Critical Graphs in Index Coding

Mehrdad Tahmasbi, Amirbehshad Shahrasbi, and Amin Gohari

***Abstract*—In this paper, we define critical graphs as minimal graphs that support a given set of rates for the index coding problem and study them for both the one-shot and asymptotic setups. For the case of equal rates, we find the critical graph with minimum number of edges for both one-shot and asymptotic cases. For the general case of possibly distinct rates, we show that for one-shot and asymptotic linear index coding, as well as asymptotic nonlinear index coding, each critical graph is a union of disjoint strongly connected subgraphs. On the other hand, we identify a non-USCS critical graph for a one-shot nonlinear index coding problem. Next, we identify a few graph structures that are critical. In addition, we show that the capacity region of the index coding is additive for union of disjoint graphs.**

***Index Terms*—Index coding, critical graphs.**

I. Introduction

INTRODUCED by Birk and Kol in [2], index coding is the problem of transmitting a set of messages to a number of receivers via public communication. Each receiver may also have some side information consisting of messages desired by some of the other receivers. In the most general form of the problem, each message can be desired by more than one destination. However the special case of each message being desired by exactly one receiver admits a graph theoretic representation in terms of directed graphs and thus has received particular attention. More specifically, if there are m receivers, we can construct a graph with m vertices. We draw a directed edge from vertex i to vertex j if and only if receiver i knows the desired message by receiver j. In this paper we work with this graph model for the index coding problem.

It is common to study the index coding problem in terms of an achievable rate region based on the size of the m messages to be decoded by the m receivers (see Section II for a formal definition). Here the rate of a receiver refers to the normalized amount of information transmitted to it. The set of all achievable rates, i.e., the capacity region, for index coding problem remains an open problem. Nonetheless, there has been some progress on this problem (e.g., see [3]–[10]). A difference between the performance of linear and non-linear codes is characterized in [11].

Manuscript received April 17, 2014; revised September 18, 2014 and November 9, 2014; accepted November 14, 2014. Date of publication December 18, 2014; date of current version March 9, 2015. This paper was presented in part at the IEEE Symposium on Information Theory, Honolulu, HI, USA, June 29–July 4, 2014. The work of A. Gohari was supported by the Institute for Research in Fundamental Sciences (IPM) under Grant 92050116.

The authors are with the Department of Electrical Engineering, Sharif University of Technology, Tehran 11365-11155, Iran (e-mail: tahmasebi_mehrdad@ee.sharif.edu; shahrasbi@ee.sharif.edu; aminzadeh@sharif.edu).

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

Digital Object Identifier 10.1109/JSAC.2014.2384294

Connections With Network Coding and Wireless Communication: The index coding problem has significant connections with network coding and wireless communications. It is shown in [9] that for both linear and non-linear case, for any instance of networking coding problem, there exists an instance of index coding problem with the same capacity region (see also [8] for another connection between network coding and index coding problem).

In [15], Jafar studies the topological interference management problem for wireless networks. More specifically, given an interference pattern in a wireless system, he constructs an imaginary index coding problem where interference among users is illustrated through a side information graph. He then shows that the degrees of freedom region of a wireless interference network is related to the corresponding index coding capacity region. For example, in [15] it is proved that the set of degrees of freedom which are available through linear schemes in the topological interference management problem is equal to the linear capacity region of an equivalent index coding problem. Moreover, the non-linear degree of freedom region of the interference management problem is related to the non-linear capacity region of the problem. Our results on index coding then imply that in certain wireless networks, adding new interference from a set of transmitters to a set of receivers does not affect the capacity region in the high SNR regime.

Our Contributions: Given a fixed set of rates, let $\mathcal{G}$ denote the set of all graphs that support the rates. We are interested in minimal members of $\mathcal{G}$ (with respect to containment of the edge set). More specifically, a graph is said to be *critical* (or *edge critical*) if (1) it belongs to $\mathcal{G}$ and (2) deletion of any edge from the graph makes it to fall outside $\mathcal{G}$. It is useful to study critical graphs since it identifies the minimum-cost architectures of the networks supporting a given set of rates. Furthermore characterizing critical graphs is equivalent with solving the index coding problem itself. When we vary the side information graph, critical graphs will be the extreme cases where coding strategies need to change to adapt to the structure of the graph.

To the best of our knowledge, critical graphs for index coding have not been studied before. We present several results in this paper regarding critical graphs. When the rates are all equal, we identify the critical graph with minimum number of edges (Theorem 1). Next we study the general case of arbitrary rates (Theorem 2; here we basically prove that a simple time division strategy is optimal). We use this result to show that critical graphs for one-shot and asymptotic linear index coding as well as those of non-linear asymptotic index coding are structured, by proving that they have to be a union of disjoint strongly connected subgraphs (USCS) (Theorem 3). On the other hand, for non-linear one-shot index coding, we construct

CaseThm	1	2 (a)	2 (b, c)	3 (a)	3 (b)	4	5	5 (sym)	6
1-Shot Non-Lin.	✓	✓	✗	✗	✓	✓	✓	✗	✓
1-Shot Lin.	✓	✓	✗	✓	✗	✓	✓	✓	✓
Asmp. Non-Lin.	✓	✓	✓	✓	✗	✓	✓	✓	✓
Asmp. Lin.	✓	✓	✓	✓	✗	✓	✓	✓	✓

Fig. 1. This table presents the coverage of different index coding scenarios in our main theorems.

a counterexample by finding a critical graph that is not USCS. Next Theorems 5 and 6 find some classes of critical graphs; these were identified after we computed a comprehensive list of symmetric critical graphs for graphs with at most five nodes.

A potential application of critical graphs is in the study of wireless broadcast networks. For example, in [16] side information of nodes in a broadcast wireless network has been employed to make the communication more efficient. In such schemes, study of critical graphs can be helpful as it identifies the side information that cannot make the communication more efficient. For instance, as our results show, those side information whose corresponding edge in the side information graph do not lie on any cycle, will not improve the efficiency of communication. Hence, these side information can be eliminated. Accordingly, the total storage resources of wireless nodes can be decreased using our results.

Lastly, we have a novel result (Theorem 2) for characterizing the index coding capacity of certain structured graphs in all scenarios except the non-linear one-shot case.

This paper is organized as follows: in Section II, basic notation and definitions are provided. The results are given in Section III, with proofs coming in the following section. Fig. 1 shows the coverage of various index coding setups in our main theorems.

II. Definitions and Preliminaries

A (unicast) index coding problem comprises of m nodes, $\{1,\cdots,m\}$, and a set of m message $\{W_1,\cdots,W_m\}$ where node i needs to decode the message W_i, $i=1,\cdots,m$. The side information of node i is assumed to be a subset of $\{W_1,\cdots,W_{i-1},W_{i+1},\cdots,W_m\}$. We can illustrate the side information of nodes by a directed graph $\mathsf{G}=(\mathcal{V},\mathcal{E})$, where $\mathcal{V}=\{1,\cdots,m\}$ and node i has an edge to node j (that is, $(i,j)\in\mathcal{E}$) if node i knows W_j. For simplicity in the rest of this paper, we use graph as a shorthand for directed graphs. Undirected graphs are referred to as "bidirectional graphs."

Definition 1: A code for an index coding problem (or an index code) consists of

1) m alphabet sets $\mathcal{W}_i$, $i=1,2,\cdots,m$ where the message intended by the i-th party, W_i, belongs to $\mathcal{W}_i$;
2) An encoding function f from $\mathcal{W}_1\times\cdots\times\mathcal{W}_m$ to $\{1,2,\cdots,N\}$ that compresses the messages $(W_1,\cdots,W_m)$ into a symbol in $\{1,2,\cdots,N\}$. $f(W_1,\cdots,W_m)$ is called the public message since it will be made available to all the nodes;
3) A set of m decoding functions at the nodes from $\{1,2,\cdots,N\}\times\prod_{(i,j)\in\mathcal{E}}\mathcal{W}_j$ to $\mathcal{W}_i$ for $i=1,2,\cdots,m$. Every node should be able to decode its message using the public message and its side information.

The rate vector associated with the code is a vector $(r_1,\cdots,r_m)$ where $r_i=\log(|\mathcal{W}_i|)/\log(N)$. We will use $\bar{\mathbf{r}}$ to indicate the rate vector $(r_1,\cdots,r_m)$.

Probability of error associated to the code is the probability that node i fails to correctly decode W_i for some $i=1,2,\cdots,m$, where rvs W_i are assumed to be uniform on their alphabet set and mutually independent of each other.

Definition 2 (Linear Codes): A linear code for an index coding problem with finite field $\mathbb{F}$ consists of

1) m positive integers $l_1,\cdots,l_m$ indicating that $W_i\in\mathbb{F}^{l_i}$ is a sequence of length l_i of symbols in $\mathbb{F}$. In other words, the alphabet set for the rv W_i is $\mathcal{W}_i=\mathbb{F}^{l_i}$;
2) A linear map f from $\mathcal{W}_1\times\cdots\times\mathcal{W}_m$ to $\mathbb{F}^n$ that compresses the messages $(W_1,\cdots,W_m)$ into a sequence of length n of symbols in $\mathbb{F}$;
3) A set of m linear decoding functions from $\mathbb{F}^n\times\prod_{(i,j)\in\mathcal{E}}\mathcal{W}_j$ to $\mathcal{W}_i$ for $i=1,2,\cdots,m$.

The rate vector associated with the code is a vector $\bar{\mathbf{r}}=(r_1,\cdots,r_m)$ where $r_i=l_i/n$.

Definition 3. Linear and Non-Linear Index Coding: In linear index coding we restrict ourselves to linear codes over an arbitrary finite field $\mathbb{F}$. However in the non-linear index coding we are allowed to use an arbitrary code.

Definition 4. One-Shot and Asymptotic Index Coding: In the one-shot problem, we have fixed message alphabets $\mathcal{W}_1,\cdots,\mathcal{W}_m$ and seek the code with the smallest alphabet size for the public message (i.e., minimum size of the range of $f(\cdot)$) that can result in a zero probability of error. On the other hand, in the asymptotic coding scheme the rate vector $\bar{\mathbf{r}}=(r_1,\cdots,r_m)$ is called achievable if and only if there exists a sequence of zero-error codes whose blocklengths converge to infinity while their rate vectors converge to $\bar{\mathbf{r}}=(r_1,\cdots,r_m)$.

Remark 1: Asymptotic index coding is generally defined for a vanishing (rather than an exactly zero) probability of error. However [13] shows that the two definitions are equivalent.

Definition 5. Critical and Symmetric Rate Critical Graphs: Given an index coding problem (linear or non-linear/one-shot or asymptotic) on a graph, we say that the graph is *critical* if removal of any edge from it *strictly* shrinks the rate region (capacity, when we are looking at asymptotics) associated to the graph.

The maximum symmetric rate achievable on a graph is the supremum of r such that $\bar{\mathbf{r}}=(r,r,\cdots,r)$ is achievable. We say that the graph is *symmetric rate critical* if removal of any edge from it *strictly* reduces the maximum symmetric rate the graph. Every symmetric rate critical graph is critical, but the reverse is not necessarily true (see Theorem 4).

