Polynomials and Computing Functions of Correlated Sources

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*—We consider the source coding problem of computing functions of correlated sources, which is an extension of the Slepian–Wolf coding problem. We observe that all the discrete functions are in fact restrictions of polynomial functions over some finite field. Based on this observation, we demonstrate how to use Elias' Lemma to enlarge the coding rate region (compared to the Slepian–Wolf region) for a certain class of polynomial functions.**

We present a classification result about polynomial functions regarding this coding problem. The result is conclusive in the two-sources scenario and, in fact, gives another interpretation of a result by Han and Kobayashi [1, Theorem 1].

I. Introduction

Let $(X_1, X_2, \cdots, X_s) \sim p$ be discrete random variables defined over sample space $\prod_{i=1}^{s} \mathcal{X}_i$. The original source coding problem for correlated sources considers the following scenario: given s sources $t_1, t_2, \cdots, t_s$ which generate correlated random data $X_1, X_2, \cdots, X_s$, what is the *coding rate region* $\mathcal{R}$, such that for any $(R_1, R_2, \cdots, R_s) \in \mathcal{R}$ and $\epsilon > 0$, there exists a big enough n, s *encoders* $\mathcal{E}_1, \mathcal{E}_2, \cdots, \mathcal{E}_s$, where $\mathcal{E}_i : \mathcal{X}_i^n \to \left[1, 2^{nR_i}\right]$, $\forall 1 \le i \le s$, and one *decoder* $\mathcal{D}$, where $\mathcal{D} : \prod_{i=1}^{s} \left[1, 2^{nR_i}\right] \to \prod_{i=1}^{s} \mathcal{X}_i^n$, such that $\Pr\left\{\prod_{i=1}^{s} \hat{X}_i^n \neq \prod_{i=1}^{s} X_i^n\right\} < \epsilon$, where X_i^n and $\prod_{i=1}^{s} \hat{X}_i^n$ are the input of $\mathcal{E}_i$ and the output of $\mathcal{D}$, respectively.

The classic Slepian–Wolf theorem [2] says that the coding rate region of the above problem is given by

$$\mathcal{R}[X_1, X_2, \cdots, X_s] = \Big\{(R_1, R_2, \cdots, R_s) \in \mathbb{R}^s \;\Big|\; \sum_{j \in J} R_j \ge H(X_J | X_{J^c}), \forall\, \emptyset \neq J \subseteq \{1, 2, \cdots, s\}\Big\}.$$

We consider a similar source coding problem in which a *discrete function* of the output data is to be recovered, instead of the original data. To be more precise, let $f : \prod_{i=1}^{s} \mathcal{X}_i \to \Omega$ be the function considered. Then what is the *coding rate region* $\mathcal{R}$, such that for any $(R_1, R_2, \cdots, R_s) \in \mathcal{R}$ and $\epsilon > 0$, there exists a big enough n, s encoders $\mathcal{E}_1, \mathcal{E}_2, \cdots, \mathcal{E}_s$, where $\mathcal{E}_i : \mathcal{X}_i^n \to \left[1, 2^{nR_i}\right]$, $\forall 1 \le i \le s$, and one decoder $\mathcal{D}$, where $\mathcal{D} : \prod_{i=1}^{s} \left[1, 2^{nR_i}\right] \to \Omega^n$, such that $\Pr\left\{\omega \neq \vec{f}\left(\prod_{i=1}^{s} X_i^n\right)\right\} < \epsilon$, where X_i^n and ω are the input of $\mathcal{E}_i$ and the output of $\mathcal{D}$, respectively, and $\vec{f}\left(\prod_{i=1}^{s} X_i^n\right) = \left\{f\left(\prod_{i=1}^{s} X_i^{(j)}\right)\right\}_{j=1}^{n}$.

Let $\mathcal{R}[f]$ be the coding rate region for computing function f. It is expected that $\mathcal{R}[f] = \mathcal{R}[X_1, X_2, \cdots, X_s]$ if f is the identity function, while in general it is easily seen that $\mathcal{R}[X_1, X_2, \cdots, X_s] \subseteq \mathcal{R}[f]$. Based on the method of Elias [3] (cf. [4]), Körner and Marton [5] show that if f is the modulo-two sum $\oplus$, $\mathcal{R}[\oplus]$ contains the convex hull of the union of $\mathcal{R}[X_1, X_2]$ and $\mathcal{R}_\oplus$, where

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

Subsequently, by combining the standard source coding technique and Elias' Lemma, Ahlswede and Han [6] gave an inner bound of $\mathcal{R}[\oplus]$ which is larger than Körner and Marton's in general.

Other variations of the general function computing problem have been studied in existing literature. As a special case, the original source coding problem is the first example with the identity function to be computed. In [7], the scenario where one of the two sources is intactly known by the decoder is considered. The idea of *characteristic graph* [8] is used in random code construction. On the other hand, the corresponding channel coding problems in which the receiver(s) is (are) interested in reproducing a function (distinct functions) of the output data of the source(s) from the channel or network output(s) are considered in [9], [10], [11]. The original MAC channel coding problem where the function to be reproduced is the identity function is a special case of this class of problems. A certain type of MAC channel is studied in [9]. [10], [11] consider the problem in the setting of network coding.

