An Equivalence Between Secure Network and Index Coding

Lawrence Ong[†], Badri N. Vellambi[‡], Jörg Kliewer[‡], and Phee Lep Yeoh[§]
[†]The University of Newcastle, Australia; [‡]New Jersey Institute of Technology, USA; [§]University of Sydney, Australia

***Abstract*—We extend the equivalence between network coding and index coding by Effros, El Rouayheb, and Langberg to the secure communication setting in the presence of an eavesdropper. Specifically, we show that the most general versions of secure network-coding setup by Chan and Grant and the secure index-coding setup by Dau, Skachek, and Chee, which also include the randomised encoding setting, are equivalent.**

I. Introduction

Recently, equivalence results in information theory and network coding have been of significant interest in the community. Such reduction results uniquely map one communication problem to another equivalent problem that is potentially easier to study than the original problem. Some of the equivalence results already established include those between instances of multiple-unicast network coding and those of (a) multiple-multicast network coding [1], (b) secure network coding [2], and (c) index coding [3].

In particular, the latter result addresses the equivalence between network coding [4] and index coding [5] in the non-secure setting. Non-secure network coding and index coding were shown to be equivalent [3, 6] in the sense that a network-coding instance can be mapped to an equivalent index-coding instance, for which a code for one instance can be translated to the other, and vice versa. Similarly, an index-coding instance can be mapped to an equivalent network-coding instance, with a suitable code translation.

While strongly-secure and weakly-secure network coding [7, 9] as well as strongly-secure and weakly-secure index coding [10] have been studied in the literature so far, an equivalence between these coding approaches for the *secure* setting has not been addressed to the best of our knowledge.

Note that the equivalence between non-secure network and index coding does not trivially apply to the secure setting. In particular, we pointed out [11] that equating the eavesdropper settings in secure network coding and secure index coding is not straightforward. We also showed that the equivalence breaks down in the randomised encoding setting—noting that randomised encoding is inevitable in some secure network-coding instances [7].

In this paper, we extend the equivalence between network and index coding by Effros, El Rouayheb, and Langberg [3] to strongly-secure and weakly-secure settings, by proposing a suitable mapping for the eavesdroppers. For the mapping from secure network coding to secure index coding, we also introduce the concept of an augmented secure network-coding instance to capture the randomness in the encoding. With this, we incidentally establish an equivalence between non-secure network and index coding with randomised encoding.

II. Problem Definition and Notation

For any positive integer n, let $[n] \triangleq \{1, \dots, n\}$. For a set $\mathcal{I} = \{i_1, \dots, i_{|\mathcal{I}|}\}$, let $\boldsymbol{X}_{\mathcal{I}} \triangleq [X_{i_1} \cdots X_{i_{|\mathcal{I}|}}]$ with an arbitrary but fixed order. The tail and head of an edge $(u, v) \in E$ in a directed graph $G = (V, E)$ refer to vertices u, and v, respectively, i.e., $u = \mathsf{tail}(e)$ and $v = \mathsf{head}(e)$. For a node $v \in V$, we let $\mathsf{in}(v)$ to be the set of all edges with $\mathsf{head}(e) = v$; similarly, $\mathsf{out}(v)$ denotes the set of all edges with $\mathsf{tail}(e) = v$.

A. Secure network coding

1) Network-coding instances: We follow Chan and Grant's secure network-coding definition [12]. It includes Bhattad and Narayanan's weakly-secure network-coding definition [9] and Cai and Yeung's strongly-secure network coding definition [7] as special cases. A secure network-coding instance, denoted by $\mathcal{I} = (G, M, W)$, is defined as follows:

- $G = (\mathcal{V}, \mathcal{E})$ is an acyclic graph with vertex set $\mathcal{V}$ and edge set $\mathcal{E}$. Each edge $e \in \mathcal{E}$ has a *capacity* given by c_e.
- $M = (\mathcal{S}, O, \mathcal{D})$ is the connection requirement. The set $\mathcal{S}$ is the collection of source-message indices, where the source messages $\{X_s : s \in \mathcal{S}\}$ are mutually independent and are each distributed on $[2^{R_s n}]$, for some positive integer n that can be chosen to suit the design of network codes. Here, R_s denotes the rate of the message X_s, $s \in \mathcal{S}$. The source-location mapping $O : \mathcal{S} \to \mathcal{V}$ specifies the originating node $O(s)$ for the source message X_s. The destination-location mapping $\mathcal{D} : \mathcal{S} \to 2^{\mathcal{V}}$ specifies the nodes $\mathcal{D}(s)$ that require the message X_s.
- $W = ((\mathcal{A}_r, \mathcal{B}_r) : r \in \mathcal{R})$ defines the eavesdropping pattern for $|\mathcal{R}|$ eavesdroppers. Each eavesdropper $r \in \mathcal{R}$ observes the set of links $\mathcal{B}_r \subseteq \mathcal{E}$ and tries to reconstruct a subset of source messages indexed by $\mathcal{A}_r \subseteq \mathcal{S}$, i.e., $\boldsymbol{X}_{\mathcal{A}_r}$.

We assume that vertices with no incoming links are originating nodes for some source messages, and vertices with no outgoing links are destinations for some source messages.

2) Deterministic network codes: Given (G, M), a network code $(\mathcal{F}, \mathcal{G})$ consists of a collection of encoding functions for the edges $\mathcal{F} = \{f_e : e \in \mathcal{E}\}$, and decoding functions for the vertices $\mathcal{G} = \{g_u : u \in \mathcal{V}\}$ satisfying the following:

This work is supported by ARC grants FT140100219, DE140100420, and DP150100903, and US NSF grants CNS-1526547 and CCF-1439465.

978-1-5090-2482-7/16/$31.00 ©2016 IEEE

The local encoding function f_e for edge e takes in random variables associated with $\mathsf{in}(\mathsf{tail}(e))$ and source messages originating at node $\mathsf{tail}(e)$, and outputs a random variable associated with link e, denoted by $X_e \in [2^{c_e n}]$.

Given that G is acyclic, each edge message X_e can be written as a function of source messages originating from its predecessors, denoted by $\bar{f}_e$. This is known as the global encoding function, and it can be recursively calculated (following the topology of the graph) using (i) $\bar{f}_e = f_e$ if $\mathsf{tail}(e)$ has no incoming links, and (ii) $\bar{f}_e = f_e(\bar{f}_{e_1}, \bar{f}_{e_2}, \ldots, \bar{f}_{e_n})$, where $\{e_1, e_2, \ldots, e_n\} = \mathsf{in}(\mathsf{tail}(e))$.

The decoding function g_u for a node $u \in \mathcal{V}$ takes in random variables associated with links $\mathsf{in}(u)$ and source messages originating at node u, and outputs $\boldsymbol{X}_{\{s\in\mathcal{S}:u\in\mathcal{D}(s)\}}$. In this paper, we only consider *zero-error* decoding.

3) Randomised network codes: A network code is said to be randomised if there exists an edge function f_e that is not a deterministic function of the random variables associated with $\mathsf{in}(\mathsf{tail}(e))$ and source messages originating at node $\mathsf{tail}(e)$.

Any randomised function can be implemented by generating an independent random variable Z_u at each node $u \in \mathcal{V}$, where Z_u takes values in an alphabet with size $\prod_{e\in\mathsf{out}(u)} 2^{c_e n}$, where $\mathsf{out}(u)$ is defined as the set of all outgoing edges from node u. These independent random variables are often referred to as *random keys*.

A randomised network code $(\mathcal{F}', \mathcal{G})$ is similar to a deterministic network code $(\mathcal{F}, \mathcal{G})$, except that each edge encoding function f'_e is a function of (i) random variables associated with $\mathsf{in}(\mathsf{tail}(e))$, (ii) source messages originating at node $\mathsf{tail}(e)$, and (iii) the random key $Z_{\mathsf{tail}(e)}$.

