Secure Network-Index Code Equivalence: Extension to Non-zero Error and Leakage

Lawrence Ong
University of Newcastle
Email: lawrence.ong@newcastle.edu.au

Jörg Kliewer
New Jersey Institute of Technology
Email: jkliewer@njit.edu

Badri N. Vellambi
Australian National University
Email: badri.vellambi@anu.edu.au

***Abstract*—A linear code equivalence between index coding and network coding was shown by El Rouayheb et al., which establishes that for any index-coding instance, there exists a network-coding instance for which any index code can be mapped to a suitable network code, and vice versa. Similarly, for any network-coding instance, there exists an index-coding instance for which a similar code equivalence can be constructed. Effros et al. extended the equivalence to include non-linear codes. Subsequently, we extended the code equivalence to the secure communication setting in the presence of an eavesdropper, in which we impose perfect decodability and secrecy. In this paper, we generalise the equivalence between secure index coding and secure network coding to include non-zero decoding error and non-zero leakage.**

I. Introduction

In this paper, we investigate an equivalence between secure index coding and secure network coding. Ong et al. [1] proved an equivalence between secure index coding and secure network coding in the special case when there is no decoding error (that is, perfect message reconstruction at the receivers) and no leakage (that is, zero mutual information between what the eavesdroppers observe and what they attempt to get). In general, moving from zero decoding error and leakage scenarios to the non-zero counterpart (for example, diminishing error and leakage as the codelength is allowed to grow) changes the underlying problem significantly. Results in these two regimes are often significantly different. In this paper, we show an equivalence between secure index coding and secure network coding for non-zero decoding error and non-zero leakage. To this end, we show that the probability distribution function of the messages in the network-coding instance is close to that in the index-coding instance.

A. Background

Index coding [2] considers a one-hop network where a sender conveys multiple messages to multiple receivers through a noiseless broadcast medium, where each receiver wants some messages from the sender, but already knows some other messages. On the other hand, network coding [3] considers a network of interconnected links with fixed capacities, where multiple senders send multiple messages to multiple receivers through these links.

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

Although these two problems appear different prima facie, the following equivalence between them has been demonstrated [4, 5]: for any index-coding instance (specified by what each receiver has and wants), one can construct an equivalence network-coding instance (specified by how the links are connected, their capacities, and all sender and receiver locations), such that any index code (specified by the encoding function of the sender and the decoding functions of all the receivers) for the index-coding instance can be mapped to a network code for the same message sizes (specified by the encoding function of every node, and the decoding functions of all receivers) for the network-coding instance, and vice versa. Similarly, for any network-coding instance, we can construct an equivalent index-coding instance with code mapping in both directions.

The equivalence was first shown for linear codes [4] and then for non-linear codes (which include linear codes as a special case) [5]. Furthermore, the equivalence has been shown for any codelength and (zero and non-zero) decoding error probability, that is, if the probability of decoding error for the network code is bounded above by a given value, the mapped index code also has this property, and vice versa.

The secure version of index coding [6] includes a number of eavesdroppers each of whom (i) knows a subset of messages; (ii) listens to the sender's broadcast; and (iii) attempts to decode some messages. The secure version of network coding [7] includes a number of eavesdropper each of whom (i) can listen to a subset of links; and (ii) attempts to decode some messages. A secure index code or a secure network code must prevent eavesdroppers from *knowing* the messages that they attempt to decode (where knowing is quantified by the information-theoretic security measure [8, Ch 22]), in addition to guaranteeing that all receivers can obtain their requested messages (by bounding the probability of decoding error).

Recently, Ong et al. [1] showed an equivalence between secure index coding and secure network coding for zero decoding error and leakage (that is, the information gained by the eavesdroppers about the messages they attempt to decode) by constructing a mapping between secure index and network coding instances and using the existing (non-secure) translation between index and network codes.by Effros et al.

B. Main contributions

In this paper, we extend the code equivalence between secure index and network coding to non-zero error and leakage

978-1-5386-4780-6/18/$31.00©2018 IEEE

by showing that the instance mapping by Ong et al. and the code translation by Effros et al. preserve both decoding and leakage criteria (which can be non-zero) to a certain extent.

Informally, in Theorem 1, we show that any secure index-coding instance $\mathbb{I}_1$ can be mapped to a secure network-coding instance $\mathbb{N}_1$, such that any code for $\mathbb{I}_1$ can be translated to a code for $\mathbb{N}_1$ (and vice versa) with the same error decoding and security criteria.

In Theorem 2 and Corollary 2.1, we show that any secure network-coding instance $\mathbb{N}_2$ can be mapped to a secure index-coding instance $\mathbb{I}_2$ such that

1) any code for $\mathbb{N}_2$ can be translated to a code for $\mathbb{I}_2$ with the same error decoding and security criteria;
2) any code for $\mathbb{I}_2$ that has
   a) zero decoding error can be translated to code for $\mathbb{N}_2$ with the same error decoding and security criteria,
   b) non-zero decoding error and is linear can be translated to a linear code for $\mathbb{N}_2$ with a security criterion that grows linearly in the codelength, and a decoding criterion that does not grow with the codelength. This implies that strongly-secure index codes map to weakly-secure network codes.

For all cases except 2b, we establish an equivalence that preserves both the decodability and security criteria.

The challenge in obtaining an equivalence for non-zero error and leakage arises due to the fact that the eavesdroppers in both instances observe different signals, i.e., messages for index coding and functions of messages transmitted on links for network coding. If decoding error at the receivers is allowed, these two types of messages do not necessarily match, making it difficult to guarantee the same amount of leakage.

This problem is even more severe for case 2b, in which we need to select certain parameters for the index code to obtain the required network code, and the parameters must simultaneously satisfy both error and leakage criteria. To obtain the above equivalence result, we use the hypothesis that decoding is correct $(1-\epsilon)$ fraction of the time for $\mathbb{I}_2$ to bound the distance between the probability mass functions (pmf) of the messages in both instances.

