Computing Polynomial Functions of Correlated Sources: Inner Bounds

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 considers the problem of source coding for computing functions of correlated i.i.d. random sources. The approach of combining standard and linear random coding for this problem was first introduced by Ahlswede and Han, in the special case of computing the modulo-two sum. In this paper, making use of an adapted version of that method, we generalize their result to more sophisticated scenarios, where the functions to be computed are polynomial functions. Since all discrete functions are fundamentally restrictions of polynomial functions, our results are universally applied.**

I. Introduction

For each $1 \leq i \leq s$, let t_i be a source that randomly generates discrete i.i.d. data $X_i^{(1)}, X_i^{(2)}, \cdots, X_i^{(n)}, \cdots$, where $X_i^{(n)}$ has sample space $\mathscr{X}_i$ and $\left(X_1^{(n)}, X_2^{(n)}, \cdots, X_s^{(n)}\right) \sim p$, $\forall\ n \in \mathbb{N}^+$. Let $f : \prod_{i=1}^{s} \mathscr{X}_i \to \Omega$ be some *discrete function*. The source coding problem of computing a function of correlated sources considers: what is the *achievable coding rate* $(R_1, R_2, \cdots, R_s) \in \mathbb{R}^s$, such that, $\forall\ \epsilon > 0$, there exists a big enough n, s encoders $\mathcal{E}_1, \mathcal{E}_2, \cdots, \mathcal{E}_s$, where $\mathcal{E}_i : \mathscr{X}_i^n \to [1, 2^{nR_i}], 1 \leq i \leq s$, and one decoder $\mathcal{D} : \prod_{i=1}^{s}[1, 2^{nR_i}] \to \Omega^n$, such that $\Pr\left\{\mathcal{D}\left[\prod_{i=1}^{s}\mathcal{E}_i\left(X_i^n\right)\right] \neq \vec{f}\left(\prod_{i=1}^{s} X_i^n\right)\right\} < \epsilon$, where $\vec{f}\left(\prod_{i=1}^{s} X_i^n\right) = \left\{f\left[X_1^{(j)}, X_2^{(j)}, \cdots, X_s^{(j)}\right]\right\}_{j=1}^{n}$.

Denote the *Slepian–Wolf region* [1] by

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

and let $\mathcal{R}[f]$ be the *achievable 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 readily seen that $\mathcal{R}[X_1, X_2, \cdots, X_s] \subseteq \mathcal{R}[f]$. Based on the method of Elias [2] (cf. [3]), Körner and Marton [4] show that if f is the *modulo-two sum* $\oplus_2$, $\mathcal{R}[\oplus_2]$ contains the convex hull of the union $\mathcal{R}[X_1, X_2] \cup \mathcal{R}_{\oplus_2}$, where

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

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

Subsequently, by combining standard source coding techniques with Elias' Lemma, Ahlswede and Han [5] give an inner bound of $\mathcal{R}[\oplus_2]$ which is larger than Körner and Marton's in general. Adopting an approach based on polynomial representations, the authors in [6] present several conditions under which the achievable coding rate region is strictly bigger than the Slepian–Wolf region.

This paper generalizes the method of Ahlswede–Han [5, Theorem 10] to more general circumstances, where the functions considered are not necessarily linear (the modulo-two sum in [5]). The fundamental idea of our approach is to combine the standard and linear random coding techniques, as first proposed by Ahlswede and Han. To make their method works in non-linear scenarios, we introduce the modifications needed and the result from Csiszár [7]. In the case of computing the modulo-two sum, some minor improvement is also gained regarding the inner bound of the achievable coding rate region presented in Ahlswede–Han (see Remark 5). In addition, the observation that all discrete functions are *restrictions* of *polynomial functions* is another important factor which works underneath our method. In essence, 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 *polynomial presentation* is of great importance.

This paper is organized as follows: some needed preliminaries about polynomial functions are given in section II. In section III, Theorem III.1 and its corollary demonstrate how the idea works. The proof of Theorem III.1 is provided in section IV showing the mechanism. In section V, Theorem V.1, a generalized result of Theorem III.1, is presented without proof because of space limitation. Section VI is the conclusion.

A. Notation

Capital letters $X, Y, Z, \cdots$ are used to denote random variables and $\mathscr{X}, \mathscr{Y}, \mathscr{Z}, \cdots$ for their corresponding sample spaces. Meanwhile, 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.

