On Linear Coding over Finite Rings and Applications to Computing

Sheng Huang, *Student Member, IEEE,* Mikael Skoglund, *Senior Member, IEEE*

Abstract

In this paper, we first present a source coding theorem for linear coding over finite rings in the Slepian–Wolf source coding problem. This result includes those given by Elias [1] and Csiszár [2] saying that linear coding over finite field is optimal, i.e., achieves the Slepian–Wolf region. In addition, we show that linear coding over several types of rings (which are not necessarily fields) is also optimal in the single source scenario.

Secondly, we propose linear coding over ring as an alternative solution to the problem of source coding for computing. A generalization of the result in Körner–Marton [3] is presented. Based on this, it is demonstrated that linear coding over ring strictly outperforms its field counterpart in terms of achieving better coding rates and reducing the required alphabet sizes of the encoders.

Index Terms

Linear Coding, Source Coding, Ring, Field, Source Coding for Computing.

I. Introduction

The problem of *source coding for computing* considers the following scenario.

Problem 1 (Source Coding for Computing)**.** Given $\mathcal{S}=\{1,2,\cdots,s\}$ and $(X_1,X_2,\cdots,X_s)\sim p$. Let t_i $(i\in\mathcal{S})$ be a *discrete memoryless source* that randomly generates i.i.d. discrete data $X_i^{(1)},X_i^{(2)},\cdots,X_i^{(n)},\cdots$, where $X_i^{(n)}$ has a finite sample space $\mathscr{X}_i$ and $\left[X_1^{(n)},X_2^{(n)},\cdots,X_s^{(n)}\right]\sim p,\forall\, n\in\mathbb{N}^+$. For a *discrete function* $g:\prod_{i\in\mathcal{S}}\mathscr{X}_i\to\Omega$, what is the largest region $\mathcal{R}[g]\subset\mathbb{R}^s$, such that, $\forall\,(R_1,R_2,\cdots,R_s)\in\mathcal{R}[g]$ and $\forall\,\epsilon>0$, there exists an $N_0\in\mathbb{N}^+$, such that for all $n>N_0$, there exist s *encoders* $\phi_i:\mathscr{X}_i^n\to\left[1,2^{nR_i}\right],i\in\mathcal{S}$, and one *decoder* $\psi:\prod_{i\in\mathcal{S}}\left[1,2^{nR_i}\right]\to\Omega^n$, with

$$\Pr\left\{\vec{g}\left(X_1^n,\cdots,X_s^n\right)\neq\psi\left[\phi_1\left(X_1^n\right),\cdots,\phi_s\left(X_s^n\right)\right]\right\}<\epsilon,$$

S. Huang and M. Skoglund are with the Communication Theory Lab, School of Electrical Engineering, KTH Royal Institute of Technology, Stockholm, 10044, Sweden e-mail: (sheng.huang@ee.kth.se; skoglund@ee.kth.se).

This work was funded in part by the Swedish Research Council.

where $X_i^n = \left[X_i^{(1)}, X_i^{(2)}, \cdots, X_i^{(n)}\right]$ and

$$\vec{g}\left(X_1^n, \cdots, X_s^n\right) = \begin{bmatrix} g\left(X_1^{(1)}, \cdots, X_s^{(1)}\right) \\ \vdots \\ g\left(X_1^{(n)}, \cdots, X_s^{(n)}\right) \end{bmatrix} \in \Omega^n ?$$

The region $\mathcal{R}[g]$ is called the *achievable coding rate region* for computing g. A rate tuple $\mathbf{R} \in \mathbb{R}^s$ is said to be *achievable* for computing g (or simply achievable) if and only if $\mathbf{R} \in \mathcal{R}[g]$. A region $\mathcal{R} \subset \mathbb{R}^s$ is said to be *achievable* for computing g (or simply achievable) if and only if $\mathcal{R} \subseteq \mathcal{R}[g]$.

If g is the *identity function*, it is obvious that the described computing problem is equivalent to the *Slepian–Wolf* (SW) *source coding* problem. Hence, $\mathcal{R}[g]$ is the *SW region* [4], namely

$$\mathcal{R}[X_1, X_2, \cdots, X_s] = \Big\{(R_1, R_2, \cdots, R_s) \in \mathbb{R}^s \;\Big|\; \sum_{j \in T} R_j > H(X_T | X_{T^c}), \forall\, \emptyset \neq T \subseteq \mathcal{S}\Big\},$$

where T^c is the *complement* of T in $\mathcal{S}$ and $X_T\,(X_{T^c})$ is the random variable array $\prod_{j \in T} X_j \left(\prod_{j \in T^c} X_j\right)$. In the original SW source coding scenario, the structure of the encoders is unclear, the corresponding mappings are chosen randomly among all feasible mappings. In the single source scenario, Elias [1] showed that *linear coding over finite field* (LCoF), where $\mathscr{X}_i$'s and Ω are embedded as subsets of this *field* and ϕ_i's are *linear mappings*, is sufficient in achieving the best coding rate. This idea is then generalized to the multiple sources scenario by Csiszár [2]. As a consequence of [2], any rate tuple in the SW region is achievable with LCoF.

Generally speaking, $\mathcal{R}[X_1, X_2, \cdots, X_s] \subseteq \mathcal{R}[g]$ for an arbitrary discrete function g. Making use of Elias' theorem on binary linear codes [1], Körner–Marton [3] shows that $\mathcal{R}[\oplus_2]$ ("$\oplus_2$" is the *modulo-two sum*) contains the region

$$\mathcal{R}_{\oplus_2} = \left\{(R_1, R_2) \in \mathbb{R}^2 \mid R_1, R_2 > H(X_1 \oplus_2 X_2)\right\}.$$

This region is not contained in the SW region for certain distributions. In other words, $\mathcal{R}[\oplus_2] \supsetneq \mathcal{R}[X_1, X_2]$. Combining the standard random coding technique and Elias' result, [5] shows that $\mathcal{R}[\oplus_2]$ can be strictly larger than the convex hull of the union $\mathcal{R}[X_1, X_2] \cup \mathcal{R}_{\oplus_2}$. However, the functions considered in these works are relatively simple.

Taking on a *polynomial* approach, [6], [7] generalize the result of Ahlswede–Han [5, Theorem 10] to the most general scenario. Making use of the fact that a discrete function is essentially a *polynomial function* [8, pp. 93] over some finite field, an achievable region is given for computing an arbitrary discrete function. Such a region contains and can be strictly larger (depending on the precise function and distribution under consideration) than the SW region. Conditions under which $\mathcal{R}[g]$ is strictly larger than the SW region are presented in [9] and [6] from different perspectives, respectively.

We observe that the linear coding (LC) technique over field by Elias and Csiszár is a key element in the results accounted for above. This observation inspires our study of searching for alternative encoding methods (coding

techniques). This paper focuses on *linear coding over finite ring* (LCoR) which serves as an alternative technique in the case of the computing problem. We will show that this approach is better in terms of achieving coding rates and reducing alphabet sizes of the encoders compared to its field counterpart. In Section III, we present a region (4) achieved with LC over several finite *rings*, namely all encoders are *linear mappings over rings* (see Definition II.9). It is proved that this region specializes to the SW region if all of these rings are fields. Thus, the results of [1], [2] become special cases of ours in this sense. In addition, Section V shows that LCoR can achieve the best coding rate in the single source scenario with certain rings (e.g., $\mathbb{Z}_4$) that are not necessary fields. To illustrate applications to computing, Problem 1 is considered in Section VI where a generalized theorem of Körner–Marton [3] is given. To conclude Section VI, Example VI.5 is constructed exhibiting the advantages of this LC technique over finite rings. In this example, LCoR achieves a strictly larger region compared to the one obtained with LCoF in the sense of Körner–Marton [3]. Additionally, the encoders using LCoR require strictly smaller alphabet sizes than when using LCoF.

In addition to the fact that LCoR outperforms LCoF, Example VI.5 also points at another circumstance. The results of Körner–Marton [3] and Ahlswede–Han [5, Theorem 10] depend on LCoF and apply to linear functions over fields only. Therefore, their methods do not apply directly to discrete functions that are neither linear nor can be linearized over any finite field (e.g., g from Example VI.5). Although the polynomial approach [6], [7], proposed by the authors of the present work, works universally for any discrete functions, it will possibly require larger alphabet size. More importantly, this approach requires that the *polynomial presentation* of the function is available in accessible form. Unfortunately, this can possibly turn out to be a very strict requirement, even though it is proved to be always possible. On the contrary, those discrete functions (e.g., g from Example VI.5) could be simply a linear function over certain finite ring. Therefore, LCoR offers an alternative solution. As a matter of fact, such an alternative solution is rather promising for, at least, functions like g from Example VI.5.

Conceptually speaking, LCoR is in fact a generalization of the LC technique proposed by Elias and Csiszár (LCoF), since a field must be a ring. However, as seen in Section IV, the analysis of decoding error for the ring version is in general substantially more challenging than in the case of the field version. Our analysis crucially relies on the concept of *ideal* of ring. A field contains no non-trivial ideal but itself. Because of this special property of fields, our general argument for finite rings deployed in later sections will render to a simple one if only finite fields are considered.

Even though our analysis for the ring scenario is more complicated than the one for field, linear encoders working over some finite rings are in general considerably easier to implement than using finite fields in practice. This is because the implementation of *finite field arithmetic* can be quite demanding. Normally, a finite field is given by its *polynomial representation*, operations are carried out based on the polynomial operation followed by *polynomial long division algorithm*. On the contrary, implementing arithmetic of many finite rings is a straightforward task. For instance, the arithmetic of *modulo integers ring* $\mathbb{Z}_q$, for any positive integer q, is simply the integer modulo q arithmetic.

II. Rings, Ideals and Linear Mappings

In this section we introduce some fundamental concepts from abstract algebra. Readers who are already familiar with this material may still choose to go through quickly to identify our notation.

Definition II.1. The tuple $[\mathfrak{R}, +, \cdot]$ is called a *ring* if the following criteria are met:

1) $[\mathfrak{R}, +]$ is an *Abelian group*;
2) There exists a *multiplicative identity*[1] $1 \in \mathfrak{R}$, namely, $1 \cdot a = a \cdot 1 = a$, $\forall\ a \in \mathfrak{R}$;
3) $\forall\ a, b, c \in \mathfrak{R}$, $a \cdot b \in \mathfrak{R}$ and $(a \cdot b) \cdot c = a \cdot (b \cdot c)$;
4) $\forall\ a, b, c \in \mathfrak{R}$, $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$ and $(b + c) \cdot a = (b \cdot a) + (c \cdot a)$.

We often write $\mathfrak{R}$ for $[\mathfrak{R}, +, \cdot]$ when the *operations* (*operators*) considered are known from the context. The operator "$\cdot$" is usually written by juxtaposition, ab for $a \cdot b$, for all $a, b \in \mathfrak{R}$.

A ring $[\mathfrak{R}, +, \cdot]$ is said to be *commutative* if $\forall\ a, b \in \mathfrak{R}$, $a \cdot b = b \cdot a$. In Definition II.1, the *identity* of the group $[\mathfrak{R}, +]$, denoted by 0, is called the *zero*. A ring $[\mathfrak{R}, +, \cdot]$ is said to be *finite* if the cardinality $|\mathfrak{R}|$ is finite, and $|\mathfrak{R}|$ is called the *order* of $\mathfrak{R}$. The set $\mathbb{Z}_q$ of integers modulo q is a commutative finite ring with respect to the *modular arithmetic*. For any ring $\mathfrak{R}$, the set of all *polynomials* of s *indeterminants*, namely $\mathfrak{R}[X_1, X_2, \cdots, X_s]$, is an infinite ring. Meanwhile, we denote the set of all the polynomial functions of s *variables* over ring $\mathfrak{R}$ by $\mathfrak{R}[s]$.

Proposition II.2. *Given s rings $\mathfrak{R}_1, \mathfrak{R}_2, \cdots, \mathfrak{R}_s$, for any non-empty set $T \subseteq \{1, 2, \cdots, s\}$, the Cartesian product $\mathfrak{R}_T = \prod_{i \in T} \mathfrak{R}_i$ forms a new ring $[\mathfrak{R}_T, +, \cdot]$ with respect to the component-wise operations defined as follows:*

$$\mathbf{r}' + \mathbf{r}'' = \left[r'_1 + r''_1, r'_2 + r''_2, \cdots, r'_{|T|} + r''_{|T|}\right],$$
$$\mathbf{r}' \cdot \mathbf{r}'' = \left[r'_1 r''_1, r'_2 r''_2, \cdots, r'_{|T|} r''_{|T|}\right],$$

$\forall\ \mathbf{r}' = \left[r'_1, r'_2, \cdots, r'_{|T|}\right], \mathbf{r}'' = \left[r''_1, r''_2, \cdots, r''_{|T|}\right] \in \mathfrak{R}_T$.