4) Secure network codes: A deterministic or randomised network code $(\mathcal{F}, \mathcal{G})$ for (G, M) is said to be secure against an eavesdropping pattern W if each eavesdropper r gains no information about $\boldsymbol{X}_{\mathcal{A}_r}$ that it attempts to reconstruct after observing $\boldsymbol{X}_{\mathcal{B}_r}$ on the links it has access to, i.e.,

$$H(\boldsymbol{X}_{\mathcal{A}_r} | \boldsymbol{X}_{\mathcal{B}_r}) = H(\boldsymbol{X}_{\mathcal{A}_r}), \quad r \in \mathcal{R}. \tag{1}$$

In other words, $(\mathcal{F}, \mathcal{G})$ is a secure network code for the secure network-coding instance I.

5) Secure network-coding rates: The secure network-coding instance is said to be $(\boldsymbol{R}_{\mathcal{S}}, n)$-feasible if and only if there exists at least one secure network code with the associated source-message rates and block size n.

B. Secure index coding

1) Secure index-coding instances: We follow Dau, Skachek, and Chee's secure index-coding definition [10]. A secure index-coding instance, denoted by $\hat{I} = (\hat{\mathcal{S}}, \hat{\mathcal{T}}, \{\hat{\mathcal{W}}_{\hat t}\}, \{\hat{\mathcal{H}}_{\hat t}\}, \hat{W})$, is defined as follows:

- $\hat{\mathcal{S}} = [k]$ is the set of indices of k source messages available at a sender. The messages $\{\hat{X}_{\hat s} : \hat s \in \hat{\mathcal{S}}\}$ are mutually independent and for $\hat s \in \hat{\mathcal{S}}$, $\hat{X}_{\hat s}$ is distributed on $[2^{\hat{R}_{\hat s} n}]$, for some non-negative message rate $\hat{R}_{\hat s}$ and positive integer n that can be chosen to suit the design of index codes.
- $\hat{\mathcal{T}} = [\ell]$ is the collection of ℓ receiver indices.
- $\hat{\mathcal{W}}_{\hat t}$ is the set of the indices of the messages required by receiver $\hat t \in \hat{\mathcal{T}}$.
- $\hat{\mathcal{H}}_{\hat t}$ is the set of indices of the messages known a priori to receiver $\hat t \in \hat{\mathcal{T}}$.
- $\hat{W} = ((\hat{\mathcal{A}}_{\hat r}, \hat{\mathcal{B}}_{\hat r}) : \hat r \in \hat{\mathcal{R}})$ is the eavesdropping pattern. Each eavesdropper $\hat r \in \hat{\mathcal{R}}$ has access to the codeword broadcast by the sender and a subset of the messages $\boldsymbol{X}_{\hat{\mathcal{B}}_{\hat r}}$, and tries to reconstruct $\boldsymbol{X}_{\hat{\mathcal{A}}_{\hat r}}$, where $\hat{\mathcal{A}}_{\hat r}, \hat{\mathcal{B}}_{\hat r} \subseteq \hat{\mathcal{S}}$.

2) Deterministic index codes: A deterministic index code $(\hat{\mathcal{F}}, \hat{\mathcal{G}}) = (\hat f, \{\hat g_{\hat t} : \hat t \in \hat{\mathcal{T}}\})$ consists of an encoding function by the sender $\hat{\mathcal{F}} = \hat f$ which takes in the random variables $\hat{\boldsymbol{X}}_{\hat{\mathcal{S}}}$ and outputs a random variable $\hat{X}_{\mathrm{b}} \in [2^{\hat c_{\mathrm{b}} n}]$, where $\hat c_{\mathrm{b}}$ is the broadcast rate, and n is the block size of the code. It also consists of a decoding function $\hat g_{\hat t}$ for receiver $\hat t$, which takes in the sender's codeword $\hat X_{\mathrm{b}}$ and its prior messages $\hat{\boldsymbol{X}}_{\hat{\mathcal{H}}_{\hat t}}$ and outputs the messages $\hat{\boldsymbol{X}}_{\hat{\mathcal{W}}_{\hat t}}$ it requires. Similar to the network-coding setup, we only consider zero-error decoding.

3) Randomised index codes: A randomised index code $(\hat{\mathcal{F}}', \hat{\mathcal{G}})$ is defined similar to the deterministic index codes except that the sender's encoding function takes in an independent random key $\hat Z \in [2^{\hat r_{\mathrm{b}} n}]$ of some positive rate $\hat r_{\mathrm{b}}$ in addition to $\hat{\boldsymbol{X}}_{\hat{\mathcal{S}}}$. Unlike the model by Mojahedian, Aref, and Gohari [8], the randomness allowed in the encoding in our setting is known only to the sender, and is not shared between the sender and the receivers.

4) Secure index codes: A deterministic or randomised index code $(\hat{\mathcal{F}}, \hat{\mathcal{G}})$ is said to be secure against the eavesdropping pattern $\hat W$ if no eavesdropper $\hat r$ gains no information about the message set $\hat{\boldsymbol{X}}_{\hat{\mathcal{A}}_{\hat r}}$ it tries to reconstruct by observing the sender's codeword $\hat X_{\mathrm{b}}$ and its side information $\hat{\boldsymbol{X}}_{\hat{\mathcal{B}}_{\hat r}}$, i.e.,

$$H(\hat{\boldsymbol{X}}_{\hat{\mathcal{A}}_{\hat r}} | \hat X_{\mathrm{b}}, \hat{\boldsymbol{X}}_{\hat{\mathcal{B}}_{\hat r}}) = H(\hat{\boldsymbol{X}}_{\hat{\mathcal{A}}_{\hat r}}), \quad \hat r \in \hat{\mathcal{R}}. \tag{2}$$

Clearly, $\hat{\mathcal{A}}_{\hat r} \cap \hat{\mathcal{B}}_{\hat r} = \emptyset$. Specifically, we say that $(\hat{\mathcal{F}}, \hat{\mathcal{G}})$ is a secure index code for the secure index-coding instance $\hat I$.

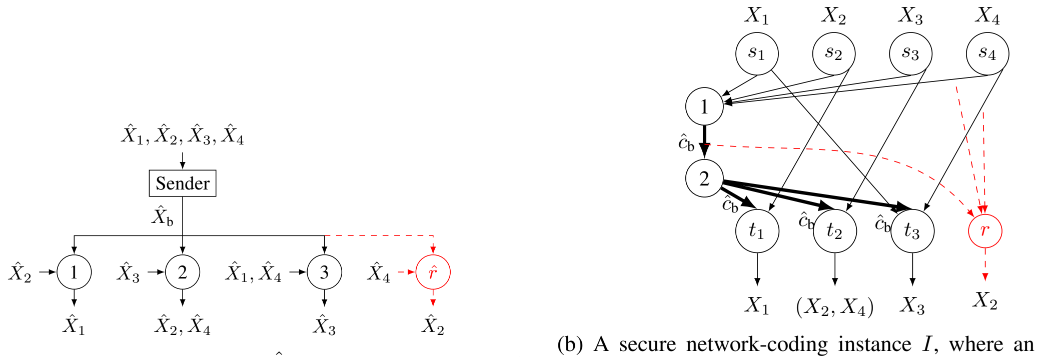
5) Secure index-coding rate: The secure index-coding instance is said to be $(\hat{\boldsymbol{R}}_{\hat{\mathcal{S}}}, \hat c_{\mathrm{b}}, n)$-feasible if and only if there exists at least one secure index code with the associated source-message rates $\hat{\boldsymbol{R}}_{\hat{\mathcal{S}}}$, broadcast rate $\hat c_{\mathrm{b}}$, and block size n.

