

A Classification of Functions in Multiterminal Distributed Computing

Shun Watanabe

Abstract

In the distributed function computation problem, dichotomy theorems, initiated by Han-Kobayashi, seek to classify functions by whether the rate regions for function computation improve on the Slepian-Wolf regions or not. In this paper, we develop a general approach to derive converse bounds on the distributed function computation problem. By using this approach, we recover the sufficiency part, i.e. the conditions such that the Slepian-Wolf regions become optimal, of the known dichotomy theorems in the two-terminal distributed computing. Furthermore, we derive an improved sufficient condition on the dichotomy theorem in the multiterminal distributed computing for the class of i.i.d. sources with positivity condition. Finally, we derive the matching sufficient and necessary condition on the dichotomy theorem in the multiterminal distributed computing for the class of smooth sources.

I. INTRODUCTION

The distributed function computing is one of the most basic but difficult problems in network information theory. In this problem, correlated sources are observed at L terminals, and separately encoded messages are sent to the decoder so that a function value of the sources can be computed at the decoder; see Fig. 1. A naive scheme to compute a function is to first reproduce the entire source at the decoder and then to compute the function value. Thus, it is apparent that the Slepian-Wolf (SW) region [28], [5] is an inner bound on the achievable rate region for the function computing problem. However, since the entire source needs not be reproduced at the decoder, the SW region can be improved in general. Then, we are interested in under what conditions the SW region can be improved. For instance, it is well known that, when the number of terminals is two and the function to be computed is the modulo-sum of the binary double symmetric source, the Körner-Martón (KM) coding improves upon the SW region [13].¹

A part of this paper is submitted to IEEE International Symposium on Information Theory 2018.

The work of S. Watanabe is supported in part by JSPS KAKENHI Grant Number 16H06091.

¹More precisely, it was shown in [13] that the structured coding improves upon the two-helper extension of the Wyner-Ahlsvede-Körner (WAK) region [33], [3]. In a similar manner as [10, Lemma 1 and Lemma 2], it can be verified that the two-helper extension of the WAK region coincides with the SW region for the modulo-sum function. It should be also noted that Ahlsvede-Han (AH) showed a hybrid use of the structured coding and the two-helper WAK scheme further improves upon the convex hull of the SW region and the KM region [2]. However, as is conjectured in [26], the minimum sum-rate of the AH region may not improve on the minimum sum-rate of the convex hull of the SW region and the KM region for the modulo-sum function.

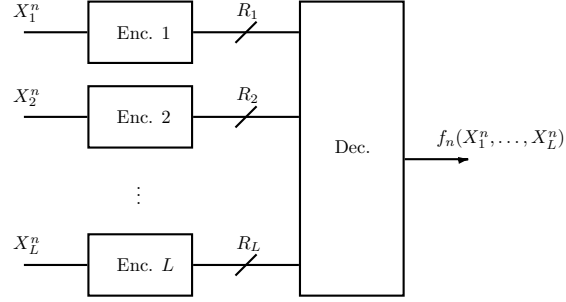


Fig. 1: A description of multiterminal distributed computing.

In [10], Han-Kobayashi initiated the study of classifying symbol-wise functions by whether the SW region can be improved or not.² When the number of terminals is two, they completely characterized the condition of classification; the derived condition, termed “dichotomy theorem”, only depends on the structure of functions in the following sense: if a given function satisfies the condition, then the achievable rate region of computing that function coincides with the SW region for any independent and identically distributed (i.i.d.) sources as long as the positivity condition is satisfied;³ on the other hand, if a given function does not satisfy the condition, then there exists an i.i.d. source satisfying positivity condition such that the SW region can be strictly improved. They also studied the classification problem for more than two terminals, and derived a partial solution to the problem; however, the complete characterization remained as an open problem.

It should be noted that, when a given function violates the Han-Kobayashi condition, the dichotomy theorem only claims existence of a source such that the SW region can be improved; it is more challenging problem to decide if the SW region can be improved or not for a given i.i.d. source with positivity condition. For instance, the modulo-sum function violates the Han-Kobayashi condition, and, as we mentioned above, the SW region can be improved for the binary double symmetric source. However, when the minimum of entropies of marginal distributions is smaller than the entropy of the modulo-sum, the optimal region for computing the modulo-sum coincides with the SW region [6, Problem 16.23(b)]. As long as the author know, it is an open problem to decide whether the optimal region for computing the modulo-sum coincides with the SW region or not for a given binary double source.

So far, the above mentioned results are for i.i.d. sources. In [16], Kuzuoka-Watanabe introduced a class of sources termed “smooth sources”. This class of sources includes i.i.d. sources with positivity condition as a special case; but it also includes sources with memory, such as Markov sources with positive transition matrices, and non-ergodic sources, such as mixtures of i.i.d. sources with positivity condition. Then, they considered the function classification problem for smooth sources and derived an alternative dichotomy theorem that is different from the one in [10]. More specifically, since the class of sources considered in [16] is broader than that considered in [10], the condition

²The symbol-wise function means that, for a given function on a single observation space, we compute the same copies of the function for a sequence of observations.

³The positivity condition is the condition such that all symbols occur with positive probability.

on functions in [16] is more strict than that in [10]. This difference stems from the fact that, even when a given function satisfies the Han-Kobayashi condition, the SW region may be improved for sources having memory, which signifies an importance of studying distributed function computing beyond i.i.d. sources.

There are three motivations in this paper:

- i) As we mentioned above, only a partial solution to the classification problem for more than two terminals was derived in [10]. Even for the simplest setting of three terminals with binary sources, there are some functions that cannot be classified by the conditions derived in [10]. In order to make a progress on this long-standing open problem, it is desirable to have a new approach to tackle this problem.
- ii) For the dichotomy theorems of two terminals, the sufficiency part⁴ of Han-Kobayashi [10] and Kuzuoka-Watanabe [16] were derived by completely different methods. The former is based on the single letter characterization; sources being i.i.d. is crucial, and it cannot be applied for general smooth sources. On the other hand, the latter is based on the argument introduced by El Gamal [7]; functions having certain structure called “total sensitivity” is crucial. The condition of total sensitivity is more restrictive than the Han-Kobayashi condition, and the proof method in [16] does not reproduce the dichotomy theorem of [10] when sources are restricted to i.i.d. sources. Thus, it is desirable to have an approach that provides both the dichotomy theorems in a unified manner; currently, when a given function satisfies the Han-Kobayashi condition but not total sensitivity, we cannot tell if the function computing region coincides with the SW region even for a simple example of smooth sources such as mixtures of i.i.d. sources.
- iii) For the two terminal distributed computing where the decoder has full side-information, a method to derive converse bounds was proposed in [17]. In that method, from the nature of distributed computing and the structure of the function to be computed, termed *informative structure*, we identify information that is inevitably conveyed to the decoder. Then, we derive a bound in terms of the optimal rate needed to send that information; when the inevitably conveyed information is the encoder’s source itself, we can conclude that the Slepian-Wolf rate is optimal. In [17], in addition to recovering some of known results in [1], [7], [24], the method was used to provide some novel results for the distributed computing with full side-information. This method has a potential to resolve problems raised in the above mentioned two motivations. However, the method crucially relies on the fact that the decoder has full side-information, and it does not apply to the case where all observations are encoded; in the conclusion of [17], the authors left such an extension as an important future research direction.

In this paper, we develop a general approach for showing dichotomy theorems, in particular, the sufficiency part. The proposed approach is build upon the proof method based on informative structure proposed in [17], but requires two new ideas described below.

As we mentioned above, the proof method based on informative structure relies on the fact that the decoder has full side-information, and it does not apply to the case where all observations are encoded. However, when the

⁴The sufficiency part of dichotomy theorems claim that there exist no codes that improve upon the SW region.

function induces conditional independence (CI) structure to the observed sources for some partition of the terminals, then we can virtually decouple the entire coding system into multiple coding systems where each decoder observe sources corresponding to parts of the partition as full side-information. Then, we can apply the proof method based on informative structure to each decoupled coding system.

It should be noted that the CI structure and its relaxation have been known as crucial for proving converse in multiterminal distributed computing or multiterminal rate-distortion problems [8], [4], [30], [20], [21], [25], [31], [32], [22], [23], [29], [27]. In fact, the proof of the dichotomy theorem by Han-Kobayashi [10] crucially uses the CI structure. More specifically, the Han-Kobayashi condition induces the CI structure for i.i.d. sources with positivity condition, which facilitate single-letter characterization. Gel'fand-Pinsker [8] also used the CI structure to derive single-letter characterization of the multiterminal source coding problem including the multiterminal distributed function computing problem as a special case. Focusing on the CI structure in this paper is inspired by ideas in these prior works. However, in our approach, the CI structure is used “operationally” to decouple the coding system into multiple parts of coding systems with full side-information, which facilitates using of the proof method based on informative structure. As long as the author know, such an operational usage of the CI structure in converse proof has never appeared in the literature, and is of independent interest.

Another new element is recursion. When we apply the above mentioned two ingredients to a given function, we can show that the function computation region of the given function is included in the function computation region of a “finer” function.⁵ If the finer function is tantamount to the identity function, then we can conclude that the function computation region of the original function coincides with the Slepian-Wolf region. However, even when the finer function is not the identity function, we can apply the same argument to that finer function to show that the function computation region of the finer function is included in the function computation region of a further finer function. By repeating this procedure recursively, we can eventually show that the function computation region of the original function coincides with the Slepian-Wolf region.

As an application of our general approach, when the number of terminals is two, we reproduce the dichotomy theorem of Han-Kobayashi [10] and that of Kuzuoka-Watanabe [16] in a unified manner. In addition to reproducing these results, our approach can provide converse results for cases that cannot be covered by neither of the methods in [10] nor [16], though we do not pursue the two terminals problem in depth in this paper; we will provide one of such examples, a mixture of i.i.d. sources, in Section V-C. We also apply our approach to derive dichotomy theorems for more than two terminals. In fact, for the the class of i.i.d. sources with positivity condition, we derive a sufficient condition that strictly subsume the sufficient condition shown in [10]. Furthermore, for the class of smooth sources, we establish the complete characterization of the classification problem. Note that the solution of classification problem for smooth sources does not imply the solution of classification problem for i.i.d. sources with positivity condition since the latter problem is not a special case of the former problem. As we mentioned above, even when the optimal function computing region coincides with the SW region for every i.i.d. sources with

⁵A function $f(x)$ is finer than function $g(x)$ when there exists a function h such that $g(x) = h(f(x))$.

positivity condition, the SW region may be improved for sources having memory.

Organization of Paper

The rest of the paper is organized as follows: In Section II, we introduce the problem formulation of multiterminal distributed computing. In Section III, we explain an overview of our approach, and in Section IV, we state and prove our general results. In Section V, we revisit known results on two-terminal distributed computing from the view point of our general approach. In Section VI and Section VII, we apply our approach to the multiterminal distributed computing for the class of i.i.d. sources with positivity condition and the class of smooth sources, respectively. We close the paper with some conclusion in Section VIII. Some of technical results and proofs are given in appendices.

Notations

Throughout this paper, random variables (e.g., X) and their realizations (e.g., x) are denoted by capital and lower case letters, respectively. All random variables take values in some finite alphabets which are denoted by the respective calligraphic letters (e.g., \mathcal{X}). The probability distribution of random variable X is denoted by P_X . Similarly, $X^n := (X_1, X_2, \dots, X_n)$ and $x^n := (x_1, x_2, \dots, x_n)$ denote, respectively, a random vector and its realization in the n th Cartesian product \mathcal{X}^n of \mathcal{X} . We will use bold lower letters to represent vectors if the length n is apparent from the context; e.g., we use \mathbf{x} instead of x^n .

For a finite set \mathcal{S} , the cardinality of \mathcal{S} is denoted by $|\mathcal{S}|$. For a subset $\mathcal{T} \subseteq \mathcal{S}$, the complement $\mathcal{S} \setminus \mathcal{T}$ is denoted by \mathcal{T}^c . For a given sequence \mathbf{s} in the n th Cartesian product \mathcal{S}^n of \mathcal{S} , the type $P_{\mathbf{s}} = (P_{\mathbf{s}}(s) : s \in \mathcal{S})$ of \mathbf{s} is defined by

$$P_{\mathbf{s}}(s) := \frac{|\{i \in [1 : n] : s_i = s\}|}{n}, \quad s \in \mathcal{S}, \quad (1)$$

where $[1 : n] := \{1, 2, \dots, n\}$. The set of all types of sequences in \mathcal{S}^n is denoted by $\mathcal{P}_n(\mathcal{S})$. The indicator function is denoted by $\mathbf{1}[\cdot]$. Information-theoretic quantities are denoted in the usual manner [5], [6]. The binary entropy is denoted by $h(\cdot)$. All logarithms are with respect to base 2. For given distributions P, Q , the variational distance is denoted by $\|P - Q\|_1 := \frac{1}{2} \sum_a |P(a) - Q(a)|$.

II. PROBLEM FORMULATION

Let

$$(\mathbf{X}_1, \dots, \mathbf{X}_L) = \{(X_1^n, \dots, X_L^n)\}_{n=1}^{\infty} \quad (2)$$

be a general correlated source with finite alphabet $\mathcal{X}_{\mathcal{L}} = \mathcal{X}_1 \times \dots \times \mathcal{X}_L$; the source is general in the sense of [11] (see also [9]), i.e., it may have memory and may not be stationary nor ergodic. Without loss of generality, we assume $\mathcal{X}_{\ell} = \{0, 1, \dots, |\mathcal{X}_{\ell}| - 1\}$ for $\ell = 1, \dots, L$. For a subset $\mathcal{A} \subseteq \mathcal{L}$ of the set $\mathcal{L} = \{1, \dots, L\}$ of all terminals, we denote $\mathbf{X}_{\mathcal{A}} = (\mathbf{X}_{\ell} : \ell \in \mathcal{A})$, $X_{\mathcal{A}}^n = (X_{\ell}^n : \ell \in \mathcal{A})$, and etc.

We consider a sequence $\mathbf{f} = \{f_n\}_{n=1}^\infty$ of functions $f_n : \mathcal{X}_\mathcal{L}^n \rightarrow \mathcal{Z}_n$. A code $\Phi_n = (\varphi_n^{(1)}, \dots, \varphi_n^{(L)}, \psi_n)$ for computing f_n is defined by encoders $\varphi_n^{(\ell)} : \mathcal{X}_\mathcal{L}^n \rightarrow \mathcal{M}_n^{(\ell)}$ for $\ell = 1, \dots, L$ and a decoder $\psi_n : \mathcal{M}_n^{(1)} \times \dots \times \mathcal{M}_n^{(L)} \rightarrow \mathcal{Z}_n$. The error probability of the code Φ_n is given by

$$P_e(\Phi_n | f_n) := \Pr \left(\psi_n(\varphi_n^{(\ell)}(X_\mathcal{L}^n) : \ell \in \mathcal{L}) \neq f_n(X_\mathcal{L}^n : \ell \in \mathcal{L}) \right). \quad (3)$$

Definition 1 For a given source $\mathbf{X}_\mathcal{L}$ and a sequence of functions \mathbf{f} , a rate tuple $R_\mathcal{L} = (R_\ell : \ell \in \mathcal{L})$ is defined to be achievable if there exists a sequence $\{\Phi_n\}_{n=1}^\infty$ of codes satisfying

$$\lim_{n \rightarrow \infty} P_e(\Phi_n | f_n) = 0 \quad (4)$$

and

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log |\mathcal{M}_n^{(\ell)}| \leq R_\ell, \quad \forall \ell \in \mathcal{L}. \quad (5)$$

The achievable rate region for computing \mathbf{f} , denoted by $\mathcal{R}(\mathbf{X}_\mathcal{L} | \mathbf{f})$, is the set of all achievable rate tuples.

One of the most simple classes of functions is the class of symbol-wise functions defined as follows; we will investigate this class intensively later in Sections V-VII.

Definition 2 (Symbol-wise Function) For a given function $f : \mathcal{X}_\mathcal{L} \rightarrow \mathcal{V}$, a function defined as $f_n(\mathbf{x}_\mathcal{L}) = (f(x_{\mathcal{L},1}), \dots, f(x_{\mathcal{L},n}))$ is called *symbol-wise* function induced by f . We sometimes denote f^n instead of f_n to emphasize that the symbol-wise function is n copies of f .

In order to compute functions, one approach is that all terminals send messages so that the decoder can reproduce the entire source, which is known as the Slepian-Wolf coding.

Definition 3 (SW Region) For a given source $\mathbf{X}_\mathcal{L}$, the achievable rate region $\mathcal{R}(\mathbf{X}_\mathcal{L} | \mathbf{f}^{\text{id}})$ for the sequence $\mathbf{f}^{\text{id}} = \{f_n^{\text{id}}\}_{n=1}^\infty$ of identity functions is called the Slepian-Wolf (SW) region, and denoted by $\mathcal{R}_{\text{SW}}(\mathbf{X}_\mathcal{L})$.