Remark 1. In Proposition II.2, it can be easily seen that $[0, 0, \cdots, 0]$ and $[1, 1, \cdots, 1]$ are the zero and multiplicative identity of $\mathfrak{R}_T$, respectively.

Definition II.3. A non-zero element a of a ring $\mathfrak{R}$ is said to be *invertible*, if and only if there exists $b \in \mathfrak{R}$, such that $ab = ba = 1$. b is called the *inverse* of a, denoted by a^{-1}. An invertible element of a ring is called a *unit*.

Remark 2. It can be proved that the inverse of a unit is unique. By definition, the multiplicative identity is the inverse of itself.

Let $\mathfrak{R}^* = \mathfrak{R} \setminus \{0\}$. The ring $[\mathfrak{R}, +, \cdot]$ is a *field* if and only if $\mathfrak{R}^*$ is an *Abelian* group with respect to the multiplicative operation "$\cdot$". In other words, all non-zero elements of $\mathfrak{R}$ are invertible. All fields are commutative

[1] Sometimes a ring without a multiplicative identity is considered. Such a structure has been called a *rng*. We consider rings with multiplicative identities in this paper. However, similar results remain valid when considering rngs instead. Although we will occasionally comment on such results, they are not fully considered in the present work.

rings. $\mathbb{Z}_q$ is a field if and only if q is a *prime*. Up to *isomorphism*, all finite fields are unique [10, pp. 549]. We use $\mathbb{F}_q$ to denote this "unique" field of order q. It is necessary that q is a power of a prime. More details of finite field can be found in [10, Ch. 14.3].

Theorem II.4 (Wedderburn's little theorem c.f. Theorem 7.13 of [8])**.** *Let $\mathfrak{R}$ be a finite ring. $\mathfrak{R}$ is a field if and only if all non-zero elements of $\mathfrak{R}$ are invertible.*

Remark 3. Wedderburn's little theorem guarantees commutativity for a finite ring if all of its non-zero elements are invertible. Hence, a finite ring is either a field or at least one of its elements has no inverse. However, a finite commutative ring is not necessary a field, e.g. $\mathbb{Z}_q$ is not a field if q is not a prime.

Definition II.5 (c.f. [10])**.** The *characteristic* of a finite ring $\mathfrak{R}$ is defined to be the smallest positive integer m, such that $\sum_{j=1}^{m} 1 = 0$, where 0 and 1 are the zero and the multiplicative identity of $\mathfrak{R}$, respectively. The characteristic of $\mathfrak{R}$ is often denoted by Char($\mathfrak{R}$).

Remark 4. Clearly, Char($\mathbb{Z}_q$) $= q$. For a finite field $\mathbb{F}_q$, Char($\mathbb{F}_q$) is always the prime q_0 such that $q = q_0^n$ for some integer n [8, Proposition 2.137].

Proposition II.6. *Let $\mathbb{F}_q$ be a finite field. For any $0 \neq a \in \mathbb{F}_q$, $m =$ Char($\mathbb{F}_q$) if and only if m is the smallest positive integer such that $\sum_{j=1}^{m} a = 0$.*

Proof: Since $a \neq 0$,

$$\sum_{j=1}^{m} a = 0 \Rightarrow a^{-1} \sum_{j=1}^{m} a = a^{-1} \cdot 0 \Rightarrow \sum_{j=1}^{m} 1 = 0 \Rightarrow \sum_{j=1}^{m} a = 0$$

The statement is proved. ■

Definition II.7. A subset $\mathfrak{I}$ of a ring $[\mathfrak{R}, +, \cdot]$ is said to be a *left ideal* of $\mathfrak{R}$, denoted by $\mathfrak{I} \leq_l \mathfrak{R}$, if and only if

1) $[\mathfrak{I}, +]$ is a subgroup of $[\mathfrak{R}, +]$;
2) $\forall\, x \in \mathfrak{I}$ and $\forall\, r \in \mathfrak{R}$, $r \cdot x \in \mathfrak{I}$.

If condition 2) is replaced by

3) $\forall\, x \in \mathfrak{I}$ and $\forall\, r \in \mathfrak{R}$, $x \cdot r \in \mathfrak{I}$,

then $\mathfrak{I}$ is called a *right ideal* of $\mathfrak{R}$, denoted by $\mathfrak{I} \leq_r \mathfrak{R}$. $\{0\}$ is a *trivial* left (right) ideal, usually denoted by 0.

The cardinality $|\mathfrak{I}|$ is called the *order* of a finite left (right) ideal $\mathfrak{I}$.

Remark 5. Let $\{a_1, a_2, \cdots, a_n\}$ be a non-empty set of elements of some ring $\mathfrak{R}$. It is easy to verify that $\langle a_1, a_2, \cdots, a_n \rangle_r = \left\{ \sum_{i=1}^{n} a_i r_i \,\middle|\, r_i \in \mathfrak{R}, \forall\, 1 \leq i \leq n \right\}$ is a right ideal. Furthermore, $\langle a_1, a_2, \cdots, a_n \rangle_r = \mathfrak{R}$ if a_i is a unit for some $1 \leq i \leq n$.

It is well-known that if $\mathfrak{I} \leq_l \mathfrak{R}$ or $\mathfrak{I} \leq_r \mathfrak{R}$, then $\mathfrak{R}$ is divided into disjoint *cosets* which are of equal size (cardinality). For any coset $\mathfrak{J}$, $\mathfrak{J} = x + \mathfrak{I} = \{x + y | y \in \mathfrak{I}\}$, $\forall\ x \in \mathfrak{J}$. The set of all cosets forms a *quotient group*, denoted by $\mathfrak{R}/\mathfrak{I}$. See [8, Ch. 1.6 and Ch. 2.9] for more details.

Proposition II.8. *Let $\mathfrak{R}_i$ $(1 \leq i \leq s)$ be a ring and $\mathfrak{R} = \prod_{i=1}^{s} \mathfrak{R}_i$. For any $\mathfrak{A} \subseteq \mathfrak{R}$, $\mathfrak{A} \leq_l \mathfrak{R}$ (or $\mathfrak{A} \leq_r \mathfrak{R}$) if and only if $\mathfrak{A} = \prod_{i=1}^{s} \mathfrak{A}_i$ and $\mathfrak{A}_i \leq_l \mathfrak{R}_i$ (or $\mathfrak{A}_i \leq_r \mathfrak{R}_i$), $\forall\ 1 \leq i \leq s$.*

Proof: It suffices to complete the proof for $\leq_l$ only. If $\mathfrak{A} \leq_l \mathfrak{R}$, then $\mathfrak{A} = \prod_{i=1}^{s} \mathfrak{A}_i$ for some $\mathfrak{A}_i \subseteq \mathfrak{R}_i$, because for any $(a_1, a_2, \cdots, a_s), (b_1, b_2, \cdots, b_s) \in \mathfrak{A}$, $\{(c_1, c_2, \cdots, c_s) \in \mathfrak{R} | c_i = a_i \text{ or } b_i, 1 \leq i \leq s\} \subseteq \mathfrak{A}$. For instance, $(b_1, a_2, \cdots, a_s) = (0, 1, \cdots, 1) \cdot (a_1, a_2, \cdots, a_s) + (1, 0, \cdots, 0) \cdot (b_1, b_2, \cdots, b_s) \in \mathfrak{A}$. Furthermore, it is easy to show that $\mathfrak{A}_i \leq_l \mathfrak{R}_i$ for all feasible i by definition. Sufficiency is obvious. ■

Remark 6. For any $\emptyset \neq T \subseteq \mathcal{S}$, Proposition II.8 states that any left (right) ideal of $\mathfrak{R}_T$ is a Cartesian product of some left (right) ideals of $\mathfrak{R}_i$, $i \in T$. Let $\mathfrak{I}_i$ be a left (right) ideal of ring $\mathfrak{R}_i$ $(1 \leq i \leq s)$. We define $\mathfrak{I}_T$ to be the left (right) ideal $\prod_{i \in T} \mathfrak{I}_i$ of $\mathfrak{R}_T$.

Remark 7. It is worthwhile to point out that Proposition II.8 does not hold for infinite index set, namely, $\mathfrak{R} = \prod_{i \in I} \mathfrak{R}_i$, where I is not finite.

Definition II.9. A mapping $f : \mathfrak{R}^n \to \mathfrak{R}^m$ given as:

$$\begin{aligned} f(x_1, x_2, \cdots, x_n) =& \left(\sum_{j=1}^{n} a_{1,j} x_j, \cdots, \sum_{j=1}^{n} a_{m,j} x_j \right)^t \\ \left(f(x_1, x_2, \cdots, x_n) =& \left(\sum_{j=1}^{n} x_j a_{1,j}, \cdots, \sum_{j=1}^{n} x_j a_{m,j} \right)^t \right), \\ & \forall\ (x_1, x_2, \cdots, x_n) \in \mathfrak{R}^n, \end{aligned} \tag{1}$$

where $a_{i,j} \in \mathfrak{R}$ for all feasible i and j, is called a *left* (*right*) *linear mapping* over ring $\mathfrak{R}$. If $m = 1$, then f is called a *left* (*right*) *linear function* over $\mathfrak{R}$.

From now on, left linear mapping (function) or right linear mapping (function) are simply called *linear mapping* (*function*). This will not lead to any confusion since the intended use can usually be clearly distinguished from the context.

Remark 8. The mapping f in Definition II.9 is called *linear* in accordance with the definition of *linear mapping* (*function*) over field. In fact, the two structures have several similar properties. Moreover, (1) is equivalent to

$$\begin{aligned} f(x_1, x_2, \cdots, x_n) =& \mathbf{A} \left(x_1, x_2, \cdots, x_n \right)^t, \\ & \forall\ (x_1, x_2, \cdots, x_n) \in \mathfrak{R}^n, \end{aligned} \tag{2}$$

where $\mathbf{A}$ is an $m \times n$ *matrix* over $\mathfrak{R}$ and $[\mathbf{A}]_{i,j} = a_{i,j}$ for all feasible i and j. $\mathbf{A}$ is named the *coefficient matrix*. It is easy to prove that a linear mapping is uniquely determined by its coefficient matrix, and vice versa. The linear mapping f is said to be *trivial*, denoted by 0, if $\mathbf{A}$ is the *zero matrix*, i.e., $[\mathbf{A}]_{i,j} = 0$ for all feasible i and j.

Let $\mathbf{A}$ be an $m \times n$ matrix over ring $\mathfrak{R}$ and $f(\mathbf{x}) = \mathbf{A}\mathbf{x}$, $\forall\, \mathbf{x} \in \mathfrak{R}^n$. For the *system of linear equations*

$$f(\mathbf{x}) = \mathbf{A}\mathbf{x} = \mathbf{0}, \text{ where } \mathbf{0} = (0, 0, \cdots, 0)^t \in \mathfrak{R}^m,$$

let $\mathfrak{S}(f)$ be the set of all *solutions*, namely $\mathfrak{S}(f) = \{\mathbf{x} \in \mathfrak{R}^n | f(\mathbf{x}) = \mathbf{0}\}$. It is obvious that $\mathfrak{S}(f) = \mathfrak{R}^n$ if f is trivial, i.e., $\mathbf{A}$ is the zero matrix. If $\mathfrak{R}$ is a field, then $\mathfrak{S}(f)$ is a *subspace* of $\mathfrak{R}^n$. We conclude this section with a lemma regarding the cardinalities of $\mathfrak{R}^n$ and $\mathfrak{S}(f)$ in the following.

Lemma II.10. *For a finite ring $\mathfrak{R}$ and a linear function*

$$f : \mathbf{x} \mapsto [a_1, a_2, \cdots, a_n]\mathbf{x}, \quad \forall\, \mathbf{x} \in \mathfrak{R}^n,$$

we have

$$\frac{|\mathfrak{S}(f)|}{|\mathfrak{R}|^n} = \frac{1}{|\mathfrak{I}|},$$

where $\mathfrak{I} = \langle a_1, a_2, \cdots, a_n \rangle_r$. In particular, if a_i is invertible for some $1 \le i \le n$, then $|\mathfrak{S}(f)| = |\mathfrak{R}|^{n-1}$.

