Linear Source Coding over Rings and Applications

Sheng Huang, Mikael Skoglund
School of Electrical Engineering
KTH Royal Institute of Technology
Stockholm, Sweden
Email: sheng.huang@ee.kth.se, skoglund@ee.kth.se

***Abstract*—This paper studies linear coding (LC) techniques in the setting of computing functions of correlated memoryless sources. Instead of linear mappings over finite fields, we consider using linear mappings over finite rings as encoders. It is shown that generally the region $c\times\mathcal{R}$, where $c\geq 1$ is a constant and $\mathcal{R}$ is the Slepian–Wolf (SW) region, is achievable with LC over ring (LCoR) when the function to compute is the identity function. $c=1$ if the ring used is a field. Hence, LCoR could be suboptimal in terms of achieving the best coding rates (the SW region) for computing the identity function.**

In spite of that, the ring version shows several advantages. It is demonstrated that there exists a function that is neither linear nor can be linearized over any finite field. Thus, LC over field (LCoF) does not apply directly for computing such a function unless the polynomial approach [1], [2] is used. On the contrary, such a function is linear over some ring. Using LCoR, an achievable region containing the SW region can be obtained for computing this function. In addition, the alphabet sizes of the encoders are strictly smaller than using LCoF. More interestingly, LCoF is not useful if some special requirement is imposed.

I. Introduction

This paper focuses on applying linear coding (LC) techniques to the coding problem of computing a *discrete function* over a *noiseless memoryless source network*. This computing problem is formally defined as follows.

Problem 1 (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* $g:\prod_{i\in\mathcal{S}}\mathscr{X}_i\rightarrow\Omega$, what is the biggest region $\mathcal{R}[g]\subset\mathbb{R}^s$, such that, $\forall\ (R_1,R_2,\cdots,R_s)\in\mathcal{R}[g]$ and $\forall\ \epsilon>0$, $\exists\ N_0>0$ such that $\forall\ n>N_0$, there exist s *encoders* $\phi_i:\mathscr{X}_i^n\rightarrow\left[1,2^{nR_i}\right],i\in\mathcal{S}$, and one *decoder* $\psi:\prod_{i\in\mathcal{S}}\left[1,2^{nR_i}\right]\rightarrow\Omega^n$, such that

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

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

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

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

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 such a computing problem is equivalent to the *Slepian–Wolf* (SW) *source coding* problem. Hence, $\mathcal{R}[g]$ is the *SW region* [3]

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

where T^c is the *complement* of T in $\mathcal{S}$ and $X_T\left(X_{T^c}\right)$ is the random variable array $\prod_{j\in T}X_j\left(\prod_{j\in T^c}X_j\right)$. Nevertheless, based on the scheme used to achieve this region, the structures of the encoders are unclear, since they are chosen randomly among all feasible mappings. In the single source scenario, Elias [4] shows 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 generalised to the multiple sources scenario by Csiszár [5]. As a consequence of [5], 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 [4], Körner–Marton [6] shows that $\mathcal{R}[\oplus_2]$ ("$\oplus_2$" is the *modulo-two sum*) contains the region

$$\mathcal{R}_{\oplus_2}=\left\{(R_1,R_2)\in\mathbb{R}^2\mid R_1,R_2\geq H(X_1\oplus_2X_2)\right\}.$$

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, [7] 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, functions considered in these literature are relatively simple (the modulo-two sum).

Taking a *polynomial* approach, [1], [2] generalise the result of Ahlswede–Han [7, Theorem 10] to the most general scenario. Making use of the fact that a discrete function is essentially a *polynomial function* [8, pp. 129] over some finite field, an achievable region is given for computing an arbitrary discrete function. Such a region contains and can be strictly bigger (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 [1] from different perspectives, respectively.

This paper focuses on *linear coding over finite ring* (LCoR) which serves as an alternative technique for the computing problem. In Section III, we present an achievable region for computing the identity function with LC over a finite *ring*, namely, all encoders are *linear mappings* over this ring (see Definition II.4). It is proved that this region resumes the SW region if this ring is a field. Thus, the results of [4], [5] become special cases of ours on this respect. As an application, Problem 1 with a non-trivial g is considered in Section IV. After giving two generalised results of [7, Theorem 10] and Körner–Marton [6], respectively, an example is constructed exhibiting several superiorities of LCoR compared to LCoF. In this example, a discrete function is defined. Since this function is neither linear nor linearizable over any finite field, the methods of [7, Theorem 10] or Körner–Martion [6] do not apply directly. However, it is linear over some finite ring. LCoR can then be used to achieve coding rates beyond the SW region for computing this function. Though LCoF can also be applied to achieve coding rates beyond the SW region if the polynomial approach [1], [2] is used, the encoders using LCoF require strictly bigger alphabet sizes than using LCoR. Under special circumstance, LCoF is not even useful (Remark 11).

