An Equivalence between Network Coding and Index Coding

M. Effros
California Institute of Technology
effros@caltech.edu

S. El Rouayheb
Princeton University
salim@princeton.edu

M. Langberg
The Open University of Israel
mikel@openu.ac.il

***Abstract*—We show that the network coding and index coding problems are equivalent. This equivalence holds in the general setting which includes *linear and non-linear* codes. Specifically, we present an efficient reduction that maps a network coding instance to an index coding instance while preserving feasibility. Previous connections were restricted to the linear case.**

I. Introduction

In the network coding paradigm, a set of source nodes transmits information to a set of terminal nodes over a network; internal nodes of the network may mix received information before forwarding it. This mixing (or encoding) of information has been extensively studied over the last decade (see e.g., [1], [2], [3], [4], [5], and references therein). While network coding in the *multicast* setting is well understood, this is far from being the case for the general multi-source, multi-terminal setting. In particular, determining the capacity of a general network coding instance remains an intriguing, central, open problem, e.g., [6], [7], [8].

A special instance of the network coding problem introduced in [9], which has seen significant interest lately, is the so-called *index coding* problem [9], [10], [11], [12], [13], [14]. The index coding problem encapsulates the "broadcast with side information" problem in which a single server wishes to communicate with several clients, each requiring potentially different information and having potentially different side information (as shown by the example in Fig. 1(a)).

One may think of the index coding problem as a *simple* and *representative* instance of the network coding problem. The instance is "simple" in the sense that any index coding instance can be represented as a topologically simple network coding instance in which only a *single* internal node has in-degree greater than one and thus only a single internal node can perform encoding (see Fig. 1(b) for an example). It is "representative" in the sense that the index coding paradigm is broad enough to characterize the network coding problem under the assumption of *linear* coding [15]. Specifically, given any instance $\mathcal{I}$ of the network coding problem, one can efficiently construct an instance $\hat{\mathcal{I}}$ of the index coding problem such that: (a) There exists a linear solution to $\mathcal{I}$ if and only if there exists a linear solution to $\hat{\mathcal{I}}$, and (b) any linear solution to $\hat{\mathcal{I}}$ can be efficiently turned into a linear solution to $\mathcal{I}$.[1]

This material is based upon work supported by ISF grant 480/08, BSF grant 2010075, and NSF grants CCF-1018741 and CCF-1016671. The work was done while Michael Langberg was visiting the California Institute of Technology. Authors appear in alphabetical order.

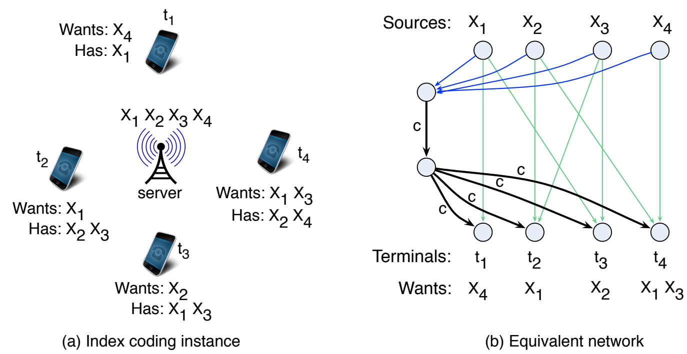


Fig. 1. (a) An instance of the index coding problem. A server has 4 binary sources $X_1, \ldots, X_4$ and there are 4 terminals with different "wants" and "has" sets (corresponding to the communication demand and side information, respectively). The server can sequentially transmit the four sources to all four terminals using four broadcasts. However, this is not optimal. It is sufficient to broadcast only 2 bits, namely $X_1 + X_2 + X_3$ and $X_1 + X_4$, where '+' denotes the xor operation. (b) Index coding is a special case of the network coding problem. In this example all links are of unit capacity (non-specified) or of capacity c. Links directly connecting between sources and terminals represent the "has" sets. Any solution to the index coding problem with c broadcast bits can be efficiently mapped to a solution to the corresponding network coding instance and vice versa. This implies that the index coding problem is a special case of the network coding problem. The focus of this work is on the opposite assertion. Namely, that the network coding problem is a special case of the index coding problem.

The results of [15] hold for (scalar and vector) linear coding functions only, and the analysis there breaks down once one allows general coding (which may be non-linear) at internal nodes. The study of non-linear coding functions is central to the study of network coding since it is shown in [16] that there exist instances in which linear codes do not suffice to achieve capacity (in the general multi-source, multi-terminal setting).

In this work, we extend the equivalence between network coding and index coding to the setting of general encoding and decoding functions. Our results effectively imply that when one wishes to solve a network coding instance $\mathcal{I}$, a possible route is to turn the network coding instance into an index coding instance $\hat{\mathcal{I}}$ (via our reduction), solve the index coding instance $\hat{\mathcal{I}}$, and turn the solution to $\hat{\mathcal{I}}$ into a solution to the original network coding instance $\mathcal{I}$. Hence, any efficient scheme to solve index coding would yield an efficient scheme

[1] Notions such as "solution," "feasibility," and "capacity" that are used in this section are defined in Section II.

to solve network coding. Stated differently, our results imply that understanding the solvability of index coding instances would imply an understanding of the solvability of network coding instances as well.

The remainder of the paper is structured as follows. In Section II, we present the models of network and index coding. In Section III, we present an example based on the "butterfly network" that illustrates our proof techniques. In Section IV, we present the main technical contribution of this work: the equivalence between network and index coding. Finally, in Section V, we conclude with some remarks and open problems. Several proofs are omitted due to space limitations and appear in an extended version [17].