II. Channel Model, Existing Instance Mapping, and Existing Code Translation

Notation: For a set $\mathcal{S}$, $\boldsymbol{X}_{\mathcal{S}} \stackrel{\text{def}}{=} (X_i : i \in \mathcal{S})$. Consider a directed graph $G = (\mathcal{V}, \mathcal{E})$ with node set $\mathcal{V}$ and edge set $\mathcal{E}$. For an edge $e = (u \to v) \in \mathcal{E}$, its tail is $\texttt{tail}(e) \stackrel{\text{def}}{=} u$, and its head is $\texttt{head}(e) \stackrel{\text{def}}{=} v$. For any node $v \in \mathcal{V}$, the set of incoming edges is denoted by $\texttt{in}(v) \stackrel{\text{def}}{=} \{e \in \mathcal{E} : \texttt{head}(e) = v\}$, and the set of outgoing edges by $\texttt{out}(v) \stackrel{\text{def}}{=} \{e \in \mathcal{E} : \texttt{tail}(e) = v\}$. For two ordered sets of discrete random variables $\boldsymbol{X}_{\mathcal{S}_1}$ and $\boldsymbol{Y}_{\mathcal{S}_2}$, $\boldsymbol{X}_{\mathcal{S}_1} \stackrel{\text{d}}{=} \boldsymbol{Y}_{\mathcal{S}_2}$ means that they have the same pmfs, and all corresponding pairs of random variables (one with index from $\mathcal{S}_1$ and another from $\mathcal{S}_2$) have the same range. For any $N \in \mathbb{Z}^+ \stackrel{\text{def}}{=} \{1, 2, \dots\}$, $[N] \stackrel{\text{def}}{=} \{1, 2, \dots, N\}$.

In *randomised* codes, at least one node i will include a random key Z_i in its encoding function. This key is independent of all messages and other random keys, and is not known to all other nodes.

A. Secure network coding

A network coding [9] instance $\mathbb{N} = (G, C, W)$ consists of the following: (i) a directed graph G with node set $\mathcal{V}$ and edge set $\mathcal{E}$, where each edge $e \in \mathcal{E}$ has a link capacity $c_e \in \mathbb{R}_0^+ \stackrel{\text{def}}{=} [0, \infty)$, meaning that node $\texttt{tail}(e)$ can send a message $x_e \in [2^{\lfloor c_e n \rfloor}]$ to node $\texttt{head}(e)$ without error in n link uses; (ii) a connection requirement $C = (\mathcal{S}, O, \mathcal{D})$, where the source messages $\{X_s : s \in \mathcal{S}\}$ are independent but can be arbitrarily distributed, $O(s) \in \mathcal{V}$ is the originating node of message X_s, and $\mathcal{D}(s) \subseteq \mathcal{V}$ is the set of destination nodes requesting X_s; (iii) $W = ((\mathcal{A}_r, \mathcal{B}_r) : r \in \mathcal{R})$ defines a set of eavesdroppers $\mathcal{R}$, where each eavesdropper $r \in \mathcal{R}$ observes the links $\mathcal{B}_r \subseteq \mathcal{E}$ and wants to obtain messages $\boldsymbol{X}_{\mathcal{A}_r}$, $\mathcal{A}_r \subseteq \mathcal{S}$.

A secure network-coding instance $\mathbb{N}$ is said to be $(\mathcal{S}^*, (p_{X_s} : s \in \mathcal{S}^*), \epsilon, \eta, n)$-feasible, for a subset $\mathcal{S}^* \subseteq \mathcal{S}$ of messages $\boldsymbol{X}_{\mathcal{S}^*}$ with a joint pmf $p_{\boldsymbol{X}_{\mathcal{S}^*}} = \prod_{s \in \mathcal{S}^*} p_{X_s}$, if and only if there exists a joint pmf $p_{\boldsymbol{X}_{\mathcal{S}\setminus\mathcal{S}^*}} = \prod_{s \in \mathcal{S}\setminus\mathcal{S}^*} p_{X_s}$ for the remaining messages $\boldsymbol{X}_{\mathcal{S}\setminus\mathcal{S}^*}$ and a secure network code,[1] which uses each link $n \in \mathbb{Z}^+$ times to satisfy a decoding-error criterion $P_{\text{e}} = \Pr\left\{\exists u \in \mathcal{U} \text{ s.t. } \boldsymbol{X}^{(u)}_{\{s \in \mathcal{S} : u \in \mathcal{D}(s)\}} \neq \boldsymbol{X}_{\{s \in \mathcal{S} : u \in \mathcal{D}(s)\}}\right\} \leq \epsilon$, and security criteria $I(\boldsymbol{X}_{\mathcal{A}_r}; \boldsymbol{X}_{\mathcal{B}_r}) \leq \eta$ for every eavesdropper $r \in \mathcal{R}$. Here, $\mathcal{U}$ is the set of destination nodes, and $\boldsymbol{X}^{(u)}_{\{s \in \mathcal{S} : u \in \mathcal{D}(s)\}}$ are the messages decoded by node u.

Note that the feasibility definition applies to a subset of messages $\mathcal{S}^*$. The reason will become clear in Section II-D when we map a randomised network code C_1 to a deterministic network code C_2, where the messages in C_2 consists of all messages and random keys in C_1. $\mathcal{S}^*$ for C_2 then corresponds to the original messages of interest in C_1.

