On a Capacity Equivalence between Network and Index Coding and the Edge Removal Problem

M. F. Wong
California Institute of Technology
Email: mwong2@caltech.edu

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

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

***Abstract*—In recent work by Effros, El Rouayheb, and Langberg, an equivalence of code feasibility between network and index coding is derived. The authors ask whether the capacity region of a network coding problem can be obtained by solving the capacity region of an index coding problem. We answer in the affirmative for the linear coding case. While the question is still open for the general case, we show that it is related to the edge removal problem, which has been studied recently.**

I. Introduction

The network coding capacity remains unsolved for most networks. The index coding problem [1] is a special case of the network coding problem that can be interpreted as a "broadcast with side information" problem: a broadcast node has access to all sources and wishes to communicate with several terminals, each having and desiring to reconstruct potentially different sets of sources.

In recent work by Effros, El Rouayheb and Langberg [2] (to appear in parallel to ISIT 2013), an equivalence of code feasibility between network coding and index coding is derived. That is, a code of rate $\mathbf{R}$, blocklength n, and error probability ϵ exists for a given network coding instance if and only if a code of the same blocklength and error probability and a rate $\hat{\mathbf{R}}$ (that is a function of $\mathbf{R}$) exists for a corresponding index coding problem. Thus, by [2], any efficient scheme to solve the index coding problem would yield an efficient scheme to solve the more general network coding problem.

Although the connection between network and index coding presented in [2] is very general, it does not resolve the question of whether the network coding capacity region can be obtained by solving the capacity region of a corresponding index coding problem. In this paper, we refer to this question as the *capacity equivalence* problem. While [2] proves capacity equivalence for networks that meet certain topological constraints, the question remains open for general topologies.

The main contribution of this work is to derive a connection between the capacity equivalence problem and the edge removal problem [3], [4]. The edge removal problem compares the rates achievable in a given network coding problem before and after an edge is removed from the network. To facilitate our proof, we use a slight variation on this question in which we compare the rates achievable on a given network coding problem $\mathcal{I}$ before and after an edge α of capacity λ is *added* to the network. The edge addition statement considered here is a weaker version of that found in [3], [4]. Nonetheless, it captures part of the essence of the edge removal question, which is to see whether removing an edge of vanishing capacity (in the blocklength) from a network always has a vanishing effect on the capacity of the network.

We also show that capacity equivalence between network and index coding holds for the case of linear coding. This establishes another connection with the edge removal problem, which is also solved for the linear coding case [3].

II. Notation and Network Models

We begin by defining instances of the network and index coding problems and their capacities. We employ the notation and models of [2], the size of a finite set S is denoted by $|S|$. We use bold letters to denote row vectors, for example $\mathbf{R} = (R_1, ..., R_{|S|})$. The i^{th} element of a vector $\mathbf{R}$ is denoted by R_i. For a vector $\mathbf{X} \in \mathbb{R}^n$, define $\mathbf{X}_+ = \max(\mathbf{0}, \mathbf{X})$, where $\mathbf{0}$ is a zero-valued vector and the max operator is applied component wise. We denote by $G^{\alpha(\lambda)}$, the graph that results from adding an edge α of capacity $\lambda > 0$ to G.

We study network coding instances and their corresponding index coding instances [2]. 'Unhatted" variables correspond to variables of network coding instances, while "hatted" variables (e.g., $\hat{x}$) correspond to the variables of index coding instances.

We denote by Tail(e) and Head(e) the tail and head nodes of edge e. We denote by $\text{In}_G(v)$ the set of edges in graph G with head node v. Similarly, $\text{Out}_G(v)$ refers to the set of edges in graph G with tail node v.

A. Network Coding

Model: An instance $\mathcal{I} = (G, S, T, B)$ of the network coding problem includes a directed acyclic network $G = (V, E)$, a set of source nodes $S \subset V$, and a set of terminal nodes $T \subset V$. Without loss of generality, we assume each source node $s \in S$ has no incoming edges and each terminal $t \in T$ has no outgoing edges. For blocklength n, source $s \in S$ holds a rate R_s random variable X_s, uniformly distributed over $\mathcal{X}_s = \mathbb{F}_2^{R_s n}$ and independent of all other sources. Edge $e \in E$ has capacity c_e and carries a message X_e in $\mathcal{X}_e = \mathbb{F}_2^{c_e n}$. Function $B : S \times T \to \mathbb{F}_2$ specifies the demands of each terminal; $B(s, t) = 1$ if terminal t requires source s and 0 otherwise. We denote by $G^{\alpha(\lambda)}$, the graph that results from adding an edge α of capacity $\lambda > 0$ to G. Let $\mathcal{I}^{\alpha(\lambda)} = (G^{\alpha(\lambda)}, S, T, B)$ be the network coding problem obtained from $\mathcal{I} = (G, S, T, B)$ by replacing G with $G^{\alpha(\lambda)}$.