II. Preliminaries

Some needed algebraic concepts and results are given in this section. More fundamentals can be found in [8], [10]. Readers who are familiar with abstract algebra can go through quickly.

Definition II.1. The tuple $[\mathfrak{R}, +, \cdot]$ is called a *ring* if and only if it satisfies the follows:

1) $[\mathfrak{R}, +]$ is a *group*;
2) $\forall\ a, b, c \in \mathfrak{R}$, $a \cdot b \in \mathfrak{R}$ and $(a \cdot b) \cdot c = a \cdot (b \cdot c)$. Moreover, there exists a *multiplicative unit*[1] $1 \in \mathfrak{R}$, i.e., $1 \cdot a = a \cdot 1 = a$, $\forall\ a \in \mathfrak{R}$;
3) $\forall\ a, b, c \in \mathfrak{R}$, $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$ 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 $[\mathfrak{R}, +]$, namely 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 $\mathfrak{R}[s]$.

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 "$\cdot$". All fields are commutative 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, 14.3].

Definition II.2 (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 unit of $\mathfrak{R}$, respectively. The characteristic of $\mathfrak{R}$ is often denoted by $\mathrm{Char}(\mathfrak{R})$.

Remark 1. Clearly, $\mathrm{Char}(\mathbb{Z}_q) = q$. For a finite field $\mathbb{F}_q$, $\mathrm{Char}(\mathbb{F}_q)$ is always the prime q_0 such that $q = q_0^n$ for some integer n [8, Proposition 3.113].

Proposition II.3. *Let $\mathbb{F}_q$ be a finite field. For any $0 \neq a \in \mathbb{F}_q$, $m = \mathrm{Char}(\mathbb{F}_q)$ 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.4. A mapping $f : \mathfrak{R}^n \to \mathfrak{R}^m$ given as:

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

where $a_{i,j} \in \mathfrak{R}$, $1 \leq i \leq m$, $1 \leq j \leq n$, is called a *linear mapping* over ring $\mathfrak{R}$. If $m = 1$, then f is called a *linear function* over $\mathfrak{R}$.

Remark 2. The mapping f in Definition II.4 is called *linear* in accordance with the definition of *linear mapping* (*linear function*) over a field. In fact, they do possess a lot of similar properties. Besides, (1) is equivalent to

$$f(x_1, x_2, \cdots, x_n) = \mathbf{A} (x_1, x_2, \cdots, x_n)^T, \quad \forall\ (x_1, x_2, \cdots, x_n) \in \mathfrak{R}^n, \tag{2}$$

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. 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. $\mathbf{A}$ is named the *coefficient matrix*.

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 $\mathfrak{S}(f)$

[1] In some literature, a ring is defined without multiplicative unit. We only consider rings with multiplicative units in this paper.

is a subspace of $\mathfrak{R}^n$. In fact, $\mathfrak{S}(f)$ is an $\mathfrak{R}$-*submodule* [10] of $\mathfrak{R}^n$ in general. We conclude this section with a lemma regarding the cardinalities of $\mathfrak{R}^n$ and its $\mathfrak{R}$-submodule $\mathfrak{S}(f)$ in the following.

Lemma II.5. *A linear function $f : \mathfrak{R}^n \to \mathfrak{R}$ over a ring $\mathfrak{R}$ is non-trivial, if and only if the set $\mathfrak{S}(f)$ is proper, i.e, $\mathfrak{S}(f) \subsetneq \mathfrak{R}^n$.*

Proof: Let $f(\mathbf{x}) = \mathbf{A}\mathbf{x}$, $\forall\ \mathbf{x} \in \mathfrak{R}^n$, where $\mathbf{A} = (a_1, a_2, \cdots, a_n)$. For necessity $\Rightarrow$, without loss of generality, assume that $a_1 \neq 0$, since f is non-trivial. Show by contradiction, suppose that $\mathfrak{S}(f) = \mathfrak{R}^n$. Then let $\mathbf{x}_0 = (x_1, x_2, \cdots, x_n) = (1, 0, \cdots, 0) \in \mathfrak{S}(f)$. We have $0 = f(\mathbf{x}_0) = \sum_{i=1}^{n} a_i x_i = a_1$, a contradiction. Sufficiency ($\Leftarrow$) is obvious. ∎

III. Random Linear Coding over Rings

This section considers Problem 1 with g being the identity function, namely the SW source coding problem. The case that g is an arbitrary function is left to Section IV.