In this paper, we focus on the source coding problem of computing a discrete function of several correlated sources. The observation that all discrete functions are *restrictions* of *polynomial functions* serves as an important factor which works underneath our method. Basically, the structure of the function considered is what distinguishes this function computing problem from the Slepian–Wolf source coding problem. Hence, the algebraic structure of a function unveiled by its

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

978-1-4673-2579-0/12/$31.00 ©2012 IEEE

polynomial presentation is of great importance. Introducing this algebraic tool to the computing problem is one of the contributions of this paper. By treating discrete functions as polynomial functions, we point out that it is possible to enlarge the coding rate region (compared to the Slepian–Wolf region) for certain types of polynomial functions (see Theorem III.3, Theorem III.4 and Corollary III.5). In particular, inspired by the research of Han and Kobayashi [1], for the two-sources scenario, we show that the coding rate region for a given function is strictly bigger than the Slepian–Wolf region if and only if this function admits a "nice" polynomial presentation (see Theorem III.6 for more details).

A. Notation

Capital letters $X, Y, Z, \cdots$ are used to denote random variables and $\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \cdots$ correspond to their respective sample spaces. In addition, lower case letters $x, y, z, \cdots$ are used to denote instances of random variables. For a fixed n, $X^n, Y^n, Z^n, \cdots$ represent i.i.d sequences of length n, respectively. $X^{(i)}$ is for the ith term of X^n, $X^{[i)}$ is for the i.i.d sequence $X^{(i)}, X^{(i+1)}, \cdots, X^{(n)}$ and $X^{[i]}$ is for the i.i.d sequence $X^{(1)}, X^{(2)}, \cdots, X^{(i-1)}, X^{(i+1)}, X^{(i+2)}, \cdots, X^{(n)}$.

Let $X_1, X_2, \cdots, X_s$ be s correlated random variables. X_J is defined to be the array of random variables $X_{j_1}, X_{j_2}, \cdots, X_{j_k}$, where $\{j_1, j_2, \cdots, j_k\} = J \subseteq \{1, 2, \cdots, s\}$. The definitions of X_J^n, $X_J^{(i)}$, $X_J^{[i)}$ and $X_J^{[i]}$ resemble such a definition.

In this paper, we assume that, for $(X_1, X_2, \cdots, X_s) \sim p$, $p_{X_i}(x) > 0, \forall\ x \in \mathcal{X}_i$ and $\forall\ 1 \leq i \leq s$, where p_{X_i} is the marginal of p, because if $p_{X_i}(x) = 0$ for some $x \in \mathcal{X}_i$, then we can assume that X_i is defined on $\mathcal{X}_i \backslash \{x\}$ rather than $\mathcal{X}_i$. In addition, X_i is not a function of $X_1, \cdots, X_{i-1}, X_{i+1}, \cdots, X_s$ for all feasible i. This implies that $|\mathcal{X}_i| \geq 2, \forall\ 1 \leq i \leq s$.

The *support* of a distribution p is set to be $\mathrm{supp}(p) = \left\{ \mathbf{x} \in \prod_{i=1}^{s} \mathcal{X}_i \;\middle|\; p(\mathbf{x}) > 0 \right\}$. The support is crucial to the coding problem for computing, since only those values (of the function) defined over the support influence the problem.

II. Polynomial Functions

In this section, we demonstrate how a discrete function can be treated as a polynomial function. Readers who are familiar with algebra could skip this section.

Definition II.1. A *polynomial function*[1] of k variables over a finite field $\mathbb{F}$ is a function $g : \mathbb{F}^k \to \mathbb{F}$ of the form

$$g(x_1, x_2, \cdots, x_k) = \sum_{j=0}^{m} a_j x_1^{m_{1j}} x_2^{m_{2j}} \cdots x_k^{m_{kj}}, \qquad (1)$$

where $a_j \in \mathbb{F}$ and m and m_{ij}'s are non-negative integers.

From now on, $\mathbb{Z}_p$ denotes the field of integers modulo prime p, $\mathbb{F}_q$ is a finite field of order[2] q, and $\mathbb{F}[k]$ is defined to be the set of all the polynomial functions of k variables over the finite field $\mathbb{F}$.

[1] polynomial and polynomial function are distinct concepts (cf. [12]).

[2] the number of elements of a finite field.

Definition II.2. Given two functions $f : \prod_{i=1}^{k} \mathcal{X}_i \to \Omega_1$ and $g : \prod_{i=1}^{k} \mathcal{Y}_i \to \Omega_2$, f and g are said to be *equivalent* if and only if there exist bijections $\phi_i : \mathcal{X}_i \to \mathcal{Y}_i$, $\forall\ 1 \leq i \leq k$, and $\psi : \Omega_1 \to \Omega_2$, such that

$$f(x_1, x_2, \cdots, x_k) = \psi^{-1}(g(\phi_1(x_1), \phi_2(x_2), \cdots, \phi_n(x_k))).$$

Remark 1. The equivalency defined does not preserve all the mathematical properties of two equivalent functions. For instance, it does not preserve orders of the domain and the codomain. However, it does preserve all the mathematical properties that concern the encoders and the decoder. In other words, it can be easily proved that two equivalent functions share the "same" coding method, consequently, their coding rate regions are the same. From now on, we will simply refer to two equivalent functions as one function.

