A Code and Rate Equivalence between Secure Network and Index Coding

Lawrence Ong, Badri N. Vellambi, Jörg Kliewer, and Phee Lep Yeoh

Abstract

Establishing code equivalences between index coding and network coding provides important insights for code design. Previous works showed an equivalence relation between any index-coding instance and a network-coding instance, for which a code for one instance can be translated to a code for the other instance with the same decoding-error performance. The equivalence also showed a surprising result that any network-coding instance can be mapped to an index-coding instance with a properly designed code translation. In this paper, we extend the existing equivalence (instance map and code translation) to one between secure index coding and secure network coding, where eavesdroppers are present in the network. In the secure setting, any code construction needs to guarantee security constraints in addition to decoding-error performance. A rate equivalence between these two problems is also established.

I. Introduction

Equivalence results in information theory and network coding are of significant interest because such reduction results uniquely map one communication problem to another equivalent problem that is potentially easier to study. Some equivalence results already established include those between instances of multiple-unicast network coding and those of (i) multiple-multicast network coding [1], (ii) secure network coding [2], and (iii) index coding [3, 4]. This paper focuses on the equivalence between index coding and network coding.

Prima facie, the two problems of index coding and network coding appear different. Index coding [5] 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

Lawrence Ong (lawrence.ong@newcastle.edu.au) is with The University of Newcastle. Badri N. Vellambi (badri.vellambi@uc.edu) is the University of Cincinnati. Jörg Kliewer (jkliewer@njit.edu) is with the New Jersey Institute of Technology. Phee Lep Yeoh (phee.yeoh@sydney.edu.au) is with The University of Sydney.

This work is supported by ARC Discovery Scheme DP190100770, and US NSF grants CCF-1908756, CNS-1815322, and CNS-1526547.

Part of this paper was presented at IEEE Globecom in 2016 and at IEEE International Symposium on Information Theory in 2018 and 2020.

knows some other messages. On the other hand, network coding [6] considers a network of interconnected links with fixed capacities, where multiple senders send multiple messages to multiple receivers through these links.

Despite the differences, an equivalence between them has been demonstrated [3, 4] in the following sense. For any index-coding instance $\mathbb{I}$, an *instance map* constructs an equivalent network-coding instance $\mathbb{N}$. This instance pair $(\mathbb{I}, \mathbb{N})$ has the following properties: Any index code for $\mathbb{I}$ can be translated to a network code for $\mathbb{N}$, and vice versa. This *code translation* preserves the message length, code length, and probability of decoding error. Similarly, for any network-coding instance $\mathbb{N}'$, an instance map constructs an equivalent index-coding instance $\mathbb{I}'$, and a code translation can translate codes between the pair $(\mathbb{N}', \mathbb{I}')$.

In this paper, we investigate the equivalence when we impose *security constraints* in addition to decodability constraints (that is, the probability of decoding error). Separately, the secure version of index coding [7] and that of network coding [8] have been studied, in which additional passive entities, called eavesdroppers, are present, and attempt to obtain some information on the messages being communicated. Codes for the secure version of these problem must prevent eavesdroppers from knowing the messages they attempt to decode (where knowing is quantified by the information-theoretic security measure [9, Ch 22]) in addition to guaranteeing that all receivers can obtain their requested messages (by bounding the probability of decoding error).

The *non-secure*[*] instance maps and code translations [3, 4] do not trivially apply to the secure version of the problems. In particular, we pointed out [10] that mapping an eavesdropper in secure network coding to secure index coding is not straightforward, as the eavesdroppers in the two problems have different characteristics. Eavesdroppers in network coding listen to transmission on certain links, while those in index coding listen to the common broadcast and have access to some subset of messages.

Also, the non-secure code translation was designed for deterministic codes. But randomised encoding is inevitable in some secure network-coding instances [8], and we have shown that the non-secure code translation breaks down for randomised encoding [10].

In this paper, we establish an equivalence between secure network coding and secure index coding. Similar to the non-secure equivalence, we construct instance maps for two directions (from secure index coding to secure network coding, and vice versa), and for each instance map, we construct two code translations (from an index code to a network code, and back).

[*]In this paper, we use the term "non-secure" to denote existing instance maps and code translations in the absence of security constraints.

This equivalence carries the practical significance of comparing secure communication against eavesdropping in wired networks with that in wireless networks. Our equivalence results reveal that a passive eavesdropper that listens to the common wireless broadcast is not advantageous—that is, more difficult to deal with—compared to a passive eavesdropper in a wired networks that taps only certain wired links. In fact, our code translation results guarantee that the same approach can be implemented in both networks with help of side information (receivers knowing some other messages a priori) in the wireless case.

A. *Contributions and Approaches*

Our approach to establish an equivalence is summarised as follows:

- First, we construct the maps for translating problem instances. We build on the existing non-secure instance maps that map legitimate receivers. This involves mapping the eavesdroppers in any secure index-coding instance (having certain messages) to those in the corresponding secure network-coding instance (listening to certain links), and vice versa.
- Second, we construct code translations for the problem-instance pairs. Again, we build on the existing non-secure code translations, which have been shown to preserve the decoding criteria. As mentioned earlier, the non-secure code translations were designed for deterministic codes, but randomised network codes are necessary for secure network coding [8]. To deal with this issue, we (i) construct a two-step code translation to convert randomised codes to deterministic codes, and (ii) restrict the randomised encoding functions to certain nodes in the mapped problem instance.
- Third, we show that although eavesdroppers in the two problems instances observe different types of messages, the code translations output codes with comparable message size, codelength, probability of decoding error, and information leakage to the eavesdroppers.
- Lastly, using the code translations, we show a rate equivalence between the two problems, that is, if a rate tuple is achievable[†] for index coding, an appropriately scaled rate tuple is also achievable for the mapped network coding, and vice versa.

II. Problem Definition and Notation

For a strictly ordered set $S = \{s_1, s_2, \dots s_{|S|}\}$, with a total order $<$ where $s_1 < s_2 \cdots < s_{|S|}$, let $X_S := (X_s)_{s\in S} = (X_{s_1}, X_{s_2}, \dots, X_{s_{|S|}})$. Consider a directed graph $G = (V, E)$ with a node set V and an link

[†]Achievability is defined in the Shannon sense, that is, both probability of decoding error and information leakage diminish, as the codelength increases.

set E. For a link $e = (u \to v) \in E$, where $u, v \in V$, its tail is $\texttt{tail}(e) := u$, and its head is $\texttt{head}(e) := v$. For any node $v \in V$, the set of incoming links is denoted by $\texttt{in}(v) := \{e \in E : \texttt{head}(e) = v\}$, and the set of outgoing links by $\texttt{out}(v) := \{e \in E : \texttt{tail}(e) = v\}$. For any $a \in \mathbb{Z}^+ := \{1, 2, \dots\}$, denote $[a] := \{1, 2, \dots, a\}$. $\mathbb{R}^+ := (0, \infty)$ and $\mathbb{R}_0^+ := [0, \infty)$.

A. Secure Network Coding

1) Problem instance: Denote a secure network-coding instance [11] by $\mathbb{N} = (G, C, P)$.

- $G = (V, E)$ is an acyclic directed graph with a node set V and a link set E. Each link $e \in E$ has capacity $c_e \in \mathbb{R}^+$, where $\texttt{tail}(e)$ can send a message $X_e \in [2^{\lfloor c_e n \rfloor}]$ to $\texttt{head}(e)$ noiselessly over $n \in \mathbb{Z}^+$ link uses.
- $C = (S, O, D)$ is the connection requirement. S contains the message indices, where the messages are $\{X_s : s \in S\}$. $O(s) \in V$ is the originating node[‡] for message X_s. $D(s) \subsetneq V$ is the set of nodes that require message X_s.
- $P = \big((A_z, B_z) : z \in Z\big)$ defines the eavesdroppers indexed by elements of Z. Each eavesdropper $z \in Z$ observes messages X_{B_z} on links $B_z \subseteq E$ and tries to reconstruct messages X_{A_z} for some $A_z \subseteq S$.

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. Otherwise, they can be deleted without any consequence. Similarly, each message is requested by at least one node.

2) Deterministic codes: Consider n uses of each link. Let the messages $\{X_s : s \in S\}$ be mutually independent, and let each X_s be uniformly distributed over $[M_s]$ for some $M_s \in \mathbb{Z}^+$. A deterministic (M_S, n)-network code consists of the following:

- A deterministic encoding function e_e for each link $e \in E$ that takes in encoded messages that are conveyed on incoming links of $\texttt{tail}(e)$ and those messages originating in $\texttt{tail}(e)$, i.e., $(X_{\texttt{in}(\texttt{tail}(e))}, X_{O^{-1}(\texttt{tail}(e))})$, and outputs message $X_e := \mathsf{e}_e(X_{\texttt{in}(\texttt{tail}(e))}, X_{O^{-1}(\texttt{tail}(e))})$ on link e taking values in $[2^{\lfloor c_e n \rfloor}]$. Here, we let $O^{-1}(v)$ to be the indices of the messages originating from node v.
- A decoding function d_v for each node $v \in V$ that takes in $(X_{\texttt{in}(v)}, X_{O^{-1}(v)})$ and outputs an estimate $X^{(v)} = \mathsf{d}_v(X_{\texttt{in}(v)}, X_{O^{-1}(v)})$ of the messages $X_{D^{-1}(v)}$ that v requires. Here $D^{-1}(v)$ is the set of indices of messages whose destinations include v.

[‡]Without loss of generality, each message is available precisely at one node.

Here, n is referred to as the blocklength of the code. It is the number of times each link is used. We assume that $c_e n \geq 1$ for every link e, such that we can transmit at least one bit.

3) Randomised codes: In this paper, we consider randomised network codes with the use of random keys. Each node v generates an independent random key Y_v that is uniformly distributed over $[K_v]$ for some $K_v \in \mathbb{Z}^+$. The key Y_v is known only to the generating node v.

A randomised network code is similar to a deterministic network code, except that each link encoding function e_e is a deterministic function that maps $(X_{\mathtt{in}(\mathtt{tail}(e))}, X_{O^{-1}(\mathtt{tail}(e))}, Y_{\mathtt{tail}(e)})$ to X_e.

Since the graph G is acyclic, any encoding function $\mathsf{e}_e(X_{\mathtt{in}(\mathtt{tail}(e))}, X_{O^{-1}(\mathtt{tail}(e))}, Y_{\mathtt{tail}(e)}))$ can be replaced by a suitable *global* encoding function $\mathsf{g}_e(X_S, Y_{\mathtt{tail}(e)})$ if all upstream links of e have deterministic encoding functions.

4) Decodability: A network code has a probability of decoding error of at most $\epsilon \in \mathbb{R}_0^+$ iff

$$P_{\mathrm{e}} := 1 - \Pr\left\{X^{(v)} = X_{D^{-1}(v)} \text{ for all } v \in V\right\} \leq \epsilon, \tag{1}$$

where $X^{(v)}$ is the set of all decoded messages whose destination is v (see Sec. II-A2). Note that when $\epsilon = 0$, the code guarantees perfect decoding.

5) Leakage: A network code has a leakage of at most $\eta \in \mathbb{R}_0^+$ iff

$$\frac{1}{n} I(X_{A_z}; X_{B_z}) \leq \eta, \qquad \text{for all } z \in Z. \tag{2}$$

With the normalisation factor of $\frac{1}{n}$, (2) is commonly referred to as weak security in the literature. When $\eta = 0$, we say that the code is perfectly secure.

6) Feasibility: A secure network-coding instance $\mathbb{N}$ is said to be (M_S, ϵ, η, n)-feasible iff there exists an (M_S, n)-network code that has a probability of decoding error of at most ϵ and a leakage of at most η.

Further, a message rate tuple of an (M_S, n)-network code is $R_S := \left(\frac{\log_2 M_s}{n} : s \in S\right)$. A rate tuple R_S is said to be achievable iff there exists a sequence of $((2^{\lceil nR_s \rceil} : s \in S), n)$-network codes, for $n \in \ell\mathbb{Z}^+$ for some $\ell \in \mathbb{Z}^+$, such that $\epsilon \to 0$ and $\eta \to 0$ as $n \to \infty$.[§]

R_V^{key} are called the key rates of the sequence, if for each code $((2^{\lceil nR_s \rceil} : s \in S), n)$ in the sequence, the alphabet size of random key Y_v is $K_v = 2^{\lceil nR_v^{\mathrm{key}} \rceil}$, for all $v \in V$.

B. Secure Index Coding

1) Problem instance: Denote a secure index-coding instance [7] by $\mathbb{I} = (\widehat{S}, \widehat{T}, \{(\widehat{W}_t, \widehat{H}_t) : t \in \widehat{T}\}, \widehat{P})$.

[§]The condition of non-consecutive codelengths is introduced to match the translation of network codes of length n to index codes of length $\sum_{e \in E} \lfloor c_e n \rfloor$. While it preserves the spirit of having infinitely many codes with sufficiently large n, it does not require codes with every codelength to exist or to satisfy the criteria.

- $\widehat{S}$ is the (ordered) message index set.
- $\widehat{T}$ is the receiver index set.
- $\widehat{W}_t \subseteq \widehat{S}$ contains the indices of messages required by receiver $t \in \widehat{T}$.
- $\widehat{H}_t \subsetneq \widehat{S}$ contains the indices of messages known a priori (known as side information) to receiver $t \in \widehat{T}$.
- $\widehat{P} = ((\widehat{A}_z, \widehat{B}_z) : z \in \widehat{Z})$ defines the eavesdroppers with indices in $\widehat{Z}$. Each eavesdropper $z \in \widehat{Z}$ has access to the codeword broadcast by the sender and messages indexed by $\widehat{B}_z \subsetneq \widehat{S}$, and attempts to reconstruct messages indexed by $\widehat{A}_z \subseteq \widehat{S}$. Note that $\widehat{A}_z \cap \widehat{B}_z = \emptyset$.

2) Deterministic codes: Let the messages $\{\widehat{X}_s : s \in \widehat{S}\}$ be mutually independent, and each $\widehat{X}_s$ be uniformly distributed over $[\widehat{M}_s]$ for some $\widehat{M}_s \in \mathbb{Z}^+$. A deterministic $(\widehat{M}_{\widehat{S}}, \widehat{n})$-index code consists of the following:

- A deterministic encoding (or broadcast) function for the sender: $\widehat{X}_{\mathrm{b}} = \widehat{\mathsf{e}}(\widehat{X}_{\widehat{S}}) \in [2^{\widehat{n}}]$, for some $\widehat{n} \in \mathbb{Z}^+$.
- A decoding function $\widehat{\mathsf{d}}_t$ for each receiver $t \in \widehat{T}$ that takes in $(\widehat{X}_{\mathrm{b}}, \widehat{X}_{\widehat{H}_t})$, and outputs an estimate $\widehat{X}^{(t)} = \widehat{\mathsf{d}}_t(\widehat{X}_{\mathrm{b}}, \widehat{X}_{\widehat{H}_t})$ of the messages $\widehat{X}_{\widehat{W}_t}$ that receiver t requires.