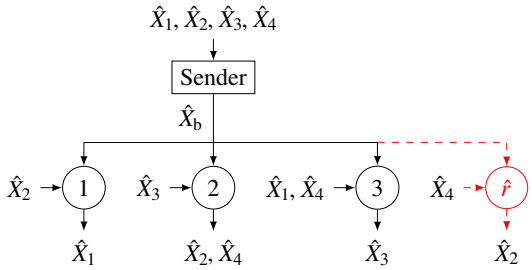
B. Secure index coding

A secure index-coding [6] instance $\mathbb{I} = (\hat{\mathcal{S}}, \hat{\mathcal{T}}, \{(\hat{\mathcal{W}}_t, \hat{\mathcal{H}}_t) : t \in \hat{\mathcal{T}}\}, \hat{W})$ consists of the following: (i) a sender having a set of messages $\{\hat{X}_s : s \in \hat{\mathcal{S}}\}$, which are mutually independent but arbitrarily distributed; (ii) a set of receivers $\hat{\mathcal{T}}$; (iii) receiver $t \in \hat{\mathcal{T}}$ requests messages $\hat{\boldsymbol{X}}_{\hat{\mathcal{W}}_t}$ and knows messages $\hat{\boldsymbol{X}}_{\hat{\mathcal{H}}_t}$ a priori; (iv) $\hat{W} = ((\hat{\mathcal{A}}_r, \hat{\mathcal{B}}_r) : r \in \hat{\mathcal{R}})$ defines a set of eavesdroppers $\hat{\mathcal{R}}$, where each eavesdropper $r \in \hat{\mathcal{R}}$ has messages $\hat{\boldsymbol{X}}_{\mathcal{B}_r}$ and wants to obtain messages $\hat{\boldsymbol{X}}_{\mathcal{A}_r}$. We assume that $\hat{\mathcal{B}}_r \cap \mathcal{A}_r = \emptyset$. Note $\hat{\mathcal{W}}_t, \hat{\mathcal{H}}_t, \hat{\mathcal{A}}_r, \hat{\mathcal{B}}_r \subseteq \hat{\mathcal{S}}$.

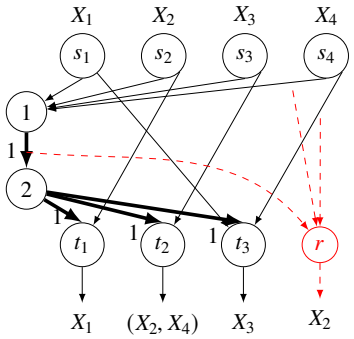
Similar to network coding, a secure index-coding instance $\mathbb{I}$ is said to be $(\hat{\mathcal{S}}^*, (p_{\hat{X}_s} : s \in \hat{\mathcal{S}}^*), \epsilon, \eta, n)$-feasible, for a subset $\hat{\mathcal{S}}^* \subseteq \hat{\mathcal{S}}$ of messages $\hat{\boldsymbol{X}}_{\hat{\mathcal{S}}^*}$ with a joint pmf $p_{\hat{\boldsymbol{X}}_{\hat{\mathcal{S}}^*}} = \prod_{s \in \hat{\mathcal{S}}^*} p_{\hat{X}_s}$, if and only if there exists a joint pmf $p_{\hat{\boldsymbol{X}}_{\hat{\mathcal{S}}\setminus\hat{\mathcal{S}}^*}} = \prod_{s \in \hat{\mathcal{S}}\setminus\hat{\mathcal{S}}^*} p_{\hat{X}_s}$ for the remaining messages $\hat{\boldsymbol{X}}_{\hat{\mathcal{S}}\setminus\hat{\mathcal{S}}^*}$ and a secure index code[2] satisfying a decoding-error criterion $\hat{P}_{\text{e}} = \Pr\{\exists t \in \hat{\mathcal{T}} \text{ s.t. } \hat{\boldsymbol{X}}^{(t)}_{\hat{\mathcal{W}}_t} \neq$

[1] A network code consists of a local encoding function $X_e = \mathsf{e}_e(\boldsymbol{X}_{\texttt{in}(\texttt{tail}(e))}, \boldsymbol{X}_{O^{-1}(\texttt{tail}(e))}) \in [2^{\lfloor c_e n \rfloor}]$ for each edge $e \in \mathcal{E}$ and a decoding function $\mathsf{d}_u(\boldsymbol{X}_{\texttt{in}(u)}, \boldsymbol{X}_{O^{-1}(u)})$ for each destination node $u \in \mathcal{U}$. Each e_e (or d_u) is a function of all incoming messages to and messages originating at $\texttt{tail}(e)$ (or u), respectively.

[2] An index code consists of an encoding function $\hat{X}_{\text{b}} = \hat{\mathsf{e}}(\hat{\boldsymbol{X}}_{\hat{\mathcal{S}}}) \in [2^n]$ at the sender and a decoding function $\hat{\mathsf{d}}_t(\hat{X}_{\text{b}}, \hat{\boldsymbol{X}}_{\hat{\mathcal{H}}_t})$ at each receiver $t \in \hat{\mathcal{T}}$.



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



(b) A secure network-coding instance $\mathbb{N}$, where an eavesdropper r has access to link $(1 \to 2)$, all outgoing links from node s_4, and tries to reconstruct X_2. The capacity of all links given by thick arrows is 1 bit per channel use

Fig. 1: A secure index-coding instance $\mathbb{I}$ and its corresponding secure network-coding instance $\mathbb{N}$ [1]

$\hat{\boldsymbol{X}}_{\hat{\mathcal{W}}_t}\} \le \epsilon$, and security criteria $I(\hat{\boldsymbol{X}}_{\hat{\mathcal{A}}_r}; \hat{X}_b, \hat{\boldsymbol{X}}_{\hat{\mathcal{B}}_r}) \le \eta$ for every eavesdropper $r \in \hat{\mathcal{R}}$. $\hat{\boldsymbol{X}}^{(t)}_{\hat{\mathcal{W}}_t}$ is the receiver t's decoded messages.

Remark 1: When each message X_s be uniformly distributed over $\mathcal{X}_s$, we define $R_s \stackrel{\text{def}}{=} (\log_2 |\mathcal{X}_s|)/n$ as the average message rate per network use. Then the above feasibility gives rise to the following notions of security criteria for fixed $\boldsymbol{R}_{\mathcal{S}}$:
(i) Strong security: $\lim_{n\to\infty} I(\boldsymbol{X}_{\mathcal{A}_r}; \boldsymbol{X}_{\mathcal{B}_r}) = 0, \forall r \in \mathcal{R}$;
(ii) Weak security: $\lim_{n\to\infty} \frac{1}{n} I(\boldsymbol{X}_{\mathcal{A}_r}; \boldsymbol{X}_{\mathcal{B}_r}) = 0, \forall r \in \mathcal{R}$.