Definition II.3. Given function $f : \mathcal{D} \to \Omega$ and $\emptyset \neq S \subseteq \mathcal{D}$, the *restriction* of f on S is defined to be the function $f|_S : S \to \Omega$ such that $f|_S : x \mapsto f(x), \forall\ x \in S$.

Lemma II.4. *Any discrete function $f(x_1, x_2, \cdots, x_k)$ defined on domain $\mathcal{D} = \prod_{i=1}^{k} \mathcal{X}_i$ is equivalent to a restrictions of some polynomial function $h \in \mathbb{Z}_p[k]$, with $p \geq \max\{|f(\mathcal{D})|, |\mathcal{X}_i| \mid 1 \leq i \leq k\}$ being a prime.*

Remark 2. In the above lemma, $\mathbb{Z}_p$ can be replaced by $\mathbb{F}_{p^m}$, where $p^m \geq \max\{|f(\mathcal{D})|, |\mathcal{X}_i| \mid 1 \leq i \leq k\}$. Interested readers can refer to [12, Lemma 7.40] for details. For completeness and to help understanding, we provide a proof of Lemma II.4.

Proof of Lemma II.4: By (1) and Fermat's little theorem, it is easy to see that the number of polynomial functions in $\mathbb{Z}_p[k]$ is p^{p^k}. Moreover, the number of distinct functions with domain $\mathbb{Z}_p^k$ and codomain $\mathbb{Z}_p$ is also p^{p^k}. Hence, any function $g : \mathbb{Z}_p^k \to \mathbb{Z}_p$ is a polynomial function.

In the mean while, it is easy to find injections $\eta_i : \mathcal{X}_i \to \mathbb{Z}_p, \forall\ 1 \leq i \leq k$ and $\theta : f(\mathcal{D}) \to \mathbb{Z}_p$, which implies that f is equivalent to a restriction of some polynomial function $h : \mathbb{Z}_p^k \to \mathbb{Z}_p$. ∎

In Lemma II.4, polynomial function h is called the *polynomial presentation* of f.

Remark 3. Lemma II.4 says that all discrete functions are polynomial functions up to the equivalent relation defined in Definition II.2. This gives a more sophisticated algebraic structure to those functions to be considered. Such a structure will facilitate our further discussion.

III. Functions with Bigger Coding Rate Regions

By Lemma II.4, we can restrict our discussion to polynomial functions, so all the functions considered later have domains and codomains as subsets of finite fields. For s random variables $X_1, X_2, \cdots, X_s$, it is easily seen that the SW region $\mathcal{R}[X_1, X_2, \cdots, X_s]$ gives an inner bound of the

coding rate region for computing any discrete function f, i.e., $\mathcal{R}[X_1, X_2, \cdots, X_s] \subseteq \mathcal{R}[f]$. We consider, in the following, under what condition(s) that $\mathcal{R}[X_1, X_2, \cdots, X_s] \subsetneq \mathcal{R}[f]$.

Lemma III.1 (Elias' Lemma [3])**.** *Given an i.i.d process $Z_1, Z_2, \cdots, Z_m, \cdots$, where $Z_i \sim p, \forall\, i$, for any $\epsilon > 0$, there exists N, such that for all $n > N$ and $k > \dfrac{nH(Z)}{\log|\mathcal{Z}|}$, there exists a $k \times n$ matrix $\mathbf{M}$ and function $\psi : \mathcal{Z}^k \to \mathcal{Z}^n$ such that*

$$\Pr\left\{\psi(\mathbf{M}\vec{Z}^n) \neq \vec{Z}^n\right\} < \epsilon,$$

where $\vec{Z}^n$ is a random vector $[Z_1, Z_2, \cdots, Z_n]^T$.

Theorem III.2. *Let $\mathcal{N} = \{1, 2, \cdots, s\}$. For the polynomial function*

$$f(x_1, x_2, \cdots, x_s) = g\left(\sum_{i=1}^{s} h_i(x_i)\right) \in \mathbb{F}[s], \tag{2}$$

where $g, h_i \in \mathbb{F}[1], \forall\, 1 \leq i \leq s$, $\mathcal{R}[f]$ is inner bounded by the region given by, $\forall\, \emptyset \neq S \subseteq \mathcal{N}$,

$$\sum_{j\in S} R_j \geq I(V_S; Y_S|V_{S^c}) + |S|\; H(Z|V_{\mathcal{N}}), \tag{3}$$

where $\forall\, 1 \leq j \leq s$, $Y_j = h_j(X_j)$, V_j's are discrete random variables such that

$$\begin{aligned} &p(y_1, y_2, \cdots, y_s, v_1, v_2, \cdots, v_s) \\ =&p(y_1, y_2, \cdots, y_s) \prod_{j=1}^{s} p(v_j|y_j), \end{aligned} \tag{4}$$

and $Z = Y_1 + Y_2 + \cdots + Y_s$.