II. Model

In what follows, we define the model for the network coding and index coding problems. Throughout this paper, "hatted" variables (e.g., $\hat{x}$) correspond to the variables of index coding instances, while "unhatted" variables correspond to the network coding instance. For any $k > 0$, $[k] = \{1, \ldots, \lfloor k \rfloor\}$.

A. Network coding

An instance $\mathcal{I} = (G, S, T, B)$ of the network coding problem includes a directed acyclic network $G = (V, E)$, a set of sources nodes $S \subset V$, a set of terminal nodes $T \subset V$, and an $|S| \times |T|$ requirement matrix B. We assume, without loss of generality, that each source $s \in S$ has no incoming edges and that each terminal $t \in T$ has no outgoing edges. Let c_e denote the capacity of each edge $e \in E$, namely for any block length n, each edge e can carry one of the messages in $[2^{c_e n}]$. In our setting, each source $s \in S$ holds a rate R_s random variable X_s uniformly distributed over $[2^{R_s n}]$ and independent from all other sources.

A network code, $(\mathcal{F}, \mathcal{G}) = (\{f_e\}, \{g_t\})$, is an assignment of an encoding function f_e to each edge $e \in E$ and a decoding function g_t to each terminal $t \in T$. For $e = (u, v)$, f_e is a function taking as input the random variables associated with incoming edges to node u; the random variable corresponding to e, $X_e \in [2^{c_e n}]$, is the random variable equal to the evaluation of f_e on its inputs. If e is an edge leaving a source node $s \in S$, then X_s is the input to f_e. The input to the decoding function g_t consists of the random variables associated with incoming edges to terminal t. The output of g_t is a vector of reproductions of all sources required by t (the latter are specified by the matrix B defined below).

Given the acyclic structure of G, edge messages $\{X_e\}$ can be defined by induction on the topological order of G. Namely, given the functions $\{f_e\}$, one can define functions $\{\bar{f}_e\}$ such that each $\bar{f}_e$ takes as its input the information sources $\{X_s\}$ and transmits as its output the random variable X_e.[2] More precisely, for $e = (u, v)$ in which u is a source node, define $\bar{f}_e \equiv f_e$. For $e = (u, v)$ in which u is an internal node with incoming edges $In(e) = \{e'_1, \ldots, e'_\ell\}$, define $\bar{f}_e \equiv f_e(\bar{f}_{e'_1}, \ldots, \bar{f}_{e'_\ell})$. Namely, the evaluation of $\bar{f}_e$ on source information $\{X_s\}$ equals the evaluation of f_e given the values of $\bar{f}_{e'}$ for $e' \in In(e)$. We use both $\{f_e\}$ and $\{\bar{f}_e\}$ in our analysis below.

[2] In the network coding literature, $\{f_e\}$ and $\{\bar{f}_e\}$ are sometimes referred to as the local and global encoding functions, respectively.

The $|S| \times |T|$ requirement matrix $B = [b_{s,t}]$ has entries in the set $\{0, 1\}$, with $b_{s,t} = 1$ if and only if terminal t requires information from source s.

A network code $(\mathcal{F}, \mathcal{G})$ is said to satisfy terminal node t under transmission $(x_s : s \in S)$ if the output of decoding function g_t equals $(x_s : b(s, t) = 1)$ when $(X_s : s \in S) = (x_s : s \in S)$. The network code $(\mathcal{F}, \mathcal{G})$ is said to satisfy the instance $\mathcal{I}$ with error probability $\varepsilon \geq 0$ if the probability that all $t \in T$ are simultaneously satisfied is at least $1 - \varepsilon$. The probability is taken over the joint distribution on random variables $(X_s : s \in S)$.

For a rate vector $R = (R_1, \ldots, R_{|S|})$, an instance $\mathcal{I}$ to the network coding problem is said to be (ε, R, n)-feasible if there exists a network code $(\mathcal{F}, \mathcal{G})$ with block length n that satisfies $\mathcal{I}$ with error probability at most ε when applied to source information $(X_1, \ldots, X_{|S|})$ uniformly distributed over $\Pi_{s=1}^{n}[2^{R_s n}]$. An instance $\mathcal{I}$ to the network coding problem is said to be R-feasible if for any $\varepsilon > 0$ and any $\delta > 0$ there exists a block length n such that $\mathcal{I}$ is $(\varepsilon, R(1-\delta), n)$-feasible. Here, $R(1-\delta) = (R_1(1-\delta), \ldots, R_{|S|}(1-\delta))$. The capacity region of an instance $\mathcal{I}$ refers to all rate vectors R for which $\mathcal{I}$ is R-feasible.