Here, the number of binary bits transmitted by the sender $\widehat{n}$ is referred to as the blocklength.

Remark 1: This index-code definition is consistent with the index-coding literature [12]–[15], but is different from that by Effros et al., where the sender transmits $\widehat{X}_{\mathrm{b}} \in [2^{\widehat{c}_{\mathrm{b}}\widehat{n}}]$, and $\widehat{c}_{\mathrm{b}}$ is then chosen to be a function of the link capacities of the equivalent network-coding instance. The difference results in a scaling factor in our rate equivalence.

3) Randomised codes: A randomised index code is similar to deterministic index codes except that the sender's encoding function takes in an independent random key $\widehat{Y} \in [\widehat{K}]$ in addition to $\widehat{X}_{\widehat{S}}$, for some $\widehat{K} \in \mathbb{Z}^+$.

4) Decodability: As with network coding, an index code has a probability of decoding error of at most $\epsilon \in \mathbb{R}_0^+$ iff

$$\widehat{P}_{\mathrm{e}} := 1 - \Pr\left\{\widehat{X}^{(t)} = \widehat{X}_{\widehat{W}_t} \text{ for all } t \in \widehat{T}\right\} \leq \epsilon. \tag{3}$$

5) Leakage: An index code has a leakage of at most $\eta \in \mathbb{R}_0^+$ iff

$$\frac{1}{\widehat{n}} I(\widehat{X}_{\widehat{A}_z}; \widehat{X}_{\mathrm{b}}, \widehat{X}_{\widehat{B}_z}) \leq \eta, \qquad \text{for all } z \in \widehat{Z}. \tag{4}$$

6) Feasibility: A secure index-coding instance $\mathbb{I}$ is said to be $(\widehat{M}_{\widehat{S}}, \epsilon, \eta, \widehat{n})$-feasible iff there exists an $(\widehat{M}_s, \widehat{n})$-index code that has a probability of decoding error of at most ϵ and a leakage of at most η. A message rate tuple of an $(\widehat{M}_s, \widehat{n})$-index code is $\widehat{R}_{\widehat{S}} := \left(\frac{\log_2 \widehat{M}_s}{\widehat{n}} : s \in \widehat{S}\right)$. A message rate tuple $\widehat{R}_{\widehat{S}}$ is said to

be achievable iff there exists a sequence of $((2^{\lceil \widehat{n}R_s \rceil} : s \in \widehat{S}), \widehat{n})$-index codes, for $\widehat{n} \in \ell\mathbb{Z}^+$ for some $\ell \in \mathbb{Z}^+$, such that $\epsilon \to 0$ and $\eta \to 0$ as $\widehat{n} \to \infty$.

C. A Note on Achievable Rates

For secure network coding, if R_S is achievable, then R'_S, where $0 \leq R'_s \leq R_s$ for all $s \in S$, is also achievable using randomised codes. This can be achieved by replacing each message of $\lceil nR_s \rceil$ bits with a new message of $\lceil nR'_s \rceil$ bits and a random key of $\lceil nR_s \rceil - \lceil nR'_s \rceil$ bits. Doing so will not affect security, as the concatenation of the shorter new message and the random key is now secure.

The above observation is not true in general for secure index coding. This is because the replacement step for a particular message cannot be replicated at the receivers having that particular message as side information. The following example illustrates this point:

Example 1: Consider a secure index coding problem with two receivers {1,2} and an eavesdropper. Receiver 1 wants X_1 and knows X_2; receiver 2 wants X_2 and knows X_1; the eavesdropper wants X_1 and knows nothing. The rate (R, R) for all $0 \leq R \leq 1$ is achievable by sending $X_1 + X_2 \mod 2^{\lceil \widehat{n}R \rceil}$. However, the rate $(R, R - \delta)$ is not achievable. To secure X_1, the sender needs to pad X_2 with $\lceil \widehat{n}R \rceil - \lceil \widehat{n}(R - \delta) \rceil$ random bits. But by doing so, receiver 1 cannot decode X_1 as it does not know these random bits.

D. Notation

Table I summarises the notation used in network coding and index coding.

III. A Summary of Results

A. Code Feasibility Equivalence

Given any secure index-coding instance $\mathbb{I}$, Section IV defines a map to obtain a corresponding secure network-coding instance $\mathbb{N}$, with the following code feasibility equivalence:

Theorem 1: (A brief version) $\mathbb{I}$ is $(\widehat{M}_{\widehat{S}}, \epsilon, \eta, \widehat{n})$-feasible iff $\mathbb{N}$ is $(\widehat{M}_{\widehat{S}}, \epsilon, \eta, \widehat{n})$-feasible.

We will prove the forward assertion in Section V and the backward assertion in Section VI.

In the other direction, given any secure network-coding instance $\mathbb{N}$, Section VII defines a map to obtain a corresponding secure index-coding instance $\mathbb{I}$, with the following code feasibility equivalence:

Theorem 2: (A brief version) Let $\widehat{n} = \sum_{e \in E} \lfloor c_e n \rfloor$ and $2^{\lfloor c_E n \rfloor} := \left(2^{\lfloor c_e n \rfloor} : e \in E\right)$.

	Secure network coding	Secure index coding
Nodes		
The set of nodes	V	Sender and $\widehat{T}$
Node j wants these messages	$X_{D^{-1}(j)}$	$\widehat{X}_{\widehat{W}_j}$
Node j has these messages	$X_{O^{-1}(j)}$	Sender has $\widehat{X}_{\widehat{S}}$ Node $j \in \widehat{T}$ has $\widehat{X}_{\hat{H}_j}$
Node j listens to these transmissions	$X_{\mathtt{in}(j)}$	$\widehat{X}_{\mathrm{b}}$
Source messages		
The set of messages	X_S	$\widehat{X}_{\widehat{S}}$
Message $X_i/\widehat{X}_i$ is known to these nodes	$O(i)$	Sender and $\{j : i \in \widehat{H}_j\}$
Message $X_i/\widehat{X}_i$ is wanted by these nodes	$D(i)$	$\{j : i \in \widehat{W}_j\}$
Links		
The set of links	E	One broadcast link
Transmission on link $e \in E$ or broadcast link	X_e	$\widehat{X}_{\mathrm{b}}$
Capacity of link $e \in E$ or broadcast link	c_e	1
Eavesdroppers		
The index set of eavesdroppers	Z	$\widehat{Z}$
Eavesdropper z wants these messages	X_{A_z}	$\widehat{X}_{\widehat{A}_z}$
Eavesdropper z has these messages	$\emptyset$	$\widehat{X}_{\widehat{B}_z}$
Eavesdropper z listens to these transmissions	X_{B_z}	$\widehat{X}_{\mathrm{b}}$
Codes		
Encoding function on link e or broadcast link	$X_e = \mathsf{e}_e\Big(X_{\mathtt{in}(\mathtt{tail}(e))}, X_{O^{-1}(\mathtt{tail}(e))}, Y_{\mathtt{tail}(e)}\Big)$	$\widehat{X}_{\mathrm{b}} = \widehat{\mathsf{e}}(\widehat{X}_{\widehat{S}}, \widehat{Y})$
Decoding function at node j	$\mathsf{d}_j(X_{\mathtt{in}(j)}, X_{O^{-1}(j)})$	$\widehat{\mathsf{d}}_j(\widehat{X}_{\mathrm{b}}, \widehat{X}_{\widehat{H}_j})$

TABLE I: List of symbols used for secure network coding and secure index coding

1) If $\mathbb{N}$ is (M_S, ϵ, η, n)-feasible, then $\mathbb{I}$ is $\Big((M_S, K_V, 2^{\lfloor c_E n \rfloor}), \epsilon, \theta_1, \widehat{n}\Big)$-feasible, where
$$\theta_1 := \frac{\eta}{\left(\sum_{e\in E} c_e - \frac{|E|}{n}\right)}.$$
2) If $\mathbb{I}$ is $\Big((M_S, K_V, 2^{\lfloor c_E n \rfloor}), \epsilon, \eta, \widehat{n}\Big)$-feasible, then $\mathbb{N}$ is $\big(M_S, (|Z|+1)\epsilon, \theta_2, n\big)$-feasible, where
$$\theta_2 := (|Z|+1)\left[\left(\frac{\eta}{1-\epsilon} + \epsilon\right)\sum_{e\in E} c_e + \frac{1}{n}\left(\frac{H_{\mathrm{b}}(\epsilon)}{1-\epsilon} + H_{\mathrm{b}}(\epsilon)\right)\right]$$
and $H_{\mathrm{b}}(\cdot)$ denotes the binary entropy function.

We will prove Part 1 in Section VIII and Part 2 in Section IX.

Remark 2:

1) For perfect decoding, that is, $\epsilon = 0$, we have that $\theta_1 = \eta\frac{1}{\left(\sum_{e\in E} c_e - \frac{|E|}{n}\right)} \leq \psi_1\eta$ and $\theta_2 = \eta(|Z|+1)\sum_{e\in E} c_e = \psi_2\eta$, for some constants ψ_1, ψ_2. The term $\sum_{e\in E} c_e$ in the scaling factor is a result of the way leakage is normalised: $\frac{1}{n}$ for network coding and $\frac{1}{\widehat{n}} = \frac{1}{\sum_{e\in E}\lfloor c_e n\rfloor}$ for index coding.
2) For perfect decoding and zero leakage ($\epsilon = \eta = 0$), we have $\theta_1 = \theta_2 = 0$, regardless of n.

3) If $n \to \infty$, $\epsilon \to 0$, and $\eta \to 0$, then $(|Z|+1)\epsilon \to 0$, $\theta_1 \to 0$, and $\theta_2 \to 0$.
4) In this direction, a network-coding instance $\mathbb{N}$ with $|S|$ sources is mapped to an index-coding instance $\mathbb{I}$ with $|S|+|V|+|E|$ sources.
5) For Theorem 2, deterministic index codes suffice.

B. Rate Equivalence

From the above code translations, we obtain the following rate-equivalence results.

Given any secure index-coding instance $\mathbb{I}$ and its corresponding secure network-coding instance $\mathbb{N}$ (via the instance map defined in Section IV), we have the following rate equivalence:

Corollary 1: (A brief version) The rate tuple $\widehat{R}_{\widehat{S}}$ is achievable for $\mathbb{I}$ iff it is achievable for $\mathbb{N}$.

For the other direction, we consider any secure network-coding instance $\mathbb{N}$, where all link capacities c_e are integers, and the corresponding secure index-coding instance $\mathbb{I}$ (via the instance map defined in Section VII).

Corollary 2: (A brief version) The rate tuple R_S is achievable for $\mathbb{N}$ (with integer link capacities), using random key with rates R_V^{key}, iff the rate tuple $\frac{1}{\sum_{e\in E} c_e}(R_S, R_V^{\text{key}}, c_E)$ is achievable for $\mathbb{I}$.

Remark 3: Corollary 2 can be extended to any secure network-coding instance $\mathbb{N}$ with rational link capacities. To this end, we consider $\lambda \in \mathbb{Z}^+$ uses of the links as a group, where λ is the least common multiple of the denominators of the link capacities. Each group of λ uses of the links is now equivalent to another network coding instance $\mathbb{N}'$ with link capacities $c'_e = \lambda c_e \in \mathbb{Z}^+$. We can now apply Corollary 2 to $\mathbb{N}'$ to get a rate equivalence between $R'_S = \lambda R_S$ for $\mathbb{N}'$ (which is R_S for $\mathbb{N}$) and $\frac{1}{\sum_{e\in E} c'_e}(R'_S, R_V'^{\text{key}}, c'_E) = \frac{1}{\sum_{e\in E} c_e}(R_S, R_V^{\text{key}}, c_E)$ for $\mathbb{I}$.

IV. An Equivalence from Secure Index Coding to Secure Network Coding

Given a secure index-coding instance $\mathbb{I} = (\widehat{S}, \widehat{T}, \{(\widehat{W}_t, \widehat{H}_t) : t \in \widehat{T}\}, \widehat{P})$. Let $\widehat{S} = [k]$ and $\widehat{T} = [\ell]$ for some positive integers k and ℓ. We will first propose a map to a secure network-coding instance $\mathbb{N} = (G, C, P)$.

A. Index-to-Network Coding Instance Map

The secure version of the index-to-network coding instance map consists of the following:

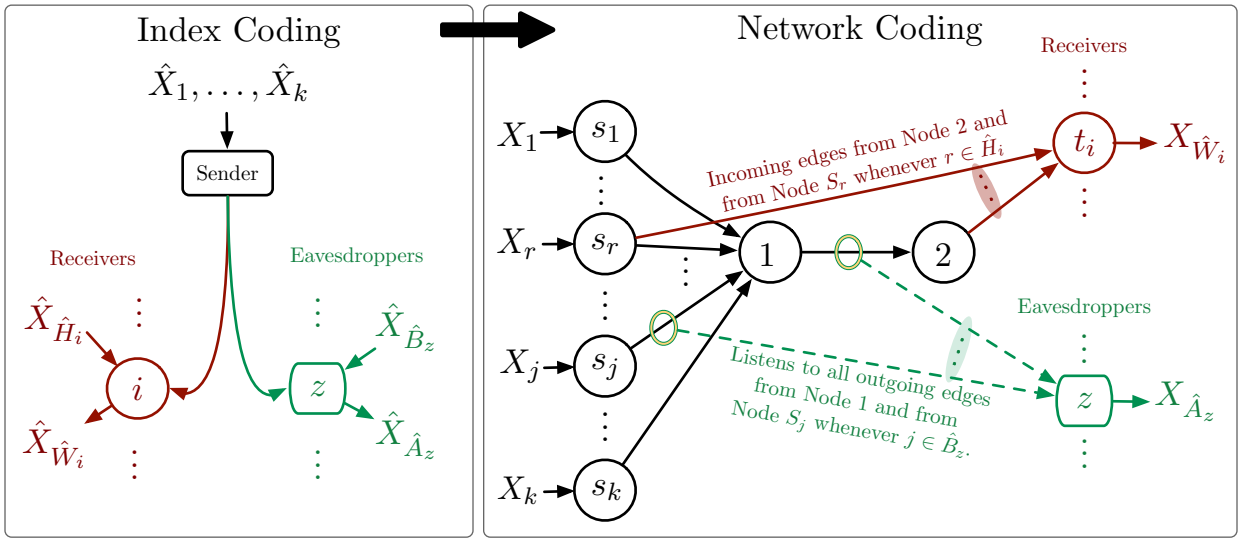


Fig. 1: The map from secure index coding to secure network coding, where red arrows indicate information associated with the receivers, and green arrows the eavesdroppers.