Given a ring $\mathfrak{R}$, it is easily seen from Remark 2 that a linear mapping $f : \mathfrak{R}^n \to \mathfrak{R}^m$ is uniquely determined by the coefficient matrix $\mathbf{A}$, and vice versa. Thus, we can define $\mathfrak{S}(\mathbf{A})$ to be $\mathfrak{S}(f)$. In particular, $\mathbf{A}$ is a row vector when $m = 1$. Let $\mathscr{L}_{\mathfrak{R}^n}$ be the set of all linear functions $f : \mathfrak{R}^n \to \mathfrak{R}$. Then $\mathscr{L}_{\mathfrak{R}^n}$ can be seen as an $\mathfrak{R}$-module with respect to the *function addition* and *scalar-function multiplication*. Consequently, every $\mathbf{x} \in \mathfrak{R}^n$ can be treated as a linear function $\mathbf{x} : \mathscr{L}_{\mathfrak{R}^n} \to \mathfrak{R}$ given by $\mathbf{x} : f \mapsto f(\mathbf{x})$, $\forall\ f \in \mathscr{L}_{\mathfrak{R}^n}$. This is the so-called *duality*.

Let $\kappa(\mathfrak{R}, n) = \max\limits_{f \in \mathscr{L}_{\mathfrak{R}^n} \setminus \{0\}} \dfrac{|\mathfrak{S}(f)|}{|\mathfrak{R}^n|} = \max\limits_{\mathbf{x} \in \mathfrak{R}^n \setminus \{\mathbf{0}\}} \dfrac{|\mathfrak{S}(\mathbf{x})|}{|\mathfrak{R}^n|}$, and $\kappa_{\mathfrak{R}} = \liminf\limits_{n \to \infty} \kappa(\mathfrak{R}, n)$.

Lemma III.1. *Let $\mathfrak{R}$ be a commutative finite ring. For any $f \in \mathscr{L}_{\mathfrak{R}^n} \setminus \{0\}$, we have $\dfrac{|\mathfrak{S}(f)|}{|\mathfrak{R}^n|} \leq \kappa(\mathfrak{R}, n) \leq \dfrac{1}{2}$. Furthermore, if $\mathfrak{R}$ is a field, then $\dfrac{|\mathfrak{S}(f)|}{|\mathfrak{R}^n|} = \dfrac{1}{|\mathfrak{R}|} = \kappa(\mathfrak{R}, n) = \kappa_{\mathfrak{R}}$.*

Proof: $\forall\ \mathbf{x}, \mathbf{y} \in \mathfrak{S}(f)$ and $\forall\ a, b \in \mathfrak{R}$, $f(a\mathbf{x} + b\mathbf{y}) = \mathbf{A}(a\mathbf{x} + b\mathbf{y}) = a(\mathbf{A}\mathbf{x}) + b(\mathbf{A}\mathbf{y}) = 0$ implies that $a\mathbf{x} + b\mathbf{y} \in \mathfrak{S}(f)$. Therefore, $\mathfrak{S}(f)$ is a $\mathfrak{R}$-*submodule* of $\mathfrak{R}^n$, hence, $\mathfrak{S}(f)$ is a *subgroup* $\mathfrak{R}^n$. By Lagrange's Theorem [11, 1.3.2], there exists a positive integer l, such that $l \times |\mathfrak{S}(f)| = |\mathfrak{R}^n|$. On the other hand, by Lemma II.5, $|\mathfrak{S}(f)| < |\mathfrak{R}^n|$ since $\mathfrak{R}$ is finite. Thus, $l \geq 2$. Therefore, $\dfrac{|\mathfrak{S}(f)|}{|\mathfrak{R}^n|} = \dfrac{1}{l} \leq \kappa(\mathfrak{R}, n) \leq \dfrac{1}{2}$.

Observe that $\mathfrak{R}^n$ is a vector space and $\mathfrak{S}(f)$ is one of its subspaces of *codimension* 1. Then the second half of the statement follows. ∎

Remark 3. The conclusion drawn in Lemma III.1 is of great importance. The achievability result Theorem III.3 deeply depends on the upper bound of $\kappa(\mathfrak{R}, n)$. Besides, 2^{-1} is a very loose upper bound on $\kappa(\mathfrak{R}, n)$ and $\kappa_{\mathfrak{R}}$ in general (e.g., $\kappa_{\mathbb{Z}_{35}} \leq 5^{-1}$).