Remark 4. The Theorem III.2 is a direct generalization of the result by Ahlswede and Han [6, Theorem 10], and its proof is similar to theirs. It resumes [6, Theorem 10] if $s = 2$, $\mathbb{F} = \mathbb{Z}_2$ and g and h_i's are identity functions.

Remark 5. A function could have different polynomial presentations over distinct finite fields. 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 $F_2(x, y) = x + y - (x + y)^2$ (on $\mathbb{Z}_3^2$) to the domain $\{0, 1\} \times \{0, 1\} \subset \mathbb{Z}_3^2$. Consequently, let $\mathscr{F}$ be the set of all polynomial presentations of f with format (2) and $\mathscr{V}$ be the set of all $V_{\mathcal{N}}$ satisfying (4). We have $\mathcal{R}[f] \supseteq \text{cov}\left(\bigcup_{F\in\mathscr{F}} \bigcup_{V_{\mathcal{N}}\in\mathscr{V}} \mathcal{R}(F, V_{\mathcal{N}})\right)$, where $\mathcal{R}(F, V_{\mathcal{N}})$ is given by (3) and $\text{cov}(D)$ is defined to be the convex hull of set $D \subseteq \mathbb{R}^s$.

Proof of Theorem III.2: By (2), if the sum $\sum_{i=1}^{s} h_i(x_i)$ is successfully recovered by the decoder, so is the function f. Therefore, achievability follows from the fact that $(R_1, R_2, \cdots, R_s)$ that satisfies (3) is achievable for computing $\sum_{i=1}^{s} h_i(x_i)$. ∎

Theorem III.3. *Given a function $f : \prod_{i=1}^{s} \mathcal{X}_i \to \Omega$ and a distribution $(X_1, X_2, \cdots, X_s) \sim p$, if $f|_{\text{supp}(p)}$ is a restriction of some $F \in \mathbb{F}[s]$ which admits the structure*

$$F(x_1, x_2, \cdots, x_s) = g(h_1(x_1), h_2(x_2), \cdots, h_s(x_s)), \tag{5}$$

where $g \in \mathbb{F}[s]$, and $(h_1, h_2, \cdots, h_s)|_{\text{supp}(p)}$ is not injective, then $\mathcal{R}[X_1, X_2, \cdots, X_s] \subsetneq \mathcal{R}[f]$.

Proof: Since $(h_1, h_2, \cdots, h_s)|_{\text{supp}(p)}$ is not injective,

$$\begin{aligned} &H(X_1, X_2, \cdots, X_s) > H(h_1(X_1), h_2(X_2), \cdots, h_s(X_s)) \\ &\qquad\Longrightarrow \\ &\mathcal{R}[X_1, X_2, \cdots, X_s] \subsetneq \mathcal{R}[h_1(X_1), h_2(X_2), \cdots, h_s(X_s)] \\ &\qquad\subseteq \mathcal{R}[f], \end{aligned}$$

which establishes the theorem. ∎

Remark 6. For $(X_1, X_2, \cdots, X_s) \sim p$ and $1 \leq l \leq s$, let

$$\text{OL}(a, b, l, p) = \left\{ \mathbf{c} \in \prod_{i\neq l} \mathcal{X}_i \,\middle|\, (a, \mathbf{c}), (b, \mathbf{c}) \in \text{supp}(p) \right\}.$$

$(h_1, h_2, \cdots, h_s)|_{\text{supp}(p)}$ is not injective in Theorem III.3 implies that there exists $1 \leq l \leq s$ and $a \neq b \in \mathcal{X}_l$, such that $\text{OL}(a, b, l, p) \neq \emptyset$ and $h_l(a) = h_l(b)$. Consequently,

$$f\left(a, \mathbf{x}_{(l)}\right) = f\left(b, \mathbf{x}_{(l)}\right), \forall\, \mathbf{x}_{(l)} \in \text{OL}(a, b, l, p). \tag{6}$$

In fact, if a discrete function f satisfies (6) for some nonempty $\text{OL}(a, b, l, p)$, then $f|_{\text{supp}(p)}$ has polynomial presentation (5) and $h_l|_{\mathcal{X}_l}$ is not injective. See Lemma A.1 for details.

Theorem III.4. *Fix the support of all distributions considered to be $\Lambda \subseteq \prod_{i=1}^{s} \mathcal{X}_i$. For function $f : \prod_{i=1}^{s} \mathcal{X}_i \to \Omega$, if $f|_\Lambda$ is a restriction of some $F = g \circ h \in \mathbb{F}[s]$, where $g \in \mathbb{F}[1]$ and $h(x_1, x_2, \cdots, x_s) = \sum_{i=1}^{s} k_i(x_i)$, with*

$$|\Lambda| > |h(\Lambda)|, \tag{7}$$

i.e., $h|_\Lambda$ is not injective, then there exists distribution $(X_1, X_2, \cdots, X_s) \sim q$ with $\text{supp}(q) = \Lambda$ such that $\mathcal{R}[X_1, X_2, \cdots, X_s] \subsetneq \mathcal{R}[f]$.