III. Mapping from Secure Index Coding to Secure Network Coding

Given an instance $\hat I = (\hat{\mathcal{S}}, \hat{\mathcal{T}}, \{\hat{\mathcal{W}}_{\hat t}\}, \{\hat{\mathcal{H}}_{\hat t}\}, \hat W)$ of a secure index-coding problem, where $\hat{\mathcal{S}} = [k]$ and $\hat{\mathcal{T}} = [\ell]$, we construct an equivalent secure network-coding instance $\mathcal{I} = (G, M, W)$ using the following rule:

Index-to-network coding mapping:

- The graph $G = (\mathcal{V}, \mathcal{E})$ consists of $k + \ell + 2$ vertices labelled as $\mathcal{V} = \{s_1, s_2, \ldots, s_k, t_1, t_2, \ldots, t_\ell, 1, 2\}$. For each $i \in \hat{\mathcal{S}}$, vertex s_i has an outgoing link to vertex 1 and to each vertex in $\{t_j : i \in \hat{\mathcal{H}}_j\}$. Each of these links from vertex s_i are of capacity $2^{\hat R_i n}$. Vertex 1 has a link of capacity $2^{\hat c_{\mathrm{b}} n}$ to vertex 2 and to each vertex in $\{t_i : i \in \hat{\mathcal{T}}\}$.
- The connection requirement M consists of the following: $\mathcal{S} = \hat{\mathcal{S}}$, and $\boldsymbol{X}_{\mathcal{S}}$ has the same distribution as $\hat{\boldsymbol{X}}_{\hat{\mathcal{S}}}$, which implies $R_i = \hat R_i$. For each message X_i, $i \in \mathcal{S}$, the source



(a) A secure index-coding instance $\hat{I}$, where an eavesdropper $\hat{r}$ has access to the broadcast message $\hat{X}_\text{b}$, side information $\hat{X}_4$, and tries to reconstruct $\hat{X}_2$

(b) A secure network-coding instance I, where an eavesdropper r has access to link $(1, 2)$, all outgoing links from node s_4, and tries to reconstruct X_2. The capacity of all links given by thick arrows is $\hat{c}_\text{b}$

Fig. 1: A secure index-coding instance $\hat{I}$ and its corresponding secure network-coding instance I

locations are $O(i) = s_i$, i.e., the message X_i originates at vertex s_i, and is destined for $\mathcal{D}(i) = \{t_j : i \in \hat{\mathcal{W}}_j\}$.
- The eavesdropping pattern W is defined as $\mathcal{R} = \hat{\mathcal{R}}$, $\mathcal{B}_r = \{(1 \to 2), \{\mathsf{out}(s_i) : i \in \hat{\mathcal{B}}_r\}\}$, and $\mathcal{A}_r = \hat{\mathcal{A}}_r$.

Note that by construction, for each $i \in \hat{\mathcal{T}}$,

- $\hat{\mathcal{W}}_i = \{j \in \mathcal{S} : t_i \in \mathcal{D}(j)\}$.
- $\hat{\mathcal{H}}_i = \{j \in \mathcal{S} : (s_j \to t_i) \in \mathcal{E}\}$.
- The vertices in $\mathcal{V} \setminus \{t_1, \dots, t_\ell\}$ are not the destinations of any source message.

Figure 1 depicts an example of such a mapping. With the above conversion, we now state an equivalence between these two instances:

Theorem 1: Let $\hat{I}$ and $\hat{c}_\text{b}$ be a secure index-coding instance and a broadcast rate, respectively. Let I be the corresponding secure network-coding instance using the index-to-network coding mapping. For any $\hat{\boldsymbol{R}}_{\hat{\mathcal{S}}}$ and $\hat{c}_\text{b}$, the instance $\hat{I}$ is $(\hat{\boldsymbol{R}}_{\hat{\mathcal{S}}}, \hat{c}_\text{b}, n)$-feasible if and only if I is $(\boldsymbol{R}_\mathcal{S}, n)$-feasible.

Proof of Theorem 1:

$\hat{I}$ is $(\hat{\boldsymbol{R}}_{\hat{\mathcal{S}}}, \hat{c}_\text{b}, n)$-feasible $\Rightarrow$ I is $(\boldsymbol{R}_\mathcal{S}, n)$-feasible:

Let $(\hat{\mathcal{F}}, \hat{\mathcal{G}})$ be a secure index code that supports $(\hat{\boldsymbol{R}}_{\hat{\mathcal{S}}}, \hat{c}_\text{b}, n)$. Then,

$$H(\hat{\boldsymbol{X}}_{\hat{\mathcal{W}}_{\hat{t}}} | \hat{X}_\text{b}, \hat{\boldsymbol{X}}_{\hat{\mathcal{H}}_{\hat{t}}}) = 0, \qquad \text{for all } \hat{t} \in \hat{\mathcal{T}}, \tag{3}$$

$$H(\hat{\boldsymbol{X}}_{\hat{\mathcal{A}}_{\hat{r}}} | \hat{X}_\text{b}, \hat{\boldsymbol{X}}_{\hat{\mathcal{B}}_{\hat{r}}}) = H(\hat{\boldsymbol{X}}_{\hat{\mathcal{A}}_{\hat{r}}}), \qquad \text{for all } \hat{r} \in \hat{\mathcal{R}}, \tag{4}$$

where $\hat{X}_\text{b} = \hat{f}(\hat{\boldsymbol{X}}_{\hat{\mathcal{S}}}, \hat{Z})$, where $\hat{Z} \in [2^{\hat{c}_\text{b} n}]$ is a random key independent of the source messages (to account for randomised index coding), and

$$\hat{\boldsymbol{X}}_{\hat{\mathcal{W}}_{\hat{t}}} = \hat{g}_{\hat{t}}(\hat{f}(\hat{\boldsymbol{X}}_{\hat{\mathcal{S}}}, \hat{Z}), \hat{\boldsymbol{X}}_{\hat{\mathcal{H}}_{\hat{t}}}), \qquad \text{for all } \hat{t} \in \hat{\mathcal{T}}. \tag{5}$$

Now, we construct a secure network code as follows:

- Set $f_e = X_{\mathsf{tail}(e)}$ for each outgoing edge e from each vertex in $\{s_i : i \in \hat{\mathcal{S}}\}$. This is possible since vertex s_i is the originating vertex for the message X_i.
- Set $f_{(1,2)} = f_e = \hat{f}(\boldsymbol{X}_\mathcal{S}, Z_1) \in [2^{\hat{c}_\text{b} n}]$ for all $e \in \mathsf{out}(2)$, where Z_1 is independent of all the source messages $\boldsymbol{X}_\mathcal{S}$ and has the same distribution as $\hat{Z}$.
- Set $g_{t_i} = \hat{g}_i$ for all $i \in \hat{\mathcal{T}}$, and $g_u = 0$ for all other vertices.

For each vertex t_i, $i \in \hat{\mathcal{T}}$, all incoming edges $\mathsf{in}(t_i)$ originate from vertex 2 and vertices $\{s_j : j \in \hat{\mathcal{H}}_i\}$. The message on edge $(2, t_i)$ is $\hat{f}(\boldsymbol{X}_\mathcal{S}, Z_1)$, and that on edge (s_j, t_i) is $f_{(s_j, t_i)} = X_{s_j}$. This means vertex t_i, $i \in \hat{\mathcal{T}}$, can decode

$$g_{t_i}(\hat{f}(\boldsymbol{X}_\mathcal{S}, Z_1), \boldsymbol{X}_{\hat{\mathcal{H}}_i}) = \hat{g}_i(\hat{f}(\boldsymbol{X}_\mathcal{S}, Z_1), \boldsymbol{X}_{\hat{\mathcal{H}}_i}) \tag{6a}$$

$$= \boldsymbol{X}_{\hat{\mathcal{W}}_i} = \boldsymbol{X}_{\{j \in \mathcal{S} : t_i \in \mathcal{D}(j)\}}. \tag{6b}$$

Here, (6b) follows from (5) as $\hat{Z}$ and Z_1 have the same distribution and both are independent of the respective sources messages. Noting that the rest of the vertices are not the destination for any source message, the network code satisfies the decoding requirements of I.

Each eavesdropper $r \in \mathcal{R} = \hat{\mathcal{R}}$ has access to messages on the edge set $\mathcal{B}_r$ consisting of

- edge $(1, 2)$, which carries $\hat{f}(\boldsymbol{X}_\mathcal{S}, Z)$, and
- edges $\{\mathsf{out}(s_i) : i \in \hat{\mathcal{B}}_r\}$, which carry messages $\boldsymbol{X}_{\hat{\mathcal{B}}_r}$.