Network Code: A network code, $(\mathcal{F}, \mathcal{G}) = (\{f_e\}, \{g_t\})$ assigns an encoding function f_e to each $e \in E$, and a decoding function g_t to each $t \in T$. For $e = (u, v)$, the inputs to f_e are the random variables associated with sources and incoming edges of node u. Random variable $X_e \in \mathbb{F}_2^{c_e n}$ equals the evaluation of f_e on its input. The inputs to g_t are the random variables associated with incoming edges of terminal t. The output of g_t is a reproduction of all sources required by t.

Linear Network Code: The definition of a linear network code, $(\mathcal{F}^l, \mathcal{G}^l) = (\{f_e^l\}, \{g_t^l\})$, is similar except that the encoding and decoding functions must be linear. For example, f_e^l takes form $\sum_{e' \in \text{In}_G(\text{Tail}(e))} X_{e'} A_{e',e} + k_e$ and g_e^l takes form $\sum_{e' \in \text{In}_G(t)} X_{e'} A_{e',t} + k_t$, where $A_{e',e}$ and $A_{e',t}$ are matrices and k_e and k_t are constant vectors with elements in $\mathbb{F}_2$.

Remark: While using linear encoders is common in the literature, the restriction to linear decoders is not. It turns out, however, that there is no loss of generality in this restriction. We roughly sketch the proof below. Consider a decoder t that wishes to decode $U \subseteq S$. Let the message received by t be $(X_e)_{\text{Head}(e)=t}$; which can be expressed as $(X_e)_{\text{Head}(e)=t} = (X_s)_{s \in U} M_1 + (X_s)_{s \notin U} M_2 + k_t$. Here M_1 and M_2 are the corresponding transfer matrices and k_t is a constant. Due to the structure of linear codes, M_1 must be full ranked and the rowspaces of M_1 and M_2 only intersect at $\{\mathbf{0}\}$, or else decoding will result in a large error probability. Hence we can decode $(X_s)_{s \in U}$ by multiplying $(X_e)_{\text{Head}(e)=t}$ with a suitable inverse matrix.

B. Index Coding

Model: An instance $\hat{\mathcal{I}} = (\hat{S}, \hat{T}, \{\hat{W}_{\hat{t}}\}, \{\hat{H}_{\hat{t}}\}, \hat{c}_B)$ of the index coding problem includes a set of terminals $\hat{T}$ and a set of sources $\hat{S}$ available to the broadcast node. For each terminal node $\hat{t} \in \hat{T}$, $\hat{W}_{\hat{t}}$ is the set of sources required by $\hat{t}$, and $\hat{H}_{\hat{t}}$ is the set of sources available at $\hat{t}$. Each terminal node $\hat{t}$ also receives information from the broadcast node at broadcast rate $\hat{c}_B$. Given a blocklength n, each source $\hat{s} \in \hat{S}$ holds a rate $\hat{R}_{\hat{s}}$ random variable $\hat{X}_{\hat{s}}$ uniformly distributed over $\hat{\mathcal{X}}_{\hat{s}} = \mathbb{F}_2^{\hat{R}_{\hat{s}} n}$. Each $\hat{X}_{\hat{s}}$ is independent of all others.

Index Code: An index code $(\hat{\mathcal{F}}, \hat{\mathcal{G}}) = (\hat{f}_B, \{\hat{g}_{\hat{t}}\})$ for $\hat{\mathcal{I}}$ assigns an encoding function $\hat{f}_B$ to the broadcast channel and a decoding function $\hat{g}_{\hat{t}}$ to each $\hat{t} \in \hat{T}$. The inputs to function $\hat{f}_B$ is the source random variables $\{\hat{X}_{\hat{s}}\}$, and the output is a rate $\hat{c}_B$ random variable $\hat{X}_B \in \mathbb{F}_2^{\hat{c}_B n}$. The inputs to the decoding function $\hat{g}_{\hat{t}}$ are the random variables in $\hat{H}_{\hat{t}}$ and the broadcast message $\hat{X}_B$. The output of $\hat{g}_{\hat{t}}$ is the vector of sources in $\hat{W}_{\hat{t}}$.