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 def-

initions of X_J^n and $X_J^{(i)}$ resemble such a definition. If $(X_1, X_2, \cdots, X_s) \sim p$, then p_{X_J} is defined to be the marginal regarding X_J, and p_{X_J} is replaced by p when X_J is clearly referred to. The set of all *ϵ-typical sequences* of length n regarding X_J is define to be

$$\mathcal{T}_\epsilon(n, X_J) = \Bigg\{ x_J^n \in \prod_{j\in J} \mathscr{X}_j^n \Bigg| \left| -\frac{\log p_{X_I^n}(x_I^n)}{n} - H(X_I) \right| < \epsilon, \forall\, \emptyset \neq I \subseteq J \Bigg\},$$

where $p_{X_I^n}(x_I^n) = \prod_{i=1}^{n} p_{X_I}\left(x_I^{(i)}\right)$. We replace $\mathcal{T}_\epsilon(n, X_J)$ with $\mathcal{T}_\epsilon$ when the length n and the random variable array X_J referred to are clear from the context.

II. Polynomial Functions

We demonstrated how a discrete function can be treated as a polynomial function in this section. Readers who are familiar with algebra may wish to go through quickly to get familiar with our notation.

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 the 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 $\mathbb{F}$.

Definition II.2. Given two functions $f : \prod_{i=1}^{k} \mathscr{X}_i \to \Omega_1$ and $g : \prod_{i=1}^{k} \mathscr{Y}_i \to \Omega_2$, f and g are said to be *equivalent* if and only if there exist bijections $\phi_i : \mathscr{X}_i \to \mathscr{Y}_i$, $\forall\, 1 \leq i \leq k$, 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 achievable coding rate regions are the same. From now on, we will simply refer to two equivalent functions as one function.

[1]Polynomial and polynomial function are distinct concepts (cf. [8]).

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

Definition II.3. Given function $f : \mathcal{D} \to \Omega$, and let $\emptyset \neq S \subseteq \mathcal{D}$. 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 (cf. Lemma 7.40 of [8])**.** *Any discrete function $f(x_1, x_2, \cdots, x_k)$ defined on domain $\mathcal{D} = \prod_{i=1}^{k} \mathscr{X}_i$ is equivalent to a restrictions of some polynomial function $h \in \mathbb{F}_{p^m}[k]$, where p is any prime that $p^m \geq \max\{|f(\mathcal{D})|, |\mathscr{X}_i| \mid 1 \leq i \leq k\}$.*

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

Remark 2. Lemma II.4 says that all discrete functions are basically polynomial functions up to the equivalency defined by Definition II.2. Therefore, we can confine all the functions considered to polynomial functions.

III. Coding for Computing Polynomial Functions

In this section, we consider those polynomial functions $f \in \mathbb{F}[s]$ with the structure

$$g\left[h(x_1, x_2, \cdots, x_t) + \sum_{j=t+1}^{s} k_j(x_j)\right], 1 \leq t < s. \qquad (2)$$

By Lemma A.1 and Lemma A.2 of [6], for any discrete function f', there always exists some finite field $\mathbb{F}'$ and polynomial function $F' \in \mathbb{F}'[s]$ with structure (2), such that F' gives a polynomial presentation of f' (from these two lemmata, the final presentation has $t = 1$). On this respect, consider polynomial functions of format (2) is to treat the computing problem universally.

Theorem III.1. *Let $\mathcal{N} = \{1, 2, \cdots, s\}$ and $T = \{1, 2, \cdots, t\}$. If f admits a polynomial presentation (2), then $\mathcal{R}[f]$ is inner bounded by the region given by, $\forall\, \emptyset \neq S \subseteq \mathcal{N}$,*

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

where $\forall\, j \in T$, $V_j = Y_j = X_j$; $\forall\, j \in \mathcal{N} \setminus T$, $Y_j = k_j(X_j)$, V_j's are discrete random variables such that

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

and $Z = h(X_1, X_2, \cdots, X_t) + \sum_{j=t+1}^{s} k_j(X_j)$.

Remark 3. In Theorem III.1, Z can be replaced by $W = \sum_{j=t+1}^{s} k_j(X_j)$ while the theorem holds true, since $H(Z|V_{\mathcal{N}}) = H(W|V_{\mathcal{N}})$ when $V_T = X_T$.

