Abelian components are compressed as discussed in Section 3. The compression then takes place in $t+1$ stages, where, in the first stage, all the non-Abelian components $(b_1^n, \ldots, b_s^n)$ are sent as such. In the k^{th} stage, $k > 1$, the Abelian component a_{k-1}^n is compressed assuming that the receiver has successfully decoded all the previous components and hence can use them as side-information. Let R_k denote the rate of compression achieved in the k^{th} stage. Here $R_1 = \sum_{i=1}^{s} \log_2|B_i|$ and R_k for $k > 1$ is calculated using (10). Then an achievable rate of compression for the group G is given by $R = \sum_{k=1}^{t+1} R_k$.

Example 4 *Consider the distributed source compression scenario with two sources as in Figure 1. Let $G \cong B \times C_2$, where B is non-Abelian and C_2 is the cyclic group with two elements. The random variable $X_1 = (X_{1B}, X_{1C_2})$ and $X_2 = (X_{2B}, X_{2C_2})$. Let $X_{1B} \perp \{X_{1C_2}, X_{2B}, X_{2C_2}\}$ and $X_{2B} \perp \{X_{1C_2}, X_{1B}, X_{2C_2}\}$. Let both X_{1B} and X_{2B} be distributed as P_B. Let X_{1C_2} and X_{2C_2} be jointly distributed as in Example 1; i.e. $P(0,0) = P(1,1) = p/2, P(0,1) = P(1,0) = (1-p)/2$. The function to be computed in the receiver is $y = x_1 \circ x_2 = (x_{1B}x_{2B}, x_{1C_2}x_{2C_2})$. The sum rate achieved using the above compression strategy is*

$$R = 2log_2|B| + 2h(p). \tag{26}$$

Slepian-Wolf coding for the same scenario would result in a sum rate

$$R_{SW} \;=\; H(X_1, X_2) = 2H(P_B) + 1 + h(p). \tag{27}$$

Clearly, if $2log_2|B| + h(p) < 2H(P_B) + 1$, the distributed compression strategy using homomorphic encoders performs better than the Slepian-Wolf encoding

method. Note that if P_B is uniformly distributed, this is always the case.

Remark 2 *If a group is decomposable, then it can be decomposed as a direct product of indecomposable groups*[1]*, which by the Krull-Schmidt theorem (see Theorem 3.8 in [8]) is unique and hence has the maximum number of Abelian components. One may carry out such a decomposition and choose to compress all the Abelian components. We note though that it is not guaranteed that such a direct product decomposition will give us the best rate of compression.*

Case 2 : Here we show how compression is possible for any non-Abelian group possessing a non-trivial center (as opposed to requiring that it be the direct product of two groups one of which is Abelian). Let $Z = \mathcal{Z}(G)$, the center of G. Then $\mathcal{Z}(G^n) = Z^n$. Let ψ be a homomorphic encoder for the Abelian group Z^n. Consider the cosets of Z^n in G^n. Let $C = \{c_1^n, \ldots, c_r^n\}$ be the coset representatives. Let the output of the two sources be x_1^n and x_2^n and the function to be computed in the receiver be $y^n = x_1^n \circ x_2^n$. Let $x_1^n = c_i^n \circ z_1^n$ and $x_2^n = c_j^n \circ z_2^n,\ z_1^n, z_2^n \in Z^n,\ c_i^n, c_j^n \in C$. The encoding operation at the two sources is then given by the map ϕ, where

$$\text{encoder 1: } \phi : x_1^n \longrightarrow (c_i^n, \psi(z_1^n)) \tag{28}$$

$$\text{encoder 2: } \phi : x_2^n \longrightarrow (c_j^n, \psi(z_2^n)), \tag{29}$$

i.e. encoding takes place in two stages; in the first stage the coset representative is sent without compression and in the second stage the center component is compressed homomorphically. The receiver multiplies the outputs of the two encoders to get $(c_i^n \circ c_j^n, \psi(z_1^n \circ z_2^n))$. The map ψ is chosen to allow reconstruction of $z_1^n \circ z_2^n$ from $\psi(z_1^n \circ z_2^n)$, using c_i^n, c_j^n as side-information (see Section 3). The receiver finally recovers the function y^n as