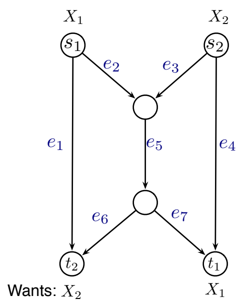
B. Index coding

An instance $\hat{\mathcal{I}} = (\hat{S}, \hat{T}, \{\hat{W}_{\hat{t}}\}, \{\hat{H}_{\hat{t}}\})$ of the index coding problem includes a set of sources $\hat{S} = \{\hat{s}_1, \hat{s}_2, \ldots, \hat{s}_{|\hat{S}|}\}$ all available at a single server, and a set of terminals $\hat{T} = \{\hat{t}_1, \ldots, \hat{t}_{|\hat{T}|}\}$. Given a block length n, each source $\hat{s} \in \hat{S}$ holds a rate $\hat{R}_{\hat{s}}$ random variable $\hat{X}_{\hat{s}}$ uniformly distributed over $[2^{\hat{R}_{\hat{s}} n}]$ (and independent from other sources). Each terminal requires information from a certain subset of sources in $\hat{S}$. In addition, information from some sources in $\hat{S}$ are available a priori as side information to each terminal. Specifically, terminal $\hat{t} \in \hat{T}$ is associated with sets:

- $\hat{W}_{\hat{t}}$ which is the set of sources required by $\hat{t}$, and
- $\hat{H}_{\hat{t}}$ which is the set of sources available at $\hat{t}$.

We refer to $\hat{W}_{\hat{t}}$ and $\hat{H}_{\hat{t}}$ as the "wants" and "has" sets of $\hat{t}$, respectively. The server uses an error-free broadcast channel to transmit information to the terminals. The objective is to design an encoding scheme that satisfies the demands of all the terminals while minimizing the number of uses of the broadcast channel. (See Fig. 1.)

An index code $(\hat{\mathcal{F}}, \hat{\mathcal{G}}) = (\hat{f}_B, \{\hat{g}_{\hat{t}}\}_{\hat{t} \in \hat{T}})$ for $\hat{\mathcal{I}}$, with broadcast rate $\hat{c}_B$, includes an encoding function $\hat{f}_B$ for the broadcast channel and a set of decoding functions $\hat{\mathcal{G}} = \{\hat{g}_{\hat{t}}\}_{\hat{t} \in \hat{T}}$ with one function for each terminal. The function $\hat{f}_B$ takes as input the source random variables $\{\hat{X}_{\hat{s}}\}$ and returns a rate $\hat{c}_B$ random variable $\hat{X}_B \in [2^{\hat{c}_B n}]$. The input to the decoding function $\hat{g}_{\hat{t}}$ consists of the random variables in $\hat{H}_{\hat{t}}$ (the source random variables available to $\hat{t}$) and the broadcast message $\hat{X}_B$. The output of $\hat{g}_{\hat{t}}$ is the reconstruction by terminal $\hat{t}$ of all sources required by $\hat{t}$ (and described by $\hat{W}_{\hat{t}}$).



(a) Butterfly Network

$\hat{X}_1, \hat{X}_2$
$\hat{X}_{e_1}, \ldots, \hat{X}_{e_7}$

Terminal	Wants	Has	Terminal	Wants	Has
$\hat{t}_{e_1}$	$\hat{X}_{e_1}$	$\hat{X}_1$	$\hat{t}_{e_6}$	$\hat{X}_{e_6}$	$\hat{X}_{e_5}$
$\hat{t}_{e_2}$	$\hat{X}_{e_2}$	$\hat{X}_1$	$\hat{t}_{e_7}$	$\hat{X}_{e_7}$	$\hat{X}_{e_5}$
$\hat{t}_{e_3}$	$\hat{X}_{e_3}$	$\hat{X}_2$	$\hat{t}_1$	$\hat{X}_1$	$\hat{X}_{e_4}\hat{X}_{e_7}$
$\hat{t}_{e_4}$	$\hat{X}_{e_4}$	$\hat{X}_2$	$\hat{t}_2$	$\hat{X}_2$	$\hat{X}_{e_1}\hat{X}_{e_6}$
$\hat{t}_{e_5}$	$\hat{X}_{e_5}$	$\hat{X}_{e_2}\hat{X}_{e_3}$	$\hat{t}_{all}$	$\hat{X}_{e_1} \ldots \hat{X}_{e_7}$	$\hat{X}_1 \hat{X}_2$

(b) Equivalent Index Coding Instance

Fig. 2. (a) The butterfly network with two sources s_1 and s_2 and two terminals t_1 and t_2. (b) The equivalent index coding instance. The server has 9 sources: one for each source, namely $\{\hat{X}_1, \hat{X}_2\}$, and one for each edge in the network, namely $\{\hat{X}_{e_1}, \ldots, \hat{X}_{e_7}\}$. There are 7 terminals corresponding to the 7 edges in the network, 2 terminals corresponding to the two terminals of the butterfly network and one extra terminal $\hat{t}_{all}$.

An index code $(\hat{\mathcal{F}}, \hat{\mathcal{G}})$ of broadcast rate $\hat{c}_B$ is said to satisfy terminal $\hat{t}$ under transmission $(\hat{x}_{\hat{s}} : \hat{s} \in \hat{S})$ if the output of decoding function $\hat{g}_{\hat{t}}$ equals $(\hat{x}_{\hat{s}} : \hat{s} \in \hat{W}_{\hat{t}})$ when $(\hat{X}_{\hat{s}} : \hat{s} \in \hat{S}) = (\hat{x}_{\hat{s}} : \hat{s} \in \hat{S})$. Index code $(\hat{\mathcal{F}}, \hat{\mathcal{G}})$ is said to satisfy instance $\hat{\mathcal{I}}$ with error probability $\varepsilon \geq 0$ if the probability that all $\hat{t} \in \hat{T}$ are simultaneously satisfied is at least $1 - \varepsilon$. The probability is taken over the joint distribution on random variables $\{\hat{X}_{\hat{s}}\}_{\hat{s} \in \hat{S}}$.

