Generalized Lexicographic Products and the Index Coding Capacity

Fatemeh Arbabjolfaei and Young-Han Kim, *Fellow, IEEE*

***Abstract*— The index coding problem studies the fundamental limit on broadcasting multiple messages to their respective receivers with different sets of side information that are represented by a directed graph. The generalized lexicographic product structure in the side information graph is introduced as a natural condition under which the corresponding index coding problem can be decomposed into multiple interacting subproblems, each consisting of vertices with the same adjacency pattern with respect to other subproblems. For side information graphs with this structure, the capacity region is characterized in terms of the subproblem capacity regions combined in the same product structure. The proof is based on dual uses of random coding—one for a new multiletter characterization of the capacity region of a general index coding problem via joint typicality decoding and the other for a construction of a new multiletter code of matching rates from a single-letter code via joint typicality encoding. Several special cases are discussed that recover and strengthen known structural properties of the index coding capacity region.**

***Index Terms*— Capacity region, broadcast rate, directed graph, random coding, multiletter characterization.**

I. Introduction

INDEX coding is a canonical problem in network information theory, in which a server broadcasts a tuple of n messages $x^n = (x_1, \ldots, x_n)$, $x_i \in \{0,1\}^{t_i}$, to n receivers by transmitting the fewest number of bits possible over a noiseless broadcast channel (see Fig. 1). Receiver $i \in [n] := \{1, 2, \ldots, n\}$ is interested in message x_i and has a set of other messages $x(A_i) := (x_j, j \in A_i), A_i \subseteq [n] \setminus \{i\}$, as *side information*. The *side information sets* $A_1, \ldots, A_n$ are known to all communicating parties. We represent the side information sets compactly by a sequence $(i|A_i), i \in [n]$. For example, the 3-message index coding problem with $A_1 = \{2,3\}, A_2 = \{1\}$, and $A_3 = \{1,2\}$ is represented as $(1|2,3), (2|1), (3|1,2)$. Each index coding problem can be also uniquely specified by a (finite, simple) directed graph with n vertices, referred to as the *side information graph*. Each vertex of the side information graph $G = (V, E)$ corresponds to a receiver (and its desired message) and there is a directed edge $j \to i$ if and only if (iff) receiver i knows message x_j as side information, i.e., $j \in A_i$ (see Fig. 2). Throughout the paper, we identify an instance of the index coding problem with its side information graph G and often write "index coding problem G."

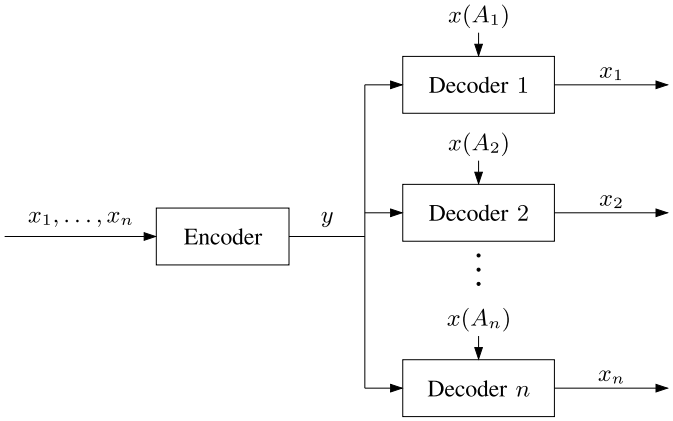

Fig. 1. The index coding problem.

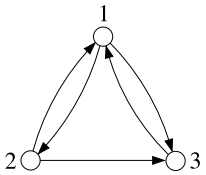

Fig. 2. The graph representation for the index coding problem with $A_1 = \{2,3\}, A_2 = \{1\}$, and $A_3 = \{1,2\}$.

We formulate the index coding problem more precisely by a $(t_1, \ldots, t_n, r)$ *index code* that consists of

- an encoder $\phi : \prod_{j\in[n]}\{0,1\}^{t_j} \to \{0,1\}^r$ that maps the message n-tuple x^n to an r-bit sequence y and
- n decoders, where the decoder at receiver $i \in [n]$, $\psi_i : \{0,1\}^r \times \prod_{j\in A_i}\{0,1\}^{t_j} \to \{0,1\}^{t_i}$, maps the received sequence y and the side information $x(A_i)$ back to x_i.

Thus, for every $x^n \in \prod_{j\in[n]}\{0,1\}^{t_j}$,

$$\psi_i(\phi(x^n), x(A_i)) = x_i, \quad i \in [n]. \tag{1}$$

Sometimes a $(t_1, \ldots, t_n, r)$ code will be written in short as a $(\mathbf{t}, r)$ code and a $(t, \ldots, t, r)$ code will be written in short as a (t, r) code.

Manuscript received August 11, 2016; revised April 8, 2019; accepted October 2, 2019. Date of publication December 24, 2019; date of current version February 14, 2020. This work was supported in part by the National Science Foundation under Grant CCF-1320895 and in part by the Korean Ministry of Science, ICT and Future Planning through the Institute for Information and Communications Technology Promotion (Development of Wired-Wireless Converged 5G Core Technologies) under Grant B0132-15-1005.

Fatemeh Arbabjolfaei is with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109 USA (e-mail: arbab@umich.edu).

Young-Han Kim is with the Department of Electrical and Computer Engineering, University of California San Diego, La Jolla, CA 92093 USA (e-mail: yhk@ucsd.edu).

Communicated by S. Watanabe, Associate Editor for Shannon Theory.

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2019.2961919

0018-9448 © 2019 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See https://www.ieee.org/publications/rights/index.html for more information.

Authorized licensed use limited to: Univ of Calif San Diego. Downloaded on February 15,2020 at 06:37:11 UTC from IEEE Xplore. Restrictions apply.

A rate tuple $\mathbf{R} = (R_1, \ldots, R_n)$ is said to be *achievable* for the index coding problem G if there exists a $(\mathbf{t}, r)$ index code such that

$$R_i \leq \frac{t_i}{r}, \quad i \in [n],$$

or equivalently, in vector notation,

$$\mathbf{R} \leq \frac{\mathbf{t}}{r}.$$

Here and henceforth, we write $\mathbf{a} \leq \mathbf{b}$ for vectors $\mathbf{a} = (a_1, \ldots, a_n)$ and $\mathbf{b} = (b_1, \ldots, b_n)$ of the same length n iff $a_i \leq b_i$, $i \in [n]$. The *capacity region* $\mathscr{C}(G)$ of the index coding problem G is defined as the closure of the set of achievable rate tuples. The *symmetric capacity* (or the *capacity* in short) of the index coding problem G is defined as

$$C_{\mathrm{sym}}(G) = \max\{R\colon (R, \ldots, R) \in \mathscr{C}(G)\}.$$

The reciprocal of the symmetric capacity, $\beta(G) = 1/C_{\mathrm{sym}}(G)$, is referred to as the *broadcast rate*.

Remark 1. The achievability of a given rate tuple can be defined alternatively by relaxing the decoding condition in (1) as

$$\lim_{r\to\infty} \mathsf{P}\{\psi_i(\phi(X^n), X(A_i)) \neq X_i,\ i \in [n]\} = 0,$$

where $X_1, \ldots, X_n$ are distributed independently and uniformly at random. The corresponding *vanishing-error capacity region* can be shown [1] (see also [2, Problem 8.11]) to be identical to the *zero-error capacity region* defined earlier, which holds in general for any single-sender network. This identity was also established in [3], [4] in the context of index coding and single-sender network coding.

The problem of broadcasting to multiple receivers with different side information traces back to the work by Celebiler and Stette [5], Willems *et al.* [6] and Wyner *et al.* [7], Yeung [8], and Birk and Kol [9], [10]. The current problem formulation is due to the last. This problem has been shown to be closely related to many other important problems in network information theory such as network coding [11]–[13], locally recoverable distributed storage [14]–[16], guessing games on directed graphs [11], [16], [17], and zero-error capacity of channels [18]. In addition, index coding has its own applications in diverse areas ranging from satellite communication [5]–[10] and multimedia distribution [19] to interference management [20] and coded caching [21], [22]. Due to this significance, the index coding problem has been studied extensively over the past two decades. We refer the reader to the dissertations of El Rouayheb [23], Blasiak [24], and the first author [25], a survey article by Byrne and Calderini [26], and a recent monograph by the authors [27].

The main information-theoretic question in studying the index coding problem is to characterize the capacity region in a computable expression. There are several inner and outer bounds on the capacity region that are tight for several interesting special cases, but the capacity region of a general n-message index coding problem is open (that is, no computable characterization is known). So far the capacity region has been characterized for all index coding problems with $n \leq 5$ messages [28]. For $n \geq 6$, the capacity region is not known in general.