Next we need the following definitions from graph theory:

Definition 6. Turán Graph: Turán Graph of order m and k, denoted by $T(m,k)$, is a bidirectional complete k-partite graph with b parts of size $a+1$ and $k-b$ parts of size a, where $m=ak+b$ for $a\geq 0, b\in\{0,1,2\cdots,k-1\}$. We denote the number of edges of $T(m,k)$ by $e(m,k)$. In [14, Ex. 5.2.18], it is shown that

$$e(m,k)=\frac{1}{2}\cdot\left(1-\frac{1}{k}\right)m^2-\frac{b(k-b)}{2k}. \tag{1}$$

Lemma 1 (Turán's Theorem): [14, Thm. 5.2.9] A bidirectional m-vertex graph G that contains no clique of size $k+1$ has at most $e(m,k)$ edges. Furthermore, the only graph (up to the class of isomorphism) with $e(m,k)$ edges which satisfies the aforementioned condition is $T(m,k)$.

Definition 7. Strongly Connected Graphs: The graph $\mathsf{G}=(\mathcal{V},\mathcal{E})$ is strongly connected if there exists a directed path between every pair of distinct vertices.

It is easy to verify that a graph is strongly connected if and only if every edge of the graphs lies on a (directed) cycle.

Definition 8. Union of Two Disjoint Graphs: The union of $\mathsf{G}=(\mathcal{V},\mathcal{E})$ and $\mathsf{G}'=(\mathcal{V}',\mathcal{E}')$ is defined as $\mathsf{G}\cup\mathsf{G}'=(\mathcal{V}\cup\mathcal{V}',\mathcal{E}\cup\mathcal{E}')$.

Definition 9. USCS Graphs: Graph G is USCS (Union of Strongly Connected Subgraphs) if there exists a set of disjoint graphs $\{\mathsf{G}_1,\mathsf{G}_2,\cdots,\mathsf{G}_k\}$ such that (1)G_i is strongly connected and (2)$\mathsf{G}=\bigcup_i \mathsf{G}_i$.

III. Main Results

Theorem 1. Minimum Number of Edges for Equal Rates: Every m-vertex graph supporting a rate vector $\bar{\mathbf{r}}=(r,\cdots,r)$ has at least:

$$g(r,m)=m(m-1)-2\cdot e\left(m,\left\lfloor\frac{1}{r}\right\rfloor\right) \tag{2}$$

edges, if $\frac{1}{m}\le r\le 1$ ($g(r,m)$ is the number of edges in the complement of $T(m,\lfloor\frac{1}{r}\rfloor)$). Moreover, there is a unique graph, up to isomorphism, that has exactly $g(r,m)$ edges and supports the rate vector $\bar{\mathbf{r}}=(r,\cdots,r)$. This theorem holds for all cases (linear or non-linear, one-shot or asymptotic).

Remark 2: This theorem shows that there is a unique (up to isomorphism) critical graph with minimum number of edges for both one-shot and asymptotic cases.

Remark 3: Theorem 1 is valid for $\frac{1}{m}\le r\le 1$. For the case $r>1$, there is no graph that supports the rate vector $\bar{\mathbf{r}}=(r,\cdots,r)$ since the rate of each node cannot be greater than one. When $r<\frac{1}{m}$, it is possible to send all messages as the public message, and hence no side information is needed. Therefore, the empty graph is sufficient in this case.

Theorem 2. Additivity of Index Coding Capacity Region:

a) Given a graph $\mathsf{G}=(\mathcal{V},\mathcal{E})$, suppose that G' and G'' are subgraphs of G induced on vertex sets $\mathcal{V}'$ and $\mathcal{V}''$. In addition, assume that $\mathcal{V}'$ and $\mathcal{V}''$ partition $\mathcal{V}$ and there exist no edge like $e=(u,v)$ in $\mathcal{E}$ that starts from $u\in\mathcal{V}'$ and ends up in $v\in\mathcal{V}''$, i.e., no directed edge from G' to G'' exists. Then, elimination of all the directed edges from G'' to G' will not change the rate region in the one-shot linear, asymptotic linear, and asymptotic non-linear index coding problems. However this statement does not hold for all one-shot non-linear index coding problems.

b) [Optimality of a simple time-division strategy]. Take an index coding problem with graph $\mathsf{G}=\mathsf{G}'\bigcup\mathsf{G}''$, such that there is no edge between G' and G''. Let C, C' and C'' denote the capacity regions of G, G' and G'' respectively (the three capacities are either all in the sense of asymptotic linear, or all in the sense of asymptotic non-linear). Then $C=\bigcup_{\alpha\in[0,1]}\alpha C'\oplus(1-\alpha)C''$ where $\oplus$ is the direct sum operator. Alternatively, the index coding region for G is of the form $\bar{\mathbf{r}}=(\alpha\bar{\mathbf{r}}',(1-\alpha)\bar{\mathbf{r}}'')$ for $\alpha\in[0,1]$ and vector $\bar{\mathbf{r}}'$ is in the region of G' and $\bar{\mathbf{r}}''$ is in the region of G'', and $(\alpha\bar{\mathbf{r}}',(1-\alpha)\bar{\mathbf{r}}'')$ is the concatenation of the vectors $\alpha\bar{\mathbf{r}}'$ and $(1-\alpha)\bar{\mathbf{r}}''$.

c) Let $\mathsf{G}_1,\mathsf{G}_2,\cdots,\mathsf{G}_l$ be strongly connected components of G (as described in part b). Then $C=\bigcup_{\alpha_1+\cdots+\alpha_l=1}\alpha_1C_1\oplus\cdots\oplus\alpha_lC_l$ where C_i denotes the capacity regions of G_i (either all in the sense of asymptotic linear, or all in the sense of asymptotic non-linear).

Theorem 3. Critical Graphs are USCS:

a) Every critical graph for linear index coding (one-shot or asymptotic) and for asymptotic non-linear index coding is USCS. In particular, removing edges not lying on a directed cycle does not change the capacity region in these cases.

b) There exists a critical graph for a one-shot non-linear index coding problem which is not USCS.

Now, we provide some results on the structure of critical graphs. The first class of critical graphs that are easy to identify are bidirectional graphs:

Theorem 4: Any bidirectional graph is critical (by a bidirectional graph we mean one in which a directed edge from node i to j implies a directed edge from node j to i). On the other hand this is not true for symmetric criticality; in particular a bidirectional cycle of size 4 is not symmetric critical.

Theorem 5. Union of Two Critical Graphs is Critical: If G and H are two critical graphs with distinct vertex sets, then $\mathsf{G}\cup\mathsf{H}$ is also a critical graph for any of linear/non-linear, one-shot/asymptotic formulations. Further, if G and H are two symmetric rate critical graphs, then $\mathsf{G}\cup\mathsf{H}$ is also a symmetric rate critical in one-shot linear, asymptotic linear, and asymptotic non-linear index coding scenarios.

Theorem 6. Two Structures That are Critical:

a) Suppose $\mathsf{G}=(\mathcal{V},\mathcal{E})$ is a directed cycle of length m, where

$$\mathcal{V}=\{1,\cdots,m\},\quad \mathcal{E}=\{(i,i+1):1\le i<m\}\cup\{(m,1)\}.$$

Now, construct a new graph $\mathsf{G}'=(\mathcal{V}',\mathcal{E}')$ so that $\mathcal{V}'=\mathcal{V}\cup\{m+1\}$ and $\mathcal{E}'=\mathcal{E}\cup\{(m+1,1),(m+1,i),(j,m+1),(k,m+1)\}$. Then, if $1\le j<i$ and $i\le k\le m$, G' is symmetric rate critical.

b) Suppose $\mathsf{G}'=(\mathcal{V}',\mathcal{E}')$ is a graph that satisfies the condition of part (a). We construct a new graph $\mathsf{G}''=(\mathcal{V}'',\mathcal{E}'')$ by replacing any vertex $u\in\mathcal{V}'$ by a complete graph (different vertices can be replaced by complete graphs of different sizes). Then, G'' is critical. More specifically, we replace vertex u with n_u vertices $u_1,u_2,\cdots,u_{n_u}$ that are mutually connected to each other. We also draw a directed edge from u_i to v_j in G'' for $i\in\{1,2,\cdots,n_u\}$ and $j\in\{1,2,\cdots,n_v\}$ if there exists a directed edge from u to v in G'.

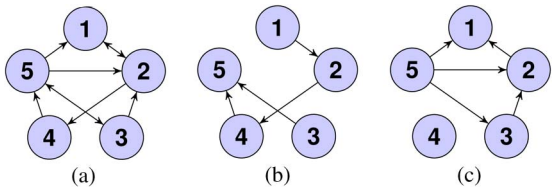

Fig. 2. An example of G, G^f, and G^b in the proof of Thm 1. (a) G. (b) G^f. (c) G^b.

IV. Proofs

A. Proof of Theorem 1

Before stating the proof, we introduce a useful lemma which is proved in [12]:

Lemma 2: Assume that $\mathcal{X}$ is a subset of the vertices of a graph $\mathsf{G}=(\mathcal{V},\mathcal{E})$ which contains no directed cycle. Then in every rate vector $\bar{\mathbf{r}}=(r_1,\cdots,r_m)$ supported by G in non-linear asymptotic case, we have $\sum_{i\in\mathcal{X}} r_i \leq 1$.

We begin the proof of Theorem 1 by proving the given lower bound on the minimum number of edges. It suffices to prove it for the non-linear asymptotic case since it implies that for all other cases (it is clear that the capacity regions of the other cases are the subset of the capacity region of non-linear asymptotic case. So, if $\bar{\mathbf{r}}$ is not in the non-linear asymptotic capacity region of the graphs with less than $g(m,r)$ edges, it is not in the capacity region of those graphs for other cases). Suppose that a given graph $\mathsf{G}=(\mathcal{V},\mathcal{E})$ supports the rate vector $\bar{\mathbf{r}}=(r,\cdots,r)$ for non-linear asymptotic case. We aim to construct two new graphs and with the help of Lemmas 1 and 2 find some bounds on the number of edges in these two graphs. Then we use these bounds to find a bound on the number of edges in G. Using Lemma 2, every subset of $\mathcal{V}(\mathsf{G})$ whose size is bigger than $\lfloor\frac{1}{r}\rfloor$, has a directed cycle, because the sum of the rates of the vertices in this subset is greater than or equal to $r\times(\lfloor\frac{1}{r}\rfloor+1)>1$. Then, we consider an arbitrary order for the vertices of G such as $1,\cdots,m$ and construct two new graphs (called "forward" and "backward" graphs) as follows: $\mathsf{G}^f=(\mathcal{V}^f,\mathcal{E}^f)$ and $\mathsf{G}^b=(\mathcal{V}^b,\mathcal{E}^b)$ where $\mathcal{V}^f=\mathcal{V}^b=\mathcal{V}$, and $\mathcal{E}^f,\mathcal{E}^b$ is a partition of $\mathcal{E}$ into two sets as follows: G^f contains those edges of G whose direction agrees with the mentioned order, that is, $\mathcal{E}^f=\{(x,y)\in\mathcal{E}|x<y\}$. G^b contains the following edges: $\mathcal{E}^b=\{(x,y)\in\mathcal{E}|x>y\}$ (for an example, see Fig. 2). Now, because every cycle in G should contain at least one edge from both G^f and G^b, every subset of size more than $\lfloor\frac{1}{r}\rfloor$ has at least one edge in both G^f and G^b.