Note that $\mathcal{R}_{\text{SW}}(\mathbf{X}_\mathcal{L})$ is a trivial inner bound on $\mathcal{R}(\mathbf{X}_\mathcal{L} | \mathbf{f})$, i.e., it holds that $\mathcal{R}_{\text{SW}}(\mathbf{X}_\mathcal{L}) \subseteq \mathcal{R}(\mathbf{X}_\mathcal{L} | \mathbf{f})$ for any \mathbf{f} .

The Slepian-Wolf coding for general sources was first studied in [19] (see also [9, Section 7.2]). Even though only the two terminal case was studied in the literature, it is straightforward to extend the result to the multiterminal case; the SW region for general sources is characterized as follows.

Proposition 1 For a given general source $\mathbf{X}_\mathcal{L}$, it holds that

$$\mathcal{R}_{\text{SW}}(\mathbf{X}_\mathcal{L}) = \left\{ R_\mathcal{L} : \sum_{\ell \in \mathcal{A}} R_\ell \geq \overline{H}(\mathbf{X}_\mathcal{A} | \mathbf{X}_{\mathcal{A}^c}), \quad \forall \mathcal{A} \subseteq \mathcal{L} \right\}, \quad (6)$$

where

$$\overline{H}(\mathbf{X}_\mathcal{A} | \mathbf{X}_{\mathcal{A}^c}) := \inf \left\{ \alpha : \lim_{n \rightarrow \infty} \Pr \left\{ \frac{1}{n} \log \frac{1}{P_{X_\mathcal{A}^n | X_{\mathcal{A}^c}^n}}(X_\mathcal{A}^n | X_{\mathcal{A}^c}^n) > \alpha \right\} = 0 \right\} \quad (7)$$

is the spectral conditional sup-entropy rate of $\mathbf{X}_\mathcal{A}$ given $\mathbf{X}_{\mathcal{A}^c}$.

TABLE I: $f : \mathcal{X}_1 \times \mathcal{X}_2 \rightarrow \mathcal{V}$

$x_1 \setminus x_2$	0	1	2
0	0	3	3
1	0	4	2
2	1	1	2

When the source is i.i.d., the characterization in Proposition 1 reduces to the well known formula by replacing $\overline{H}(\mathbf{X}_{\mathcal{A}}|\mathbf{X}_{\mathcal{A}^c})$ with the conditional entropy $H(X_{\mathcal{A}}|X_{\mathcal{A}^c})$ in (6).

In this paper, we shall develop an approach to characterize the function computation region $\mathcal{R}(\mathbf{X}_{\mathcal{L}}|\mathbf{f})$ for general sources. The function computation region can be sensitive to the support of distributions (cf. the comment at the end of [10, Sec. III]). To avoid such a complication, we consider the following class of smooth sources introduced in [16].

Definition 4 (Smooth Source) A general source $\mathbf{X}_{\mathcal{L}}$ is said to be *smooth* if there exists a constant $0 < q < 1$, which does not depend on n , satisfying

$$P_{X_{\mathcal{L}}^n}(\mathbf{x}_{\mathcal{L}}) \geq q P_{X_{\mathcal{L}}^n}(\hat{\mathbf{x}}_{\mathcal{L}}) \quad (8)$$

for every $\mathbf{x}_{\mathcal{L}}, \hat{\mathbf{x}}_{\mathcal{L}} \in \mathcal{X}_{\mathcal{L}}^n$ with $d_H(\mathbf{x}_{\mathcal{L}}, \hat{\mathbf{x}}_{\mathcal{L}}) = 1$, where $d_H(\cdot, \cdot)$ is the Hamming distance.

The class of smooth sources is a natural generalization of i.i.d. sources with positivity condition studied in [10], [1], and enables us to study distributed computation for a variety of sources in a unified manner. Indeed, this class contains sources with memory, such as Markov sources with positive transition matrices, or non-ergodic sources, such as mixtures of i.i.d. sources with positivity condition; see [16] for the detail.

III. AN OVERVIEW OF THE APPROACH

In this section, we shall explain an overview of our approach for the special case of the two terminals and symbol-wise functions. In our approach, we use three ingredients: the proof method based on informative structure [17], the conditional independence structure, and the recursion. Since the third one, recursion, is not needed in the two terminals case, we only explain the first two ingredients here. For concreteness, let us consider symbol-wise function given by Table I. Since this function satisfies the Han-Kobayashi condition [10] (see also Section V), the function computation region $\mathcal{R}(\mathbf{X}_{\mathcal{L}}|\mathbf{f})$ for this symbol-wise function coincides with the Slepian-Wolf region $\mathcal{R}_{\text{SW}}(\mathbf{X}_{\mathcal{L}})$ for any i.i.d. sources with positivity condition. As an overview of our approach, we shall provide an alternative proof of this known fact.

Let us first consider the special case where X_2^n is observed at the decoder as full side-information (see Fig. 2). For this special case, the following approach to prove $R_1 \geq H(X_1|X_2)$ was proposed in [17]. Let $\Phi_n = (\varphi_n^{(1)}, \psi_n)$ be a code to compute f_n with vanishing error probability ε_n . Since message $\varphi_n^{(1)}(X_1^n)$ is encoded without knowing the realization of side-information X_2^n , if we input $\pi_i^b(X_2^n)$ to $\psi_n(\varphi_n^{(1)}(X_1^n), \cdot)$ instead of X_2^n , then we expect it

will outputs $f_n(X_1^n, \pi_i^b(X_2^n))$ with high probability, where π_i^b shifts i th symbol of X_2^n by b , i.e., $X_{2,i} \mapsto X_{2,i} + b \pmod{|\mathcal{X}_2|}$. In fact, by noting the smoothness of the source, i.e., the fact that the probability does not change drastically when one symbol is shifted, we can show

$$\Pr\left(\psi_n(\varphi_n^{(1)}(X_1^n), \pi_i^b(X_2^n)) \neq f_n(X_1^n, \pi_i^b(X_2^n))\right) \leq \frac{\varepsilon_n}{q}. \quad (9)$$

This implies that, via union bound, the decoder can reproduce the list $(f(X_{1,i}, x_2) : x_2 \in \mathcal{X}_2)$ with vanishing error probability. Here, note that every rows in Table I are different. Thus, the decoder can distinguish $X_{1,i}$ from the list. By doing this procedure for $1 \leq i \leq n$, the decoder can reproduce the entire source X_1^n with vanishing error probability.⁶ This means that the rate R_1 must be as large as the Slepian-Wolf rate $H(X_1|X_2)$. A similar argument can be applied to prove $R_2 \geq H(X_2|X_1)$ since every columns in Table I are different. This is the first ingredient, informative structure, i.e., the function has a structure that reveals information about $X_{\ell,i}$.

In order to consider the entire region $\mathcal{R}(\mathbf{X}_{\mathcal{L}}|\mathbf{f})$, we need the second ingredient, conditional independence structure. In the above argument to bound individual rates, a crucial step is to input $\pi_i^b(X_2^n)$ to $\psi_n(\varphi_n^{(1)}(X_1^n), \cdot)$; this is possible because X_2^n is directly observed at the decoder in the system of Fig. 2. When we consider the entire region, the same argument cannot be used since the decoder only gets $\varphi_n^{(1)}(X_1^n)$ and $\varphi_n^{(2)}(X_2^n)$. In order to circumvent this problem, we note that the function in Table I induces conditional independence, i.e., the Markov chain $X_1 - \circ - f(X_1, X_2) - \circ - X_2$ holds.⁷ In fact, it can be verified from the fact that either x_1 or x_2 is determined from $f(x_1, x_2)$. Because of this observation, once the decoder can reproduce $V^n = f_n(X_1^n, X_2^n)$ (with vanishing error probability), the decoder can generate \tilde{X}_1^n and \tilde{X}_2^n via channels $P_{X_1|V}^n$ and $P_{X_2|V}^n$ so that (X_1^n, \tilde{X}_2^n) and (\tilde{X}_1^n, X_2^n) have (approximately) the same joint distribution as (X_1^n, X_2^n) . By this argument, we can decouple the entire system (see Fig. 3) into two virtual systems (see Fig. 4). Then, we can use the previous argument of the full side-information case, and show that the decoder can reproduce the entire sources X_1^n and X_2^n with vanishing error probability. This means that the function computation region $\mathcal{R}(\mathbf{X}_{\mathcal{L}}|\mathbf{f})$ must be included in the Slepian-Wolf region $\mathcal{R}_{\text{SW}}(\mathbf{X}_{\mathcal{L}})$.

In the above argument, the function being symbol-wise and the source being i.i.d. are not necessary. In fact, the argument goes thorough as long as the function reveals (partial) information about $X_{\ell,i}$ and the function induces conditional independence for the given source. In the next section, we will introduce an approach to derive converse bounds on the multiterminal distributed computing by generalizing the above argument.

IV. GENERAL RESULT

For $\ell \in \mathcal{L}$, let $\overline{\mathcal{X}}_{\ell}$ be a partition of \mathcal{X}_{ℓ} ; i.e., $\overline{\mathcal{X}}_{\ell} = \{\mathcal{C}_1, \dots, \mathcal{C}_t\}$ is a set of nonempty subsets $\mathcal{C}_i \subseteq \mathcal{X}_{\ell}$ ($i = 1, \dots, t$) satisfying $\mathcal{C}_i \cap \mathcal{C}_j = \emptyset$ ($i \neq j$) and $\mathcal{X}_{\ell} = \cup_{\mathcal{C} \in \overline{\mathcal{X}}_{\ell}} \mathcal{C}$. We say that partition $\overline{\mathcal{X}}_{\ell}$ is nontrivial if $t \geq 2$. For each $x_{\ell} \in \mathcal{X}_{\ell}$, the subset $\mathcal{C} \in \overline{\mathcal{X}}_{\ell}$ satisfying $x_{\ell} \in \mathcal{C}$ is uniquely determined and denoted by $[x_{\ell}]_{\overline{\mathcal{X}}_{\ell}}$. For a sequence

⁶For simplicity, we have omitted one step in which we boost the small symbol error probability to the small block error probability (see Lemma 3).

⁷Note that this CI structure is also used in the proof of [10] to facilitate single letter characterization.

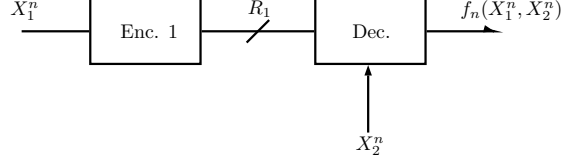


Fig. 2: A description of distributed computing with full side-information.

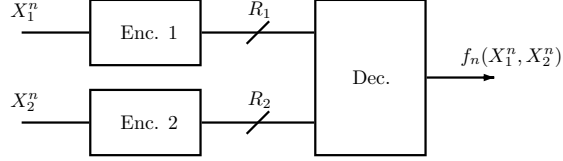


Fig. 3: A description of two-terminal distributed computing.

$\mathbf{x}_\ell = (x_{\ell,1}, \dots, x_{\ell,n}) \in \mathcal{X}_\ell^n$, we apply $[\cdot]_{\overline{\mathcal{X}}_\ell}$ to each symbol, i.e., $[\mathbf{x}_\ell]_{\overline{\mathcal{X}}_\ell} = ([x_{\ell,1}]_{\overline{\mathcal{X}}_\ell}, \dots, [x_{\ell,n}]_{\overline{\mathcal{X}}_\ell}) \in \overline{\mathcal{X}}_\ell^n$. For a subset $\mathcal{A} \subseteq \mathcal{L}$ and a tuple $\overline{\mathcal{X}}_{\mathcal{A}} = (\overline{\mathcal{X}}_\ell : \ell \in \mathcal{A})$ of partitions, we denote $[\mathbf{x}_{\mathcal{A}}]_{\overline{\mathcal{X}}_{\mathcal{A}}} = ([x_\ell]_{\overline{\mathcal{X}}_\ell} : \ell \in \mathcal{A})$ and $[\mathbf{x}_{\mathcal{A}}]_{\overline{\mathcal{X}}_{\mathcal{A}}} = ([\mathbf{x}_\ell]_{\overline{\mathcal{X}}_\ell} : \ell \in \mathcal{A})$.

For a subset $\mathcal{A} \subseteq \mathcal{L}$, a symbol $a_{\mathcal{A}} \in \mathcal{X}_{\mathcal{A}}$, a sequence $\mathbf{x}_{\mathcal{A}} \in \mathcal{X}_{\mathcal{A}}^n$, and an index $i \in [1 : n]$, let $a_{\mathcal{A}} \mathbf{x}_{\mathcal{A}}^{(-i)}$ be the sequence such that elements $x_{\mathcal{A},i}$ of $\mathbf{x}_{\mathcal{A}}$ are replaced by $a_{\mathcal{A}}$.

In the following, we introduce a few structures of functions; we say that a sequence $\mathbf{f} = \{f_n\}_{n=1}^\infty$ of functions has certain structure if f_n has that structure for every n .

Definition 5 (Local Function) For a given tuple of partitions $\overline{\mathcal{X}}_{\mathcal{L}} = (\overline{\mathcal{X}}_\ell : \ell \in \mathcal{L})$, the symbol-wise function $f_{n, \overline{\mathcal{X}}_{\mathcal{L}}} : \mathcal{X}_{\mathcal{L}}^n \rightarrow \prod_{\ell=1}^L \overline{\mathcal{X}}_\ell^n$ defined by

$$f_{n, \overline{\mathcal{X}}_{\mathcal{L}}}(\mathbf{x}_\ell : \ell \in \mathcal{L}) := ([\mathbf{x}_\ell]_{\overline{\mathcal{X}}_\ell} : \ell \in \mathcal{L}). \quad (10)$$

is called *local function*.

Since a local function is a symbol-wise function, we sometimes just say local function $f_{\overline{\mathcal{X}}_{\mathcal{L}}} : \mathcal{X}_{\mathcal{L}} \rightarrow \prod_{\ell=1}^L \overline{\mathcal{X}}_\ell$. As is clear from the definition, ℓ th component of local functions can be locally computable at ℓ th terminal. For a sequence of local function $\mathbf{f}_{\overline{\mathcal{X}}_{\mathcal{L}}} = \{f_{n, \overline{\mathcal{X}}_{\mathcal{L}}}\}_{n=1}^\infty$, the achievable rate region $\mathcal{R}(\mathbf{X}_{\mathcal{L}} | \mathbf{f}_{\overline{\mathcal{X}}_{\mathcal{L}}})$ plays an important role. Note that $\mathcal{R}(\mathbf{X}_{\mathcal{L}} | \mathbf{f}_{\overline{\mathcal{X}}_{\mathcal{L}}}) = \mathcal{R}_{\text{SW}}(\mathbf{X}_{\mathcal{L}})$ when all the partitions are the finest partitions, i.e., $\overline{\mathcal{X}}_\ell = \{\{x_\ell\} : x_\ell \in \mathcal{X}_\ell\}$ for every $\ell \in \mathcal{L}$. In the rest of the paper, we write $\overline{\mathcal{X}}_\ell \equiv \mathcal{X}_\ell$ when $\overline{\mathcal{X}}_\ell$ is the finest partition.

The following class of functions is a multiterminal generalization of the function class introduced in [17]; for the motivation of the definition, see [17, Sections III and IV].

Definition 6 (Informative Function) For a subset $\mathcal{A} \subseteq \mathcal{L}$, let $\overline{\mathcal{X}}_{\mathcal{A}} = (\overline{\mathcal{X}}_\ell : \ell \in \mathcal{A})$ be a tuple of partitions. A function f_n is said to be $\overline{\mathcal{X}}_{\mathcal{A}}$ -informative if f_n satisfies the following conditions:

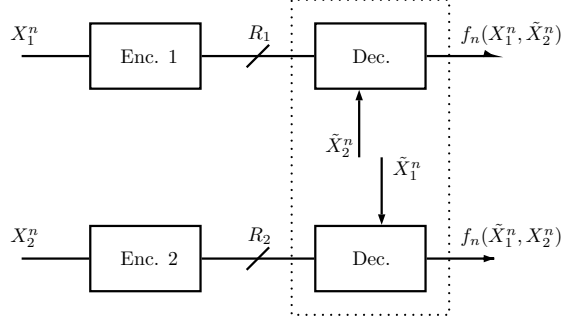


Fig. 4: A description of two virtual systems decoupled from the system in Fig. 3.