Proof: Let $(X_1, X_2, \cdots, X_s) \sim q$ and $\text{supp}(q) = \Lambda$. (7) implies that $H(Z) < H(X_1, X_2, \cdots, X_s)$, where $Z = h(X_1, X_2, \cdots, X_s)$. Furthermore, there exists some small $\delta > 0$ and q_0 with $\text{supp}(q_0) = \Lambda$, such that $sH(Z) + \delta < H(X_1, X_2, \cdots, X_s)$ if $q = q_0$. Meanwhile, it can be shown that by Elias' Lemma $R_i = \dfrac{\delta}{s+1} + H(Z), \forall 1 \leq i \leq s$, is achievable (see [5]). Therefore, $\sum_{i=1}^{s} R_i < sH(Z) + \delta < H(X_1, X_2, \cdots, X_s)$ which implies $\mathcal{R}[X_1, X_2, \cdots, X_s] \subsetneq \mathcal{R}[f]$ if $q = q_0$. ∎

Remark 7. In fact, any discrete function can be written as a restriction of some $F = g \circ h \in \mathbb{F}[s]$ with $h(x_1, x_2, \cdots, x_s) =$

$\sum_{i=1}^{s} k_i(x_i)$ (see Appendix A for more details). However, (7) is not satisfied in general, and examples include the identity function and the function $f\{0,1\} \times \{0,1\} \to \{0,1,2\}$ given by $f(a,b) = \begin{cases} 0; & \text{if } b = 0, \\ a+b; & \text{if } b = 1. \end{cases}$

Corollary III.5. *In Theorem III.4, condition (7) can be replaced by*

$$\mathrm{Deg}(g) < \max_{\omega \in \Omega} \left\{ \left| f^{-1}(\omega) \bigcap \Lambda \right| \right\}, \tag{8}$$

while the statement holds true.

Proof: Let $\omega_0 = \arg\max_{\omega \in \Omega} \left\{ \left| f^{-1}(\omega) \bigcap \Lambda \right| \right\}$ and $\rho = \left| f^{-1}(\omega_0) \bigcap \Lambda \right|$. We know that the polynomial function $g - \omega_0$ has at most $\mathrm{Deg}(g - \omega_0) = \mathrm{Deg}(g)$ zeros. Thus, $h|_\Lambda$ is not injective, otherwise, $g - \omega_0$ has at least ρ zeros. $\rho > \mathrm{Deg}(g)$, a contradiction. ∎

Remark 8. The min function in Remark 5 together with its polynomial presentation F_2 gives an example of functions illustrated in Theorem III.4 and Corollary III.5.

Theorem III.6. *Fix the support of all distributions considered to be $\Lambda \subseteq \mathcal{X}_1 \times \mathcal{X}_2$. For function $f : \mathcal{X}_1 \times \mathcal{X}_2 \to \Omega$, there exists joint distribution $(X_1, X_2) \sim p$ with $\mathrm{supp}(p) = \Lambda$ such that $\mathcal{R}[X_1, X_2] \subsetneq \mathcal{R}[f]$, if and only if $f|_\Lambda$ is a restriction of some $F \in \mathbb{F}[2]$ which admits the structure*

$$F(x_1, x_2) = g(k_1(x_1) + k_2(x_2)), \tag{9}$$

where $g \in \mathbb{F}[1]$, and $(k_1 + k_2)|_\Lambda$ is not injective.

Remark 9. Theorem III.6 says that, in the case $s = 2$, the necessary condition of Theorem III.4 is also sufficient.

Proof of Theorem III.6: $\Leftarrow$ follows from Theorem III.4. The other direction, $\Rightarrow$, is proved in Appendix B. ∎

Remark 10. Essentially, Theorem III.6 gives another interpretation of [1, Theorem 1]. In the case of $\Lambda = \mathcal{X}_1 \times \mathcal{X}_2$, one way to prove Theorem III.6 is to show that its sufficient condition is equivalent to the inverses of (3.1), (3.11) and (3.13) of [1].

IV. Conclusion

Considering discrete functions as polynomial functions provides several advantages. A polynomial presentation gives a function a much more sophisticated algebraic structure compared to its matrix presentation. This could facilitate investigating properties of the function and unearthing useful information. In Theorem III.4, we have seen that a "nice" summing structure can be utilized via the linear random coding technique to obtain a bigger coding rate region. For the two-sources scenario, Theorem III.6 claims that such a summing structure is the only valuable one in terms of enlarging the coding rate region. To the best of our knowledge, for the multiple-sources scenario, the question of what is the sufficient and necessary condition so that the coding rate region coincides with the SW region is unanswered (see [1]). We expect that this polynomial approach will provide additional insight into this problem in the future.