Lemma III.2. *Given a commutative finite ring $\mathfrak{R}$, for the linear mapping $f : \mathfrak{R}^n \to \mathfrak{R}^k$ randomly chosen according to a uniform distribution, $\Pr\{f(\mathbf{x}) = \mathbf{0}\} = \left(\dfrac{|\mathfrak{S}(\mathbf{x})|}{|\mathfrak{R}^n|}\right)^k \leq [\kappa(\mathfrak{R}, n)]^k \leq 2^{-k}, \forall\ \mathbf{0} \neq \mathbf{x} \in \mathfrak{R}^n$. Moreover, if $\mathfrak{R}$ is a field, then $\Pr\{f(\mathbf{x}) = \mathbf{0}\} = |\mathfrak{R}|^{-k}, \forall\ \mathbf{0} \neq \mathbf{x} \in \mathfrak{R}^n$.*

Proof: Let $f = (f_1, f_2, \cdots, f_k)^T$, where $f_i : \mathfrak{R}^n \to \mathfrak{R}$ is a linear function over $\mathfrak{R}$ for all feasible i. By a simple counting argument, it is easy to verify that

$$\Pr\{f(\mathbf{x}) = \mathbf{0}\} = \Pr\left\{\bigcap_{1 \leq i \leq k} \{f_i(\mathbf{x}) = 0\}\right\} = \left(\frac{|\mathfrak{S}(\mathbf{x})|}{|\mathfrak{R}^n|}\right)^k$$

The lemma follows from Lemma III.1 since $\mathbf{x} \neq \mathbf{0}$. ∎

Remark 4. Consider using f in Lemma III.2 as the encoder, and assume that $\mathbf{x}_0$ is the original data generated by the source. The probability that the decoder mistakes $\mathbf{x}' \neq \mathbf{x}_0$ as the original data, namely $\Pr\{f(\mathbf{x}_0 - \mathbf{x}') = \mathbf{0}\}$, is then bounded by $[\kappa(\mathfrak{R}, n)]^k \leq 2^{-k}$. Thus, if k is big enough, then the total probability of decoding error for decoding all feasible $\mathbf{x}_0$ is very small. For such a reason, we name the *lower limit* $\kappa_{\mathfrak{R}}$ the *fundamental error factor*, because the error exponent of a random linear code over $\mathfrak{R}$ is closely related to it (see Theorem III.3 for more details).

Theorem III.3. *In Problem 1, if g is the identity function, then $\left(-\dfrac{\log|\mathfrak{R}|}{\log \kappa_{\mathfrak{R}}}\right)\mathcal{R}$, where $\mathcal{R} = \mathcal{R}[X_1, X_2, \cdots, X_s]$, is achievable with linear coding over commutative finite ring $\mathfrak{R}$ of order $|\mathfrak{R}| \geq \max\limits_{1 \leq i \leq s} |\mathscr{X}_i|$. Furthermore, $\mathcal{R}$ is achievable with linear coding over $\mathfrak{R}$ if $\mathfrak{R}$ is a field.*

Sketch of the Proof: Since $|\mathfrak{R}| \geq \max\limits_{1 \leq i \leq s} |\mathscr{X}_i|$, $\mathscr{X}_i$ can be seen as a subset of $\mathfrak{R}$ for each $1 \leq i \leq s$. If $\mathfrak{R}$ is a finite field, then $\left(-\dfrac{\log|\mathfrak{R}|}{\log \kappa_{\mathfrak{R}}}\right)\mathcal{R} = \mathcal{R}$. Hence, it suffices to prove that $\left(-\dfrac{\log|\mathfrak{R}|}{\log \kappa_{\mathfrak{R}}}\right)\mathcal{R}$ is achievable. Since the lower limit $\liminf\limits_{m \to \infty} \kappa(\mathfrak{R}, m)$ always exists, we can define $\mathcal{Q}_{\mathfrak{R}} = \{n_{l_1}, n_{l_2}, \cdots, n_{l_j}, \cdots\} \subseteq \mathbb{N}^+$, such that $\lim\limits_{j \to \infty} \kappa(\mathfrak{R}, n_{l_j}) = \kappa_{\mathfrak{R}}$. $\forall\ (R_1, R_2, \cdots, R_s) \in \mathcal{R}$, let $k_i = \dfrac{nR_i}{-\log \kappa(\mathfrak{R}, n)}$, $1 \leq i \leq s$, where $n \in \mathcal{Q}_{\mathfrak{R}}$ is the length of the codewords. Construct the encoders and decoder as for LC over a finite field. In particular, choose randomly and independently for each source t_i a linear mapping $f_i : \mathfrak{R}^n \to \mathfrak{R}^{k_i}$ as its encoders. For all $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_s) \in \prod\limits_{i=1}^{s} \mathscr{X}_i^n$ and $\forall\ \emptyset \neq S \subseteq \mathcal{S}$, let $D(\mathbf{x}, S)$ be the set

$$\left\{ (\mathbf{x}_1', \mathbf{x}_2', \cdots, \mathbf{x}_s') \in \prod_{i=1}^{s} \mathscr{X}_i^n \,\middle|\, \mathbf{x}_i' \neq \mathbf{x}_i \text{ if and only if } i \in S \right\}.$$