Remark 4. 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 $F_1(x, y) = xy \in \mathbb{Z}_2[2]$ or be treated as the restriction of $F_2(x, y) = x + y - (x + y)^2 \in \mathbb{Z}_3[2]$ to the domain $\{0, 1\} \times \{0, 1\} \subsetneq \mathbb{Z}_3^2$. Let $\mathcal{F}$ be the set of all polynomial

presentations of f with format (2) and $\mathcal{V}$ be the set of all $V_{\mathcal{N}}$ satisfying (4). We have

$$\mathcal{R}[f] \supseteq \mathrm{cov}\left(\bigcup_{F\in\mathcal{F}}\bigcup_{V_{\mathcal{N}}\in\mathcal{V}}\mathcal{R}(F,V_{\mathcal{N}})\right), \tag{5}$$

where $\mathcal{R}(F,V_{\mathcal{N}})$ is given by (3) and $\mathrm{cov}(D)$ is defined to be the convex hull of the set $D\subseteq\mathbb{R}^s$.

Based on Theorem III.1, one simple conclusion is derived as follows. For two discrete random variables Y and V, let $\mathscr{I}(Y,V)=\begin{cases}0, & \text{if } p_{Y,V}(y,v)=p_V(v) \text{ or } 0;\\ 1, & \text{otherwise.}\end{cases}$ We have

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

$$f(x_1,x_2,\cdots,x_s)=g\left[\sum_{i=1}^{s}h_i(x_i)\right]\in\mathbb{F}[s], \tag{6}$$

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

$$\sum_{j\in S}R_j\ge I(V_S;Y_S|V_{S^c})+H(Z|V_{\mathcal{N}})\sum_{j\in S}\mathscr{I}(Y_j,V_j), \tag{7}$$

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

$$\begin{aligned}&p(y_1,y_2,\cdots,y_s,v_1,v_2,\cdots,v_s)\\ =&p(y_1,y_2,\cdots,y_s)\prod_{j=1}^{s}p(v_j|y_j),\end{aligned} \tag{8}$$

and $Z=\sum_{i=1}^{s}Y_i$.

Remark 5. Corollary III.2 resumes the result of Ahlswede–Han [5, Theorem 10] if $s=2$, $\mathbb{F}=\mathbb{Z}_2$ and g and h_i's are identity functions. Actually, (7) gives a potentially bigger region than the one given by (6.4), (6.5) and (6.6) of [5], since $\mathscr{I}$ will remove one copy of $H(Z|V_{\mathcal{N}})$ whenever Y_j is a function of V_j for any $1\le j\le s$.

IV. Proof of Theorem III.1

It suffices to prove only the strict inequality of (3). Let $(R_1,R_2,\cdots,R_s)\in\mathbb{R}^s$ satisfy (3) strictly. Then there exist $\delta>6\epsilon>0$, such that $R_j=R'_j+R''_j$, $\forall\ j\in\mathcal{N}$, $\sum_{j\in S}R'_j>I(Y_S;V_S|V_{S^c})+2|S|\delta$, $\forall\ \emptyset\neq S\subseteq\mathcal{N}$, and $R''_j>H(Z|V_{\mathcal{N}})+2\delta$, $\forall\ j\in\mathcal{N}\setminus T$.

A. Codebook:

Fix the joint distribution p which satisfies (4). For all $j\in T$, let $\mathscr{V}_{j,\epsilon}$ be the set of all the ϵ-typical sequences in $\mathscr{X}_j^n$ (Note: $\mathscr{X}_j,\mathscr{Y}_j,\mathscr{Z}_j\subseteq\mathbb{F},\forall\ j\in\mathcal{N}$). For all $j\in\mathcal{N}\setminus T$, generate randomly $2^{n[I(Y_j;V_j)+\delta]}$ ϵ-typical sequences according to distribution $p_{V_j^n}$ and let $\mathscr{V}_{j,\epsilon}$ be the set of these generated sequences. Define mapping $\phi_j:\mathscr{X}_j^n\to\mathscr{V}_{j,\epsilon}$ as follows:

1) If $j\in T$, then, $\forall\ \mathbf{x}\in\mathscr{X}_j^n$, $\phi_j(\mathbf{x})=\begin{cases}\mathbf{x}, & \text{if } \mathbf{x}\in\mathcal{T}_\epsilon;\\ \mathbf{x}_0, & \text{otherwise,}\end{cases}$ where $\mathbf{x}_0$ is some fixed ϵ-typical sequence.
2) If $j\in\mathcal{N}\setminus T$, then for every $\mathbf{x}\in\mathscr{X}_j^n$, let $\mathcal{L}_{\mathbf{x}}=\{\mathbf{v}\in\mathscr{V}_{j,\epsilon}|(\vec{k}_j(\mathbf{x}),\mathbf{v})\in\mathcal{T}_\epsilon\}$. If $\mathbf{x}\in\mathcal{T}_\epsilon$ and $\mathcal{L}_{\mathbf{x}}\neq\emptyset$, then $\phi_j(\mathbf{x})$ is set to be some element in $\mathcal{L}_{\mathbf{x}}$; otherwise $\phi_j(\mathbf{x})$ is some fixed $\mathbf{v}_0\in\mathscr{V}_{j,\epsilon}$.