- 1) For each $i \in [1 : n]$, there exists a mapping $\xi_{n,\mathcal{A}}^{(i)} : \mathcal{Z}_n^{|\mathcal{X}_{\mathcal{A}^c}|} \rightarrow \overline{\mathcal{X}}_{\mathcal{A}}$ such that,⁸ for any $a_{\mathcal{A}} = (a_{\ell} : \ell \in \mathcal{A}) \in \mathcal{X}_{\mathcal{A}}$ and $\mathbf{x}_{\mathcal{L}} \in \mathcal{X}_{\mathcal{L}}^n$,

$$\xi_{n,\mathcal{A}}^{(i)} \left(\left(f_n(a_{\mathcal{A}} \mathbf{x}_{\mathcal{A}}^{(-i)}, a_{\mathcal{A}^c} \mathbf{x}_{\mathcal{A}^c}^{(-i)}) : a_{\mathcal{A}^c} \in \mathcal{X}_{\mathcal{A}^c} \right) \right) = [a_{\mathcal{A}}]_{\overline{\mathcal{X}}_{\mathcal{A}}} = ([a_{\ell}]_{\overline{\mathcal{X}}_{\ell}} : \ell \in \mathcal{A}). \quad (11)$$

- 2) For every $\mathbf{x}_{\mathcal{L}} \in \mathcal{X}_{\mathcal{L}}^n$, every $\ell \in \mathcal{A}$, and any permutation σ on $[1 : n]$ satisfying $[\sigma(\mathbf{x}_{\ell})]_{\overline{\mathcal{X}}_{\ell}} = [\mathbf{x}_{\ell}]_{\overline{\mathcal{X}}_{\ell}}$,

$$f_n(\mathbf{x}_1, \dots, \sigma(\mathbf{x}_{\ell}), \dots, \mathbf{x}_L) = f_n(\mathbf{x}_1, \dots, \mathbf{x}_{\ell}, \dots, \mathbf{x}_L). \quad (12)$$

On the other hand, when only Condition 1 is satisfied, f_n is said to be *semi $\overline{\mathcal{X}}_{\mathcal{A}}$ -informative*.

Remark 1 In contrast to the two terminal case in [17], when we say “ $\overline{\mathcal{X}}_{\mathcal{A}}$ -informative”, it also designates the coordinates \mathcal{A}^c we can vary in (11).

Remark 2 Condition (1) means that, for each $i \in [1 : n]$ and for any $a_{\mathcal{A}} = (a_{\ell} : \ell \in \mathcal{A}) \in \mathcal{X}_{\mathcal{A}}$, the tuple of equivalence classes $(\mathcal{K}_{\ell} : \ell \in \mathcal{A})$ satisfying $a_{\ell} \in \mathcal{K}_{\ell}$ for each $\ell \in \mathcal{A}$ can be uniquely determined from the list

$$\left(f_n(a_{\mathcal{A}} \mathbf{x}_{\mathcal{A}}^{(-i)}, a_{\mathcal{A}^c} \mathbf{x}_{\mathcal{A}^c}^{(-i)}) : a_{\mathcal{A}^c} \in \mathcal{X}_{\mathcal{A}^c} \right) \quad (13)$$

irrespective of $\mathbf{x}_{\mathcal{L}} \in \mathcal{X}_{\mathcal{L}}^n$. From the definition, it is apparent that, if a function is $\overline{\mathcal{X}}_{\mathcal{A}}$ -informative, then it is $\overline{\mathcal{X}}_{\mathcal{A}'}$ -informative for every $\mathcal{A}' \subseteq \mathcal{A}$.

For two terminals setting, three examples of informative functions are given in [17, Section IV]; the first one is symbol-wise functions, and the latter two are not necessarily symbol-wise functions. For later convenience, we would like to review the first one here; the multiterminal generalization will be studied in Sections VI and VII.

For a given function $f : \mathcal{X}_1 \times \mathcal{X}_2 \rightarrow \mathcal{V}$, let $\overline{\mathcal{X}}_{1,f}$ be the partition of \mathcal{X}_1 such that two symbols x_1 and \hat{x}_1 are in the same partition of \mathcal{X}_1 if and only if $f(x_1, x_2) = f(\hat{x}_1, x_2)$ for every $x_2 \in \mathcal{X}_2$. Similarly, we define partition $\overline{\mathcal{X}}_{2,f}$ of \mathcal{X}_2 by interchanging the role of the two terminals. It is known that the equivalence class given by the partition $\overline{\mathcal{X}}_{i,f}$ plays an important role to characterize the optimal rate for computing the function; it represents the

⁸By a slight abuse of notation, we also use notation $\overline{\mathcal{X}}_{\mathcal{A}}$ to describe $\times_{\ell \in \mathcal{A}} \overline{\mathcal{X}}_{\ell}$ though the same notation is used for a tuple of partitions.

TABLE II: $f : \mathcal{X}_1 \times \mathcal{X}_2 \rightarrow \mathcal{V}$

$x_1 \setminus x_2$	0	1	2
0	0	1	2
1	3	4	2
2	3	4	2

equivalence class of the encoder's observation that is learnable by the decoder when the function is computed [24], [14]. For the function given by Table II, the partitions are given by $\overline{\mathcal{X}}_{1,f} = \{\{0\}, \{1, 2\}\}$ and $\overline{\mathcal{X}}_{2,f} \equiv \mathcal{X}_2$. For two terminal symbol-wise functions, we have the following characterization.

Proposition 2 ([17]) A symbol-wise function $f_n : \mathcal{X}_1^n \times \mathcal{X}_2^n \rightarrow \mathcal{V}^n$ defined from $f : \mathcal{X}_1 \times \mathcal{X}_2 \rightarrow \mathcal{V}$ is $\overline{\mathcal{X}}_{1,f}$ -informative and $\overline{\mathcal{X}}_{2,f}$ -informative, respectively.

Next, we shall introduce the class of functions that induce conditional independence.

Definition 7 (Conditional Independence) For a given source $\mathbf{X}_{\mathcal{L}}$ and a nontrivial partition $\overline{\mathcal{L}}$ of \mathcal{L} , we say that $\mathbf{f} = \{f_n\}_{n=1}^\infty$ induces (approximate) conditional independence for $(\mathbf{X}_{\mathcal{L}}, \overline{\mathcal{L}})$ if there exists a sequence of random variables $\mathbf{S} = \{S_n\}_{n=1}^\infty$ such that, for every $n \geq 1$,

$$X_{\mathcal{A}}^n \text{---} S_n \text{---} X_{\mathcal{A}^c}^n \quad (14)$$

for every $\mathcal{A} \in \overline{\mathcal{L}}$ and

$$\Pr(\gamma_n(Z_n) \neq S_n) \leq \mu_n \quad (15)$$

for some function $\gamma_n : \mathcal{Z}_n \rightarrow \mathcal{S}_n$ with $\mu_n \rightarrow 0$ as $n \rightarrow \infty$.

Note that whether conditional independence is induced or not depend on both functions and sources in general. When the source $\mathbf{X}_{\mathcal{L}}$ is independent across a partition $\overline{\mathcal{L}}$, then any function induces conditional independence for $(\mathbf{X}_{\mathcal{L}}, \overline{\mathcal{L}})$ by taking \mathbf{S} to be a constant. In latter sections, we will show some conditions of functions such that those functions induce conditional independence for $(\mathbf{X}_{\mathcal{L}}, \overline{\mathcal{L}})$ for every sources in the class of i.i.d. sources or the class of smooth sources.

Finally, we shall introduce product functions.

Definition 8 (Product Function) For given functions $\mathbf{f}_1 = \{f_{1,n}\}_{n=1}^\infty$ and $\mathbf{f}_2 = \{f_{2,n}\}_{n=1}^\infty$, let us define the product function as follows:

$$(f_{1,n}, f_{2,n})(\mathbf{x}_{\mathcal{L}}) := (f_{1,n}(\mathbf{x}_{\mathcal{L}}), f_{2,n}(\mathbf{x}_{\mathcal{L}})). \quad (16)$$

For the product function $(\mathbf{f}_1, \mathbf{f}_2) = \{(f_{1,n}, f_{2,n})\}_{n=1}^\infty$, the achievable rate region is denoted by $\mathcal{R}(\mathbf{X}_{\mathcal{L}} | \mathbf{f}_1, \mathbf{f}_2)$. The product of multiple functions is defined similarly.

Now, we are ready to state our main results.

Theorem 1 For a given smooth source $\mathbf{X}_{\mathcal{L}}$ and a nontrivial partition $\overline{\mathcal{L}}$ of \mathcal{L} , suppose that $\mathbf{f} = \{f_n\}_{n=1}^{\infty}$ induces conditional independence for $(\mathbf{X}_{\mathcal{L}}, \overline{\mathcal{L}})$. Furthermore, for a given tuple $\overline{\mathcal{X}}_{\mathcal{L}} = (\overline{\mathcal{X}}_{\ell} : \ell \in \mathcal{L})$ of partitions, suppose that \mathbf{f} is $\overline{\mathcal{X}}_{\mathcal{A}}$ -informative for every $\mathcal{A} \in \overline{\mathcal{L}}$, where $\overline{\mathcal{X}}_{\mathcal{A}} = (\overline{\mathcal{X}}_{\ell} : \ell \in \mathcal{A})$. Then, we have

$$\mathcal{R}(\mathbf{X}_{\mathcal{L}}|\mathbf{f}) = \mathcal{R}(\mathbf{X}_{\mathcal{L}}|\mathbf{f}_{\overline{\mathcal{X}}_{\mathcal{L}}}), \quad (17)$$

where $\mathbf{f}_{\overline{\mathcal{X}}_{\mathcal{L}}}$ is the local function defined by (10) for the tuple $\overline{\mathcal{X}}_{\mathcal{L}}$. On the other hand, if \mathbf{f} is semi $\overline{\mathcal{X}}_{\mathcal{A}}$ -informative for every $\mathcal{A} \in \overline{\mathcal{L}}$, then we have

$$\mathcal{R}(\mathbf{X}_{\mathcal{L}}|\mathbf{f}) = \mathcal{R}(\mathbf{X}_{\mathcal{L}}|\mathbf{f}, \mathbf{f}_{\overline{\mathcal{X}}_{\mathcal{L}}}). \quad (18)$$

We can apply the latter claim (18) of Theorem 1 recursively to obtain the following corollaries.

Corollary 1 (Recursion) Let $\mathbf{X}_{\mathcal{L}}$ be a smooth source. For a given sequence of nontrivial partitions $\{\overline{\mathcal{L}}^{(i)}\}_{i=1}^k$ of \mathcal{L} and a given sequence of tuples of partitions $\{\overline{\mathcal{X}}_{\mathcal{L}}^{(i)} = (\overline{\mathcal{X}}_{\ell}^{(i)} : \ell \in \mathcal{L})\}_{i=1}^k$, suppose that, for every $i = 1, \dots, k$, $(\mathbf{f}, \mathbf{f}_{\overline{\mathcal{X}}_{\mathcal{L}}^{(1)}}, \dots, \mathbf{f}_{\overline{\mathcal{X}}_{\mathcal{L}}^{(i-1)}})$ induces⁹ conditional independence for $(\mathbf{X}_{\mathcal{L}}, \overline{\mathcal{L}}^{(i)})$ and semi $\overline{\mathcal{X}}_{\mathcal{A}}^{(i)}$ -informative for every $\mathcal{A} \in \overline{\mathcal{L}}^{(i)}$. Then, we have

$$\mathcal{R}(\mathbf{X}_{\mathcal{L}}|\mathbf{f}) = \mathcal{R}(\mathbf{X}_{\mathcal{L}}|\mathbf{f}, \mathbf{f}_{\overline{\mathcal{X}}_{\mathcal{L}}^{(1)}}, \dots, \mathbf{f}_{\overline{\mathcal{X}}_{\mathcal{L}}^{(k)}}). \quad (19)$$

Corollary 2 Under the same condition as Corollary 1, if $\overline{\mathcal{X}}_{\ell}^{(k)} \equiv \mathcal{X}_{\ell}$ for every $\ell \in \mathcal{L}$, then we have

$$\mathcal{R}(\mathbf{X}_{\mathcal{L}}|\mathbf{f}) = \mathcal{R}_{\text{SW}}(\mathbf{X}_{\mathcal{L}}). \quad (20)$$

A. Proof of Converse Part of Theorem 1

We first prove \subseteq part of (17). Let $\{\Phi_n = (\varphi_n^{(1)}, \dots, \varphi_n^{(L)}, \psi_n)\}_{n=1}^{\infty}$ be a code for computing \mathbf{f} , and suppose that it achieves a rate tuple $R_{\mathcal{L}} = (R_{\ell} : \ell \in \mathcal{L}) \in \mathcal{R}(\mathbf{X}_{\mathcal{L}}|\mathbf{f})$. From this code, we shall construct a modified code for computing local function $\mathbf{f}_{\overline{\mathcal{X}}_{\mathcal{L}}}$. By letting $Z_n = f_n(X_{\ell}^n : \ell \in \mathcal{L})$ and $\hat{Z}_n = \psi_n(\varphi_n^{(\ell)}(X_{\ell}^n) : \ell \in \mathcal{L})$, we have

$$\Pr(Z_n \neq \hat{Z}_n) \leq \varepsilon_n \quad (21)$$

for some $\varepsilon_n \rightarrow 0$ as $n \rightarrow \infty$. Since \mathbf{f} induces conditional independence for $(\mathbf{X}_{\mathcal{L}}, \overline{\mathcal{L}})$, we can find $\mathbf{S} = \{S_n\}_{n=1}^{\infty}$ such that

$$P_{X_{\mathcal{L}}^n S_n} = P_{S_n} \prod_{\mathcal{A} \in \overline{\mathcal{L}}} P_{X_{\mathcal{A}}^n | S_n}. \quad (22)$$

Also, there exists a mapping γ_n satisfying (15). Furthermore, since \mathbf{f} is $\overline{\mathcal{X}}_{\mathcal{A}}$ -informative for every $\mathcal{A} \in \overline{\mathcal{L}}$, there exists function $\xi_{n,\mathcal{A}}^{(i)} : \mathcal{Z}_n^{|\mathcal{X}_{\mathcal{A}^c}|} \rightarrow \overline{\mathcal{X}}_{\mathcal{A}}$ satisfying (11) for each $\mathcal{A} \in \overline{\mathcal{L}}$ and $i \in [1 : n]$.

⁹For $i = 1$, $(\mathbf{f}, \mathbf{f}_{\overline{\mathcal{X}}_{\mathcal{L}}^{(1)}}, \dots, \mathbf{f}_{\overline{\mathcal{X}}_{\mathcal{L}}^{(i-1)}})$ means \mathbf{f} .

Upon receiving messages $m_{\mathcal{L}} = (m_{\ell} : \ell \in \mathcal{L})$ from the encoders, the modified decoder first computes $\hat{s}_n = \gamma_n(\psi_n(m_{\mathcal{L}}))$, and, for each $\mathcal{A} \in \bar{\mathcal{L}}$, locally generate $\tilde{X}_{\mathcal{A}}^n \sim P_{X_{\mathcal{A}}^n|S_n}(\cdot|\hat{s}_n)$. Then, for each $\mathcal{A} \in \bar{\mathcal{L}}$, the decoder computes

$$\xi_{n,\mathcal{A}}^{(i)} \left(\left(\psi_n(m_{\mathcal{A}}, \varphi_n^{(\mathcal{A}^c)}(a_{\mathcal{A}^c} \tilde{X}_{\mathcal{A}^c}^{(-i)})) : a_{\mathcal{A}^c} \in \mathcal{X}_{\mathcal{A}^c} \right) \right), \quad (23)$$

where $m_{\mathcal{A}} = (m_{\ell} : \ell \in \mathcal{A})$, $\varphi_n^{(\mathcal{A}^c)}(a_{\mathcal{A}^c} \tilde{X}_{\mathcal{A}^c}^{(-i)}) = (\varphi_n^{(\ell)}(a_{\ell} \tilde{X}_{\ell}^{(-i)}) : \ell \in \mathcal{A}^c)$, and $a_{\ell} \tilde{X}_{\ell}^{(-i)}$ is the sequence obtained by replacing i th coordinate $\tilde{X}_{\ell,i}$ of \tilde{X}_{ℓ}^n with a_{ℓ} . We shall show that, for each $\mathcal{A} \in \bar{\mathcal{L}}$,

$$W_{\mathcal{A},i} = \xi_{n,\mathcal{A}}^{(i)} \left(\left(\psi_n(\varphi_n^{(\mathcal{A})}(X_{\mathcal{A}}^n), \varphi_n^{(\mathcal{A}^c)}(a_{\mathcal{A}^c} \tilde{X}_{\mathcal{A}^c}^{(-i)})) : a_{\mathcal{A}^c} \in \mathcal{X}_{\mathcal{A}^c} \right) \right) \quad (24)$$

coincides with $[X_{\mathcal{A},i}]_{\bar{\mathcal{X}}_{\mathcal{A}}}$ with high probability.