By Lemma III.2, for any two sequences $\mathbf{x} = (\mathbf{x}_1, \cdots, \mathbf{x}_s) \in \prod\limits_{i=1}^{s} \mathscr{X}_i^n$ and $(\mathbf{x}_1', \mathbf{x}_2', \cdots, \mathbf{x}_s') \in D(\mathbf{x}, S)$, the probability that

$[f_1(\mathbf{x}'_1), f_2(\mathbf{x}'_2), \cdots, f_s(\mathbf{x}'_s)] = [f_1(\mathbf{x}_1), f_2(\mathbf{x}_2), \cdots, f_s(\mathbf{x}_s)]$, i.e., $[f_1(\mathbf{x}'_1 - \mathbf{x}_1), f_2(\mathbf{x}'_2 - \mathbf{x}_2), \cdots, f_s(\mathbf{x}'_s - \mathbf{x}_s)]^T = \mathbf{0}$, is

$$\begin{aligned}\prod_{i\in S}\left(\frac{|\mathfrak{S}(\mathbf{x}'_i - \mathbf{x}_i)|}{|\mathfrak{R}^n|}\right)^{k_i} \leq & [\kappa(\mathfrak{R}, n)]^{\sum_{i\in S} k_i} \\ = & 2^{-\sum_{i\in S} -k_i \log \kappa(\mathfrak{R}, n)} \\ = & 2^{-n\sum_{i\in S} R_i}.\end{aligned}$$

Thus, the probability of decoding error of decoding all *strongly typical sequences* [12] is bounded around 0 when n is big enough. Meanwhile, the rate of this code is $(R'_1, R'_2, \cdots, R'_s)$, where $R'_i = \frac{k_i}{n}\log|\mathfrak{R}| = -\frac{R_i \log|\mathfrak{R}|}{\log \kappa(\mathfrak{R}, n)}$ for all feasible i. Since $\kappa_{\mathfrak{R}} = \liminf_{m\to\infty} \kappa(\mathfrak{R}, m)$, the theorem follows. ∎

Remark 5. Let $q = \max_{1\leq i\leq s} |\mathscr{X}_i|$ and p_1 be the smallest prime that divides q. In the above theorem, $\mathfrak{R}$ can be set to $\mathbb{Z}_q$. Thus, $\left(\frac{\log q}{\log p_1}\right)\mathcal{R}$ is achievable with LC over $\mathbb{Z}_q$, since $\kappa_{\mathbb{Z}_q} \leq p_1^{-1}$. If $-\frac{\log q}{\log \kappa_{\mathbb{Z}_q}} > 1$, then $\left(-\frac{\log q}{\log \kappa_{\mathbb{Z}_q}}\right)\mathcal{R}$ is strictly smaller than $\mathcal{R}$. However, this can be improved if a tighter bound on $\kappa_{\mathfrak{R}}$ is obtained. For the special case where $\mathfrak{R}$ is a field, making use of the tightest bound, we have the whole region $\mathcal{R}$ achievable with LC as it is stated in Theorem III.3 ([4] as well).

IV. Application: Source Coding for Computing

Problem 1 with an arbitrary g is considered in this section. A source coding theorem of LCoR for computing an arbitrary function will be presented. Meanwhile, an example is given to demonstrate some advantages of LCoR compared to LCoF.

Theorem IV.1. *Let $\mathfrak{R}$ be a commutative finite ring. In Problem 1, if $g \in \mathfrak{R}[s]$ admits the structure*

$$g(x_1, \cdots, x_s) = h\left[k_0(x_1, \cdots, x_t) + \sum_{j=t+1}^{s} k_j(x_j)\right], \quad (3)$$

where $k_0 \in \mathfrak{R}[t]$ and $h, k_j \in \mathfrak{R}[1]$ for all $1 \leq t < j \leq s$, then

$$\begin{aligned}\mathcal{R}[g] \supseteq \mathcal{R} = \Big\{&(R_1, R_2, \cdots, R_s) \in \mathbb{R}^s \Big| \forall\, \emptyset \neq S \subseteq \mathcal{S}, \\ &\sum_{j\in S} R_j > I(Y_S; V_S|V_{S^c}) - \frac{|S\setminus T|\log|\mathfrak{R}|}{\log \kappa_{\mathfrak{R}}} H(Z|V_{\mathcal{S}})\Big\},\end{aligned}$$

where $T = \{1, 2, \cdots, t\}$; $\forall\, j \in T$, $V_j = Y_j = X_j$; $\forall\, t < j \leq s$, $Y_j = k_j(X_j)$, V_j's are discrete random variables such that

$$p(y_1, v_1, \cdots, y_s, v_s) = p(y_1, \cdots, y_s)\prod_{j=t+1}^{s} p(v_j|y_j);$$

and $Z = k_{t+1}(X_{t+1}) + k_{t+2}(X_{t+2}) + \cdots + k_s(X_s)$.