Next, we briefly review the mapping between secure network coding and index coding instances, and the code translation between them. We refer the reader to Effros et al. [5] and Ong et al. [1] for a more detailed and descriptive account.

C. Mapping a secure index coding instance $\mathbb{I}$ to a secure network coding instance $\mathbb{N}$

Consider $\mathbb{I} = (\hat{\mathcal{S}}, \hat{\mathcal{T}}, \{(\hat{\mathcal{W}}_t, \hat{\mathcal{H}}_t) : t \in \hat{\mathcal{T}}\}, \hat{W})$. Let $\hat{\mathcal{S}} = [k]$ and $\hat{\mathcal{T}} = [\ell]$. We construct an equivalent $\mathbb{N} = (G, C, W)$ as follows: $G = (\mathcal{V}, \mathcal{E})$ consists of nodes $\mathcal{V} = \{s_1, s_2, \ldots, s_k, t_1, t_2, \ldots, t_\ell, 1, 2\}$. Each node s_i has an outgoing link with a sufficiently large capacity to node 1 and to each node in $\{t_j : i \in \hat{\mathcal{H}}_j\}$. Node 1 has a link of capacity 1 to node 2, and node 2 has an outgoing link of capacity 1 to every node t_i. For C, we set $\mathcal{S} = \hat{\mathcal{S}}$; X_i originates from $O(i) = s_i$ and is requested by $\mathcal{D}(i) = \{t_j : i \in \hat{\mathcal{W}}_j\}$ for all $i \in \mathcal{S}$. For W, $\mathcal{R} = \hat{\mathcal{R}}$, $\mathcal{B}_r = \{(1 \to 2), \{\texttt{out}(s_i) : i \in \hat{\mathcal{B}}_r\}\}$ and $\mathcal{A}_r = \hat{\mathcal{A}}_r$ for each $r \in \hat{\mathcal{R}}$. Figure 1 depicts an example of such a mapping.

Next, we describe the code translation in both directions:

1) Translating a code for $\mathbb{I}$ to a code for $\mathbb{N}$: Let $\hat{\mathsf{e}}(\hat{\boldsymbol{X}}_{\hat{\mathcal{S}}}, \hat{Z})$ be the sender's encoding function in $\mathbb{I}$, where $\hat{Z}$ is the random key in the randomised encoding function. We translate to the following code for $\mathbb{N}$: For every outgoing link $(s_i \to v)$ from every s_i, we set $X_{s_i \to v} = \mathsf{e}_{s_i \to v}(\cdot) = X_i$. For every outgoing link from nodes 1 and 2, we set $X_{1\to 2} = \hat{\mathsf{e}}(\boldsymbol{X}_{\mathcal{S}}, Z_1)$ and $X_e = X_{1\to 2}$ for all $e \in \texttt{out}(2)$, where $Z_1 \stackrel{\text{d}}{=} \hat{Z}$.

2) Translating a code for $\mathbb{N}$ to a code for $\mathbb{I}$: Suppose the global encoding function for edge $1 \to 2$ in $\mathbb{N}$ can be written as $\mathsf{g}_{1\to 2}(\boldsymbol{X}_{\mathcal{S}}, Z_1)$. We translate to the following code for $\mathbb{I}$: $\hat{X}_b = \hat{\mathsf{e}}(\cdot) = \mathsf{g}_{1\to 2}(\hat{\boldsymbol{X}}_{\hat{\mathcal{S}}}, \hat{Z})$, where $\hat{Z} \stackrel{\text{d}}{=} Z_1$.

D. Mapping a secure network coding instance $\mathbb{N}$ to a secure index coding instance $\mathbb{I}$

Consider $\mathbb{N} = (G, C, W)$, where without loss of generality, we assume that each message is requested by at least one destination. Let $\mathcal{S} = [S]$ and $\mathcal{V} = [V]$. We first map $\mathbb{N}$ to an *augmented* secure network-coding instance $\mathbb{N}' = (G', C', W')$ where $G' = (\mathcal{V}', \mathcal{E}') = G = (\mathcal{V}, \mathcal{E})$ with link capacities $c'_e = c_e$, and $W' = W$; thus $\mathcal{R}' = \mathcal{R}$, $\mathcal{B}'_r = \mathcal{B}_r$, and $\mathcal{A}'_r = \mathcal{A}_r$, for all $r \in \mathcal{R}$. Set $\mathcal{S}' = \mathcal{S} \cup \{S+1, S+2, \ldots, S+V\}$. For $s \in \mathcal{S}$, $O'(s) = O(s)$, and $\mathcal{D}'(s) = \mathcal{D}(s)$. For the newly added messages, $O'(S+v) = v$ and $\mathcal{D}'(S+v) = \emptyset$, for all $v \in [V]$. This means each X'_{S+v} originates at node v, and is not requested by any node. It takes the role of the random key Z_v in the randomised encoding at node v in $\mathbb{N}$. So for any vertex $v \in [V]$ that has no outgoing edge, we set $X'_{S+v} = \alpha$ to be a constant. With this, any (deterministic or randomised) code for $\mathbb{N}$ can be mapped to a deterministic code for $\mathbb{N}'$.