1) An existing non-secure map [3] for $G = (V, E)$ and $C = (S, O, D)$:

- The vertex set $V = \{s_1, s_2, \ldots, s_k, t_1, t_2, \ldots, t_\ell, 1, 2\}$.
- The link set E contains the following links:
  - $(s_i \to 1)$ for each $i \in [k]$, with sufficiently large capacities.
  - $(s_i \to t_j)$ iff $i \in \widehat{H}_j$, with sufficiently large capacities.
  - $(1 \to 2)$ with link capacity 1.
  - $(2 \to t_j)$ for each $j \in [\ell]$, with link capacity 1.
- The message indices $S = \widehat{S} = [k]$.
- The originating node for message $X_i, i \in S$, is $O(i) = s_i$.
- The destinations for message X_i is $D(i) = \{t_j : i \in \widehat{W}_j\}$.

2) Our proposed map for $P = ((A_z, B_z) : z \in Z)$:

- $Z = \widehat{Z}$
- For each $z \in Z$: $A_z = \widehat{A}_z$.
- For each $z \in Z$, B_z comprises of links included by the following rules:
  - $(1 \to 2) \in B_z$.
  - For any $i \in \hat{B}_z$, all outgoing links from node s_i are in B_z.

Remark 4: The number of source messages in both instances is the same. Side information in $\mathbb{I}$ manifests itself as links $(s_i \to t_j)$ in $\mathbb{N}$. In the constructed $\mathbb{N}$, sources nodes are $\{s_i : i \in \widehat{S}\}$, and destination nodes

are $\{t_i : i \in \widehat{T}\}$.

The index-to-network coding instance map is summarised in Figure 1.

B. Equivalence Results

With the above map, we prove an equivalence between these two instances.

Theorem 1: Let $\mathbb{I}$ be a secure index-coding instance, and $\mathbb{N}$ be the corresponding secure network-coding instance using the index-to-network coding map. For any $\epsilon, \eta \in \mathbb{R}_0^+$ and $\widehat{n} \in \mathbb{Z}^+$, $\mathbb{I}$ is $(\widehat{M}_{\widehat{S}}, \epsilon, \eta, \widehat{n})$-feasible iff $\mathbb{N}$ is $(\widehat{M}_{\widehat{S}}, \epsilon, \eta, \widehat{n})$-feasible with deterministic encoding functions at vertices 2 and $\{s_i : i \in \widehat{S}\}$, and a randomised encoding function at node 1.

Proof: See Sec. V for the forward direction and Sec. VI for the backward direction. ■

The theorem above preserves the message size, as well as the decodability and security criteria. The proof of the theorem utilises the non-secure code translations, which has been shown to preserve the decoding error criterion ϵ when the codes are deterministic in the absent of eavesdroppers. As an equivalence for the secure instances is required here, our main contribution in this direction for the equivalence is to show that this code translation

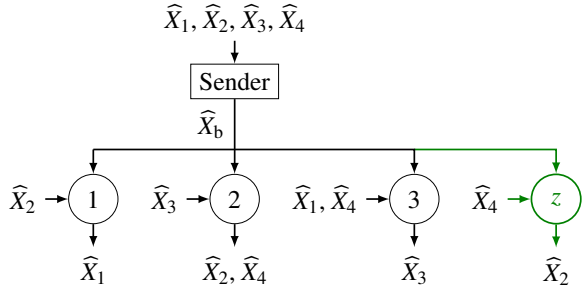
- works with the addition of eavesdroppers,
- works for all randomised index codes,
- preserves the security criterion η,
- still preserves the decoding criterion ϵ.

With Theorem 1, we can show the following rate equivalence:

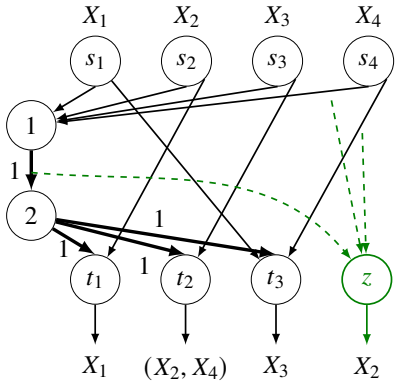
Corollary 1: Let $\mathbb{I}$ be a secure index-coding instance, and $\mathbb{N}$ be the corresponding secure network-coding instance using the index-to-network coding map. A rate tuple $\widehat{R}_{\widehat{S}}$ is achievable for $\mathbb{I}$ iff it is also achievable for $\mathbb{N}$.

Proof of Corollary 1: If a rate tuple $\widehat{R}_{\widehat{S}}$ is achievable for $\mathbb{I}$, then there exists a sequence of $\left(\left(2^{\lceil \widehat{n}\widehat{R}_s \rceil}\right), \widehat{n}\right)$-index codes, $\widehat{n} \in \ell\mathbb{Z}^+$, with probability of decoding error $\epsilon \to 0$ and leakage $\eta \to 0$ as $\widehat{n} \to \infty$. From Theorem 1, it follows that there exists a sequence of network codes with the same properties, and hence $\widehat{R}_{\widehat{S}}$ is also achievable for $\mathbb{N}$.

The other direction follows exactly the same argument. ■



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



(b) A secure network-coding instance $\mathbb{N}$, where an eavesdropper z 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. 2: A secure index-coding instance and its corresponding secure network-coding instance

C. An Example

Consider the index-coding instance $\mathbb{I}$ depicted in Figure 2a, and its mapped network-coding instance $\mathbb{N}$ in Figure 2b. An example of index codes with $\epsilon = 0$ and $\eta = 0$ is $\widehat{\mathsf{e}}(\widehat{X}_{[4]}, \widehat{Y}) = (\widehat{X}_1 + \widehat{X}_2, \widehat{X}_2 + \widehat{X}_3, \widehat{X}_4)$. One can verify that each user can decode their intended messages, and the eavesdropper z has no information about $\widehat{X}_2$. In the translated network code, each source node s_i, $i \in [4]$, transmits X_i on every outgoing link; node 1 transmits $X_{1\to 2} = \mathsf{g}_{1\to 2}(X_{[4]}, Y_1) = \widehat{\mathsf{e}}(X_{[4]}, Y_1) = (X_1 + X_2, X_2 + X_3, X_4)$; node 2 forwards $X_{1\to 2}$ on every outgoing link. Clearly, each destination node in $\mathbb{N}$ can decode its required messages using the same decoding function in $\mathbb{I}$, and the eavesdropper z gains no information about X_2.

In the other direction, consider a network code with $\epsilon = 0$ and $\eta = 0$ as follows: each source node s_i, $i \in [4]$, transmits X_i on every outgoing link; node 1 transmits $X_{1\to 2} = \mathsf{g}_{1\to 2}(X_{[4]}, Y_1) = (X_1 + X_3, X_2 + X_3, X_4)$; node 2 forwards $X_{1\to 2}$ on every outgoing link. The translated index code is $\widehat{\mathsf{e}}(\widehat{X}_{[4]}, \widehat{Y}) = \mathsf{g}_{1\to 2}(\widehat{X}_{[4]}, \widehat{Y}) = (\widehat{X}_1 + \widehat{X}_3, \widehat{X}_2 + \widehat{X}_3, \widehat{X}_4)$. One can verify that for both instances, all users can decode their requested messages, and the eavesdropper gains no information of the messages it attempts to decode.

V. Proof of Theorem 1 – the forward direction

We will now prove Theorem 1 for the forward direction: if $\mathbb{I}$ is $(\widehat{M}_{\widehat{\mathcal{S}}}, \epsilon, \eta, \widehat{n})$-feasible, then $\mathbb{N}$ is $(\widehat{M}_{\widehat{\mathcal{S}}}, \epsilon, \eta, \widehat{n})$-feasible. This is achieved by showing that for any index code that satisfies the feasibility condition for $\mathbb{I}$, we can translate it to a network code that satisfies the feasibility condition for $\mathbb{N}$.

A. Code Translation

We start with any randomised index code $(\widehat{\mathsf{e}}, \widehat{\mathsf{D}})$ for $\mathbb{I}$ that is $(\widehat{M}_{\widehat{S}}, \epsilon, \eta, \widehat{n})$-feasible. We will show that the network code obtained by the existing non-secure code translation satisfies both the decoding and security criteria even for randomised codes. In the following, we modify the existing non-secure code translation to translate randomised index codes.

- For all outgoing links from s_i, $i \in [k]$: Set a deterministic link function $X_e = \mathsf{e}_e(X_{O^{-1}(s_i)}) = \mathsf{e}_e(X_i) = X_i \in [\widehat{M}_i]$, for each $e \in \mathtt{out}(s_i)$. This is possible since vertex s_i is the originating vertex for the message X_i, and the link capacity is sufficiently large.
- For link $(1 \to 2)$: Set $X_{1\to 2} = \mathsf{e}_{1\to 2}(X_{\mathtt{in}(1)}, Y_1) = \widehat{\mathsf{e}}(X_{[k]}, Y_1) \in [2^{\widehat{n}}]$. Here, we have set the cardinality of Y_1 (which is the random key used in the encoding function of vertex 1 in $\mathbb{N}$) to be the same as that of the random key $\widehat{Y}$ used in the encoding function $\widehat{\mathsf{e}}(\widehat{X}_{\widehat{S}}, \widehat{Y})$ of the sender in $\mathbb{I}$.
- For all outgoing links from 2: Set a deterministic function $X_e = \mathsf{e}_e(X_{1\to 2}) = X_{1\to 2} \in [2^{\widehat{n}}]$, for each $e \in \mathtt{out}(2)$. This is possible as every outgoing link from vertex 2 has capacity $\widehat{n}$.
- Set $\mathsf{d}_{t_i}(X_{\mathtt{in}(t_i)}) = \widehat{\mathsf{d}}_i(X_{2\to t_i}, X_{\widehat{H}_i})$ for all $i \in [\ell]$, and $\mathsf{d}_v = 0$ for all other vertices v.

B. Decoding Criterion

See Appendix A for the proof of the decoding criterion, which is similar to that of the non-secure equivalence.

C. Security Criteria

Each eavesdropper $z \in Z$ in $\mathbb{N}$ has access to links B_z consisting of (i) link $(1 \to 2)$, which carries $X_{1\to 2} = \widehat{\mathsf{e}}(X_S, Y_1)$, and (ii) outgoing links from $\{s_i : i \in \widehat{B}_z\}$, which carry messages $X_{\widehat{B}_z}$, because each outgoing link from node s_i carries X_i by construction.

Now, we bound the leakage for the network code as follows:

$$\frac{1}{\widehat{n}} I(X_{A_z}; X_{B_z}) = \frac{1}{\widehat{n}} I\left(X_{\widehat{A}_z}; \widehat{\mathsf{e}}(X_S, Y_1), X_{\widehat{B}_z}\right) \overset{\text{(a)}}{=} \frac{1}{\widehat{n}} I\left(\widehat{X}_{\widehat{A}_z}; \widehat{\mathsf{e}}(\widehat{X}_{\widehat{S}}, \widehat{Y}), \widehat{X}_{\widehat{B}_z}\right) \overset{\text{(b)}}{\leq} \eta, \tag{5}$$

where (a) follows from a change of random variables by noting that $(\widehat{X}_{\widehat{S}}, \widehat{Y})$ has the same distribution as (X_S, Y_1); (b) follows from the premise that the leakage for the index code is at most η. This completes the security proof for $\mathbb{N}$.

VI. Proof of Theorem 1 – the backward direction

We will now prove Theorem 1 for the backward direction: if $\mathbb{N}$ is $(\widehat{M}_{\widehat{S}}, \epsilon, \eta, \widehat{n})$-feasible, then $\mathbb{I}$ is $(\widehat{M}_{\widehat{S}}, \epsilon, \eta, \widehat{n})$-feasible.

A. Code Construction

Similar to the forward direction, we start with any network code (E, D) for $\mathbb{N}$ that satisfies the feasibility conditions. Recall the encoding functions at node 2 and nodes $\{s_i : i \in [k]\}$ are deterministic, and that at node 1 randomised. In the following, we modify the existing non-secure code translation to translate a network code with a random key Y_1 used at node 1.

- Set the sender's transmitted code to be $\widehat{X}_{\mathrm{b}} = \widehat{\mathsf{e}}(\widehat{X}_{\widehat{S}}, \widehat{Y}) = \mathsf{e}_{1\to 2}\Big(\big(\mathsf{e}_{s_i\to 1}(\widehat{X}_i) : i \in [k]\big), \widehat{Y}\Big) \in [2^{\widehat{n}}]$, where $\widehat{Y}$ has the same cardinality as Y_1 (the random key used in the network code).
- Set the decoding function of receiver $i \in [\ell]$ to be $\widehat{\mathsf{d}}_i(\widehat{X}_{\mathrm{b}}, \widehat{X}_{\widehat{H}_i}) = \mathsf{d}_{t_i}\Big(\mathsf{e}_{2\to t_i}(\widehat{X}_{\mathrm{b}}), \big(\mathsf{e}_{s_j\to t_i}(\widehat{X}_j) : j \in \widehat{H}_i\big)\Big)$.

B. Restricting Network Codes

Next, we show that we only need to consider a specific class of network codes without loss of generality.[¶] Specifically, we only need to consider network codes such that $\mathsf{e}_{2\to t_i}(X_{1\to 2}) = X_{1\to 2}$, for all $i \in [\ell]$. First, observe that

$$X_{\widehat{W}_i} - \Big(X_{1\to 2}, \big(\mathsf{e}_{s_j\to t_i}(X_j) : j \in \widehat{H}_i\big)\Big) - \Big(\mathsf{e}_{2\to t_i}(X_{1\to 2}), \big(\mathsf{e}_{s_j\to t_i}(X_j) : j \in \widehat{H}_i\big)\Big) \tag{6}$$

forms a Markov chain, where $X_{1\to 2} = \mathsf{e}_{1\to 2}\Big(\big(\mathsf{e}_{s_i\to 1}(X_i) : i \in [k]\big), Y_1\Big)$, and $\widehat{W}_i \cap \widehat{H}_i = \emptyset$.