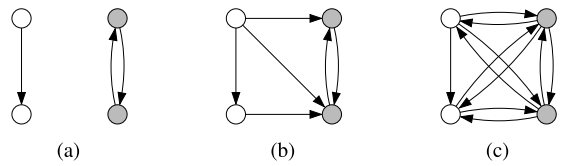

Fig. 3. Side information graphs with (a) no interaction, (b) one-way interaction, and (c) complete two-way interaction among the two parts (white and gray).

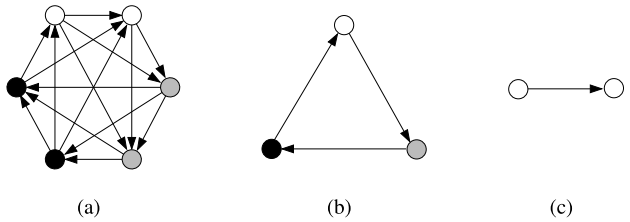

Fig. 4. (a) The lexicographic product $G_0 \circ G_1$ of (b) the 3-vertex graph G_0 and (c) the 2-vertex graph G_1.

For some cases, however, the side information can be decomposed into subgraphs with some connectivity (interaction) pattern, and this structure can be used to characterize the capacity region in terms of those of the subproblems. Consider the three side information graphs illustrated in Fig. 3, in which each graph has two parts and the interaction between them is none, one-way, and complete two-way. These *union* structures were investigated earlier in [29]–[31], and it was shown that the capacity region of a given index coding problem is characterized as the "sum" of the subproblem capacity regions for the first two cases [30], [31], and as the "maximum" of the subproblem capacity regions for the third case [31]; see Sections III-A through III-C for details.

As another example, consider the side information graph in Fig. 4(a), which can be generated by replacing each vertex of the graph in Fig. 4(b) by the graph in Fig. 4(c). This *lexicographic product* structure was investigated in [29]. Although the capacity in this case was not characterized in terms of those of the subproblems, a "product" of the capacities of the subproblems was shown to be a nontrivial lower bound on the capacity, and this lower bound was utilized to establish a strong separation result between different capacity bounds [29].

This paper identifies the *generalized lexicographic product* structure as a natural decomposition of the side information graph into subgraphs, which includes the aforementioned union and product structures studied in [29]–[31] as simple special cases. The main contribution, presented in Theorem 1 in the next section, shows that the capacity region of a generalized lexicographic product has a natural lexicographic product structure itself, and can be characterized in terms of the subgraph capacity regions as well as the capacity region of the connectivity graph. Although this generalized lexicographic product structure is rather special, its relaxation

Authorized licensed use limited to: Univ of Calif San Diego. Downloaded on February 15,2020 at 06:37:11 UTC from IEEE Xplore. Restrictions apply.

in Corollary 1 provides straightforward inner and outer bounds on the capacity region for general side information graphs.

The proof of Theorem 1 uses standard Shannon-theoretic arguments. The main challenge is the proof of the converse, which relies on two key ideas. The first idea, Theorem 2, is a construction of an index code based on Shannon's *random coding* and *joint typicality decoding*, the achievable rate region of which is characterized as a multiletter expression by the *packing lemma*. The second idea, Lemma 2, is a construction of a new multiletter index code with relaxed decoding conditions from a single-letter index code, which is based on random coding and joint typicality *encoding*. The achievable rate region of this code is characterized by the *covering lemma*. The converse proof matches the corresponding rate regions from the two ideas carefully to establish the desired structure of the capacity region.

The rest of the paper is organized as follows. Section II introduces the generalized lexicographic product of graphs and presents the capacity region of a generalized lexicographic product in terms of those of the subgraphs (Theorem 1). Section III presents several examples and special cases of Theorem 1 and its relaxation (Corollary 1). Section IV establishes the Shannon-theoretic multiletter characterization of the capacity region, which may be of independent interest. Section V presents the proof of Theorem 1. Section VI concludes the paper with a few remarks on applications of the main result. Technical proofs used in the proof of the converse are relegated to the Appendices.

II. Main Result

In this section, we first define the generalized lexicographic product of graphs and then state the main theorem of the paper.

A. Generalized Lexicographic Product of Graphs

Consider the following graph product, first considered by Schwenk [32] in the context of spectral graph theory.

Definition 1 (Generalized lexicographic product [32], [33])**.** Let $G_0 = (V(G_0), E(G_0))$ be a directed graph on m vertices and let $G_i = (V(G_i), E(G_i))$, $i \in [m]$, be directed graphs on disjoint sets of vertices, i.e., $V(G_i) \cap V(G_j) = \emptyset$, $i \neq j$. The *generalized lexicographic product* $G = G_0 \circ (G_1, \ldots, G_m)$ is defined by the set of vertices $V(G) = \cup_{i\in[m]} V(G_i)$ and the set of edges $E(G)$ consisting of directed edges (i, j) such that

$$i, j \in V(G_k) \text{ for some } k \text{ and } (i, j) \in E(G_k)$$

or

$$i \in V(G_k), j \in V(G_l) \text{ for some } k \neq l \text{ and } (k, l) \in E(G_0).$$

In other words, vertex $i \in V(G_0)$ is replaced by a copy of G_i and every vertex in the copy of G_k is connected to every vertex in the copy of G_l according to $E(G_0)$; see Fig. 5 for an illustration.

Remark 2. This notion of generalized lexicographic product extends that of lexicographic product $G_0 \circ G_1$ [34], [35], which

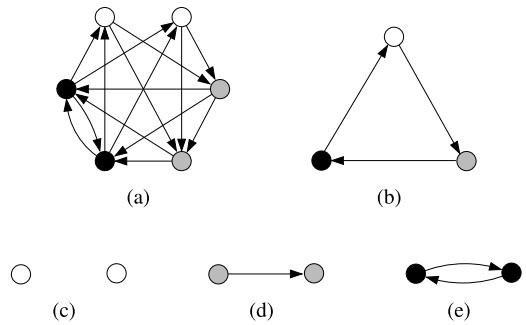

Fig. 5. (a) A 6-vertex graph that is the generalized lexicographic product $G_0 \circ (G_1, G_2, G_3)$, (b) the 3-vertex graph G_0, (c) the 2-vertex graph G_1, (d) the 2-vertex graph G_2, and (e) the 2-vertex graph G_3.

is a graph with vertex set $V(G_0) \times V(G_1)$ and (i_1, i_2) is connected to (j_1, j_2) iff

$$(i_1, j_1) \in E(G_0) \quad \text{or} \quad (i_1 = j_1 \text{ and } (i_2, j_2) \in E(G_1)).$$

By relabeling the vertices, $G_0 \circ G_1 = G_0 \circ (G_1^{(1)}, \ldots, G_1^{(m)})$, where $G_1^{(1)}, \ldots, G_1^{(m)}$ are copies of G_1 over disjoint vertex sets.

Remark 3. To verify whether a graph $G = ([n], E)$ is a generalized lexicographic product of smaller graphs, it suffices to go over all subsets of vertices $S \subseteq [n]$ with $2 \leq |S| \leq n-1$ and check if the vertices in S have the same adjacency pattern with respect to all the vertices in $[n] \setminus S$.

B. Capacity Region of a Generalized Lexicographic Product

The main contribution of the paper is the following characterization of the capacity region of the index coding problem $G_0 \circ (G_1, \ldots, G_m)$ in terms of the capacity regions of smaller problems $G_0, G_1, \ldots, G_m$.

Theorem 1. *Let $G_0 = ([m], E)$ be the side information graph of an index coding problem with m messages and capacity region $\mathscr{C}_0$. Let $G_1, \ldots, G_m$ be the side information graphs of m index coding problems with capacity regions $\mathscr{C}_1, \ldots, \mathscr{C}_m$, respectively. Then the capacity region of the index coding problem with side information graph $G = G_0 \circ (G_1, \ldots, G_m)$ is*

$$\begin{aligned}\mathscr{C}(G) &= \mathscr{C}_0 \circ (\mathscr{C}_1, \ldots, \mathscr{C}_m) \\ &:= \left\{(\rho_1 \mathbf{R}_1, \ldots, \rho_m \mathbf{R}_m) \colon \boldsymbol{\rho} \in \mathscr{C}_0,\ \mathbf{R}_i \in \mathscr{C}_i, i \in [m]\right\}\end{aligned} \tag{2}$$

and its broadcast rate is

$$\beta(G) = \min_{R:(R,\ldots,R)\in\mathscr{C}(G)} \frac{1}{R}. \tag{3}$$

Remark 4. Since $\mathscr{C}_0, \mathscr{C}_1, \ldots, \mathscr{C}_m$ are compact, so is the RHS of (2).