Remark 6. Obviously, if $V_j = Y_j$, $\forall\, j \in \mathcal{S}$, in Theorem IV.1, then $\mathcal{R} = \mathcal{R}[X_1, X_2, \cdots, X_s]$, the SW region. Thus, $\mathcal{R}$ contains the SW region if V_j's are chosen properly. Theorem IV.1 can be identified as a ring version of [2, Theorem III.1] (a generalisation of [7, Theorem 10]). They can be proved with parallel techniques. The proof of the above theorem is omitted because of space limitation. However, we provide the proof of its corollary (Corollary IV.4) showing the mechanism of how LCoR works in the computing problem.

Definition IV.2. 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 IV.3. 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 7. 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 $F_1(x, y) = xy \in \mathbb{Z}_2[2]$ or be treated as the restriction of $F_2(x, y) = x+y-(x+y)^2 \in \mathbb{Z}_3[2]$ to the domain $\{0,1\}\times\{0,1\} \subset \mathbb{Z}_3^2$. We refer to each presented format of a function as a *presentation* of this function.

The following is a corollary (special case) of Theorem IV.1. It can also be seen as a generalisation of Körner–Martion [6].

Corollary IV.4. *Given a discrete function $g : \prod_{i=1}^{s} \mathscr{X}_i \to \Omega$, and let $q = \max\left\{|\mathscr{X}_1|, |\mathscr{X}_2|, \cdots, |\mathscr{X}_s|, \left|g\left(\prod_{i=1}^{s}\mathscr{X}_i\right)\right|\right\}$, $\mathfrak{R}$ be a communicative finite ring of order $|\mathfrak{R}| \geq q$. If g is equivalent to some restriction of a polynomial function*

$$\hat{g} = h \circ k, \text{ where } k(x_1, x_2, \cdots, x_s) = \sum_{i=1}^{s} k_i(x_i), \quad (4)$$

and $h, k_i \in \mathfrak{R}[1]$ for all feasible i, then $(r, r, \cdots, r) \in \mathcal{R}[g]$, for all $r > -\frac{\log|\mathfrak{R}|}{\log \kappa_{\mathfrak{R}}} H(Z)$, where $Z = k(X_1, X_2, \cdots, X_s)$.

Proof: Let $\mathfrak{R}$ be a finite commutative ring, such that $|\mathfrak{R}| \geq q$. Up to equivalence, $\mathscr{X}_i$ can be seen as a subset of $\mathfrak{R}$ for $1 \leq i \leq s$. By Theorem III.3, $\forall\, \epsilon > 0$, there exists a big enough n, an $m \times n$ matrix $\mathbf{A} \in \mathfrak{R}^{m\times n}$ and a decoder ψ, such that $\Pr\{Z^n \neq \psi(\mathbf{A}Z^n)\} < \epsilon$, if $m > -\frac{nH(Z)}{\log \kappa_{\mathfrak{R}}}$. 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 $\vec{h}\left(\hat{Z}^n\right)$, where $\hat{Z}^n = \psi\left[\sum_{i=1}^{s}\phi_i\left(X_i^n\right)\right]$, is the function required to compute. The probability of decoding error is

$$\begin{aligned}&\Pr\left\{\vec{h}\left[\vec{k}\left(X_1^n, X_2^n, \cdots, X_s^n\right)\right] \neq \vec{h}\left(\hat{Z}^n\right)\right\} \\ \leq &\Pr\left\{Z^n \neq \hat{Z}^n\right\}\end{aligned}$$

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

Therefore, all $(r, r, \cdots, r) \in \mathbb{R}^s$, where $r = \dfrac{m \log |\mathfrak{R}|}{n} > -\dfrac{\log |\mathfrak{R}|}{\log \kappa_{\mathfrak{R}}} H(Z)$ are achievable, i.e., $(r, r, \cdots, r) \in \mathcal{R}[g]$. ∎

Remark 8. By Theorem III.3, LCoR may not be as good as LCoF in terms of achieving optimal coding rates in the SW source coding problem. Nevertheless, LCoR can also be applied to obtain achievable region larger than the SW region in the computing problem. It is easy to find such a distribution $(X_1, X_2, \cdots, X_s) \sim p$, such that, in Corollary IV.4, $sr < H(X_1, X_2, \cdots, X_s)$. Hence $(r, r, \cdots, r) \notin \mathcal{R}[X_1, X_2, \cdots, X_s]$, i.e., Corollary IV.4 offers coding rates beyond the SW region.