Define mapping $\eta_j:\mathscr{V}_{j,\epsilon}\to[1,2^{nR'_j}]$ by randomly choosing the value for each $\mathbf{v}\in\mathscr{V}_{j,\epsilon}$ according to a uniform distribution.

Let $k=\min_{j\in\mathcal{N}\setminus T}\left\{\left\lfloor\frac{nR''_j}{\log|\mathbb{F}|}\right\rfloor\right\}$. When n is big enough, $k>\frac{n[H(Z|V_{\mathcal{N}})+\delta]}{\log|\mathbb{F}|}$. Randomly generate a $k\times n$ matrix $\mathbf{M}\in\mathbb{F}^{k\times n}$, and for $j\in\mathcal{N}\setminus T$, let $\theta_j:\mathscr{X}_j^n\to\mathbb{F}^k$ be the funciton $\theta_j:\mathbf{x}\mapsto\mathbf{M}\vec{k}_j(\mathbf{x}),\forall\ \mathbf{x}\in\mathscr{X}_j^n$.

B. Encoding:

For $j\in T$, the encoder $\mathcal{E}_j=\eta_j\circ\phi_j$, a scale-valued function. For $j\in\mathcal{N}\setminus T$, $\mathcal{E}_j=(\eta_j\circ\phi_j,\theta_j)$, a vector-valued function. Assume that X_j^n is the data generated by the jth source, then $\mathscr{A}=(a_1,a_2,\cdots,a_t,(a_{t+1},b_{t+1}),\cdots,(a_s,b_s))=(\mathcal{E}_1(X_1^n),\mathcal{E}_2(X_2^n),\cdots,\mathcal{E}_s(X_s^n))$ is sent.

C. Decoding:

Upon observing $\mathscr{A}$ at the decoder, the decoder claims that $\vec{g}(\hat{Z}^n)$ is the function of the generated data, i.e., $\vec{g}(\hat{Z}^n)=\vec{f}(X_1^n,X_2^n,\cdots,X_s^n)$, if and only if there exists one and only one $\hat{\mathbf{V}}=(\hat{\mathbf{v}}_1,\hat{\mathbf{v}}_2,\cdots,\hat{\mathbf{v}}_s)\in\prod_{j=1}^{s}\mathscr{V}_{j,\epsilon}$, such that $a_j=\eta_j(\hat{\mathbf{v}}_j),\forall\ j\in\mathcal{N}$ and $\hat{Z}^n$ is the only element in the set

$$\mathscr{L}_{\hat{\mathbf{V}}}=\Big\{\mathbf{z}\in\mathscr{Z}^n\Big|(\mathbf{z},\hat{\mathbf{V}})\in\mathcal{T}_\epsilon,\ \mathbf{M}\mathbf{z}=\mathbf{M}\vec{h}(\hat{\mathbf{v}}_1,\hat{\mathbf{v}}_2,\cdots,\hat{\mathbf{v}}_t)+\sum_{j=t+1}^{s}b_j\Big\}.$$

D. Error:

Assume that X_j^n is the data generated by the jth source and let $Z^n=\vec{h}\,(X_1^n,X_2^n,\cdots,X_t^n)+\sum_{j=t+1}^{s}\vec{k}_j(X_j^n)$. An error happens if and only if one of the following events happens.