$$y^n = c_i^n \circ c_j^n \circ z_1^n \circ z_2^n = c_i^n \circ z_1^n \circ c_j^n \circ z_2^n = x_1^n \circ x_2^n, \tag{30}$$

where the second equality follows since z_1^n, z_2^n belong to the center of G^n. Note that the map ϕ is in general not a homomorphism. Yet the scheme allows for distributed function compression as with homomorphic encoders. The rate of compression can be calculated in a fashion as was done in Case

[1] A group G is indecomposable (Definition 3.1 in [8]) if $G \neq \{1_G\}$ and G is not the direct product of two of its proper subgroups.

1.

Case 3: Here we give achievable rates for any non-Abelian group, in a two-source distributed coding setting. Let G be the source alphabet and let A be any Abelian subgroup of G. Such an Abelian subgroup always exists (for example, subgroup generated by a non-identity element). Let x_1^n and x_2^n be the source outputs and the function to be computed be $y_1^n = x_1^n \circ x_2^n$. The first source represents x_1^n an element of left coset of A^n, while the second source represents its output x_2^n as an element of the right coset of A^n. Let C be the set of coset representatives for both the left and right cosets. Let $x_1^n = c_i^n \circ a_1^n$ and $x_2^n = a_2^n \circ c_j^n, a_1^n, a_2^n \in A^n, c_i^n, c_j^n \in C$. Then encoders ϕ_1 and ϕ_2 at the two sources are described as

$$\text{encoder 1: } \phi_1 : x_1^n \longrightarrow (c_i^n, \psi(a_1^n)) \tag{31}$$

$$\text{encoder 2: } \phi_2 : x_2^n \longrightarrow (\psi(a_2^n), c_j^n), \tag{32}$$

where ψ is a homomorphic encoder for A^n. Note that unlike in the previous two cases, the encoder is different for the two sources. The receiver can recover the product $a_1^n \circ a_2^n$ from $\psi(a_1^n \circ a_2^n)$ which in turn is obtained by multiplying $\psi(a_1^n)$ and $\psi(a_2^n)$. The function y^n is then calculated as $y^n = c_i^n \circ (a_1^n \circ a_2^n) \circ c_j^n$. The rate calculation can be carried out similar to what was done in Case 1.

References

[1] R. Puri, A. Majumdar, and K. Ramchandran, "Prism: A video coding paradigm with motion estimation at the decoder," *Image Processing, IEEE Transactions on*, vol. 16, no. 10, pp. 2436–2448, Oct. 2007.

[2] D. Slepian and J. Wolf, "Noiseless coding of correlated information sources," *Information Theory, IEEE Transactions on*, vol. 19, no. 4, pp. 471–480, Jul 1973.

[3] D. Krithivasan and S. Pradhan, "Distributed Source Coding using Abelian Group Codes," *Available: arxiv: 0808.2659v1[cs.IT].*

[4] J. Korner and K. Marton, "How to encode the modulo-two sum of binary sources (corresp.)," *Information Theory, IEEE Transactions on*, vol. 25, no. 2, pp. 219–221, Mar 1979.

[5] D. Krithivasan and S.S. Pradhan, "Distributed source coding using abelian group codes: Extracting performance from structure," in *Communication, Control, and Computing, 2008 46th Annual Allerton Conference on*, Sept. 2008, pp. 1538–1545.

[6] J.C. Interlando, Jr. Palazzo, R., and M. Elia, "Group block codes over nonabelian groups are asymptotically bad," *Information Theory, IEEE Transactions on*, vol. 42, no. 4, pp. 1277–1280, Jul 1996.

[7] E. Biglieri and M. Elia, "On the construction of group block codes," *Annals of Telecommunications*, vol. 50, no. 9, pp. 817–823, 1995.

[8] T.W. Hungerford, *Abstract Algebra: An Introduction.* Saunders College Publishing, 1997.