Linear Index Code: The definition of a linear index code $(\hat{\mathcal{F}}^l, \hat{\mathcal{G}}^l)$ for $\hat{\mathcal{I}}$ is similar except that the encoding and decoding functions must be linear. Thus $\hat{f}_B^l$ takes the form $\sum_{\hat{s} \in \hat{S}} \hat{X}_{\hat{s}} \hat{A}_{\hat{s},B} + \hat{k}_B$ and $\hat{g}_{\hat{t}}^l$ takes the form $\sum_{\hat{s} \in \hat{H}_{\hat{t}}} \hat{X}_{\hat{s}} \hat{A}_{\hat{s},\hat{t}} + \hat{X}_B \hat{A}_{B,\hat{t}} + \hat{k}_{\hat{t}}$. Here $\hat{A}_{\hat{s},B}$, $\hat{A}_{\hat{s},\hat{t}}$, $\hat{A}_{B,\hat{t}}$, $\hat{k}_{\hat{t}}$ and $\hat{k}_B$ are constant matrices and vectors over $\mathbb{F}_2$.

Remark: Recall that index coding is a special case of network coding (see, e.g., [2]). Under our definitions, $\hat{\mathcal{I}} = (\hat{S}, \hat{T}, \{\hat{W}_{\hat{t}}\}, \{\hat{H}_{\hat{t}}\}, \hat{c}_B)$ may also be expressed in the form of a network coding problem, and results for general network coding problems also apply to index coding problems.

C. Code Feasibility and Capacity Regions

Here, we give the feasibility definition and notations for both network and index coding.

Definition 1 *($(\epsilon, \mathbf{R}, n)$-feasibility):* For a rate vector $\mathbf{R} = (R_1, ..., R_{|S|})$, an instance $\mathcal{I}$ of the network coding problem is $(\epsilon, \mathbf{R}, n)-$feasible if there exists a network code $(\mathcal{F}, \mathcal{G})$ with blocklength n and rate $\mathbf{R}$ such that each of the decoding functions g_t outputs a vector of all sources required by t, with an overall failure probability at most ϵ.

Similarly, a network coding problem $\mathcal{I}$ is $(\epsilon, \mathbf{R}, n)$-linear feasible if there exists a linear code $(\mathcal{F}^l, \mathcal{G}^l)$ with blocklength n, such that rate $\mathbf{R}$ is achieved with an overall failure probability at most ϵ.

Definition 2 *($\mathbf{R}$-feasibility):* A network coding problem $\mathcal{I}$ is $\mathbf{R}$-feasible if for any $\epsilon > 0$ and any $\delta > 0$ there exists a blocklength n such that $\mathcal{I}$ is $(\epsilon, \mathbf{R}(1-\delta), n)-$feasible.

A network coding problem $\mathcal{I}$ is $\mathbf{R}$-linear feasible if for any $\epsilon > 0$ and any $\delta > 0$ there exists a blocklength n such that $\mathcal{I}$ is $(\epsilon, \mathbf{R}(1-\delta), n)$-linear feasible.

Definition 3 *(Capacity region):* The capacity region of $\mathcal{I}$, denoted by $\mathcal{R}(\mathcal{I})$, is the set of all $\mathbf{R}$ such that $\mathcal{I}$ is $\mathbf{R}$-feasible.

Definition 4 *(Linear coding capacity region):* The linear coding capacity region of $\mathcal{I}$, denoted as $\mathcal{R}_l(\mathcal{I})$, is the set of all $\mathbf{R}$ such that $\mathcal{I}$ is $\mathbf{R}$-linear feasible.

For a rate vector $\hat{\mathbf{R}} = (\hat{R}_1, ..., \hat{R}_{|\hat{S}|})$, an index coding problem $\hat{\mathcal{I}}$ is $(\epsilon, \hat{\mathbf{R}}, n)-$feasible if there exists an index code $(\hat{\mathcal{F}}, \hat{\mathcal{G}})$ with blocklength n, rate $\hat{\mathbf{R}}$ such that each decoding function $\hat{g}_{\hat{t}}$ outputs a vector of all sources required by $\hat{t}$, with an overall failure probability at most ϵ. The definitions of $(\epsilon, \hat{\mathbf{R}}, n)$-linear feasibility, $\hat{\mathbf{R}}$-feasibility, $\hat{\mathbf{R}}$-linear feasibility, capacity region $\mathcal{R}(\hat{\mathcal{I}})$, and linear capacity region $\mathcal{R}_l(\hat{\mathcal{I}})$ for index coding instance $\hat{\mathcal{I}}$ follow similarly.