Proof: It is obvious that the image $f(\mathfrak{R}^n) = \mathfrak{I}$ by definition. Moreover, $\forall\ x \neq y \in \mathfrak{I}$, the pre-images $f^{-1}(x) \cap f^{-1}(y) = \emptyset$ and $\left|f^{-1}(x)\right| = \left|f^{-1}(y)\right| = |\mathfrak{S}(f)|$. Therefore, $|\mathfrak{I}|\,|\mathfrak{S}(f)| = |\mathfrak{R}|^n$, i.e., $\frac{|\mathfrak{S}(f)|}{|\mathfrak{R}|^n} = \frac{1}{|\mathfrak{I}|}$. Moreover, if a_i is a unit, then $\mathfrak{I} = \mathfrak{R}$, thus, $|\mathfrak{S}(f)| = |\mathfrak{R}|^n / |\mathfrak{R}| = |\mathfrak{R}|^{n-1}$. ∎

Remark 9. For linear function

$$f : \mathbf{x} \mapsto \mathbf{x}^t[a_1, a_2, \cdots, a_n]^t, \quad \forall\, \mathbf{x} \in \mathfrak{R}^n,$$

Lemma II.10 holds true, if

$$\mathfrak{I} = \langle a_1, a_2, \cdots, a_n \rangle_l = \left\{ \sum_{i=1}^{n} r_i a_i \,\middle|\, r_i \in \mathfrak{R}, \forall\, 1 \le i \le n \right\}$$

which is a left ideal of $\mathfrak{R}$.

III. Linear Coding over Finite Rings

In this section, we will present a coding rate region achieved with LCoR for the SW source coding problem, i.e., g is the identity function in Problem 1. This region is exactly the SW region if all the rings considered are fields. However, being field is not necessary as seen in Section V.

Before proceeding, a subtleness needs to be cleared out. It is assumed that a source, say t_i, generates data taking values from a finite sample space $\mathscr{X}_i$, while $\mathscr{X}_i$ does not necessarily admit any algebraic structure. We have to either assume that $\mathscr{X}_i$ is with a certain algebraic structure, for instance $\mathscr{X}_i$ is a ring, or injectively map elements of $\mathscr{X}_i$ into some algebraic structure. In our subsequent discussions, we assume that $\mathscr{X}_i$ is mapped into a finite ring

$\mathfrak{R}_i$ of order at least $|\mathscr{X}_i|$ by some injection Φ_i. Hence, $\mathscr{X}_i$ can simply be treated as a subset $\Phi_i(\mathscr{X}_i) \subseteq \mathfrak{R}_i$ for a fixed Φ_i. When required, Φ_i can also be selected to obtain desired outcomes (see Remark 11).

To simplify our discussion, the following notation is used. For $\emptyset \neq T \subseteq \mathcal{S}$, X_T $(x_T, \mathscr{X}_T)$ is defined to be the Cartesian product

$$\prod_{i\in T} X_i \left(\prod_{i\in T} x_i, \prod_{i\in T} \mathscr{X}_i \right),$$

where $x_i \in \mathscr{X}_i$ is a realization of X_i. If $(X_1, X_2, \cdots, X_s) \sim p$, we denote the marginal of p with respect to X_T by p_{X_T} and define

$$\text{supp}(p) = \left\{ \prod_{i=1}^{s} x_i \in \prod_{i=1}^{s} \mathscr{X}_i \middle| \, p\left(\prod_{i=1}^{s} x_i \right) > 0 \right\}.$$

Besides, $\mathscr{M}(\mathscr{X}_{\mathcal{S}}, \mathfrak{R}_{\mathcal{S}})$ is defined to be

$$\{ [\Phi_1, \Phi_2, \cdots, \Phi_s] | \, \Phi_i : \mathscr{X}_i \to \mathfrak{R}_i \text{ is injective}, \forall\, i \in \mathcal{S} \}$$

($|\mathfrak{R}_i| \geq |\mathscr{X}_i|$ is implicitly assumed).

Theorem III.1. *Given $\Phi \in \mathscr{M}(\mathscr{X}_{\mathcal{S}}, \mathfrak{R}_{\mathcal{S}})$ and let*

$$\mathcal{R}_\Phi = \left\{ [R_1, R_2, \cdots, R_s] \in \mathbb{R}^s \middle| \sum_{i\in T} \frac{R_i \log|\mathfrak{I}_i|}{\log|\mathfrak{R}_i|} > r(T, \mathfrak{I}_T), \right.$$
$$\left. \forall\, \emptyset \neq T \subseteq \mathcal{S}, \forall\, 0 \neq \mathfrak{I}_i \leq_l \mathfrak{R}_i \right\}, \tag{3}$$

where $r(T, \mathfrak{I}_T) = H(X_T|X_{T^c}) - H(Y_{\mathfrak{R}_T/\mathfrak{I}_T}|X_{T^c})$ and $Y_{\mathfrak{R}_T/\mathfrak{I}_T}$ is a random variable with sample space $\mathfrak{R}_T/\mathfrak{I}_T$ and

$$\Pr\left\{ Y_{\mathfrak{R}_T/\mathfrak{I}_T} = a + \mathfrak{I}_T \middle| \, X_T = x_T \right\} = \begin{cases} 1; & \textit{if } x_T \in a + \mathfrak{I}_T, \\ 0; & \textit{otherwise}. \end{cases}$$

Any rate in $\mathcal{R}_\Phi$ is achievable with linear coding over finite rings $\mathfrak{R}_1, \mathfrak{R}_2, \cdots, \mathfrak{R}_s$.

A concrete example can be helpful in the interpretation of this theorem.

Example III.2. Consider the single source scenario, where $X_1 \sim p$ and $\mathscr{X}_1 = \mathbb{Z}_6$, satisfying the follows.

X_1	0	1	2	3	4	5
$p(X_1)$	0.05	0.1	0.15	0.2	0.2	0.3

By Theorem III.1,

$$\begin{aligned} \mathcal{R} &= \{R_1 \in \mathbb{R} | R_1 > \max\{2.40869, 2.34486, 2.24686\}\} \\ &= \{R_1 \in \mathbb{R} | R_1 > 2.40869 = H(X_1)\} \end{aligned}$$

is achievable with linear coding over ring $\mathbb{Z}_6$. Obviously, $\mathcal{R}$ is just the SW region $\mathcal{R}[X_1]$. Optimality is claimed.

Remark 10. Without much effort, one can see that $\mathcal{R}_\Phi$ in the above theorem depends on Φ via random variables $Y_{\mathfrak{R}_T/\mathfrak{I}_T}$'s whose distributions are determined by Φ. For each $i \in \mathcal{S}$, there exist $\binom{|\mathfrak{R}_i|}{|\mathscr{X}_i|}$ distinct injections from $\mathscr{X}_i$ to a ring $\mathfrak{R}_i$ of order at least $|\mathscr{X}_i|$. Let $\mathrm{cov}(A)$ be the convex hull of a set $A \subseteq \mathbb{R}^s$. By a straightforward time sharing argument, we have that

$$\mathcal{R}_l = \mathrm{cov}\left(\bigcup_{\Phi \in \mathscr{M}(\mathscr{X}_\mathcal{S}, \mathfrak{R}_\mathcal{S})} \mathcal{R}_\Phi\right) \tag{4}$$

is achievable with LCoR.

Remark 11. From Theorem V.1, one will see that (3) and (4) are the same when all the rings are fields. Actually, both are identical to the SW region. However, (4) can be strictly larger than (3) (see Theorem V.2), when not all the rings are fields. This implies that, in order to achieve desired rate, a suitable injection is required. However, be reminded that taking convex hull (4) is not always needed for optimality as shown in Example III.2. Much more sophisticated analysis on this issue is found in Section V.

The rest of this section provides key supporting lemmata and concepts used to prove Theorem III.1. The final proof is given in Section IV.

Lemma III.3. *Given a finite ring $\mathfrak{R}$, two distinct sequences $\mathbf{x}, \mathbf{y} \in \mathfrak{R}^n$, and let $\mathbf{y} - \mathbf{x} = [a_1, a_2, \cdots, a_n]^t$ and $f : \mathfrak{R}^n \to \mathfrak{R}^k$ be a linear mapping chosen uniformly randomly, i.e., generate the $k \times n$ coefficient matrix $\mathbf{A}$ of f by independently choosing each entry of $\mathbf{A}$ uniformly at random. Then*

$$\Pr\{f(\mathbf{x}) = f(\mathbf{y})\} = |\mathfrak{I}|^{-k}, \tag{5}$$

where $\mathfrak{I} = \langle a_1, a_2, \cdots, a_n \rangle_l$.

Proof: Assume that $f = (f_1, f_2, \cdots, f_k)^t$, where $f_i : \mathfrak{R}^n \to \mathfrak{R}$ is a random linear function. Then

$$\begin{aligned}\Pr\{f(\mathbf{x}) = f(\mathbf{y})\} &= \Pr\left\{\bigcap_{i=1}^k \{f_i(\mathbf{x}) = f_i(\mathbf{y})\}\right\} \\ &= \prod_{i=1}^k \Pr\{f_i(\mathbf{x} - \mathbf{y}) = 0\},\end{aligned}$$

since f_i's are independent to each other. The statement follows from Lemma II.10 and Remark 9 which assure that $\Pr\{f_i(\mathbf{x} - \mathbf{y}) = 0\} = |\mathfrak{I}|^{-1}$. ∎

Remark 12. In Lemma III.3, if $\mathfrak{R}$ is a field and $\mathbf{x} \neq \mathbf{y}$, then $\mathfrak{I} = \mathfrak{R}$ because every non-zero a_i is a unit. Thus, $\Pr\{f(\mathbf{x}) = f(\mathbf{y})\} = |\mathfrak{R}|^{-k}$.

Definition III.4 (c.f. [11])**.** Let $X \sim p_X$ be a discrete random variable with sample space $\mathscr{X}$. The set $\mathcal{T}_\epsilon(n, X)$ of *strongly ϵ-typical sequences* of length n with respect to X is defined to be

$$\left\{\mathbf{x} \in \mathscr{X}^n \left| \left| \frac{N(x; \mathbf{x})}{n} - p_X(x) \right| \leq \epsilon, \forall\, x \in \mathscr{X} \right.\right\},$$

where $N(x;\mathbf{x})$ is the number of occurrences of x in the sequence $\mathbf{x}$.
$\mathcal{T}_\epsilon(n,X)$ is sometimes replaced by $\mathcal{T}_\epsilon$ when the length n and the random variable X referred to are clear from the context.

Let $X\sim p_X$ be a discrete random variable with finite sample space $\mathscr{X}$ and $H(p_X)=H(X)$. It is well-known that H is a concave function, i.e., $cH(p_1)+(1-c)H(p_2)\le H(cp_1+(1-c)p_2)$, $\forall\ 0\le c\le 1$. Equality holds if and only if $p_1=p_2$. For convenience, define $\mathrm{emp}(\mathbf{x})$ to be the *empirical distribution* of $\mathbf{x}$ and $H(\mathbf{x})=H(\mathrm{emp}(\mathbf{x}))$, $\forall\ \mathbf{x}\in\mathscr{X}^n$.

Lemma III.5. *Let $(X_1,X_2)\sim p$ be a jointly random variable whose sample space is a finite ring $\mathfrak{R}=\mathfrak{R}_1\times\mathfrak{R}_2$. For any $\eta>0$, there exists $\epsilon>0$, such that, $\forall\ (\mathbf{x}_1,\mathbf{x}_2)^t\in\mathcal{T}_\epsilon(n,(X_1,X_2))$ and $\forall\ \mathfrak{I}\le_l\mathfrak{R}_1$,*

$$|D_\epsilon(\mathbf{x}_1,\mathfrak{I}|\mathbf{x}_2)|<2^{n\left[H(X_1|X_2)-H(Y_{\mathfrak{R}_1/\mathfrak{I}}|X_2)+\eta\right]}, \tag{6}$$

where

$$D_\epsilon(\mathbf{x}_1,\mathfrak{I}|\mathbf{x}_2)=\left\{(\mathbf{y},\mathbf{x}_2)^t\in\mathcal{T}_\epsilon\middle|\,\mathbf{y}-\mathbf{x}_1\in\mathfrak{I}^n\right\}$$

and $Y_{\mathfrak{R}_1/\mathfrak{I}}$ is a random variable with sample space $\mathfrak{R}_1/\mathfrak{I}$ such that

$$\Pr\left\{Y_{\mathfrak{R}_1/\mathfrak{I}}=a+\mathfrak{I}\middle|\,X_1=r_1\right\}=\begin{cases}1; & \textit{if } r_1\in a+\mathfrak{I},\\ 0; & \textit{otherwise}.\end{cases}$$