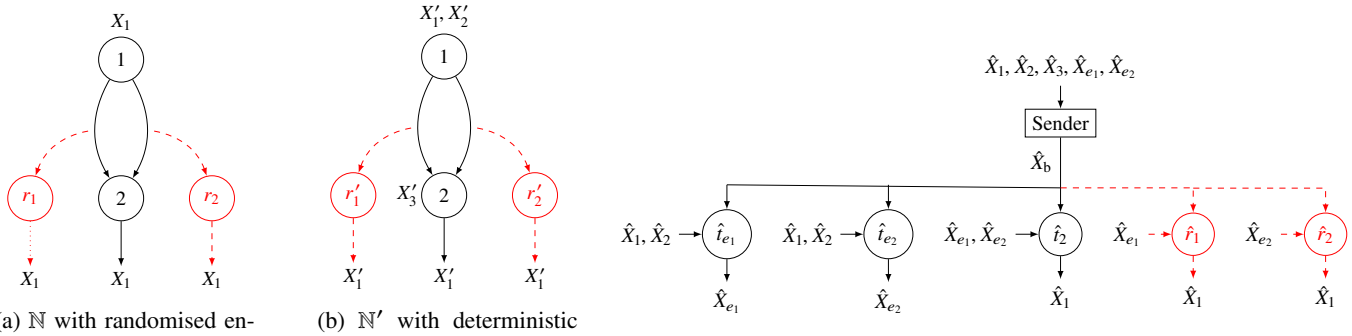
Then, we map $\mathbb{N}'$ to $\mathbb{I}$ as follows: $\hat{\mathcal{S}} = \mathcal{S}' \cup \mathcal{E}'$. $\hat{\mathcal{T}} = \{\hat{t}_i\}_{i\in\mathcal{U}'} \cup \{\hat{t}_e\}_{e\in\mathcal{E}'}$, where $\mathcal{U}' = \mathcal{U}$ is the set of destination nodes in $\mathbb{N}'$. For each $\hat{t}_e \in \hat{\mathcal{T}}$ where $e \in \mathcal{E}'$, we set $\hat{\mathcal{H}}_{\hat{t}_e} = \texttt{in}(\texttt{tail}(e)) \cup O'^{-1}(\texttt{tail}(e))$, and $\hat{\mathcal{W}}_{\hat{t}_e} = \{e\}$. For each $\hat{t}_i \in \hat{\mathcal{T}}$ where $i \in \mathcal{U}'$, we set $\hat{\mathcal{H}}_{\hat{t}_i} = \texttt{in}(i) \cup O'^{-1}(i)$, and $\hat{\mathcal{W}}_{\hat{t}_i} = \{s \in [S] : i \in \mathcal{D}'(s)\}$. $\hat{\mathcal{R}} = \mathcal{R}'$. For each $\hat{r} \in \hat{\mathcal{R}}$, $\hat{\mathcal{B}}_{\hat{r}} = \mathcal{B}'_{\hat{r}}$, and $\hat{\mathcal{A}}_{\hat{r}} = \mathcal{A}'_{\hat{r}}$. Figure 2 depicts an example of such a mapping.

1) Translating a code for $\mathbb{N}$ to a code for $\mathbb{I}$: Let $\{\mathsf{g}'_e : e \in \mathcal{E}\}$ be deterministic global encoding functions in $\mathbb{N}'$. For $\mathbb{I}$, we set $\hat{X}_b = [\hat{X}_{b,e} : e \in \mathcal{E}']$, where $\hat{X}_{b,e} = \hat{X}_e + \mathsf{g}'_e(\hat{\boldsymbol{X}}_{\mathcal{S}'}) \bmod 2^{\lfloor c_e n \rfloor}$, and $\hat{X}_e$ is independently and uniformly distributed over $[2^{\lfloor c'_e n \rfloor}]$.

2) Translating a code for $\mathbb{I}$ to a code for $\mathbb{N}$: Let $\{\hat{\mathsf{d}}_v(\hat{X}_b, \hat{\boldsymbol{X}}_{\hat{\mathcal{H}}_v}) : v \in \hat{\mathcal{T}}\}$ be the decoding functions in $\mathbb{I}$. For $\mathbb{N}'$, we first choose some broadcast message σ, and then set $\mathsf{e}_e(\cdot) = \hat{\mathsf{d}}_{\hat{t}_e}(\sigma, \boldsymbol{X}'_{\texttt{in}(\texttt{tail}(e))\cup O'^{-1}(\texttt{tail}(e))})$ for each edge $e \in \mathcal{E}'$.

III. Results

The main results of this paper are as follows:



(a) $\mathbb{N}$ with randomised encoding

(b) $\mathbb{N}'$ with deterministic encoding

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

Fig. 2: A secure network-coding instance $\mathbb{N}$, its augmented version $\mathbb{N}'$ (with additional messages X_2' and X_3'), and the corresponding secure index-coding instance $\mathbb{I}$ [1]

Theorem 1: Let $\mathbb{I}$ be a secure index-coding instance, and $\mathbb{N}$ be the corresponding secure network-coding instance. For any $\epsilon, \eta \in \mathbb{R}_0^+$, $n \in \mathbb{Z}^+$, the instance $\mathbb{I}$ is $(\hat{\mathcal{S}}, (p_{\hat{X}_s} : s \in \hat{\mathcal{S}}), \epsilon, \eta, n)$-feasible if and only if $\mathbb{N}$ is $(\mathcal{S}, (p_{X_s} : s \in \mathcal{S}), \epsilon, \eta, n)$-feasible with deterministic coding functions for nodes $\{s_i : i \in \hat{\mathcal{S}}\}$, where $\hat{\boldsymbol{X}}_{\hat{\mathcal{S}}} \overset{\mathrm{d}}{=} \boldsymbol{X}_{\mathcal{S}}$.

The theorem above preserves the message size, as well as the decodability and security criteria from a secure index-coding instance to a secure network-coding instance.

For the other direction, we first define $\hat{n} = \sum_{e \in \mathcal{E}} \lfloor c_e n \rfloor$. Let the total variation distance between two pmfs p and q on Ω be expressed in L^1 norms as $\delta(p, q) = \frac{1}{2}\|p - q\|_1 = \frac{1}{2}\sum_{\sigma \in \Omega} |p(\sigma) - q(\sigma)|$. Also, denote the uniform distribution on a finite set Ω by $\mathtt{unif}(\Omega)$.