Remark 9. [1, Lemma A.2] implies that every discrete function is equivalent to some restriction of a polynomial function of format (4) ((3) as well) over some finite field (or ring). Hence, Theorem IV.1 and Corollary IV.4 are universally applied for computing any discrete function.

In the rest of this section, a discrete function is constructed. This function is equivalent to a linear function over some finite ring. It is proved that this function is neither linear nor can be linearized over any finite field (see Lemma IV.7). Consequently, the results of Körner–Martion [6] or Ahlswede–Han [7, Theorem 10] do not apply directly. However, it will be shown that LCoR offers a promising solution.

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

$$
\begin{array}{ll}
g : (\alpha_0, \alpha_0, \alpha_0) \mapsto \beta_0; & g : (\alpha_0, \alpha_0, \alpha_1) \mapsto \beta_3; \\
g : (\alpha_0, \alpha_1, \alpha_0) \mapsto \beta_2; & g : (\alpha_0, \alpha_1, \alpha_1) \mapsto \beta_1; \\
g : (\alpha_1, \alpha_0, \alpha_0) \mapsto \beta_1; & g : (\alpha_1, \alpha_0, \alpha_1) \mapsto \beta_0; \\
g : (\alpha_1, \alpha_1, \alpha_0) \mapsto \beta_3; & g : (\alpha_1, \alpha_1, \alpha_1) \mapsto \beta_2.
\end{array} \tag{5}
$$

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

$$
\begin{array}{l}
\mu : \alpha_j \mapsto j, \quad \forall\, j \in \{0, 1\}, \text{ and} \\
\nu : \beta_j \mapsto j, \quad \forall\, j \in \{0, 1, 2, 3\},
\end{array} \tag{6}
$$

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 IV.6, there exists no $\hat{g} \in \mathbb{F}_4[3]$ of format (4) so that g is equivalent to any restriction of $\hat{g}$. Although, by [1, 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 (4), the size q must be strictly bigger than 4. For instance, let

$$
\hat{h}(x) = \sum_{a \in \mathbb{Z}_5} a\left[1 - (x - a)^4\right] - \left[1 - (x - 4)^4\right] \in \mathbb{Z}_5[1]. \tag{7}
$$

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 (6).

Proposition IV.6. *There exists no polynomial function $\hat{g} \in \mathbb{F}_4[3]$ of format* (4)*, such that a restriction of $\hat{g}$ is equivalent to the function g defined by* (5)*.*

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 IV.7. ∎

Lemma IV.7. *No matter which finite field $\mathbb{F}_q$ is chosen, g given by* (5) *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 (5), we have

$$
\begin{aligned}
\nu(\beta_1) =& (k_1 \circ \mu_1)(\alpha_1) + (k_2 \circ \mu_2)(\alpha_0) + (k_3 \circ \mu_3)(\alpha_0) \\
=& (k_1 \circ \mu_1)(\alpha_0) + (k_2 \circ \mu_2)(\alpha_1) + (k_3 \circ \mu_3)(\alpha_1) \text{ and} \\
\nu(\beta_3) =& (k_1 \circ \mu_1)(\alpha_1) + (k_2 \circ \mu_2)(\alpha_1) + (k_3 \circ \mu_3)(\alpha_0) \\
=& (k_1 \circ \mu_1)(\alpha_0) + (k_2 \circ \mu_2)(\alpha_0) + (k_3 \circ \mu_3)(\alpha_1) \\
\Longrightarrow& \nu(\beta_1) - \nu(\beta_3) = \tau = -\tau \\
\Longrightarrow& \tau + \tau = 0,
\end{aligned} \tag{8}
$$

where $\tau = k_2(\mu_2(\alpha_0)) - k_2(\mu_2(\alpha_1))$. Since μ_2 is injective, (8) implies that either $\tau = 0$ or $\mathrm{Char}(\mathbb{F}_q) = 2$ by Proposition II.3. 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, $\mathrm{Char}(\mathbb{F}_q) = 2$. 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 $\mathrm{Char}(\mathbb{F}_q) = 2$. 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 10. This lemma says that no matter which finite field $\mathbb{F}_q$ is chosen, g defined by (5) has no presentation that is linear