Proof: Let $\mathfrak{R}_1/\mathfrak{I}=\{a_1+\mathfrak{I},a_2+\mathfrak{I},\cdots,a_m+\mathfrak{I}\}$, where $m=|\mathfrak{R}_1|/|\mathfrak{I}|$. For arbitrary $\epsilon>0$ and integer n, without loss of generality, assume that

$$\begin{bmatrix}\mathbf{x}_1\\ \mathbf{x}_2\end{bmatrix}=\begin{bmatrix}\mathbf{x}_{1,1},\mathbf{x}_{1,2},\cdots,\mathbf{x}_{1,m}\\ \mathbf{x}_{2,1},\mathbf{x}_{2,2},\cdots,\mathbf{x}_{2,m}\end{bmatrix}=\begin{bmatrix}x_1^{(1)}, & x_1^{(2)}, & \cdots, & x_1^{(n)}\\ x_2^{(1)}, & x_2^{(2)}, & \cdots, & x_2^{(n)}\end{bmatrix}$$

and

$$\mathbf{z}_j=\begin{bmatrix}\mathbf{x}_{1,j}\\ \mathbf{x}_{2,j}\end{bmatrix}=\begin{bmatrix}x_1^{(\sum_{k=0}^{j-1}c_k+1)}, & x_1^{(\sum_{k=0}^{j-1}c_k+2)}, & \cdots, & x_1^{(\sum_{k=0}^{j}c_k)}\\ x_2^{(\sum_{k=0}^{j-1}c_k+1)}, & x_2^{(\sum_{k=0}^{j-1}c_k+2)}, & \cdots, & x_2^{(\sum_{k=0}^{j}c_k)}\end{bmatrix}\in(a_j+\mathfrak{I}\times\mathfrak{R}_2)^{c_j},$$

where $c_0=0$ and $c_j=\sum_{r\in a_j+\mathfrak{I}\times\mathfrak{R}_2}N\left(r,(\mathbf{x}_1,\mathbf{x}_2)^t\right)$, $1\le j\le m$. For any $\mathbf{y}=\left[y^{(1)},y^{(2)},\cdots,y^{(n)}\right]$ with $(\mathbf{y},\mathbf{x}_2)^t\in D_\epsilon(\mathbf{x}_1,\mathfrak{I}|\mathbf{x}_2)$, we have $y^{(i)}-x_1^{(i)}\in\mathfrak{I}$, $\forall\ 1\le i\le n$, by definition. Thus, $y^{(i)}$ and $x_1^{(i)}$ belong to the same coset, i.e., $y^{(\sum_{k=0}^{j-1}c_k+1)},y^{(\sum_{k=0}^{j-1}c_k+2)},\cdots,y^{(\sum_{k=0}^{j}c_k)}\in a_j+\mathfrak{I},\forall\ 1\le j\le m$. Furthermore, $\forall\ r\in\mathfrak{R}$,

$$\begin{aligned}&\left|N\left(r,(\mathbf{x}_1,\mathbf{x}_2)^t\right)/n-p(r)\right|\le\epsilon \text{ and}\\ &\left|N\left(r,(\mathbf{y},\mathbf{x}_2)^t\right)/n-p(r)\right|\le\epsilon\\ \Longrightarrow&\left|\frac{N\left(r,(\mathbf{y},\mathbf{x}_2)^t\right)}{n}-\frac{N\left(r,(\mathbf{x}_1,\mathbf{x}_2)^t\right)}{n}\right|\le 2\epsilon,\end{aligned}$$

since $(\mathbf{x}_1,\mathbf{x}_2)^t,(\mathbf{y},\mathbf{x}_2)^t\in\mathcal{T}_\epsilon$. As a consequence,

$$\mathbf{z}'_j=\begin{bmatrix}y^{(\sum_{k=0}^{j-1}c_k+1)}, & y^{(\sum_{k=0}^{j-1}c_k+2)}, & \cdots, & y^{(\sum_{k=0}^{j}c_k)}\\ x_2^{(\sum_{k=0}^{j-1}c_k+1)}, & x_2^{(\sum_{k=0}^{j-1}c_k+2)}, & \cdots, & x_2^{(\sum_{k=0}^{j}c_k)}\end{bmatrix}\in(a_j+\mathfrak{I}\times\mathfrak{R}_2)^{c_j}$$

is a strongly 2ϵ-typical sequences of length c_j with respect to the random variable $Z_j \sim p_j = \text{emp}(\mathbf{z}_j)$. The sample space of Z_j is $a_j + \mathfrak{I} \times \mathfrak{R}_2$. Therefore, the number of such a $\mathbf{z}'_j$ of length c_j (all elements $\begin{bmatrix} \mathbf{w}_1 \\ \mathbf{w}_2 \end{bmatrix} \in \mathcal{T}_{2\epsilon}(c_j, Z_j)$ such that $\mathbf{w}_2 = \mathbf{x}_{2,j}$) is upper bounded by $2^{c_j[H(p_j)-H(p_{j,2})+2\epsilon]}$, where $p_{j,2}$ is the marginal of p_j with respect to the second coordinate. Consequently,

$$|D_\epsilon(\mathbf{x}_1, \mathfrak{I}|\mathbf{x}_2)| \le 2^{\sum_{j=1}^m c_j[H(p_j)-H(p_{j,2})+2\epsilon]}. \tag{7}$$

Direct computation yields

$$\begin{aligned}
&\frac{1}{n}\sum_{j=1}^m c_j H(p_j) \\
=&\sum_{j=1}^m \frac{c_j}{n} \sum_{r\in a_j+\mathfrak{I}\times\mathfrak{R}_2} \frac{N\left(r,(\mathbf{x}_1,\mathbf{x}_2)^t\right)}{c_j} \log \frac{c_j}{N\left(r,(\mathbf{x}_1,\mathbf{x}_2)^t\right)} \\
=&\sum_{r\in\mathfrak{R}} \frac{N\left(r,(\mathbf{x}_1,\mathbf{x}_2)^t\right)}{n} \log \frac{n}{N\left(r,(\mathbf{x}_1,\mathbf{x}_2)^t\right)} - \sum_{j=1}^m \frac{c_j}{n}\log\frac{n}{c_j}
\end{aligned}$$

and

$$\begin{aligned}
&\frac{1}{n}\sum_{j=1}^m c_j H(p_{j,2}) \\
=&\sum_{j=1}^m \frac{c_j}{n}\Bigg[\sum_{r_2\in\mathfrak{R}_2} \frac{\sum_{r_1\in a_j+\mathfrak{I}} N\left((r_1,r_2),(\mathbf{x}_1,\mathbf{x}_2)^t\right)}{c_j} \\
&\times \log \frac{c_j}{\sum_{r_1\in a_j+\mathfrak{I}} N\left((r_1,r_2),(\mathbf{x}_1,\mathbf{x}_2)^t\right)}\Bigg] \\
=&\sum_{j=1}^m \sum_{r_2\in\mathfrak{R}_2} \frac{\sum_{r_1\in a_j+\mathfrak{I}} N((r_1,r_2),(\mathbf{x}_1,\mathbf{x}_2)^t)}{n} \\
&\times \log \frac{n}{\sum_{r_1\in a_j+\mathfrak{I}} N((r_1,r_2),(\mathbf{x}_1,\mathbf{x}_2)^t)} - \sum_{j=1}^m \frac{c_j}{n}\log\frac{n}{c_j}.
\end{aligned}$$

Since the entropy H is a continuous function, there exists some small $0 < \epsilon < \eta/4$, such that

$$\left|\sum_{r\in\mathfrak{R}} \frac{N\left(r,(\mathbf{x}_1,\mathbf{x}_2)^t\right)}{n} \log \frac{n}{N\left(r,(\mathbf{x}_1,\mathbf{x}_2)^t\right)} - H(X_1,X_2)\right| < \eta/8,$$

$$\left|\sum_{j=1}^m \frac{c_j}{n}\log\frac{n}{c_j} - H(Y_{\mathfrak{R}_1/\mathfrak{I}})\right| < \eta/8 \text{ and}$$

$$\left|\sum_{j=1}^m \sum_{r_2\in\mathfrak{R}_2} \frac{\sum_{r_1\in a_j+\mathfrak{I}} N((r_1,r_2),(\mathbf{x}_1,\mathbf{x}_2)^t)}{n} \times \log \frac{n}{\sum_{r_1\in a_j+\mathfrak{I}} N((r_1,r_2),(\mathbf{x}_1,\mathbf{x}_2)^t)} - H(X_2,Y_{\mathfrak{R}_1/\mathfrak{I}})\right| < \eta/8.$$

Therefore,

$$\frac{1}{n}\sum_{j=1}^{m} c_j H(p_j) < H(X_1, X_2) - H(Y_{\mathfrak{R}_1/\mathfrak{I}}) + \eta/4 \tag{8}$$

$$= H(X_1|X_2) - H(X_2) - H(Y_{\mathfrak{R}_1/\mathfrak{I}}) + \eta/4, \tag{9}$$

$$\frac{1}{n}\sum_{j=1}^{m} c_j H(p_{j,2}) > H(X_2, Y_{\mathfrak{R}_1/\mathfrak{I}}) - H(Y_{\mathfrak{R}_1/\mathfrak{I}}) - \eta/4 \tag{10}$$

$$= H(Y_{\mathfrak{R}_1/\mathfrak{I}}|X_2) - H(X_2) - H(Y_{\mathfrak{R}_1/\mathfrak{I}}) - \eta/4, \tag{11}$$

where (8) and (10) are guaranteed for some small $0 < \epsilon < \eta/4$. Substituting (9) and (11) into (7), (6) follows. ■

Remark 13. Assume that $\mathbf{y} - \mathbf{x} = [a_1, a_2, \cdots, a_n]^t$, then $\mathbf{y} - \mathbf{x} \in \mathfrak{I}^n$ is equivalent to $\langle a_1, a_2, \cdots, a_n\rangle_l \subseteq \mathfrak{I}$.

IV. Proof of Theorem III.1

As mentioned, $\mathscr{X}_i$ can be seen as a subset of $\mathfrak{R}_i$ for a fixed $\Phi = [\Phi_1, \cdots, \Phi_s]$. In this section, we assume that X_i has sample space $\mathfrak{R}_i$, which makes sense since Φ_i is injective.

Let $\mathbf{R} = [R_1, R_2, \cdots, R_s]$ and $k_i = \left\lfloor \frac{nR_i}{\log|\mathfrak{R}_i|} \right\rfloor$, $\forall\ i \in \mathcal{S}$, where n is the length of the data sequences. If $\mathbf{R} \in \mathcal{R}_\Phi$, then $\sum_{i\in T} \frac{R_i \log|\mathfrak{I}_i|}{\log|\mathfrak{R}_i|} > r(T, \mathfrak{I}_T)$, (this implies that $\frac{1}{n}\sum_{i\in T} k_i \log|\mathfrak{I}_i| - r(T, \mathfrak{I}_T) > 2\eta$ for some small constant $\eta > 0$ and large enough n), $\forall\ \emptyset \neq T \subseteq \mathcal{S}, \forall\ 0 \neq \mathfrak{I}_i \leq_l \mathfrak{R}_i$. We claim that $\mathbf{R}$ is achievable.

A. Encoding:

For every $i \in \mathcal{S}$, randomly generate a $k_i \times n$ matrix $\mathbf{A}_i$ based on a uniform distribution, i.e., independently choose each entry of $\mathbf{A}_i$ uniformly at random. Define a linear encoder $\phi_i : \mathfrak{R}_i^n \to \mathfrak{R}_i^{k_i}$ such that

$$\phi_i : \mathbf{x} \mapsto \mathbf{A}_i\mathbf{x}, \forall\ \mathbf{x} \in \mathfrak{R}_i^n.$$

Obviously the coding rate of this encoder is $\frac{1}{n}\log|\phi_i(\mathfrak{R}_i^n)| \leq \frac{1}{n}\log|\mathfrak{R}_i|^{k_i} = \frac{\log|\mathfrak{R}_i|}{n}\left\lfloor \frac{nR_i}{\log|\mathfrak{R}_i|} \right\rfloor \leq R_i$.

B. Decoding:

Subject to observing $\mathbf{y}_i \in \mathfrak{R}_i^{k_i}$ ($i \in \mathcal{S}$) from the ith encoder, the decoder claims that $\mathbf{x} = [\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_s]^t \in \prod_{i=1}^{s} \mathfrak{R}_i^n$ is the array of the encoded data sequences, if and only if:

1) $\mathbf{x} \in \mathcal{T}_\epsilon$; and
2) $\forall\ \mathbf{x}' = [\mathbf{x}'_1, \mathbf{x}'_2, \cdots, \mathbf{x}'_s]^t \in \mathcal{T}_\epsilon$, if $\mathbf{x}' \neq \mathbf{x}$, then $\phi_j(\mathbf{x}'_j) \neq \mathbf{y}_j$, for some j.

C. Error:

Assume that $\mathbf{X}_i \in \mathfrak{R}_i^n$ ($i \in \mathcal{S}$) is the original data sequence generated by the ith source. It is readily seen that an error occurs if and only if:

E_1: $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_s]^t \notin \mathcal{T}_\epsilon$; or

E_2: There exists $\mathbf{X} \neq [\mathbf{x}'_1, \mathbf{x}'_2, \cdots, \mathbf{x}'_s]^t \in \mathcal{T}_\epsilon$, such that $\phi_i(\mathbf{x}'_i) = \phi_i(\mathbf{X}_i)$, $\forall\ i \in \mathcal{S}$.