Theorem 2: Let $\mathbb{N}$ be a secure network-coding instance and $\mathbb{I}$ be the corresponding secure index-coding instance. For any $\eta \in \mathbb{R}_0^+$, $\epsilon \in [0, 0.5]$, $n \in \mathbb{Z}^+$, we have the following:

1) If $\mathbb{N}$, in which all messages $\boldsymbol{X}_{\mathcal{S}}$ are independent and uniformly distributed, is $(\mathcal{S}, (p_{X_s} : s \in \mathcal{S}), \epsilon, \eta, n)$-feasible, then $\mathbb{I}$ is $(\mathcal{S}, (p_{\hat{X}_s} : s \in \mathcal{S}), \epsilon, \eta, \hat{n})$-feasible with a deterministic code, where $\hat{\boldsymbol{X}}_{\mathcal{S}} \overset{\mathrm{d}}{=} \boldsymbol{X}_{\mathcal{S}}$.
2) If $\mathbb{I}$, in which all messages $(\hat{\boldsymbol{X}}_{\mathcal{S}}, \hat{\boldsymbol{X}}_{\mathcal{E}})$ are independent and uniformly distributed, is $(\mathcal{S}, (p_{\hat{X}_s} : s \in \mathcal{S}), \epsilon, \eta, \hat{n})$-feasible with a deterministic code, then
   a) For $\epsilon = 0$, $\mathbb{N}$ is $(\mathcal{S}, (p_{X_s} : s \in \mathcal{S}), \epsilon, \eta, n)$-feasible; and
   b) Otherwise, $\epsilon \in (0, 0.5]$, $\mathbb{N}$ is $(\mathcal{S}, (p_{X_s} : s \in \mathcal{S}), |\mathcal{R}|\eta + \zeta, \gamma, n)$-feasible,

   where $\boldsymbol{X}_{\mathcal{S}} \overset{\mathrm{d}}{=} \hat{\boldsymbol{X}}_{\mathcal{S}}$, ζ is a function of (ϵ, n), and γ is a function of $(\zeta, \epsilon, \eta, n)$, defined as follows:

$$\gamma \overset{\mathrm{def}}{=} \min\left\{(|\mathcal{R}|\eta + \zeta)\left(\frac{1}{1-\epsilon} + \frac{\log e + \hat{n}}{1 - (|\mathcal{R}|\eta + \zeta)} + \log|\mathcal{X}_{\mathcal{S}'}|\right) + \frac{1}{1-\epsilon}|\mathcal{R}|H_{\mathrm{b}}(\epsilon) - \log\left(1 - (|\mathcal{R}|\eta + \zeta)\right),\ \hat{n}\right\},$$

$$\zeta \overset{\mathrm{def}}{=} \min\left\{\epsilon[1 + 2\delta(p_{\hat{X}_{\mathrm{b}}}, \mathtt{unif}([2^n]))],\ \epsilon[1 + \epsilon 2^{\hat{n}}],\ 1\right\}.$$

In Part 2b of Theorem 2, the upper bounds on decoding error and leakage increase exponentially with n. We can tighten the bounds for linear codes:

Corollary 2.1: Let $\mathbb{N}$ be a secure network-coding instance and $\mathbb{I}$ be the corresponding secure index-coding instance. For any $\eta \in \mathbb{R}_0^+$, $\epsilon \in (0, 0.5]$, $n \in \mathbb{Z}^+$, we have the following: If $\mathbb{I}$ is $(\mathcal{S}, (p_{\hat{X}_s} : s \in \mathcal{S}), \epsilon, \eta, \hat{n})$-feasible using a linear deterministic index code with cardinality $2^{\hat{n}}$, where $(\hat{\boldsymbol{X}}_{\mathcal{S}}, \hat{\boldsymbol{X}}_{\mathcal{E}})$ are independent and uniformly distributed, then $\mathbb{N}$ is $(\mathcal{S}, (p_{X_s} : s \in \mathcal{S}), |\mathcal{R}|\eta + \epsilon, \gamma', n)$-feasible, where

$$\gamma' \overset{\mathrm{def}}{=} \min\left\{(|\mathcal{R}|\eta + \epsilon)\left(\frac{1}{1-\epsilon} + \frac{\log e + \hat{n}}{1 - (|\mathcal{R}|\eta + \epsilon)} + \log|\mathcal{X}_{\mathcal{S}'}|\right) + \frac{1}{1-\epsilon}|\mathcal{R}|H_{\mathrm{b}}(\epsilon) - \log\left(1 - (|\mathcal{R}|\eta + \epsilon)\right),\ \hat{n}\right\}.$$

Note here that, for linear codes, the error probability for $\mathbb{N}$ is independent of n, and is solely a function of ϵ, η, and the number of eavesdroppers $|\mathcal{R}|$; the leakage for $\mathbb{N}$ is a linear function of n, and the coefficient of n can be made arbitrarily small by choosing arbitrarily small η and ϵ. This means a sequence of strongly-secure index codes for $\mathbb{I}$ translates to a sequence of weakly-secure network codes for $\mathbb{N}$ (with appropriate rate scaling).

Proof of Corollary 2.1: Using linear codes for $\mathbb{I}$, if the messages are uniformly distributed, then the codeword $\hat{X}_{\mathrm{b}}$ is uniformly distributed over its support. So, $\delta(p_{\hat{X}_{\mathrm{b}}}, \mathtt{unif}([2^n])) = 0$, which implies $\zeta = \epsilon$, and Corollary 2.1 follows directly from Part 2b of Theorem 2. ■

IV. Sketch of Proofs for Theorems 1 and 2

Due to space constraints, we present here the sketch of proofs for Theorems 1 and 2, and refer the reader to the longer version of this paper [10] for complete proofs.

A. For Theorem 1 – the forward direction

The proof for the forward direction, that is $\mathbb{I}$ is $(\hat{\mathcal{S}}, (p_{\hat{X}_s} : s \in \hat{\mathcal{S}}), \epsilon, \eta, n)$-feasible $\Rightarrow$ $\mathbb{N}$ is $(\mathcal{S}, (p_{X_s} : s \in \mathcal{S}), \epsilon, \eta, n)$-feasible, is rather straightforward and is omitted here.

B. For Theorem 1 – the backward direction

For the other direction, the decodability criteria is also straightforward. Here, we prove the security criteria. Recall that we have chosen $(\hat{\boldsymbol{X}}_{\hat{\mathcal{S}}}, \hat{Z}) \overset{\mathrm{d}}{=} (\boldsymbol{X}_{\mathcal{S}}, Z_1)$.