For a rate vector $\hat{R} = (\hat{R}_1, \ldots, \hat{R}_{|\hat{S}|})$, an instance $\hat{\mathcal{I}}$ to the index coding problem is said to be $(\varepsilon, \hat{R}, \hat{c}_B, n)$-feasible if there exists an index code $(\hat{\mathcal{F}}, \hat{\mathcal{G}})$ with broadcast rate $\hat{c}_B$ and block length n that satisfies $\hat{\mathcal{I}}$ with error probability at most ε when applied to source information $(\hat{X}_{\hat{s}} : \hat{s} \in \hat{S})$ uniformly and independently distributed over $\Pi_{\hat{s} \in \hat{S}}[2^{\hat{R}_{\hat{s}} n}]$. An instance $\hat{\mathcal{I}}$ to the network coding problem is said to be $(\hat{R}, \hat{c}_B)$-feasible if for any $\varepsilon > 0$ and $\delta > 0$ there exists a block length n such that $\hat{\mathcal{I}}$ is $(\varepsilon, \hat{R}(1-\delta), \hat{c}_B, n)$-feasible. As before, $\hat{R}(1-\delta) = (\hat{R}_1(1-\delta), \ldots, \hat{R}_{|\hat{S}|}(1-\delta))$. The capacity region of an instance $\hat{\mathcal{I}}$ with broadcast rate $\hat{c}_B$ refers to all rate vectors $\hat{R}$ for which $\hat{\mathcal{I}}$ is $(\hat{R}, \hat{c}_B)$-feasible.

III. Example

Our main result, formally stated as Theorem 1 in Section IV, states that the network coding and index coding problems are equivalent. The proof is based on a reduction that constructs, for any given network coding problem, an equivalent index coding problem. In this section, we explain the main elements of our proof by applying it to the butterfly network example [1] shown in Fig. 2(a). For simplicity, our example does not consider any error in communication. Our reduction is similar to the construction in [15]; our analysis differs to capture the case of non-linear encoding.

Following the notation in Section II-A, let $X_{e_i} = \bar{f}_{e_i}(X_1, X_2)$ be the one-bit message on edge e_i of the butterfly network. Then, the following is a network code that satisfies the demands of the terminals: $X_{e_1} = X_{e_2} = X_1$, $X_{e_3} = X_{e_4} = X_2$, $X_{e_5} = X_{e_6} = X_{e_7} = X_1 + X_2$, where '+' denotes the xor operation. Terminal t_1 can decode X_1 by computing $X_1 = X_{e_4} + X_{e_7}$, and t_2 can decode X_2 by computing $X_2 = X_{e_1} + X_{e_6}$. Thus, the butterfly network is $(\epsilon, R, n) = (0, (1,1), 1)$-feasible.

The problem now is to construct an index coding instance that is "equivalent" to the butterfly network; equivalence here means that any index code for that instance would imply a network code for the butterfly network, and vice versa. We propose the construction, based on that presented in [15], in which the server has 9 sources split into two sets, and 10 terminals, as described in Fig. 2.

Next, we explain how the solutions are mapped between these two instances. "Direction 1" strongly follows the analysis appearing in [15]; our major innovation is in "Direction 2". (Both proof directions are presented below for completeness.)

Direction 1: Network code to index code. Suppose we are given a network code with local encoding functions f_{e_i} and global encoding functions $\bar{f}_{e_i}, i = 1, \ldots, 7$. In our index coding solution the server broadcasts the 7-bit vector $\hat{X}_B = (\hat{X}_B(e_1), \ldots, \hat{X}_B(e_7))$, where

$$\hat{X}_B(e_i) = \hat{X}_{e_i} + \bar{f}_{e_i}(\hat{X}_1, \hat{X}_2), \quad i = 1, \ldots, 7. \tag{1}$$

One can check that this index code allows each terminal to recover the sources in its "wants" set using the broadcast $\hat{X}_B$ and the information in its "has" set (see [15]).

Direction 2: Index code to network code. Let $\hat{c}_B$ equal the total capacity of edges in the butterfly network, i.e., $\hat{c}_B = 7$. Suppose we are given an index code with broadcast rate $\hat{c}_B$ that allows each terminal to decode the sources it requires (with no errors). We want to show that any such code can be mapped to a network code for the butterfly network. Let us denote by $\hat{X}_B = (\hat{X}_{B,1}, \ldots, \hat{X}_{B,7})$ the broadcast information where $\hat{X}_B$ is a (possibly non-linear) function of the 9 sources available at the server: $\hat{X}_1, \hat{X}_2$ and $\hat{X}_{e_1}, \ldots, \hat{X}_{e_7}$.

For every terminal $\hat{t}$, there exists a decoding function $\hat{g}_{\hat{t}}$ that takes as input the broadcast information $\hat{X}_B$ and the sources in its "has" set and returns the sources it requires. For example $\hat{g}_{\hat{t}_{e_1}}(\hat{X}_B, \hat{X}_1) = \hat{X}_{e_1}$. We use these decoding functions to construct the network code for the butterfly network. Consider for example edge e_5. Its incoming edges are e_2 and e_3, so we need to define a local encoding f_{e_5} which is a function of the information X_{e_2} and X_{e_3} they are carrying. In our approach, we fix a specific value σ for $\hat{X}_B$, and define

$$f_{e_5}(X_{e_2}, X_{e_3}) = \hat{g}_{\hat{t}_{e_5}}(\sigma, X_{e_2}, X_{e_3}).$$

Similarly, we define the encoding functions for every edge in the butterfly network, and the decoding functions for the two terminals t_1 and t_2 by applying the corresponding decoder to the received inputs and the fixed value of σ. The crux of our proof lies in showing that there exists a value of σ for which the corresponding network code allows correct decoding. In the example at hand, one may choose σ to be the all zero vector $\mathbf{0}$. (In this example, all values of σ are equally good.)