Now let us construct a bidirectional graph $\widetilde{\mathsf{G}}^f$ on the same set of vertices as follows: x is connected to y in $\widetilde{\mathsf{G}}^f$ for $x\neq y$ if an only if $(\min(x,y),\max(x,y))\notin\mathcal{E}^f$. Observe that $\widetilde{\mathsf{G}}^f$ is like the complement of G^f if we ignore the edge arrows of G^f. Similarly, $\widetilde{\mathsf{G}}^b$ is constructed as the complement of G^b if we ignore the direction of arrows in it. Since every subset of size more than $\lfloor\frac{1}{r}\rfloor$ has at least one edge in both G^f and G^b, we can conclude that $\widetilde{\mathsf{G}}^f$ and $\widetilde{\mathsf{G}}^b$ do not have a clique of size $\lfloor\frac{1}{r}\rfloor+1$. Using Lemma 1, the number of edges of both G^f and G^b is at least

$$\binom{m}{2}-e\left(m,\left\lfloor\frac{1}{r}\right\rfloor\right)=\frac{g(r,m)}{2}. \tag{3}$$

Hence, G itself has at least $g(r,m)$ edges.

Next, we will show that the complement of $T(m,\lfloor\frac{1}{r}\rfloor)$ supports the rate $\bar{\mathbf{r}}$. It suffices to show this for one-shot linear coding and it implies that for all cases there exists a graph which supports the rate $\bar{\mathbf{r}}$. Let $m=a\lfloor\frac{1}{r}\rfloor+b$ for some $a\geq 0, b\in\{0,1,2\cdots,\lfloor\frac{1}{r}\rfloor-1\}$. Then we construct G as complement of $T(m,\lfloor\frac{1}{r}\rfloor)$ so that it consists of b cliques of size $a+1$, and $\lfloor\frac{1}{r}\rfloor-b$ cliques of size a.[1] Then one can verify that G has $g(r,m)$ edges. In addition, if every node desires only one bit and we transmit the XOR of the bits in every clique, every vertex can decode its message, and the rate of every message equals to $\frac{1}{\lfloor\frac{1}{r}\rfloor}\geq r$. Furthermore, it is obvious that this is a one-shot linear coding. Thus there is a graph which supports the rate $\bar{\mathbf{r}}$.

Lastly, to show that no other graph with exactly $g(r,m)$ edges supports $\bar{\mathbf{r}}$, consider a graph G that has $g(r,m)$ edges and supports the rate vector $\bar{\mathbf{r}}=(r,\cdots,r)$ in non-linear asymptotic case (it suffices to show this for the non-linear asymptotic case and it will imply other cases). According to our previous argument, if we construct $\widetilde{\mathsf{G}}^f$ and $\widetilde{\mathsf{G}}^b$ as discussed before, each of them should have exactly $e(m,\lfloor\frac{1}{r}\rfloor)$ edges. As they cannot have a clique of size $\lfloor\frac{1}{r}\rfloor+1$, Lemma 1 gives that they should have the structure mentioned in this lemma. So, the only remaining step is to show that the independent sets in $\widetilde{\mathsf{G}}^f$ and $\widetilde{\mathsf{G}}^b$ coincide on each other. Suppose this does not hold, that is, there are two vertices where there is an edge between them in $\widetilde{\mathsf{G}}^b$, but not in $\widetilde{\mathsf{G}}^f$. Let us call these two vertices u and v. (equivalently, there is an edge between u and v in G^f, but not in G^b). Choose one vertex from each of the $\lfloor\frac{1}{r}\rfloor$ independent sets of $\widetilde{\mathsf{G}}^f$ such that u is chosen and let us denote this set by $\mathcal{X}$. Then we claim that $\mathcal{X}\cup\{v\}$ does not contain any cycle in G. Note that if a cycle exists, it should include the edge between u and v, because it is the only edge in $\mathcal{X}\cup\{v\}$ in G^f and the cycle should have at least one edge from G^f. Now the other edges in the cycle form a path from v to u in G^b. As every component of G^b is a clique then u and v should have an edge, which contradicts our assumption that u and v are disconnected in G^b. ■

B. Proof of Theorem 2

The proof of part (a) for one-shot non-linear index coding follows from part (b) of Theorem 3. Other cases are considered below:

Proof of Part (a) for Asymptotic Non-Linear Index Coding: Consider an arbitrary code on the original graph with zero probability of error. Let $K=f(W_1,W_2,\cdots,W_m)$ be the public message. The rate of this code is $\bar{\mathbf{r}}=(r_1,r_2,\cdots,r_m)$ where $r_i=\log(|\mathcal{W}_i|)/\log(|\mathcal{K}|)$.

The union of G' and G'' corresponds to the graph G after elimination of directed edges from G'' to G'. Take an arbitrary $\varepsilon>0$. We create a code for the union of G' and G'' that achieves the rate vector $\bar{\mathbf{r}}'=(r'_1,r'_2,\cdots,r'_m)$ where $r'_i\geq r_i-\varepsilon$, with the probability of error being less than ε. This concludes the proof (see Remark 1 on index coding with a vanishing probability of error).

[1] A clique is a graph where every vertex has a directed edge to every other vertex.

We can conceive n i.i.d. repetitions of the given code with $(W_1^n, W_2^n, \cdots, W_m^n)$ and public message K^n. The rate of the i.i.d. code is the same as the original one since

$$\log\left(\left|\mathcal{W}_i^n\right|\right) = n\log\left(|\mathcal{W}_i|\right), \qquad \log\left(|\mathcal{K}^n|\right) = n\log\left(|\mathcal{K}|\right).$$

Since the original code had zero error probability, the i.i.d. code has also a zero probability of error.

We define $W_{\mathsf{G}'}$ as a shorthand for $W_i,\ i \in \mathsf{G}'$, and $W_{\mathsf{G}'}^n$ as a shorthand for $W_i^n,\ i \in \mathsf{G}'$. We define a new code that uses (K', K'') instead of K^n where nodes in G' decode their messages using K' and nodes in G'' decode their messages using K'':

- Size of the alphabet of K', i.e., $|\mathcal{K}'|$, is less than or equal to $2^{n(I(K;W_{\mathsf{G}'})+\delta)}$. Furthermore, the nodes in G' can use K' and their side information (which is inside G') to recover their message with probability $1-\varepsilon$.
- Size of the alphabet of K'', i.e., $|\mathcal{K}''|$, is less than or equal to $2^{n(H(K|W_{\mathsf{G}'})+\delta)}$. Furthermore, the nodes in G'' can use K'' and part of their side information of messages inside G'' to recover their message with probability $1-\varepsilon$.

This would finish the proof since $\log(|\mathcal{K}'|\cdot|\mathcal{K}''|)$ is equal to $n(\log(|\mathcal{K}|)+2\delta)$ and by choosing δ small enough we can ensure that the rate of the new code is within ε of the original code.

Construction of K'': We have $\min_{w_{\mathsf{G}'}} H(K|W_{\mathsf{G}'}=w_{\mathsf{G}'}) \le H(K|W_{\mathsf{G}'})$. Thus, it suffices to construct K'' whose alphabet size is less than or equal to $2^{n(H(K|W_{\mathsf{G}'}=w_{\mathsf{G}'})+\delta)}$ where $w_{\mathsf{G}'}$ is the one that minimizes $H(K|W_{\mathsf{G}'}=w_{\mathsf{G}'})$.

Let us first assume in the original problem that $W_{\mathsf{G}'}=w_{\mathsf{G}'}$ has occurred and the nodes in $\mathsf{G}''=\mathsf{G}-\mathsf{G}'$ are all aware of this (thus, if some of the nodes in G'' had partial information about messages of nodes in G', we are giving all of them a full access to $W_{\mathsf{G}'}$ and this should only help them in decoding their message). Thus the nodes in G'' should be able to recover their intended messages using K and their side information inside G'' with probability one, when $W_{\mathsf{G}'}=w_{\mathsf{G}'}$ is fixed. We can use the conditional joint pmf $p(K, W_{\mathsf{G}''}|W_{\mathsf{G}'}=w_{\mathsf{G}'})$ as a joint pmf $q(k, w_{\mathsf{G}''})$ on $K, W_{\mathsf{G}''}$ and think of it as an index code on nodes in G'', since $W_{\mathsf{G}''}$ is independent of $W_{\mathsf{G}'}$, the marginal distribution of $q(w_{\mathsf{G}''})$ is uniform and coordinatewise mutually independent; furthermore K will be a function of $W_{\mathsf{G}''}$ when $W_{\mathsf{G}'}=w_{\mathsf{G}'}$ is fixed. The public message in the index coding problem on G'' is produced from $q(k|W_{\mathsf{G}''})=p(k|W_{\mathsf{G}''}, W_{\mathsf{G}'}=w_{\mathsf{G}'})$ and it leads to zero error probability.

If we have n i.i.d. copies of the pmf q (still a code with zero error probability), the corresponding public message can be compressed using Shannon's source coding theorem and sent to the parties, where nodes in G'' can first decompress it and then use it to run their decoding algorithm. Compression can be achieved at a rate of $H_q(K)+\delta = H(K|W_H=w_H)+\delta$ bits at the cost of a probability of error of ε, which is tolerated.

Note that the public message K'' is only meant for the use of subgraph G''; to construct the code for G'' we have pretended that $W_{\mathsf{G}'}=w_{\mathsf{G}'}$ has happened in each copy of G'. It is clear that K'' contains no useful information about $W_{\mathsf{G}'}^n$ that has actually occurred, and nodes in G' can ignore K''.

Construction of K': Let $p(k, w_{\mathsf{G}'})$ denote the joint distribution of K and $W_{\mathsf{G}'}$ in the original code. The decoding function used by node $i \in \mathsf{G}'$ can be expressed as the conditional pmf $p(\hat{w}_i|k, (w_j)_{j:(i,j)\in\mathcal{E}})$ where $\hat{W}_i$ is the reconstruction of node i. Of course $\hat{W}_i = W_i$ since perfect reconstruction is assumed. Therefore the joint pmf

$$p(k, w_{\mathsf{G}'}, \hat{w}_{\mathsf{G}'}) = p(k, w_{\mathsf{G}'}) \prod_{i\in\mathsf{G}'} p(\hat{w}_i|k, (w_j)_{j:(i,j)\in\mathcal{E}})$$

has the property that the marginal distribution on $W_{\mathsf{G}'}$ and $\hat{W}_{\mathsf{G}'}$ is equal to

$$p(W_{\mathsf{G}'}=w_{\mathsf{G}'}, \hat{W}_{\mathsf{G}'}=\hat{w}_{\mathsf{G}'}) = \prod_{i\in\mathsf{G}'} \mathbf{1}[w_i=\hat{w}_i]. \tag{4}$$

We use the covering lemma (rate-distortion coding) to create a code for nodes in G'. Let $\delta>0$ be an arbitrary small real.

Codebook Generation: Assume that the transmitter and the receivers initially share a codebook of $2^{n(I(K;W_{\mathsf{G}'})+\delta)}$ sequences

$$K^n(1), K^n(2), \cdots, K^n\left(2^{n(I(K;W_{\mathsf{G}'})+\delta)}\right)$$

each being an i.i.d. sequence according to $p(k)$.