From the hypothesis, we have the security condition $I(\boldsymbol{X}_{\mathcal{A}_r}; \boldsymbol{X}_{\mathcal{B}_r}) < \eta$ for $\mathbb{N}$. Showing that the index code also satisfy a similar security condition is not trivial, as the eavesdroppers in $\mathbb{I}$ can access the messages themselves, instead of just functions of the messages as in $\mathbb{N}$. These functions may not necessarily allow one to recover the messages, as we allow non-zero error decoding probability. So, it seems that the eavesdroppers in $\mathbb{I}$ have "better" observations, which may lead to a larger leakage in the code.

We will show that this is not the case. First, note the following: (i) $\{\boldsymbol{X}_{\mathcal{S}}, Z_1\}$ are mutually independent; (ii) $\boldsymbol{X}_{\mathtt{out}(s_i)}$, for each $i \in \mathcal{S}$, is a deterministic function of X_i; (iii) $\hat{\mathcal{B}}_r \cap \mathcal{A}_r = \emptyset$. With these, we have the following Markov chain for every r, $\boldsymbol{X}_{\hat{\mathcal{B}}_r} - \boldsymbol{X}_{\{\mathtt{out}(s_i): i \in \hat{\mathcal{B}}_r\}} - (Z_1, \boldsymbol{X}_{\mathcal{A}_r}, \boldsymbol{X}_{\mathcal{S}\setminus(\mathcal{A}_r \cup \hat{\mathcal{B}}_r)})$, from which we can then establish $I(\boldsymbol{X}_{\hat{\mathcal{B}}_r}; \boldsymbol{X}_{\mathcal{A}_r} | \boldsymbol{X}_{\mathcal{B}_r}) = 0$. This means that eavesdropper r, having observed the links $\boldsymbol{X}_{\mathcal{B}_r}$, does not gain any more information about $\boldsymbol{X}_{\mathcal{A}_r}$ even if it can further observe the sources messages $\boldsymbol{X}_{\hat{\mathcal{B}}_r}$. With this, we get

$$\begin{aligned} &I(\boldsymbol{X}_{\hat{\mathcal{B}}_r}, X_{1\to 2}; \boldsymbol{X}_{\mathcal{A}_r}) = I(\boldsymbol{X}_{\hat{\mathcal{B}}_r}, X_{1\to 2}, \boldsymbol{X}_{\{\mathtt{out}(s_i): i \in \hat{\mathcal{B}}_r\}}; \boldsymbol{X}_{\mathcal{A}_r}) \\ &= I(\boldsymbol{X}_{\hat{\mathcal{B}}_r}, \boldsymbol{X}_{\mathcal{B}_r}; \boldsymbol{X}_{\mathcal{A}_r}) = I(\boldsymbol{X}_{\mathcal{B}_r}; \boldsymbol{X}_{\mathcal{A}_r}) + I(\boldsymbol{X}_{\hat{\mathcal{B}}_r}; \boldsymbol{X}_{\mathcal{A}_r} | \boldsymbol{X}_{\mathcal{B}_r}) \\ &= I(\boldsymbol{X}_{\mathcal{B}_r}; \boldsymbol{X}_{\mathcal{A}_r}) \le \eta. \end{aligned}$$

Since we set $\hat{X}_{\mathrm{b}} = \mathsf{g}_{1\to 2}(\hat{\boldsymbol{X}}_{\hat{\mathcal{S}}}, \hat{Z})$, we have $(\hat{\boldsymbol{X}}_{\hat{\mathcal{S}}}, \hat{Z}, \hat{X}_{\mathrm{b}}) \overset{\mathrm{d}}{=} (\boldsymbol{X}_{\mathcal{S}}, Z_1, X_{1\to 2})$. Recalling that $\mathcal{A}_r = \hat{\mathcal{A}}_r$ and $\mathcal{B}_r = \{(1 \to 2), \{\mathtt{out}(s_i) : i \in \hat{\mathcal{B}}_r\}\}$, we have $I(\hat{\boldsymbol{X}}_{\hat{\mathcal{B}}_r}, \hat{X}_{\mathrm{b}}; \hat{\boldsymbol{X}}_{\hat{\mathcal{A}}_r}) \le \eta$, which is the required security criteria for $\mathbb{I}$.

C. For Theorem 2 – Part 1 (the forward direction)

Due to space constraints, we omit the proof of this part. The proof mainly relies on the observations that $\{\hat{X}_{\mathrm{b}}(e) : e \notin \hat{\mathcal{B}}_{\hat{r}}\}$ are independent of $(\hat{\boldsymbol{X}}_{\hat{\mathcal{A}}_{\hat{r}}}, \{\hat{X}_{\mathrm{b},e} : e \in \hat{\mathcal{B}}_{\hat{r}}\}, \{\hat{X}_{e'} : e' \in \hat{\mathcal{B}}_{\hat{r}}\})$ and that $\hat{X}_{\mathrm{b},e}$ is a deterministic function of $(\hat{X}_e, \mathsf{g}'_e(\hat{\boldsymbol{X}}_{\mathcal{S}'}))$.

D. For Theorem 2 – Part 2 (the backward direction)

Recall that for this direction, we need to choose a fixed σ for all the encoding functions in $\mathbb{N}'$.

When $\epsilon = 0$ (for Part 2a), any σ for $\mathbb{N}'$ must guarantee perfect decoding for $\mathbb{N}'$, and we only need to select a good σ to guarantee the security criterion. We can show that such a candidate exists.