To prove correct decoding, we show that for any fixed values $\hat{X}_1 = \hat{x}_1$ and $\hat{X}_2 = \hat{x}_2$, there exists a unique value for the vector $(\hat{X}_{e_1}, \ldots, \hat{X}_{e_7})$ that results in the broadcast transmission of $\hat{X}_B = \mathbf{0}$. (Recall that $\hat{X}_B$ is a function of $\hat{X}_1, \hat{X}_2$ and $\hat{X}_{e_1}, \ldots, \hat{X}_{e_7}$.) Otherwise, since $\hat{c}_B = 7$, $\hat{t}_{all}$ cannot decode correctly. Roughly speaking, this correspondence allows us to reduce the analysis of correct decoding in the network code to correct decoding in the index code. Details on this reduction and the choice of σ appear in the next section.

IV. Main Result

Theorem 1: For any instance $\mathcal{I}$ of the network coding problem, one can efficiently construct an instance $\hat{\mathcal{I}}$ of the index coding problem and an integer $\hat{c}_B$ such that for any rate vector R, any integer n, and any $\varepsilon \geq 0$ it holds that $\mathcal{I}$ is (ε, R, n) feasible if and only if $\hat{\mathcal{I}}$ is $(\varepsilon, \hat{R}, \hat{c}_B, n)$ feasible. Here, the rate vector $\hat{R}$ for $\hat{\mathcal{I}}$ can be efficiently computed from R and $\mathcal{I}$, and the network and index codes that imply feasibility in the reduction can be efficiently constructed from one another.

In words, Theorem 1 states that for any network coding instance $\mathcal{I}$ one can efficiently construct an index coding instance $\hat{\mathcal{I}}$ that *preserves feasibility*. Specifically, for any feasible rate vector R, our reduction allows the construction of rate R network codes for $\mathcal{I}$ by studying index codes for $\hat{\mathcal{I}}$.

Proof: Let $G = (V, E)$ and $\mathcal{I} = (G, S, T, B)$. Let n be any integer, and let $R = (R_1, \ldots, R_{|S|})$. We start by defining $\hat{\mathcal{I}} = (\hat{S}, \hat{T}, \{\hat{W}_{\hat{t}}\}, \{\hat{H}_{\hat{t}}\})$, the integer $\hat{c}_B$, and the rate vector $\hat{R}$. See Fig. 2 for an example. To simplify notation, we use the notation $\hat{X}_{\hat{s}}$ to denote both the source $\hat{s} \in \hat{S}$ and the corresponding random variable. For $e = (u, v)$ in E, let $In(e)$ be the set of edges entering u in G. If u is a source s let $In(e) = \{s\}$. For $t_i \in T$, let $In(t_i)$ be the set of edges entering t_i in G.

Set $\hat{S}$ consists of $|S|+|E|$ sources: one source, denoted $\hat{X}_s$, for each source $s \in S$ from $\mathcal{I}$, and one source, denoted $\hat{X}_e$, for each edge $e \in E$. Thus, $\hat{S} = \{\hat{X}_s\}_{s \in S} \cup \{\hat{X}_e\}_{e \in E}$. Set $\hat{T}$ consists of $|E| + |T| + 1$ terminals: $|E|$ terminals, denoted $\hat{t}_e$, corresponding to the edges in E, $|T|$ terminals, denoted $\hat{t}_i$, corresponding to the terminals in $\mathcal{I}$, and a single terminal, denoted $\hat{t}_{all}$. Thus, $\hat{T} = \{\hat{t}_e\}_{e \in E} \cup \{\hat{t}_i\}_{i \in [|T|]} \cup \{\hat{t}_{all}\}$. For each $\hat{t}_e \in \hat{T}$, we set $\hat{H}_{\hat{t}_e} = \{\hat{X}_{e'}\}_{e' \in In(e)}$ and $\hat{W}_{\hat{t}_e} = \{\hat{X}_e\}$. For each $\hat{t}_i \in \hat{T}$, let t_i be the corresponding terminal in T. We set $\hat{H}_{\hat{t}_i} = \{\hat{X}_{e'}\}_{e' \in In(t_i)}$ and $\hat{W}_{\hat{t}_i} = \{\hat{X}_s\}_{s : b(s, t_i) = 1}$. For $\hat{t}_{all}$ set $\hat{H}_{\hat{t}_{all}} = \{\hat{X}_s\}_{s \in S}$ and $\hat{W}_{\hat{t}_{all}} = \{\hat{X}_e\}_{e \in E}$. Let $\hat{R}$ be a vector of length $|S|+|E|$ consisting of two parts: $(\hat{R}_s : \ s \in S)$ represents the rate $\hat{R}_s$ of each $\hat{X}_s$ and $(\hat{R}_e : \ e \in E)$ represents the rate $\hat{R}_e$ of each $\hat{X}_e$. Set $\hat{R}_s = R_s$ for each $s \in S$ and $\hat{R}_e = c_e$ for each $e \in E$. (Here R_s is the entry corresponding to s in the vector R, and c_e is the capacity of the edge e in G.) Finally, set $\hat{c}_B = \sum_{e \in E} c_e = \sum_{e \in E} \hat{R}_e$.