Recall that receiver t_i, for each $i \in [\ell]$, attempts to decode $X_{\widehat{W}_i}$ from $\Big(\mathsf{e}_{2\to t_i}(X_{1\to 2}), \big(\mathsf{e}_{s_j\to t_i}(X_j) : j \in \widehat{H}_i\big)\Big)$. By the data-processing inequality, the probability of decoding error P_{e} cannot increase if we set $\mathsf{e}_{2\to t_i}(X_{1\to 2}) = X_{1\to 2}$ in each receiver t_i's observations. Also, by definition, none of the links $\{2 \to t_i : i \in [\ell]\}$ can be accessed by any eavesdropper. So, this choice will not affect the leakage of the code.

Consequently, for any network code $(\mathsf{E}', \mathsf{D}')$ for $\mathbb{N}$ with at most ϵ probability of error and η leakage, we can always obtain another network code (E, D) by modifying $(\mathsf{E}', \mathsf{D}')$ such that $\mathsf{e}_{2\to t_i}(X_{1\to 2}) = X_{1\to 2}$, for all $i \in [\ell]$. The modified network code also has an error-decoding probability of at most ϵ and a leakage of at most η. For the subsequent subsections, we will only consider network codes of this form.

[¶]This was not shown in existing works on non-secure equivalence.

C. Decoding Criterion

See Appendix B for the proof of the decoding criterion, which is similar to that of the non-secure equivalence.

D. Security Criteria

From the security criteria of $\mathbb{N}$, we have $\frac{1}{\tilde{n}} I(X_{A_z}; X_{B_z}) < \eta$ for each z. Eavesdropper z observes links B_z, which consists of all outgoing links from source nodes $\{s_i : i \in \widehat{B}_z\}$ as well as link $1 \to 2$. It attempts to decode messages indexed by $A_z = \widehat{A}_z$.

Showing that the translated index code also satisfies a similar security condition as the original network code 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 decoding error probability. So, it seems that the eavesdroppers in $\mathbb{I}$ have "better" observations, which may lead to a larger leakage of the code.

We will show that this is not the case. First, note the following: (i) $\{X_S, Y_1\}$ are mutually independent; (ii) $X_{\mathtt{out}(s_i)}$, for each $i \in S$, are each a deterministic function of X_i; (iii) $\widehat{B}_z \cap A_z = \emptyset$. With these, we have the following Markov chain for every $z \in Z$:

$$X_{\widehat{B}_z} - X_{\{\mathtt{out}(s_i):i\in\widehat{B}_z\}} - \left(Y_1, X_{A_z}, X_{S\setminus(A_z\cup\widehat{B}_z)}\right), \tag{7}$$

which is equivalent to

$$\begin{aligned}
0 &= I(X_{\widehat{B}_z}; Y_1, X_{A_z}, X_{S\setminus(A_z\cup\widehat{B}_z)} | X_{\{\mathtt{out}(s_i):i\in\widehat{B}_z\}}) && \text{(8a)}\\
&= I(X_{\widehat{B}_z}; Y_1, X_{A_z}, X_{S\setminus(A_z\cup\widehat{B}_z)}, X_{\{\mathtt{out}(s_i):i\in\widehat{B}_z\}} | X_{\{\mathtt{out}(s_i):i\in\widehat{B}_z\}}) && \text{(8b)}\\
&= I(X_{\widehat{B}_z}; Y_1, X_{S\setminus\widehat{B}_z}, X_{\{\mathtt{out}(s_i):i\in\widehat{B}_z\}}, X_{\{s_i\to 1:i\in S\}} | X_{\{\mathtt{out}(s_i):i\in\widehat{B}_z\}}) && \text{(8c)}\\
&= I(X_{\widehat{B}_z}; Y_1, X_{S\setminus\widehat{B}_z}, X_{\{\mathtt{out}(s_i):i\in\widehat{B}_z\}}, X_{\{s_i\to 1:i\in S\}}, X_{1\to 2} | X_{\{\mathtt{out}(s_i):i\in\widehat{B}_z\}}) && \text{(8d)}\\
&\geq I(X_{\widehat{B}_z}; X_{A_z}, X_{1\to 2} | X_{\{\mathtt{out}(s_i):i\in\widehat{B}_z\}}) && \text{(8e)}\\
&\geq I(X_{\widehat{B}_z}; X_{A_z} | X_{\{\mathtt{out}(s_i):i\in\widehat{B}_z\}}, X_{1\to 2}) && \text{(8f)}\\
&= I(X_{\widehat{B}_z}; X_{A_z} | X_{B_z}) \geq 0. && \text{(8g)}
\end{aligned}$$

This means that eavesdropper z, having observed the links X_{B_z}, does not gain any more information about $X_{\widehat{A}_z}$ even if it can also observe the source messages $X_{\widehat{B}_z}$. Now, we show that the eavesdropper cannot do

better if we replace its observation of the outgoing links from the sources with the source messages:

$$\frac{1}{\widehat{n}} I(X_{\widehat{B}_z}, X_{1\to 2}; X_{A_z}) = \frac{1}{\widehat{n}} I(X_{\widehat{B}_z}, X_{1\to 2}, X_{\{\mathtt{out}(s_i):i\in\widehat{B}_z\}}; X_{A_z}) \tag{9a}$$

$$= \frac{1}{\widehat{n}} I(X_{\widehat{B}_z}, X_{B_z}; X_{A_z}) \tag{9b}$$

$$= \frac{1}{\widehat{n}} I(X_{B_z}; X_{A_z}) + I(X_{\widehat{B}_z}; X_{A_z}|X_{B_z}) \tag{9c}$$

$$= \frac{1}{\widehat{n}} I(X_{B_z}; X_{A_z}) \tag{9d}$$

$$\le \eta, \tag{9e}$$

where (9d) follows from (8g), and (9e) follows from the premise of the network code.

Since $(\widehat{X}_{[k]}, \widehat{Y}, \widehat{X}_{\mathrm{b}})$ and $(X_{[k]}, Y_1, X_{1\to 2})$ have the same distribution, we have $\frac{1}{\widehat{n}} I(\widehat{X}_{\widehat{B}_z}, \widehat{X}_{\mathrm{b}}; \widehat{X}_{\widehat{A}_z}) \le \eta$ for $\mathbb{I}$. This completes the proof. ∎

VII. An Equivalence from Secure Network Coding to Secure Index Coding

In the other direction, consider a secure network-coding instance $\mathbb{N} = (G, C, P)$. For simplicity, let the source indices be $S = [k]$ and the vertex indices be $V = [\ell]$. Recall that each message is requested by at least one destination.

We will construct the following map to obtain an index-coding instance $\mathbb{I}$:

1) Construct an augmented secure network-coding instance $\mathbb{N}'$ from any (possibly randomised) secure network-coding instance $\mathbb{N}$. This step converts any possibly randomised secure network code to a deterministic network code.
2) Map $\mathbb{N}'$ to $\mathbb{I}$:
   a) For legitimate receivers: We use the existing non-secure instance map (which only works for deterministic codes), except that we omit one receiver in $\widehat{T}$. We will show that omitting this receiver will not affect the result.
   b) For eavesdroppers: We construct a map from the eavesdroppers in $\mathbb{N}'$ to those in $\mathbb{I}$.

For $\mathbb{I}$, we set

$$\widehat{n} = \sum_{e\in E} \lfloor c_e n \rfloor. \tag{10}$$

This means the number of bits that the sender can transmit in $\mathbb{I}$ equals the total number of bits that can be transmitted on all the links in $\mathbb{N}$.

A. Network-to-Index Coding Map

Now, we describe the instance map in detail:

1) Augmented secure network coding: We construct an augmented secure network-coding instance $\mathbb{N}' = (G', C', P')$ as follows:

- $G' = (V, E) = G$, where each link e in G' has the same capacity c_e as that in G.
- $C' = (S', O', D')$: Here, we introduce an additional independent source $X'_{k+i} \in [K_i]$ originating at each vertex $i \in [\ell]$ that takes the role of and has the same distribution as the random key Y_i used in the randomised encoding at vertex i in $\mathbb{N}$.
  - $S' = [k + \ell]$, where the message alphabet sizes are $M'_i = M_i$ for each $i \in [k]$, and $M'_{k+i} = K_i$ for each $i \in [\ell]$.
  - $O'(i) = O(i)$ for each $i \in [k]$, and $O'(k + i) = i$ for each $i \in [\ell]$
  - $D'(i) = D(i)$ for each $i \in [k]$, and $D'(k + i) = \emptyset$ for each $i \in [\ell]$.
- $P' = ((A_z, B_z) : z \in Z) = P$.

Note that by construction, $X'_{[k+\ell]}$ and $(X_{[k]}, Y_{[\ell]})$ have the same distribution. So, there is a bijective map from a deterministic or randomised secure network code for $\mathbb{N}$ to a deterministic secure network code for $\mathbb{N}'$. Note that in $\mathbb{N}'$, the additional sources $\{X'_{k+i} : i \in [\ell]\}$ are not required to be decoded by any node. Also, they are neither known to any eavesdropper nor required to be protected.

Denote the set of vertices in $\mathbb{N}'$ that are the destinations for some source messages by $U = \{j \in [\ell] : j \in D'(i) \text{ for some } i \in [k]\}$.

2) Network-to-index coding map: Now, we map $\mathbb{N}'$ to a secure index-coding instance $\mathbb{I}$.

- $\widehat{S} = [k + \ell] \cup E$. It consists of one message $\widehat{X}_i \in [\widehat{M}_i] = [M'_i]$ for each $i \in [k + \ell]$ and one message $\widehat{X}_e \in [\widehat{M}_e] = [2^{\lfloor c_e n \rfloor}]$ for each $e \in E$ in $\mathbb{N}'$.
- $\widehat{T} = \{\widehat{t}_i\}_{i \in U} \cup \{\widehat{t}_e\}_{e \in E}$. This means $\mathbb{I}$ has $|U|+|E|$ receivers: the first set corresponds to each destination node in $\mathbb{N}'$, and the second set corresponds to each link in $\mathbb{N}'$.
- For each $\widehat{t}_e \in \widehat{T}$ where $e \in E$, we set $\widehat{H}_{\widehat{t}_e} = \mathtt{in}(\mathtt{tail}(e)) \cup O'^{-1}(\mathtt{tail}(e))$, and $\widehat{W}_{\widehat{t}_e} = \{e\}$.
- For each $\widehat{t}_i \in \widehat{T}$ where $i \in U$, we set $\widehat{H}_{\widehat{t}_i} = \mathtt{in}(i) \cup O'^{-1}(i)$, and $\widehat{W}_{\widehat{t}_i} = \{j \in [k + \ell] : i \in D'(j)\} = D'^{-1}(i)$.
- The eavesdropper setting $\widehat{P}$: $\widehat{Z} = Z$. For each $z \in \widehat{Z}$, $\widehat{B}_z = B_z$, and $\widehat{A}_z = A_z$.

The two steps in the map from a network coding instance to the corresponding index-coding instance are summarised in Figures 3 and 4 respectively.

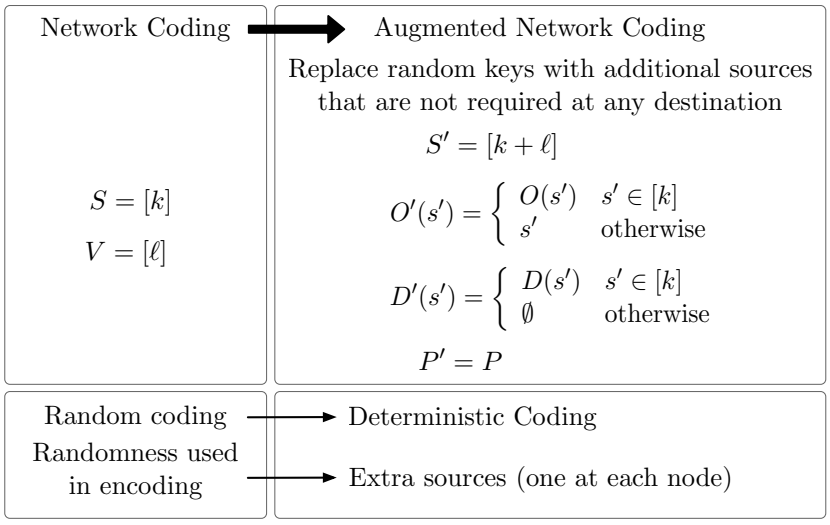


Fig. 3: From secure network coding to augmented secure network coding.

Remark 5: The mapping to the receivers in $\mathbb{I}$ from $\mathbb{N}'$ is slightly different from that in the non-secure instance map [3], which includes an additional receiver $\widehat{t}_{\text{all}}$ in $\mathbb{I}$. The receiver was included to guarantee a useful property, which, as we will see in Proposition 1, remains true even without receiver $\widehat{t}_{\text{all}}$.

B. Equivalence Results

With the above-mentioned conversion, we now state an equivalence between $\mathbb{N}$ and $\mathbb{I}$ through $\mathbb{N}'$. Recall that $\widehat{n} = \sum_{e\in E} \lfloor c_e n \rfloor$ and $2^{\lfloor c_E n \rfloor} := \left(2^{\lfloor c_e n \rfloor} : e \in E\right)$.

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

1) If $\mathbb{N}$ is (M_S, ϵ, η, n)-feasible, then $\mathbb{I}$ is $\left((M_S, K_V, 2^{\lfloor c_E n \rfloor}), \epsilon, \theta_1, \widehat{n}\right)$-feasible with a deterministic index code for some $K_V \in (\mathbb{Z}^+)^\ell$, where $\theta_1 := \frac{\eta}{\left(\sum_{e\in E} c_e - \frac{|E|}{n}\right)}$.
2) If $\mathbb{I}$ is $\left((M_S, K_V, 2^{\lfloor c_E n \rfloor}), \epsilon, \eta, \widehat{n}\right)$-feasible with a deterministic index code, then $\mathbb{N}$ is $\left(M_S, (|Z| + 1)\epsilon, \theta_2, n\right)$-feasible, where $\theta_2 := (|Z| + 1)\left[\left(\frac{\eta}{1-\epsilon} + \epsilon\right)\sum_{e\in E} c_e + \frac{1}{n}\left(\frac{H_b(\epsilon)}{1-\epsilon} + H_b(\epsilon)\right)\right]$.

Proof: See Section VIII for the proof of Part 1 and Section IX for Part 2. ■

Unlike the index-to-network map, here $\widehat{X}_E$ and X'_E have different distributions. In $\mathbb{I}$, $\widehat{X}_E$ are the source messages, which are mutually independent; in $\mathbb{N}'$, X'_E are the link messages, which are functions of the source messages $X'_{[k+\ell]}$ and may be correlated.