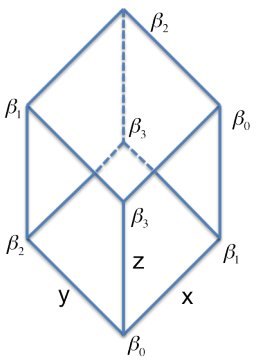

Fig 1: $g : \{\alpha_0, \alpha_1\}^3 \to \{\beta_0, \beta_1, \beta_2, \beta_3\}$

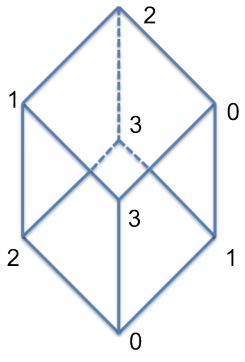

Fig 2: $x + 2y + 3z \in \mathbb{Z}_4[3]$

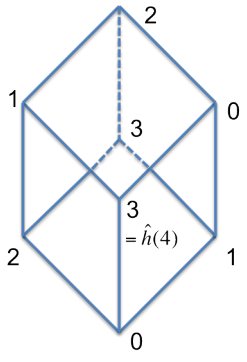

Fig 3: $\hat{h}(x + 2y + 4z) \in \mathbb{Z}_5[3]$

over $\mathbb{F}_q$. Thus, [6] or [7, Theorem 10] do not directly apply. In contrast, g is equivalent to linear function $x+2y+3z \in \mathbb{Z}_4[3]$. Theorem IV.1 and Corollary IV.4 apply if $\mathfrak{R}$ is set to be $\mathbb{Z}_4$. Although a solution of using LCoF is guaranteed by taking the polynomial approach [1], [2], the alphabet sizes required by the encoders are at least 5 strictly bigger then using LCoR.

Remark 11. As mentioned, using the polynomial approach, LCoF can be applied for computing g defined by (5). However, there is a critical drawback of the polynomial approach. To use this approach for computing a discrete function $\overline{g}$, one needs to search for a polynomial presentation $h \circ (k_0, k_{t+1}, \cdots, k_s) \in \mathfrak{R}[s]$, where $k_0 \in \mathfrak{R}[t]$ and $h, k_j \in \mathfrak{R}[1]$ $(t < j \leq s)$, of $\overline{g}$ over some ring (or field) $\mathfrak{R}$, then unveil the structure of h to the decoder. For instance, g admits presentation $\hat{h}(x + 2y + 4z) \in \mathbb{Z}_5[3]$, where $\hat{h}$ is given by (7). In order to use the polynomial approach in the setting of $\mathbb{Z}_5$, formula (7) must be known by the decoder. Unfortunately, there exist circumstances forbidding releasing (partial) information of $\overline{g}$. Lemma IV.7 implies that any polynomial presentation $h \circ (k_1, k_2, k_3)$ of g over any finite field admits a non-trivial h (h is not the superposition of k_1, k_2, k_3). Assume that the formula defining h can not be released to the decoder for some reason (security, communication impossible, etc). Then LCoF can not be applied for computing g with or without the polynomial approach. On the contrary, g is equivalent to $x + 2y + 3z \in \mathbb{Z}_4[3]$. Using LC over $\mathbb{Z}_4$, all the decoder needs to do is to add the messages observed from the encoders together then apply the decoding mapping. Neither the formula defining g, namely (5), nor the fact that g is equivalent to a linear function over $\mathbb{Z}_4$ is known by the decoder.

V. Conclusion

This paper considers LCoR and its application to the problem of coding for computing (see Problem 1). Source coding theorems of LCoR regarding computing the identity function or an arbitrary function are given. In addition, a non-linear (over any finite field) function is constructed. It is shown that methods of [6] and [7, Theorem 10], which depend on LCoF, do not directly applied for computing this function. Even though when the polynomial approach [1], [2] is used, LCoF can be applied for computing this function. The decoders require strictly bigger alphabet sizes compared to applying LCoR.

We mentioned that LCoR could be suboptimal in SW source coding scenario. We have not compared LCoF and LCoR with respect to their abilities of achieving larger achievable region for computing a non-identity function. However, we strongly believe that there exist non-field rings over which LC is optimal for the SW source coding scenario. Furthermore, there exists an example demonstrating that LCoR outperforms LCoF in achieving larger achievable region for computing some discrete function.

References

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

[2] S. Huang and M. Skoglund, "Computing polynomial functions of correlated sources: Inner bounds," in *International Symposium on Information Theory and its Applications*, October 2012.

[3] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Transactions on Information Theory*, vol. 19, pp. 471–480, July 1973.

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

[5] I. Csiszár, "Linear codes for sources and source networks: Error exponents, universal coding," *IEEE Transactions on Information Theory*, vol. 28, pp. 585–592, July 1982.

[6] J. Körner and K. Marton, "How to encode the modulo-two sum of binary sources," *IEEE Transactions on Information Theory*, vol. 25, pp. 219–221, Mar. 1979.

[7] 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, pp. 396–411, May 1983.

[8] J. J. Rotman, *Advanced Modern Algebra*. Prentice Hall, 2002.

[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, pp. 69–76, Jan. 1987.

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

[11] R. B. Ash, *Abstract Algebra: The Basic Graduate Year*. Robert B. Ash, 2000.

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