D. Error Probability:

By the joint AEP, $\Pr\{E_1\} \to 0$, $n \to \infty$.

Additionally, for $\emptyset \neq T \subseteq \mathcal{S}$, let

$$D_\epsilon(\mathbf{X}; T) = \Big\{ [\mathbf{x}'_1, \mathbf{x}'_2, \cdots, \mathbf{x}'_s]^t \in \mathcal{T}_\epsilon \Big| \\ \mathbf{x}'_i \neq \mathbf{X}_i \text{ if and only if } i \in T \Big\}.$$

We have

$$D_\epsilon(\mathbf{X}; T) = \bigcup_{0 \neq \mathfrak{I} \leq_l \mathfrak{R}_T} \left[D_\epsilon(\mathbf{X}_T, \mathfrak{I} | \mathbf{X}_{T^c}) \setminus \{\mathbf{X}\} \right], \tag{12}$$

where $\mathbf{X}_T = \prod_{i \in T} \mathbf{X}_i$ and $\mathbf{X}_{T^c} = \prod_{i \in T^c} \mathbf{X}_i$, since $\mathfrak{I}$ goes over all possible non-trivial left ideals. Consequently,

$$\begin{aligned}
&\Pr\{E_2 | E_1^c\} \\
&= \sum_{[\mathbf{x}'_1, \cdots, \mathbf{x}'_s]^t \in \mathcal{T}_\epsilon \setminus \{\mathbf{X}\}} \prod_{i \in \mathcal{S}} \Pr\{\phi_i(\mathbf{x}'_i) = \phi_i(\mathbf{X}_i) | E_1^c\} \\
&= \sum_{\emptyset \neq T \subseteq \mathcal{S}} \sum_{\substack{[\mathbf{x}'_1, \cdots, \mathbf{x}'_s]^t \\ \in D_\epsilon(\mathbf{X}; T)}} \prod_{i \in T} \Pr\{\phi_i(\mathbf{x}'_i) = \phi_i(\mathbf{X}_i) | E_1^c\} && (13) \\
&\leq \sum_{\emptyset \neq T \subseteq \mathcal{S}} \sum_{0 \neq \mathfrak{I} \leq_l \mathfrak{R}_T} \sum_{\substack{[\mathbf{x}'_1, \cdots, \mathbf{x}'_s]^t \\ \in D_\epsilon(\mathbf{X}_T, \mathfrak{I} | \mathbf{X}_{T^c}) \setminus \{\mathbf{X}\}}} \prod_{i \in T} \Pr\{\phi_i(\mathbf{x}'_i) = \phi_i(\mathbf{X}_i) | E_1^c\} && (14) \\
&< \sum_{\emptyset \neq T \subseteq \mathcal{S}} \sum_{\substack{0 \neq \prod_{i \in T} \mathfrak{I}_i \\ \leq_l \mathfrak{R}_T}} \left(2^{n[r(T, \mathfrak{I}) + \eta]} - 1 \right) \prod_{i \in T} |\mathfrak{I}_i|^{-k_i} && (15) \\
&< (2^s - 1) \left(2^{|R_{\mathcal{S}}|} - 2 \right) \times \max_{\substack{\emptyset \neq T \in \mathcal{S}, \\ 0 \neq \prod_{i \in T} \mathfrak{I}_i \leq_l \mathfrak{R}_T}} 2^{-n \left[\frac{1}{n} \sum_{i \in T} k_i \log |\mathfrak{I}_i| - [r(T, \mathfrak{I}) + \eta] \right]}, && (16)
\end{aligned}$$

where

(13) is from the fact that $\mathcal{T}_\epsilon \setminus \{\mathbf{X}\} = \coprod_{\emptyset \neq T \subseteq \mathcal{S}} D_\epsilon(\mathbf{X}; T)$ (disjoint union);

(14) follows from (12) by the union bound (Boole's inequality);

(15) is from Lemma III.3 and Lemma III.5, as well as the fact that every left ideal of $\mathfrak{R}_T$ is a Cartesian product of some left ideals $\mathfrak{I}_i$ of $\mathfrak{R}_i$, $i \in T$ (see Proposition II.8). At the same time, ϵ is required to be sufficiently small;

(16) is due to the facts that the number of non-empty subsets of $\mathcal{S}$ is $2^s - 1$ and the number of non-trivial left ideals of the finite ring $\mathfrak{R}_T$ is less than $2^{|R_{\mathcal{S}}|} - 1$, which is the number of non-empty subsets of $\mathfrak{R}_{\mathcal{S}}$ ($\supseteq \mathfrak{R}_T$).

Thus, $\Pr\{E_2 | E_1^c\} \to 0$, when $n \to \infty$, from (16), since for sufficiently large n and small ϵ, $\frac{1}{n} \sum_{i \in T} k_i \log |\mathfrak{I}_i| - [r(T, \mathfrak{I}) + \eta] > \eta > 0$.

Therefore, $\Pr\{E_1 \cup E_2\} = \Pr\{E_1\} + \Pr\{E_2 | E_1^c\} \to 0$ as $\epsilon \to 0$ and $n \to \infty$.

V. Optimality

Some optimality results regarding region (4) are presented in this section. First of all, it is proved that Theorem III.1 specializes to its field counterpart in the following result.

Theorem V.1. *Region* (3) *is the SW region if $\mathfrak{R}_i$ contains no proper non-trivial left ideal, equivalently*[2]*, $\mathfrak{R}_i$ is a field, for all $i \in \mathcal{S}$. As a consequence, region* (4) *is the SW region.*

Proof: In Theorem III.1, random variable $Y_{\mathfrak{R}_T/\mathfrak{I}_T}$ admits a sample space of cardinality 1 for all $\emptyset \neq T \subseteq \mathcal{S}$, since the only non-trivial left ideal of $\mathfrak{R}_i$ is itself for all feasible i. Thus, $0 = H(Y_{\mathfrak{R}_T/\mathfrak{I}_T}) \geq H(Y_{\mathfrak{R}_T/\mathfrak{I}_T}|X_{T^c}) \geq 0$. Consequently,

$$\mathcal{R}_\Phi = \Big\{ [R_1, R_2, \cdots, R_s] \in \mathbb{R}^s \Big| \sum_{i\in T} R_i > H(X_T|X_{T^c}), \\ \forall\, \emptyset \neq T \subseteq \mathcal{S} \Big\},$$

which is the SW region $\mathcal{R}[X_1, X_2, \cdots, X_s]$. Therefore, region (4) is also the SW region.

If $\mathfrak{R}_i$ is a field, then obviously it has no proper non-trivial left (right) ideal. Conversely, $\forall\ 0 \neq a \in \mathfrak{R}_i$, $\langle a \rangle_l = \mathfrak{R}_i$ implies that $\exists\ 0 \neq b \in \mathfrak{R}_i$, such that $ba = 1$. Similarly, $\exists\ 0 \neq c \in \mathfrak{R}_i$, such that $cb = 1$. Moreover, $c = c \cdot 1 = cba = 1 \cdot a = a$. Hence, $ab = cb = 1$. b is the inverse of a. By Wedderburn's little theorem, $\mathfrak{R}_i$ is a field. ∎

Remark 14. Theorem V.1 states that LCoF of Elias [1] and Csiszár [2] are special cases of Theorem III.1 in the sense of achieving the optimal coding rate region, the SW region. However, $\mathfrak{R}_1, \mathfrak{R}_2, \cdots, \mathfrak{R}_s$ being fields is not necessary. As shown in Theorem V.2, the SW region can be achieved using rings such as $\mathbb{Z}_4$ and $\mathbb{M}_L = \left\{ \begin{bmatrix} x & 0 \\ y & x \end{bmatrix} \middle| x, y \in \mathbb{Z}_2 \right\}$ ($\mathbb{M}_L$ is a ring with respect to matrix addition and multiplication). Clearly, neither $\mathbb{Z}_4$ nor $\mathbb{M}_L$ is a field.

A. Single Source

In the single source scenario, showing that the convex hull (4) is the SW region is equivalent to proving that

$$\max_{0\neq\mathfrak{I}_1\leq_l\mathfrak{R}_1} \frac{\log|\mathfrak{R}_1|}{\log|\mathfrak{I}_1|} \left[H(X) - H(Y_{\mathfrak{R}_1/\mathfrak{I}_1})\right] = H(X) \tag{17}$$

for some $\Phi \in \mathscr{M}(\mathscr{X}_\mathcal{S}, \mathfrak{R}_\mathcal{S})$. As mentioned before, $Y_{\mathfrak{R}_1/\mathfrak{I}_1}$ depends on Φ_1 (note that $\Phi = [\Phi_1]$ for single source). Hence, an appropriate Φ_1 is a crucial.

Theorem V.2. *If $s = 1$ and $\mathfrak{R}_1$ is of order 4 containing one and only one proper non-trivial left ideal (equivalently*[3]*, $\mathfrak{R}_1$ is isomorphic to either $\mathbb{Z}_4$ or $\mathbb{M}_L$), then region* (4) *is the SW region, i.e., there exists $\Phi = [\Phi_1] \in \mathscr{M}(\mathscr{X}_1, \mathfrak{R}_1)$, such that* (17) *holds.*

[2]Equivalency does not necessarily hold for rngs.

[3]Equivalency does not hold necessarily for rngs.

Proof: $\mathfrak{R}_1$ has only two non-trivial left ideals, $\mathfrak{R}_1$ and the unique proper non-trivial left ideal $\mathfrak{J}$. Moreover, $|\mathfrak{J}| = 2$ since $|\mathfrak{R}_1| = 4$. If $\mathfrak{I}_1 = \mathfrak{R}_1$, then $\Pr\left\{Y_{\mathfrak{R}_1/\mathfrak{I}_1} = \mathfrak{R}_1\right\} = 1$, thus, $H(Y_{\mathfrak{R}_1/\mathfrak{I}_1}) = 0$ and

$$\frac{\log|\mathfrak{R}_1|}{\log|\mathfrak{I}_1|}\left[H(X) - H(Y_{\mathfrak{R}_1/\mathfrak{I}_1})\right] = H(X).$$

Thus, (17) is valid if and only if

$$\begin{aligned}\frac{\log|\mathfrak{R}_1|}{\log|\mathfrak{J}|}\left[H(X) - H(Y_{\mathfrak{R}_1/\mathfrak{J}})\right] \leq & H(X)\\ \Leftrightarrow H(X) \leq & 2H(Y_{\mathfrak{R}_1/\mathfrak{J}}).\end{aligned} \tag{18}$$

for some injection $\Phi_1 : \mathscr{X}_1 \to \mathfrak{R}_1$.

Without loss of generality, assume that $\mathfrak{R}_1 = \{1, 0, a, b\}$, $\mathfrak{J} = \{0, a\}$, $\mathscr{X}_1 = \{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$ and $0 \leq p(\alpha_4) \leq p(\alpha_3) \leq p(\alpha_2) \leq p(\alpha_1) \leq 1$ (note if $|\mathscr{X}_1| < 4$, then $p(\alpha_4)$, $p(\alpha_3)$ or $p(\alpha_2)$ is simply 0). Let

$$\Phi_1 : \alpha_1 \mapsto 1;\ \ \Phi_1 : \alpha_2 \mapsto 0;\ \ \Phi_1 : \alpha_3 \mapsto a;\ \ \Phi_1 : \alpha_4 \mapsto b.$$

Then $\Pr\left\{Y_{\mathfrak{R}_1/\mathfrak{J}} = \mathfrak{J}\right\} = p(0) + p(a) = p(\alpha_2) + p(\alpha_3)$ and $\Pr\left\{Y_{\mathfrak{R}_1/\mathfrak{J}} = 1 + \mathfrak{J}\right\} = p(1) + p(b) = p(\alpha_1) + p(\alpha_4)$ (note that after fixing Φ_1, $\mathscr{X}_1$ is seen as a subset of $\mathfrak{R}_1$). Consequently, (18) is equivalent to

$$\begin{aligned}&-\sum_{j=1}^{4} p(\alpha_j)\log p(\alpha_j)\\ \leq &-2\left[p(\alpha_2) + p(\alpha_3)\right]\log\left[p(\alpha_2) + p(\alpha_3)\right]\\ &-2\left[p(\alpha_1) + p(\alpha_4)\right]\log\left[p(\alpha_1) + p(\alpha_4)\right],\end{aligned}$$

which is established by Lemma A.1. Therefore, (18) holds, so does (17). Furthermore, that $\mathfrak{R}_1$ is isomorphic to either $\mathbb{Z}_4$ or $\mathbb{M}_L$ is a well-known fact [12]. The theorem is proved. ■

Remark 15. Notice that Theorem V.2 is valid for any $X \sim p$. However, for other rings (in particular for those containing only one left (right) ideal besides the trivial ideal and itself, for instance, $\mathbb{Z}_{q^2}$ with prime q), no conclusive result is so far available. However, that optimality, i.e., (17) holds, can still be shown for certain distribution p.