Encoding: Having $W_{\mathsf{G}'}^n$ at the transmitter, it finds an index j such that $K^n(j)$ is jointly typical with $W_{\mathsf{G}'}^n$ (i.e., $(K^n(j), W_{\mathsf{G}'}^n) \in \mathcal{T}_\delta^n(p(k, w_{\mathsf{G}'}))$), where we use the notion of typicality given in [1, 2.4]. Since the number of generated $K^n(\cdot)$ sequences is larger than $2^{n(I(K;W_{\mathsf{G}'})+\delta)}$ by the covering lemma [1, Lemma 3.3], this can be done with high probability. The transmitter then sends the index j as K' to the receiver (the cardinality of the alphabet of K' allows it to send the index j).

Decoding: Having received $K'=j$, nodes $i \in \mathsf{G}'$ create $\hat{W}_i^n$ as a function of $K^n(j)$ and their side information (they use the same decoding functions of the original code). More precisely, if we denote the joint pmf of $K^n(j)$ and $W_{\mathsf{G}'}^n$ by $q_{K^n(j),W_{\mathsf{G}'}^n}(k, w_{\mathsf{G}'}^n)$, the joint pmf of the constructed rv's is equal to

$$q_{K^n(j),W_{\mathsf{G}'}^n}(k, w_{\mathsf{G}'}^n) \prod_{i\in\mathsf{G}'}\prod_{s=1}^{n} p\left(\hat{w}_{is}|k_s, (w_{js})_{j:(i,j)\in\mathcal{E}}\right)$$

If $(K^n(j), W_{\mathsf{G}'}^n) \in \mathcal{T}_\delta^n(p(k, w_{\mathsf{G}'}))$, with high probability we will have $(K^n(j), W_{\mathsf{G}'}^n, \hat{W}_{\mathsf{G}'}^n) \in \mathcal{T}_{\delta'}^n(p(k, w_{\mathsf{G}'}, \hat{w}_{\mathsf{G}'}))$ for any $\delta'>\delta$, as we have passed $K^n(j)$, $W_{\mathsf{G}'}^n$ through the i.i.d. conditional pmf of $p(\hat{w}_{\mathsf{G}'}|k, w_{\mathsf{G}'})$ (Conditional typicality lemma [1, 2.5]). Therefore $K^n(j)$, $W_{\mathsf{G}'}^n$, $\hat{W}_{\mathsf{G}'}^n$ will be joint typical with high probability. Thus for any $i \in \mathsf{G}'$, with high probability $(W_i^n, \hat{W}_i^n)$ will be jointly typical. We claim that two sequences $(W_i^n, \hat{W}_i^n)$ jointly typicality in the sense of [1, 2.4] is equivalent with their equality. Equation (4) implies that $p(W_{\mathsf{G}'}=w_{\mathsf{G}'}, \hat{W}_{\mathsf{G}'}=\hat{w}_{\mathsf{G}'})>0$ if and only if $w_{\mathsf{G}'}=\hat{w}_{\mathsf{G}'}$, and hence for any pair $(w_{\mathsf{G}'}, \hat{w}_{\mathsf{G}'})$ where $w_{\mathsf{G}'} \neq \hat{w}_{\mathsf{G}'}$ we have (using notation of [1]) that

$$\left|\Pi\left(w_{\mathsf{G}'}, \hat{w}_{\mathsf{G}'}|W_i^n, \hat{W}_i^n\right) - p(w_{\mathsf{G}'}, \hat{w}_{\mathsf{G}'})\right| \le \delta' \cdot p(w_{\mathsf{G}'}, \hat{w}_{\mathsf{G}'}) = 0.$$

Hence $\Pi(w_{\mathsf{G}'}, \hat{w}_{\mathsf{G}'}|W_i^n, \hat{W}_i^n) = p(w_{\mathsf{G}'}, \hat{w}_{\mathsf{G}'}) = 0$ for any $w_{\mathsf{G}'} \neq \hat{w}_{\mathsf{G}'}$, implying that $W_i^n = \hat{W}_i^n$. Therefore with high probability the decoders will successfully decode their intended messages. □

Proof of Part (a) for One-Shot Linear Index Coding: Assume that there exists a valid one-shot linear coding scheme for a graph G with $|\mathcal{V}| = m$ vertices such that $W_i = (w_{i1}, w_{i2}, \cdots, w_{il_i})$, where $w_{ij} \in \mathbb{F}$ for some field $\mathbb{F}$. Additionally, assume that $f(W_1, W_2, \cdots, W_m) = (t_1, t_2, \cdots, t_n)$ where t_k is equal to

$$t_k = \sum_{i=1}^{m} \sum_{j=1}^{l_i} c_{ijk} \cdot w_{ij}, \qquad \forall 1 \le k \le n, \tag{5}$$

for some coefficients c_{ijk} in the field $\mathbb{F}$. In other words, the following matrix is used for the linear map:

$$\mathsf{C} = \begin{bmatrix} c_{111} & c_{121} & \cdots & c_{1l_11} & c_{211} & \cdots & c_{ml_m1} \\ & & & \vdots & & & \\ c_{11n} & c_{12n} & \cdots & c_{1l_1n} & c_{21n} & \cdots & c_{ml_mn} \end{bmatrix}.$$

Without loss of generality we can assume that C is in the row echelon form, since elementary row operation on C is equivalent to using *invertible* linear combinations of $t_1, t_2, \cdots, t_n$ instead of these variables. We represent the first non-zero elements of each row of C by a sequence of indices

$$(\mathtt{i}_k, \mathtt{j}_k), \qquad k = 1, 2, \cdots, n, \mathtt{j}_k \le l_{\mathtt{i}_k} \tag{6}$$

that are increasing in a lexicographical order, i.e., either $\mathtt{i}_k < \mathtt{i}_{k+1}$ holds or both $\mathtt{i}_k = \mathtt{i}_{k+1}$ and $\mathtt{j}_k < \mathtt{j}_{k+1}$ hold. Further we must have $c_{ijk} = 0$ if (i, j) is less than $(\mathtt{i}_k, \mathtt{j}_k)$ in the lexical order.

Since all nodes are able to decode their messages via $(t_1, \cdots, t_n)$ and their side information, there should exist coefficients $\alpha_{ij1}, \alpha_{ij2}, \cdots, \alpha_{ijn}$ for each message w_{ij} $(1 \le j \le l_i)$ so that:

$$\sum_{k=1}^{n} \alpha_{ijk} t_k \tag{7}$$

is equal to w_{ij} plus a linear combination of $w_{i'j'}$ that are available to node i as side information, i.e.,

$$\sum_{k=1}^{n} \alpha_{ijk} t_k = w_{ij} + \sum_{i', j' : (i,i') \in \mathcal{E}} w_{i'j'} \cdot \gamma_{i'j'}, \tag{8}$$

for some coefficients $\gamma_{i'j'}$.

Now, we turn to the proof. Without loss of generality, suppose that the vertices of G' are $m - |\mathcal{V}'| + 1, m - |\mathcal{V}'| + 2, \cdots, m$, where G' was defined in the statement of the theorem. Note that this assumption and the assumption that C is in the row echelon form do not contradict the generality together. One can simply label the vertices of G such that nodes in G' be labeled with $m - |\mathcal{V}'| + 1, m - |\mathcal{V}'| + 2, \cdots, m$ and then applies some elementary row operations to find C in row echelon form. The statement of the theorem basically asks us to show that there is no need for nodes in G'' to know (as side information) any of the messages for nodes in G', i.e., W_i, $i \in \mathsf{G}'$. To show this, we first define a new encoding linear map f' and then prove that it enables nodes in G'' to recover their intended messages without any need to have access to W_i, $i \in \mathsf{G}'$. Nodes in G' are also shown to be still able to decode their messages with the encoding function f' using their side information (nodes in G' do not know any of the messages of nodes in G'' since G' does not have any outgoing edge). Thus, the edges between G' and G'' can be removed.

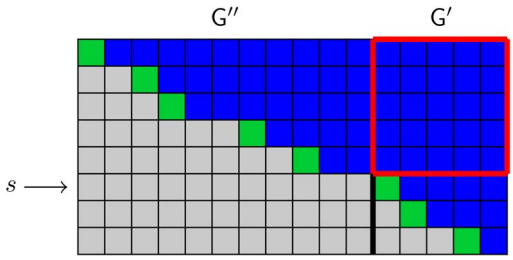

Fig. 3. A pictorial representation of the row echelon form of C that clarifies the definition of s. Gray elements are zero and green elements are non-zero. To define c'_{ijk} we set the values inside the red box to be zero.

Part 1: Definition of a new linear encoding function f': Let $s = \min\{k \mid c_{ijk} = 0, \forall 1 \le i \le m - |\mathcal{V}'|, 1 \le j \le l_i\}$. Fig. 3 clarifies the definition of the s. Note that $\{k \mid c_{ijk} = 0, \forall 1 \le i \le m - |\mathcal{V}'|, 1 \le j \le l_i\}$ cannot be empty. We prove this by contradiction. Suppose that the mentioned set is empty, then for each $1 \le k \le n$ there exists a $c_{ijk} \ne 0$ that $i \le m - |\mathcal{V}'|$. As a result,

$$\forall 1 \le k \le n : \mathtt{i}_k \le m - |\mathcal{V}'|. \tag{9}$$

Now assume that θ is the smallest number that $\alpha_{m1\theta} \ne 0$, then:

$$\begin{aligned} \sum_{k=1}^{n} \alpha_{m1k} \cdot t_k &= \sum_{k=\theta}^{n} \alpha_{m1k} \cdot t_k \\ &= \sum_{k=\theta}^{n} \alpha_{m1k} \cdot \left(\sum_{p=1}^{m} \sum_{q=0}^{l_p} c_{pqk} \cdot w_{pq} \right) \end{aligned}$$

Note that the coefficient of $w_{\mathtt{i}_\theta \mathtt{j}_\theta}$ in the above statement is $\sum_{k=\theta}^{n} \alpha_{m1k} \cdot c_{\mathtt{i}_\theta \mathtt{j}_\theta k}$. Because $(\mathtt{i}_\theta, \mathtt{j}_\theta)$ is lexicographically smaller than $(\mathtt{i}_k, \mathtt{j}_k)$ for any $k > \theta$, $c_{\mathtt{i}_\theta \mathtt{j}_\theta k} = 0$. Then

$$\sum_{k=\theta}^{n} \alpha_{m1k} \cdot c_{\mathtt{i}_\theta \mathtt{j}_\theta k} = \alpha_{m1\theta} \cdot c_{\mathtt{i}_\theta \mathtt{j}_\theta \theta} \ne 0 \tag{10}$$

Note that $\mathtt{i}_\theta \le m - |\mathcal{V}'|$ by (9), so $\mathtt{i}_\theta \in \mathsf{G}''$. Hence, $w_{\mathtt{i}_\theta \mathtt{j}_\theta}$ is not provided as side information to node m, which is in G'. As a result, the non-zero coefficient of $w_{\mathtt{i}_\theta \mathtt{j}_\theta}$ in $\sum_{k=1}^{n} \alpha_{m1k} \cdot t_k$ is in contradiction to (8). So we have proved that $\{k \mid c_{ijk} = 0, \quad \forall 1 \le i \le m - |\mathcal{V}'|, \quad 1 \le j \le l_i\}$ is a non-empty set and thus s is well-defined.

Further let

$$c'_{ijk} = \begin{cases} 0 & k < s \quad \text{and} \quad i > m - |\mathcal{V}'| \\ c_{ijk} & \text{otherwise} \end{cases} \tag{11}$$

and $t'_k = \sum_{i=1}^{m} \sum_{j=1}^{l_i} c'_{ijk} \cdot w_{ij}$ for all $1 \le k \le n$. Set $f'(W_1, W_2, \cdots, W_m) = (t'_1, t'_2, \cdots, t'_n)$. Observe that (11) implies that $t'_k = t_k$ for all $k \ge s$.