III. Prior Work

In recent work by Effros, El Rouayheb and Langberg [2], an equivalence of code feasibility between network coding and index coding is derived. Roughly speaking, it is shown that any instance $\mathcal{I}$ of the network problem can be transformed to a corresponding instance $\hat{\mathcal{I}}$ of the index coding problem such that any efficient scheme to solve the instance $\hat{\mathcal{I}}$ yields an efficient scheme for $\mathcal{I}$. We restate the transformation from $\mathcal{I}$ to $\hat{\mathcal{I}}$ and the equivalence result below.

A. Transformation from $\mathcal{I}$ to $\hat{\mathcal{I}}$

Given $\mathcal{I} = (G, S, T, B)$, $G = (V, E)$, and rate vector $\mathbf{R}$, the corresponding index coding problem $\hat{\mathcal{I}} = (\hat{S}, \hat{T}, \{\hat{W}_{\hat{t}}\}, \{\hat{T}_{\hat{t}}\}, \hat{c}_B)$ and rate $\hat{\mathbf{R}}$ are constructed as follows:

- $\hat{S} = \{\hat{X}_s\}_{s \in S} \cup \{\hat{X}_e\}_{e \in E}$. Set $\hat{S}$ has $|S|+|E|$ sources: one source $\hat{X}_s$ corresponds to each original source $s \in S$ and one source $\hat{X}_e$ corresponds to each original edge $e \in E$.
- $\hat{T} = \{\hat{t}_e\}_{e \in E} \cup \{\hat{t}_i\}_{t_i \in T} \cup \{\hat{t}_{all}\}$. Set $\hat{T}$ has $|E| + |T| + 1$ terminals: one terminal $\hat{t}_e$ corresponds to each $e \in E$, one terminal $\hat{t}_i$ corresponding to each $t_i \in T$ and a single terminal $\hat{t}_{all}$.
- For $e \in E : \hat{H}_{\hat{t}_e} = \{\hat{X}_{e'}\}_{e' \in \text{In}_G(\text{Tail}(e))}$, $\hat{W}_{\hat{t}_e} = \{\hat{X}_e\}$.

- For $t_i \in T$: $\hat{H}_{\hat{t}_i} = \{\hat{X}_{e'}\}_{e' \in \text{In}_G(t_i)}$, $\hat{W}_{\hat{t}_i} = \{\hat{X}_s\}_{s:B(s,t_i)=1}$.
- For $\hat{t}_{all}$: $\hat{H}_{\hat{t}_{all}} = \{\hat{X}_s\}_{s \in S}$ and $\hat{W}_{\hat{t}_{all}} = \{\hat{X}_e\}_{e \in E}$.
- $\hat{c}_B = \sum_{e \in E} c_e$.
- $\hat{\mathbf{R}} = (\mathbf{R}, (c_e)_{e \in E})$.

B. Equivalence in Code Feasibility

[2, Theorem 1]*:* Consider a network coding problem $\mathcal{I}$ and corresponding index coding problem $\hat{\mathcal{I}}$. For any rate vector R, integer n and $\epsilon \geq 0$, $\mathcal{I}$ is $(\epsilon, \mathbf{R}, n)$-feasible if and only if $\hat{\mathcal{I}}$ is $(\epsilon, \hat{\mathbf{R}}, n)$-feasible [2, Theorem 1].

C. Code Feasibility and Capacity Regions

Throughout the remainder of this paper, let $\hat{\mathcal{I}}$ be the index coding instance corresponding to $\mathcal{I}$.

Definition 5 *($\hat{\mathcal{R}}$-capacity region):* The $\hat{\mathcal{R}}$-capacity region of $\mathcal{I}$, denoted by $\hat{\mathcal{R}}(\mathcal{I})$, is the set of all $\mathbf{R}$ such that $\hat{\mathcal{I}}$ is $(\mathbf{R}, (c_e)_{e \in E})$-feasible.

Definition 6 *($\hat{\mathcal{R}}$-linear coding capacity region):* The $\hat{\mathcal{R}}$-linear coding capacity region of $\mathcal{I}$, denoted by $\hat{\mathcal{R}}_l(\mathcal{I})$, is the set of all $\mathbf{R}$ such that $\hat{\mathcal{I}}$ is $(\mathbf{R}, (c_e)_{e \in E})$-linear feasible.