We now present our proof. The fact that $\mathcal{I}$ is (ε, R, n) feasible implies that $\hat{\mathcal{I}}$ is $(\varepsilon, \hat{R}, \hat{c}_B, n)$ feasible is essentially shown in [15] and is omitted here due to space limitations. Full details appear in [17]. The other direction is the major technical contribution of this work.

$\hat{\mathcal{I}}$ is $(\varepsilon, \hat{R}, \hat{c}_B, n)$ feasible implies that $\mathcal{I}$ is (ε, R, n) feasible: We assume that $\hat{\mathcal{I}}$ is $(\varepsilon, \hat{R}, \hat{c}_B, n)$ feasible with $\hat{c}_B = \sum_{e \in E} c_e = \sum_{e \in E} \hat{R}_e$ (as defined above). Thus, there exists an index code $(\hat{\mathcal{F}}, \hat{\mathcal{G}}) = (\hat{f}_B, \{\hat{g}_{\hat{t}}\})$ for $\hat{\mathcal{I}}$ with block length n and success probability at least $1-\varepsilon$. In what follows, we obtain a network code $(\mathcal{F}, \mathcal{G}) = \{f_e\} \cup \{g_t\}$ for $\mathcal{I}$. The key observation we use is summarized below by Claim 1 which follows from our definition of $\hat{c}_B = \sum_{e \in E} \hat{R}_e$. The proof of Claim 1 appears in the full version of this work [17].

We start with some notation. For each realization $\hat{\mathbf{x}}_{\mathbf{S}} = \{\hat{x}_s\}_{s \in S}$ of source information $\{\hat{X}_s\}_{s \in S}$ in $\hat{\mathcal{I}}$, let $A_{\hat{\mathbf{x}}_{\mathbf{S}}}$ be the realizations $\hat{\mathbf{x}}_{\mathbf{E}} = \{\hat{x}_e\}_{e \in E}$ of $\{\hat{X}_e\}_{e \in E}$ for which all terminals decode correctly. That is, if we use the term "good" to refer to any source realization pair $(\hat{\mathbf{x}}_{\mathbf{S}}, \hat{\mathbf{x}}_{\mathbf{E}})$ for which all terminals decode correctly, then $A_{\hat{\mathbf{x}}_{\mathbf{S}}} = \{\hat{\mathbf{x}}_{\mathbf{E}} \mid$ the pair $(\hat{\mathbf{x}}_{\mathbf{S}}, \hat{\mathbf{x}}_{\mathbf{E}})$ is good$\}$.

Claim 1: There exists a $\sigma \in [2^{\hat{c}_B n}]$ such that at least a $(1-\varepsilon)$ fraction of source realizations $\hat{\mathbf{x}}_{\mathbf{S}}$ satisfy $\hat{f}_B(\hat{\mathbf{x}}_{\mathbf{S}}, \hat{\mathbf{x}}_{\mathbf{E}}) = \sigma$ for some $\hat{\mathbf{x}}_{\mathbf{E}} \in A_{\hat{\mathbf{x}}_{\mathbf{S}}}$.

We now define the encoding and decoding functions of $(\mathcal{F}, \mathcal{G})$ for the network code instance $\mathcal{I}$. Specifically, we define the encoding functions $\{f_e\}$ and the decoding functions $\{g_t\}$ for the edges e in E and terminals t in T (where E and T are the edge set and terminal set of the network coding instance $\mathcal{I}$). We next formally define the functions and then prove that they yield an (ε, R, n) feasible network code for $\mathcal{I}$.

Let σ be the value whose existence is proven in Claim 1. Let A_σ be the set of realizations $\hat{\mathbf{x}}_{\mathbf{S}}$ for which there exists a realization $\hat{\mathbf{x}}_{\mathbf{E}} \in A_{\hat{\mathbf{x}}_{\mathbf{S}}}$ with $\hat{f}_B(\hat{\mathbf{x}}_{\mathbf{S}}, \hat{\mathbf{x}}_{\mathbf{E}}) = \sigma$. By Claim 1, the size of A_σ is at least $(1-\varepsilon)2^{n(\sum_{s \in S} \hat{R}_s)} = (1-\varepsilon)2^{n(\sum_{s \in S} R_s)}$.

For $e \in E$ let $f_e : \left[2^{n \sum_{e' \in In(e)} c_{e'}}\right] \to [2^{n c_e}]$ be the function that takes as input the random variables $(X_{e'} : e' \in In(e))$ and returns as output $X_e = \hat{g}_{\hat{t}_e}(\sigma, (X_{e'} : e' \in In(e)))$. Here, $X_{e'}$ is a random variable of support $[2^{c_{e'} n}]$.

For terminals $t_i \in T$ in $\mathcal{I}$ let $g_{t_i} : \left[2^{n \sum_{e' \in In(t_i)} c_{e'}}\right] \to \left[2^{n \sum_{s \in S : b(s, t_i) = 1} R_s}\right]$ be the function that takes as input the random variables $(X_{e'} : e' \in In(t_i))$ and returns as output $\hat{g}_{\hat{t}_i}(\sigma, (X_{e'} : e' \in In(t_i)))$.