Part 2: Showing that nodes in G' are able to decode their message by using $(t'_1, t'_2, \cdots, t'_n)$ and their side information: Consider the coefficients α_{ijk} for decoding of the original linear mapping given in (7). We claim for any $r \in \mathsf{G}'$ (i.e.,

$r > m - |\mathcal{V}'|$) that $\alpha_{rj1} = \cdots = \alpha_{rj(s-1)} = 0$ for every $1 \le j \le l_r$. This completes the proof since for every $r \in \mathsf{G}'$:

$$\sum_{k=1}^{n} \alpha_{rjk} t_k = \sum_{k=s}^{n} \alpha_{rjk} t_k = \sum_{k=s}^{n} \alpha_{rjk} t'_k. \tag{12}$$

Equations (8) and (12) illustrate that every node $r \in \mathsf{G}'$ is able to obtain w_{rj} in the new coding scheme by calculating $\sum_{k=s}^{n} \alpha_{rjk} t'_k$.

We prove $\alpha_{rj1} = \cdots = \alpha_{rj(s-1)} = 0$ for every $1 \le j \le l_r$ by contradiction. Suppose that x is the smallest index that $\alpha_{rjx} \neq 0$ and $x < s$. By the definition of s, $\mathtt{i}_x \le m - |\mathcal{V}'|$ ($\mathtt{i}_x \in \mathsf{G}''$). It is also clear that the definition of x results in the fact that $\alpha_{rjy} = 0$ for all $y < x$. Additionally, because $(\mathtt{i}_k, \mathtt{j}_k), k = 1, 2, \cdots, n$ is strictly increasing $c_{\mathtt{i}_x \mathtt{j}_x y} = 0$ for all $y > x$. Thus, as

$$\sum_{k=1}^{n} \alpha_{rjk} \cdot t_k = \sum_{k=1}^{n} \alpha_{rjk} \cdot \left(\sum_{p=1}^{m} \sum_{q=1}^{l_p} c_{pqk} \cdot w_{pq} \right), \tag{13}$$

the coefficient of $w_{\mathtt{i}_x \mathtt{j}_x}$ in $\sum_{k=1}^{n} \alpha_{rjk} \cdot t_k$ is:

$$\begin{aligned}
\bar{\mathbf{r}}_\varepsilon &= \left(\frac{\log_{\mathbb{F}} |\mathcal{W}_1|}{n' + n''}, \frac{\log_{\mathbb{F}} |\mathcal{W}_2|}{n' + n''}, \cdots, \frac{\log_{\mathbb{F}} \left| \mathcal{W}_{|\mathcal{V}'|+|\mathcal{V}'''|} \right|}{n' + n''} \right) \\
&= \left(\frac{n'}{n' + n''} \left(\frac{\log_{\mathbb{F}} |\mathcal{W}_1|}{n'}, \cdots, \frac{\log_{\mathbb{F}} \left| \mathcal{W}_{|\mathcal{V}'|} \right|}{n'} \right), \frac{n''}{n' + n''} \right. \\
&\qquad \left. \times \left(\frac{\log_{\mathbb{F}} \left| \mathcal{W}_{|\mathcal{V}'|+1} \right|}{n''}, \cdots, \frac{\log_{\mathbb{F}} \left| \mathcal{W}_{|\mathcal{V}'|+|\mathcal{V}'''|} \right|}{n''} \right) \right) \\
&= \left(\frac{n'}{n' + n''} \bar{\mathbf{r}}', \left(1 - \frac{n'}{n' + n''} \right) \bar{\mathbf{r}}'' \right).
\end{aligned} \tag{14}$$

$$\begin{aligned}
&\sum_{k=1}^{n} \alpha_{rjk} \cdot c_{\mathtt{i}_x \mathtt{j}_x k} \\
&\quad = \sum_{k=1}^{x-1} 0 \cdot c_{\mathtt{i}_x \mathtt{j}_x k} + \alpha_{rjx} \cdot c_{\mathtt{i}_x \mathtt{j}_x x} + \sum_{k=x+1}^{n} \alpha_{rjk} \cdot 0 && (15) \\
&\quad = \alpha_{rjx} \cdot c_{\mathtt{i}_x \mathtt{j}_x x} \neq 0 && (16)
\end{aligned}$$

which is in contradiction to the independency of $\sum_{k=1}^{n} \alpha_{rjk} t_k$ from $w_{\mathtt{i}_x \mathtt{j}_x}$, that was guaranteed by (8). (Note that $r \in \mathsf{G}'$ and $\mathtt{i}_x \in \mathsf{G}''$, so r does not have $w_{\mathtt{i}_x \mathtt{j}_x}$ as side information).

Part 3: Showing that under f' decoding is possible without the need for nodes in G'' to know messages for nodes in G': For every $i \in \mathsf{G}''$, let

$$\beta_{ijk} = \begin{cases} 0 & k \ge s; \\ \alpha_{ijk} & k < s. \end{cases} \tag{17}$$

We claim that for every $i \in \mathsf{G}''$:

$$\sum_{k=1}^{n} \beta_{ijk} t'_k = w_{ij} + \sum_{i', j' : (i,i') \in \mathcal{E} \text{ and } i' \notin \mathsf{G}'} w_{i'j'} \cdot \gamma_{i'j'}, \tag{18}$$

where $\gamma_{i'j'}$ is given in (8). This shows that nodes at G'' are able to decode their messages using $(t'_1, t'_2, \cdots, t'_n)$ and their side information in G'' (excluding side information from nodes at G'). We have:

$$\begin{aligned}
\sum_{k=1}^{n} \beta_{ijk} t'_k &= \sum_{k=1}^{s-1} \alpha_{ijk} t'_k + \sum_{k=s}^{n} 0 \cdot t'_k && (19) \\
&= \sum_{k=1}^{s-1} \alpha_{ijk} \cdot \left(\sum_{p=1}^{m} \sum_{q=0}^{l_p} c'_{pqk} \cdot w_{pq} \right) && (20) \\
&= \sum_{k=1}^{s-1} \alpha_{ijk} \cdot \left(\sum_{p=1}^{m-|\mathcal{V}'|} \sum_{q=0}^{l_p} c'_{pqk} \cdot w_{pq} \right) && (21) \\
&= \sum_{k=1}^{s-1} \alpha_{ijk} \cdot \left(\sum_{p=1}^{m-|\mathcal{V}'|} \sum_{q=0}^{l_p} c_{pqk} \cdot w_{pq} \right), && (22)
\end{aligned}$$

where (21) and (22) follow from the definition of c'_{ijk} in (11). Note that the expression of (22) does not include any of w_{ij} for $i > m - |\mathcal{V}'|$. Moreover, the coefficient of w_{ij} for $i \le m - |\mathcal{V}'|$ are the same as those in $\sum_{k=1}^{n} \alpha_{ijk} t_k$. This establishes (18). □

Proof of Part (a) for Asymptotic Linear Index Coding: Assume that $\bar{\mathbf{r}} = (r_1, \cdots, r_m)$ is an achievable rate vector in asymptotic linear index coding problem associated with G. Then there exists a sequence of codes like $C_1, C_2, \cdots$ whose rate vector converges to $\bar{\mathbf{r}}$. Using Theorem 2 part a for one-shot linear index coding, for each C_i there exists a code like C'_i for the index coding problem introduced by $\mathsf{G}' \cup \mathsf{G}''$ whose rate vector equals the rate vector of C_i. Thus, the rate vector of $C'_1, C'_2, \cdots$ converges to $\bar{\mathbf{r}}$. This results in the fact that $\bar{\mathbf{r}} = (r_1, \cdots, r_m)$ is achievable in asymptotic linear index coding problem associated with $\mathsf{G}' \cup \mathsf{G}''$. □

Proof of Part (b): The proof has two parts: first we show that $\bigcup_{\alpha \in [0,1]} \alpha C' \oplus (1-\alpha) C'' \subseteq C$ and then we will finish the proof by showing that $C \subseteq \bigcup_{\alpha \in [0,1]} \alpha C' \oplus (1-\alpha) C''$.

Before starting the proof, let us label the vertices of G so that the vertices of G' come first.

Proving $\bigcup_{\alpha \in [0,1]} \alpha C' \oplus (1-\alpha) C'' \subseteq C$: Take an arbitrary vector $\bar{\mathbf{r}}'$ in C'. Then we can allocate all of our resources for G' and do not send anything for G''. This shows that $(\bar{\mathbf{r}}', 0)$ is in C. Similarly for any $\bar{\mathbf{r}}''$ in C'', we have that $(0, \bar{\mathbf{r}}'')$ is in C. Using the standard time-sharing techniques, one can show that the capacity region of the index coding problem is a convex set. Therefore for any $\alpha \in [0,1]$ the rate $(\alpha \bar{\mathbf{r}}', (1-\alpha) \bar{\mathbf{r}}'') \in C$. This completes the proof.

Proving $C \subseteq \bigcup_{\alpha \in [0,1]} \alpha C' \oplus (1-\alpha) C''$: For any rate vector $\bar{\mathbf{r}} \in C$, there exist a sequence of codes like $C_1, C_2, \cdots$ whose rates converge to $\bar{\mathbf{r}}$. Take some $\varepsilon > 0$ and a code described by encoding function f whose rate $\bar{\mathbf{r}}_\varepsilon$ is within ε distance of $\bar{\mathbf{r}}$.

Linear Case: Suppose that $f : \mathcal{W}_1 \times \mathcal{W}_2 \times \cdots \times \mathcal{W}_{|\mathcal{V}'|+|\mathcal{V}'''|} \to \mathbb{F}^n$. In the proof of the part (a), we showed that there exist encoding functions $f' : \mathcal{W}_1 \times \mathcal{W}_2 \times \cdots \times \mathcal{W}_{|\mathcal{V}'|} \to \mathbb{F}^{n'}$ and $f'' : \mathcal{W}_{|\mathcal{V}'|+1} \times \mathcal{W}_{|\mathcal{V}'|+2} \times \cdots \times \mathcal{W}_{|\mathcal{V}'|+|\mathcal{V}'''|} \to \mathbb{F}^{n''}$ which are respectively valid for G' and G''. Additionally, the size of the range of the concatenation of f' and f'' equals the size of the range of f, i.e., $n = n' + n''$. Hence, if we call the rates of f' and f'', $\bar{\mathbf{r}}'$ and $\bar{\mathbf{r}}''$, $\bar{\mathbf{r}}_\varepsilon$ equals the expression given in (14).

Since $\bar{\mathbf{r}}' \in C'$ and $\bar{\mathbf{r}}'' \in C''$, above statement results in the fact that $\bar{\mathbf{r}}_\varepsilon$ lies in $\bigcup_{\alpha \in [0,1]} \alpha C' \oplus (1-\alpha) C''$. By the definition of the

asymptotic capacity region, C' and C'' are closed sets. We are done with the proof by noting that $\bar{\mathbf{r}}_\varepsilon$ can be made arbitrarily close to $\bar{\mathbf{r}}$.