Appendix A
Discrete Functions as Polynomial Functions

Lemma A.1. *Let $(X_1, X_2, \cdots, X_s) \sim p$. A discrete function f satisfies (6) for some nonempty $\mathrm{OL}(a, b, l, p)$ with $a \neq b \in \mathcal{X}_l$ $(1 \leq l \leq s)$, if and only if $f|_{\mathrm{supp}(p)}$ has polynomial presentation (5) and $(h_1, h_2, \cdots, h_s)|_{\mathrm{supp}(p)}$ is not injective.*

Proof: Without loss of generality, assume that $l = 1$, and let $\Lambda = \mathrm{supp}(p)$. If f satisfies (6) for a nonempty $\mathrm{OL}(a, b, 1, p)$ with $a \neq b \in \mathcal{X}_1$, then there exists $\tilde{f} : \prod_{i=1}^{s} \mathcal{X}_i \to \Omega$ such that $\tilde{f}(a, \mathbf{c}) = \tilde{f}(b, \mathbf{c})$, $\forall\ \mathbf{c} \in \prod_{i=2}^{s} \mathcal{X}_i$, and $\tilde{f}|_\Lambda = f|_\Lambda$. Let g be any polynomial presentation of $\tilde{f}$, and define function $h_1 : \mathbb{F} \to \mathbb{F}$ by $h_1(c) = \begin{cases} a; & \text{if } c \in \{a, b\}, \\ c; & \text{otherwise.} \end{cases}$ Then $F = g(h_1, h_2, \cdots, h_s)$, where $h_i(x) = x$, $\forall\ 2 \leq i \leq s$, is the required polynomial presentation of $f|_\Lambda$. Moreover, $(h_1, h_2, \cdots, h_s)|_\Lambda$ is not injective, since $h_1(a) = h_1(b)$. Sufficiency is obvious. ∎

Lemma A.2. *Given any discrete function $f : \prod_{i=1}^{k} [0, m_i - 1] \to \Omega$, and let $p \geq \prod_{i=1}^{k} m_i$ be a prime. There exists $g \in \mathbb{Z}_p[1]$, such that f is equivalent to $g \circ h_p|_{\mathcal{D}_1 \times \mathcal{D}_2 \times \cdots \times \mathcal{D}_k}$, where $h_p(x_1, x_2, \cdots, x_k) = \sum_{i=1}^{k} x_i \in \mathbb{Z}_p[k]$ and*

$$\mathcal{D}_i = \{0, d_i, 2d_i, \cdots, (m_i - 1)d_i\} \subseteq \mathbb{Z}_p,\ d_i = \prod_{j=1}^{i-1} m_j.$$

Moreover, if $f(a_1, a_2, \cdots, a_k) = f(b_1, b_2, \cdots, b_k)$ and $a_i \neq b_i, \forall\ 1 \leq i \leq k$, then f is equivalent to $g' \circ h_p|_{\mathcal{D}_1 \times \mathcal{D}_2 \times \cdots \times \tilde{\mathcal{D}}_k}$ for some $g' \in \mathbb{Z}_p[1]$, where

$$\tilde{\mathcal{D}}_k = \{0, d_k - 1, 2d_k, \cdots, (m_k - 1)d_k\} \subseteq \mathbb{Z}_p.$$

Proof: Let $\mathcal{D} = \mathcal{D}_1 \times \mathcal{D}_2 \times \cdots \times \mathcal{D}_k$ and η_i be the bijection from $[0, m_i - 1]$ to $\mathcal{D}_i$ given by $\eta_i : t \mapsto td_i$. It is easy to verify that $h_p|_{\mathcal{D}}$ is injective. Besides, there exists injection $\theta : f\left(\prod_{i=1}^{k} [0, m_i - 1]\right) \to \mathbb{Z}_p$, since $\left| f\left(\prod_{i=1}^{k} [0, m_i - 1]\right) \right| \leq p$. Let g be a function satisfying

$$g(h_p(\eta_1(t_1), \eta_2(t_2), \cdots, \eta_k(t_k))) = \theta(f(t_1, t_2, \cdots, t_k)), \\ \forall\ t_i \in [0, m_i - 1], \forall\ 1 \leq i \leq k. \tag{10}$$

By Lemma II.4, $g \in \mathbb{Z}_p[1]$. By definition, f is equivalent to $g \circ h_p|_{\mathcal{D}_1 \times \mathcal{D}_2 \times \cdots \times \mathcal{D}_k}$.