Remark 5. If $\mathscr{C}_0, \mathscr{C}_1, \ldots, \mathscr{C}_m$ are polytopes of the form $\mathscr{C}_i = \{\mathbf{R} : T_i \mathbf{R} \leq \mathbf{1} = (1, \ldots, 1)^T\}$, $i = 0, 1, \ldots, m$, then $\mathscr{C}$ is also a polytope characterized by Fourier–Motzkin elimination

Authorized licensed use limited to: Univ of Calif San Diego. Downloaded on February 15,2020 at 06:37:11 UTC from IEEE Xplore. Restrictions apply.

of m variables $\boldsymbol{\rho} = (\rho_1, \ldots, \rho_m)$ from the linear inequalities

$$T_0\boldsymbol{\rho} \leq \mathbf{1},$$
$$T_i\mathbf{R}_i \leq \rho_i\mathbf{1}, \quad i \in [m].$$

Remark 6. Theorem 1 can be specialized to the broadcast rate of $G = G_0 \circ (G_1, \ldots, G_m)$. If $\boldsymbol{\beta} = (\beta(G_1), \ldots, \beta(G_m))$, then

$$\beta(G) = \frac{1}{C_0(\boldsymbol{\beta})} \leq \max_{i\in[m]} \beta(G_0)\beta(G_i),$$

where $C_0(\boldsymbol{\beta}) = \max\{R\colon R\boldsymbol{\beta} \in \mathscr{C}_0\}$.

The following sandwich argument extends the application of Theorem 1 beyond index coding instances with side information graph in the form of a generalized lexicographic product.

Corollary 1. For $i = 0, 1, \ldots, m$, let G_i' and G_i'' be side information graphs of index coding problems with capacity regions $\mathscr{C}_i'$ and $\mathscr{C}_i''$, respectively, such that $V(G_i') = V(G_i'')$ and $E(G_i') \subseteq E(G_i'')$. Suppose that $|V(G_0')| = |V(G_0'')| = m$ and let

$$G' = G_0' \circ (G_1', \ldots, G_m')$$

and

$$G'' = G_0'' \circ (G_1'', \ldots, G_m'').$$

Then the capacity region of any index coding problem G such that

$$V(G) = V(G') = V(G'')$$

and

$$E(G') \subseteq E(G) \subseteq E(G'')$$

is bounded as

$$\begin{aligned}\mathscr{C}_0' \circ (\mathscr{C}_1', \ldots, \mathscr{C}_m') &= \mathscr{C}(G') \\ &\subseteq \mathscr{C}(G) \\ &\subseteq \mathscr{C}(G'') = \mathscr{C}_0'' \circ (\mathscr{C}_1'', \ldots, \mathscr{C}_m'').\end{aligned}$$

In particular, if $\mathscr{C}_i' = \mathscr{C}_i'' = \mathscr{C}_i$, $i = 0, 1, \ldots, m$, then

$$\mathscr{C}(G) = \mathscr{C}(G') = \mathscr{C}(G'') = \mathscr{C}_0 \circ (\mathscr{C}_1, \ldots, \mathscr{C}_m).$$

Remark 7. For any side information graph G, the bounding graphs G' and G'' can be easily constructed by considering any vertex subset S, say $[k]$, with $2 \leq |S| = k \leq n-1$, and taking the intersection and union of the neighbors from/to S to/from $[n]\backslash S$, respectively. Now that the adjacency pattern is the same for all vertices in S, we can identify G_0' and G_0'' by replacing $G|_S$ with a single vertex and keeping the other vertices. The resulting G' and G'' are generalized lexicographic products of $n - k + 1$ graphs.

Remark 8. An index coding problem is said to be *critical* if removal of any of the edges of its side information graph strictly reduces the capacity region [30], [31]. Note that in Corollary 1, $\mathscr{C}_i' = \mathscr{C}_i''$, $i = 0, 1, \ldots, m$, implies that the index coding problem G is not critical, as those edges of the side information graph G that are not in G' can be removed from G without reducing the capacity region. Thus, Corollary 1 provides a necessary condition for criticality of a side information graph (see [27] for other necessary conditions).

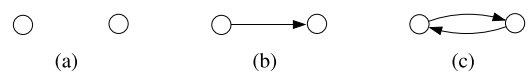

Fig. 6. (a) A 2-vertex graph with no edge, (b) a 2-vertex graph with one edge, and (c) a 2-vertex graph with two edges.

III. Examples

A. No Interaction Between Partitions

Consider the side information graph G depicted in Fig. 3(a), which has two noninteracting parts G_1 and G_2, i.e., there is no edge between G_1 and G_2. Then G can be viewed as $G_0 \circ (G_1, G_2)$, where G_0 is the two-vertex graph in Fig. 6(a). Since the capacity region of G_0 is $\{(R_1, R_2)\colon R_1 + R_2 \leq 1\}$, by Theorem 1,

$$\mathscr{C}(G) = \big\{(\rho\mathbf{R}_1, (1-\rho)\mathbf{R}_2)\colon \mathbf{R}_1 \in \mathscr{C}(G_1), \mathbf{R}_2 \in \mathscr{C}(G_2), \rho \in [0,1]\big\}. \tag{4}$$

Moreover, the maximum symmetric rate in $\mathscr{C}(G)$ is attained when $\rho/\beta(G_1) = (1-\rho)/\beta(G_2)$, or equivalently, $\rho = \beta(G_1)/(\beta(G_1) + \beta(G_2))$, which implies

$$\beta(G) = \beta(G_1) + \beta(G_2). \tag{5}$$

More generally, consider a side information graph G that consists of m vertex-induced subgraphs $G_1, \ldots, G_m$ with no edges among them. Then G can be viewed as $G_0 \circ (G_1, \ldots, G_m)$, where G_0 is a graph with m vertices and no edge. By Theorem 1 (or by applying (4) and (5) inductively),

$$\mathscr{C}(G) = \left\{(\rho_1\mathbf{R}_1, \ldots, \rho_m\mathbf{R}_m)\colon \mathbf{R}_i \in \mathscr{C}(G_i), i \in [m], \sum_{i\in[m]} \rho_i \leq 1\right\}$$

and

$$\beta(G) = \sum_{i\in[m]} \beta(G_i).$$

In other words, when G is partitioned into noninteracting parts $G_1, \ldots, G_m$, the capacity region of G is achieved by *time division* among the optimal coding schemes for subproblems $G_1, \ldots, G_m$ [30].

B. One-way Interaction Between Partitions

Consider the side information graph G depicted in Fig. 3(b), which has one-way interaction between its two parts G_1 and G_2, i.e., there is no edge from G_2 to G_1. Let G_0' and G_0'' be the two graphs on two vertices as depicted in Figs. 6(a) and (b), respectively. Then, $\mathscr{C}(G_0') = \mathscr{C}(G_0'') = \{(R_1, R_2)\colon R_1 + R_2 \leq 1\}$ and $E(G_0' \circ (G_1, G_2)) \subseteq E(G) \subseteq E(G_0'' \circ (G_1, G_2))$. Thus, by Corollary 1,

$$\mathscr{C}(G) = \big\{(\rho\mathbf{R}_1, (1-\rho)\mathbf{R}_2)\colon \mathbf{R}_1 \in \mathscr{C}(G_1), \mathbf{R}_2 \in \mathscr{C}(G_2), \rho \in [0,1]\big\}$$

Authorized licensed use limited to: Univ of Calif San Diego. Downloaded on February 15,2020 at 06:37:11 UTC from IEEE Xplore. Restrictions apply.

and

$$\beta(G) = \beta(G_1) + \beta(G_2).$$

More generally, suppose that the graph G consists of m vertex-induced subgraphs $G_1, \ldots, G_m$ such that there exists no edge from G_j to G_i for $i < j$. Let G_0' and G_0'' be directed graphs with m vertices such that $E(G_0') = \emptyset$ and $E(G_0'') = \{(i,j)\colon i < j\}$. Note that $\mathscr{C}(G_0') = \mathscr{C}(G_0'') = \{(R_1, \ldots, R_m)\colon \sum_{i\in[m]} R_i \leq 1\}$. Since $E(G_0' \circ (G_1, \ldots, G_m)) \subseteq E(G) \subseteq E(G_0'' \circ (G_1, \ldots, G_m))$, by Corollary 1,

$$\mathscr{C}(G) = \Bigg\{(\rho_1 \mathbf{R}_1, \ldots, \rho_m \mathbf{R}_m)\colon \mathbf{R}_i \in \mathscr{C}(G_i), i \in [m], \sum_{i\in[m]} \rho_i \leq 1\Bigg\}$$

and

$$\beta(G) = \sum_{i\in[m]} \beta(G_i).$$

In words, the capacity region of a graph with one-way interaction among its parts is no larger than the capacity region of a graph with noninteracting parts. Thus the edges connecting the parts $G_1, \ldots, G_m$ in one way, or equivalently by the Farkas lemma [36, Th. 2.2], the edges that are not on a directed cycle can be removed without affecting the capacity region (cf. Remark 8) and the graph is not critical [30].