Now,

$$H(\boldsymbol{X}_{\mathcal{A}_r} | \boldsymbol{X}_{\mathcal{B}_r}) = H(\boldsymbol{X}_{\hat{\mathcal{A}}_r} | \hat{f}(\boldsymbol{X}_\mathcal{S}, Z), \boldsymbol{X}_{\hat{\mathcal{B}}_r}) \tag{7a}$$

$$= H(\boldsymbol{X}_{\hat{\mathcal{A}}_r}) = H(\boldsymbol{X}_{\mathcal{A}_r}), \tag{7b}$$

where (7b) follows from (4) with a change of variables (from non-hatted to hatted).

This completes the proof that the network code is also secure against all the eavesdroppers described by W.

I is $(\boldsymbol{R}_\mathcal{S}, n)$-feasible $\Rightarrow$ $\hat{I}$ is $(\hat{\boldsymbol{R}}_{\hat{\mathcal{S}}}, \hat{c}_\text{b}, n)$-feasible:

Let $\{\mathcal{F}, \mathcal{G}\}$ be a secure network code that supports $(\boldsymbol{R}_\mathcal{S}, n)$. Then,

$$H(\boldsymbol{X}_{\{s \in \mathcal{S} : u \in \mathcal{D}(s)\}} | \boldsymbol{X}_{\mathsf{in}(u)}, X_{O^{-1}(u)}) = 0, \quad u \in \mathcal{V}, \tag{8}$$

and

$$H(\boldsymbol{X}_{\mathcal{A}_r} | \boldsymbol{X}_{\mathcal{B}_r}) = H(\boldsymbol{X}_{\mathcal{A}_r} | \{f_e : e \in \mathcal{B}_r\})$$

$$= H(\boldsymbol{X}_{\mathcal{A}_r}), \quad r \in \mathcal{R}. \tag{9}$$

By the construction of I, we know that $\{s \in \mathcal{S} : u \in \mathcal{D}(s)\} = \emptyset$, for any $u \in \{s_1, s_2, \dots, s_k, 1, 2\}$, and $\{s \in \mathcal{S} : t_i \in \mathcal{D}(s)\} = \hat{\mathcal{W}}_i$ for $i \in \{1, 2, \dots, \ell\}$. Also, none of the vertices in $\{t_1, t_2, \dots, t_\ell\}$ is the originating node for any source message.

So, (8) becomes

$$H(\boldsymbol{X}_{\hat{\mathcal{W}}_i}|\boldsymbol{X}_{\mathsf{in}(t_i)}) = H(\boldsymbol{X}_{\hat{\mathcal{W}}_i}|X_{(2,t_i)}, X_{l_1}, \ldots, X_{l_L}) = 0, \tag{10}$$

for $i \in [\ell]$, where $\mathsf{in}(t_i) = \{(2, t_i), l_1, l_2, \ldots, l_L\}$, $L = |\hat{\mathcal{H}}_i|$ and

$$X_{(2,t_i)} = f_{2\to t_i}(\bar{f}_{(2,t_i)}(\boldsymbol{X}_{\mathcal{S}}, \boldsymbol{Z}), Z_2) \tag{11}$$

$$\boldsymbol{X}_{\{l_1,\ldots,l_L\}} = [f_{(s_{h_1},t_i)}(X_{h_1}, Z_{s_{h_1}}), \ldots, f_{(s_{h_L},t_i)}(X_{h_L}, Z_{s_{h_L}})], \tag{12}$$

where $\hat{\mathcal{H}}_i = \{h_1, \ldots, h_L\}$, $\boldsymbol{Z}$ is the collection of all $\{Z_i\}$ generated by nodes $\{1, s_1, \ldots, s_{|\mathcal{S}|}\}$, and Z_2 is independent of all messages and $\boldsymbol{Z}$.

Decoding at all t_i's must succeed for any realisation of $\{Z_j\}$. Hence, (10) must also hold when all $Z_{s_j} = 0$. This gives

$$H(\boldsymbol{X}_{\hat{\mathcal{W}}_i}|X'_{(2,t_i)}, X'_{l_1}, \ldots, X'_{l_L}) = 0, \tag{13}$$

where

$$X'_{(2,t_i)} = f_{(2,t_i)}(\bar{f}_{(1,2)}(\boldsymbol{X}_{\mathcal{S}}, Z_1), Z_2), \tag{14}$$

$$\boldsymbol{X}'_{\{\ell_1,\ldots,\ell_L\}} = [f_{(s_{h_1},t_i)}(X_{h_1}, 0), \ldots, f_{(s_{h_L},t_i)}(X_{h_L}, 0)], \tag{15}$$

by setting $Z_{s_i} = 0$ for all vertices s_i, $i = 1, \ldots, k$.

Now, we construct a secure index code as follows:

$$\hat{X}_{\mathsf{b}} = \hat{f}(\hat{\boldsymbol{X}}_{\hat{\mathcal{S}}}, \hat{Z}) = \bar{f}_{(1,2)}(\hat{\boldsymbol{X}}_{\hat{\mathcal{S}}}, \hat{Z}), \tag{16}$$

where $\hat{Z} \in [2^{\hat{c}_{\mathsf{b}} n}]$ is a random variable independent of the messages, and having the same distribution as Z_1.

Now, for each receiver $\hat{t} \in \hat{\mathcal{T}}$,

$$\begin{aligned}
&H(\hat{\boldsymbol{X}}_{\hat{\mathcal{W}}_{\hat{t}}}|\hat{X}_{\mathsf{b}}, \hat{\boldsymbol{X}}_{\hat{\mathcal{H}}_{\hat{t}}}) \\
&= H(\hat{\boldsymbol{X}}_{\hat{\mathcal{W}}_{\hat{t}}}|\hat{X}_{\mathsf{b}}, \hat{\boldsymbol{X}}_{\hat{\mathcal{H}}_{\hat{t}}}, f_{(s_{h_1},t_{\hat{t}})}(\hat{X}_{h_1}, 0), \ldots, f_{(s_{h_L},t_{\hat{t}})}(\hat{X}_{h_L}, 0)) \\
&= H(\hat{\boldsymbol{X}}_{\hat{\mathcal{W}}_{\hat{t}}}|\hat{X}_{\mathsf{b}}, \hat{\boldsymbol{X}}_{\hat{\mathcal{H}}_{\hat{t}}}, f_{(s_{h_1},t_{\hat{t}})}(\hat{X}_{h_1}, 0), \ldots, f_{(s_{h_L},t_{\hat{t}})}(\hat{X}_{h_L}, 0), \\
&\qquad f_{(2,t_{\hat{t}})}(\hat{X}_{\mathsf{b}}, \hat{Z}_2)) \qquad (17a) \\
&\le H(\hat{\boldsymbol{X}}_{\hat{\mathcal{W}}_{\hat{t}}}|f_{(s_{h_1},t_{\hat{t}})}(\hat{X}_{h_1}, 0), \ldots, f_{(s_{h_L},t_{\hat{t}})}(\hat{X}_{h_L}, 0), \\
&\qquad f_{(2,t_{\hat{t}})}(\hat{X}_{\mathsf{b}}, \hat{Z}_2)) \\
&= 0, \qquad (17b)
\end{aligned}$$

where (17b) follows from (13) with a change of variables (from hatted to non-hatted), and (17a) follows from the following Markov chains

$$\hat{\boldsymbol{X}}_{\hat{\mathcal{W}}_{\hat{t}}} - (\hat{X}_{\mathsf{b}}, \hat{\boldsymbol{X}}_{\hat{\mathcal{H}}_{\hat{t}}}, f_{s_{h_1}\to t_{\hat{t}}}(\hat{X}_{h_1}, 0), \ldots, f_{s_{h_L}\to t_{\hat{t}}}(\hat{X}_{h_L}, 0)) - f_{2\to t_{\hat{t}}}(\hat{X}_{\mathsf{b}}, \hat{Z}_2), \tag{18}$$

as $\hat{Z}_2$ is independent of all other random variables, and has the same distribution as Z_2.

Since conditional entropy is non-negative, it follows from (17b) that each receiver $\hat{t} \in \hat{\mathcal{T}}$ can decode the messages $\hat{\boldsymbol{X}}_{\hat{\mathcal{W}}_{\hat{t}}}$ that it requires, given $(\hat{X}_{\mathsf{b}}, \hat{\boldsymbol{X}}_{\hat{\mathcal{H}}_{\hat{t}}})$.