Furthermore, if $f(\mathbf{a}) = f(b_1, b_2, \cdots, b_k)$, where $\mathbf{a} = (a_1, a_2, \cdots, a_k)$, and $a_i \neq b_i, \forall\ 1 \leq i \leq k$, without loss of generality, assume that $a_i = m_i - 1, \forall\ 1 \leq i \leq k - 1$, $a_k = b_1 = b_2 = \cdots = b_{k-1} = 0$ and $b_k = 1$. Redefine

$\eta_k : [0, m_k - 1] \to \tilde{\mathcal{D}}_k$ by $\eta_k(t) = \begin{cases} td_k; & \text{if } t \neq 1, \\ d_k - 1; & \text{if } t = 1. \end{cases}$ We have $h_p|_{\{\mathcal{D}_1 \times \mathcal{D}_2 \times \cdots \times \tilde{D}_k\} \setminus \{\mathbf{a}\}}$ is injective and

$$\begin{aligned} &h_p(\eta_1(a_1), \eta_2(a_2), \cdots, \eta_k(a_k)) \\ =&h_p(\eta_1(b_1), \eta_2(b_2), \cdots, \eta_k(b_k)), \end{aligned}$$

i.e., $h_p|_{\mathcal{D}_1 \times \mathcal{D}_2 \times \cdots \times \tilde{D}_k}$ is non-injective. Similarly, there exists $g' \in \mathbb{Z}_p[1]$, such that (10) holds true for $g = g'$. Therefore, f is equivalent to $g' \circ h_p|_{\mathcal{D}_1 \times \mathcal{D}_2 \times \cdots \times \tilde{D}_k}$. ∎

APPENDIX B
PROOF OF SUFFICIENCY OF THEOREM III.6

$\mathcal{R}[X_1, X_2] \subsetneq \mathcal{R}[f]$ implies that there exists $(R_1, R_2) \in \mathcal{R}[f]$ while $(R_1, R_2) \notin \mathcal{R}[X_1, X_2]$. That is, for some fixed $\delta > 0$, $R_1 < H(X_1|X_2) - \delta$, or $R_2 < H(X_2|X_1) - \delta$ or $R_1 + R_2 < H(X_1, X_2) - \delta$. Moreover, $\forall\ \epsilon > 0$, there exists a big enough n, two encoders $\mathcal{E}_1, \mathcal{E}_2$ and one decoder $\mathcal{D}$, such that $\Pr\{Z^n \neq \mathcal{D}(\mathcal{E}_1(X_1^n), \mathcal{E}_2(X_2^n))\} < \epsilon$, where $Z = f(X_1, X_2)$.

Let $W_i = \mathcal{E}_i(X_i^n), i = 1, 2$. We have, for $\emptyset \neq J \subseteq \{1, 2\}$,

$$\begin{aligned} &n\sum_{j \in J} R_j \geq H(W_J) \geq H(W_J|X_{J^c}^n) \geq I(W_J; X_J^n|X_{J^c}^n) \\ =&H(X_J^n|X_{J^c}^n) - H(X_J^n|W_J, X_{J^c}^n) \\ =&H\left(X_J^n \middle| X_{J^c}^n\right) - H\left(X_J^n \middle| W_J, W_{J^c}, X_{J^c}^n\right) \\ =&\sum_{i=1}^n H\left(X_J^i \middle| X_{J^c}^i\right) - \sum_{i=1}^n H\left(X_J^i \middle| W_J, X_J^{i-1}, W_{J^c}, X_{J^c}^n\right) \\ =&nH\left(X_J^{(A)} \middle| X_{J^c}^{(A)}, A\right) \\ &- nH\left(X_J^{(A)} \middle| W_J, X_J^{A-1}, W_{J^c}, X_{J^c}^n, A\right), \end{aligned}$$

where A is a random variable which is uniformly distributed over $\{1, 2, \cdots, n\}$ and independent to all the other random variables. Suppose $\sum_{j \in J} R_j < H(X_J|X_{J^c}) - \delta$, then

$$H(X_J^{(A)}|W_J, X_J^{A-1}, W_{J^c}, X_{J^c}^{A-1}, X_{J^c}^{(A)}, X_{J^c}^{[A]}, A) > \delta. \quad (11)$$

Besides, by Fano's inequality,

$$\begin{aligned} n\epsilon \geq&H\left(Z^n \middle| W_J, W_{J^c}\right) = \sum_{i=1}^n H\left(Z^{(i)} \middle| W_J, W_{J^c}, Z^{i-1}\right) \\ \geq&\sum_{i=1}^n H\left(Z^{(i)} \middle| W_J, W_{J^c}, Z^{i-1}, X_J^{i-1}, X_{J^c}^n\right) \\ =&nH\left(Z^{(A)} \middle| W_J, X_J^{A-1}, W_{J^c}, X_{J^c}^n, A\right). \end{aligned}$$

Thus,

$$\epsilon > H(Z^{(A)}|W_J, X_J^{A-1}, W_{J^c}, X_{J^c}^{A-1}, X_{J^c}^{(A)}, X_{J^c}^{[A]}, A). \quad (12)$$