C. Complete Two-way Interaction Between Partitions

Consider the side information graph G in Fig. 3(c). Since there are two-way edges between every vertex in G_1 and every vertex in G_2, G can be written as $G_0 \circ (G_1, G_2)$, where G_0 is the complete graph with two vertices depicted in Fig. 6(c). More generally, suppose that G_0 is a complete graph with m vertices. Then its capacity region is characterized as $\mathscr{C}(G_0) = \{(R_1, \ldots, R_m)\colon R_i \leq 1, i \in [m]\}$. Thus, by Theorem 1, the capacity region of $G = G_0 \circ (G_1, \ldots, G_m)$ is

$$\mathscr{C}(G) = \big\{(\mathbf{R}_1, \ldots, \mathbf{R}_m)\colon \mathbf{R}_i \in \mathscr{C}(G_i), i \in [m]\big\}. \quad (6)$$

Moreover, (6) implies

$$\begin{aligned}\max\{R : (R, \ldots, R) \in \mathscr{C}(G)\} &= \min_{i\in[m]} \max\{R : (R, \ldots, R) \in \mathscr{C}_i\} \\ &= \min_{i\in[m]} \frac{1}{\beta(G_i)}\end{aligned}$$

and thus

$$\beta(G) = \max_{i\in[m]} \beta(G_i).$$

In words, the capacity region of a graph with complete two-way interaction among its parts is achieved by simultaneously using the optimal coding schemes for individual parts [31].

D. Lexicographic Products

We revisit the side information graph G in Fig. 4(a), which is the lexicographic product of the two graphs in Fig. 4(b) and Fig. 4(c). By Theorem 1, the capacity region of problem $G_0 \circ G_1$ is

$$\mathscr{C}(G) = \big\{(\rho_1 \mathbf{R}_1, \ldots, \rho_m \mathbf{R}_m)\colon \boldsymbol{\rho} \in \mathscr{C}(G_0), \mathbf{R}_i \in \mathscr{C}(G_1), i \in [m]\big\},$$

which implies

$$\beta(G_0 \circ G_1) = \beta(G_0)\beta(G_1). \quad (7)$$

In words, the broadcast rate is multiplicative under the lexicographic product of index coding side information graphs. We note that one direction ($\leq$) in (7) was established earlier in [29].

E. Beyond Generalized Lexicographic Products

In Section III-B, we have seen a simple application of Corollary 1. We now present a more substantial example. Consider the side information graph depicted in Fig. 7(a), which cannot be viewed as the generalized lexicographic product of smaller graphs. Let G' and G'' be the graphs depicted in Figs. 7(b) and (c), respectively. Since the graph G satisfies $V(G) = V(G') = V(G'')$ and $E(G') \subseteq E(G) \subseteq E(G'')$, its capacity region is sandwiched between the capacity regions $\mathscr{C}(G')$ and $\mathscr{C}(G'')$. Now the graphs G' and G'' are generalized lexicographic products of smaller graphs as $G' = G_0' \circ (G_1', G_2', G_3')$ and $G'' = G_0'' \circ (G_1'', G_2'', G_3'')$, where G_0', G_0'', $G_1' = G_1''$, G_2', G_2'', and $G_3' = G_3''$ are the graphs depicted in Fig. 8. Note that for each $i = 0, 1, 2, 3$, $V(G_i') = V(G_i'')$ and $E(G_i') \subseteq E(G_i'')$. Furthermore, the capacity regions of problems G_i' and G_i'' can be shown to be identical as

$$\begin{aligned}\mathscr{C}_0 &= \mathscr{C}(G_0') = \mathscr{C}(G_0'') = \\ &\qquad \{(\rho_a, \rho_b, \rho_c) : \rho_a + \rho_b \leq 1, \rho_b + \rho_c \leq 1\}, \\ \mathscr{C}_1 &= \mathscr{C}(G_1') = \mathscr{C}(G_1'') = \{R_1 : R_1 \leq 1\}, \\ \mathscr{C}_2 &= \mathscr{C}(G_2') = \mathscr{C}(G_2'') = \{(R_2, R_3) : R_2 + R_3 \leq 1\}, \\ \mathscr{C}_3 &= \mathscr{C}(G_3') = \mathscr{C}(G_3'') = \{(R_4, R_5) : R_4 + R_5 \leq 1\}.\end{aligned}$$

Hence, by Corollary 1, the capacity region $\mathscr{C}(G)$ of index coding problem G is equal to $\mathscr{C}(G') = \mathscr{C}(G'')$, which is the set of all rate tuples $(R_1, R_2, R_3, R_4, R_5)$ such that

$$\begin{aligned}R_1 &\leq \rho_a, \\ R_2 + R_3 &\leq \rho_b, \\ R_4 + R_5 &\leq \rho_c\end{aligned}$$

for some (ρ_a, ρ_b, ρ_c) such that $\rho_a + \rho_b \leq 1$ and $\rho_b + \rho_c \leq 1$. By Remark 5, this region simplifies to the set of $(R_1, \ldots, R_5)$ such that

$$\begin{aligned}R_1 + R_2 + R_3 &\leq 1, \\ R_2 + R_3 + R_4 + R_5 &\leq 1.\end{aligned}$$

Authorized licensed use limited to: Univ of Calif San Diego. Downloaded on February 15,2020 at 06:37:11 UTC from IEEE Xplore. Restrictions apply.

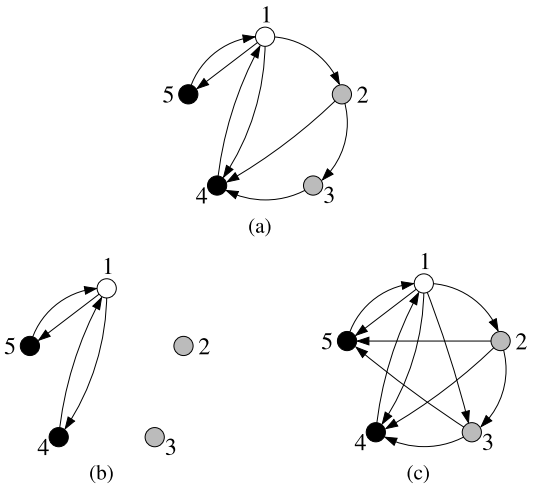

Fig. 7. (a) The 5-vertex graph G is sandwiched between (b) $G' = G'_0 \circ (G'_1, G'_2, G'_3)$ and (c) $G'' = G''_0 \circ (G''_1, G''_2, G''_3)$.

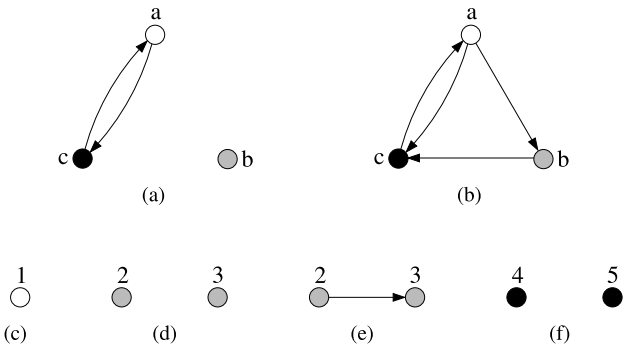

Fig. 8. (a) The 3-vertex graph G'_0, (b) the 3-vertex graph G''_0, (c) the 1-vertex graph $G'_1 = G''_1$, (d) the 2-vertex graph G'_2, (e) the 2-vertex graph G''_2, and (f) the 2-vertex graph $G'_3 = G''_3$.

IV. A Multiletter Characterization of the Index Coding Capacity Region

Alon, Hassidim, Lubetzky, Stav, and Weinstein [37] established a graph-theoretic characterization of the broadcast rate as the limit of multiletter expressions involving the chromatic number of the confusion graph [37], [38]. This characterization was later strengthened in [39] by replacing the chromatic number with the *fractional* chromatic number and also extended to the capacity region.

In this section, we use Shannon's random coding idea [40] to establish the following information theoretic multiletter characterization of the capacity region of the index coding problem.