By applying Lemma 1, (21) and (15) imply

$$\|P_{X_{\mathcal{A}}^n S_n} - P_{X_{\mathcal{A}}^n \hat{S}_n}\|_1 \leq \Pr(S_n \neq \hat{S}_n) \quad (25)$$

$$\leq \Pr(Z_n \neq \hat{Z}_n \text{ or } \gamma_n(Z_n) \neq S_n) \quad (26)$$

$$\leq \varepsilon_n + \mu_n, \quad (27)$$

where $\hat{S}_n = \gamma_n(\hat{Z}_n)$. Then, since taking marginal does not increase the variational distance, we have

$$\|P_{X_{\mathcal{A}}^n X_{\mathcal{A}^c}^n} - P_{X_{\mathcal{A}}^n \tilde{X}_{\mathcal{A}^c}^n}\|_1 \leq \|P_{X_{\mathcal{A}}^n X_{\mathcal{A}^c}^n S_n} - P_{X_{\mathcal{A}}^n \tilde{X}_{\mathcal{A}^c}^n \hat{S}_n}\|_1 \quad (28)$$

$$= \left\| P_{X_{\mathcal{A}}^n S_n} \prod_{\mathcal{B} \in \bar{\mathcal{L}} \setminus \{\mathcal{A}\}} P_{X_{\mathcal{B}}^n | S_n} - P_{X_{\mathcal{A}}^n \hat{S}_n} \prod_{\mathcal{B} \in \bar{\mathcal{L}} \setminus \{\mathcal{A}\}} P_{X_{\mathcal{B}}^n | S_n} \right\|_1 \quad (29)$$

$$= \|P_{X_{\mathcal{A}}^n S_n} - P_{X_{\mathcal{A}}^n \hat{S}_n}\|_1 \quad (30)$$

$$\leq \varepsilon_n + \mu_n. \quad (31)$$

By noting (11), if

$$\left(\psi_n(\varphi_n^{(\mathcal{A})}(X_{\mathcal{A}}^n), \varphi_n^{(\mathcal{A}^c)}(a_{\mathcal{A}^c} \tilde{X}_{\mathcal{A}^c}^{(-i)})) : a_{\mathcal{A}^c} \in \mathcal{X}_{\mathcal{A}^c} \right) \quad (32)$$

coincides with

$$\left(f_n(X_{\mathcal{A}}^n, a_{\mathcal{A}^c} \tilde{X}_{\mathcal{A}^c}^{(-i)}) : a_{\mathcal{A}^c} \in \mathcal{X}_{\mathcal{A}^c} \right), \quad (33)$$

then $W_{\mathcal{A},i}$ coincides with $[X_{\mathcal{A},i}]_{\bar{\mathcal{X}}_{\mathcal{A}}}$. Thus, we have

$$\Pr(W_{\mathcal{A},i} \neq [X_{\mathcal{A},i}]_{\bar{\mathcal{X}}_{\mathcal{A}}}) \quad (34)$$

$$\leq \Pr \left(\left(\psi_n(\varphi_n^{(\mathcal{A})}(X_{\mathcal{A}}^n), \varphi_n^{(\mathcal{A}^c)}(a_{\mathcal{A}^c} \tilde{X}_{\mathcal{A}^c}^{(-i)})) : a_{\mathcal{A}^c} \in \mathcal{X}_{\mathcal{A}^c} \right) \neq \left(f_n(X_{\mathcal{A}}^n, a_{\mathcal{A}^c} \tilde{X}_{\mathcal{A}^c}^{(-i)}) : a_{\mathcal{A}^c} \in \mathcal{X}_{\mathcal{A}^c} \right) \right) \quad (35)$$

$$\leq \Pr \left(\left(\psi_n(\varphi_n^{(\mathcal{A})}(X_{\mathcal{A}}^n), \varphi_n^{(\mathcal{A}^c)}(a_{\mathcal{A}^c} \tilde{X}_{\mathcal{A}^c}^{(-i)})) : a_{\mathcal{A}^c} \in \mathcal{X}_{\mathcal{A}^c} \right) \neq \left(f_n(X_{\mathcal{A}}^n, a_{\mathcal{A}^c} \tilde{X}_{\mathcal{A}^c}^{(-i)}) : a_{\mathcal{A}^c} \in \mathcal{X}_{\mathcal{A}^c} \right) \right) + \varepsilon_n + \mu_n, \quad (36)$$

where the second inequality follows from Lemma 2 and (31). Furthermore, by using the smoothness of the source $X_{\mathcal{L}}^n$, we can show that (see Appendix B)

$$\Pr \left(\left(\psi_n(\varphi_n^{(\mathcal{A})}(X_{\mathcal{A}}^n), \varphi_n^{(\mathcal{A}^c)}(a_{\mathcal{A}^c} X_{\mathcal{A}^c}^{(-i)})) : a_{\mathcal{A}^c} \in \mathcal{X}_{\mathcal{A}^c} \right) \neq \left(f_n(X_{\mathcal{A}}^n, a_{\mathcal{A}^c} X_{\mathcal{A}^c}^{(-i)}) : a_{\mathcal{A}^c} \in \mathcal{X}_{\mathcal{A}^c} \right) \right) \leq \frac{|\mathcal{X}_{\mathcal{A}^c}|}{q} \varepsilon_n. \quad (37)$$

Thus, we have

$$\Pr(W_{\mathcal{A},i} \neq [X_{\mathcal{A},i}]_{\overline{\mathcal{X}}_{\mathcal{A}}}) \leq \left(1 + \frac{|\mathcal{X}_{\mathcal{A}^c}|}{q} \right) \varepsilon_n + \mu_n, \quad (38)$$

which implies

$$\Pr(W_{\ell,i} \neq [X_{\ell,i}]_{\overline{\mathcal{X}}_{\ell}}) \leq \left(1 + \frac{|\mathcal{X}_{\mathcal{A}^c}|}{q} \right) \varepsilon_n + \mu_n \quad (39)$$

for each $\ell \in \mathcal{A}$. By conducting the above procedures for $i \in [1 : n]$, $W_{\ell}^n = (W_{\ell,1}, \dots, W_{\ell,n})$ satisfies

$$\mathbb{E} \left[\frac{1}{n} d_H(W_{\ell}^n, [X_{\ell}^n]_{\overline{\mathcal{X}}_{\ell}}) \right] = \frac{1}{n} \sum_{i=1}^n \Pr(W_{\ell,i} \neq [X_{\ell,i}]_{\overline{\mathcal{X}}_{\ell}}) \quad (40)$$

$$\leq \left(1 + \frac{|\mathcal{X}_{\mathcal{A}^c}|}{q} \right) \varepsilon_n + \mu_n. \quad (41)$$

By the Markov inequality, for any $\beta > 0$, we have

$$\Pr \left(\frac{1}{n} d_H(W_{\ell}^n, [X_{\ell}^n]_{\overline{\mathcal{X}}_{\ell}}) \geq \beta \right) \leq \frac{1}{\beta} \left[\left(1 + \frac{|\mathcal{X}_{\mathcal{A}^c}|}{q} \right) \varepsilon_n + \mu_n \right]. \quad (42)$$

Then, by Lemma 3, there exists a code $(\kappa_n^{(\ell)}, \tau_n^{(\ell)})$ of size $2^{n\delta}$ such that

$$\Pr \left(\tau_n^{(\ell)}(\kappa_n^{(\ell)}([X_{\ell}^n]_{\overline{\mathcal{X}}_{\ell}}), W_{\ell}^n) \neq [X_{\ell}^n]_{\overline{\mathcal{X}}_{\ell}} \right) \leq \frac{1}{\beta} \left[\left(1 + \frac{|\mathcal{X}_{\mathcal{A}^c}|}{q} \right) \varepsilon_n + \mu_n \right] + \nu_n(\beta, |\mathcal{X}_{\ell}|) 2^{-n\delta}. \quad (43)$$

By taking $\delta > 0$ appropriately compared to $\beta > 0$, the error probability in the right hand side of (43) converges to 0 for every $\ell \in \mathcal{A}$ and $\mathcal{A} \in \overline{\mathcal{L}}$. This means that, if each encoder sends additional message of rate δ , then the modified decoder can reproduce $[X_{\mathcal{L}}^n]_{\overline{\mathcal{X}}_{\mathcal{L}}}$ with vanishing error probability. Thus, $(R_{\ell} + \delta : \ell \in \mathcal{L}) \in \mathcal{R}(\mathbf{X}_{\mathcal{L}} | \mathbf{f}_{\overline{\mathcal{X}}_{\mathcal{L}}})$. Since $\beta > 0$ can be arbitrarily small, and we can make $\delta > 0$ arbitrarily small accordingly, we have

$$\mathcal{R}(\mathbf{X}_{\mathcal{L}} | \mathbf{f}) \subseteq \mathcal{R}(\mathbf{X}_{\mathcal{L}} | \mathbf{f}_{\overline{\mathcal{X}}_{\mathcal{L}}}). \quad (44)$$

We also have (18) since the modified code can also compute \mathbf{f} with vanishing error probability. \blacksquare

B. Proof of Achievability Part of Theorem 1

The proof of the achievability part, i.e., \supseteq part of (17), is essentially the same as [17]. First, we claim that, given $[\mathbf{x}_{\ell}]_{\overline{\mathcal{X}}_{\ell}}$ and $P_{\mathbf{x}_{\ell}}$ of a sequence $\mathbf{x}_{\ell} \in \mathcal{X}_{\ell}^n$, we can construct a sequence $\hat{\mathbf{x}}_{\ell} \in \mathcal{X}_{\ell}^n$ satisfying $[\hat{\mathbf{x}}_{\ell}]_{\overline{\mathcal{X}}_{\ell}} = [\mathbf{x}_{\ell}]_{\overline{\mathcal{X}}_{\ell}}$ and $\hat{\mathbf{x}}_{\ell} = \sigma_{\ell}(\mathbf{x}_{\ell})$ for some permutation σ_{ℓ} on $[1 : n]$. Indeed, we can construct $\hat{\mathbf{x}}_{\ell} = \hat{\mathbf{x}}_{\ell}([\mathbf{x}_{\ell}]_{\overline{\mathcal{X}}_{\ell}}, P_{\mathbf{x}_{\ell}})$ as follows. From $[\mathbf{x}_{\ell}]_{\overline{\mathcal{X}}_{\ell}}$, we can determine a partition $\{\mathcal{I}_{\mathcal{C}} : \mathcal{C} \in \overline{\mathcal{X}}_{\ell}\}$ of $[1 : n]$ as

$$\mathcal{I}_{\mathcal{C}} := \{i \in [1 : n] : [x_{\ell,i}]_{\overline{\mathcal{X}}_{\ell}} = \mathcal{C}\}, \quad \mathcal{C} \in \overline{\mathcal{X}}_{\ell}. \quad (45)$$

Then, for given $P_{\mathbf{x}_\ell}$, we can divide each $\mathcal{I}_\mathcal{C}$ into a partition $\{\mathcal{I}_a : a \in \mathcal{C}\}$ so that¹⁰

$$|\mathcal{I}_a| = nP_{\mathbf{x}_\ell}(a), \quad \forall a \in \mathcal{C} \subseteq \mathcal{X}_\ell. \quad (46)$$

Note that $\{\mathcal{I}_a : a \in \mathcal{X}_\ell\}$ is also a partition of $[1 : n]$; i.e., for each $i \in [1 : n]$, there exists only one $\hat{x}_{\ell,i} \in \mathcal{X}_\ell$ such that $i \in \mathcal{I}_{\hat{x}_{\ell,i}}$. Then, it is not hard to see that $\hat{\mathbf{x}}_\ell = (\hat{x}_{\ell,1}, \dots, \hat{x}_{\ell,n})$ satisfies the desired property.

Now, suppose that we are given a code $\hat{\Phi}_n = (\hat{\varphi}_n^{(1)}, \dots, \hat{\varphi}_n^{(L)}, \hat{\psi}_n)$ for computing $[X_\mathcal{L}^n]_{\bar{\mathcal{X}}_\mathcal{L}} = f_{n,\bar{\mathcal{X}}_\mathcal{L}}(X_\mathcal{L}^n)$ with error probability ε_n . From this code, we shall construct a code for computing f_n as follows. In the new code, upon observing $X_\ell^n = \mathbf{x}_\ell$, the ℓ th encoder sends the marginal type $P_{\mathbf{x}_\ell}$ of \mathbf{x}_ℓ by using $|\mathcal{X}_\ell| \log(n+1)$ bits in addition to the codeword $\hat{\varphi}_n^{(\ell)}(\mathbf{x}_\ell)$ of the original code. Suppose that the decoder obtained $[\mathbf{x}_\ell]_{\bar{\mathcal{X}}_\ell}$ from the codewords $(\hat{\varphi}_n^{(\ell)}(\mathbf{x}_\ell) : \ell \in \mathcal{L})$. Then, since $P_{\mathbf{x}_\ell}$ is also sent from the ℓ th encoder, for each $\ell \in \mathcal{L}$, the decoder can construct a sequence $\hat{\mathbf{x}}_\ell = \hat{\mathbf{x}}_\ell([\mathbf{x}_\ell]_{\bar{\mathcal{X}}_\ell}, P_{\mathbf{x}_\ell})$ satisfying $[\hat{\mathbf{x}}_\ell]_{\bar{\mathcal{X}}_\ell} = [\mathbf{x}_\ell]_{\bar{\mathcal{X}}_\ell}$ and $\hat{\mathbf{x}}_\ell = \sigma_\ell(\mathbf{x}_\ell)$ for some permutation σ_ℓ on $[1 : n]$ as shown above. Since f_n is $\bar{\mathcal{X}}_\mathcal{A}$ -informative for every $\mathcal{A} \in \bar{\mathcal{L}}$, we have

$$f_n(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_L) = f_n(\hat{\mathbf{x}}_1, \mathbf{x}_2, \dots, \mathbf{x}_L) \quad (47)$$

$$= f_n(\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2, \dots, \mathbf{x}_L) \quad (48)$$

$$\vdots \quad (49)$$

$$= f_n(\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2, \dots, \hat{\mathbf{x}}_L). \quad (50)$$

This implies that the decoder can compute $f_n(\mathbf{x}_\mathcal{L})$ with error probability ε_n , and thus we have

$$\mathcal{R}(\mathbf{X}_\mathcal{L} | \mathbf{f}_{\bar{\mathcal{X}}_\mathcal{L}}) \subseteq \mathcal{R}(\mathbf{X}_\mathcal{L} | \mathbf{f}). \quad (51)$$

■

V. TWO TERMINALS

In this section, we apply Theorem 1 for the two terminal setting.

A. Han-Kobayashi's Dichotomy

In [10], in order to classify symbol-wise functions by whether the function computation regions coincide with the Slepian-Wolf region, Han and Kobayashi introduced the following class of functions (see also [16]).

Definition 9 (HK Function) A function f_n is called a Han-Kobayashi (HK) function if f_n is a symbol-wise function defined by some $f : \mathcal{X}_1 \times \mathcal{X}_2 \rightarrow \mathcal{V}$ such that

- 1) for every $a \neq a'$ in \mathcal{X}_1 , the functions $f(a, \cdot)$ and $f(a', \cdot)$ are distinct,
- 2) for every $b \neq b'$ in \mathcal{X}_2 , the functions $f(\cdot, b)$ and $f(\cdot, b')$ are distinct, and

¹⁰Although there are several partitions which satisfy (46), the choice of a partition does not affect the argument; we may choose a partition so that, for $a, \hat{a} \in \mathcal{C}$ satisfying $a < \hat{a}$, $i \in \mathcal{I}_a$ and $j \in \mathcal{I}_{\hat{a}}$ imply $i < j$.

3) $f(a, b) \neq f(a', b')$ whenever $a \neq a'$ and $b \neq b'$.

In Proposition 2, we have seen that the symbol-wise function is $\overline{\mathcal{X}}_{1,f}$ -informative and $\overline{\mathcal{X}}_{2,f}$ -informative for partitions induced by f . It is not difficult to verify the following.

Proposition 3 If symbol-wise function f^n satisfies Conditions 1 and 2 of the HK function, respectively, then it is \mathcal{X}_1 -informative and \mathcal{X}_2 -informative, respectively.¹¹

We can find that Condition 3 of the HK function can be rephrased as follows: for every $v \in \mathcal{V}$, the inverse image satisfies either $f^{-1}(v) \subseteq \mathcal{X}_1 \times \{x_2\}$ for some $x_2 \in \mathcal{X}_2$ or $f^{-1}(v) \subseteq \{x_1\} \times \mathcal{X}_2$ for some $x_1 \in \mathcal{X}_1$. Thus, for a given correlated random variables (X_1, X_2) on $\mathcal{X}_1 \times \mathcal{X}_2$, we have $X_1 \text{---} f(X_1, X_2) \text{---} X_2$, which implies the following.

Proposition 4 If symbol-wise function $\mathbf{f} = \{f_n\}_{n=1}^\infty$ satisfies Conditions 3 of the HK function and $(\mathbf{X}_1, \mathbf{X}_2)$ is an i.i.d. source, then \mathbf{f} induces conditional independence for $((\mathbf{X}_1, \mathbf{X}_2), \{\{1\}, \{2\}\})$.