E_1: $(X_1^n,X_2^n,\cdots,X_s^n,Y_1^n,Y_2^n,\cdots,Y_s^n,Z^n)\notin\mathcal{T}_\epsilon$;
E_2: There exists some $j_0\in\mathcal{N}$, such that $\mathcal{L}_{X_{j_0}^n}=\emptyset$;
E_3: $(Y_1^n,Y_2^n,\cdots,Y_s^n,Z^n,\mathbf{V})\notin\mathcal{T}_\epsilon$, where $\mathbf{V}=(\mathbf{v}_1,\mathbf{v}_2,\cdots,\mathbf{v}_s)$ and $\mathbf{v}_j=\phi_j(X_j^n),\forall\ j\in\mathcal{N}$;
E_4: There exists $\mathbf{V}'=(\mathbf{v}'_1,\mathbf{v}'_2,\cdots,\mathbf{v}'_s)\in\prod_{j=1}^{s}\mathscr{V}_{j,\epsilon}$, $\mathbf{V}'\neq\mathbf{V}$, such that $\eta_j(\mathbf{v}'_j)=\eta_j(\mathbf{v}_j)$, $\forall\ j\in\mathcal{N}$ and $\mathbf{V}'\in\mathcal{T}_\epsilon$;
E_5: $Z^n\notin\mathscr{L}_{\mathbf{V}}$ or $|\mathscr{L}_{\mathbf{V}}|>1$, i.e., there exists $Z_0^n\in\mathscr{Z}^n$, $Z^n\neq Z_0^n$, such that $\mathbf{M}Z_0^n=\mathbf{M}Z^n$ and $(Z_0^n,\mathbf{V})\in\mathcal{T}_\epsilon$.

Let $\gamma=\Pr\left\{\bigcup_{l=1}^{5}E_l\right\}=\sum_{l=1}^{5}\Pr\{E_l|E_{l,c}\}$, where $E_{1,c}=\emptyset$ and $E_{l,c}=\bigcap_{\tau=1}^{l-1}E_\tau^c$ for $1<l\le5$. In the following, we show that $\gamma\to0$, $n\to\infty$.

(a). By the joint AEP, $\Pr\{E_1\}\to0$, $n\to\infty$.

(b). Let $E_{2,j}=\left\{\mathcal{L}_{X_j^n}=\emptyset\right\}$, $\forall j\in\mathcal{N}$. Then

$$\Pr\{E_2|E_{2,c}\}\le\sum_{j\in\mathcal{N}}\Pr\{E_{2,j}|E_{2,c}\}. \tag{9}$$

If $j\in T$, then $\mathcal{L}_{X_j^n}=\{X_j^n\}$, thus $\Pr\{E_{2,j}\}=0$. If $j\in\mathcal{N}\backslash T$, then

$$\begin{aligned}\Pr\{E_{2,j}|E_{2,c}\}=&\Pr\left\{\mathcal{L}_{X_j^n}=\emptyset\mid E_{2,c}\right\}\\=&\prod_{\mathbf{v}\in\mathscr{V}_{j,\epsilon}}\Pr\left\{\left(\vec{k}_j(X_j^n),\mathbf{v}\right)\notin\mathcal{T}_\epsilon\right\}\\<&\left\{1-2^{-n[I(Y_j;V_j)+\delta/2]}\right\}^{2^{n[I(Y_j;V_j)+\delta]}}\\&\to 0,n\to\infty.\end{aligned}\tag{10}$$

Notice that the sequence $\mathbf{v}\in\mathscr{V}_{j,\epsilon}$ and $Y_j^n=\vec{k}_j(X_j^n)$ are drawn independently, therefore,

$$\begin{aligned}\Pr\{(Y_j^n,\mathbf{v})\in\mathcal{T}_\epsilon\}\ge&(1-\epsilon)2^{-n[I(Y_j;V_j)+3\epsilon]}\\=&(1-\epsilon)2^{-n[I(Y_j;V_j)+\delta/2]+n(\delta/2-3\epsilon)}\\>&2^{-n[I(Y_j;V_j)+\delta/2]}\end{aligned}$$

when n is big enough, thus, (10) holds true for all big enough n. Now, we use the fact that $(1-1/a)^a\to e^{-1},a\to\infty$, to show that the sum on the right hand side of inequality (10) has limit 0. By (9), $\Pr\{E_2|E_{2,c}\}\to 0$, $n\to\infty$.