Theorem 2. *The capacity region of the index coding problem $(i|A_i)$, $i \in [n]$, with side information graph G is the closure of*

$$\bigcup_{r=1}^{\infty} \mathscr{C}_r(G),$$

where $\mathscr{C}_r(G)$ is the set of all rate tuples $(R_1, \ldots, R_n)$ satisfying

$$R_i \le \frac{1}{r} I(U_i; V|U(A_i)), \quad i \in [n],$$

for some pmf $p(u_1)\cdots p(u_n)$ and function $f : \mathcal{U}_1 \times \cdots \times \mathcal{U}_n \to \mathcal{V}$ that maps the n-tuple $(U_1, \ldots, U_n)$ to V such that the cardinality of the auxiliary random variable V is upper bounded by 2^r.

In Theorem 2, $I(U_i; V|U(A_i))$ denotes the conditional mutual information [40], [41] between U_i and V given $U(A_i)$. Since $U_1, \ldots, U_n$ are mutually independent,

$$I(U_i; V|U(A_i)) = I(U_i; V,\ U(A_i)), \quad i \in [n].$$

In the following, we prove the theorem in two steps.

A. Proof of Achievability

We follow the standard arguments in the random coding proof of Shannon's channel coding theorem using the notion of typicality [2], [40], [41]. Here and henceforth, we define the set of ϵ-typical k-sequences $u^k = (u_1, \ldots, u_k)$ with respect to $U \sim p(u)$ for $\epsilon \in (0, 1)$ as

$$\mathcal{T}_\epsilon^{(k)}(U) = \{u^k : |\pi(u|u^k) - p(u)| \le \epsilon p(u) \text{ for all } u \in \mathcal{U}\},$$

where

$$\pi(u|u^k) = \frac{|i : u_i = u|}{k}, \quad u \in \mathcal{U},$$

is the empirical pmf of u^k. Elementary properties of the typical set and typical sequences can be found in [2], [42].

Now we prove the achievability of the rate tuples in $\mathscr{C}_r$ for each $r = 1, 2, \ldots$, based on random coding. For simplicity of presentation, we assume throughout the proof that krR_i is an integer for every $i \in [n]$.

Codebook generation. Fix a pmf $p(u_1)\cdots p(u_n)$ and a function $v = f(u_1, \ldots, u_n)$ under the prescribed cardinality constraint. For each $i \in [n]$, randomly and independently generate 2^{krR_i} sequences $u_i^k(x_i)$, $x_i \in [2^{krR_i}]$, each according to $\prod_{j=1}^k p_{U_i}(u_{ij})$. These codewords constitute the codebook, which is shared among all communicating parties.

Encoding. To communicate the message tuple $(x_1, \ldots, x_n)$, we transmit $y = v^k(u_1^k(x_1), \ldots, u_n^k(x_n)) \in [2^{kr}]$, where $v_j = f(u_{1j}(x_1), \ldots, u_{nj}(x_n))$, $j \in [k]$.

Decoding. We use *joint typicality decoding* (see, for example, [2, Sec. 3.1]). Let v^k be the received sequence and $u_J^k(x(J)) = (u_j^k(x_j), j \in J)$. Decoder $i \in [n]$ declares that $\hat{x}_i$ is sent if it is the unique message such that

$$(u_i^k(\hat{x}_i), u_{A_i}^k(x(A_i)), v^k) \in \mathcal{T}_\epsilon^{(k)}.$$

Otherwise it declares an error.

Analysis of the probability of error. By the symmetry of codebook generation, the probability of error averaged over the messages and the random codebook generation satisfies

$$\begin{aligned}\mathsf{P}(\mathcal{E}) &= \mathsf{P}\{(X_1, \ldots, X_n) \ne (\hat{X}_1, \ldots, \hat{X}_n)\} \\ &= \mathsf{P}\{\mathcal{E}|(X_1, \ldots, X_n) = (1, \ldots, 1)\}.\end{aligned}$$

Hence, we assume without loss of generality that $X_i = 1$, $i \in [n]$, is sent, and suppress the condition $\{(X_1, \ldots, X_n) = (1, \ldots, 1)\}$ in the subsequent probability expressions for

Authorized licensed use limited to: Univ of Calif San Diego. Downloaded on February 15,2020 at 06:37:11 UTC from IEEE Xplore. Restrictions apply.

brevity. Let $\mathsf{P}(\mathcal{E}_i)$ be the probability of error for decoder i. Then, by the union of events bound

$$\mathsf{P}(\mathcal{E}) \leq \sum_{i\in[n]} \mathsf{P}(\mathcal{E}_i). \tag{8}$$

Note that decoder i makes an error iff one or more of the following events occur:

$$\mathcal{E}_{i1} = \{(U_i^k(1), U_{A_i}^k((1,\ldots,1)), V^k) \notin \mathcal{T}_\epsilon^{(k)}\},$$
$$\mathcal{E}_{i2} = \{(U_i^k(x_i), U_{A_i}^k((1,\ldots,1)), V^k) \in \mathcal{T}_\epsilon^{(k)} \text{ for some } x_i \neq 1\}.$$

Thus, by the union of events bound, the probability of error for decoder i is upper bounded as

$$\mathsf{P}(\mathcal{E}_i) \leq \mathsf{P}(\mathcal{E}_{i1}) + \mathsf{P}(\mathcal{E}_{i2}).$$

By the law of large numbers, $\mathsf{P}(\mathcal{E}_{i1})$ tends to zero as $k \to \infty$. If $x_i \neq 1$, $U_i^k(x_i)$ is independent of V^k and $U^k(A_i)$. Hence, by the packing lemma [2, Lemma 3.1], $\mathsf{P}(\mathcal{E}_{i2})$ tends to zero as $k \to \infty$ if

$$rR_i < I(U_i; V, U(A_i)) - \delta(\epsilon) = I(U_i; V|U(A_i)) - \delta(\epsilon), \tag{9}$$

where $\delta(\epsilon)$ tends to zero as $\epsilon \to 0$ and the last identity follows since U_i and $U(A_i)$ are independent. Therefore, by (8), if the specified rate constraints in (9) are satisfied simultaneously for all messages, the probability of error $\mathsf{P}(\mathcal{E})$ averaged over messages and codebooks tends to zero as $k \to \infty$, and there must exist a sequence of $(\lceil krR_1\rceil, \ldots, \lceil krR_n\rceil, kr)$ index codes such that the probability of error averaged over the messages tends to zero as $k \to \infty$. Letting $\epsilon \to 0$ shows that any rate tuple $(R_1, \ldots, R_n) \in \mathscr{C}_r$ is achievable with vanishing probability of error. By Remark 1, this error probability can be made to be exactly zero without sacrificing the rates and thus $\mathscr{C}_r$ is contained in the capacity region. This completes the proof of achievability.

B. Proof of the Converse

We show that any achievable rate tuple $(R_1, \ldots, R_n)$ lies in some $\mathscr{C}_r$. First note that for any $(t_1, \ldots, t_n, r)$ index code,

$$H(X_i|Y, X(A_i)) = 0, \quad i \in [n].$$

Hence,

$$rR_i \leq t_i = H(X_i) = I(X_i; Y|X(A_i)), \quad i \in [n].$$

By identifying $U_i = X_i$, $i \in [n]$, and $V = Y$, the cardinalities of which are all upper bounded by 2^r, we can conclude that

$$R_i \leq \frac{1}{r} I(U_i; V|U(A_i)), \quad i \in [n],$$

for some $p(u_1)\cdots p(u_n)$ and $v = f(u_1, \ldots, u_n)$ such that the cardinalities are bounded by 2^r. This completes the proof of the converse.

V. Proof of Theorem 1

In this section, we use the information theoretic characterization of index coding capacity region in Theorem 2 to prove the main result of the paper.

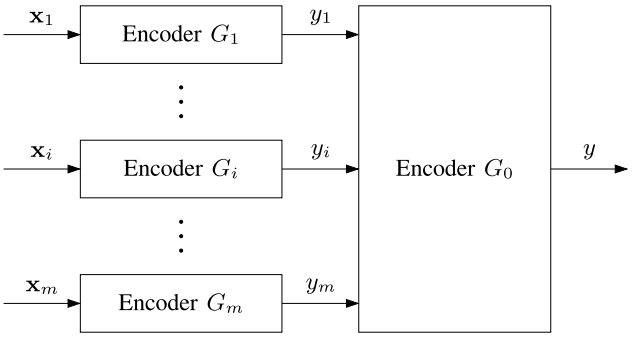

Fig. 9. Construction of an index code for index coding problem $G_0 \circ (G_1, \ldots, G_m)$ by concatenating the index codes for problems $G_1, \ldots, G_m$ as the inner codes and the index code for problem G_0 as the outer code. The message tuple $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_m)$ is encoded by index codes for $G_1, \ldots, G_m$ part by part. The outputs $y_1, \ldots, y_m$ are then encoded by the index code for G_0.