By noting Propositions 3 and 4, we can recover the sufficiency part of Han-Kobayashi's dichotomy theorem as a corollary of Theorem 1.

Corollary 3 ([10]) If $\mathbf{f} = \{f_n\}_{n=1}^\infty$ is an HK function, then, for any i.i.d. sources satisfying the positivity condition, we have $\mathcal{R}(\mathbf{X}_1, \mathbf{X}_2 | \mathbf{f}) = \mathcal{R}_{\text{SW}}(\mathbf{X}_1, \mathbf{X}_2)$.

B. Kuzuoka-Watanabe's Dichotomy

In [16], the class of joint sensitive functions was introduced.

Definition 10 A function $f_n : \mathcal{X}_1^n \times \mathcal{X}_2^n \rightarrow \mathcal{Z}_n$ is said to be *jointly sensitive* if $f_n(\mathbf{x}_1, \mathbf{x}_2) \neq f_n(\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2)$ whenever $\mathbf{x}_1 \neq \hat{\mathbf{x}}_1$ and $\mathbf{x}_2 \neq \hat{\mathbf{x}}_2$.

By a similar reasoning as Proposition 4,¹² we have the following property of jointly sensitive functions.

Proposition 5 If $\mathbf{f} = \{f_n\}_{n=1}^\infty$ is jointly sensitive, then, for any given source $(\mathbf{X}_1, \mathbf{X}_2)$ (not necessarily i.i.d.), \mathbf{f} induces conditional independence for $((\mathbf{X}_1, \mathbf{X}_2), \{\{1\}, \{2\}\})$.

By noting Proposition 5, Theorem 1 implies the following.

Corollary 4 If \mathbf{f} is $\overline{\mathcal{X}}_1$ -informative and $\overline{\mathcal{X}}_2$ -informative for a given partition $\overline{\mathcal{X}}_{\mathcal{L}} = (\overline{\mathcal{X}}_1, \overline{\mathcal{X}}_2)$ and \mathbf{f} is jointly sensitive, then, for any smooth source $(\mathbf{X}_1, \mathbf{X}_2)$, we have $\mathcal{R}(\mathbf{X}_1, \mathbf{X}_2 | \mathbf{f}) = \mathcal{R}(\mathbf{X}_1, \mathbf{X}_2 | \mathbf{f}_{\overline{\mathcal{X}}_{\mathcal{L}}})$.

¹¹By a slight abuse of terminology, we say \mathcal{X}_ℓ -informative to mean $\overline{\mathcal{X}}_\ell$ -informative for the finest partition $\overline{\mathcal{X}}_\ell = \{\{x_\ell\} : x_\ell \in \mathcal{X}_\ell\}$.

¹²Instead of one symbol function f , we consider the same argument for f_n .

In general, a function being HK function does not imply the same function being jointly sensitive (cf. [16, Proposition 1]). Because of this fact, Han-Kobayashi's dichotomy theorem is not valid for the class of smooth sources, and an alternative dichotomy theorem was derived in [16]; the sufficiency part of the dichotomy theorem in [16] can be recovered as a special case of Corollary 4.

Corollary 5 ([16]) If f is HK function and jointly sensitive, then, for any smooth sources, we have $\mathcal{R}(\mathbf{X}_1, \mathbf{X}_2|f) = \mathcal{R}_{\text{SW}}(\mathbf{X}_1, \mathbf{X}_2)$.¹³

C. Mixed Sources

When a source has memory, the statement of Proposition 4 may not hold in general. Thus, it was not clear if a function being HK (but not jointly sensitive) implies $\mathcal{R}(\mathbf{X}_1, \mathbf{X}_2|f) = \mathcal{R}_{\text{SW}}(\mathbf{X}_1, \mathbf{X}_2)$ when the source has memory. Here, we show that this claim holds for a mixture of i.i.d. sources.

Let $(\mathbf{X}_1, \mathbf{X}_2) = \{(X_1^n, X_2^n)\}_{n=1}^\infty$ be a mixture of two i.i.d. sources with positivity condition such that

$$P_{X_1^n X_2^n}(\mathbf{x}_1, \mathbf{x}_2) = \Pr(A = 0)P_{X_1 X_2, 0}^n(\mathbf{x}_1, \mathbf{x}_2) + \Pr(A = 1)P_{X_1 X_2, 1}^n(\mathbf{x}_1, \mathbf{x}_2) \quad (52)$$

for some random variable A taking values in $\{0, 1\}$. Let $f_n : \mathcal{X}_1^n \times \mathcal{X}_2 \rightarrow \mathcal{V}^n$ be an HK function. Suppose that $P_{V, 0} \neq P_{V, 1}$, where $P_{V, a}$ is the distribution of $V = f(X_1, X_2)$ under $P_{X_1 X_2, a}$. Then, there exists an estimator γ_n such that the estimation error probability $\Pr(A \neq \gamma_n(V^n))$ vanishes asymptotically. Thus, by taking $S_n = (A, V^n)$, the function $\mathbf{f} = \{f_n\}_{n=1}^\infty$ induces conditional independence for $((\mathbf{X}_1, \mathbf{X}_2), \{\{1\}, \{2\}\})$. Consequently, Theorem 1 and Proposition 3 imply $\mathcal{R}(\mathbf{X}_1, \mathbf{X}_2|f) = \mathcal{R}_{\text{SW}}(\mathbf{X}_1, \mathbf{X}_2)$.

VI. MORE THAN TWO TERMINALS WITH I.I.D. SOURCES

In [10], Han-Kobayashi considered classification of functions for more than two terminals; let $\mathcal{F}_{\text{SW}}^{\text{iid}}$ be the set of all functions f such that the symbol-wise function induced by f satisfies $\mathcal{R}(\mathbf{X}_{\mathcal{L}}|f) = \mathcal{R}_{\text{SW}}(\mathbf{X}_{\mathcal{L}})$ for any i.i.d. sources with positivity condition. They derived a necessary condition and sufficient conditions such that a given function belongs to $\mathcal{F}_{\text{SW}}^{\text{iid}}$. In this section, as an application of the general results in Section IV, we derive a novel sufficient condition that strictly subsumes the sufficient conditions in [10].

First, we shall review the necessary condition and the sufficient conditions derived in [10]. For that purpose, it is convenient to introduce some “geometrical” notation of functions.

Definition 11 (Projection) For $f : \mathcal{X}_{\mathcal{L}} \rightarrow \mathcal{V}$ and a subset $\mathcal{A} \subseteq \mathcal{L}$, we define projected function $f_{\mathcal{A}} : \mathcal{X}_{\mathcal{A}} \rightarrow \mathcal{V}^{|\mathcal{X}_{\mathcal{A}^c}|}$ by¹⁴

$$f_{\mathcal{A}}(x_{\mathcal{A}}) = (f(x_{\mathcal{A}}, x_{\mathcal{A}^c}) : x_{\mathcal{A}^c} \in \mathcal{X}_{\mathcal{A}^c}). \quad (53)$$

¹³In [16], the dichotomy theorem was stated in terms of “totally sensitive function”; however, a function being HK and jointly sensitive is equivalent to HK and totally sensitive.

¹⁴For $\mathcal{A} = \mathcal{L}$, note that $f_{\mathcal{L}} = f$.

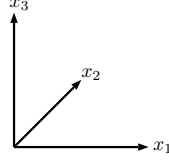


Fig. 5: Coordinates.

Definition 12 (Span) For $f : \mathcal{X}_{\mathcal{L}} \rightarrow \mathcal{V}$ and $v \in \mathcal{V}$, we define $\text{span}f^{-1}(v)$ as the minimal subset $\mathcal{B} \subseteq \mathcal{L}$ such that there exists $x_{\mathcal{B}^c} \in \mathcal{X}_{\mathcal{B}^c}$ satisfying $f^{-1}(v) \subseteq \mathcal{X}_{\mathcal{B}} \times \{x_{\mathcal{B}^c}\}$. In particular, when $|f^{-1}(v)| \leq 1$, then $\text{span}f^{-1}(v) = \emptyset$. For a subset $\mathcal{A} \subseteq \mathcal{L}$, $\text{span}f_{\mathcal{A}}^{-1}(v)$ is defined similarly by replacing f with projected function $f_{\mathcal{A}} : \mathcal{X}_{\mathcal{A}} \rightarrow \mathcal{V}^{|\mathcal{X}_{\mathcal{A}^c}|}$.

Essentially, $\text{span}f^{-1}(v)$ is the set of coordinates ℓ such that the value of x_{ℓ} cannot be uniquely determined from v . Let us verify Definition 11 and Definition 12 with some examples.

Example 1 Let us consider three terminal function $f(x_1, x_2, x_3)$, where values of (x_1, x_2, x_3) are ordered as in Fig. 5.¹⁵ For the function in Fig. 6(a), $f_{\{3\}}(0) = (0, 0, 3, 4)$ and $f_{\{3\}}(1) = (1, 2, 3, 2)$. For the same function, $\text{span}f^{-1}(1) = \emptyset$ and $\text{span}f^{-1}(0) = \{2\}$, respectively. For the function in Fig. 6(f), $\text{span}f^{-1}(0) = \{1, 2, 3\}$.

Now, we are ready to review the conditions derived by Han and Kobayashi [10], where we rephrase their conditions by using the notations introduced above.

Proposition 6 ([10]; Necessary Condition) A function $f : \mathcal{X}_{\mathcal{L}} \rightarrow \mathcal{V}$ belongs to $\mathcal{F}_{\text{SW}}^{\text{iid}}$ only if the following conditions holds: for every $\mathcal{A} \subseteq \mathcal{L}$,

$$f_{\mathcal{A}}(x_{\mathcal{A}}) \neq f_{\mathcal{A}}(\hat{x}_{\mathcal{A}}) \quad (54)$$

holds for any pair $x_{\mathcal{A}}, \hat{x}_{\mathcal{A}} \in \mathcal{X}_{\mathcal{A}}$ satisfying $x_{\ell} \neq \hat{x}_{\ell}$ for every $\ell \in \mathcal{A}$.

Proposition 7 ([10]; Sufficient Condition) A function $f : \mathcal{X}_{\mathcal{L}} \rightarrow \mathcal{V}$ belongs to $\mathcal{F}_{\text{SW}}^{\text{iid}}$ if the following two conditions hold:

- 1) For any $v \in \mathcal{V}$, the inverse image satisfies $|\text{span}f^{-1}(v)| \leq 1$.
- 2) For arbitrary $\ell \in \mathcal{L}$, the projected function $f_{\mathcal{L} \setminus \{\ell\}}$ is injective.

Proposition 8 ([10]; Sufficient Condition) A function $f : \mathcal{X}_{\mathcal{L}} \rightarrow \mathcal{V}$ belongs to $\mathcal{F}_{\text{SW}}^{\text{iid}}$ if, in addition to Condition 2 in Proposition 7, the following condition holds:

1')

$$\bigcup_{v \in \mathcal{V}} \text{span}f^{-1}(v) \subsetneq \mathcal{L}. \quad (55)$$

¹⁵This method of describing functions was introduced in [10].

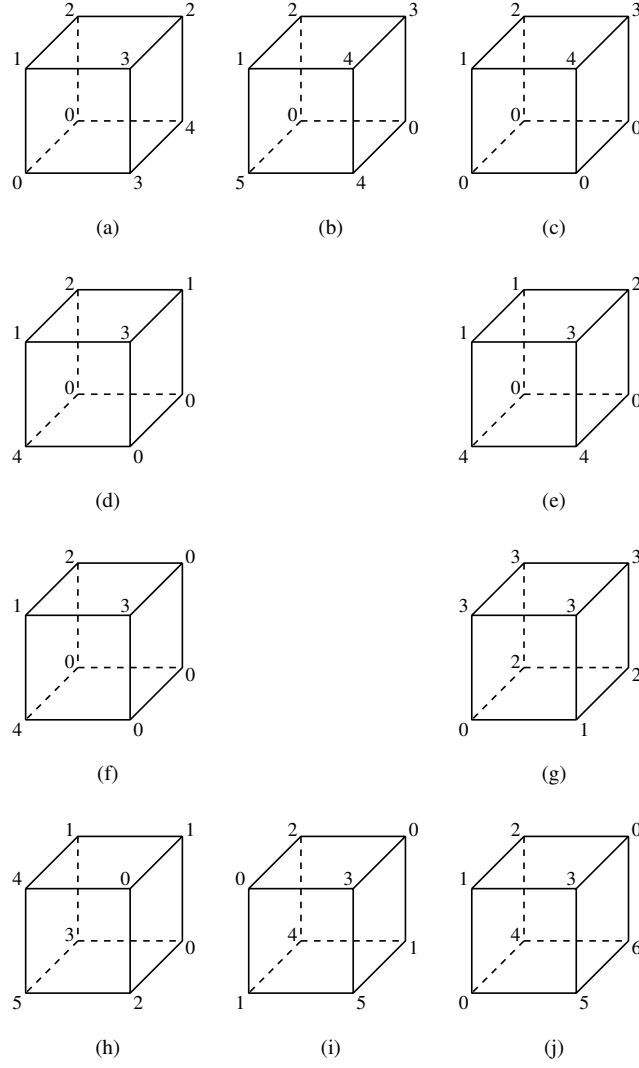


Fig. 6: Representative patterns for $f(x_1, x_2, x_3)$.

In the following example, we examine if functions in Fig. 6 satisfy the conditions in Propositions 6, 7, and 8, which is summarized in Table III.¹⁶ In the rightmost column of the table, we also described if we can decide $f \in \mathcal{F}_{SW}^{iid}$ or not by using Propositions 6, 7, and 8, and Theorem 2 shown later in this section.

Example 2 All the functions in Fig. 6(a)-(h) satisfy the necessary condition in Proposition 6; the functions in Fig. 6(i) and (j) violate the necessary condition in Proposition 6. From Proposition 6, we can determine that the functions in (i) and (j) do not belong to \mathcal{F}_{SW}^{iid} ; on the other hand, from Proposition 7 and Proposition 8, we can determine that the functions in (a), (b), and (d) belong to \mathcal{F}_{SW}^{iid} . In [10], it is claimed that, by some inspection, the

¹⁶All the functions in Fig. 6 except (g) and (h) are the same as those in [10, Fig. 4] and [10, Fig. 5].

TABLE III: Necessary/Sufficient Conditions for Functions in Fig. 6

$f(x_1, x_2, x_3)$	Neces. Cond. (Prop. 6)	Suff. Cond. 1 (Prop. 7)	Suff. Cond. 1' (Prop. 8)	Suff. Cond. 2 (Prop. 7)	$f \in \mathcal{F}_{SW}^{\text{iid}}$
(a)	✓	✓		✓	✓ (Prop. 7)
(b)	✓	✓		✓	✓ (Prop. 7)
(c)	✓		✓		✓ (Thm. 2)
(d)	✓		✓	✓	✓ (Prop. 8)
(e)	✓	✓			✓ (Thm. 2)
(f)	✓			✓	
(g)	✓		✓		✓ (Thm. 2)
(h)	✓			✓	✓ (Thm. 2)
(i)		✓			× (Prop. 6)
(j)				✓	× (Prop. 6)

function in (c) is found to belong to $\mathcal{F}_{SW}^{\text{iid}}$; later in Example 3, we will verify this fact from more general sufficient condition.

The conditions in Propositions 7 and 8 can be rephrased in terms of the general framework in Section IV as follows. When function f satisfies Conditions 1 and 2 of Proposition 7, then it induces conditional independence for $(\mathbf{X}_{\mathcal{L}}, \bar{\mathcal{L}})$ with the finest partition $\bar{\mathcal{L}} \equiv \mathcal{L}$ and it is semi \mathcal{X}_{ℓ} -informative for every $\ell \in \mathcal{L}$. Thus, Corollary 2 with $k = 1$ implies $f \in \mathcal{F}_{SW}^{\text{iid}}$. On the other hand, when function f satisfies Condition 2 of Proposition 7 and Condition 1' of Proposition 8, then it induces conditional independence for $(\mathbf{X}_{\mathcal{L}}, \bar{\mathcal{L}})$ with partition $\bar{\mathcal{L}} = \{\{\ell_0\}, \mathcal{L} \setminus \{\ell_0\}\}$ for some fixed $\ell_0 \in \mathcal{L}$ and it is semi $\mathcal{X}_{\mathcal{L} \setminus \{\ell_0\}}$ -informative and semi \mathcal{X}_{ℓ_0} -informative. Thus, Corollary 2 with $k = 1$ again implies $f \in \mathcal{F}_{SW}^{\text{iid}}$. In fact, Condition 2 of Proposition 7 is quite strong since it implies that f is semi $\mathcal{X}_{\mathcal{L} \setminus \{\ell\}}$ -informative for every $\ell \in \mathcal{L}$.