Theorem 2 leads to the following rate equivalence.

Corollary 2: Let $\mathbb{N}$ be a secure network-coding instance with $c_e \in \mathbb{Z}^+$, and $\mathbb{I}$ be the corresponding secure index-coding instance obtained using the network-to-index coding map. A rate tuple R_S is achievable for

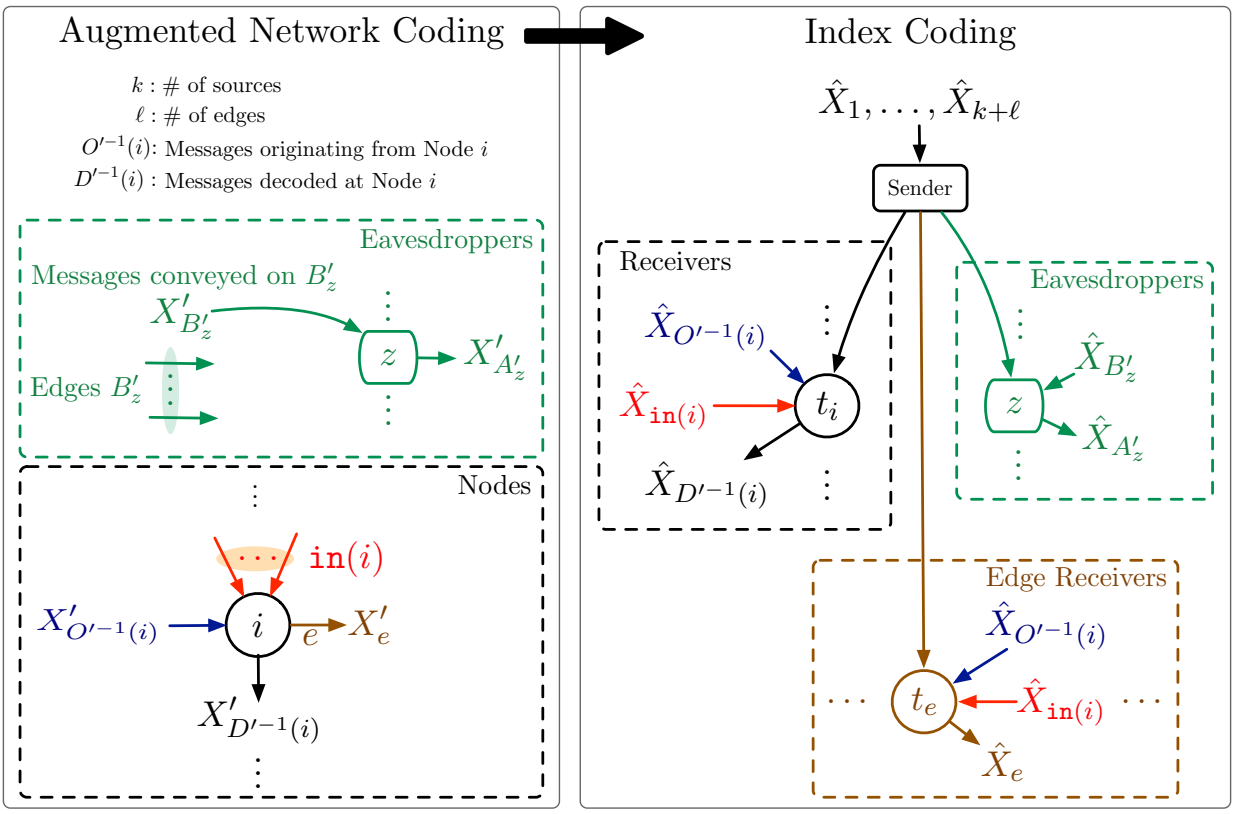


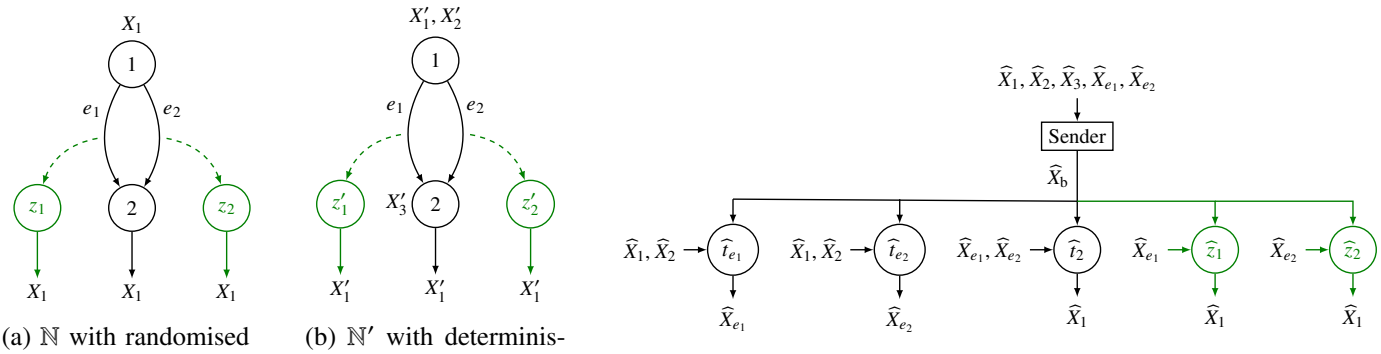
Fig. 4: From augmented secure network coding to secure index coding: For a node in augmented secure network coding, blue arrows indicate the originating messages, red arrows the incoming links, brown arrows the outgoing link(s), and black arrows the messages requested by the node. Green arrows are associated with the eavesdroppers. The same colours are used in secure index coding to indicate how the messages are mapped from the corresponding augmented secure network coding.

$\mathbb{N}$ (with codelengths $\ell\mathbb{Z}^+$) using random key rates R_V^{key} iff the rate tuple $\frac{1}{\sum_{e\in E} c_e}(R_S, R_V^{\text{key}}, c_E)$ is achievable for $\mathbb{I}$ (with codelengths $\ell(\sum_{e\in E} c_e)\mathbb{Z}^+$).

Proof of Corollary 2: Suppose that a rate tuple $R_S \in (\mathbb{R}_0^+)^{|S|}$ is achievable for $\mathbb{N}$. Then there exists a sequence of network codes $\left(2^{\lceil nR_s\rceil}, n\right)$, $n \in \ell\mathbb{Z}^+$, where each code has messages sizes $2^{\lceil nR_s\rceil}$, for $s \in S$, with probability of decoding error $\epsilon \to 0$ and leakage $\eta \to 0$ as $n \to \infty$.

From Theorem 2, it follows that there exists a sequence of index codes $\left(\left[(2^{\lceil \widehat{n}\frac{n}{\widehat{n}}R_s\rceil}, s \in S), (2^{\left\lceil \widehat{n}\frac{n}{\widehat{n}}R_v^{\text{key}}\right\rceil}, v \in V), (2^{\widehat{n}\frac{\lfloor nc_e\rfloor}{\widehat{n}}}, e \in E)\right], \widehat{n}\right)$, $\widehat{n} = \sum_{e\in E}\lfloor c_e n\rfloor = n\sum_{e\in E} c_e \in \ell(\sum_{e\in E} c_e)\mathbb{Z}^+$, with probability of decoding error $\epsilon \to 0$ and leakage $\theta_{1,n} \to 0$ as $\eta \to 0$ and $n \to \infty$. Hence, the rate $\frac{1}{\sum_{e\in E} c_e}(R_S, R_V^{\text{key}}, c_E)$ is achievable for $\mathbb{I}$.

The other direction follows a similar argument. Suppose that $\frac{1}{\sum_{e\in E} c_e}(R_S, R_V^{\text{key}}, c_E)$ is achievable for $\mathbb{I}$. Then, there exists a sequence of index codes $\left(\left[(2^{\left\lceil \widehat{n}\left(\frac{R_s}{\sum_{e\in E} c_e}\right)\right\rceil}, s \in S), (2^{\left\lceil \widehat{n}\left(\frac{R_v^{\text{key}}}{\sum_{e\in E} c_e}\right)\right\rceil}, v \in V), (2^{\left\lceil \widehat{n}\left(\frac{c_i}{\sum_{e\in E} c_e}\right)\right\rceil}, i \in E)\right], \widehat{n}\right)$, $\widehat{n} \in \ell(\sum_{e\in E} c_e)\mathbb{Z}^+$, with probability of decoding $\eta \to 0$ and leakage $\eta \to 0$ as $\widehat{n} \to \infty$.



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

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

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

Fig. 5: 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 $\mathbb{I}$.

From Theorem 2, it follows that there exists a sequence of network codes $\left(2^{\lceil nR_s \rceil}, n\right)$, $n = \frac{\widehat{n}}{\sum_{e \in E} c_e} \in \ell\mathbb{Z}^+$, with probability of decoding error of $(|Z|+1)\epsilon \to 0$ and leakage $\theta_2 \to 0$, $\eta \to 0$ and $n \to \infty$. Hence, the rate tuple R_S is achievable for $\mathbb{N}$. ■

C. An Example

Consider the network-coding instance $\mathbb{N}$ depicted in Figure 5a, its augmented version $\mathbb{N}'$ in Figure 5b, and the mapped index-coding instance $\mathbb{I}$ in Figure 5c.

One randomised zero-error zero-leakage network code for $\mathbb{N}$ is $X_{e_1} = \mathsf{e}_{e_1}(X_1, Y_1) = Y_1$ and $X_{e_2} = \mathsf{e}_{e_2}(X_1, Y_1) = X_1 + Y_1$, where $Y_1 \in [M_1]$ (that is, $K_1 = M_1$). Note that without using the random key Y_1, it is not possible to protect X_1 from the eavesdroppers.

This network code is then translated to a deterministic network code for $\mathbb{N}'$ with three messages $\{X'_1, X'_2, X'_3\}$, with $X'_3 = \alpha$ set to be a constant, as follows: $X'_{e_1} = \mathsf{g}'_{e_1}(X'_1, X'_2, X'_3) = X'_2$ and $X'_{e_2} = \mathsf{g}'_{e_2}(X'_1, X'_2, X'_3) = X'_1 + X'_2$. The augmentation from $\mathbb{N}$ to $\mathbb{N}'$ changes neither the decodability of the users nor the leakage to the eavesdroppers.

In the index-coding instance $\mathbb{I}$, two receivers $\{\widehat{t}_{e_1}, \widehat{t}_{e_2}\}$ correspond to the links in $\mathbb{N}'$ and one receiver $\widehat{t}_2$ corresponds to the destination node in $\mathbb{N}'$. The translated index code is

$$\widehat{X}_{\mathrm{b}} = \widehat{\mathsf{e}}(\widehat{X}_{[3]\cup E}) = (\widehat{X}_{\mathrm{b},e_1}, \widehat{X}_{\mathrm{b},e_2}) = (\widehat{X}_{e_1} + \mathsf{g}'_{e_1}(\widehat{X}_{[3]}), \widehat{X}_{e_1} + \mathsf{g}'_{e_2}(\widehat{X}_{[3]})) = (\widehat{X}_{e_1} + \widehat{X}_2, \widehat{X}_{e_2} + \widehat{X}_1 + \widehat{X}_2).$$

One can verify that all receivers can decode their required messages, and each eavesdropper gains no information about $\widehat{X}_1$.

In the other direction, consider the following secure index code (recall that $\widehat{X}_3 = \alpha$ is a constant):

$$\widehat{X}_{\mathrm{b}} = \widehat{\mathsf{e}}(\widehat{X}_{[3]\cup E}) = (\widehat{\mathsf{e}}_1(\widehat{X}_{[3]\cup E}), \widehat{\mathsf{e}}_2(\widehat{X}_{[3]\cup E})) = (\widehat{X}_{e_1} + \beta_1\widehat{X}_1 + \beta_2\widehat{X}_2, \widehat{X}_{e_2} + \gamma_1\widehat{X}_1 + \gamma_2\widehat{X}_2) := (\widehat{X}_{\mathrm{b},1}, \widehat{X}_{\mathrm{b},2}),$$

for some non-zero β_2, γ_2 such that $\beta_1/\beta_2 \neq \gamma_1/\gamma_2$. The decoding function of nodes $\widehat{t}_{e_1}, \widehat{t}_{e_2}, \widehat{t}_2$ are respectively

$$\begin{aligned}
&\widehat{\mathsf{d}}_{\widehat{t}_{e_1}}(\widehat{X}_{\mathrm{b}}, \widehat{X}_{\widehat{H}_{\widehat{t}_{e_1}}}) = \widehat{\mathsf{d}}_{\widehat{t}_{e_1}}(\widehat{X}_{\mathrm{b}}, \widehat{X}_1, \widehat{X}_2) = \widehat{X}_{\mathrm{b},1} - \beta_1\widehat{X}_1 - \beta_2\widehat{X}_2,\\
&\widehat{\mathsf{d}}_{\widehat{t}_{e_2}}(\widehat{X}_{\mathrm{b}}, \widehat{X}_{\widehat{H}_{\widehat{t}_{e_2}}}) = \widehat{\mathsf{d}}_{\widehat{t}_{e_2}}(\widehat{X}_{\mathrm{b}}, \widehat{X}_1, \widehat{X}_2) = \widehat{X}_{\mathrm{b},2} - \gamma_1\widehat{X}_1 - \gamma_2\widehat{X}_2,\\
&\widehat{\mathsf{d}}_{\widehat{t}_2}(\widehat{X}_{\mathrm{b}}, \widehat{X}_{\widehat{H}_{\widehat{t}_2}}) = \widehat{\mathsf{d}}_{\widehat{t}_2}(\widehat{X}_{\mathrm{b}}, \widehat{X}_{e_1}, \widehat{X}_{e_2}) = \left(\frac{\widehat{X}_{\mathrm{b},1} - \widehat{X}_{e_1}}{\beta_2} - \frac{\widehat{X}_{\mathrm{b},2} - \widehat{X}_{e_2}}{\gamma_2}\right)\left(\frac{\beta_1}{\beta_2} - \frac{\gamma_1}{\gamma_2}\right)^{-1}.
\end{aligned}$$

The translated deterministic network code for $\mathbb{N}'$ is as follows:

$$X'_{e_1} = \mathsf{e}_{e_1}(X'_{\mathtt{in}(\mathtt{tail}(e_1))}, X'_{O'^{-1}(\mathtt{tail}(e_1))}) = \widehat{\mathsf{d}}_{\widehat{t}_{e_1}}(0, X'_1, X'_2) = -\beta_1 X'_1 - \beta_2 X'_2, \tag{11}$$

$$X'_{e_2} = \mathsf{e}_{e_2}(X'_{\mathtt{in}(\mathtt{tail}(e_2))}, X'_{O'^{-1}(\mathtt{tail}(e_2))}) = \widehat{\mathsf{d}}_{\widehat{t}_{e_2}}(0, X'_1, X'_2) = -\gamma_1 X'_1 - \gamma_2 X'_2. \tag{12}$$

In $\mathbb{N}'$, the destination (node 2) can recover X'_1 by choosing $\mathsf{d}_2(X'_{\mathtt{in}(2)}, X'_{O'^{-1}(2)}) = \widehat{\mathsf{d}}_{\widehat{t}_2}(0, X'_{e_1}, X'_{e_2})$. Also, each eavesdropper gains no information about X'_1 as β_2 and γ_2 are non-zero. Lastly, by replacing X'_2 with Y_1, we obtain a zero-error zero-leakage network code for $\mathbb{N}$.