We now prove the security constraints in $\hat{I}$. First, consider I. If X_i is requested by some nodes, then observing all outgoing links from s_i must enable one to reconstruct X_i. If X_j is not requested by any node, we assume that observing all outgoing links from s_j also enables one to reconstruct X_j. The rationale behind this assumption is as follows: If $R_j = R'$ is in the feasible region, then all $R_j > R'$ are also in the feasible region by having node s_j transmitting the first nR' bits of X_j. Hence, we only need to consider the smallest feasible rate for X_j, denoted by R_{min}, when all the other rates are kept fixed. Now, if after observing all outgoing links from s_j, one can obtain only nR''_j bits of information of X_j (where $R''_j < R_{\text{min}}$), then node s_j could have transmitted X_j at rate R''_j, which contradicts that R_{min} is the smallest feasible rate for X_j.

So, from (9), we have

$$\begin{aligned}
H(\boldsymbol{X}_{\mathcal{A}_r}) &= H(\boldsymbol{X}_{\mathcal{A}_r}|\{X_e : e \in \mathcal{B}_r\}) && (19a) \\
&= H(\boldsymbol{X}_{\mathcal{A}_r}|\bar{f}_{(1,2)}(\boldsymbol{X}_{\mathcal{S}}, Z_1), \boldsymbol{X}_{\{\mathsf{out}(s_i):i\in\hat{\mathcal{B}}_r\}}) && (19b) \\
&= H(\boldsymbol{X}_{\mathcal{A}_r}|\bar{f}_{(1,2)}(\boldsymbol{X}_{\mathcal{S}}, Z_1), \boldsymbol{X}_{\{\mathsf{out}(s_i):i\in\hat{\mathcal{B}}_r\}}, \boldsymbol{X}_{\hat{\mathcal{B}}_r}) && (19c) \\
&\le H(\boldsymbol{X}_{\mathcal{A}_r}|\bar{f}_{(1,2)}(\boldsymbol{X}_{\mathcal{S}}, Z_1), \boldsymbol{X}_{\hat{\mathcal{B}}_r}) && (19d) \\
&\le H(\boldsymbol{X}_{\mathcal{A}_r}), && (19e)
\end{aligned}$$

where (19c) is derived because observing all outgoing links from s_i allows one to reconstruct X_i. Thus, it follows that (19d) and (19e) must hold with equality.

Now, consider an eavesdropper $\hat{r} \in \hat{\mathcal{R}}$ in the index-coding equivalence $\hat{I}$.

$$\begin{aligned}
H(\hat{\boldsymbol{X}}_{\hat{\mathcal{A}}_{\hat{r}}}|\hat{\boldsymbol{X}}_{\mathsf{b}}, \hat{\boldsymbol{X}}_{\hat{\mathcal{B}}_{\hat{r}}}) &= H(\hat{\boldsymbol{X}}_{\hat{\mathcal{A}}_{\hat{r}}}|\bar{f}_{(1,2)}(\hat{\boldsymbol{X}}_{\hat{\mathcal{S}}}, \hat{Z}), \hat{\boldsymbol{X}}_{\hat{\mathcal{B}}_{\hat{r}}}) \\
&= H(\hat{\boldsymbol{X}}_{\mathcal{A}_{\hat{r}}}) = H(\hat{\boldsymbol{X}}_{\hat{\mathcal{A}}_{\hat{r}}}), \qquad (20a)
\end{aligned}$$

where (20a) follows from (19a)–(19e) with a change of variables (from non-hatted to hatted), by noting that $(\hat{\boldsymbol{X}}_{\hat{\mathcal{S}}}, \hat{Z})$ and $(\boldsymbol{X}_{\mathcal{S}}, Z_1)$ have the same distribution.

So, $\hat{X}_{\mathsf{b}}$ is a secure index code for $\hat{I}$. ■

IV. Mapping from Secure Network Coding to Secure Index Coding

Given a secure network-coding instance $\mathcal{I} = (G, M, W)$, we first construct an augmented secure network-coding instance with deterministic encoding, and then construct an equivalent secure index-coding instance $\hat{I} = (\hat{\mathcal{S}}, \hat{\mathcal{T}}, \{\hat{\mathcal{W}}_{\hat{t}}\}, \{\hat{\mathcal{H}}_{\hat{t}}\}, \hat{W})$.

Augmented secure network coding: We construct an *augmented* secure network-coding instance $I' = (G', M', W')$ as follows:

- $G' = (\mathcal{V}', \mathcal{E}') = G = (\mathcal{V}, \mathcal{E})$, and $c'_e = c_e$ for all $e \in \mathcal{E}'$. The vertices, the edges, and the edge capacities remain the same.
- Let $\mathcal{S} = [S]$, where $S \triangleq |\mathcal{S}|$. The connection requirement is augmented as follows: $\mathcal{S}' = \mathcal{S} \cup \{S+1, S+2, \ldots, S+|\mathcal{V}'|\}$, where we introduce an additional source $X'_{S+v} \in [\prod_{e\in\mathsf{out}(v)} 2^{c_e n}] \triangleq [2^{k_v n}]$, originating at each vertex $v \in [|\mathcal{V}'|]$, that takes the role of and has the same distribution as the random key Z_v used in the randomised encoding at vertex v in I. So, $O'(S+v) = v$, i.e., X'_{S+v} originates at vertex v, and we define $\mathcal{D}'(S+v) = \emptyset$, i.e., X'_{S+v} is not requested by any vertex. For $s \in \mathcal{S}$, $\boldsymbol{X}'_{\mathcal{S}}$ has the same distribution as $\boldsymbol{X}_{\mathcal{S}}$, $O'(s) = O(s)$, and $\mathcal{D}'(s) = \mathcal{D}(s)$. Note that $R'_s = R_s$ for all $s \in [S]$, and $R'_{S+v} = k_v$ for all $v \in [|\mathcal{V}'|]$.

- $W' = W$, i.e., $\mathcal{R}' = \mathcal{R}$, $\mathcal{B}'_r = \mathcal{B}_r$, and $\mathcal{A}'_r = \mathcal{A}_r$. The adversarial setting remains the same. Thus, messages $\{X'_{S+v} : v \in [|\mathcal{V}'|]\}$ are neither known to the adversaries nor need to be protected.

Any deterministic or randomised (i.e., using an independent random key Z_v at vertex v) secure network code for I is equivalent to a deterministic secure network code for I', where each node v gets an additional source X'_{S+v} that is not required to be decoded by any node.

Denote the set of vertices in $\mathcal{I}'$ that are destinations for some source messages by $\mathcal{T}' = \{j \in \mathcal{V}' : j \in \mathcal{D}'(i) \text{ for some } i \in \mathcal{S}'\}$. Note that $O'(\cdot)$ can map different source indices to one vertex, and hence, $O'^{-1}(j)$ returns a set of indices of messages originating at vertex j.

Network-to-index coding mapping:

- $\hat{\mathcal{S}} = \mathcal{S}' \cup \mathcal{E}'$. It consists of one source message $\hat{X}_s$ for each $s \in \mathcal{S}'$ in I', and one $\hat{X}_e$ for each edge $e \in \mathcal{E}'$ in I'. $\hat{\boldsymbol{X}}_{\mathcal{S}'}$ has the same distribution as $\boldsymbol{X}'_{\mathcal{S}'}$. The rates of the messages are $\hat{R}_s = R'_s$ and $\hat{R}_e = c'_e$.
- $\hat{\mathcal{T}} = \{\hat{t}_i\}_{i \in \mathcal{T}'} \cup \{\hat{t}_e\}_{e \in \mathcal{E}'}$. This means $\hat{I}$ has $|\mathcal{T}'| + |\mathcal{E}'|$ receivers, one for each destination node in I' and one for each edge in I'.
- For each $\hat{t}_e \in \hat{\mathcal{T}}$, $\hat{\mathcal{H}}_{\hat{t}_e} = \mathsf{in}(\mathsf{tail}(e)) \cup O'^{-1}(\mathsf{tail}(e))$, and $\hat{\mathcal{W}}_{\hat{t}_e} = \{e\}$.
- For each $\hat{t}_i \in \hat{\mathcal{T}}$, $\hat{\mathcal{H}}_{\hat{t}_i} = \mathsf{in}(i) \cup O'^{-1}(i)$, and $\hat{\mathcal{W}}_{\hat{t}_i} = \{s \in [S] : i \in \mathcal{D}'(s)\}$.
- The eavesdropper setting W': $\hat{\mathcal{R}} = \mathcal{R}'$. For each $\hat{r} \in \hat{\mathcal{R}}$, $\hat{\mathcal{B}}_{\hat{r}} = \mathcal{B}'_{\hat{r}}$, and $\hat{\mathcal{A}}_{\hat{r}} = \mathcal{A}'_{\hat{r}}$.
- We set the broadcast rate as $\hat{c}_{\mathsf{b}} = \sum_{e \in \mathcal{E}'} c'_e$.