As we discussed above, Propositions 7 and 8 are implied as special cases of Corollary 2. In order to leverage Corollary 2 in a full generality, let us rephrase conditional independence structure and informative structure in terms of the basis function $f : \mathcal{X}_{\mathcal{L}} \rightarrow \mathcal{V}$, which will be proved at the end of this section.

Theorem 2 For a given nontrivial partition $\bar{\mathcal{L}}$ of \mathcal{L} and tuple of partitions $\bar{\mathcal{X}}_{\mathcal{L}} = (\bar{\mathcal{X}}_{\ell} : \ell \in \mathcal{L})$, suppose that $f : \mathcal{X}_{\mathcal{L}} \rightarrow \mathcal{V}$ satisfies the following condition:

- 1) For every $v \in \mathcal{V}$, the inverse image satisfies $\text{span} f^{-1}(v) \subseteq \mathcal{A}$ for some $\mathcal{A} \in \bar{\mathcal{L}}$.

Then, for any i.i.d. source $\mathbf{X}_{\mathcal{L}}$, the symbol-wise function f^n induces conditional independence for $(\mathbf{X}_{\mathcal{L}}, \bar{\mathcal{L}})$. On the other hand, suppose that f satisfies the following condition:

- 2) For every $\mathcal{A} \in \bar{\mathcal{L}}$ and every $x_{\mathcal{A}}, \hat{x}_{\mathcal{A}} \in \mathcal{X}_{\mathcal{A}}$ with $[x_{\mathcal{A}}]_{\bar{\mathcal{X}}_{\mathcal{A}}} \neq [\hat{x}_{\mathcal{A}}]_{\bar{\mathcal{X}}_{\mathcal{A}}}$, the projected function satisfies

$$f_{\mathcal{A}}(x_{\mathcal{A}}) \neq f_{\mathcal{A}}(\hat{x}_{\mathcal{A}}). \quad (56)$$

Then, the symbol-wise function f^n is semi $\bar{\mathcal{X}}_{\mathcal{A}}$ -informative for every $\mathcal{A} \in \bar{\mathcal{L}}$.

We can use Theorem 2 and Corollary 2 to verify if a given function f is included in $\mathcal{F}_{\text{SW}}^{\text{iid}}$. More specifically, if a given function satisfies the two conditions in Theorem 2, then we can use Corollary 2 to show that that function is included in $\mathcal{F}_{\text{SW}}^{\text{iid}}$. We will illustrate the utility of this approach with some examples in the following.

Example 3 Let us consider the function in Fig. 6(c). For the partition $\overline{\mathcal{L}} = \{\{1, 2\}, \{3\}\}$, since $\text{span}f^{-1}(0) = \{1, 2\}$ and $\text{span}f^{-1}(v) = \emptyset$ for $v = 1, \dots, 4$, we can verify that Condition 1 of Theorem 2 is satisfied. Furthermore, for the finest partitions $\overline{\mathcal{X}}_i \equiv \mathcal{X}_i$ for $i = 1, 2, 3$, since the projected functions $f_{\{1, 2\}}$ and $f_{\{3\}}$ are both injective, we can verify that Condition 2 is satisfied. Thus, the function in Fig. 6(c) belongs to $\mathcal{F}_{\text{SW}}^{\text{iid}}$.

Example 4 Let us consider the function in Fig. 6(e). There are two manners to show that this function belongs to $\mathcal{F}_{\text{SW}}^{\text{iid}}$. The first manner is the same as Example 3: For the partition $\overline{\mathcal{L}} = \{\{1, 2\}, \{3\}\}$, we can verify that Condition 1 of Theorem 2 is satisfied. Furthermore, for the finest partitions $\overline{\mathcal{X}}_i \equiv \mathcal{X}_i$ for $i = 1, 2, 3$, we can verify that Condition 2 is satisfied. Thus, it belongs to $\mathcal{F}_{\text{SW}}^{\text{iid}}$. The second manner is as follows: For the partition $\overline{\mathcal{L}} = \{\{1\}, \{2\}, \{3\}\}$, we can verify that Condition 1 of Theorem 2 is satisfied. Furthermore, for the finest partitions $\overline{\mathcal{X}}_i \equiv \mathcal{X}_i$ for $i = 1, 2, 3$, we can verify that Condition 2 is satisfied. Thus, it belongs to $\mathcal{F}_{\text{SW}}^{\text{iid}}$.

Example 5 Let us consider the function in Fig. 6(h). For the partition $\overline{\mathcal{L}} = \{\{1\}, \{2, 3\}\}$, since $\text{span}f^{-1}(0) = \{2, 3\}$, $\text{span}f^{-1}(1) = \{1\}$, and $\text{span}f^{-1}(v) = \emptyset$ for $v = 2, \dots, 5$, we can verify that Condition 1 of Theorem 2 is satisfied. Furthermore, for the finest partitions $\overline{\mathcal{X}}_i \equiv \mathcal{X}_i$ for $i = 1, 2, 3$, we can verify that Condition 2 is satisfied. Thus, it belongs to $\mathcal{F}_{\text{SW}}^{\text{iid}}$.

For some examples, we need to use recursion of Corollary 2.

Example 6 Let us consider the function in Fig. 6(g). To prove that this function belongs to $\mathcal{F}_{\text{SW}}^{\text{iid}}$, we need to use recursion of Corollary 2. For the partition $\overline{\mathcal{L}}^{(1)} = \{\{1, 2\}, \{3\}\}$, we can verify that f satisfies Condition 1 of Theorem 2. Furthermore, for the partitions $\overline{\mathcal{X}}_1^{(1)} = \{\{0, 1\}\}$, $\overline{\mathcal{X}}_2^{(1)} = \{\{0\}, \{1\}\}$, and $\overline{\mathcal{X}}_3^{(1)} = \{\{0\}, \{1\}\}$, we can verify that f satisfies Condition 2 of Theorem 2. Thus, we have $\mathcal{R}(X_{\mathcal{L}}|f) = \mathcal{R}(X_{\mathcal{L}}|f, f_{\overline{\mathcal{X}}_{\mathcal{L}}^{(1)}})$, where $(f, f_{\overline{\mathcal{X}}_{\mathcal{L}}^{(1)}})$ is the sequence of symbol-wise functions that consists of the product of f and local function $f_{\overline{\mathcal{X}}_{\mathcal{L}}^{(1)}}$. In the recursion step, we shall show that the product function $(f, f_{\overline{\mathcal{X}}_{\mathcal{L}}^{(1)}})$ belongs to $\mathcal{F}_{\text{SW}}^{\text{iid}}$. Since $\overline{\mathcal{X}}_1$ is the trivial partition and $\overline{\mathcal{X}}_2$ and $\overline{\mathcal{X}}_3$ are the finest partition, $(f, f_{\overline{\mathcal{X}}_{\mathcal{L}}^{(1)}})$ is equivalent to $(f(x_1, x_2, x_3), x_2, x_3)$; see Fig. 7. Then, for the partition $\overline{\mathcal{L}}^{(2)} = \{\{1\}, \{2\}, \{3\}\}$, we can verify that $(f, f_{\overline{\mathcal{X}}_{\mathcal{L}}^{(1)}})$ satisfies Condition 1 of Theorem 2. Furthermore, for the finest partitions $\overline{\mathcal{X}}_i^{(2)} \equiv \mathcal{X}_i$ for $i = 1, 2, 3$, we can verify that $(f, f_{\overline{\mathcal{X}}_{\mathcal{L}}^{(1)}})$ satisfies Condition 2 of Theorem 2. Thus, the function $(f, f_{\overline{\mathcal{X}}_{\mathcal{L}}^{(1)}})$ belongs to $\mathcal{F}_{\text{SW}}^{\text{iid}}$, which implies that f belongs to $\mathcal{F}_{\text{SW}}^{\text{iid}}$.

Example 7 Let us consider the function in Fig. 8(a), where $\mathcal{X}_1 = \mathcal{X}_2 = \{0, 1\}$ and $\mathcal{X}_3 = \{0, 1, 2\}$. To prove that this function belongs to $\mathcal{F}_{\text{SW}}^{\text{iid}}$, we need to use recursion of Corollary 2. For the partition $\overline{\mathcal{L}}^{(1)} = \{\{1\}, \{2, 3\}\}$, we can verify that f satisfies Condition 1 of Theorem 2. Furthermore, for the partitions $\overline{\mathcal{X}}_1^{(1)} = \{\{0\}, \{1\}\}$, $\overline{\mathcal{X}}_2^{(1)} = \{\{0, 1\}\}$, and $\overline{\mathcal{X}}_3^{(1)} = \{\{0\}, \{1, 2\}\}$, we can verify that f satisfies Condition 2 of Theorem 2. Thus, we

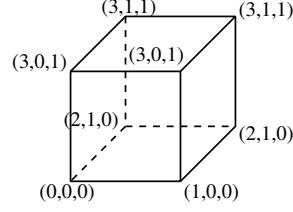


Fig. 7: The function in recursion step of Example 6.

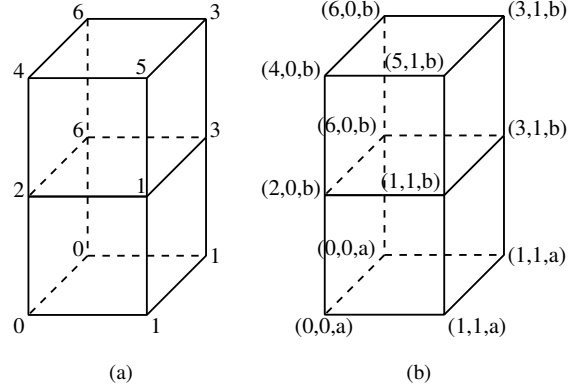


Fig. 8: (a) Description of $f(x_1, x_2, x_3)$ in Example 7; (b) Description of the function in recursion step in the same example, where $a = \{0\}$ and $b = \{1, 2\}$.

have $\mathcal{R}(\mathbf{X}_{\mathcal{L}}|\mathbf{f}) = \mathcal{R}(\mathbf{X}_{\mathcal{L}}|\mathbf{f}, \mathbf{f}_{\overline{\mathcal{X}}_{\mathcal{L}}^{(1)}})$. In the recursion step, since $\overline{\mathcal{X}}_2^{(1)}$ is the trivial partition, $(f, \mathbf{f}_{\overline{\mathcal{X}}_{\mathcal{L}}^{(1)}})$ is equivalent to $(f(x_1, x_2, x_3), x_1, [x_3]_{\overline{\mathcal{X}}_3^{(1)}})$; see Fig. 8(b). Thus, for the partition $\overline{\mathcal{L}}^{(2)} = \{\{1\}, \{2\}, \{3\}\}$, we can verify that $(f, \mathbf{f}_{\overline{\mathcal{X}}_{\mathcal{L}}^{(1)}})$ satisfies Condition 1 of Theorem 2. Furthermore, for the partitions $\overline{\mathcal{X}}_i^{(2)} = \{\{0\}, \{1\}\}$ for $i = 1, 2$ and $\overline{\mathcal{X}}_3^{(2)} = \{\{0\}, \{1\}, \{2\}\}$, we can verify that $(f, \mathbf{f}_{\overline{\mathcal{X}}_{\mathcal{L}}^{(1)}})$ satisfies Condition 2 of Theorem 2. Thus, the function $(f, \mathbf{f}_{\overline{\mathcal{X}}_{\mathcal{L}}^{(1)}})$ belongs to $\mathcal{F}_{\text{SW}}^{\text{iid}}$, which implies that f belongs to $\mathcal{F}_{\text{SW}}^{\text{iid}}$.

So far, we have seen examples such that the classification problem can be solved by our approach. However, there are still some cases such that the classification problem cannot be solved by our approach. The following is one of such examples.

Example 8 Let us consider the function in Fig. 6(f). In this case, Condition 1 of Theorem 2 is not satisfied for any nontrivial partition $\overline{\mathcal{L}}$ of $\{1, 2, 3\}$; in fact, $\text{span} f^{-1}(0) = \{1, 2, 3\}$. Thus, we cannot determine whether f belongs to $\mathcal{F}_{\text{SW}}^{\text{iid}}$ or not.

In the case of three terminals, a nontrivial partition $\overline{\mathcal{L}}$ inevitably include a singleton part. In the case of more than three terminals, this is not the case. Even when the function induces conditional independence among non-singleton groups of parties, the classification problem can be solved, which is illustrated in the following example.

TABLE IV: Description of $f(x_1, x_2, x_3, x_4)$ in Example 9.

$(x_1, x_2) \setminus (x_3, x_4)$	00	01	10	11
00	0	0	4	0
01	1	6	4	2
10	1	6	4	5
11	1	3	3	3

Example 9 Let us consider the function in Table IV. For the partition $\bar{\mathcal{L}} = \{\{1, 2\}, \{3, 4\}\}$, we can verify that Condition 1 of Theorem 2 is satisfied. Furthermore, for the finest partitions $\bar{\mathcal{X}}_i = \mathcal{X}_i$ for $i = 1, 2, 3, 4$, we can verify that Condition 2 is satisfied. Thus, the function in Table IV belongs to $\mathcal{F}_{SW}^{\text{iid}}$.

Proof of Theorem 2: Since the source is i.i.d. and the function is symbol-wise, in order to prove that the function induces conditional independence, it suffices to show that, for $X_{\mathcal{L}} \sim P_{X_{\mathcal{L}}}$ and $V = f(X_{\mathcal{L}})$, the joint distribution $P_{X_{\mathcal{L}}V}$ factorizes as

$$P_V \prod_{\mathcal{A} \in \bar{\mathcal{L}}} P_{X_{\mathcal{A}}|V}. \quad (57)$$

Since function f satisfies Condition 1, for each $v \in \mathcal{V}$, we have $f^{-1}(v) \subseteq \mathcal{X}_{\mathcal{A}} \times \{x_{\mathcal{A}^c}\}$ for some $\mathcal{A} \in \bar{\mathcal{L}}$ and $x_{\mathcal{A}^c} \in \mathcal{X}_{\mathcal{A}^c}$, i.e., the value of $X_{\mathcal{A}^c}$ is determined from $V = v$. Thus, $P_{X_{\mathcal{L}}V}$ factorizes as in (57).

On the other hand, since function f satisfies Condition 2, for each $\mathcal{A} \in \bar{\mathcal{L}}$ and any $a_{\mathcal{A}} \in \mathcal{X}_{\mathcal{A}}$, the value $[a_{\mathcal{A}}]_{\bar{\mathcal{X}}_{\mathcal{A}}}$ is uniquely determined from the list $(f(a_{\mathcal{A}}, b_{\mathcal{A}^c}) : b_{\mathcal{A}^c} \in \mathcal{X}_{\mathcal{A}^c})$. Thus, we can construct a mapping satisfying Condition 1 of Definition 6. Thus, the symbol-wise function f^n is semi $\bar{\mathcal{X}}_{\mathcal{A}}$ -informative for every $\mathcal{A} \in \bar{\mathcal{L}}$. ■

VII. MORE THAN TWO TERMINALS WITH SMOOTH SOURCES

In Section VI, we considered whether $\mathcal{R}(X_{\mathcal{L}}|f) = \mathcal{R}_{SW}(X_{\mathcal{L}})$ holds or not for the class of i.i.d. sources with positivity condition. In this section, we consider the same problem for the class of smooth sources. In contrast to the classification problem for the class of i.i.d. sources in Section VI, we can completely solve the classification problem for the class of smooth sources in this section. Let $\mathcal{F}_{SW}^{\text{smt}}$ be the set of all functions f such that the symbol-wise function induced by f satisfies $\mathcal{R}(X_{\mathcal{L}}|f) = \mathcal{R}_{SW}(X_{\mathcal{L}})$ for any smooth sources. Since the class of smooth sources is broader than the class of i.i.d. sources with positivity condition, apparently we have $\mathcal{F}_{SW}^{\text{smt}} \subseteq \mathcal{F}_{SW}^{\text{iid}}$. Throughout this section, we use the notations introduced in Section VI such as Definition 11 and Definition 12.

Since the class of smooth sources include i.i.d. sources with positivity condition, the necessary condition in Proposition 6 is valid for the classification problem for the class of smooth sources. In fact, since the class of smooth sources includes sources with memory, we can apply the necessary condition in Proposition 6 for functions on extended alphabets, which will be used in the proof of “only if” part of Theorem 3.

For the class of smooth sources, we can rephrase the conditions in Corollary 2 in a similar manner as Theorem 2. However, for the class of smooth sources, we can show the matching necessary and sufficient condition in more

compact manner. For that purpose, let us introduce the concept of pseudo identity function, which is recursively defined as follows.