The following argument shows that the network code $(\mathcal{F}, \mathcal{G})$ defined above decodes correctly with probability $1 - \varepsilon$. Consider any rate-R realization $\mathbf{x}_{\mathbf{S}} = \{x_s\}_{s \in S}$ of the source information in $\mathcal{I}$, where $R = (R_1, \ldots, R_{|S|})$. Consider the source information $\hat{\mathbf{x}}_{\mathbf{S}}$ of $\hat{\mathcal{I}}$ corresponding to $\mathbf{x}_{\mathbf{S}}$, namely let $\hat{\mathbf{x}}_{\mathbf{S}} = \mathbf{x}_{\mathbf{S}}$. If $\hat{\mathbf{x}}_{\mathbf{S}} \in A_\sigma$, then there exists a realization $\hat{\mathbf{x}}_{\mathbf{E}}$ of source information $\{\hat{X}_e\}$ in $\hat{\mathcal{I}}$ for which $\hat{f}_B(\hat{\mathbf{x}}_{\mathbf{S}}, \hat{\mathbf{x}}_{\mathbf{E}}) = \sigma$. Recall that, by our definitions, all terminals of $\hat{\mathcal{I}}$ decode correctly given source realization $(\hat{\mathbf{x}}_{\mathbf{S}}, \hat{\mathbf{x}}_{\mathbf{E}})$. For any $s \in S$, let $\hat{\mathbf{x}}_{\mathbf{S}}(s) = x_s$ be the entry in $\hat{\mathbf{x}}_{\mathbf{S}}$ that corresponds to $\hat{X}_s$. For $e \in E$, let $\hat{\mathbf{x}}_{\mathbf{E}}(e)$ be the entry in $\hat{\mathbf{x}}_{\mathbf{E}}$ that corresponds to $\hat{X}_e$.

We show by induction on the topological order of G that, for source information $\mathbf{x}_{\mathbf{S}}$, the evaluation of f_e in the network code above results in the value x_e which is equal to $\hat{\mathbf{x}}_{\mathbf{E}}(e)$. For the base case, consider any edge $e = (u, v)$ in which u is a source; recall that any source has no incoming edges. In that case, by our definitions, the information x_e on edge e equals $f_e(x_s) = \hat{g}_{\hat{t}_e}(\sigma, x_s) = \hat{g}_{\hat{t}_e}(\hat{f}_B(\hat{\mathbf{x}}_{\mathbf{S}}, \hat{\mathbf{x}}_{\mathbf{E}}), \hat{\mathbf{x}}_{\mathbf{S}}(s)) = \hat{\mathbf{x}}_{\mathbf{E}}(e)$. The last equality follows from the fact that the index code $(\hat{\mathcal{F}}, \hat{\mathcal{G}})$ succeeds on source realization $(\hat{\mathbf{x}}_{\mathbf{S}}, \hat{\mathbf{x}}_{\mathbf{E}})$. Thus all terminals (and, in particular, terminal $\hat{t}_e$) decode correctly.

Next, consider any edge $e = (u, v)$ with incoming edges $e' \in In(e)$. In that case, by our definitions, the information x_e on edge e equals $f_e(x_{e'} : e' \in In(e))$. However, by induction, each $x_{e'}$ for which $e' \in In(e)$ satisfies $x_{e'} = \hat{\mathbf{x}}_{\mathbf{E}}(e')$. Thus $x_e = \hat{g}_{\hat{t}_e}(\sigma, (x_{e'} : e' \in In(e))) = \hat{g}_{\hat{t}_e}(\hat{f}_B(\hat{\mathbf{x}}_{\mathbf{S}}, \hat{\mathbf{x}}_{\mathbf{E}}), (\hat{\mathbf{x}}_{\mathbf{E}}(e') : e' \in In(e))) = \hat{\mathbf{x}}_{\mathbf{E}}(e)$. Again, the last equality follows because the index code $(\hat{\mathcal{F}}, \hat{\mathcal{G}})$ succeeds on $(\hat{\mathbf{x}}_{\mathbf{S}}, \hat{\mathbf{x}}_{\mathbf{E}})$.

Finally, we address the value of the decoding functions g_t for any $t \in T$. By definition, the outcome of g_t is exactly $\hat{g}_{\hat{t}_i}(\sigma, (x_{e'} : e' \in In(t_i))) = \hat{g}_{\hat{t}_i}(\hat{f}_B(\hat{\mathbf{x}}_{\mathbf{S}}, \hat{\mathbf{x}}_{\mathbf{E}}), (\hat{\mathbf{x}}_{\mathbf{E}}(e') : e' \in In(t_i))) = (\hat{\mathbf{x}}_{\mathbf{S}}(s) : b(s, t_i) = 1) = (x_s : b(s, t_i) = 1)$. This suffices to show that the proposed network code $(\mathcal{F}, \mathcal{G})$ succeeds with probability $1 - \varepsilon$ on a source input with rate vector R. Namely, we have presented correct decoding for $\mathcal{I}$ when $\mathbf{x}_{\mathbf{S}} = \hat{\mathbf{x}}_{\mathbf{S}} \in A_\sigma$ and have shown that $|A_\sigma| \geq (1 - \varepsilon)2^{n(\sum_{s \in S} R_s)}$. ■

V. Conclusions

In this work, we address the equivalence between the network and index coding paradigms. Following the line of proof presented in [15] for a restricted equivalence in the case of linear encoding, we present an equivalence for general (not necessarily linear) encoding and decoding functions. Our results show that the study and understanding of the index coding paradigm imply a corresponding understanding of the network coding paradigm.