Figure 2 depicts an example of such a mapping.

Remark 1: This network-to-index coding mapping is slightly different from that of Effros et al. [3], since we do not require the use of an additional receiver $\hat{t}_{\text{all}}$ for the corresponding index-coding instance. We will show that omitting this receiver will not affect the equivalence.

Note that unlike the index-to-network mapping, here $\hat{\boldsymbol{X}}_{\mathcal{E}'}$ and $\boldsymbol{X}'_{\mathcal{E}'}$ have different distributions, where the latter are functions of $\boldsymbol{X}'_{\mathcal{S}'}$. For the corresponding secure index-coding instance, we choose $\{\hat{X}_e : e \in \mathcal{E}'\}$ to be mutually independent, independent of all other messages, and each $\hat{X}_e$ is uniformly distributed over $[2^{\hat{R}_e n}]$. We will see that using uniformly distributed $\hat{X}_e$ is the key to ensuring security.

With the above conversion, we now state an equivalence between I and $\hat{I}$ through I':

Theorem 2: Let I be a secure network-coding instance and I' be its augmented instance. Let $\hat{I}$ and $\hat{c}_{\mathsf{b}}$ be the corresponding secure index-coding instance and a broadcast rate, respectively, obtained using the network-to-index coding mapping from I'. For any $\boldsymbol{R}_{\mathcal{S}}$, the instance I is $(\boldsymbol{R}_{\mathcal{S}}, n)$-feasible if and only if the instance $\hat{I}$ is $(\hat{\boldsymbol{R}}_{\hat{\mathcal{S}}}, \hat{c}_{\mathsf{b}}, n)$-feasible.

Proof of Theorem 2:

I is $(\boldsymbol{R}_{\mathcal{S}}, n)$-feasible $\Rightarrow$ $\hat{I}$ is $(\hat{\boldsymbol{R}}_{\hat{\mathcal{S}}}, \hat{c}_{\mathsf{b}}, n)$-feasible:

Note that I is $(\boldsymbol{R}_{\mathcal{S}}, n)$-feasible if and only if I' is $(\boldsymbol{R}'_{\mathcal{S}'}, n)$-feasible using *deterministic* network encoding functions $\{f'_e\}$ derived from $\{f_e\}$ for I, where all the randomness $\{Z_v\}$ in the network code for I is realised using $\{X_{S+v}\}$ in I'.

Since the network code for I' is deterministic, we use the same code mapping as that proposed by Effros et al. [3]: The sender's broadcast message is $\hat{X}_{\mathsf{b}} = [\hat{X}_{\mathsf{b}}(e)]_{e \in \mathcal{E}'}$, where

$$\hat{X}_{\mathsf{b}}(e) = \hat{X}_e + \bar{f}'_e(\hat{\boldsymbol{X}}_{\mathcal{S}'}). \tag{21}$$

Note that $\hat{X}_e, \bar{f}'_e \in [2^{\hat{R}_e n}] = [2^{c'_e n}] = [2^{c_e n}]$.

In I', each vertex $v \in \mathcal{T}'$ can decode all messages that it requires from the message on all incoming edges and messages originating at v, meaning that

$$\begin{aligned}\boldsymbol{X}'_{\{s \in \mathcal{S}' : v \in \mathcal{D}'(s)\}} &= \boldsymbol{X}'_{\{s \in [S] : v \in \mathcal{D}'(s)\}} = g'_v(\boldsymbol{X}'_{\mathsf{in}(v) \cup O'^{-1}(v)}) \\ &= g'_v(\boldsymbol{X}'_{\mathsf{in}(v)}, \boldsymbol{X}'_{O'^{-1}(v)}) \\ &= g'_v([\bar{f}'_e(\boldsymbol{X}'_{\mathcal{S}'})]_{e \in \mathsf{in}(v)}), \boldsymbol{X}'_{O'^{-1}(v)}).\end{aligned}$$

As mentioned above, while messages $\boldsymbol{X}'_{O'^{-1}(v)}$ and $\hat{\boldsymbol{X}}_{O'^{-1}(v)}$ (with node subscripts) have the same distribution, messages $\boldsymbol{X}'_{\mathsf{in}(v)}$ and $\hat{\boldsymbol{X}}_{\mathsf{in}(v)}$ (with edge subscripts) may not. To deal with this issue, consider the broadcast message $\hat{X}_{\mathsf{b}}$. From (21), any receiver that knows $\hat{X}_e$ can obtain $\bar{f}'_e(\hat{\boldsymbol{X}}_{\mathcal{S}'})$, where $[\bar{f}'_e(\hat{\boldsymbol{X}}_{\mathcal{S}'})]_{e \in \mathcal{E}'}$ and $[\bar{f}'_e(\boldsymbol{X}'_{\mathcal{S}'})]_{e \in \mathcal{E}'}$ have the same distribution.

So, with a change of variables (from non-hatted to hatted), receiver $\hat{t}_i \in \hat{\mathcal{T}}$ can decode the messages it requires using

$$\hat{\boldsymbol{X}}_{\hat{\mathcal{W}}_{\hat{t}_i}} = \hat{\boldsymbol{X}}_{\{s \in [S] : i \in \mathcal{D}'(s)\}} = g'_i([\bar{f}'_e(\hat{\boldsymbol{X}}_{\mathcal{S}'})]_{e \in \mathsf{in}(i)}), \hat{\boldsymbol{X}}_{O'^{-1}(i)}).$$

As receiver $\hat{t}_i$ knows $\hat{\mathcal{H}}_{\hat{t}_i} = \mathsf{in}(i) \cup O'^{-1}(i)$ by the mapping, it knows $\hat{\boldsymbol{X}}_{O'^{-1}(i)}$ and can obtain $[\bar{f}'_e(\hat{\boldsymbol{X}}_{\mathcal{S}'})]_{e \in \mathsf{in}(i))}$ from $\hat{X}_{\mathsf{b}}$ and $\hat{\boldsymbol{X}}_{\mathsf{in}(i)}$.

Receiver $\hat{t}_e \in \hat{\mathcal{T}}$ uses (21) to obtain the required $\hat{X}_e$ from $\hat{X}_{\mathsf{b}}(e) - \bar{f}'_e(\hat{\boldsymbol{X}}_{\mathcal{S}'})$, where the first term is the broadcast message available to the receiver $\hat{t}_e$. To obtain the second term, express the global encoding function as its local encoding function, $\bar{f}'_e(\hat{\boldsymbol{X}}_{\mathcal{S}'}) = f'_e([\bar{f}'_{e'}(\hat{\boldsymbol{X}}_{\mathcal{S}'})]_{e' \in \mathsf{in}(\mathsf{tail}(e))}, \hat{\boldsymbol{X}}_{O'^{-1}(\mathsf{tail}(e))})$, where $\hat{\boldsymbol{X}}_{O'^{-1}(\mathsf{tail}(e))}$ is available to receiver $\hat{t}_e$ as side information. From the broadcast message, receiver $\hat{t}_e$ can obtain $\bar{f}'_{e'}(\hat{\boldsymbol{X}}_{\mathcal{S}'}) = \hat{X}_{\mathsf{b}}(e') - \hat{X}_{e'}$, as it has $\hat{X}_{e'}$, $e' \in \mathsf{in}(\mathsf{tail}(e))$, as side information. With this, we have shown that each $\hat{t} \in \hat{\mathcal{T}}$ can decode the messages that it requires.