A. Proof of Achievability

The proof of achievability extends the arguments in [29] and uses the simple construction of an index code for $G = G_0 \circ (G_1, \ldots, G_m)$ from index codes for subproblems as illustrated in Fig. 9. To be more precise, consider any rate tuple $(\rho_1\mathbf{R}_1, \ldots, \rho_m\mathbf{R}_m)$, where $\mathbf{R}_i \in \mathscr{C}_i$, $i \in [m]$, and $(\rho_1, \ldots, \rho_m) \in \mathscr{C}_0$. Let $\epsilon > 0$. Then, by the definition of the capacity region, there exists a $(\lceil(\rho_1-\epsilon)r\rceil, \ldots, \lceil(\rho_m-\epsilon)r\rceil, r)$ index code for problem G_0 for r sufficiently large. Also for each $i \in [m]$, there exists a $(\lceil(\mathbf{R}_i - \epsilon\mathbf{1})r_i\rceil, r_i) = (\mathbf{t}_i, r_i)$ index code for problem G_i for r_i sufficiently large. Let $r_i = \lceil(\rho_i - \epsilon)r\rceil$, $i \in [m]$. Then, by concatenating the $(\mathbf{t}_i, r_i)$ index codes, $i \in [m]$, with $(r_1, \ldots, r_m, r)$ index code as shown in Fig. 9, we can construct a code for problem G. The rate of message i of this code is

$$\begin{aligned}\frac{\mathbf{t}_i}{r} &= \frac{r_i}{r}\frac{\mathbf{t}_i}{r_i} \\ &= \frac{\lceil(\rho_i - \epsilon)r\rceil}{r}\frac{\lceil(\mathbf{R}_i - \epsilon\mathbf{1})r_i\rceil}{r_i} \\ &\geq (\rho_i - \epsilon)(\mathbf{R}_i - \epsilon\mathbf{1}), \quad i \in [m].\end{aligned}$$

Letting $\epsilon \to 0$ completes the proof.

B. Proof of the Converse

Our proof is inspired by the proof for the one-way interaction in [30], but significantly extends the arguments therein. Let $A_j \subseteq V(G)$ denote the side information set of receiver $j \in V(G)$ for the index coding problem G, and let $A_i' \subseteq [m]$ denote the side information set of receiver $i \in [m]$ for the index coding problem G_0. In this notation, if $j \in V(G_i)$ for some $i \in [m]$, then by the definition of generalized lexicographic product, the side information set of receiver j can be decomposed as

$$A_j = \big(A_j \cap V(G_i)\big) \cup \Big(\bigcup_{l\in A_i'} V(G_l)\Big), \tag{10}$$

where the first term denotes the side information from within the subproblem G_i and the second term denotes the side information from other subproblems. As in Fig. 9, we write $\mathbf{x}_i$

Authorized licensed use limited to: Univ of Calif San Diego. Downloaded on February 15,2020 at 06:37:11 UTC from IEEE Xplore. Restrictions apply.

for $(x_j : j \in V(G_i))$ and $\mathbf{x}$ for $(\mathbf{x}_1, \ldots, \mathbf{x}_m)$. We also write $\mathbf{x}(A_i')$ for $(x_j : j \in \cup_{l \in A_i'} V(G_l))$.

To prove the converse ($\mathscr{C} \subseteq \mathscr{C}_0 \circ (\mathscr{C}_1, \ldots, \mathscr{C}_m)$), for any $(\mathbf{t}_1, \ldots, \mathbf{t}_m, r)$ index code for $G = G_0 \circ (G_1, \ldots, G_m)$, we argue that the corresponding rate tuple can be factored as

$$\frac{\mathbf{t}_i}{r} = \frac{s_i}{r}\frac{\mathbf{t}_i}{s_i}, \quad i \in [m],$$

for some $(s_1, \ldots, s_m)$, so that

$$\left(\frac{s_1}{r}, \ldots, \frac{s_m}{r}\right) \in \mathscr{C}_0, \tag{11a}$$

and for any $\epsilon > 0$,

$$\frac{(1-\epsilon)\mathbf{t}_i}{s_i} \in \mathscr{C}_i, \quad i \in [m]. \tag{11b}$$

Consequently,

$$(1-\epsilon)\left(\frac{\mathbf{t}_1}{r}, \ldots, \frac{\mathbf{t}_m}{r}\right) \in \mathscr{C}_0 \circ (\mathscr{C}_1, \ldots, \mathscr{C}_m).$$

Since $\epsilon > 0$ is arbitrary, this would establish the desired proof of the converse.

We now verify (11) for an appropriate $(s_1, \ldots, s_m)$. Let $Y = \phi(\mathbf{X}_1, \ldots, \mathbf{X}_m) \in \{0,1\}^r$ be the encoder output of the given index code for independent and uniformly distributed messages, which induces the joint distribution of the form

$$p(\mathbf{x}_1, \ldots, \mathbf{x}_m, y) = p(\mathbf{x}_1) \cdots p(\mathbf{x}_m) p(y|\mathbf{x}_1, \ldots, \mathbf{x}_m) \tag{12}$$

such that Y is a function of $(\mathbf{X}_1, \ldots, \mathbf{X}_m)$ and X_j is a function of $(Y, X(A_j))$ for every $j \in V(G)$, namely, $X_j = \psi_j(Y, X(A_j))$, $j \in V(G)$. Now let

$$s_i = I(\mathbf{X}_i; Y|\mathbf{X}(A_i')), \quad i \in [m],$$

where the mutual information is evaluated under the joint distribution in (12). Then, by Theorem 2 (with $U_i = \mathbf{X}_i$ and $V = Y$), we have (11a). For (11b), we first state two lemmas, the proofs of which are presented in Appendices A and B.

Lemma 1. *For any n-message index coding problem $(i|A_i)$, $i \in [n]$, with side information graph G, let*

$$(\phi(x^n), \psi_1(y, x(A_1)), \ldots, \psi_n(y, x(A_n)))$$

be the encoder and decoders of a $(t_1, \ldots, t_n, r)$ index code under a relaxed *decoding condition*

$$\psi_i(\phi(x^n), x(A_i)) = x_i, \quad i \in J,$$

for some subset $J \subseteq [n]$ of the messages. Then,

$$\frac{(t_i : i \in J)}{r} \in \mathscr{C}(G|_J).$$

Lemma 2. *Let $\epsilon > 0$ and $s_i = I(\mathbf{X}_i; Y|\mathbf{X}(A_i'))$. Then there exist mappings*

$$\phi_i'(\mathbf{x}^k) \in \{0,1\}^{ks_i/(1-\epsilon)}, \quad i \in [m], \tag{13a}$$

and

$$\psi_j'(w_i, x^k(A_j)) \in \{0,1\}^{kt_j}, \quad j \in V(G_i), \tag{13b}$$

TABLE I

THE NUMBERS OF n-MESSAGE INDEX CODING PROBLEMS WHOSE CAPACITY REGIONS CAN BE CHARACTERIZED BY THE DIVIDE-AND-CONQUER APPROACH BASED ON THEOREM 1 AND COROLLARY 1

Number of messages	N	N_{GLP}	N_{Sand}	$\frac{N_{\mathrm{GLP}}+N_{\mathrm{Sand}}}{N}$
2	3	3	0	100%
3	16	11	3	87.5%
4	218	110	70	82.6%
5	9,608	2,511	4,054	68.3%
6	1,540,944	161,989	607,161	49.9%

such that

$$\psi_j'(\phi_i'(\mathbf{x}^k), x^k(A_j)) = x_j^k, \quad i \in [m],\ j \in V(G_i), \tag{13c}$$

for k sufficiently large.

Now we are ready to verify (11b). We first apply Lemma 2 for each $i \in [m]$. The mappings $\phi_i'(\mathbf{x}^k)$ and $\psi_j'(w_i, x^k(A_j))$, $j \in V(G_i)$, form a $(k\mathbf{t}_1, \ldots, k\mathbf{t}_m, ks_i/(1-\epsilon))$ index code for G under the relaxed decoding condition that only $\mathbf{x}_i^k = (x_j^k : j \in V(G_i))$ is required to be recovered correctly. Hence, by Lemma 1, we can conclude that (11b) holds. This completes the proof of Theorem 1.