Definition 13 (Pseudo Identity) For a given subset $\mathcal{A} \subseteq \mathcal{L}$, the projected function $f_{\mathcal{A}} : \mathcal{X}_{\mathcal{A}} \rightarrow \mathcal{V}^{|\mathcal{X}_{\mathcal{A}^c}|}$ is said to be *pseudo identity* if either of the following conditions is satisfied:

- 1) $f_{\mathcal{A}}$ is injective;
- 2) It holds that

$$\tilde{\mathcal{A}} := \bigcup_{v \in \mathcal{V}^{|\mathcal{X}_{\mathcal{A}^c}|}} \text{span} f_{\mathcal{A}}^{-1}(v) \subsetneq \mathcal{A}, \quad (58)$$

and, for every $\mathcal{B} \in \{\tilde{\mathcal{A}}, \{\ell\} : \ell \in \mathcal{A} \setminus \tilde{\mathcal{A}}\}$, projected function $f_{\mathcal{B}} : \mathcal{X}_{\mathcal{B}} \rightarrow \mathcal{V}^{|\mathcal{X}_{\mathcal{B}^c}|}$ is pseudo identity.

For the two terminal case, the pseudo identity function is equivalent to the totally sensitive function in [16].

Remark 3 Since (58) implies that the values of x_{ℓ} for $\ell \in \mathcal{A} \setminus \tilde{\mathcal{A}}$ are uniquely determined from $f_{\mathcal{A}}(x_{\mathcal{A}})$, $f_{\{\ell\}}$ for $\ell \in \mathcal{A} \setminus \tilde{\mathcal{A}}$ are injective, i.e., pseudo identity. Thus, we only need to verify if $f_{\tilde{\mathcal{A}}}$ is pseudo identity in the recursion step.

The following theorem, which will be proved in Sections VII-A and VII-B, completely solve the classification problem for the class of smooth sources.

Theorem 3 A symbol-wise function $f : \mathcal{X}_{\mathcal{L}} \rightarrow \mathcal{V}$ belongs to $\mathcal{F}_{\text{SW}}^{\text{smt}}$ if and only if f is pseudo identity.

Now, we shall illustrate Theorem 3 by some examples.

Example 10 The functions in Figs. 6(b), (c), (d), (e), and (g) are pseudo identity, and thus are included in $\mathcal{F}_{\text{SW}}^{\text{smt}}$. Since it is not difficult to verify that (b), (c), (d), and (e) are pseudo identity, we only verify (g). Since

$$\bigcup_{v \in \mathcal{V}} \text{span} f^{-1}(v) = \{1, 2\} \subsetneq \{1, 2, 3\}, \quad (59)$$

it suffices to verify $f_{\{1,2\}}$ is pseudo identity (cf. Remark 3). Since

$$\bigcup_{v \in \mathcal{V}^{|\mathcal{X}_3|}} \text{span} f_{\{1,2\}}^{-1}(v) = \{1\} \subsetneq \{1, 2\}, \quad (60)$$

it suffices to verify $f_{\{1\}}$ is pseudo identity; it is pseudo identity since it is injective.

Example 11 The functions in Figs. 6(a), (f), (h), (i), and (j) are not pseudo identity, and thus are not included in $\mathcal{F}_{\text{SW}}^{\text{smt}}$. For (a), (f), (h), (j), since

$$\bigcup_{v \in \mathcal{V}} \text{span} f^{-1}(v) = \{1, 2, 3\}, \quad (61)$$

those functions violate the conditions of pseudo identity. On the other hand, for (i), even though

$$\bigcup_{v \in \mathcal{V}} \text{span} f^{-1}(v) = \{1, 2\} \subsetneq \{1, 2, 3\}, \quad (62)$$

the projected function $f_{\{1,2\}}$ is not pseudo identity since

$$\bigcup_{v \in \mathcal{V}^{|\mathcal{X}_3|}} \text{span} f_{\{1,2\}}^{-1}(v) = \{1, 2\}. \quad (63)$$

Note that $\mathcal{F}_{SW}^{\text{smt}} \subsetneq \mathcal{F}_{SW}^{\text{id}}$. In fact, the functions in (a) and (h) are included in $\mathcal{F}_{SW}^{\text{id}} \setminus \mathcal{F}_{SW}^{\text{smt}}$. Although it is shown that the function in (f) does not belong to $\mathcal{F}_{SW}^{\text{smt}}$, it is not clear whether it is included in $\mathcal{F}_{SW}^{\text{id}} \setminus \mathcal{F}_{SW}^{\text{smt}}$ or not (see also Example 8).

Example 12 The function in Fig. 8(a) is not pseudo identity. In fact, even though

$$\bigcup_{v \in \mathcal{V}} \text{span} f^{-1}(v) = \{2, 3\} \subsetneq \{1, 2, 3\}, \quad (64)$$

$f_{\{2,3\}}$ is not pseudo identity since

$$\bigcup_{v \in \mathcal{V}^{|\mathcal{X}_1|}} f_{\{2,3\}}^{-1}(v) = \{2, 3\}. \quad (65)$$

Finally, let us consider an example of function such that multiple rounds of recursion are needed to verify that is pseudo identity.

Example 13 We define function f recursively as follows, where $\mathcal{X}_\ell = \{0, 1\}$ for every $\ell \in \mathcal{L}$. Let $f_{[1]}(x_1) := x_1$. Then, for $\ell = 2, \dots, L$, let

$$f_{[\ell]}(x_1, \dots, x_{\ell-1}, 0) := f_{[\ell-1]}(x_1, \dots, x_{\ell-1}), \quad (66)$$

$$f_{[\ell]}(x_1, \dots, x_{\ell-1}, 1) := \ell. \quad (67)$$

Then, the function $f = f_{[L]} : \mathcal{X}_L \rightarrow \mathcal{V} = \{0, 1, \dots, L\}$ is pseudo identity. In fact, even though f is not injective, we can verify that

$$\bigcup_{v \in \mathcal{V}} \text{span} f^{-1}(v) = \mathcal{L} \setminus \{L\} \subsetneq \mathcal{L}. \quad (68)$$

Similarly, even though $f_{\mathcal{L} \setminus \{\ell+1, \dots, L\}}$ is not injective for $\ell = L-1, \dots, 2$, we can verify that

$$\bigcup_{v \in \mathcal{V}^{2^{L-\ell}}} \text{span} f_{\mathcal{L} \setminus \{\ell+1, \dots, L\}}^{-1}(v) = \{1, \dots, \ell-1\} \subsetneq \{1, \dots, \ell\}. \quad (69)$$

Eventually, we can verify that $f_{\{1\}}$ is injective. The special case for $L = 3$ is the function in Fig. 6(g).

It should be emphasized that the function in Example 13 belongs to $\mathcal{F}_{SW}^{\text{smt}}$ even though the image size $|\mathcal{V}| = L+1$ of the function is much smaller (exponential in the number of terminals) than the input size $|\mathcal{X}_L| = 2^L$.

A. Proof of “if” part of Theorem 3

If f is injective, then it is trivial that $f \in \mathcal{F}_{SW}^{\text{smt}}$; thus, we assume that f is not injective. Then, since f is pseudo identity, there exists a sequence of subsets

$$\mathcal{A}_k \subsetneq \mathcal{A}_{k-1} \subsetneq \dots \subsetneq \mathcal{A}_1 \subsetneq \mathcal{A}_0 = \mathcal{L} \quad (70)$$

for some $k \geq 1$ such that (58) holds with $\mathcal{A} = \mathcal{A}_{i-1}$ and $\tilde{\mathcal{A}} = \mathcal{A}_i$ for $i = 1, \dots, k$, and $f_{\mathcal{A}_k}$ is injective. From this $\{\mathcal{A}_i\}_{i=1}^k$, we shall construct a sequence of partitions $\{\bar{\mathcal{L}}^{(i)}\}_{i=1}^k$ and a sequence of tuples of partitions $\{\bar{\mathcal{X}}_{\mathcal{L}}^{(i)}\}_{i=1}^k$ satisfying the requirement of Corollary 2. In fact, we set

$$\bar{\mathcal{L}}^{(i)} = \{\mathcal{A}_i, \{\ell\} : \ell \in \mathcal{L} \setminus \mathcal{A}_i\} \quad (71)$$

for $i = 1, \dots, k$;

$$\bar{\mathcal{X}}_{\ell}^{(i)} \equiv \mathcal{X}_{\ell} \text{ for } \ell \in \mathcal{L} \setminus \mathcal{A}_{i+1} \text{ and } \bar{\mathcal{X}}_{\ell}^{(i)} = \{\mathcal{X}_{\ell}\} \text{ for } \ell \in \mathcal{A}_{i+1} \quad (72)$$

for $i = 1, \dots, k-1$; and

$$\bar{\mathcal{X}}_{\ell}^{(k)} \equiv \mathcal{X}_{\ell} \text{ for } \ell \in \mathcal{L}. \quad (73)$$

Then, we can verify the requirement of Corollary 2, i.e., $(\mathbf{f}, \mathbf{f}_{\bar{\mathcal{X}}_{\mathcal{L}}^{(1)}}, \dots, \mathbf{f}_{\bar{\mathcal{X}}_{\mathcal{L}}^{(i-1)}})$ induces conditional independence for $(\mathbf{X}_{\mathcal{L}}, \bar{\mathcal{L}}^{(i)})$ and semi $\bar{\mathcal{X}}_{\mathcal{A}}^{(i)}$ -informative for every $\mathcal{A} \in \bar{\mathcal{L}}^{(i)}$ for each $i = 1, \dots, k$, as follows.

For $i = 1$, since

$$\mathcal{A}_1 = \bigcup_{v \in \mathcal{V}} \text{span} f^{-1}(v), \quad (74)$$

the values of x_{ℓ} for $\ell \in \mathcal{L} \setminus \mathcal{A}_1$ are uniquely determined from $f(x_{\mathcal{L}})$. Thus, \mathbf{f} induces conditional independence for $(\mathbf{X}_{\mathcal{L}}, \bar{\mathcal{L}}^{(1)})$. Furthermore, for $\ell \in \mathcal{L} \setminus \mathcal{A}_1$, \mathbf{f} is $\bar{\mathcal{X}}_{\ell}^{(1)}$ -informative.¹⁷ If $k = 1$, then \mathbf{f} is semi $\bar{\mathcal{X}}_{\mathcal{A}_1}^{(1)}$ -informative since $f_{\mathcal{A}_1}$ is injective. If $k \geq 2$, since

$$\mathcal{A}_2 = \bigcup_{v \in \mathcal{V}^{|\mathcal{X}_{\mathcal{A}_1^c}|}} \text{span} f_{\mathcal{A}_1}^{-1}(v), \quad (75)$$

the values of x_{ℓ} for $\ell \in \mathcal{A}_2 \setminus \mathcal{A}_1$ are uniquely determined from $f_{\mathcal{A}_1}(x_{\mathcal{A}_1})$. Furthermore, $\bar{\mathcal{X}}_{\ell}^{(1)} = \{\mathcal{X}_{\ell}\}$ is the trivial partition for $\ell \in \mathcal{A}_1$. Thus, \mathbf{f} is $\bar{\mathcal{X}}_{\mathcal{A}_1}^{(1)}$ -informative.

For $i \geq 2$, since $\bar{\mathcal{X}}_{\ell}^{(i-1)} \equiv \mathcal{X}_{\ell}$ for $\ell \in \mathcal{L} \setminus \mathcal{A}_i$, the values of x_{ℓ} is trivially determined from local function $f_{\bar{\mathcal{X}}_{\mathcal{L}}^{(i-1)}}(x_{\mathcal{L}})$ for $\ell \in \mathcal{L} \setminus \mathcal{A}_i$. Thus, $(\mathbf{f}, \mathbf{f}_{\bar{\mathcal{X}}_{\mathcal{L}}^{(1)}}, \dots, \mathbf{f}_{\bar{\mathcal{X}}_{\mathcal{L}}^{(i-1)}})$ induces conditional independence for $(\mathbf{X}_{\mathcal{L}}, \bar{\mathcal{L}}^{(i)})$. Furthermore, for $\ell \in \mathcal{L} \setminus \mathcal{A}_i$, $(\mathbf{f}, \mathbf{f}_{\bar{\mathcal{X}}_{\mathcal{L}}^{(1)}}, \dots, \mathbf{f}_{\bar{\mathcal{X}}_{\mathcal{L}}^{(i-1)}})$ is semi $\bar{\mathcal{X}}_{\ell}^{(i)}$ -informative¹⁸. If $i = k$, then $(\mathbf{f}, \mathbf{f}_{\bar{\mathcal{X}}_{\mathcal{L}}^{(1)}}, \dots, \mathbf{f}_{\bar{\mathcal{X}}_{\mathcal{L}}^{(i-1)}})$ is semi $\bar{\mathcal{X}}_{\mathcal{A}_i}^{(i)}$ -informative since $f_{\mathcal{A}_i}$ is injective. If $i \leq k-1$, since

$$\mathcal{A}_{i+1} = \bigcup_{v \in \mathcal{V}^{|\mathcal{X}_{\mathcal{A}_i^c}|}} \text{span} f_{\mathcal{A}_i}^{-1}(v), \quad (76)$$

the values of x_{ℓ} for $\ell \in \mathcal{A}_{i+1} \setminus \mathcal{A}_i$ are uniquely determined from $f_{\mathcal{A}_i}(x_{\mathcal{A}_i})$. Furthermore, $\bar{\mathcal{X}}_{\ell}^{(i)} = \{\mathcal{X}_{\ell}\}$ is the trivial partition for $\ell \in \mathcal{A}_{i+1}$. Thus, $(\mathbf{f}, \mathbf{f}_{\bar{\mathcal{X}}_{\mathcal{L}}^{(1)}}, \dots, \mathbf{f}_{\bar{\mathcal{X}}_{\mathcal{L}}^{(i-1)}})$ is $\bar{\mathcal{X}}_{\mathcal{A}_i}^{(i)}$ -informative. ■

¹⁷Note that $\bar{\mathcal{X}}_{\ell}^{(1)}$ for $\ell \in \mathcal{L} \setminus \mathcal{A}_1$ is defined by (72) when $k \geq 2$ and defined by (73) when $k = 1$.

¹⁸Note that $\bar{\mathcal{X}}_{\ell}^{(i)}$ for $\ell \in \mathcal{L} \setminus \mathcal{A}_i$ is defined by (72) when $i \leq k-1$ and defined by (73) when $i = k$.

B. Proof of “only if” part of Theorem 3

Since f is not pseudo identity (cf. Definition 13), either

- (i) there exists $\ell \in \mathcal{L}$ such that $f_{\{\ell\}}$ is not injective;
- (ii) there exists $\mathcal{A} \subseteq \mathcal{L}$ such that

$$\bigcup_{v \in \mathcal{V}^{|\mathcal{A}^c|}} \text{span} f_{\mathcal{A}}^{-1}(v) = \mathcal{A}. \quad (77)$$

We construct a function such that $\mathcal{R}_{\text{SW}}(\mathbf{X}_{\mathcal{L}}) \subsetneq \mathcal{R}(\mathbf{X}_{\mathcal{L}}|\mathbf{f})$ in a similar manner as the proof of “only if part” of two terminal dichotomy theorem in [16, Theorem 3].

When Case (i) is true, the function f violates the necessary condition in Proposition 6. Thus, there exists an i.i.d. source $\mathbf{X}_{\mathcal{L}}$ such that the function computation region $\mathcal{R}(\mathbf{X}_{\mathcal{L}}|\mathbf{f})$ is strictly broader than the Slepian-Wolf region $\mathcal{R}_{\text{SW}}(\mathbf{X}_{\mathcal{L}})$.

When Case (ii) is true, for each $\ell \in \mathcal{A}$, there exist $a_{\mathcal{A}}^{(\ell)}, \hat{a}_{\mathcal{A}}^{(\ell)} \in \mathcal{X}_{\mathcal{A}}$ such that

$$a_{\ell}^{(\ell)} \neq \hat{a}_{\ell}^{(\ell)} \quad (78)$$

and

$$f_{\mathcal{A}}(a_{\mathcal{A}}^{(\ell)}) = f_{\mathcal{A}}(\hat{a}_{\mathcal{A}}^{(\ell)}). \quad (79)$$

By denoting $m = |\mathcal{A}|$, let us consider m -length copy of function f , i.e.,

$$f^m(x_{\mathcal{L}}^m) := (f(x_{\mathcal{L},1}), \dots, f(x_{\mathcal{L},m})) \quad (80)$$

for $x_{\mathcal{L}}^m = (x_{\mathcal{L},1}, \dots, x_{\mathcal{L},m})$. Then, this function f^m violates the necessary condition in Proposition 6 as a function from $\mathcal{X}_{\mathcal{L}}^m$ to \mathcal{V}^m . In fact, because of (78) and (79), the pair $(a_{\mathcal{A}}^{(1)}, \dots, a_{\mathcal{A}}^{(m)})$ and $(\hat{a}_{\mathcal{A}}^{(1)}, \dots, \hat{a}_{\mathcal{A}}^{(m)})$ violate the requirement of the necessary condition. Thus, there exists a block i.i.d. source $\tilde{X}_{\mathcal{L}}^{m,N} = (\tilde{X}_{\mathcal{L}}^m[1], \dots, \tilde{X}_{\mathcal{L}}^m[N])$ on extended alphabet $\mathcal{X}_{\mathcal{L}}^m$ such that $\tilde{\mathbf{X}}_{\mathcal{L}}^m = \{\tilde{X}_{\mathcal{L}}^{m,N}\}_{N=1}^{\infty}$ and $\mathbf{f}^m = \{(f^m)^N\}_{N=1}^{\infty}$ satisfy

$$\mathcal{R}_{\text{SW}}(\tilde{\mathbf{X}}_{\mathcal{L}}^m) \subsetneq \mathcal{R}(\tilde{\mathbf{X}}_{\mathcal{L}}^m|\mathbf{f}^m). \quad (81)$$

Now, let $\mathbf{X} = \tilde{\mathbf{X}}_{\mathcal{L}}^m$, which is a smooth source on $\mathcal{X}_{\mathcal{L}}$,¹⁹ and $\mathbf{f} = \mathbf{f}^m$. Then, (81) implies (with scaling by the common factor $\frac{1}{m}$)

$$\mathcal{R}_{\text{SW}}(\mathbf{X}_{\mathcal{L}}) \subsetneq \mathcal{R}(\mathbf{X}_{\mathcal{L}}|\mathbf{f}). \quad (82)$$

■

¹⁹When $n \neq mN$ for any integer N , we appropriately pad some random variables, which does not affect the rate regions asymptotically.