Remark 16. Up to isomorphism, there are 4 rings (and 11 rngs) of order 4 [12]. On the contrary, there is one and only one finite field of any finite order. Theorem V.2 suggests that LC over rings, e.g., $\mathbb{F}_4$, $\mathbb{Z}_4$ and $\mathbb{M}_L$ can be as good as LC over fields in the single source SW source coding problem. Actually, due to its flexible structure (a ring does not have to be commutative, can have a non-prime characteristic, etc), LCoR offers certain advantages. One example is shown in Section VI demonstrating this.

B. Multiple Sources

Based on Theorem V.2, the following corollary can be verified immediately.

Corollary V.3. *Region* (4) *contains the region*

$$\left\{(R_1, R_2, \cdots, R_s) \in \mathbb{R}^s \,|R_i > H(X_i), \forall\, i \in \mathcal{S}\right\},$$

if $\mathfrak{R}_i$ is isomorphic to either $\mathbb{Z}_4$, $\mathbb{M}_L$ or $\mathbb{F}_4$ for all $1 \leq i \leq s$.

For the multiple sources scenario, we do not have conclusive result like Theorem V.2 that is valid for any distribution. However, the following theorem makes it more plausible that there exists some set of non-field rings over which LC is optimal.

Theorem V.4. *Regardless which set of rings $\mathfrak{R}_1, \mathfrak{R}_2, \cdots, \mathfrak{R}_s$ is chosen, as long as $|\mathfrak{R}_i| = |\mathscr{X}_i|$ for all feasible i, region* (3) *is the SW region if $(X_1, X_2, \cdots, X_s) \sim p$ is a uniform distribution.*

Proof: If p is uniform, then, for any $\emptyset \neq T \subseteq \mathcal{S}$ and $0 \neq \mathfrak{I}_T \leq_l \mathfrak{R}_T$, $Y_{\mathfrak{R}_T/\mathfrak{I}_T}$ is uniformly distributed on $\mathfrak{R}_T/\mathfrak{I}_T$. Moreover, X_T and X_{T^c} are independent, so are $Y_{\mathfrak{R}_T/\mathfrak{I}_T}$ and X_{T^c}. Therefore, $H(X_T|X_{T^c}) = H(X_T) = \log|\mathfrak{R}_T|$ and $H(Y_{\mathfrak{R}_T/\mathfrak{I}_T}|X_{T^c}) = H(Y_{\mathfrak{R}_T/\mathfrak{I}_T}) = \log\dfrac{|\mathfrak{R}_T|}{|\mathfrak{I}_T|}$. Consequently,

$$r(T, \mathfrak{I}_T) = H(X_T|X_{T^c}) - H(Y_{\mathfrak{R}_T/\mathfrak{I}_T}|X_{T^c}) = \log|\mathfrak{I}_T|.$$

Region (3) is the SW region. ∎

Remark 17. When p is uniform, it is obvious that the uncoded strategy (all encoders are one-to-one mappings) is optimal in the SW source coding problem. However, optimality stated in Theorem V.4 does not come from deliberately fixing the encoding mappings, but generating them randomly.

VI. Application: Source Coding for Computing

Some advantages of LCoR are demonstrated in this section. In Example VI.5 below, we show that LCoR achieves better coding rates compared to LCoF in the sense of Körner–Marton [3] for some function g in Problem 1. At the same time, the encoders using LCoR require strictly smaller alphabet sizes than using LCoF. We first present the following theorem which is seen as a generalization of Körner–Marton [3].

Theorem VI.1. *In Problem 1, if $\hat{g}$ is a polynomial function in $\mathfrak{R}[s]$ admitting that*

$$\hat{g} = h \circ k, \text{ where } k(x_1, x_2, \cdots, x_s) = \sum_{i=1}^{s} k_i(x_i), \tag{19}$$

and $h, k_i \in \mathfrak{R}[1]$ for all feasible i, then

$$\begin{aligned}\mathcal{R}_{\hat{g}} = &\Big\{(r, r, \cdots, r) \in \mathbb{R}^s \Big| r > \\ &\max_{0 \neq \mathfrak{I} \leq_l \mathfrak{R}} \frac{\log|\mathfrak{R}|}{\log|\mathfrak{I}|}\left[H(X) - H(Y_{\mathfrak{R}/\mathfrak{I}})\right]\Big\} \subseteq \mathcal{R}[\hat{g}],\end{aligned} \tag{20}$$

where $X = k(X_1, X_2, \cdots, X_s)$ and $Y_{\mathfrak{R}/\mathfrak{I}}$ is a random variable with sample space $\mathfrak{R}/\mathfrak{I}$ and

$$\Pr\left\{Y_{\mathfrak{R}/\mathfrak{I}} = a + \mathfrak{I} \middle| X = x\right\} = \begin{cases} 1; & \text{if } x \in a + \mathfrak{I}, \\ 0; & \text{otherwise}. \end{cases}$$

Proof: By Theorem III.1, $\forall\, \epsilon > 0$, there exists a large enough n, an $m \times n$ matrix $\mathbf{A} \in \mathfrak{R}^{m \times n}$ and a decoder ψ, such that $\Pr\{X^n \neq \psi(\mathbf{A}X^n)\} < \epsilon$, if $m > \max\limits_{0 \neq \mathfrak{I} \leq_l \mathfrak{R}} \dfrac{n(H(X) - H(Y_{\mathfrak{R}/\mathfrak{I}}))}{\log|\mathfrak{I}|}$. Let $\phi_i = \mathbf{A} \circ \vec{k}_i$ $(1 \leq i \leq s)$ be

the encoder of the ith source. Upon receiving $\phi_i(X_i^n)$ from the ith source, the decoder claims that $\vec{h}\left(\hat{X}^n\right)$, where $\hat{X}^n = \psi\left[\sum_{i=1}^{s}\phi_i\left(X_i^n\right)\right]$, is the function, namely $\hat{g}$, subject to computation. The probability of decoding error is

$$\begin{aligned}
&\Pr\left\{\vec{h}\left[\vec{k}\left(X_1^n, X_2^n, \cdots, X_s^n\right)\right] \neq \vec{h}\left(\hat{X}^n\right)\right\} \\
\leq &\Pr\left\{X^n \neq \hat{X}^n\right\} \\
= &\Pr\left\{X^n \neq \psi\left[\sum_{i=1}^{s}\phi_i\left(X_i^n\right)\right]\right\} \\
= &\Pr\left\{X^n \neq \psi\left[\sum_{i=1}^{s}\mathbf{A}\vec{k}_i\left(X_i^n\right)\right]\right\} \\
= &\Pr\left\{X^n \neq \psi\left[\mathbf{A}\sum_{i=1}^{s}\vec{k}_i\left(X_i^n\right)\right]\right\} \\
= &\Pr\left\{X^n \neq \psi\left[\mathbf{A}\vec{k}\left(X_1^n, X_2^n, \cdots, X_s^n\right)\right]\right\} \\
= &\Pr\left\{X^n \neq \psi\left(\mathbf{A}X^n\right)\right\} < \epsilon.
\end{aligned}$$

Therefore, all $(r, r, \cdots, r) \in \mathbb{R}^s$, where $r = \dfrac{m\log|\mathfrak{R}|}{n} > \max\limits_{0\neq\mathfrak{I}\leq_l\mathfrak{R}}\dfrac{\log|\mathfrak{R}|}{\log|\mathfrak{I}|}\left[H(X) - H(Y_{\mathfrak{R}/\mathfrak{I}})\right]$, is achievable, i.e., $\mathcal{R}_{\hat{g}} \subseteq \mathcal{R}[\hat{g}]$. ∎

Corollary VI.2. *In Theorem VI.1, let $X = k(X_1, X_2, \cdots, X_s) \sim p_X$. We have*

$$\mathcal{R}_{\hat{g}} = \left\{(r, r, \cdots, r) \in \mathbb{R}^s \middle| \, r > H(X)\right\} \subseteq \mathcal{R}[\hat{g}],$$

if either of the following conditions holds:

1) *$\mathfrak{R}$ is isomorphic to some finite field;*
2) *$\mathfrak{R}$ is isomorphic to $\mathbb{Z}_4$ and*

$$p_X(0) = p_1, p_X(1) = p_2, p_X(3) = p_3 \text{ and } p_X(2) = p_4$$

satisfying (A.1)*;*

3) *$\mathfrak{R}$ is isomorphic to $\mathbb{M}_L$ and*

$$p_X\left(\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}\right) = p_1, p_X\left(\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\right) = p_2, p_X\left(\begin{bmatrix}1 & 0\\ 1 & 1\end{bmatrix}\right) = p_3 \text{ and } p_X\left(\begin{bmatrix}0 & 0\\ 1 & 0\end{bmatrix}\right) = p_4$$

satisfying (A.1)*.*

Proof: If either 1), 2) or 3) holds, then Theorem V.1 or Lemma A.1 guarantee that

$$\max_{0\neq\mathfrak{I}\leq_l\mathfrak{R}}\frac{\log|\mathfrak{R}|}{\log|\mathfrak{I}|}\left[H(X) - H(Y_{\mathfrak{R}/\mathfrak{I}})\right] = H(X)$$

in Theorem VI.1. The statement follows. ∎

Remark 18. If $\mathfrak{R}$ is isomorphic to $\mathbb{Z}_2$ and $\hat{g}$ is the modulo-two sum, then Corollary VI.2 recovers the theorem of Körner–Marton [3]. While if $\mathfrak{R}$ is (isomorphic to) a field, it becomes a special case of [7, Theorem III.1]. Actually, almost all the results in [6] and [7] can be reproved in the setting of rings in a parallel fashion.

Definition VI.3. Let $g_1 : \prod_{i=1}^{s} \mathscr{X}_i \to \Omega_1$ and $g_2 : \prod_{i=1}^{s} \mathscr{Y}_i \to \Omega_2$ be two functions. If there exist bijections $\mu_i : \mathscr{X}_i \to \mathscr{Y}_i$, $\forall\ 1 \leq i \leq s$, and $\nu : \Omega_1 \to \Omega_2$, such that

$$g_1(x_1, x_2, \cdots, x_s) = \nu^{-1}(g_2(\mu_1(x_1), \mu_2(x_2), \cdots, \mu_s(x_s))),$$

then g_1 and g_2 are said to be *equivalent* (via $\mu_1, \mu_2, \cdots, \mu_s$ and ν).

Definition VI.4. Given function $g : \mathscr{D} \to \Omega$, and let $\emptyset \neq \mathscr{S} \subseteq \mathscr{D}$. The *restriction* of g on $\mathscr{S}$ is defined to be the function $g|_{\mathscr{S}} : \mathscr{S} \to \Omega$ such that $g|_{\mathscr{S}} : x \mapsto g(x), \forall\ x \in \mathscr{S}$.

Remark 19. Up to equivalence, a function can be presented in many different formats. For example, the function $\min\{x, y\}$ defined on $\{0,1\} \times \{0,1\}$ can either be seen as $F_1(x, y) = xy$ on $\mathbb{Z}_2^2$ or be treated as the restriction of $F_2(x, y) = x + y - (x + y)^2$ defined on $\mathbb{Z}_3^2$ to the domain $\{0,1\} \times \{0,1\} \subsetneq \mathbb{Z}_3^2$. We refer to each presented format of a function as a *presentation* of this function.

Assume that g has presentation $\hat{g} \in \mathfrak{R}[s]$ for some finite ring $\mathfrak{R}$, we say that the region $\mathcal{R}_{\hat{g}}$ given by (20) is achievable for computing g in the sense of Körner–Marton [3]. From [7], we know that $\mathcal{R}_{\hat{g}}$ might not the largest achievable region one can obtain for computing g. However, $\mathcal{R}_{\hat{g}}$ still captures the ability of LC over $\mathfrak{R}$ when used for computing g. More precisely, $\mathcal{R}_{\hat{g}}$ is the region purely achieved with LC over $\mathfrak{R}$ for computing g. On the other hand, regions from [7] are achieved by combining the LC and the standard random coding techniques. Therefore, it is reasonable to compare LCoR with LCoF in the sense of Körner–Marton[4].

Making use of Theorem VI.1 and Corollary VI.2, we show that LCoR strictly outperforms its field counterpart in the following example.