VI. CONCLUDING REMARKS

The generalized lexicographic product structure investigated in this paper provides a natural method of building a larger index coding problem from smaller problems so that the capacity region of the larger problem can be expressed in the same generalized lexicographic product structure from the subproblem capacity regions. This leads to a divide-and-conquer approach to computing the index coding capacity region, either through direct decomposition (Theorem 1) or by sandwiching between two generalized lexicographic products (Corollary 1).

Since the capacity region of a general n-message index coding problem is known for $n \leq 5$, we can test this divide-and-conquer approach for all problems with six or less messages. Table I lists the number of all nonisomorphic n-message index coding problems N, along with the number of problems that are generalized lexicographic products of smaller graphs (N_{GLP}), the number of problems that are sandwiched between two generalized lexicographic products of the same capacity region (N_{Sand}), and the percentage of the problems whose capacity regions can be characterized by this *divide-and-conquer* approach. This simple approach solves about one half of the 6-message problems without explicitly computing any inner and outer bounds on the capacity region.

Identifying the generalized lexicographic product structure in a general side information graph is a computationally challenging problem (see Remarks 3 and 7). We offer the following algorithmic questions that would shed some light on the current line of investigation:

- Given a graph G, can we efficiently determine whether G is a generalized lexicographic product of smaller

Authorized licensed use limited to: Univ of Calif San Diego. Downloaded on February 15,2020 at 06:37:11 UTC from IEEE Xplore. Restrictions apply.

graphs? Although only a very small number of graphs are generalized lexicographic products, the capacity regions of many other graphs can be tightly sandwiched by the capacity regions of these graphs.

- Can we efficiently transform a graph G into a generalized lexicographic product by adding or removing a few edges? A recursive application of this procedure can yield a general outer or inner bound on the capacity region.

Appendix A
Proof of Lemma 1

We construct an index code for problem $G|_J$ by setting $x_i = 0$, $i \notin J$, in $\phi(x^n)$ and ψ_i, $i \in J$. For every $x^n \in \Pi_{i=1}^n \{0,1\}^{t_i}$, define $\tilde{x}^n = \tilde{x}^n(x^n)$ by

$$\tilde{x}_i = \begin{cases} x_i, & i \in J, \\ 0, & i \notin J, \end{cases}$$

represented in the same t_i bits. Note that the side information set of receiver $i \in J$ for problem $G|_J$ is $A_i \cap J$. Let

$$\phi'(x(J)) = \phi(\tilde{x}^n) \in \{0,1\}^r,$$

and

$$\psi_i'(y, x(A_i \cap J)) = \psi_i(y, \tilde{x}(A_i)).$$

Then, by the given decoding condition, for all $i \in J$ we have

$$\psi_i'(\phi'(x(J)), x(A_i \cap J)) = \psi_i(\phi(\tilde{x}^n), \tilde{x}(A_i)) = \tilde{x}_i = x_i.$$

Hence, the mappings $\phi'(x(J))$ and $\psi_i'(y, x(A_i \cap J))$, $i \in J$, form a valid index code for the problem $G|_J$. This completes the proof of the lemma.

Appendix B
Proof of Lemma 2

At a high level, the proof is based on random coding for rate–distortion theory [43] and joint typicality encoding [2, Sec. 3.6] over k copies of $(\mathbf{X}_1, \ldots, \mathbf{X}_m, Y)$. For each $i \in [m]$, consider the joint distribution $p(\mathbf{x}_i, \mathbf{x}(A_i'), y)$ from (12) and fix the conditional distribution $p(y|\mathbf{x}(A_i'))$. For each $\mathbf{x}^k(A_i')$, generate kr-bit sequences $y_i^k(w_i|\mathbf{x}^k(A_i'))$, $w_i \in [2^{ks_i/(1-\epsilon)}]$, each i.i.d. according to $p(y|\mathbf{x}(A_i'))$. Then by the covering lemma [2, Lemma 3.3], with high probability there exists at least one w_i such that

$$(\mathbf{x}_i^k, y_i^k(w_i|\mathbf{x}^k(A_i')), \mathbf{x}^k(A_i')) \in \mathcal{T}_\epsilon^{(k)}(\mathbf{X}_i, Y, \mathbf{X}(A_i')), \quad (14)$$

provided that k is sufficiently large and

$$s_i/(1-\epsilon) > I(\mathbf{X}_i; Y|\mathbf{X}(A_i')).$$

If there is such a w_i (if there is more than one, choose one arbitrarily), then we set

$$\phi_i'(\mathbf{x}^k) = w_i.$$

Note by (14) that the chosen w_i is a function of $\mathbf{x}_i^k$ and $\mathbf{x}^k(A_i')$ (and thus of $\mathbf{x}^k$). If there is no such index, set $\phi_i'(\mathbf{x}^k) = 1$.

We now define ψ_j' for each $j \in V(G_i)$. Let

$$\psi_j'(w_i, x^k(A_j)) = \psi_j(y_i^k(w_i|\mathbf{x}^k(A_i')), x^k(A_j)),$$

where ψ_j is the decoding function of the given index code for problem G that is employed k times. Suppose that the joint typicality in (14) holds among $\mathbf{x}_i^k$, $y_i^k(w_i|\mathbf{x}^k(A_i'))$, and $\mathbf{x}^k(A_i')$. Then by the properties of joint typicality [2, Section 2.5], any functional relationship for them should hold, namely

$$x_j^k = \psi_j(y_i^k, x^k(A_j)) = \psi_j'(\phi_i'(\mathbf{x}^k), x^k(A_j)), \quad j \in V(G_i).$$

Therefore, as long as w_i satisfying (14) is found, which happens with high probability, the mappings ϕ_i' and ψ_j' defined above satisfy the desired properties in (13) with high probability. Finally, by Remark 1, we can come up with mappings for which these properties hold for every sequence with a negligible decrease in the rates. This completes the proof of the lemma.

Acknowledgments

The authors would like to thank the Associate Editor and anonymous reviewers for their constructive comments, which improved the readability of the article significantly. They would also like to thank E. Grigorescu and M. Zhu for pointing out an error in an earlier proof of Theorem 1 based on the clique number of confusion graphs, and acknowledge P. Sadeghi and J. Verstraete for helpful discussions.

References

[1] F. M. J. Willems, "The maximal-error and average-error capacity region of the broadcast channel are identical: A direct proof," *Probl. Control Inf. Theory*, vol. 19, no. 4, pp. 339–347, 1990.
[2] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge, U.K.: Cambridge Univ. Press, 2011.
[3] T. Chan and A. Grant, "On capacity regions of non-multicast networks," in *Proc. IEEE Int. Symp. Inf. Theory*, Austin, TX, USA, Jun. 2010, pp. 2378–2382.
[4] M. Langberg and M. Effros, "Network coding: Is zero error always possible?" in *Proc. 49th Ann. Allerton Conf. Commun. Control Comput.*, Monticello, IL, USA, 2011, pp. 1478–1485.
[5] M. Celebiler and G. Stette, "On increasing the down-link capacity of a regenerative satellite repeater in point-to-point communications," *Proc. IEEE*, vol. 66, no. 1, pp. 98–100, Jan. 1978.
[6] F. M. J. Willems, J. K. Wolf, and A. D. Wyner, "Communicating via a processing broadcast satellite," in *Proc. IEEE/CAM Inf. Theory Workshop*, Cornell, NY, USA, Jun. 1989, p. 3_1.
[7] A. D. Wyner, J. K. Wolf, and F. M. J. Willems, "Communicating via a processing broadcast satellite," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1243–1249, Jun. 2002.
[8] R. W. Yeung, "Multilevel diversity coding with distortion," *IEEE Trans. Inf. Theory*, vol. 41, no. 2, pp. 412–422, Mar. 1995.
[9] Y. Birk and T. Kol, "Informed-source coding-on-demand (ISCOD) over broadcast channels," in *Proc. 17th Annu. IEEE Int. Conf. Comput. Commun. (INFOCOM)*, San Francisco, CA, USA, Mar./Apr. 1998, pp. 1257–1264.
[10] Y. Birk and T. Kol, "Coding on demand by an informed source (ISCOD) for efficient broadcast of different supplemental data to caching clients," *IEEE Trans. Inf. Theory*, vol. 52, no. 6, pp. 2825–2830, Jun. 2006.
[11] S. Riis, "Information flows, graphs and their guessing numbers," *Electr. J. Combin.*, vol. 14, no. 1, pp. 1–17, 2007.
[12] S. El Rouayheb, A. Sprintson, and C. Georghiades, "On the relation between the index coding and the network coding problems," in *Proc. IEEE Int. Symp. Inf. Theory*, Toronto, ON, Canada, Jul. 2008, pp. 1823–1827.
[13] M. Effros, S. El Rouayheb, and M. Langberg, "An equivalence between network coding and index coding," *IEEE Trans. Inf. Theory*, vol. 61, no. 5, pp. 2478–2487, May 2015.
[14] A. Mazumdar, "On a duality between recoverable distributed storage and index coding," in *Proc. IEEE Int. Symp. Inf. Theory*, Honolulu, HI, USA, Jun./Jul. 2014, pp. 1977–1981.