VIII. CONCLUSION

In this paper, we developed a general approach to derive converse bounds on multiterminal distributed function computing. As an application of the proposed approach, we considered the function classification problem for the class of i.i.d. sources in multiterminal setting, and derived a novel sufficient condition that strictly subsume the sufficient conditions derived by Han-Kobayashi. As we have seen in Example 8, there are some functions that cannot be classified by our approach. Currently, it is not clear whether the sufficient condition is loose or the necessary condition is loose. If the necessary condition could be improved, a novel achievability scheme that is more sophisticated than the Körner-Marton type scheme used in [10] may be needed; possibly, the approaches studied in the literature [15], [12] might be useful.

APPENDIX

A. Technical Lemmas

The following lemma is a simple property of the variational distance (e.g. see [18, Lemma 11.3]).

Lemma 1 For a pair of random variables (X, Y) on the same alphabet, we have

$$\|P_X - P_Y\|_1 \leq \Pr(X \neq Y). \quad (83)$$

The following lemma is an immediate consequence of the alternative definition of the variational distance.

Lemma 2 For a pair of random variables (X, Y) on the same alphabet and any event \mathcal{E} , we have

$$\Pr(X \in \mathcal{E}) \leq \Pr(Y \in \mathcal{E}) + \|P_X - P_Y\|_1. \quad (84)$$

In the converse proof of Theorem 1, we need to boost a code with small symbol error probability to a code with small block error probability. For that purpose, we use the following lemma shown in [17, Lemma 4]. For $0 < \beta < 1/2$ and an integer m , let

$$\nu_n(\beta, m) := \sum_{i=0}^{\lceil n\beta \rceil - 1} (m-1)^i \binom{n}{i} \leq nm^{n\beta} 2^{nh(\beta)}. \quad (85)$$

Lemma 3 [17] Suppose that (W^n, X^n) on $\mathcal{X}^n \times \mathcal{X}^n$ satisfies

$$\Pr\left(\frac{1}{n}d_H(W^n, X^n) \geq \beta\right) \leq \varepsilon_n. \quad (86)$$

Then, there exists encoders $\kappa_n : \mathcal{X}^n \rightarrow \mathcal{K}_n$ with $|\mathcal{K}_n| \leq 2^{n\delta}$ and a decoder $\tau_n : \mathcal{K}_n \times \mathcal{X}^n \rightarrow \mathcal{X}^n$ such that

$$\Pr(\tau_n(\kappa_n(X^n), W^n) \neq X^n) \leq \varepsilon_n + \nu_n(\beta, |\mathcal{X}|)2^{-n\delta}. \quad (87)$$

B. Proof of (37)

The proof is essentially the same as [17, Proof of Theorem 1] (see also [17, Sec. III] for the high level idea of the argument). For convenience of notation, we identify $\mathcal{X}_{\mathcal{A}^c}$ as $\{0, 1, \dots, |\mathcal{X}_{\mathcal{A}^c}| - 1\}$. Let $\pi_i : \mathcal{X}_{\mathcal{A}^c}^n \rightarrow \mathcal{X}_{\mathcal{A}^c}^n$ be the permutation that shift only i th symbol of $\mathbf{x}_{\mathcal{A}^c} \in \mathcal{X}_{\mathcal{A}^c}^n$, i.e., $x_{\mathcal{A}^c, i} \mapsto x_{\mathcal{A}^c, i} + 1 \pmod{|\mathcal{X}_{\mathcal{A}^c}|}$. By using this cyclic permutation π_i , we can rewrite (37) as follows:

$$\Pr \left(\left(\psi_n(\varphi_n^{(\mathcal{A})}(X_{\mathcal{A}}^n), \varphi_n^{(\mathcal{A}^c)}(a_{\mathcal{A}^c} X_{\mathcal{A}^c}^{(-i)})) : a_{\mathcal{A}^c} \in \mathcal{X}_{\mathcal{A}^c} \right) \neq \left(f_n(X_{\mathcal{A}}^n, a_{\mathcal{A}^c} X_{\mathcal{A}^c}^{(-i)}) : a_{\mathcal{A}^c} \in \mathcal{X}_{\mathcal{A}^c} \right) \right) \quad (88)$$

$$= \Pr \left(\exists a_{\mathcal{A}^c} \in \mathcal{X}_{\mathcal{A}^c} \text{ s.t. } \psi_n(\varphi_n^{(\mathcal{A})}(X_{\mathcal{A}}^n), \varphi_n^{(\mathcal{A}^c)}(a_{\mathcal{A}^c} X_{\mathcal{A}^c}^{(-i)})) \neq f_n(X_{\mathcal{A}}^n, a_{\mathcal{A}^c} X_{\mathcal{A}^c}^{(-i)}) \right) \quad (89)$$

$$= \sum_{\mathbf{x}_{\mathcal{A}}, \mathbf{x}_{\mathcal{A}^c}} P_{X_{\mathcal{L}}^n}(\mathbf{x}_{\mathcal{A}}, \mathbf{x}_{\mathcal{A}^c}) \mathbf{1} \left[\exists a_{\mathcal{A}^c} \in \mathcal{X}_{\mathcal{A}^c} \text{ s.t. } \psi_n(\varphi_n^{(\mathcal{A})}(\mathbf{x}_{\mathcal{A}}), \varphi_n^{(\mathcal{A}^c)}(a_{\mathcal{A}^c} \mathbf{x}_{\mathcal{A}^c}^{(-i)})) \neq f_n(\mathbf{x}_{\mathcal{A}}, a_{\mathcal{A}^c} \mathbf{x}_{\mathcal{A}^c}^{(-i)}) \right] \quad (90)$$

$$= \sum_{\mathbf{x}_{\mathcal{A}}, \mathbf{x}_{\mathcal{A}^c}} P_{X_{\mathcal{L}}^n}(\mathbf{x}_{\mathcal{A}}, \mathbf{x}_{\mathcal{A}^c}) \mathbf{1} \left[\exists b \in \mathcal{X}_{\mathcal{A}^c} \text{ s.t. } \psi_n(\varphi_n^{(\mathcal{A})}(\mathbf{x}_{\mathcal{A}}), \varphi_n^{(\mathcal{A}^c)}(\pi_i^b(\mathbf{x}_{\mathcal{A}^c}))) \neq f_n(\mathbf{x}_{\mathcal{A}}, \pi_i^b(\mathbf{x}_{\mathcal{A}^c})) \right] \quad (91)$$

$$\leq \sum_{b \in \mathcal{X}_{\mathcal{A}^c}} \sum_{\mathbf{x}_{\mathcal{A}}, \mathbf{x}_{\mathcal{A}^c}} P_{X_{\mathcal{L}}^n}(\mathbf{x}_{\mathcal{A}}, \mathbf{x}_{\mathcal{A}^c}) \mathbf{1} \left[\psi_n(\varphi_n^{(\mathcal{A})}(\mathbf{x}_{\mathcal{A}}), \varphi_n^{(\mathcal{A}^c)}(\pi_i^b(\mathbf{x}_{\mathcal{A}^c}))) \neq f_n(\mathbf{x}_{\mathcal{A}}, \pi_i^b(\mathbf{x}_{\mathcal{A}^c})) \right], \quad (92)$$

where π_i^b means b times application of π_i . Then, since $X_{\mathcal{L}}$ is smooth, for every $b \in \mathcal{X}_{\mathcal{A}^c}$ and $i \in [1 : n]$, we can bound each term of (92) as follows:

$$\sum_{\mathbf{x}_{\mathcal{A}}, \mathbf{x}_{\mathcal{A}^c}} P_{X_{\mathcal{L}}^n}(\mathbf{x}_{\mathcal{A}}, \mathbf{x}_{\mathcal{A}^c}) \mathbf{1} \left[\psi_n(\varphi_n^{(\mathcal{A})}(\mathbf{x}_{\mathcal{A}}), \varphi_n^{(\mathcal{A}^c)}(\pi_i^b(\mathbf{x}_{\mathcal{A}^c}))) \neq f_n(\mathbf{x}_{\mathcal{A}}, \pi_i^b(\mathbf{x}_{\mathcal{A}^c})) \right] \quad (93)$$

$$\leq \sum_{\mathbf{x}_{\mathcal{A}}, \mathbf{x}_{\mathcal{A}^c}} \frac{1}{q} P_{X_{\mathcal{L}}^n}(\mathbf{x}_{\mathcal{A}}, \pi_i^b(\mathbf{x}_{\mathcal{A}^c})) \mathbf{1} \left[\psi_n(\varphi_n^{(\mathcal{A})}(\mathbf{x}_{\mathcal{A}}), \varphi_n^{(\mathcal{A}^c)}(\pi_i^b(\mathbf{x}_{\mathcal{A}^c}))) \neq f_n(\mathbf{x}_{\mathcal{A}}, \pi_i^b(\mathbf{x}_{\mathcal{A}^c})) \right] \quad (94)$$

$$= \frac{1}{q} \Pr \left(\psi_n(\varphi_n^{(\mathcal{A})}(X_{\mathcal{A}}^n), \varphi_n^{(\mathcal{A}^c)}(X_{\mathcal{A}^c}^n)) \neq f_n(X_{\mathcal{A}}^n, X_{\mathcal{A}^c}^n) \right) \quad (95)$$

$$\leq \frac{\varepsilon_n}{q}. \quad (96)$$

Substituting this bound into (92), we have the desired bound. \blacksquare

ACKNOWLEDGEMENTS

The author would like to thank Shigeaki Kuzuoka for helpful comments.

REFERENCES

- [1] R. Ahlswede and I. Csiszár, “To get a bit of information may be as hard as to get full information,” *IEEE Trans. Inform. Theory*, vol. 27, no. 4, pp. 398–408, July 1981.
- [2] 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 Trans. Inform. Theory*, vol. 29, no. 3, pp. 396–412, May 1983.
- [3] R. Ahlswede and J. Körner, “Source coding with side information and a converse for the degraded broadcast channel,” *IEEE Trans. Inform. Theory*, vol. 21, no. 6, pp. 629–637, November 1975.
- [4] T. Berger, Z. Zhang, and H. Viswanathan, “The CEO problem,” *IEEE Trans. Inform. Theory*, vol. 42, no. 3, pp. 887–902, May 1996.
- [5] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. John Wiley & Sons, 2006.
- [6] I. Csiszár and J. Körner, *Information Theory, Coding Theorems for Discrete Memoryless Systems*, 2nd ed. Cambridge University Press, 2011.

- [7] A. ElGamal, "A simple proof of the Ahlswede-Csiszár one-bit theorem," *IEEE Trans. Inform. Theory*, vol. 29, no. 6, pp. 931–933, November 1983.
- [8] S. I. Gel'fand and M. S. Pinsker, "Coding of sources on the basis of observations with incomplete information," *Problems of Information Transmission*, vol. 15, pp. 115–125, 1979.
- [9] T. S. Han, *Information-Spectrum Methods in Information Theory*. Springer, 2003.
- [10] 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 Trans. Inform. Theory*, vol. 33, no. 1, pp. 69–76, January 1987.
- [11] T. S. Han and S. Verdú, "Approximation theory of output statistics," *IEEE Trans. Inform. Theory*, vol. 39, no. 3, pp. 752–772, May 1993.
- [12] S. Huang and M. Skoglund, "On linear coding over finite rings and applications to computing," *Entropy*, vol. 19, no. 5, p. 233, May 2017.
- [13] J. Körner and K. Marton, "How to encode the modulo-two sum of binary sources," *IEEE Trans. Inform. Theory*, vol. 25, no. 2, pp. 219–221, March 1979.
- [14] H. Kowshik and P. R. Kumar, "Optimal function computation in directed and undirected graphs," *IEEE Trans. Inform. Theory*, vol. 58, no. 6, pp. 3407–3418, June 2012.
- [15] D. Krithivasan and S. Sandeep, "Distributed source coding using abelian group codes: A new achievable rate-distortion region," *IEEE Trans. Inform. Theory*, vol. 57, no. 3, pp. 1495–1519, March 2011.
- [16] S. Kuzuoka and S. Watanabe, "A dichotomy of functions in distributed coding: An information-spectrum approach," *IEEE Trans. Inform. Theory*, vol. 61, no. 9, pp. 5028–5041, September 2015.
- [17] —, "On distributed computing for functions with certain structures," *IEEE Trans. Inform. Theory*, vol. 63, no. 11, pp. 7003–7017, November 2017.
- [18] M. Mitzenmacher and E. Upfal, *Probability and Computing*. Cambridge University Press, 2005.
- [19] S. Miyake and F. Kanaya, "Coding theorems on correlated general sources," *IEICE Trans. Fundamentals*, vol. E78-A, no. 9, pp. 1063–1070, September 1995.
- [20] Y. Oohama, "The rate-distortion function for the quadratic Gaussian CEO problem," *IEEE Trans. Inform. Theory*, vol. 44, no. 3, pp. 1057–1070, May 1998.
- [21] —, "Rate-distortion theory for Gaussian multiterminal source coding systems with several side information at the decoder," *IEEE Trans. Inform. Theory*, vol. 51, no. 7, pp. 2577–2593, July 2005.
- [22] —, "Gaussian multiterminal source coding with several side informations at the decoder," in *Proceedings of IEEE International Symposium on Information Theory*, Seattle, USA, July 2006, pp. 1409–1413.
- [23] —, "Many-help-one problem for Gaussian sources with a tree structure on their correlation," 2009, arXiv:0901.1988.
- [24] A. Orlitsky and R. Roche, "Coding for computing," *IEEE Trans. Inform. Theory*, vol. 47, no. 3, pp. 903–917, March 2001.
- [25] V. Prabhakaran, D. Tse, and K. Ramchandran, "Rate region of the quadratic Gaussian CEO problem," in *Proceedings of IEEE International Symposium on Information Theory*, Chicago, USA, June 2004, p. 117.
- [26] M. Sefidgaran, A. Gohari, and M. R. Aref, "On Körner-Martón's sum modulo two problem," in *Proceedings of 2015 Iran Workshop on Communication and Information Theory (IWCIT)*, Tehran, Iran, May 2015.
- [27] M. Sefidgaran and A. Tchamkerten, "Distributed function computation over a rooted directed tree," *IEEE Trans. Inform. Theory*, vol. 62, no. 12, pp. 7135–7152, December 2016.
- [28] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inform. Theory*, vol. 19, no. 4, pp. 471–480, July 1973.
- [29] S. Tavildar, P. Viswanath, and A. B. Wagner, "The Gaussian many-help-one distributed source coding problem," *IEEE Trans. Inform. Theory*, vol. 56, no. 1, pp. 564–581, January 2010.
- [30] H. Viswanathan and T. Berger, "The quadratic Gaussian CEO problem," *IEEE Trans. Inform. Theory*, vol. 43, no. 5, pp. 1549–1559, September 1997.
- [31] A. B. Wagner and V. Anantharam, "An improved outer bound for multiterminal source coding," *IEEE Trans. Inform. Theory*, vol. 54, no. 5, pp. 1919–1937, May 2008.
- [32] A. B. Wagner, S. Tavildar, and P. Viswanath, "Rate region of the quadratic Gaussian two-encoder source-coding problem," *IEEE Trans. Inform. Theory*, vol. 54, no. 5, pp. 1938–1961, May 2008.

- [33] A. D. Wyner, "On source coding with side information at the decoder," *IEEE Trans. Inform. Theory*, vol. 21, no. 3, pp. 294–300, May 1975.