Non-Linear Case: Suppose that $f : \mathcal{W}_1 \times \mathcal{W}_2 \times \cdots \times \mathcal{W}_{|\mathcal{V}'|+|\mathcal{V}''|} \to \{1,2,\cdots,N\}$. In the proof of the part (a), we showed that we can find encoding functions $f' : \mathcal{W}_1^n \times \mathcal{W}_2^n \times \cdots \times \mathcal{W}_{|\mathcal{V}'|}^n \to \{1,2,\cdots,2^{n(K'+\delta)}\}$ and $f'' : \mathcal{W}_{|\mathcal{V}'|+1}^n \times \mathcal{W}_{|\mathcal{V}'|+2}^n \times \cdots \times \mathcal{W}_{|\mathcal{V}'|+|\mathcal{V}''|}^n \to \{1,2,\cdots,2^{n(K''+\delta)}\}$ in which K' and K'' satisfy $N = 2^{K'+K''}$, δ is an arbitrary positive real number, and an appropriate n can be found for any fixed δ so that such functions exist. Moreover, f' and f'' are respectively valid for G' and G'' over the alphabet sets of $\mathcal{W}_1^n, \mathcal{W}_2^n, \cdots, \mathcal{W}_{|\mathcal{V}'|}^n$ and $\mathcal{W}_{|\mathcal{V}'|+1}^n, \mathcal{W}_{|\mathcal{V}'|+2}^n, \cdots, \mathcal{W}_{|\mathcal{V}'|+|\mathcal{V}''|}^n$. Hence, if we call the rates of f' and f'', $\bar{\mathbf{r}}'$ and $\bar{\mathbf{r}}''$, $\bar{\mathbf{r}}_\varepsilon$ equals:

$$\begin{aligned}
&\left(\frac{\log|\mathcal{W}_1|}{\log N}, \frac{\log|\mathcal{W}_2|}{\log N}, \cdots, \frac{\log\left|\mathcal{W}_{|\mathcal{V}'|+|\mathcal{V}''|}\right|}{\log N}\right)\\
&= \left(\frac{\log|\mathcal{W}_1^n|}{n\log N}, \frac{\log|\mathcal{W}_2^n|}{n\log N}, \cdots, \frac{\log\left|\mathcal{W}_{|\mathcal{V}'|+|\mathcal{V}''|}^n\right|}{n\log N}\right)\\
&= \left(\frac{\log|\mathcal{W}_1^n|}{n(K'+K'')}, \frac{\log|\mathcal{W}_2^n|}{n(K'+K'')}, \cdots, \frac{\log\left|\mathcal{W}_{|\mathcal{V}'|+|\mathcal{V}''|}^n\right|}{n(K'+K'')}\right)\\
&= \left(\frac{K'+\delta}{K'+K''}\left(\frac{\log|\mathcal{W}_1^n|}{n(K'+\delta)}, \cdots, \frac{\log\left|\mathcal{W}_{|\mathcal{V}'|}^n\right|}{n(K'+\delta)}\right),\right.\\
&\qquad \left.\frac{K''+\delta}{K'+K''}\left(\frac{\log\left|\mathcal{W}_{|\mathcal{V}'|+1}^n\right|}{n(K''+\delta)}, \cdots, \frac{\log\left|\mathcal{W}_{|\mathcal{V}'|+|\mathcal{V}''|}^n\right|}{n(K''+\delta)}\right)\right)\\
&= \left(\frac{K'+\delta}{K'+K''}\bar{\mathbf{r}}', \frac{K''+\delta}{K'+K''}\bar{\mathbf{r}}''\right)\\
&= \frac{K'+K''+2\delta}{K'+K''}\left(\frac{K'+\delta}{K'+K''+2\delta}\bar{\mathbf{r}}', \frac{K''+\delta}{K'+K''+2\delta}\bar{\mathbf{r}}''\right)
\end{aligned}$$

Thus,

$$\frac{K'+K''}{K'+K''+2\delta}\bar{\mathbf{r}}_\varepsilon = \left(\frac{K'+\delta}{K'+K''+2\delta}\bar{\mathbf{r}}', \frac{K''+\delta}{K'+K''+2\delta}\bar{\mathbf{r}}''\right).$$

As $\bar{\mathbf{r}}' \in C'$ and $\bar{\mathbf{r}}'' \in C''$, $\frac{K'+K''}{K'+K''+2\delta}\bar{\mathbf{r}}_\varepsilon$ lies in $\bigcup_{\alpha\in[0,1]} \alpha C' \oplus (1-\alpha)C''$ for any $\delta > 0$. Since $\bigcup_{\alpha\in[0,1]} \alpha C' \oplus (1-\alpha)C''$ is a closed set and we can make δ and ε as close to zero as we want, we will be done. ∎

Proof of Part (c): We prove this part using induction on l. If $l = 1$, then $\mathsf{G} = \mathsf{G}_1$ and $C = C_1$ so we are done. Now assume that $l > 1$. Lets call the vertex set of G_i, $\mathcal{V}_i$. It is straightforward to see that every directed graph has a strongly connected component that does not contain any outgoing edges, i.e., there exists i so that there is no edge in G that starts from the vertices of $\mathcal{V}_i$ and ends in $\mathcal{V} - \mathcal{V}_i$. Set $\mathcal{V}' = \mathcal{V}_i$ and $\mathcal{V}'' = \mathcal{V} - \mathcal{V}_i$. We denote the subgraph of G induced on $\mathcal{V}''$ by $\tilde{\mathsf{G}}$. Using parts (a) or (b) on $\mathcal{V}'$ and $\mathcal{V}''$ we have:

$$C = \bigcup_{\alpha_i\in[0,1]} \alpha_i C_i \oplus (1-\alpha_i)\tilde{C} \tag{23}$$

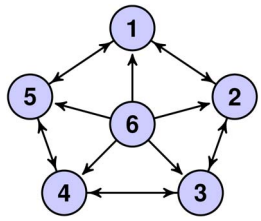

Fig. 4. In the index coding problem associated with this graph, removing the edges which belong to no cycle implies a larger public message rate.

where $\tilde{C}$ is the capacity region of $\tilde{\mathsf{G}}$. Since $\tilde{\mathsf{G}}$ has exactly $l-1$ strongly connected components $(\mathsf{G}_1, \cdots, \mathsf{G}_{i-1}, \mathsf{G}_{i+1}, \cdots, \mathsf{G}_l)$, by induction we have:

$$C = \bigcup_{\sum_{1\le j\le n, j\neq i}\alpha_j = 1} \left(\bigoplus_{1\le j\le n, j\neq i} \alpha_j C_i\right) \tag{24}$$

(23) and (24) clearly complete the proof.

C. Proof of Theorem 3

1) Proof of Part (a)—Linear one-shot case: If G is USCS, then proof is finished. Otherwise, G contains an edge like $e = (u, v)$ which is not located in any cycle. Let $\mathcal{V}_1$ be the set of vertices that can be reached from v. Moreover, let $\mathcal{V}_2$ be the set of vertices that cannot be reached from v. It is easy to verify that there will be no edge that starts from $\mathcal{V}_1$ and finishes in $\mathcal{V}_2$. Using part (a) of Theorem 2, we can remove all edges between $\mathcal{V}_1$ and $\mathcal{V}_2$ including e, so that the rate region does not shrink. As the number of the edges of G is finite, by repeating this process we can find a USCS subgraph of G like G' whose rate region equals the rate region of G. Hence, if G is a critical graph, it should be equal to G' which is USCS. In other words, any critical graph for one-shot linear index coding is USCS.

The proof for *Linear asymptotic* and *Non-linear asymptotic index coding* using part (a) of Theorem 2 is similar.

2) Proof of Part (b): To prove this part we need to show that a critical graph exists for one-shot non-linear case that is not USCS.

Consider the graph given in Fig. 4. We call this graph $\mathsf{G} = (\mathcal{V}, \mathcal{E})$. Assume that $\mathcal{W}_i = \{0,1\}$ for all $1 \le i \le 5$ and $\mathcal{W}_6 = \{0,1,2,3,4\}$. We have the following claim.

Claim 1: Sending a symbol from $\{1,2,\cdots,32\}$ as the public message suffices for every node to decode its message. However, if we remove the edges connected to node 6, which do not belong to any cycle, we need at least 35 symbols to have a successful transmission of the messages.

This claim establishes the desired result, since if G is critical it would be an instance of a non-USCS graph that is critical. If G is not critical, there is a subgraph G' of it (obtained by removing edges from G) that is critical; that is the graph G' is such that sending a symbol from $\{1,2,\cdots,32\}$ as the public message suffices for every node to decode its message. However any further removal of edges from G' results in a graph that does not have this property. By the above claim, the minimal graph G' should contain at least one of the edges connected to the node 6; since if not, G' would be a subgraph of the graph shown in the claim to need at least 35 symbols. Therefore, G' contains

an edge that is not on any cycle. Hence it is a non-USCS and critical graph.

We now turn to the proof of the claim. In order to construct the coding scheme using 32 symbols for G, first note that $W_1W_2W_3W_4W_5$ forms a binary sequence of the length 5. Based on the value of W_6, we XOR this sequence with one the following sequence: 00000, 10001, 01111, 01100, 10111, that is, if W_6 is 0 we XOR the sequence with 00000, if it is 1 we XOR it with 10001, and so on. Then, we transmit the result as the public message (the public message has 32 different possibilities and can be transmitted). Let us denote the 5-bit public message by $\tilde{W}_1\tilde{W}_2\tilde{W}_3\tilde{W}_4\tilde{W}_5$. It is sufficient to show that every node can decode its message with the help of the public message and its side information. First of all, because the node 6 knows the message of 1 to 5, it can XOR their message by the public message and from the XOR decode its message. For the other nodes, note that $W_i \oplus \tilde{W}_i$ for $1 \leq i \leq 5$ is a function of the side information of node i, and therefore, node i can decode its message. We explain the decoding process for node 1; the decoding process for other nodes is similar. Node 1 knows W_2 and W_5. By comparing these two bits with $\tilde{W}_2$ and $\tilde{W}_5$, node 1 can exactly recover W_6 if it is equal to 0, 2 or 3. If W_6 is equal to 1 or 4, node 1 cannot find the exact value of W_6. However in both cases of $W_6 = 1, 4$ we have $\tilde{W}_1 = \neg W_1$, and by flipping $\tilde{W}_1$ the first node can recover its intended bit.

In order to prove that if we remove the edges connected to node 6, at least 35 symbols are needed, suppose that there exists a coding scheme which requires at most 34 symbols. First notice that the sets $\{f(w_1,w_2,\cdots,w_6) : w_1,w_2,\cdots,w_5 \in \{0,1\}\}$ for different $w_6 \in \{0,1,2,3,4\}$ are mutually disjoint (Otherwise node 6 cannot decode its message using its side information and public message). According to Pigeonhole Principle, we can conclude that there exists $w_6 \in \{0,1,2,3,4\}$ so that if $W_6 = w_6$, public message gets at most 6 distinct different values when we vary the $w_1,w_2,\cdots,w_5$, i.e., the cardinality of the set $\{f(w_1,w_2,\cdots,w_6) : w_1,w_2,\cdots,w_5 \in \{0,1\}\}$ is at most 6 (If not, we would have 5 sets of the size at least 7, and it contradicts with our assumption that the public message has 34 different symbols). For this value of w_6, consider the following function over five variables $w_1,w_2,\cdots,w_5$: $\tilde{f}(w_1,w_2,\cdots,w_5) = f(w_1,w_2,\cdots,w_5,w_6)$. Since W_6 was independent of $(W_1,\cdots,W_5)$ and we have zero probability of error, $\tilde{f}$ is a valid encoding function for a cycle of length 5. This contradicts Lemma 3 below. ■

Lemma 3: The bidirectional cycle of length 5 with $\mathcal{W}_i' = \{0,1\}$ needs a public message of alphabet size 7 to achieve a zero probability of error for the one-shot problem.