VIII. Proof of Theorem 2 – Part 1 (the forward direction)

We will now prove Theorem 2 for the forward direction: if $\mathbb{N}$ is (M_S, ϵ, η, n)-feasible, then $\mathbb{I}$ is $((M_S, K_V, 2^{\lfloor c_E n \rfloor}), \epsilon, \theta_1, \widehat{n})$-feasible with a deterministic index code.

A. Code Construction

Recall that $S = [k]$ and $V = [\ell]$. Also recall that each $\widehat{X}_e$, $e \in E$, is chosen to be independently and uniformly distributed over $[\widehat{M}_e] = [2^{\lfloor c_e n \rfloor}]$.

Note that $\mathbb{N}$ is (M_S, ϵ, η, n)-feasible iff $\mathbb{N}'$ is $((M_S, K_V), \epsilon, \eta, n)$–feasible, where the random key Y_i at each node $i \in V$ in the network code for $\mathbb{N}$ is realised using an additional independent source X'_{k+i} at the same node $\mathbb{N}'$.

From a network code for $\mathbb{N}'$, we will use the non-secure code translation to construct an index code for $\mathbb{I}$, where the sender broadcasts $\widehat{X}_{\mathrm{b}} = (\widehat{X}_{\mathrm{b},e} : e \in E)$, consisting of

$$\widehat{X}_{\mathrm{b},e} = \widehat{X}_e + \mathsf{g}'_e(\widehat{X}_{[k+\ell]}) \mod 2^{\lfloor c_e n \rfloor}. \tag{13}$$

Note that each $\widehat{X}_e, \mathsf{g}'_e \in [2^{\lfloor c_e n \rfloor}]$, and therefore $\widehat{X}_{\mathrm{b}} \in [\prod_{e\in E} 2^{\lfloor c_e n \rfloor}] = [2^{\sum_{e\in E}\lfloor c_e n \rfloor}] = [2^{\widehat{n}}]$.

B. Decoding Criterion

See Appendix C for the proof of the decoding criterion, which is similar to that of the non-secure equivalence.

C. Security Criteria

Given $\frac{1}{n} I(X_{A_z}; X_{B_z}) \le \eta$ for $\mathbb{N}'$, we need to show $\frac{1}{n} I(\widehat{X}_{\widehat{A}_z}; \widehat{X}_{\mathrm{b}}, \widehat{X}_{\widehat{B}_z}) \le \eta$ for $\mathbb{I}$.

We now consider the security constraints. For each $z \in \widehat{Z}$,

$$
\begin{aligned}
H(\widehat{X}_{\widehat{A}_z}|\widehat{X}_{\mathrm{b}}, \widehat{X}_{\widehat{B}_z}) &= H(\widehat{X}_{\widehat{A}_z}|\{\widehat{X}_{\mathrm{b},e} : e \in E\}, \{\widehat{X}_{e'} : e' \in \widehat{B}_z\}) && \text{(14a)}\\
&= H(\widehat{X}_{\widehat{A}_z}|\{\widehat{X}_{\mathrm{b},e} : e \in \widehat{B}_z\}, \{\widehat{X}_{e'} : e' \in \widehat{B}_z\}) && \text{(14b)}\\
&= H(\widehat{X}_{\widehat{A}_z}|\{\widehat{X}_{\mathrm{b},e}, \widehat{X}_e, \mathsf{g}'_e(\widehat{X}_{[k+\ell]}) : e \in \widehat{B}_z\}) && \text{(14c)}\\
&= H(\widehat{X}_{\widehat{A}_z}|\{\widehat{X}_e, \mathsf{g}'_e(\widehat{X}_{[k+\ell]}) : e \in \widehat{B}_z\}) && \text{(14d)}\\
&= H(\widehat{X}_{\widehat{A}_z}|\{\mathsf{g}'_e(\widehat{X}_{[k+\ell]}) : e \in \widehat{B}_z\}) && \text{(14e)}\\
&= H(\widehat{X}_{A_z}|\{\mathsf{g}'_e(\widehat{X}_{[k+\ell]}) : e \in B_z\}) && \text{(14f)}\\
&= H(X'_{A_z}|\{\mathsf{g}'_e(X'_{[k+\ell]}) : e \in B_z\}) && \text{(14g)}\\
&= H(X'_{A_z}|X'_{B_z}) = H(X_{A_z}|X_{B_z}), && \text{(14h)}
\end{aligned}
$$

where (14b) follows from the Markov chain

$$
\widehat{X}_{\widehat{A}_z} - \Big(\{\widehat{X}_{\mathrm{b},e} : e \in \widehat{B}_z\}, \{\widehat{X}_{e'} : e' \in \widehat{B}_z\}\Big) - \Big(\{\widehat{X}_{\mathrm{b},e} : e \notin \widehat{B}_z\}\Big),
$$

where $\{\widehat{X}_{\mathrm{b}}(e) : e \notin \widehat{B}_z\}$ are independent of $\Big(\widehat{X}_{\widehat{A}_z}, \{\widehat{X}_{\mathrm{b},e} : e \in \widehat{B}_z\}, \{\widehat{X}_{e'} : e' \in \widehat{B}_z\}\Big)$, because the former has been randomised by independently and uniformly distributed $\{\widehat{X}_e : e \notin \widehat{B}_z\}$ (which are independent of $(\widehat{X}_{\widehat{A}_z}, \widehat{X}_{\widehat{B}_z}, \widehat{X}_{[k+\ell]})$, see (13));

(14c) follows from (13);

(14d) is derived because $\widehat{X}_{\mathrm{b},e}$ is a deterministic function of $(\widehat{X}_e, \mathsf{g}'_e(\widehat{X}_{[k+\ell]}))$;

(14e) follows from the Markov chain $\widehat{X}_{\widehat{A}_z} - \{\mathsf{g}'_e(\widehat{X}_{[k+\ell]}) : e \in \widehat{B}_z\} - \{\widehat{X}_e : e \in \widehat{B}_z\}$, which can be derived from noting that $\{\widehat{X}_e : e \in E\}$ are independent of $(\widehat{X}_{\widehat{A}_z}, \widehat{X}_{[k+\ell]})$;

(14g) follows from a change of variables (from hatted to dashed) as $(\widehat{X}_{[k+\ell]}, \widehat{X}_{A_z})$ and $(X'_{[k+\ell]}, X'_{A_z})$ have

the same distribution;

(14h) is obtained from noting that $\{\mathsf{g}'_e(X'_{[k+\ell]}) : e \in B_z\} = X'_{B_z}$.

Now, for $\mathbb{N}$, if $I(X_{A_z}; X_{B_z}) \leq \eta$, then

$$\frac{1}{\widehat{n}} I(\widehat{X}_{\widehat{A}_z}; \widehat{X}_{\mathrm{b}}, \widehat{X}_{\widehat{B}_z}) = \frac{1}{\widehat{n}} \left[H(\widehat{X}_{\widehat{A}_z}) - H(\widehat{X}_{\widehat{A}_z} | \widehat{X}_{\mathrm{b}}, \widehat{X}_{\widehat{B}_z}) \right] \tag{15a}$$

$$= \frac{1}{\widehat{n}} \left[H(\widehat{X}_{\widehat{A}_z}) - H(X_{A_z} | X_{B_z}) \right] \tag{15b}$$

$$= \frac{1}{\widehat{n}} \left[H(X_{A_z}) - H(X_{A_z} | X_{B_z}) \right] \tag{15c}$$

$$= \frac{1}{\sum_{e \in E} \lfloor c_e n \rfloor} I(X_{A_z}; X_{B_z}) \tag{15d}$$

$$\leq \frac{1}{(n \sum_{e \in E} c_e - |E|)} I(X_{A_z}; X_{B_z}) \tag{15e}$$

$$< \frac{\eta}{\left(\sum_{e \in E} c_e - \frac{|E|}{n} \right)} := \theta_1, \tag{15f}$$

where (15b) follows from (14h), and (15c) follows from $X_{[k]}$ and $\widehat{X}_{[k]}$ having the same distribution. Recall that $c_e n \geq 1$, which ensures that none of the links is degenerated (that is, cannot carry any information).

Combining the decodability and the security results, the index code is $((M_S, K_V, 2^{\lfloor c_E n \rfloor}), \epsilon, \theta_1, \widehat{n})$-feasible.

IX. Proof of Theorem 2 – Part 2 (the backward direction)

We will now prove Theorem 2 for the backward direction: if $\mathbb{I}$ is $\left((M_S, K_V, 2^{\lfloor c_E n \rfloor}), \epsilon, \eta, \widehat{n}\right)$-feasible with a deterministic index code, then $\mathbb{N}$ is $(M_S, (|Z|+1)\epsilon, \theta_2, n)$-feasible.

We only need to show the result from deterministic index codes for $\mathbb{I}$ to deterministic network codes for $\mathbb{N}'$. The code translation from $\mathbb{N}'$ to $\mathbb{N}$ is straightforward. By substituting each $X'_{[k+i]} \in [K_i]$, $i \in [\ell]$, with a random key $Y_i \in [K_i]$, we conclude that if $\mathbb{N}'$ is $((M_{[k]}, K_{[\ell]}), \epsilon, \eta, n)$-feasible using a deterministic code, then $\mathbb{N}$ is $(M_{[k]}, \epsilon, \eta, n)$-feasible using a randomised code. This is possible as $X'_{[k+i]}$ originates at node i and is not required by any node or any eavesdropper.

A. Code construction

We will start with the non-secure code translation, which translates any index code for $\mathbb{I}$ to a network code for $\mathbb{N}'$ using a parameter σ. The translated network code consists of the following:

- An encoding function for each link $e \in E$ such that $\mathsf{e}_e(x_{\widehat{H}_{\widehat{t}_e}}) = \widehat{\mathsf{d}}_{\widehat{t}_e}(\sigma, x_{\widehat{H}_{\widehat{t}_e}}) \in [\widehat{M}_e] = [2^{\lfloor c_e n \rfloor}]$,

- A decoding function for each destination $i \in U$ such that $\mathsf{d}_i(x_{\mathtt{in}(i)\cup O^{-1}(i)}) = \widehat{\mathsf{d}}_{\widehat{t}_i}(\sigma, x_{\mathtt{in}(i)\cup O^{-1}(i)}) \in [\widehat{M}_i]$. Recall that U is the set of destination nodes in $\mathbb{N}'$.

The blocklength of the network code is n.

In $\mathbb{I}$, σ is the broadcast message $\widehat{x}_{\mathrm{b}} = \widehat{\mathsf{e}}(\widehat{x}_{[k+\ell]\cup E})$, which depends on *all* messages $\widehat{x}_{[k+\ell]\cup E}$.

B. A Summary of Our Approach for the Choice of σ

This direction of the equivalence is the most challenging amongst the four. The difficulty here is to select a suitable $\sigma \in [2^{\widehat{n}}]$ for $\mathbb{N}'$ that satisfies both decodability and secrecy, but cannot be made dependent on the messages—because each node only knows a subset of messages.

Without security, a good candidate σ exists. Using this fixed σ (which is independent of the message realisation), the translated network code has a probability of decoding error of at most ϵ. This particular choice was found by considering network codes with the parameter σ randomly chosen according to a random variable Σ uniformly distributed over $[2^{\widehat{n}}]$.

With security, for the special case of perfect decoding and no leakage, we established [16] that the same σ found using the above method can be used to translate an index code with $\epsilon = \eta = 0$ to a network code with $\epsilon = \eta = 0$. This approach relies on an observation that if for message realisations that can be decoded correctly in $\mathbb{I}$, not only is decoding correct in $\mathbb{N}$, but the leakage is also preserved.

In general, when decoding is incorrect in $\mathbb{I}$, the leakage in $\mathbb{N}'$ does not match that in $\mathbb{I}$. To avoid this problem, we consider instances of correct decoding in $\mathbb{I}$ and look at the distribution of the broadcast message conditioned on correct decoding.

Our approach is to design a collection of network codes, where σ for the network codes is randomly chosen based on this probability. We can show that, on average, $\mathbb{N}'$ will have the desired probability of decoding error and leakage. From there, we will show the existence of a good candidate for σ.

C. An Important Property of the Broadcast Message in $\mathbb{I}$

Central to the proof of the code translation in this direction is the following properties of the broadcast message in $\mathbb{I}$:

Proposition 1: Fix any broadcast message $\widehat{x}_{\mathrm{b}} \in [2^{\widehat{n}}]$ and any realisation $\widehat{x}_{[k+\ell]}$. If all receivers $\widehat{T}$ can decode their requested messages correctly, then there can be at most one realisation $\widehat{x}_E$ for which $\widehat{\mathsf{e}}(\widehat{x}_{[k+\ell]}, \widehat{x}_E) = \widehat{x}_{\mathrm{b}}$.

Proposition 1 was proven for a slightly different network-to-index coding map [3], which includes an additional receiver $\widehat{t}_{\text{all}}$ in $\mathbb{I}$, where $\widehat{H}_{\widehat{t}_{\text{all}}} = [k+\ell]$ and $\widehat{W}_{\widehat{t}_{\text{all}}} = E$. We will show that the proposition remains true even without $\widehat{t}_{\text{all}}$, by taking into account the fact that graph G (which was defined for $\mathbb{N}$ from which $\mathbb{I}$ has been mapped) is acyclic.

Proof of Proposition 1: Fix any $\widehat{x}_{[k+\ell]}$ and $\widehat{x}_{\text{b}}$. The decoding function of each receiver $\widehat{t}_e$, $e \in E$, is $\widehat{\mathsf{d}}_{\widehat{t}_e}(\widehat{x}_{\text{b}}, \widehat{x}_{\texttt{in}(\texttt{tail}(e))\cup O'^{-1}(\texttt{tail}(e))})$, where $\texttt{in}(\texttt{tail}(e)) \subsetneq E$ are in the upstream of e, and $O'^{-1}(\texttt{tail}(e)) \subseteq [k+\ell]$.