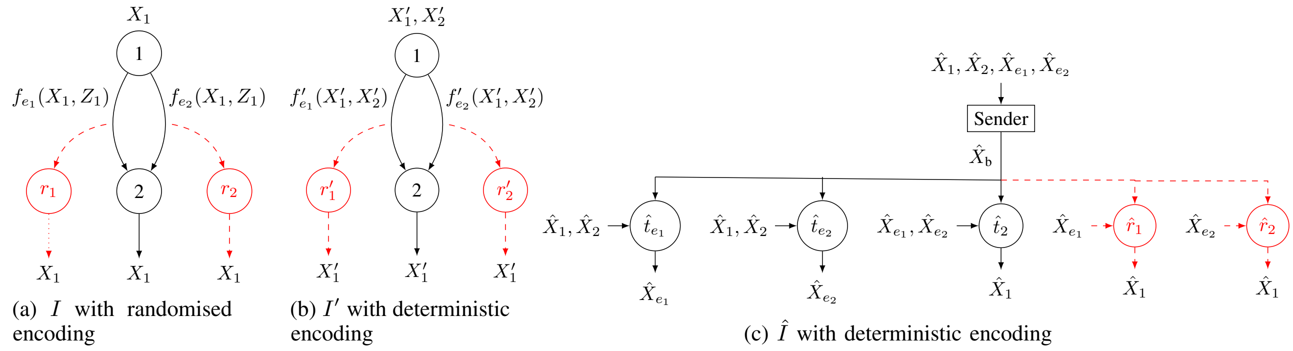
We now consider the security constraints. For each $\hat{r} \in \hat{\mathcal{R}}$,

$$\begin{aligned}&H(\hat{\boldsymbol{X}}_{\hat{\mathcal{A}}_{\hat{r}}} | \hat{X}_{\mathsf{b}}, \hat{\boldsymbol{X}}_{\hat{\mathcal{B}}_{\hat{r}}}) && \text{(22a)}\\ &= H(\hat{\boldsymbol{X}}_{\hat{\mathcal{A}}_{\hat{r}}} | \{\hat{X}_{\mathsf{b}}(e) : e \in \mathcal{E}'\}, \{\hat{X}_{e'} : e' \in \hat{\mathcal{B}}_{\hat{r}}\}) && \text{(22b)}\\ &= H(\hat{\boldsymbol{X}}_{\hat{\mathcal{A}}_{\hat{r}}} | \{\hat{X}_{\mathsf{b}}(e) : e \in \hat{\mathcal{B}}_{\hat{r}}\}, \{\hat{X}_{e'} : e' \in \hat{\mathcal{B}}_{\hat{r}}\}) && \text{(22c)}\\ &= H(\hat{\boldsymbol{X}}_{\hat{\mathcal{A}}_{\hat{r}}} | \{\hat{X}_{\mathsf{b}}(e), \hat{X}_{e'}, \bar{f}'_e(\hat{\boldsymbol{X}}_{\mathcal{S}'}) : e \in \hat{\mathcal{B}}_{\hat{r}}\}) && \text{(22d)}\\ &= H(\hat{\boldsymbol{X}}_{\hat{\mathcal{A}}_{\hat{r}}} | \{\bar{f}'_e(\hat{\boldsymbol{X}}_{\mathcal{S}'}) : e \in \hat{\mathcal{B}}_{\hat{r}}\}) && \text{(22e)}\\ &= H(\hat{\boldsymbol{X}}_{\mathcal{A}'_{\hat{r}}} | \{\bar{f}'_e(\hat{\boldsymbol{X}}_{\mathcal{S}'}) : e \in \mathcal{B}'_{\hat{r}}\}) && \text{(22f)}\\ &= H(\hat{\boldsymbol{X}}_{\mathcal{A}'_{\hat{r}}}) = H(\hat{\boldsymbol{X}}_{\hat{\mathcal{A}}_{\hat{r}}}), && \text{(22g)}\end{aligned}$$

where (22c) follows from the Markov chain

$$\begin{aligned}\hat{\boldsymbol{X}}_{\hat{\mathcal{A}}_{\hat{r}}} - \Big(\{\hat{X}_{\mathsf{b}}(e) : e \in \hat{\mathcal{B}}_{\hat{r}}\}, \{\hat{X}_{e'} : e' \in \hat{\mathcal{B}}_{\hat{r}}\}\Big) \\ - (\{\hat{X}_{\mathsf{b}}(e) : e \notin \hat{\mathcal{B}}_{\hat{r}}\}),\end{aligned}$$

where $\{\hat{X}_{\mathsf{b}}(e) : e \notin \hat{\mathcal{B}}_{\hat{r}}\}$ has been randomised by independently and uniformly distributed $\{\hat{X}_e : e \notin \hat{\mathcal{B}}_{\hat{r}}\}$, which are independent of $(\hat{\boldsymbol{X}}_{\hat{\mathcal{A}}_{\hat{r}}}, \hat{\boldsymbol{X}}_{\hat{\mathcal{B}}_{\hat{r}}}, \hat{\boldsymbol{X}}_{\mathcal{S}'})$ (see (21)); (22d) follows



(a) I with randomised encoding

(b) I' with deterministic encoding

(c) $\hat{I}$ with deterministic encoding

Fig. 2: A secure network-coding instance I, its augmented version I', and the corresponding secure index-coding instance $\hat{I}$, where $r_1, r_2, r'_1, r'_2, \hat{r}_1, \hat{r}_2$ are eavesdroppers

from (21); (22e) follows from the Markov chain

$$\hat{\boldsymbol{X}}_{\hat{\mathcal{A}}_{\hat{r}}} - \{\bar{f}'_e(\hat{\boldsymbol{X}}_{\mathcal{S}'}) : e \in \hat{\mathcal{B}}_{\hat{r}}\} - \{\hat{X}_{e'}, \hat{X}_\mathsf{b}(e) : e \in \hat{\mathcal{B}}_{\hat{r}}\},$$

which can be derived from (21) and noting that $\{\hat{X}_e : e \in \mathcal{E}'\}$ are independent of $(\hat{\boldsymbol{X}}_{\hat{\mathcal{A}}_{\hat{r}}}, \hat{\boldsymbol{X}}_{\hat{\mathcal{B}}_{\hat{r}}}, \hat{\boldsymbol{X}}_{\mathcal{S}'})$; and (22g) follows from (1) by a change of variables (from hatted to non-hatted) and noting that $\{\bar{f}'_e(\boldsymbol{X}'_{\mathcal{S}'}) : e \in \mathcal{B}'_{\hat{r}}\} = \boldsymbol{X}'_{\mathcal{B}'_{\hat{r}}}$

So, the index code is secure.

$\hat{I}$ is $(\hat{\boldsymbol{R}}_{\hat{\mathcal{S}}}, \hat{c}_\mathsf{b}, n)$-feasible $\Rightarrow$ I is $(\boldsymbol{R}_\mathcal{S}, n)$-feasible:

We will show that if $\hat{I}$ is $(\hat{\boldsymbol{R}}_{\hat{\mathcal{S}}}, \hat{c}_\mathsf{b}, n)$-feasible, then I' is $(\boldsymbol{R}'_{\mathcal{S}'}, n)$-feasible, which implies that I is $(\boldsymbol{R}_\mathcal{S}, n)$-feasible.