Proof: We prove this lemma by contradiction. Assume otherwise that there exists a coding scheme that uses a public message with 6 possibilities. As there are 32 combinations of the messages of the nodes, from the Pigeonhole Principle, we conclude that the encoding function maps at least 6 combinations of the messages to one symbol, i.e., there are six sequences of $(w_{1i},w_{2i},\cdots,w_{5i}) \in \{0,1\}^5$, $i = 1,2,\cdots,6$, whose $f(w_{1i},w_{2i},\cdots,w_{5i})$ are equal, i.e., their corresponding public message is the same. Thus, the nodes should be able to recover their own messages using their side information. In other words, for instance for node 1, if $w_{1i} \neq w_{1i'}$ for some i and i', then we should have $(w_{2i},w_{5i}) \neq (w_{2i'},w_{5i'})$. Thus the six sequences should be distinguishable, where we call two sequences $(w_1,w_2,\cdots,w_5)$ and $(w_1',w_2',\cdots,w_5')$ distinguishable if for each $i \in \{1,2,\cdots,5\}$ either $w_i = w_i'$ or the $w_j \neq w_j'$ for some $j : (i,j) \in \mathcal{E}$.

Given a sequence $(w_1,w_2,\cdots,w_5)$, consider the graph induced on the set of vertices $\{j : w_j = 1\}$. We call the sequence $(w_1,w_2,\cdots,w_5)$ "good" if the induced graph does not contain a vertex of degree zero (i.e., is connected). For instance, in a cycle of size 5 if we take $(w_1,w_2,\cdots,w_5) = (1,1,1,0,0)$, the induced graph would be on nodes 1, 2, 3 which is connected. However $(w_1,w_2,\cdots,w_5) = (1,1,0,1,0)$ corresponds to the induced graph on nodes 1, 2, 4 which is not connected since node 4 is not connected to nodes 1 and 2. It is easy to verify that $(w_1,w_2,\cdots,w_5)$ and $(w_1',w_2',\cdots,w_5')$ are distinguishable if and only if their bitwise XOR is good. For instance (1,1,0,1,0) and (0,0,0,0,0) are not distinguishable (by node 4) since their XOR, (1,1,0,1,0) is not good.

Now, we know that the XOR of any two of $(w_{1i},w_{2i},\cdots,w_{5i})$, $1 \leq i \leq 6$ is good. We show that this cannot happen. Without loss of generality, we can assume that one of the six sequences is the all zero sequence. Thus, we should look for 5 sequences that are individually good, and their pairwise bitwise XOR is also good. Non-existence of such sequences is verified by a computer program. □

D. Proof of Theorem 4

Suppose that $\mathsf{G} = (\mathcal{V},\mathcal{E})$ is a bidirectional graph where $\mathcal{V} = \{1,2,\cdots,m\}$. To show that G is critical, we need to show that there exists a rate vector $\bar{\mathbf{r}} = (r_1,r_2,\cdots,r_m)$ for every $e = (u,v) \in \mathcal{E}$ that is achievable in G, but not achievable in $\mathsf{G}-e$. We define $\bar{\mathbf{r}}$ in the following manner:

$$r_i = \begin{cases} 1 & \text{if } i = u \text{ or } i = v \\ 0 & \text{otherwise.} \end{cases}$$

To show that $\bar{\mathbf{r}}$ is achievable in G, suppose that W_i is the message of node i, and $W_i \in \mathcal{W}_i'$ where:

$$\begin{cases} \mathcal{W}_i' = \{0,1\} & \text{if } i = u \text{ or } i = v \\ \mathcal{W}_i' = \{0\} & \text{otherwise.} \end{cases}$$

Now, if we send $W_u \oplus W_v$ as public message, then u and v can decode their message, because they have the message of each other as side information and the sum of their message. As $\mathcal{W}_i'$ has only one element for $i \neq u,v$, the other vertices can trivially decode their message. Therefore, $\bar{\mathbf{r}}$ is supported by G.

Additionally, since the set $\{u,v\}$ in $\mathsf{G}-e$ has no directed cycle, Lemma 2 implies that for every $\bar{\mathbf{r}'} = (r_1',r_2',\cdots,r_m')$ supported by $\mathsf{G}-e$, $r_u' + r_v' \leq 1$. Thus, $\bar{\mathbf{r}}$ cannot be supported by $\mathsf{G}-e$.

Next we show that a cycle of size four with vertices {1, 2, 3, 4} and edges {(1, 2), (2, 1), (2, 3), (3, 2), (3, 4), (4, 3), (4, 1), (1, 4)}) is not symmetric rate critical. If this graph supports the rate vector $\bar{\mathbf{r}} = (r,r,r,r)$, as the set {1, 3} has no directed cycle, Lemma 2 gives that $r \leq \frac{1}{2}$. In addition, consider the subgraph H of the cycle with edges {(1, 2), (2, 1), (3, 4), (4, 3)}. If we send two bits $(W_1 \oplus W_2, W_3 \oplus W_4)$ as public message, then all nodes can decode their message. Hence, the rate $(\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2})$ is

achievable in H. Now, since removing edges (2, 3), (3, 2), (4, 1), (1, 4) do not change the symmetric capacity region of cycle of size four, it is not symmetric rate critical. ■

E. Proof of Theorem 5

1) Criticality of $\mathsf{G}\cup\mathsf{H}$*:* In order to show the criticality of $\mathsf{G}\cup\mathsf{H}$ we need to show that by eliminating every edge like e from $\mathsf{G}\cup\mathsf{H}$, the capacity region of the index coding problem related to $\mathsf{G}\cup\mathsf{H}$ shrinks strictly. Without loss of generality assume that $e\in\mathcal{E}_{\mathsf{G}}$. As G is a critical graph, there exists a rate vector like $\bar{\mathbf{r}}$ that supports G, but not $\mathsf{G}'(\mathcal{V}_{\mathsf{G}},\mathcal{E}_{\mathsf{G}}-\{e\})$. Now, consider a rate vector for the index coding problem introduced by $\mathsf{G}\cup\mathsf{H}$ in which the rates of nodes in H are all zero and rates of the nodes in G equals $\bar{\mathbf{r}}$. This rate vector is evidently admissible for $\mathsf{G}\cup\mathsf{H}$, but not for $\mathsf{G}'\cup\mathsf{H}$ (which is $\mathsf{G}\cup\mathsf{H}$ after elimination of e).

2) Symmetric Criticality of $\mathsf{G}\cup\mathsf{H}$ *in Asymptotic Scenarios:* Showing that the maximal symmetric rate also reduces after we remove an edge from $\mathsf{G}\cup\mathsf{H}$ is more challenging. Let r_1 and r_2 be the maximal symmetric rate for G and H respectively. It is clear that concatenation of these two coding functions with proportion of $\frac{r_2}{r_1+r_2}$ and $\frac{r_1}{r_1+r_2}$ for G and H respectively, results in a coding function for $\mathsf{G}\cup\mathsf{H}$ with the symmetric rate of $r=(r_1r_2)/(r_1+r_2)$.

We claim that this symmetric rate would not be achievable if any edge like e is removed from $\mathsf{G}\cup\mathsf{H}$. This will prove the symmetric criticality of $\mathsf{G}\cup\mathsf{H}$. We are going to prove this claim by contradiction. Without loss of generality, assume that e is an edge of G. We refer to the graph obtained by G after elimination of e as G'. Suppose that there exists a coding function like f for $\mathsf{G}'\cup\mathsf{H}$ with the symmetric rate of r. From Theorem 2 then, there exist some $\alpha\in[0,1]$ and symmetric rates r'_1 and r'_2 for G' and H such that $r=\alpha r'_1=\bar{\alpha}r'_2$. A simple calculation and using the fact that $r=(r_1r_2)/(r_1+r_2)$ gives us $(1/r'_1)+(1/r'_2)=(1/r_1)+(1/r_2)$. However this cannot happen since $r'_1<r_1$ and $r'_2\leq r_2$, by the symmetric criticality of G and by the definitions of r_1 and r_2. This completes the proof by contradiction.

3) Symmetric Criticality of $\mathsf{G}\cup\mathsf{H}$ *in One-Shot Linear Scenario:* Consider the symmetric one-shot linear index coding problems defined over G, H, and $\mathsf{G}\cup\mathsf{H}$. Assume that the alphabet of each node in each of these problems is $\mathbb{F}^l$ for some finite field $\mathbb{F}$. Let n_1 be the minimum possible positive integer number such that there exists a valid linear coding function with the output size of n_1 symbols over $\mathbb{F}$. Define n_2 for H in the same manner. It is clear that there exists an encoding function for the problem related to $\mathsf{G}\cup\mathsf{H}$ that uses a public message of n_1+n_2 symbols by concatenation of the encoding functions that use n_1 symbols for G and n_2 symbols for H. Hence, the symmetric rate of $\frac{l}{n_1+n_2}$ is achievable for the index coding problem introduced by $\mathsf{G}\cup\mathsf{H}$.

We are going to show that if G and H are both symmetric critical, then $\mathsf{G}\cup\mathsf{H}$ is symmetric critical too. To prove the symmetric criticality of $\mathsf{G}\cup\mathsf{H}$, we will prove that any valid coding function for $\mathsf{G}\cup\mathsf{H}$ needs at least n_1+n_2+1 public symbols after removal of any edge like e from $\mathsf{G}\cup\mathsf{H}$. Without loss of generality, we assume that e is removed from the G component of $\mathsf{G}\cup\mathsf{H}$. We refer to the graph obtained by G after the removal of e as G'. Then, the graph obtained from $\mathsf{G}\cup\mathsf{H}$ after the removal of e would be $\mathsf{G}'\cup\mathsf{H}$. Let f be a valid encoding function for $\mathsf{G}'\cup\mathsf{H}$, we have shown in the proof of part (a) of Theorem 2 that there exist two valid encoding functions f' and f'' for G' and H so that the concatenation of f' and f'' has the same output size to f. As f'' is a valid encoding function for H, its output size is at least n_2. In addition, because of the criticality of G, we know that every valid encoding function for G', including f', needs at least n_1+1 symbols. Accordingly, the concatenation of f' and f'' has an output size of at least n_1+n_2+1. Consequently, the output size of the f is at least n_1+n_2+1. This means that $\mathsf{G}\cup\mathsf{H}$ is symmetric critical because it cannot support the symmetric rate of $l/(n_1+n_2)$ after removal of any of its edges.

F. Proof of Theorem 6

1) Proof of Part (a): In the first step of the proof, we will show that the new graph G' supports the symmetric rate of $\bar{\mathbf{r}}=\left(\frac{1}{m-1},\frac{1}{m-1},\cdots,\frac{1}{m-1}\right)$. In the next step, we will show that this rate will not be achievable if any edge is eliminated. These two steps will clearly prove this theorem.