(c). By (4), it is obvious that $V_{J_1}-Y_{J_1}-Y_{J_2}-V_{J_2}$ forms a Markov chain for any two disjoint nonempty sets $J_1,J_2\subsetneq\mathcal{N}$. Thus, if $(Y_j^n,\mathbf{v}_j)\in\mathcal{T}_\epsilon$ for all $j\in\mathcal{N}$ and $(Y_1^n,Y_2^n,\cdots,Y_s^n)\in\mathcal{T}_\epsilon$, then $(Y_1^n,Y_2^n,\cdots,Y_s^n,\mathbf{V})\in\mathcal{T}_\epsilon$. In the meantime, $Z-(Y_1,Y_2,\cdots,Y_s)-(V_1,V_2,\cdots,V_s)$ is a Markov chain, hence, $(Y_1^n,Y_2^n,\cdots,Y_s^n,Z^n,\mathbf{V})\in\mathcal{T}_\epsilon$ if $(Y_1^n,Y_2^n,\cdots,Y_s^n,Z^n)\in\mathcal{T}_\epsilon$. Therefore, $\Pr\{E_3|E_{3,c}\}=0$.

(d). For all $\emptyset\ne J\subseteq\mathcal{N}$, let $J=\{j_1,j_2,\cdots,j_{|J|}\}$ and

$$\begin{aligned}\Gamma_J=\Big\{\mathbf{V}'=(\mathbf{v}_1',\mathbf{v}_2',\cdots,\mathbf{v}_s')\in\prod_{j=1}^s\mathscr{V}_{j,\epsilon}\Big|\\\mathbf{v}_j'=\mathbf{v}_j\text{ if and only if }j\in\mathcal{N}\setminus J\Big\}.\end{aligned}$$

Then $|\Gamma_J|=\left|\prod_{j\in J}\mathscr{V}_{j,\epsilon}\right|-1=2^{n\left[\sum_{j\in J}I(Y_j;V_j)+|J|\delta\right]}-1$ and

$$\begin{aligned}&\Pr\{E_4|E_{4,c}\}\\=&\sum_{\emptyset\ne J\subseteq\mathcal{N}}\sum_{\mathbf{V}'\in\Gamma_J}\Pr\left\{\eta_j(\mathbf{v}_j')=\eta_j(\mathbf{v}_j),\forall\,j\in J,\mathbf{V}'\in\mathcal{T}_\epsilon|E_{4,c}\right\}\\=&\sum_{\emptyset\ne J\subseteq\mathcal{N}}\sum_{\mathbf{V}'\in\Gamma_J}\Pr\left\{\eta_j(\mathbf{v}_j')=\eta_j(\mathbf{v}_j),\forall\,j\in J\right\}\\&\times\Pr\{\mathbf{V}'\in\mathcal{T}_\epsilon|E_{4,c}\}\qquad(11)\\<&\sum_{\emptyset\ne J\subseteq\mathcal{N}}\sum_{\mathbf{V}'\in\Gamma_J}2^{-n\sum_{j\in J}R_j'}\\&\times2^{-n\left[\sum_{i=1}^{|J|}I(V_{j_i};V_{J^c},V_{j_1},\cdots,V_{j_{i-1}})-|J|\delta\right]}\qquad(12)\\<&\sum_{\emptyset\ne J\subseteq\mathcal{N}}2^{n\left[\sum_{j\in J}I(Y_j;V_j)+|J|\delta\right]}\times2^{-n\sum_{j\in J}R_j'}\\&\times2^{-n\left[\sum_{i=1}^{|J|}I(V_{j_i};V_{J^c},V_{j_1},\cdots,V_{j_{i-1}})-|J|\delta\right]}\\\le&C\max_{\emptyset\ne J\subseteq\mathcal{N}}2^{-n\left[\sum_{j\in J}R_j'-I(Y_J;V_J|V_{J^c})-2|J|\delta\right]}\qquad(13)\\&\to 0,n\to\infty,\end{aligned}$$

where $C=2^s-1$. Equality (11) holds because the processes of choosing η_j's and generating $\mathbf{V}'$ are done independently. (12) follows from Lemma IV.1 and the definitions of η_j's. (13) is from Lemma IV.2.

Lemma IV.1. *Let $X_1,X_2,\cdots,X_l,Y$ be $l+1$ independent random variables. For any $\epsilon>0$ and positive integer n, if $\mathbf{y}\in\mathscr{Y}^n$ is an ϵ-typical sequence, then*

$$\begin{aligned}&\Pr\{(X_1^n,X_2^n,\cdots,X_l^n,Y^n)\in\mathcal{T}_\epsilon|Y^n=\mathbf{y}\}\\\le&2^{-n\left[\sum_{j=1}^lI(X_j;Y,X_1,X_2,\cdots,X_{j-1})-3l\epsilon\right]}.\end{aligned}$$