Again, we use the network-code construction proposed by Effros et al. [3]. Note that for a secure index code, there exists a decoding function at receiver $\hat{t}_i$ for each $i \in \mathcal{T}'$, such that

$$\hat{g}_{\hat{t}_i}(\hat{X}_\mathsf{b}, \hat{\boldsymbol{X}}_{\hat{\mathcal{H}}_{\hat{t}_i}}) = \hat{g}_{\hat{t}_i}(\hat{X}_\mathsf{b}, \hat{\boldsymbol{X}}_{\mathsf{in}(i) \cup O'^{-1}(i)}) \tag{23a}$$

$$= \hat{\boldsymbol{X}}_{\hat{\mathcal{W}}_{\hat{t}_i}} = \hat{\boldsymbol{X}}_{\{s \in [S] : i \in \mathcal{D}'(s)\}}, \tag{23b}$$

and a decoding function at receiver $\hat{t}_e$, $e \in \mathcal{E}'$, such that

$$\hat{g}_{\hat{t}_e}(\hat{X}_\mathsf{b}, \hat{\boldsymbol{X}}_{\hat{\mathcal{H}}_{\hat{t}_e}}) = \hat{g}_{\hat{t}_e}(\hat{X}_\mathsf{b}, \hat{\boldsymbol{X}}_{\mathsf{in}(\mathsf{tail}(e)) \cup O'^{-1}(\mathsf{tail}(e))}) \tag{24a}$$

$$= \hat{\boldsymbol{X}}_{\hat{\mathcal{W}}_{\hat{t}_e}} = \hat{X}_e. \tag{24b}$$

In the secure index-coding instance $\hat{I}$, messages $\hat{\boldsymbol{X}}_{\mathcal{E}'}$ are independent of messages $\hat{\boldsymbol{X}}_{\mathcal{S}'}$, and the broadcast message $\hat{X}_\mathsf{b}$ is a function of these messages. However, given $\hat{X}_\mathsf{b}$, the messages $\hat{\boldsymbol{X}}_{\mathcal{E}'}$ and $\hat{\boldsymbol{X}}_{\mathcal{S}'}$ are dependent.

We set $\hat{X}_\mathsf{b} = \sigma$ (which is an arbitrary but valid realisation of $\hat{X}_\mathsf{b}$ in the network code) for all $\hat{g}_{\hat{t}_i}(\cdot)$ and $\hat{g}_{\hat{t}_e}(\cdot)$, and choose

$$X'_e = g'_e(\boldsymbol{X}'_{\mathsf{in}(\mathsf{tail}(e)) \cup O'^{-1}(\mathsf{tail}(e))}) \tag{25}$$

$$= \hat{g}_{\hat{t}_e}(\sigma, \boldsymbol{X}'_{\mathsf{in}(\mathsf{tail}(e)) \cup O'^{-1}(\mathsf{tail}(e))}), \tag{26}$$

for all edges $e \in \mathcal{E}'$, and

$$g'_i(\boldsymbol{X}'_{\mathsf{in}(i) \cup O'^{-1}(i)}) = \hat{g}_{\hat{t}_i}(\sigma, \boldsymbol{X}'_{\mathsf{in}(i) \cup O'^{-1}(i)}), \tag{27}$$

for each destination vertex $i \in \mathcal{T}'$. By fixing the first argument in the functions to be σ, $\boldsymbol{X}'_{\mathcal{E}'}$ are now functions of the source messages $\boldsymbol{X}'_{\mathcal{S}'}$, and they can be generated following the (acyclic) graph topology of I'. Now, (23a)–(24b) hold for any realisation of the variables $\hat{\boldsymbol{X}}_{\mathcal{S}' \cup \mathcal{E}'}$. So, for any realisation of the messages $\boldsymbol{x}'_{\mathcal{S}'}$, using (26) with the chosen σ, and following the topology of G', we can generate the correct and unique realisation of x'_e for every edge e. This will ensure that the decoding step (27) for each destination $i \in \mathcal{T}'$ gives the correct $\boldsymbol{x}'_{\{s \in [S] : i \in \mathcal{D}'(s)\}}$. Thus, correct decoding can be achieved without using the additional receiver $\hat{t}_\mathsf{all}$ proposed by Effros et al. [3].

Finally, consider the security constraints of I'. Security for the index code implies that

$$H(\hat{\boldsymbol{X}}_{\hat{\mathcal{A}}_{\hat{r}}}) = H(\hat{\boldsymbol{X}}_{\hat{\mathcal{A}}_{\hat{r}}} | \hat{X}_\mathsf{b}, \hat{\boldsymbol{X}}_{\hat{\mathcal{B}}_{\hat{r}}}) \leq H(\hat{\boldsymbol{X}}_{\hat{\mathcal{A}}_{\hat{r}}} | \hat{X}_\mathsf{b}) \leq H(\hat{\boldsymbol{X}}_{\hat{\mathcal{A}}_{\hat{r}}}),$$

which implies that $\hat{\boldsymbol{X}}_{\hat{\mathcal{A}}_{\hat{r}}}$ and $\hat{\boldsymbol{X}}_{\hat{\mathcal{B}}_{\hat{r}}}$ are independent given $\hat{X}_b$. In particular, they are independent given $\hat{X}_b = \sigma$. As $\hat{\boldsymbol{X}}_{\mathcal{S}'}$ and $\boldsymbol{X}'_{\mathcal{S}'}$ have the same distribution, in the event that $\hat{X}_\mathsf{b} = \sigma$, we see from (23a)–(27) that $p(\boldsymbol{x}'_{\mathcal{S}'}, \boldsymbol{x}'_{\mathcal{E}'}) = p(\hat{\boldsymbol{x}}_{\mathcal{S}'}, \hat{\boldsymbol{x}}_{\mathcal{E}'} | \hat{x}_\mathsf{b} = \sigma)$. Since

$$I(\hat{\boldsymbol{X}}_{\hat{\mathcal{A}}_r}; \hat{\boldsymbol{X}}_{\hat{\mathcal{B}}_r} | \hat{X}_\mathsf{b} = \sigma) = 0 = I(\hat{\boldsymbol{X}}_{\mathcal{A}_r}; \hat{\boldsymbol{X}}_{\mathcal{B}_r} | \hat{X}_\mathsf{b} = \sigma), \tag{28}$$

we have $I(\boldsymbol{X}'_{\mathcal{A}_r}; \boldsymbol{X}'_{\mathcal{B}_r}) = 0$, which gives the required security constraint (1) for I'. ■

References

[1] R. Dougherty, K. Zeger, "Nonreversibility and equivalent constructions of multiple-unicast networks," *IEEE Trans. Inf. Theory*, vol. 52, no. 11, pp. 5067–5077, Nov. 2006.

[2] W. Huang, T. Ho, M. Langberg, J. Kliewer, "On secure network coding with uniform wiretap sets," in *Proc. IEEE NetCod*, 2013.

[3] M. Effros, S. El Rouayheb, M. Langberg, "An equivalence between network coding and index coding," *IEEE Trans. Inf. Theory*, vol. 61, no. 5, pp. 2478–2487, May 2015.

[4] R. Ahlswede, N. Cai, S. R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Inf. Theory*, vol. 46, no. 4, pp. 1204–1216, July 2000.

[5] Z. Bar-Yossef, Y. Birk, T. S. Jayram, and T. Kol, "Index coding with side information," *IEEE Trans. Inf. Theory*, vol. 57, no. 3, pp. 1479–1494, Mar. 2011.

[6] S. El Rouayheb, A. Sprintson, and C. Georghiades, "On the index coding problem and its relation to network coding and matroid theory," *IEEE Trans. Inf. Theory*, vol. 56, no. 7, pp. 3187–3195, July 2010.

[7] N. Cai and R. W. Yeung, "Secure network coding on wiretap network," *IEEE Trans. Inf. Theory*, vol. 57, no. 1, pp. 424–435, Jan. 2011.

[8] M. M. Mojahedian, A. Gohari, and M. R. Aref. "Perfectly secure index coding," in *Proc. IEEE ISIT*, 2015.

[9] K. Bhattad and K. R. Narayanan, "Weakly secure network coding," in *Proc. Netcod*, 2005.

[10] S. H. Dau, V. Skachek, and Y. M. Chee, "On the security of index coding with side information," *IEEE Trans. Inf. Theory*, vol. 58, no. 6, pp. 3975–3988, June 2012.

[11] L. Ong, B. N. Vellambi, P. L. Yeoh, J. Kliewer, and J. Yuan, "Secure index coding: Existence and construction," in *Proc. IEEE ISIT*, 2016.

[12] T. Chan and A. Grant, "Capacity bounds for secure network coding," in *Proc. AusCTW*, 2008.