(I) Achievability: Let W_i denote the message of node i, and $W_i\in\mathcal{W}_i=\{0,1\}$. Consider the encoding function $f(W_1,\cdots,W_m)=(f_1,\cdots,f_m)=(W_1\oplus W_2,\cdots,W_{j-1}\oplus W_j,W_j\oplus W_{j+1}\oplus W_{m+1},W_{j+1}\oplus W_{j+2},\cdots,W_{k-1}\oplus W_k,W_k\oplus W_{k+1}\oplus W_{m+1},W_{k+1}\oplus W_{k+2},\cdots,W_{m-1}\oplus W_m,W_m\oplus W_1)$.

For $1\leq l\leq m$, f_l is the sum of the side information of node l and W_l; therefore node l (for $1\leq l\leq m$) can decode its message. Node $m+1$ can also consider:

$$\bigoplus_{l=1}^{i-1}f_l=\bigoplus_{l=1,l\neq j}^{i-1}(W_l\oplus W_{l+1})\oplus(W_j\oplus W_{j+1}\oplus W_{m+1})$$
$$=W_1\oplus W_i\oplus W_{m+1}.$$

Since node $m+1$ has W_1 and W_i as side information, it can decode its message with the help public message.

As $f(W_1,\cdots,W_m)\in\{0,1\}^m$, it shows the achievability of rate $(\frac{1}{m},\frac{1}{m},\cdots,\frac{1}{m})$. For $t\neq j,k$, if omit f_t from f and construct a new encoding function f', then due to the fact that:

$$\bigoplus_{1\leq l\leq m,l\neq t}f_l=f_t,$$

each node can obtain f from f'. So, they can decode their message by the help of f' and their side information. It shows the achievability of the rate $\bar{\mathbf{r}}=\left(\frac{1}{m-1},\frac{1}{m-1},\cdots,\frac{1}{m-1}\right)$.

(II) Unachievability after Edge Removal: To show that G' is critical, it suffices to prove that after removing any edge of $\mathcal{E}'$, we will need at least m bits of public message. Lemma 2 implies that if there exists a subset of length m in a graph which does not contain any cycle, the rate $\bar{\mathbf{r}}$ would not be achievable in the graph (otherwise the sum of the rates would be $\frac{m}{m-1}$ which is greater that 1). Thus, it suffices to show that for every $e\in\mathcal{E}'$, there exists a subset of $\mathcal{V}'$, say A, of length at least m such that the induced subgraph of A in $\mathsf{G}'-e$ has no directed cycle.

First, suppose that $e\in\mathcal{E}$. Then we can choose $A=\mathcal{V}$. As A contains exactly m vertices and the induced graph is a directed

path which contains no cycle then these edges are critical. For $e=(m+1,1)$, A can be chosen as $\mathcal{V}\cup\{m+1\}\setminus\{i\}$. Same argument can be made for $e=(m+1,i)$. For $e=(j,m+1)$, A can be chosen as $\mathcal{V}\cup\{m+1\}\setminus\{k\}$. Same argument can be made for $e=(k,m+1)$.

2) Proof of Part (b): We use the same approach from part (a) to show that the rate $\bar{\mathbf{r}}=\left(\frac{1}{m-1},\frac{1}{m-1},\cdots,\frac{1}{m-1}\right)$ is achievable in G'', and by removing every edge the rate would not be achievable.

(I) Achievability: Suppose that the message of node $u_i\in\mathcal{V}''$ is $W_{u,i}$. Then, define: $W_u=\bigoplus_{i=1}^{n_u}W_{u,i}$. Similar to part (a), consider the encoding function $f=(f_1,f_2,\cdots,f_m)=(W_1\oplus W_2,\cdots,W_{j-1}\oplus W_j,W_j\oplus W_{j+1}\oplus W_{m+1},W_{j+1}\oplus W_{j+2},\cdots,W_{k-1}\oplus W_k,W_k\oplus W_{k+1}\oplus W_{m+1},W_{k+1}\oplus W_{k+2},\cdots,W_{m-1}\oplus W_m)$. Again, for $1\le l\le m$ and $t\in\{1,2,\cdots,n_l\}$, the l^{th} element of f, f_l, is the sum of the side information and $W_{l,t}$. So, these nodes can decode their message. For $t\in\{1,2,\cdots,n_{m+1}\}$, node $(m+1)_t$ can consider:

$$\begin{aligned}\bigoplus_{l=1}^{i-1}f_l&=\bigoplus_{l=1,l\neq j}^{i-1}(W_l\oplus W_{l+1})\oplus(W_j\oplus W_{j+1}\oplus W_{m+1})\\&=W_1\oplus W_i\oplus W_{m+1}\\&=\left(\bigoplus_{l=1}^{n_1}W_{1,l}\right)\oplus\left(\bigoplus_{l=1}^{n_i}W_{i,l}\right)\oplus\left(\bigoplus_{l=1}^{n_{m+1}}W_{m+1,l}\right).\end{aligned}$$

By definition of G'', node $(m+1)_t$ knows $W_{1,1},\cdots,W_{1,n_1}$, $W_{i,1},\cdots,W_{i,n_i},W_{m+1,1},\cdots,W_{m+1,t-1},W_{m+1,t+1},\cdots,W_{m+1,n_{m+1}}$ as side information. Therefore, with the help of public message and its side information $(m+1)_t$ can decode its message. As the range of f is subset of $\{0,1\}^m$, the rate $(\frac{1}{m},\frac{1}{m},\cdots,\frac{1}{m})$ is achievable. Again, for $t\neq j,k$, we have that $f_t=\bigoplus_{1\le l\le m,l\neq t}f_l$. So, by omitting f_t from f, we can achieve the rate $\bar{\mathbf{r}}=(\frac{1}{m-1},\frac{1}{m-1},\cdots,\frac{1}{m-1})$.

(II) Unachievability after Edge Removal: Now, we want to show that by elimination of any edge in $\mathcal{E}''$, $\bar{\mathbf{r}}$ will not be achievable anymore. As discussed in part (a), it suffices to show that after removing any edge in $\mathcal{E}''$, there will be $A\subset\mathcal{V}''$ with at least m vertices which does not contain any directed cycle. As we have two different types of edges in G'', we analyze the impact of edge removal on the capacity region in two different cases.

Case 1) $e=(u_s,v_t)$ where $u\neq v$: By definition of G'', we have $e'=(u,v)\in\mathcal{E}'$. In part (a), we proved that there exists $A'\subset\mathcal{V}''$ of size m which does not contain any cycle in $\mathsf{G}'-e'$. Now, choose $A=\{l_1:l\in A',l\neq u,v\}\cup\{u_s,v_t\}$. If x_y and $x'_{y'}\in A$, then x, $x'\in A'$ and $x\neq x'$. Thus, x_y has edge to $x'_{y'}$ in G'' if and only if x has edge to x' in G' and because A' has no cycle in G' then A has no cycle in G''.

Case 2) $e=(u_s,u_t)$: For $1\le u\le m$, choose $A=\{u_s,u_t\}\cup\{((u+l)\bmod m)_1,1\le l\le m-1\}$, and for $u=m+1$. Choose $A=\{l_1:1\le l\le m,l\le j,k\}\cup\{u_s,u_t\}$. It can be checked that these two sets contain no cycle.

REFERENCES

[1] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge, U.K.: Cambridge Univ. Press, 2011.

[2] Y. Birk and T. Kol, "Informed-Source Coding-on-Demand (ISCOD) over broadcast channels," in *Proc. 17th Annu. IEEE INFOCOM*, San Francisco, CA, USA, Mar. 1998, pp. 1257–1264.

[3] S. H. Dau, V. Skachek, and Y. M. Chee, "Secure index coding with side information," arXiv:1011.5566.

[4] A. Blasiak, R. Kleinberg, and E. Lubetzky, "Lexicographic products and the power of non-linear network coding," in *Proc. 52nd Annu. IEEE Symp. FOCS*, 2011, pp. 609–618.

[5] F. Arbabjolfaei, B. Bandemer, Y. Kim, E. Sasoglu, and L. Wang, "On the capacity region for index coding," arXiv:1302.1601.

[6] H. Sun and S. A. Jafar, "Index coding capacity: How far can one go with only Shannon inequalities?" arXiv:1303.7000.

[7] K. Shanmugam, A. G. Dimakis, and M. Langberg, "Local graph coloring and index coding," in *Proc. IEEE Int. Symp. IT*, 2013, pp. 1152–1156.

[8] 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. IT*, 2008, pp. 1823–1827.

[9] M. Effros, S. Rouayheb, and M. Langberg, "An equivalence between network coding and index coding," arXiv:1211.6660, 2012.

[10] H. Maleki, V. Cadambe, and S. Jafar, "Index coding: An interference alignment perspective," in *Proc. IEEE Int. Symp. IT*, 2012, pp. 2236–2240.

[11] E. Lubetzky and U. Stav, "Non-linear index coding outperforming the linear optimum," in *Proc. 48th Annu. IEEE Symp. FOCS*, 2007, pp. 161–167.

[12] Z. Bar-Yossef, Y. Birk, T. S. Jayram, and T. Kol, "Index coding with side information," in *Proc. 47th Annu. IEEE Symp. FOCS*, 2006, pp. 197–206.

[13] M. Langberg and M. Effros, "Network coding: Is zero error always possible?" in *Proc. 49th Annu. Allerton Conf. Commun., Control Comput.*, 2011, pp. 1478–1485.

[14] D. B. West, *Introduction to Graph Theory*. Upper Saddle River, NJ, USA: Prentice-Hall, 2001.

[15] S. A. Jafar, "Topological interference management through index coding," arXiv:1301.3106 2013.

[16] M. J. Neely, A. Tehrani, and Z. Zhang, "Dynamic index coding for wireless broadcast networks," in *Proc. IEEE INFOCOM*, 2012, pp. 316–324.

[17] A. Tehrani, A. G. Dimakis, and M. Neely, "Bipartite index coding," in *Proc. IEEE Int. Symp. Inf. Theory*, Cambridge, MA, USA, Jul. 2012, pp. 2246–2250.

[18] K. Shanmugam, A. Dimakis, and M. Langberg, "Graph theory versus minimum rank for index coding," arXiv:1402.3898.

Mehrdad Tahmasbi received the B.Sc. degree in electrical engineering and pure mathematics from Sharif University of Technology, Tehran, Iran. He is currently with the Department of Electrical Engineering, Sharif University of Technology. He was a recipient of a silver medal at the 22nd International Olympiad in Informatics (IOI 2010).

Amirbehshad Shahrasbi is currently working toward the double major B.Sc. degree in electrical engineering and computer science with Sharif University of Technology, Tehran, Iran. He was a recipient of a gold medal at the 18th Iranian National Olympiad in Informatics in 2009.

Amin Gohari received the B.Sc. degree from Sharif University of Technology, Tehran, Iran, in 2004 and the Ph.D. degree in electrical engineering from the University of California, Berkeley (UC Berkeley), CA, USA, in 2010. He is currently an Assistant Professor at Sharif University of Technology. He was a recipient of the 2010 Eli Jury Award from UC Berkeley, Department of Electrical Engineering, for "outstanding achievement in the area of communication networks," and the 2009–2010 Bernard Friedman Memorial Prize in Applied Mathematics from UC Berkeley, Department of Mathematics, for "demonstrated ability to do research in applied mathematics." He was also a recipient of the Gold Medal from the 41st International Mathematical Olympiad (IMO 2000) and the First Prize from the 9th International Mathematical Competition for University Students (IMC 2002).