Authorized licensed use limited to: Univ of Calif San Diego. Downloaded on February 15,2020 at 06:37:11 UTC from IEEE Xplore. Restrictions apply.

[15] K. Shanmugam and A. G. Dimakis, "Bounding multiple unicasts through index coding and locally repairable codes," in *Proc. IEEE Int. Symp. Inf. Theory*, Honolulu, HI, USA, Jun./Jul. 2014, pp. 296–300.
[16] F. Arbabjolfaei and Y.-H. Kim, "Three stories on a two-sided coin: Index coding, locally recoverable distributed storage, and guessing games on graphs," in *Proc. 53rd Annu. Allerton Conf. Commun. Control Comput.*, Monticello, IL, USA, Sep./Oct. 2015, pp. 843–850.
[17] X. Yi, H. Sun, S. A. Jafar, and D. Gesbert, "TDMA is optimal for all-unicast DoF region of TIM if and only if topology is chordal bipartite," *IEEE Trans. Inf. Theory*, vol. 64, no. 3, pp. 2065–2076, Mar. 2018.
[18] K. Shanmugam, M. Asteris, and A. G. Dimakis, "On approximating the sum-rate for multiple-unicasts," in *Proc. IEEE Int. Symp. Inf. Theory*, Hong Kong, Jun. 2015, pp. 381–385.
[19] M. J. Neely, A. S. Tehrani, and Z. Zhang, "Dynamic index coding for wireless broadcast networks," in *Proc. 31st Annu. IEEE Int. Conf. Comput. Commun. (INFOCOM)*, Orlando, FL, USA, Mar. 2012, pp. 316–324.
[20] S. A. Jafar, "Topological interference management through index coding," *IEEE Trans. Inf. Theory*, vol. 60, no. 1, pp. 529–568, Jan. 2014.
[21] M. A. Maddah-Ali and U. Niesen, "Fundamental limits of caching," *IEEE Trans. Inf. Theory*, vol. 60, no. 5, pp. 2856–2867, May 2014.
[22] M. Ji, G. Caire, and A. F. Molisch, "Fundamental limits of caching in wireless D2D networks," *IEEE Trans. Inf. Theory*, vol. 62, no. 2, pp. 849–869, Feb. 2016.
[23] S. Y. El Rouayheb, "Network and index coding with application to robust and secure communications," Ph.D. dissertation, Dept. Elect. Comput. Eng., Texas A&M Univ., College Station, TX, USA, 2009.
[24] A. Blasiak, "A graph-theoretic approach to network coding," Ph.D. dissertation, Dept. Comput. Sci., Cornell Univ., Ithaca, NY, USA, 2013.
[25] F. Arbabjolfaei, "Index coding: Fundamental limits, coding schemes, and structural properties," Ph.D. dissertation, Dept. Elect. Comput. Eng., Univ. California, San Diego, La Jolla, CA, USA, 2017.
[26] E. Byrne and M. Calderini, "Index coding, network coding and broadcast with side-information," in *Network Coding and Subspace Designs*, M. Greferath, M. O. Pavčević, N. Silberstein, and M. Á. Vázquez-Castro, Eds. Cham, Switzerland: Springer, 2018, pp. 247–293.
[27] F. Arbabjolfaei and Y.-H. Kim, "Fundamentals of index coding," *Found. Trends Commun. Inf. Theory*, vol. 14, nos. 3–4, pp. 164–344, 2018.
[28] F. Arbabjolfaei, B. Bandemer, Y.-H. Kim, E. Şaşoğlu, and L. Wang, "On the capacity region for index coding," in *Proc. IEEE Int. Symp. Inf. Theory*, Istanbul, Turkey, Jul. 2013, pp. 962–966.
[29] A. Blasiak, R. Kleinberg, and E. Lubetzky, "Lexicographic products and the power of non-linear network coding," in *Proc. 52nd Annu. IEEE Symp. Found. Comput. Sci.*, Palm Springs, CA, USA, Oct. 2011, pp. 609–618.
[30] M. Tahmasbi, A. Shahrasbi, and A. Gohari, "Critical graphs in index coding," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 2, pp. 225–235, Feb. 2015.
[31] F. Arbabjolfaei and Y.-H. Kim, "Structural properties of index coding capacity using fractional graph theory," in *Proc. IEEE Int. Symp. Inf. Theory*, Hong Kong, Jun. 2015, pp. 1034–1038.
[32] A. J. Schwenk, "Computing the characteristic polynomial of a graph," in *Graphs and Combinatorics* (Lecture Notes in Mathematics), vol. 406. Berlin, Germany: Springer, 1974.
[33] C. Godsil and B. McKay, "A new graph product and its spectrum," *Bull. Austral. Math. Soc.*, vol. 18, no. 1, pp. 21–28, Feb. 1978.
[34] R. Hammack, W. Imrich, and S. Klavzar, *Handbook of Product Graphs*, 2nd ed. Boca Raton, FL, USA: CRC Press, 2011.
[35] E. R. Scheinerman and D. H. Ullman, *Fractional Graph Theory: A Rational Approach to the Theory of Graphs*. New York, NY, USA: Dover, 2011.
[36] A. Bachem and W. Kern, *Linear Programming Duality: An Introduction to Oriented Matroids*. Berlin, Germany: Springer, 1992.
[37] N. Alon, A. Hassidim, E. Lubetzky, U. Stav, and A. Weinstein, "Broadcasting with side information," in *Proc. 49th Ann. IEEE Symp. Found. Comput. Sci.*, Philadelphia, PA, USA, Oct. 2008, pp. 823–832.
[38] M. Gadouleau and S. Riis, "Graph-theoretical constructions for graph entropy and network coding based communications," *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6703–6717, Oct. 2011.
[39] F. Arbabjolfaei and Y.-H. Kim, "Approximate capacity of index coding for some classes of graphs," in *Proc. IEEE Int. Symp. Inf. Theory*, Barcelona, Spain, Jul. 2016, pp. 2154–2158.
[40] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 3, pp. 379–423, Jul./Oct. 1948.
[41] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. New York, NY, USA: Wiley, 2006.
[42] A. Orlitsky and J. R. Roche, "Coding for computing," *IEEE Trans. Inf. Theory*, vol. 47, no. 3, pp. 903–917, Mar. 2001.
[43] C. E. Shannon, "Coding theorems for a discrete source with a fidelity criterion," *Proc. IRE*, vol. 7, no. 1, pp. 142–163, Mar. 1959.

Fatemeh Arbabjolfaei received the B.S. and M.S. degrees (Hons.) in electrical engineering from the Isfahan University of Technology, Iran, in 2007 and 2010, respectively, and the Ph.D. degree in electrical engineering from the University of California San Diego in 2017. She performed her postdoctoral research at the University of Southern California. She is currently a Lecturer with the University of Michigan, Ann Arbor. She coauthored the monograph *Fundamentals of Index Coding* (Now Publishers, 2019). Her research interests are in information theory, digital signal processing, and theoretical data science.

Young-Han Kim (Fellow, IEEE) received the B.S. degree (Hons.) in electrical engineering from Seoul National University, South Korea, in 1996, and the M.S. degrees in electrical engineering and in statistics and the Ph.D. degree in electrical engineering from Stanford University in 2001, 2006, and 2006, respectively. In 2006, he joined the University of California San Diego, where he is currently a Professor with the Department of Electrical and Computer Engineering. He coauthored the book *Network Information Theory* (Cambridge University Press, 2011) and the monograph *Fundamentals of Index Coding* (Now Publishers, 2019). His research interests are in information theory, communication engineering, and data science. He was a recipient of the 2008 NSF Faculty Early Career Development (CAREER) Award, the 2009 U.S.-Israel Binational Science Foundation Bergmann Memorial Award, the 2012 IEEE Information Theory Paper Award, and the 2015 IEEE Information Theory Society James L. Massey Research and Teaching Award for Young Scholars. He served as an Associate Editor for the IEEE TRANSACTIONS ON INFORMATION THEORY and a Distinguished Lecturer for the IEEE Information Theory Society.

Authorized licensed use limited to: Univ of Calif San Diego. Downloaded on February 15,2020 at 06:37:11 UTC from IEEE Xplore. Restrictions apply.