Although our connection between network and index coding is very general, it does not directly imply a tool for determining the network coding capacity region as defined in Section II for general network coding instances. In the full version of this work [17], we show that for the case of colocated sources, one can determine the capacity region of network coding using that of index coding. However, a naive attempt to reduce the problem of determining whether a certain rate vector R is in the capacity region of a general network coding instance $\mathcal{I}$ to the corresponding problem in the context of index coding, shows that a stronger, more robust connection between index and network coding is needed. Specifically, a connection which allows some flexibility in the value of $\hat{c}_B$ might suffice and is subject of continuing research.

Recently, it has been shown [18], [19] that certain intriguing open questions in the context of network coding are well understood in the context of index coding (or the so-called "super-source" setting of network coding). These include the "zero-vs-ε error" question [20], [18], the "edge removal" question [21], [22], and the "δ-dependent source" question [19]. At first, it may seem that the equivalence presented in this work implies a full understanding of the open questions above in the context of network coding. Although this may be the case, a naive attempt to use our results with those presented in [18], [19] again shows the need for a stronger connection between index and network coding that (as above) allows some flexibility in the value of $\hat{c}_B$.

VI. Acknowledgements

S. El Rouayheb would like to thank Prof. Vincent Poor for his continuous support, and Curt Schieler for interesting discussions on the index coding problem.

References

[1] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung. Network Information Flow. *IEEE Transactions on Information Theory*, 46(4):1204–1216, 2000.

[2] S.-Y. R. Li, R. W. Yeung, and N. Cai. Linear Network Coding. *IEEE Transactions on Information Theory*, 49(2):371 – 381, 2003.

[3] R. Koetter and M. Medard. An Algebraic Approach to Network Coding. *IEEE/ACM Transactions on Networking*, 11(5):782 – 795, 2003.

[4] S. Jaggi, P. Sanders, P. A. Chou, M. Effros, S. Egner, K. Jain, and L. Tolhuizen. Polynomial Time Algorithms for Multicast Network Code Construction. *IEEE Transactions on Information Theory*, 51(6):1973–1982, June 2005.

[5] T. Ho, M. Médard, R. Koetter, D. R. Karger, M. Effros, J. Shi, and B. Leong. A Random Linear Network Coding Approach to Multicast. *IEEE Transactions on Information Theory*, 52(10):4413–4430, 2006.

[6] T. Cover and J. Thomas. *Elements of Information Theory, second edition*. John Wiley and Sons, 2006.

[7] R. W. Yeung, S.-Y. R. Li, N. Cai, and Z. Zhang. Network Coding Theory. *Foundations and Trends in Communications and Information Theory. Now Publishers*, 2006.

[8] T. Chan and A. Grant. Dualities Between Entropy Functions and Network Codes. *IEEE Transactions on Information Theory*, 54(10):4470–4487, 2008.

[9] Z. Bar-Yossef, Y. Birk, T. S. Jayram, and T. Kol. Index Coding with Side Information. *In Proceedings of 47th Annual IEEE Symposium on Foundations of Computer Science*, pages 197–206, 2006.

[10] E. Lubetzky and U. Stav. Non-linear Index Coding Outperforming the Linear Optimum. *In Proceedings of 48th Annual IEEE Symposium on Foundations of Computer Science*, pages 161–168, 2007.

[11] S. El Rouayheb, M. A. R. Chaudhry, and A. Sprintson. On the minimum number of transmissions in single-hop wireless coding networks. In *Information Theory Workshop (ITW)*, 2007.

[12] N. Alon, E. Lubetzky, U. Stav, A. Weinstein, and A. Hassidim. Broadcasting with side information. *In Proceedings of 49th Annual IEEE Symposium on Foundations of Computer Science*, pages 823–832, 2008.

[13] M. Langberg and A. Sprintson. On the Hardness of Approximating the Network Coding Capacity. *IEEE Transactions on Information Theory*, 57(2):1008–1014, 2011.

[14] A. Blasiak, R. Kleinberg, and E. Lubetzky. Lexicographic products and the power of non-linear network coding. *In Proceedings of 52nd Annual IEEE Symposium on Foundations of Computer Science*, pages 609 – 618, 2011.

[15] S. Y. El Rouayheb, A. Sprintson, and C. Georghiades. On the Relation Between the Index Coding and the Network Coding Problems. *In proceedings of IEEE International Symposium on Information Theory (ISIT)*, 2008.

[16] R. Dougherty, C. Freiling, and K. Zeger. Insufficiency of Linear Coding in Network Information Flow. *IEEE Transactions on Information Theory*, 51(8):2745–2759, 2005.

[17] M. Effros, S. El Rouayheb, and M. Langberg. An Equivalence between Network Coding and Index Coding. *Manuscript, available at http://arxiv.org/abs/1211.6660*, 2012.

[18] M. Langberg and M. Effros. Network coding: Is zero error always possible? In *In proceedings of Forty-Ninth Annual Allerton Conference on Communication, Control, and Computing*, pages 1478–1485, 2011.

[19] M. Langberg and M. Effros. Source coding for dependent sources. *In proceedings of IEEE Information Theory Workshop (ITW).*, 2012.

[20] T. Chan and A. Grant. On capacity regions of non-multicast networks. In *International Symposium on Information Theory*, pages 2378 – 2382, 2010.

[21] T. Ho, M. Effros, and S. Jalali. On equivalences between network topologies. In *Forty-Eighth Annual Allerton Conference on Communication, Control, and Computing*, 2010.

[22] S. Jalali, M. Effros, and T. Ho. On the impact of a single edge on the network coding capacity. In *Information Theory and Applications Workshop (ITA)*, 2011.