Since the decoding for all receivers are correct, for each $e \in E$, $\widehat{t}_e$ recovers $\widehat{x}_e$ by using the function $\widehat{\mathsf{d}}_{\widehat{t}_e}$. As G is acyclic, by considering decoding functions $\widehat{\mathsf{d}}_{\widehat{t}_e}$ starting from *root* nodes, that is, links e where $\texttt{in}(\texttt{tail}(e)) = \emptyset$, and traversing the links in the directions of the links, all *link messages* $\widehat{x}_E$ are completely determined by $\widehat{x}_{[k+\ell]}$ and $\widehat{x}_{\text{b}}$. ■

Suppose in $\mathbb{I}$ that a message realisation $\widehat{x}_{[k+\ell]\cup E}$ results in correct decoding. From the proof of Proposition 1, we know that given $\widehat{x}_{[k+\ell]}$ and $\widehat{\mathsf{e}}(\widehat{x}_{[k+\ell]\cup E})$, a sequence of receiver decoding functions can collectively recover $\widehat{x}_E$. So, using the above network-code translation, we have the following observation:

Observation 1: Suppose that $\widehat{x}_{[k+\ell]\cup E}$ results in correct decoding in $\mathbb{I}$. We use the translated network code in $\mathbb{N}'$. For the message realisation $x_{[k+\ell]} = \widehat{x}_{[k+\ell]}$ in $\mathbb{N}'$, if $\sigma = \widehat{\mathsf{e}}(\widehat{x}_{[k+\ell]\cup E})$ had been chosen for the network code, then each link $e \in E$ will send $x_e = \widehat{x}_e$ (since its encoding function is derived from the decoding function in $\mathbb{I}$), and the decoding of all receivers in $\mathbb{N}'$ will be correct (since decoding in $\mathbb{I}$ is all correct).

D. Choosing a Distribution for σ

Recall that each network code is specified by the choice of σ. We first randomly select σ, independent of the messages, according to some probability mass function p_Σ.

Remark 6: This approach requires all nodes V to know the selected σ. This can be implemented by a random public key, which the eavesdroppers may also access. However, the use of a randomised σ is only an intermediate step for us to prove the existence of a good candidate. The final result will be based on a particular pre-chosen σ.

To simplify notation, let

- $m := \prod_1^{k+\ell} M_i$ denote the total number of message realisations of $\widehat{X}_{[k+\ell]}$.
- $d := 2^{\widehat{n}} = 2^{\sum_{e\in E}\lfloor c_e n\rfloor}$ denote the total number of message realisations of $\widehat{X}_E$.

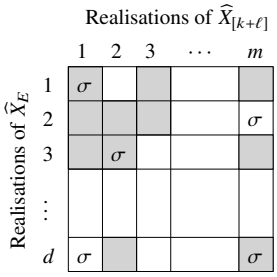


Fig. 6: A table showing the value of $\widehat{\mathsf{e}}(\widehat{x}_{[k+\ell]\cup E}) \in [2^{\widehat{n}}]$ for each message realisation $(\widehat{x}_{[k+\ell]}, \widehat{x}_E)$. Shaded cells indicate message realisations that result in correct decoding for all receivers in $\mathbb{I}$.

Recall that $\widehat{\mathsf{e}}(x_{[k+\ell]\cup E}) \in [2^{\widehat{n}}]$. Figure 6 shows the broadcast message $\widehat{\mathsf{e}}(x_{[k+\ell]\cup E})$ for each message realisation $(\widehat{x}_{[k+\ell]}, \widehat{x}_E)$. Note that, due to Proposition 1, $\widehat{\mathsf{e}}(x_{[k+\ell]\cup E})$ corresponding to all shaded cells in any column must be distinct.

Let N_σ be the number of realisations $\widehat{x}_{[k+\ell]\cup E}$ that results in correct decoding for all receivers and $\widehat{\mathsf{e}}(x_{[k+\ell]\cup E}) = \sigma$. In Figure 6, N_σ is the total number of shaded cells labelled as σ. Define $\bar{\epsilon}$ as the fraction of unshaded cells. As the messages are uniformly distributed, $\bar{\epsilon}$ is also the probability of decoding error $\widehat{P}_{\mathrm{e}}$ in $\mathbb{I}$. It is easy to see the following:

$$\sum_{\sigma\in[d]} N_\sigma = (1-\bar{\epsilon})md, \tag{16}$$

$$\sum_{\sigma\in[d]} \frac{1}{d}\frac{N_\sigma}{m} = 1-\bar{\epsilon}. \tag{17}$$

Define a new random variable $\widehat{C} \in \{0,1\}$ in $\mathbb{I}$, where $\widehat{C} = 1$ if decoding is correct, and $\widehat{C} = 0$ otherwise. Now, consider a translated network code, where σ is the realisation of a random variable Σ whose distribution is given by

$$p_\Sigma(\sigma) = \frac{N_\sigma}{(1-\bar{\epsilon})md} = p_{\widehat{X}_{\mathrm{b}}|\widehat{C}}(\sigma|1), \tag{18}$$

where the second equality is obtained by observing that the messages in $\mathbb{I}$ are uniformly distributed.

E. Decodability Criterion using Randomly Chosen σ

Let $P'_{\mathrm{e},\sigma}$ be the probability of decoding error in $\mathbb{N}'$ when σ is chosen for the network code. From Observation 1, if σ is chosen, for any message realisation $x_{[k+\ell]}$ such that the column $\widehat{x}_{[k+\ell]} = x_{[k+\ell]}$ in Figure 6 contains σ in a shaded cell, decoding in $\mathbb{N}'$ is correct. From Proposition 1, each σ can appear at most once over the shaded cells in each column. So, the probability of correct decoding in $\mathbb{N}'$ is

$$1 - P'_{\mathrm{e},\sigma} \geq \frac{N_\sigma}{m}. \tag{19}$$

Then, averaged over σ, the probability of correct decoding in $\mathbb{N}'$ is

$$1 - P'_{\mathrm{e}} = \sum_{\sigma \in [d]} p_{\Sigma}(\sigma)(1 - P'_{\mathrm{e},\sigma}) \tag{20a}$$

$$\geq \sum_{\sigma \in [d]} \frac{N_\sigma}{(1-\bar{\epsilon})md} \frac{N_\sigma}{m} = \frac{1}{1-\bar{\epsilon}} \sum_{\sigma \in [d]} \frac{1}{d} \left(\frac{N_\sigma}{m} \right)^2 \tag{20b}$$

$$\overset{\text{(a)}}{\geq} \frac{1}{1-\bar{\epsilon}} \left(\sum_{\sigma \in [d]} \frac{1}{d} \frac{N_\sigma}{m} \right)^2 \overset{\text{(b)}}{=} 1 - \bar{\epsilon} \overset{\text{(c)}}{\geq} 1 - \epsilon, \tag{20c}$$

where (a) is obtained using Jensen's inequality; (b) follows from (17); (c) follows from the fact that the probability of decoding error in $\mathbb{I}$ is $\widehat{P}_{\mathrm{e}} = \bar{\epsilon} \leq \epsilon$.

Thus, the randomised translated network code has a probability of decoding error $P'_{\mathrm{e}} \leq \epsilon$.

F. Security Criteria using Randomly Chosen σ

From the identity $I(\mathtt{P};\mathtt{Q}|\mathtt{X}) + I(\mathtt{P};\mathtt{R}|\mathtt{Q},\mathtt{X}) = I(\mathtt{P};\mathtt{R}|\mathtt{X}) + I(\mathtt{P};\mathtt{Q}|\mathtt{R},\mathtt{X})$, we get

$$I(\mathtt{P};\mathtt{Q}|\mathtt{X}) = I(\mathtt{P};\mathtt{Q}|\mathtt{R},\mathtt{X}) + I(\mathtt{P};\mathtt{R}|\mathtt{X}) - I(\mathtt{P};\mathtt{R}|\mathtt{Q},\mathtt{X}) \tag{21}$$

$$\geq I(\mathtt{P};\mathtt{Q}|\mathtt{R},\mathtt{X}) - I(\mathtt{P};\mathtt{R}|\mathtt{Q},\mathtt{X}) \geq I(\mathtt{P};\mathtt{Q}|\mathtt{R},\mathtt{X}) - H(\mathtt{R}). \tag{22}$$

Similarly,

$$I(\mathtt{P};\mathtt{Q}|\mathtt{R},\mathtt{X}) \geq I(\mathtt{P};\mathtt{Q}|\mathtt{X}) - H(\mathtt{R}). \tag{23}$$

Consider eavesdropper $z \in Z$. Let $B := \widehat{B}_z = B_z$ and $A := \widehat{A}_z = A_z$, where we drop the subscripts to ease notation. Starting with an index code that has a leakage of at most η,

$$\begin{aligned} \widehat{n}\eta &\geq I(\widehat{X}_A; \widehat{X}_{\mathrm{b}}, \widehat{X}_B) \geq I(\widehat{X}_A; \widehat{X}_B|\widehat{X}_{\mathrm{b}}) \geq I(\widehat{X}_A; \widehat{X}_B|\widehat{X}_{\mathrm{b}}, \widehat{C}) - H(\widehat{C}) \geq I(\widehat{X}_A; \widehat{X}_B|\widehat{X}_{\mathrm{b}}, \widehat{C}) - H_{\mathrm{b}}(\epsilon) \\ &= \sum_{\sigma} p_{\widehat{X}_{\mathrm{b}},\widehat{C}}(\sigma, 1) I(\widehat{X}_A; \widehat{X}_B|\widehat{X}_{\mathrm{b}} = \sigma, \widehat{C} = 1) + \sum_{\sigma} p_{\widehat{X}_{\mathrm{b}},\widehat{C}}(\sigma, 0) I(\widehat{X}_A; \widehat{X}_B|\widehat{X}_{\mathrm{b}} = \sigma, \widehat{C} = 0) - H_{\mathrm{b}}(\epsilon). \end{aligned} \tag{24}$$

To map the above to $\mathbb{N}'$, we define $C \in \{0, 1\}$ as a random variable for $\mathbb{N}'$ as follows. For a specific chosen σ and message realisation $x_{[k+\ell]}$, $C = 1$ iff σ appears in a shaded cell in the column $\widehat{x}_{[k+\ell]} = x_{[k+\ell]}$ in Figure 6. $C = 1$ implies that the decoding in $\mathbb{N}'$ is correct (but the reverse is not always true). Noting that N_σ/m is the fraction of message realisations whose column contains σ in a shaded cell in Figure 6, we have

$$p_C(1) = \sum_{\sigma} p_{\Sigma}(\sigma) N_\sigma/m \geq 1 - \bar{\epsilon} = p_{\widehat{C}}(1) \geq 1 - \epsilon, \tag{25}$$

where (25) follows from (20a)–(20c). This means

$$\frac{p_{\widehat{C}}(1)}{1-\epsilon} \geq 1 \geq P_C(1) \quad \Rightarrow \quad p_{\widehat{C}}(1) \geq (1-\epsilon)p_C(1). \tag{26}$$

From Proposition 1, if a message realisation $(\widehat{x}_{[k+\ell]}, \widehat{x}_E)$ results in correct decoding in $\mathbb{I}$ (that is, it maps to a shaded cell in Figure 6), then $\widehat{x}_E = \phi_\sigma(\widehat{x}_{[k+\ell]})$ for some deterministic function ϕ_σ, where $\sigma = \widehat{\mathsf{e}}(\widehat{x}_{[k+\ell]}, \widehat{x}_E)$. Now, suppose that $\sigma = \widehat{\mathsf{e}}(\widehat{x}_{[k+\ell]}, \widehat{x}_E)$ is chosen for the network code. From Observation 1, if $x_{[k+\ell]} = \widehat{x}_{[k+\ell]}$ is transmitted, then $x_E = \phi_\sigma(x_{[k+\ell]})$. Then, for any (a, σ) such that $p_{\widehat{X}_{[k+\ell]}, \widehat{X}_\mathrm{b}, \widehat{C}}(a, \sigma, 1) > 0$, and any b, we have

$$p_{X_E|X_{[k+\ell]},\Sigma,C}(b|a,\sigma,1) = p_{\widehat{X}_E|\widehat{X}_{[k+\ell]},\widehat{X}_\mathrm{b},\widehat{C}}(b|a,\sigma,1). \tag{27}$$

As the messages $X_{[k+\ell]}$ in $\mathbb{I}$ and $\widehat{X}_{[k+\ell]}$ in $\mathbb{N}'$ are uniformly distributed, we see from Figure 6 that for any σ with $N_\sigma > 0$,

$$p_{X_{[k+\ell]}|\Sigma,C}(a|\sigma,1) = p_{\widehat{X}_{[k+\ell]}|\widehat{X}_\mathrm{b},\widehat{C}}(a|\sigma,1) = \begin{cases} 1/N_\sigma, & \text{if a shaded cell in column } a \text{ contains } \sigma, \\ 0, & \text{otherwise.} \end{cases}$$

So, for any σ with $N_\sigma > 0$, a, and b, we get the following:

$$p_{X_{[k+\ell]},X_E|\Sigma,C}(a,b|\sigma,1) = p_{\widehat{X}_{[k+\ell]},\widehat{X}_E|\widehat{X}_\mathrm{b},\widehat{C}}(a,b|\sigma,1), \tag{28}$$

which then necessitates that

$$I(\widehat{X}_A; \widehat{X}_B|\widehat{X}_\mathrm{b} = \sigma, \widehat{C} = 1) = I(X_A; X_B|\Sigma = \sigma, C = 1). \tag{29}$$

Substituting (29) into (24), we get

$$\widehat{n}\eta + H_\mathrm{b}(\epsilon) \geq \sum_\sigma p_{\widehat{X}_\mathrm{b},\widehat{C}}(\sigma,1) I(X_A; X_B|\Sigma = \sigma, C = 1) \tag{30a}$$

$$= p_{\widehat{C}}(1) \sum_\sigma p_{\widehat{X}_\mathrm{b}|\widehat{C}}(\sigma|1) I(X_A; X_B|\Sigma = \sigma, C = 1) \tag{30b}$$

$$= p_{\widehat{C}}(1) I(X_A; X_B|\Sigma, C = 1) \tag{30c}$$

$$\geq (1-\epsilon) p_C(1) I(X_A; X_B|\Sigma, C = 1), \tag{30d}$$

where (30c) is obtained noting (18); (30d) follows from (26).