Let $X_i = X_i^{(A)}$, $Z = Z^{(A)}$, $Q_i = (X_i^{[A]}, A)$ and $V_i = (W_i, X_i^{A-1})$ for $i = 1, 2$. Then $H(X_J|V_1, V_2, X_{J^c}, Q_{J^c}) > \delta$ and $\epsilon > H(Z|V_1, V_2, X_{J^c}, Q_{J^c})$ by (11) and (12). Moreover, $V_1 - X_1 - X_2 - V_2\ |Q_j$ forms a Markov chain for all $j = 1, 2$. By the Caratheodory theorem, Q_1, Q_2, V_1 and V_2 can be replaced, respectively, by new random variables $Q_{1,\epsilon}, Q_{2,\epsilon}$, $V_{1,\epsilon}$ and $V_{2,\epsilon}$ whose sample spaces are bounded by a fixed finite number ρ. At the same time, $H(X_J|V_{1,\epsilon}, V_{2,\epsilon}, X_{J^c}, Q_{J^c,\epsilon}) > \delta$, $\epsilon > H(Z|V_{1,\epsilon}, V_{2,\epsilon}, X_{J^c}, Q_{J^c,\epsilon})$ and $V_{1,\epsilon}, -X_1 - X_2 - V_{2,\epsilon}$, $|Q_{j,\epsilon}$ forms a Markov chain for all $j = 1, 2$.

Now, let ϵ go to 0, the continuity of entropy guarantees that $H(X_J|\tilde{V}_1, \tilde{V}_2, X_{J^c}, \tilde{Q}_{J^c}) \geq \delta$ and $0 = H(Z|\tilde{V}_1, \tilde{V}_2, X_{J^c}, \tilde{Q}_{J^c})$ for new random variables $\tilde{Q}_1, \tilde{Q}_2$, $\tilde{V}_1$ and $\tilde{V}_2$ whose *pmf*'s are the limit *pmf*'s of $Q_{1,\epsilon}, Q_{2,\epsilon}$, $V_{1,\epsilon}$ and $V_{2,\epsilon}$, respectively. Furthermore, $\tilde{V}_1 - X_1 - X_2 - \tilde{V}_2\ |\tilde{Q}_j$ forms a Markov chain for all $j = 1, 2$.

1) If $J = \{1\}$, i.e., $H(X_1|\tilde{V}_1, \tilde{V}_2, X_2, \tilde{Q}_2) \geq \delta$, then there exist $a \neq b \in \mathcal{X}_1$ and $c \in \mathcal{X}_2$, such that $f(a, c) = f(b, c)$ since $0 = H(Z|\tilde{V}_1, \tilde{V}_2, X_2, \tilde{Q}_2)$. On the other hand, since $\tilde{V}_1 - X_1 - X_2 - \tilde{V}_2\ |\tilde{Q}_2$ is a Markov chain, then $\forall\ x_2 \in \mathcal{X}_2$ with $p(a, x_2)p(b, x_2) > 0$, $f(a, x_2) = f(b, x_2)$. Otherwise, $0 < H(Z|\tilde{V}_1, \tilde{V}_2, X_2, \tilde{Q}_2)$, a contradiction. By Lemma A.1 and the first half of Lemma A.2, $f|_\Lambda$ has polynomial presentation (9) with a non-injective k_1. Hence, $k_1 + k_2|_\Lambda$ is not injective. Similar conclusion can be drawn if $J = \{2\}$.
2) If $J = \{1, 2\}$, i.e., $H(X_1, X_2|\tilde{V}_1, \tilde{V}_2) \geq \delta$, then there exist $(a, c) \neq (b, d) \in \Lambda$, such that $f(a, c) = f(b, d)$. If $c = d$ (or $a = b$), then $H(X_1|\tilde{V}_1, \tilde{V}_2, X_2, \tilde{Q}_2) \geq \delta$ (or $H(X_2|\tilde{V}_1, \tilde{V}_2, X_1, \tilde{Q}_1) \geq \delta$). Then the same conclusion is reached by the same argument as in 1). Assume that $a \neq b$ and $c \neq d$. By the second half of Lemma A.2, $f|_\Lambda$ processes a polynomial presentation (9) with non-injective $k_1 + k_2|_\Lambda$.

REFERENCES

[1] 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.

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

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

[4] R. G. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 1968.

[5] 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.

[6] 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.

[7] A. Orlitsky and R. Roche, "Coding for computing," *IEEE Transactions on Information Theory*, vol. 47, Mar. 2001.

[8] H. Witsenhausen, "The zero-error side information problem and chromatic numbers," *IEEE Transactions on Information Theory*, vol. 22, pp. 592–593, September 1976.

[9] B. Nazer and M. Gastpar, "Computation over multiple-access channels," *IEEE Transactions on Information Theory*, vol. 53, Oct. 2007.

[10] R. Appuswamy, M. Franceschetti, N. Karamchandani, and K. Zeger, "Network coding for computing: Cut-set bounds," *IEEE Transactions on Information Theory*, vol. 57, Feb. 2011.

[11] R. Appuswamy, M. Franceschetti, N. Karamchandani, and K. Zeger, "Linear codes, target function classes, and network computing capacity," *IEEE Transactions on Information Theory*, Submitted.

[12] R. Lidl and H. Niederreiter, *Finite Fields*. New York: Gambridge University Press, 2nd ed., 1997.