Proof: Let F_j be the event $\{(X_1^n,X_2^n,\cdots,X_j^n,Y^n)\in\mathcal{T}_\epsilon\}$, $1\le j\le l$, and $F_0=\emptyset$. We have

$$\begin{aligned}&\Pr\{(X_1^n,X_2^n,\cdots,X_l^n,Y^n)\in\mathcal{T}_\epsilon|Y^n=\mathbf{y}\}\\=&\prod_{j=1}^l\Pr\{F_j|Y^n=\mathbf{y},F_{j-1}\}\\\le&\prod_{j=1}^l2^{-n[I(X_j;Y,X_1,X_2,\cdots,X_{j-1})-3\epsilon]}\\=&2^{-n\left[\sum_{j=1}^lI(X_j;Y,X_1,X_2,\cdots,X_{j-1})-3l\epsilon\right]},\end{aligned}$$

since $X_1,X_2,\cdots,X_l,Y$ are independent. ■

Lemma IV.2. *If $(Y_1,V_1,Y_2,V_2,\cdots,Y_s,V_s)\sim q$, and*

$$q(y_1,v_1,y_2,v_2,\cdots,y_s,v_s)=q(y_1,y_2,\cdots,y_s)\prod_{i=1}^sq(v_i|y_i),$$

then, $\forall\,J=\{j_1,j_2,\cdots,j_{|J|}\}\subseteq\{1,2,\cdots,s\}$,

$$I(Y_J;V_J|V_{J^c})=\sum_{i=1}^{|J|}I(Y_{j_i};V_{j_i})-I(V_{j_i};V_{J^c},V_1,\cdots,V_{j_{i-1}}).$$

(e). Let $E_{5,1}=\{\mathscr{L}_{\mathbf{V}}=\emptyset\}$ and $E_{5,2}=\{|\mathscr{L}_{\mathbf{V}}|>1\}$. We have $\Pr\{E_{5,1}|E_{5,c}\}=0$, because $E_{5,c}$ contains the event that $\mathbf{V}$ is the only element in $\prod_{j=1}^s\mathscr{V}_{j,\epsilon}$ satisfying $(Z^n,\mathbf{V})\in\mathcal{T}_\epsilon$ and

$$\begin{aligned}\mathbf{M}Z^n=&\mathbf{M}\left[\vec{h}(X_1^n,X_2^n,\cdots,X_t^n)+\sum_{j=t+1}^s\vec{k}_j(X_j^n)\right]\\=&\mathbf{M}\vec{h}(X_1^n,X_2^n,\cdots,X_t^n)+\sum_{j=t+1}^s\mathbf{M}\vec{k}_j(X_j^n)\\=&\mathbf{M}\vec{h}(\mathbf{v}_1,\mathbf{v}_2,\cdots,\mathbf{v}_t)+\sum_{j=t+1}^s\theta_j(X_j^n).\end{aligned}$$

Therefore,

$$\begin{aligned}\Pr\{E_5|E_{5,c}\}=&\Pr\{E_{5,2}|E_{5,c}\}\\=&\sum_{Z_0^n\ne Z^n,(Z_0^n,\mathbf{V})\in\mathcal{T}_\epsilon}\Pr\{\mathbf{M}Z_0^n=\mathbf{M}Z^n\}\\<&2^{n[H(Z|V_{\mathcal{N}})+\delta]}\times2^{-k\log|\mathbb{F}|}\end{aligned}$$

$$=2^{-n[k\log|\mathbb{F}|/n-H(Z|V_{\mathcal{N}})-\delta]}\to 0, n\to\infty,$$

where $\Pr\{\mathbf{M}Z_0^n=\mathbf{M}Z^n\}<2^{-k\log|\mathbb{F}|}$ is from [2] (cf. [3]).

To summarize, by (a)–(e), we have $\gamma\to 0, n\to\infty$. The theorem is established.

Remark 6. The main idea of the above proof is to combine the standard and linear random coding techniques, as originally proposed by Ahlswede and Han [5]. In order to deal with the non-linear phenomenon, a trick is introduced. Observe that, in the proof, before decoding Z^n, $\mathbf{V}=(V_1^n,V_2^n,\cdots,V_s^n)$ has to be decoded correctly. Here, we gain the opportunity to "eliminate" the non-linear effect from the function, if V_j is set to be X_j for all $1\le j\le t$ when constructing the codebook. Once $\mathbf{V}$ is correctly decoded, the value of $\vec{h}$ is known to the decoder. At this point, to decode Z^n is seen as a linear computation problem from the decoder's point of view. However, it will be difficult to cope with the non-linear effect without making use of a polynomial presentation.