Example VI.5 ([13])**.** Let $g : \{\alpha_0, \alpha_1\}^3 \to \{\beta_0, \beta_1, \beta_2, \beta_3\}$ (Fig 1) be a function such that

$$\begin{aligned}
&g : (\alpha_0, \alpha_0, \alpha_0) \mapsto \beta_0; \quad g : (\alpha_0, \alpha_0, \alpha_1) \mapsto \beta_3;\\
&g : (\alpha_0, \alpha_1, \alpha_0) \mapsto \beta_2; \quad g : (\alpha_0, \alpha_1, \alpha_1) \mapsto \beta_1;\\
&g : (\alpha_1, \alpha_0, \alpha_0) \mapsto \beta_1; \quad g : (\alpha_1, \alpha_0, \alpha_1) \mapsto \beta_0;\\
&g : (\alpha_1, \alpha_1, \alpha_0) \mapsto \beta_3; \quad g : (\alpha_1, \alpha_1, \alpha_1) \mapsto \beta_2.
\end{aligned} \tag{21}$$

Define $\mu : \{\alpha_0, \alpha_1\} \to \mathbb{Z}_4$ and $\nu : \{\beta_0, \beta_1, \beta_2, \beta_3\} \to \mathbb{Z}_4$ by

$$\begin{aligned}
&\mu : \alpha_j \mapsto j, \quad \forall\ j \in \{0,1\}, \text{ and}\\
&\nu : \beta_j \mapsto j, \quad \forall\ j \in \{0,1,2,3\},
\end{aligned} \tag{22}$$

[4]In fact, the authors believe that LCoR outperforms LCoF even in the sense of combining the LC and the standard random coding techniques [7].

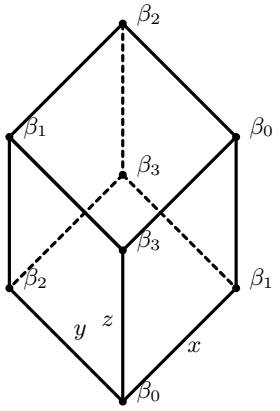

Fig 1: $g : \{\alpha_0, \alpha_1\}^3 \to \{\beta_0, \beta_1, \beta_2, \beta_3\}$

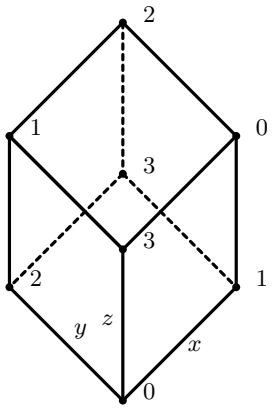

Fig 2: $x + 2y + 3z \in \mathbb{Z}_4[3]$

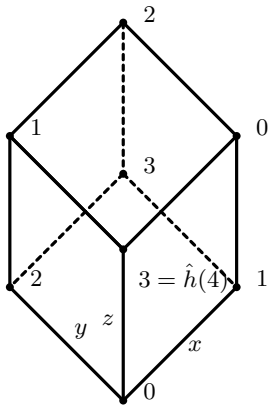

Fig 3: $\hat{h}(x + 2y + 4z) \in \mathbb{Z}_5[3]$

respectively. Obviously, g is equivalent to $x + 2y + 3z \in \mathbb{Z}_4[3]$ (Fig 2) via $\mu_1 = \mu_2 = \mu_3 = \mu$ and ν. However, by Proposition VI.6, there exists no $\hat{g} \in \mathbb{F}_4[3]$ of format (19) so that g is equivalent to any restriction of $\hat{g}$. Although, by [6, Lemma A.2], there always exists a bigger field $\mathbb{F}_q$ such that g admits a presentation for some $\hat{g} \in \mathbb{F}_q[3]$ of format (19), the size q must be strictly bigger than 4. For instance, let $\hat{h}(x) = \sum_{a \in \mathbb{Z}_5} a \left[1 - (x-a)^4\right] - \left[1 - (x-4)^4\right] \in \mathbb{Z}_5[1]$. Then, g has presentation $\hat{h}(x + 2y + 4z) \in \mathbb{Z}_5[3]$ (Fig 3) via $\mu_1 = \mu_2 = \mu_3 = \mu : \{\alpha_0, \alpha_1\} \to \mathbb{Z}_5$ and $\nu : \{\beta_0, \beta_1, \beta_2, \beta_3\} \to \mathbb{Z}_5$ defined (symbolic-wise) by (22).

Proposition VI.6. *There exists no polynomial function $\hat{g} \in \mathbb{F}_4[3]$ of format* (19)*, such that a restriction of $\hat{g}$ is equivalent to the function g defined by* (21)*.*

Proof: Suppose $\nu \circ g = \hat{g} \circ (\mu_1, \mu_2, \mu_3)$, where $\mu_1, \mu_1, \mu_3 : \{\alpha_0, \alpha_1\} \to \mathbb{F}_4$, $\nu : \{\beta_0, \cdots, \beta_3\} \to \mathbb{F}_4$ are injections and $\hat{g} = h \circ (k_1 + k_2 + k_3)$ with $h, k_i \in \mathbb{F}_4[1]$ for all feasible i. We claim that $\hat{g}$ and h are both surjective, since $\left|g\left(\{\alpha_0, \alpha_1\}^3\right)\right| = |\{\beta_0, \beta_1, \beta_2, \beta_3\}| = 4 = |\mathbb{F}_4|$. In particular, h is bijective. Therefore, $h^{-1} \circ \nu \circ g = k_1 \circ \mu_1 + k_2 \circ \mu_2 + k_3 \circ \mu_3$, i.e., g admits a presentation $k_1(x) + k_2(y) + k_3(z) \in \mathbb{F}_4[3]$. A contradiction to Lemma A.2. ∎

As a consequence of Proposition VI.6, in order to use LCoF in the sense of Körner–Marton to compute function g, the alphabet sizes of the three encoders need to be at least 5. However, LCoR offers a solution in which the alphabet sizes are 4, strictly smaller than using LCoF. In addition, in the sense of Körner–Marton, the region achieved with LC over a finite field $\mathbb{F}_q$, is always a subset of the one achieved with LC over $\mathbb{Z}_4$. This is proved in the following proposition.

Proposition VI.7. *Let g be the function defined by* (21)*, $(X_1, X_2, X_3) \sim p$ and p_X be the distribution of $X = g(X_1, X_2, X_3)$. If*

$$p_X(\beta_0) = p_1, p_X(\beta_1) = p_2, p_X(\beta_3) = p_3 \text{ and } p_X(\beta_2) = p_4$$

satisfying (A.1)*, then, in the sense of Körner–Marton, the region $\mathcal{R}_1$ achieved with LC over $\mathbb{Z}_4$ contains the one,*

that is $\mathcal{R}_2$, obtained with LC over any finite field $\mathbb{F}_q$ for computing g. Moreover, if $\mathrm{supp}(p)$ *is the whole domain of g, then $\mathcal{R}_1 \supsetneq \mathcal{R}_2$.*

Proof: Let $\hat{g} = h \circ k \in \mathbb{F}_q[3]$ be a polynomial presentation of g with format (19) guaranteed by [6, Lemma A.2]. By Corollary VI.2, we have

$$\mathcal{R}_1 = \left\{ (R_1, R_2, R_3) \in \mathbb{R}^3 \middle| \, R_i > H(X_1 + 2X_2 + 3X_3) \right\},$$
$$\mathcal{R}_2 = \left\{ (R_1, R_2, R_3) \in \mathbb{R}^3 \middle| \, R_i > H(k(X_1, X_2, X_3)) \right\}.$$

Assume that $\nu \circ g = h \circ k \circ (\mu_1, \mu_2, \mu_3)$, where $\mu_1, \mu_1, \mu_3 : \{\alpha_0, \alpha_1\} \to \mathbb{F}_q$ and $\nu : \{\beta_0, \cdots, \beta_3\} \to \mathbb{F}_q$ are injections. Obviously, $g(X_1, X_2, X_3)$ is a function of $k(X_1, X_2, X_3)$. Hence,

$$H(k(X_1, X_2, X_3)) \geq H(g(X_1, X_2, X_3)). \tag{23}$$

On the other hand, $H(X_1 + 2X_2 + 3X_3) = H(g(X_1, X_2, X_3))$, since $x + 2y + 3z \in \mathbb{Z}_4[3]$ is equivalent to g. Therefore,

$$H(k(X_1, X_2, X_3)) \geq H(X_1 + 2X_2 + 3X_3), \tag{24}$$

and $\mathcal{R}_1 \supseteq \mathcal{R}_2$. In addition, we claim that $h|_{\mathscr{S}}$, where $\mathscr{S} = k\left(\prod_{j=1}^{3} \mu_j\{\alpha_0, \alpha_1\}\right)$, is not injective. Otherwise, $h : \mathscr{S} \to \mathscr{S}'$, where $\mathscr{S}' = h(\mathscr{S})$, is bijective, hence, $(h|_{\mathscr{S}'})^{-1} \circ \nu \circ g = k \circ (\mu_1, \mu_2, \mu_3) = k_1 \circ \mu_1 + k_2 \circ \mu_2 + k_3 \circ \mu_3$. A contradiction to Lemma A.2. Consequently, $|\mathscr{S}| > |\mathscr{S}'| = |\nu(\{\beta_0, \cdots, \beta_3\})| = 4$. If $\mathrm{supp}(p) = \{\alpha_0, \alpha_1\}^3$, then (23) as well as (24) hold strictly, thus, $\mathcal{R}_1 \supsetneq \mathcal{R}_2$. ∎

A more intuitive comparison can be identified from the presentations of g given in Fig 2 and Fig 3. According to Corollary VI.2, LC over field $\mathbb{Z}_5$ achieves the region

$$\left\{ (R_1, R_2, R_3) \in \mathbb{R}^3 \middle| \, R_i > H(X_1 + 2X_2 + 4X_3) \right\}. \tag{25}$$

The one achieved by LC over ring $\mathbb{Z}_4$ is

$$\left\{ (R_1, R_2, R_3) \in \mathbb{R}^3 \middle| \, R_i > H(X_1 + 2X_2 + 3X_3) \right\}. \tag{26}$$

Clearly, $H(X_1 + 2X_2 + 3X_3) \leq H(X_1 + 2X_2 + 4X_3)$, thus, (26) contains (25). Furthermore, as long as

$$0 < \Pr(\alpha_0, \alpha_0, \alpha_1), \Pr(\alpha_1, \alpha_1, \alpha_0) < 1$$

(e. g., $(X_1, X_2, X_3) \sim p$ is a uniform distribution), (26) is strictly larger than (25), since $H(X_1 + 2X_2 + 3X_3) < H(X_1 + 2X_2 + 4X_3)$.

Based on Proposition VI.6 and Proposition VI.7, we conclude that LCoR dominates LCoF, in terms of achieving better coding rates with smaller alphabet sizes of the encoders for computing g.

Remark 20. With proof similar to Proposition VI.7, one can show that LCoR outperforms LCoF in computing g given by (21) in the sense of [7], namely combining the standard and LC techniques.

VII. Conclusion

Careful readers might have noticed that the encoders used in Theorem III.1 are left linear mappings. An almost identical theorem (Theorem VII.1) can be easily proved when using right linear mappings as encoders.

Theorem VII.1. *Given $\Phi \in \mathscr{M}\left(\mathscr{X}_{\mathcal{S}}, \mathfrak{R}_{\mathcal{S}}\right)$ and let*

$$\mathcal{R}'_{\Phi} = \left\{ [R_1, R_2, \cdots, R_s] \in \mathbb{R}^s \left| \sum_{i \in T} \frac{R_i \log |\mathfrak{I}_i|}{\log |\mathfrak{R}_i|} > r\left(T, \mathfrak{I}_T\right), \right. \right.$$
$$\left. \left. \forall\, \emptyset \neq T \subseteq \mathcal{S}, \forall\, 0 \neq \mathfrak{I}_i \leq_r \mathfrak{R}_i \right\}, \tag{27}$$

where $r\left(T, \mathfrak{I}_T\right) = H(X_T|X_{T^c}) - H(Y_{\mathfrak{R}_T/\mathfrak{I}_T}|X_{T^c})$ and $Y_{\mathfrak{R}_T/\mathfrak{I}_T}$ is a random variable with sample space $\mathfrak{R}_T/\mathfrak{I}_T$ and

$$\Pr\left\{ Y_{\mathfrak{R}_T/\mathfrak{I}_T} = a + \mathfrak{I}_T \middle| \, X_T = x_T \right\} = \begin{cases} 1; & \text{if } x_T \in a + \mathfrak{I}_T, \\ 0; & \text{otherwise}. \end{cases}$$

Any rate in $\mathcal{R}'_{\Phi}$ is achievable with linear coding over finite rings $\mathfrak{R}_1, \mathfrak{R}_2, \cdots, \mathfrak{R}_s$.