We are now ready to bound the leakage in $\mathbb{N}'$, where Σ is known to all nodes. First, we have

$$I(X_A; X_B, \Sigma) = I(X_A; X_B|\Sigma) \tag{31a}$$

$$\leq I(X_A; X_B|\Sigma, C) + H(C) \tag{31b}$$

$$= p_C(1) I(X_A; X_B|\Sigma, C=1) + p_C(0) I(X_A; X_B|\Sigma, C=0) + H(C) \tag{31c}$$

$$\leq p_C(1) I(X_A; X_B|\Sigma, C=1) + \epsilon n \sum_{e\in E} c_e + H_\mathrm{b}(\epsilon) \tag{31d}$$

$$\leq \frac{\widehat{n}\eta + H_\mathrm{b}(\epsilon)}{1-\epsilon} + \epsilon n \sum_{e\in E} c_e + H_\mathrm{b}(\epsilon), \tag{31e}$$

where (31a) follows since Σ is chosen independent of the messages; (31b) follows from (23); (31d) follows from $p_C(0) \leq \epsilon$, $I(X_A; X_B|\Sigma, C=0) \leq H(X_B) \leq H(X_E) = \sum_{e\in E} \log_2 M_e \leq n \sum_{e\in E} c_e$, and $H(C) \leq H_\mathrm{b}(\epsilon)$; (31e) follows from (30d).

Thus, the randomised translated network code has a leakage of

$$\frac{1}{n} I(X_A; X_B|\Sigma) < \frac{1}{n}\left(\frac{\widehat{n}\eta + H_\mathrm{b}(\epsilon)}{1-\epsilon} + H_\mathrm{b}(\epsilon)\right) + \epsilon \sum_{e\in E} c_e \tag{32a}$$

$$\leq \left(\frac{\eta}{1-\epsilon} + \epsilon\right) \sum_{e\in E} c_e + \frac{1}{n}\left(\frac{H_\mathrm{b}(\epsilon)}{1-\epsilon} + H_\mathrm{b}(\epsilon)\right) := \theta, \tag{32b}$$

where $\widehat{n} = \sum_{e\in E} \lfloor c_e n \rfloor \leq n \sum_{e\in E} c_e$, and $n \geq 1$.

G. Existence of a Candidate σ for Decodability and Security

Averaged over all realisations of Σ, inequalities (20c) and (32b) give

$$\sum_{\sigma\in[d]} p_\Sigma(\sigma) P'_{\mathrm{e},\sigma} \leq \epsilon, \tag{33}$$

$$\sum_{\sigma\in[d]} p_\Sigma(\sigma) \frac{1}{n} I(X_{A_z}; X_{B_z}|\Sigma=\sigma) \leq \theta, \quad \text{for each } z \in Z. \tag{34}$$

Invoking Markov's inequality, we get the following for some $\lambda > 0$:

$$\Pr\left[\underbrace{P'_{\mathrm{e},\sigma} \geq (|Z|+1+\lambda)\epsilon}_{:=V_0(\lambda)}\right] \leq \frac{1}{|Z|+1+\lambda}, \tag{35}$$

$$\Pr\left[\underbrace{\frac{1}{n} I(X_{A_z}; X_{B_z}|\Sigma=\sigma) \geq (|Z|+1+\lambda)\theta}_{:=V_z(\lambda)}\right] \leq \frac{1}{|Z|+1+\lambda}, \quad \text{for each } z \in Z. \tag{36}$$

Using the union bound, we get

$$\Pr\left[\bigcup_{z\in Z\cup\{0\}} V_z(\lambda)\right] \leq \frac{|Z|+1}{|Z|+1+\lambda}, \tag{37}$$

$$\text{or equivalently,} \qquad \Pr\left[\bigcap_{z\in Z\cup\{0\}} V_z^{\mathrm{c}}(\lambda)\right] \geq 1 - \frac{|Z|+1}{|Z|+1+\lambda} > 0. \tag{38}$$

Now, the alphabet of Σ is $[d]$, which is finite. Also, for each $\lambda > 0$, there exists a σ such that $\bigcap_{z\in Z\cup\{0\}} V_z(\lambda)^{\mathrm{c}}$ holds. Consequently, there must exist a σ for which $\bigcap_{z\in Z\cup\{0\}} V_z(0)^{\mathrm{c}}$ holds, that is, there exists a network code (using a particular σ) for which

$$P'_{\mathrm{e}} < (|Z|+1)\epsilon, \tag{39}$$

$$\frac{1}{n} I(X_{A_z}; X_{B_z}) < (|Z|+1)\theta := \theta_2, \tag{40}$$

for all $z \in Z$. This completes the proof of Part 2 of Theorem 2. ■

REFERENCES

[1] R. Dougherty and K. Zeger, "Nonreversibility and equivalent constructions of multiple-unicast networks," IEEE Trans. Inf. Theory, vol. 52, no. 11, pp. 1982–1986, Nov. 2006.

[2] W. Huang, T. Ho, M. Langberg, and J. Kliewer, "On secure network coding with uniform wiretap sets," in Proc. IEEE Int. Symp. on Netw. Coding (NetCod), Calgary, Canada, June 7–9 2013.

[3] 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.

[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] 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] 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.

[7] 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.

[8] 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.

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

[10] L. Ong, B. N. Vellambi, P. L. Yeoh, J. Kliewer, and J. Yuan, "Secure index coding: Existence and construction," in Proc. IEEE Int. Symp. Inf. Theory (ISIT), Barcelona Spain, July 10–15 2016, pp. 2834–2838.

[11] 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.

[12] A. Blasiak, R. Kleinberg, and E. Lubetzky, "Broadcasting with side information: Bounding and approximating the broadcast rate," IEEE Trans. Inf. Theory, vol. 59, no. 9, pp. 292–298, Sept. 2013.

[13] S. Unal and A. B. Wagner, "A rate-distortion approach to index coding," IEEE Trans. Inf. Theory, vol. 62, no. 11, pp. 6359–6378, Nov. 2016.

[14] F. Arbabjolfaei, B. Bandemer, Y.-H. Kim, E. Şaşoğlu, and L. Wang, "On the capacity region for index coding," in Proc. IEEE Int. Symp. Inf. Theory (ISIT), Istanbul, Turkey, July 7–12 2013, pp. 962–966.

[15] K. Shanmugam, A. G. Dimakis, and M. Langberg, "Local graph coloring and index coding," in Proc. IEEE Int. Symp. Inf. Theory (ISIT), Istanbul, Turkey, July 7–12 2013, pp. 1152–1156.

[16] 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.

Appendix A
Proof of the Decodability Criterion for Part 1 of Theorem 1

Note that in the network-coding instance $\mathbb{N}$, only receivers $\{t_i\colon i \in [\ell]\}$ need to decode messages, and each t_i receives $X_{\mathtt{in}(t_i)} = (X_{\widehat{H}_i}, X_{2\to t_i})$ over its incoming links, where $X_{2\to t_i} = X_{1\to 2} = \widehat{\mathsf{e}}(X_S, Y_1)$. These are the same functions that each receiver $i \in \widehat{T}$ receives in the index-coding instance $\mathbb{I}$. By using the same decoding functions for receivers $\{t_i\colon i \in \widehat{T}\}$ in $\mathbb{N}$, if $\widehat{P}_{\mathrm{e}} \le \epsilon$ for $\mathbb{I}$, we also must have $P_{\mathrm{e}} \le \epsilon$ for $\mathbb{N}$.

Appendix B
Proof of the Decodability Criterion for Part 2 of Theorem 1

In $\mathbb{N}$, receiver t_i, for each $i \in [\ell]$, tries to decode $X_{\widehat{W}_i}$ using the decoding function $\mathsf{d}_{t_i}\left(X_{\mathtt{in}(t_i)}\right) = \mathsf{d}_{t_i}\Big(X_{2\to t_i}, \Big(X_{s_j\to t_i} : j \in \widehat{H}_i\Big)\Big)$, where

- $X_{2\to t_i} = X_{1\to 2} = \mathsf{e}_{1\to 2}\left(\left(\mathsf{e}_{s_j\to 1}(X_j) : j \in [k]\right), Y_1\right)$,
- $X_{s_j\to t_i} = \mathsf{e}_{s_j\to t_i}(X_j)$ for $j \in \widehat{H}_i$.

The premise states that $P_{\mathrm{e}} \le \epsilon$ for $\mathbb{N}$.

In $\mathbb{I}$, according to the code translation, receiver i, for each $i \in [\ell]$, tries to decode $\widehat{X}_{\widehat{W}_i}$ using $\widehat{\mathsf{d}}_i(\widehat{X}_{\mathrm{b}}, \widehat{X}_{\widehat{H}_i})$, where

- $\widehat{X}_{\mathrm{b}} = \mathsf{e}_{1\to 2}\left(\left(\mathsf{e}_{s_j\to 1}(\widehat{X}_j) : j \in [k]\right), \widehat{Y}\right)$,
- $\widehat{X}_{\widehat{H}_i} = \left(\mathsf{e}_{s_j\to t_i}(\widehat{X}_j) : j \in \widehat{H}_i\right)$

Since the decoding functions in $\mathbb{I}$ exactly match those in $\mathbb{N}$, and since $(\widehat{X}_{[k]}, \widehat{Y})$, and $(X_{[k]}, Y_1)$ have the same distribution, we must have $\widehat{P}_{\mathrm{e}} \le \epsilon$ for $\mathbb{I}$.

Appendix C
Proof of the Decodability Criterion for Part 1 of Theorem 2

In $\mathbb{N}$, according to definition (1), with probability of at least $(1-\epsilon)$ (over the message realisations $x_{\mathcal{S}}$), every vertex $v \in U$ can decode all messages that it requires from the message on all incoming links

and messages originating at v. Recall that only messages $X'_{[k]}$ of all messages $X'_{[k+\ell]}$ in $\mathbb{N}'$ need to be decoded. Suppose that, every node $v \in U$ can decode its required messages correctly with probability of at least $(1-\epsilon_v)$, that is,

$$\Pr\left\{X'_{D'^{-1}(v)} = \mathsf{d}'_v(X'_{\mathtt{in}(v)}, X'_{O'^{-1}(v)})\right\} \geq 1-\epsilon_v, \tag{41}$$

or equivalently,

$$\Pr\left\{X'_{D'^{-1}(v)} = \mathsf{d}'_v([\mathsf{g}'_e(X'_{[k+\ell]})]_{e\in\mathtt{in}(v))}, X'_{O'^{-1}(v)})\right\} \geq 1-\epsilon_v. \tag{42}$$

For $\mathbb{I}$, we first consider receivers $\hat{t}_i \in \widehat{T}$ where $i \in U$. As mentioned above, while source messages $X'_{O'^{-1}(v)}$ in $\mathbb{N}'$ and $\hat{X}_{O'^{-1}(v)}$ in $\mathbb{I}$ have the same distribution, link messages $X'_{\mathtt{in}(v)}$ in $\mathbb{N}'$ and $\hat{X}_{\mathtt{in}(v)}$ in $\mathbb{I}$ may not. So, although node $\hat{t}_i \in \hat{T}$ in $\mathbb{I}$ has side information $(\hat{X}_{\mathtt{in}(i)}, \hat{X}'_{O'^{-1}(i)})$, using (41) in $\mathbb{I}$ will not work, as the pmf of $X'_{[k+\ell]\cup E}$ is different from that of $\hat{X}_{[k+\ell]\cup E}$.

To deal with this issue, we use (42), which requires $[\mathsf{g}'_e(\hat{X}_{[k+\ell]})]_{e\in\mathtt{in}(i)}$ instead. This will work because $(X'_{[k+\ell]}, [\mathsf{g}'_e(X'_{[k+\ell]})]_{e\in E}) \overset{\mathrm{d}}{=} (\hat{X}_{[k+\ell]}, [\mathsf{g}'_e(\hat{X}_{[k+\ell]})]_{e\in E})$.

In $\mathbb{I}$, $\widehat{H}_{\hat{t}_i} = \mathtt{in}(i) \cup O'^{-1}(i)$. Receiver $\hat{t}_i$ knows $\hat{X}_{O'^{-1}(i)}$ and calculates $[\mathsf{g}'_e(\hat{X}_{[k+\ell]})]_{e\in\mathtt{in}(i)}$ from the broadcast message $\hat{X}_{\mathrm{b}}$ and side information $\hat{X}_{\mathtt{in}(i)}$ using (13). So, using (42) with a change of variables (from non-hatted to hatted), we have

$$\Pr\left\{\hat{X}_{\widehat{W}_{\hat{t}_i}} = \hat{X}_{D'^{-1}(i)} = \mathsf{d}'_i([\mathsf{g}'_e(\hat{X}_{[k+\ell]})]_{e\in\mathtt{in}(i))}, \hat{X}_{O'^{-1}(i)})\right\} \geq 1-\epsilon_i.$$

This means every receiver $\hat{t}_i \in \widehat{T}$ where $i \in U$ can correctly decode its required messages with probability of at least $(1-\epsilon_i)$.

Now, we consider receivers $\hat{t}_e \in \widehat{T}$ where $e \in E$. Recall that $\widehat{H}_{\hat{t}_e} = \mathtt{in}(\mathtt{tail}(e)) \cup O'^{-1}(\mathtt{tail}(e))$, and $\widehat{W}_{\hat{t}_e} = \{e\}$. Receiver $\hat{t}_e$ performs the following steps:

1) Knowing $\{\hat{X}_d : d \in \mathtt{in}(\mathtt{tail}(e))\}$, it obtains $\{\mathsf{g}'_d(\hat{X}_{[k+\ell]}) : d \in \mathtt{in}(\mathtt{tail}(e))\}$ from (13).
2) Knowing $\hat{X}_{O'^{-1}(\mathtt{tail}(e))}$ as side information, it calculates $\mathsf{g}'_e(\hat{X}_{[k+\ell]}) = \mathsf{e}'_e([\mathsf{g}'_d(\hat{X}_{[k+\ell]}) : d \in \mathtt{in}(\mathtt{tail}(e))], \hat{X}_{O'^{-1}}$
3) With $\mathsf{g}'_e(\hat{X}_{[k+\ell]})$ and the broadcast message $\hat{X}_{\mathrm{b},e}$, it obtains the required $\hat{X}_e$ using (13).

So, every receiver $\hat{t}_e \in \widehat{T}$, $e \in E$, must be able to correctly decode the required $\hat{X}_e$ without error.

Combining these two classes of receivers, we have shown that all receivers in $\mathbb{I}$ can correctly decode their required messages with probability of at least $(1-\epsilon)$.