V. Generalization

The approach used to obtaining a inner bound of $\mathcal{R}[f]$ for computing a function f with presentation (2) is applicable to many other scenarios. Later on, we present another theorem for functions with presentation

$$g\left[h_{\vec{m}}(x_{01},x_{02},\cdots,x_{0r}),\sum_{j=1}^{l_1}k_{1j}(x_{1j}),\right.$$
$$\left.\sum_{j=1}^{l_2}k_{2j}(x_{2j}),\cdots,\sum_{j=1}^{l_t}k_{tj}(x_{tj})\right]\in\mathbb{F}\left[r+\sum_{i=1}^{t}l_i\right],\quad(14)$$

where $h_{\vec{m}}$ is defined to be a vector-valued function $(h_1,h_2,\cdots,h_m)$ and $h_i\in\mathbb{F}[r],\forall\ 1\le i\le m$. It is easily seen that (2) is a special case of (14).

Theorem V.1. *Let* $T_0=\{01,02,\cdots,0r\}$, $T_i=\{i1,i2,\cdots,il_i\}$ *and* $\mathcal{N}=\bigcup_{i=0}^{t}T_i$. *If* $f\in\mathbb{F}\left[r+\sum_{i=1}^{t}l_i\right]$ *has presentation (14), then* $\mathcal{R}[f]$ *contains the region given by*

$$\sum_{j\in S}R_j\ge I(Y_S;V_S|V_{S^c})+\sum_{i=1}^{t}\rho_i\sum_{j\in S\cap T_i}\mathscr{I}(Y_j,V_j),$$
$$\forall\ \emptyset\ne S\subseteq\mathcal{N},\quad(15)$$

where $\forall\ j\in T_0$, $V_j=Y_j=X_j$; $\forall\ j\in\mathcal{N}\setminus T_0$, $Y_j=k_j(X_j)$, V_j*'s are discrete random variables such that*

$$p\left((y_j)_{j\in\mathcal{N}},(v_j)_{j\in\mathcal{N}}\right)=p\left((y_j)_{j\in\mathcal{N}}\right)\prod_{j\in\mathcal{N}\setminus T_0}p(v_j|y_j);\quad(16)$$

$Z_i=\sum_{j\in T_i}Y_j$, $\forall\ 1\le i\le t$; *and* $\sum_{i\in\mathcal{I}}\rho_i=H(Z_{\mathcal{I}}|V_{\mathcal{N}},Z_{\mathcal{I}^c})$, $\forall\ \emptyset\ne\mathcal{I}\subseteq\{1,2,\cdots,t\}$.

Remark 7. Instead of using Elias' Lemma, the linear technique used to prove the above theorem is from Csiszár [7]. The proof is omitted because of limitation of space.

Remark 8. Let $\mathcal{F}'$ be the set of all polynomial presentations of f with format (14) and $\mathcal{V}'$ be the set of all $V_{\mathcal{N}}$ satisfying (16). We have

$$\mathcal{R}[f]\supseteq\mathrm{cov}\left(\bigcup_{F\in\mathcal{F}'}\bigcup_{V_{\mathcal{N}}\in\mathcal{V}'}\mathcal{R}'(F,V_{\mathcal{N}})\right),\quad(17)$$

where $\mathcal{R}'(F,V_{\mathcal{N}})$ is given by (15). It is mentioned before that any discrete function has a polynomial presentation with structure (2). However, it does not seem straightforward that inner bound (5) is as good as (17), nor (17) is strictly better than (5).

VI. Conclusion

The essence that separates the function computing problem from the Slepian–Wolf source coding problem is the structure of the function considered. One of the most important structure regarding a function is its polynomial structure. As shown in Theorem III.1, Corollary III.2 and Theorem V.1, the polynomial presentation unveils valuable information which can be used in the construction of codebooks, encoders and decoder. It is likely that similar polynomial approaches can also be employed to reach performance gains for function computing problems in the general channel or network coding settings.

The situation that is hard to handle with our approach is either when a function is not given by a polynomial presentation, for instance $\min(x,y)$, or when known coding methods are not directly applicable to the given polynomial structure, for example xy. Therefore, reformulating the function is required in order to apply existing methods. However, this seems to be a less handy task.

Acknowledgment

The authors would like to thank the reviewer for pointing out a mistake in a previous manuscript of this paper.

References

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

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

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

[4] 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.

[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, 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] 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.

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