By time sharing,

$$\mathcal{R}_{\mathfrak{R}_{\mathcal{S}}} = \mathrm{cov}\left(\bigcup_{\Phi \in \mathscr{M}\left(\mathscr{X}_{\mathcal{S}}, \mathfrak{R}_{\mathcal{S}}\right)} \left(\mathcal{R}_{\Phi} \cup \mathcal{R}'_{\Phi}\right) \right), \tag{28}$$

where $\mathcal{R}_{\Phi}$ and $\mathcal{R}'_{\Phi}$ are given by (3) and (27), respectively, is achievable with (left and right) LCoR.

As mentioned before, [6], [7] consider the computing problem, Problem 1, by treating a discrete function as a polynomial function over some finite field. Naturally, we believe that similar results can be obtained for polynomial functions over finite rings. In addition, both [9] and [6] considered the problem that under what circumstances $\mathcal{R}[g] \supsetneq \mathcal{R}[X_1, X_2, \cdots, X_s]$. However, a conclusive result is proved only for the case $s = 2$. It will be interesting to know whether the ring approach can provide further insight regarding this problem.

In order to focus on the basics of the new ideas, we did not consider the computing problem in the context of noisy channels (e.g., [14]) or network coding (e.g., [15]). However, it is clear that the polynomial (over fields or rings) approach can be applied to such scenarios as well. This will be considered in our further work.

As another suggestion for further work, it will be very interesting to consider instead linear coding over rngs. It will be even more intriguing should it turn out that the rng version outperforms the ring version in the computing problem, in the same manner that the ring version outperforms the field counterpart. It will also be interesting to see whether the idea of using rng provides more understanding of the problems from [9] and [6].

Regarding algebraic structure coding, some authors [16], [17], [18] propose to implement coding over a simpler structure, group. Seemingly, this is a more universal approach since a field or a ring must be a group. However, one subtle issue is often overlooked in this context. Strictly speaking, the set of rings (or rngs) is not a subset of the set of groups. Several non-isomorphic rings (or rngs) can be defined on one and the same group. For instance, given two distinct primes p and q, up to isomorphism,

1) there are 2 finite rngs of order p, while there is only one group of order p;
2) there are 4 finite rngs of order pq;
3) there are 11 finite rngs of order p^2 (if $p = 2$, 4 of them are rings, namely $\mathbb{F}_4$, $\mathbb{Z}_4$, $\mathbb{Z}_2 \times \mathbb{Z}_2$ and $\mathbb{M}_L$ [12]);
4) there are 22 finite rngs of order p^2q;
5) there are 52 finite rngs of order 8;
6) there are $3p + 50$ finite rngs of order p^3 ($p > 2$). (More can be found from [19].)

Therefore, there is no one-to-one correspondence from rings (or rngs) to groups, in either direction. Furthermore, from the point of view of formulating a *multivariate function*, group, associated with a single operator, is in a subordinate position compared to ring (rng or field). On the contrary, it is well-known that every discrete function is essentially a restriction of some polynomial function over some finite ring (rng or field). Although non-Abelian structures (non-Abelian groups) possess the potential to offer prominent results [20], [21], they are very difficult to handle theoretically and in practice. The performance of non-Abelian group block codes is sometimes bad [22].

APPENDIX A
SUPPORTING LEMMATA

Lemma A.1. *If both*

$$0 \leq \max\{p_2, p_3\} \not< \min\{p_1, p_4\} \leq 1 \text{ and } 0 \leq \max\{p_1, p_4\} \not< \min\{p_2, p_3\} \leq 1 \tag{A.1}$$

are valid, and $\sum_{j=1}^{4} p_j = 1$, *then*

$$\begin{aligned} -\sum_{j=1}^{4} p_j \log p_j \leq & -2\big[(p_2 + p_3) \log (p_2 + p_3) \\ & + (p_1 + p_4) \log (p_1 + p_4) \big]. \end{aligned} \tag{A.2}$$

Proof [23]: Without loss of generality, we assume that $0 \leq \max\{p_4, p_3\} \leq \min\{p_2, p_1\} \leq 1$ which implies that $p_1 + p_2 - 1/2 \geq |p_1 + p_4 - 1/2|$. Let $H_2(c) = -c \log c - (1-c)\log(1-c)$, $0 \leq c \leq 1$, be the binary entropy function. (A.2) equals to

$$\begin{aligned} & (p_1 + p_4) \left(\frac{p_1}{p_1 + p_4} \log \frac{p_1 + p_4}{p_1} + \frac{p_4}{p_1 + p_4} \log \frac{p_1 + p_4}{p_4} \right) \\ & + (p_2 + p_3) \left(\frac{p_2}{p_2 + p_3} \log \frac{p_2 + p_3}{p_2} + \frac{p_3}{p_2 + p_3} \log \frac{p_2 + p_3}{p_3} \right) \\ & \leq - (p_2 + p_3) \log (p_2 + p_3) - (p_1 + p_4) \log (p_1 + p_4) \\ & \quad \Leftrightarrow \\ A = & (p_1 + p_4) H_2 \left(\frac{p_1}{p_1 + p_4} \right) + (p_2 + p_3) H_2 \left(\frac{p_2}{p_2 + p_3} \right) \\ & \leq H_2(p_1 + p_4). \end{aligned}$$

Since H_2 is a concave function and $\sum_{j=1}^{4} p_j = 1$, then

$$A \leq H_2\left(p_1 + p_2\right).$$

Moreover, $p_1 + p_2 - 1/2 \geq |p_1 + p_4 - 1/2|$ guarantees that

$$H_2\left(p_1 + p_2\right) \leq H_2\left(p_1 + p_4\right),$$

because $H_2(c) = H_2(1-c), \forall\ 0 \leq c \leq 1$, and $H_2(c') \leq H_2(c'')$ if $0 \leq c' \leq c'' \leq 1/2$. Therefore, $A \leq H_2\left(p_1 + p_4\right)$ and (A.2) holds. ■

Lemma A.2. *No matter which finite field $\mathbb{F}_q$ is chosen, g given by* (21) *admits no presentation $k_1(x)+k_2(y)+k_3(z)$, where $k_i \in \mathbb{F}_q[1]$ for all feasible i.*

Proof: Suppose otherwise, i.e., $k_1 \circ \mu_1 + k_2 \circ \mu_2 + k_3 \circ \mu_3 = \nu \circ g$ for some injections $\mu_1, \mu_1, \mu_3 : \{\alpha_0, \alpha_1\} \to \mathbb{F}_q$ and $\nu : \{\beta_0, \cdots, \beta_3\} \to \mathbb{F}_q$. By (21), we have

$$\begin{aligned}
\nu(\beta_1) =& (k_1 \circ \mu_1)(\alpha_1) + (k_2 \circ \mu_2)(\alpha_0) + (k_3 \circ \mu_3)(\alpha_0) \\
=& (k_1 \circ \mu_1)(\alpha_0) + (k_2 \circ \mu_2)(\alpha_1) + (k_3 \circ \mu_3)(\alpha_1) \\
\nu(\beta_3) =& (k_1 \circ \mu_1)(\alpha_1) + (k_2 \circ \mu_2)(\alpha_1) + (k_3 \circ \mu_3)(\alpha_0) \\
=& (k_1 \circ \mu_1)(\alpha_0) + (k_2 \circ \mu_2)(\alpha_0) + (k_3 \circ \mu_3)(\alpha_1) \\
\Longrightarrow& \nu(\beta_1) - \nu(\beta_3) = \tau = -\tau \\
\Longrightarrow& \tau + \tau = 0,
\end{aligned} \tag{A.3}$$

where $\tau = k_2(\mu_2(\alpha_0)) - k_2(\mu_2(\alpha_1))$. Since μ_2 is injective, (A.3) implies that either $\tau = 0$ or Char$(\mathbb{F}_q) = 2$ by Proposition II.6. Noticeable that $k_2(\mu_2(\alpha_0)) \neq k_2(\mu_2(\alpha_1))$, i.e., $\tau \neq 0$, otherwise, $\nu(\beta_1) = \nu(\beta_3)$ which contradicts the assumption ν is injective. Thus, Char$(\mathbb{F}_q) = 2$. Let $\rho = (k_3 \circ \mu_3)(\alpha_0) - (k_3 \circ \mu_3)(\alpha_1)$. Obviously, $\rho \neq 0$ because of the same reason that $\tau \neq 0$, and $\rho + \rho = 0$ since Char$(\mathbb{F}_q) = 2$. Therefore,

$$\begin{aligned}
\nu(\beta_0) =& (k_1 \circ \mu_1)(\alpha_0) + (k_2 \circ \mu_2)(\alpha_0) + (k_3 \circ \mu_3)(\alpha_0) \\
=& (k_1 \circ \mu_1)(\alpha_0) + (k_2 \circ \mu_2)(\alpha_0) + (k_3 \circ \mu_3)(\alpha_1) + \rho \\
=& \nu(\beta_3) + \rho \\
=& (k_1 \circ \mu_1)(\alpha_1) + (k_2 \circ \mu_2)(\alpha_1) + (k_3 \circ \mu_3)(\alpha_0) + \rho \\
=& (k_1 \circ \mu_1)(\alpha_1) + (k_2 \circ \mu_2)(\alpha_1) + (k_3 \circ \mu_3)(\alpha_1) + \rho + \rho \\
=& \nu(\beta_2) + 0 = \nu(\beta_2).
\end{aligned}$$

This contradicts that ν is injective. ■

Remark 21. As a special case, this lemma implies that no matter which finite field $\mathbb{F}_q$ is chosen, g defined by (21) has no presentation that is linear over $\mathbb{F}_q$. In contrast, g is equivalent to linear function $x + 2y + 3z \in \mathbb{Z}_4[3]$.

ACKNOWLEDGMENT

The authors would like to thank their colleagues Jinfeng Du and Mattias Andersson for assistance in proving Lemma A.1.

REFERENCES

[1] P. Elias, "Coding for noisy channels," *IRE Conv. Rec.*, pp. 37–46, March 1955.

[2] I. Csiszár, "Linear codes for sources and source networks: Error exponents, universal coding," *IEEE Transactions on Information Theory*, vol. 28, no. 4, pp. 585–592, July 1982.

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

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

[5] R. Ahlswede and T. S. Han, "On source coding with side information via a multiple-access channel and related problems in multi-user information theory," *IEEE Transactions on Information Theory*, vol. 29, no. 3, pp. 396–411, May 1983.

[6] S. Huang and M. Skoglund, "Polynomials and computing functions of correlated sources," in *IEEE International Symposium on Information Theory*, July 2012.

[7] ——, "Computing polynomial functions of correlated sources: Inner bounds," in *International Symposium on Information Theory and its Applications*, October 2012.

[8] J. J. Rotman, *Advanced Modern Algebra*, 2nd ed. American Mathematical Society, August 2010.

[9] T. S. Han and K. Kobayashi, "A dichotomy of functions f(x, y) of correlated sources (x, y) from the viewpoint of the achievable rate region," *IEEE Transactions on Information Theory*, vol. 33, no. 1, pp. 69–76, Jan. 1987.

[10] D. S. Dummit and R. M. Foote, *Abstract Algebra*, 3rd ed. Wiley, 2003.

[11] R. W. Yeung, *Information Theory and Network Coding*, 1st ed. Springer Publishing Company, Incorporated, Sep. 2008.

[12] D. Singmaster and D. M. Bloom, "Rings of order four," *Mathematical Association of America*, vol. 71, no. 8, pp. 918–920, October 1964.

[13] S. Huang and M. Skoglund, "Linear source coding over rings and applications," in *IEEE Swedish Communication Technologies Workshop*, October 2012.

[14] B. Nazer and M. Gastpar, "Computation over multiple-access channels," *IEEE Transactions on Information Theory*, vol. 53, no. 10, Oct. 2007.

[15] R. Appuswamy, M. Franceschetti, N. Karamchandani, and K. Zeger, "Network coding for computing: Cut-set bounds," *IEEE Transactions on Information Theory*, vol. 57, no. 2, Feb. 2011.

[16] D. Slepian, "Group codes for the gaussian channel," *The Bell System Technical Journal*, 1968.

[17] R. Ahlswede, "Group codes do not achieve shannon's channel capacity for general discrete channels," *The Annals of Mathematical Statistics*, vol. 42, no. 1, pp. 224–240, February 1971.

[18] G. D. Forney, Jr., "On the hamming distance properties of group codes," *IEEE Transactions on Information Theory*, vol. 38, no. 6, pp. 1797–1801, November 1992.

[19] Wikimedia Foundation, Inc. (2012, July) Finite ring. [Online]. Available: http://en.wikipedia.org/wiki/Finite_ring

[20] T. H. Chan and A. Grant, "Entropy vector and network codes," in *IEEE International Symposium on Information Theory*, Jun. 2007.

[21] S. Huang, "Some properties of $\overline{\Gamma}_n^*$ and error control with group network codes," Master's thesis, The University of Hong Kong, Jun. 2011.

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

[23] J. Du and M. Andersson, Private Communication, May 2012.