When $\epsilon > 0$ (for Part 2b), we need to choose a good σ that simultaneously guarantees the decodability and the security criteria. This requires us to express both the security and decodability expressions for $\mathbb{I}$ in terms of $\hat{X}_{\mathrm{b}}$. Then we need to relate the random variables in the two instances $\mathbb{I}$ and $\mathbb{N}'$ through ϵ and η. Our solution consists of the following steps:

S.1 We relate security expressions for $\mathbb{N}'$ (in which the messages are denoted by X'_i, $i \in \mathcal{S}' \cup \mathcal{E}'$) to that for $\mathbb{I}$. We show that for any $\sigma \in [2^{\hat{n}}]$,

$$\begin{aligned} I(\boldsymbol{X}'_{\mathcal{A}'_r}; \boldsymbol{X}'_{\mathcal{B}'_r}) \le\ & I(\hat{\boldsymbol{X}}_{\hat{\mathcal{A}}_r}; \hat{\boldsymbol{X}}_{\hat{\mathcal{B}}_r} | \hat{D} = 1, \hat{X}_{\mathrm{b}} = \sigma) \\ &+ \epsilon' \log |\mathcal{X}_{\mathcal{S}'}| - \log(1-\epsilon') + \frac{\epsilon'}{1-\epsilon'} (\log e + \hat{n})), \quad (2) \end{aligned}$$

where $\hat{D} = 1$ denotes the event of correct decoding in $\mathbb{I}$, and $\mathcal{X}_{\mathcal{S}'} = \prod_{s \in \mathcal{S}'} \mathcal{X}_s$ is the set of all message realisations.

S.2 We express security in $\mathbb{I}$ in terms of the expression obtained in S.1 (in the RHS of (2)) averaged over $\hat{X}_{\mathrm{b}}$:

$$\begin{aligned} |\mathcal{R}|\eta \ge & \sum_{\sigma \in [2^{\hat{n}}]} p_{\hat{X}_{\mathrm{b}}}(\sigma) \sum_{r \in \mathcal{R}} \Big[I(\hat{\boldsymbol{X}}_{\hat{\mathcal{A}}_r}; \hat{D} | \hat{X}_{\mathrm{b}} = \sigma) \\ &- I(\hat{\boldsymbol{X}}_{\hat{\mathcal{A}}_r}; \hat{D} | \hat{\boldsymbol{X}}_{\hat{\mathcal{B}}_r}, \hat{X}_{\mathrm{b}} = \sigma) \\ &+ \sum_{i \in \{0,1\}} p_{\hat{D}}(i) I(\hat{\boldsymbol{X}}_{\hat{\mathcal{A}}_r}; \hat{\boldsymbol{X}}_{\hat{\mathcal{B}}_r} | \hat{D} = i, \hat{X}_{\mathrm{b}} = \sigma) \\ \overset{\text{def}}{=} & \sum_{\sigma \in [2^{\hat{n}}]} p_{\hat{X}_{\mathrm{b}}}(\sigma) \Gamma(\sigma). \end{aligned} \quad (3)$$

S.3 We relate the decoding criterion in $\mathbb{N}'$ to that in $\mathbb{I}$: We derive $\epsilon \ge \sum_{\sigma \in [2^{\hat{n}}]} \mathtt{unif}([2^{\hat{n}}]) \frac{|\mathcal{G}^{\mathrm{c}}_\sigma|}{|\mathcal{X}_{\mathcal{S}'}|}$ and $P_{\mathrm{e},\sigma} \le \frac{|\mathcal{G}^{\mathrm{c}}_\sigma|}{|\mathcal{X}_{\mathcal{S}'}|}$, where $P_{\mathrm{e},\sigma}$ is the probability of decoding error in $\mathbb{N}'$ when σ is chosen, $\mathcal{G}^{\mathrm{c}}_\sigma \overset{\text{def}}{=} \mathcal{X}_{\mathcal{S}'} \setminus \mathcal{G}_\sigma$, and $\mathcal{G}_\sigma$ is the set of all message realisations $\mathcal{X}_{\mathcal{S}'}$ in $\mathbb{I}$ that result in both correct decoding and the broadcast message $\hat{x}_{\mathrm{b}} = \sigma$.

S.4 Using S.3, we express decodability in $\mathbb{I}$ as an average over $\hat{X}_{\mathrm{b}}$:

$$\begin{aligned} \sum_{\sigma \in [2^{\hat{n}}]} p_{\hat{X}_{\mathrm{b}}}(\sigma) \frac{|\mathcal{G}^{\mathrm{c}}_\sigma|}{|\mathcal{X}_{\mathcal{S}'}|} \le \min \Big\{ & \epsilon[1 + 2\delta(p_{\hat{X}_{\mathrm{b}}}, \mathtt{unif}([2^n]))], \\ & \epsilon[1 + \epsilon 2^{\hat{n}}],\ 1 \Big\} \overset{\text{def}}{=} \zeta. \end{aligned} \quad (4)$$

S.5 We combine the results from S.2 and S.4 to find a σ that is simultaneously good for security and decodability. Combining (4) and (3), there exists a σ such that

$$|\mathcal{R}|\eta + \zeta \ge \frac{|\mathcal{G}^{\mathrm{c}}_\sigma|}{|\mathcal{X}_{\mathcal{S}'}|} + \Gamma(\sigma), \quad (5)$$

which, when combined with (2), leads to the required decodability constraint $P_{\mathrm{e},\sigma} \le \frac{|\mathcal{G}^{\mathrm{c}}_\sigma|}{|\mathcal{X}_{\mathcal{S}'}|} \le |\mathcal{R}|\eta + \zeta$ and security constraint $I(\boldsymbol{X}'_{\mathcal{A}'_{r'}}; \boldsymbol{X}'_{\mathcal{B}'_{r'}}) \le \gamma$.

References

[1] L. Ong, B. N. Vellambi, J. Kliewer, and P. L. Yeoh, "An equivalence between secure network and index coding," in *Proc. IEEE Globecom – NetCod*, Washington, USA, Dec. 4 2016.

[2] 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.

[3] 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.

[4] 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.

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

[6] 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.

[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] A. El Gamal and Y. Kim, *Network Information Theory*, 1st ed. Cambridge University Press, 2011.

[9] T. Chan and A. Grant, "Capacity bounds for secure network coding," in *Proc. Australian Commun. Theory Workshop (AusCTW)*, Christchurch, New Zealand, Jan. 30–Feb. 1 2008, pp. 95–100.

[10] L. Ong, J. Kliewer, B. N. Vellambi, and P. L. Yeoh. (2018, Apr. 26) A code equivalence between secure network and index coding. [Online]. Available: https://arxiv.org/pdf/1804.09888v1.pdf