Definition 7 *($\mathcal{D}$-function):* Given rate vectors $\mathbf{R}_a = (R_{a1}, ..., R_{am})$ and $\mathbf{R}_b = (R_{b1}, ..., R_{bm})$, define distance $\mathrm{d}_\infty(\mathbf{R}_a, \mathbf{R}_b) = \max_i |R_{ai} - R_{bi}|$. Given rate regions $\mathcal{R}_a$ and $\mathcal{R}_b$ such that $\mathcal{R}_a \subseteq \mathcal{R}_b$, define

$$\mathcal{D}(\mathcal{R}_a, \mathcal{R}_b) = \max_{\mathbf{R}_b \in \mathcal{R}_b} \min_{\mathbf{R}_a \in \mathcal{R}_a} \mathrm{d}_\infty(\mathbf{R}_a, \mathbf{R}_b).$$

D. Our starting point from [2]

In addressing the question of capacity equivalence which asks whether $\mathcal{R}(I) = \hat{\mathcal{R}}(I)$, [2] studies the two directions needed, whether $\mathcal{R}(I) \subseteq \hat{\mathcal{R}}(I)$ and whether $\hat{\mathcal{R}}(I) \subseteq \mathcal{R}(I)$. The first direction is shown to hold for all network coding instances $\mathcal{I}$. The second direction is more subtle. It is proven only for networks $\mathcal{I}$ with co-located sources. The major technical claim used in [2] for that proof is stated below and is the starting point of our analysis.

Roughly speaking, it is shown in [2] that when the index coding instance $\hat{\mathcal{I}}$ is $(\epsilon, (\mathbf{R}, (c_e)_{e \in E}), n)$-feasible, then one can use the corresponding index code $(\hat{\mathcal{F}}, \hat{\mathcal{G}})$ to obtain a network code $(\mathcal{F}, \mathcal{G})$ for $\mathcal{I}$ by carefully fixing the value of the encoding function $\hat{f}_B$ to a specific value $\sigma \in \mathbb{F}_2^{\hat{c}_B n}$. This σ has certain properties specified below and is central to the design of the network code $(\mathcal{F}, \mathcal{G})$. Assuming that $(\mathbf{R}, (c_e)_{e \in E}) \in \mathcal{R}(\hat{\mathcal{I}})$, one can only guarantee that $\hat{\mathcal{I}}$ is $(\epsilon, (\mathbf{R}, (c_e)_{e \in E})(1-\delta), n)$-feasible for any $\delta > 0$. As a result, a single value σ with the desired properties cannot be found, but a relatively small set of σ values can. In [2] this small set is used to construct a network code $(\mathcal{F}, \mathcal{G})$ for the collocated source case, and in what follows we use this small set to prove our claims.

We begin with some notation and then describe the technical claim of [2] we use. Let $\hat{\mathcal{I}}$ be the index coding instance corresponding to $\mathcal{I}$. Consider any $(\mathbf{R}, (c_e)_{e \in E}) \in \mathcal{R}(\hat{\mathcal{I}})$. Fix some $\epsilon > 0$ and $\delta > 0$ for which $\hat{\mathcal{I}}$ is $(\epsilon, (\mathbf{R}, (c_e)_{e \in E})(1-\delta), n)$-feasible under a blocklength-n index code $(\hat{\mathcal{F}}, \hat{\mathcal{G}})$. Denote by $\hat{\mathbf{x}}_\mathbf{S} = (\hat{x}_s)_{s \in S}$ and $\hat{\mathbf{x}}_\mathbf{E} = (\hat{x}_e)_{e \in E}$ the vector of source and edge messages for $\hat{\mathcal{I}}$, respectively. Let $E(\hat{\mathbf{x}}_\mathbf{S}, \hat{\mathbf{x}}_\mathbf{E})$ be the error indicator function under the code $(\hat{\mathcal{F}}, \hat{\mathcal{G}})$; then $E(\hat{\mathbf{x}}_\mathbf{S}, \hat{\mathbf{x}}_\mathbf{E})$ is equal to 1 if any of the terminals decode incorrectly when $(\hat{\mathbf{X}}_\mathbf{S}, \hat{\mathbf{X}}_\mathbf{E}) = (\hat{\mathbf{x}}_\mathbf{S}, \hat{\mathbf{x}}_\mathbf{E})$ and 0 otherwise. We also define $A_{\hat{\mathbf{x}}_\mathbf{S}} = \{\hat{\mathbf{x}}_\mathbf{E} | E(\hat{\mathbf{x}}_\mathbf{S}, \hat{\mathbf{x}}_\mathbf{E}) = 0\}$.

[2, Claim 3]*:* There exists a set $\Sigma \subset \mathbb{F}_2^{\hat{c}_B n}$ of cardinality

$$|\Sigma| = n \log(4/3)(1-\delta)(\sum_{s \in S} \hat{R}_s) 2^{n\delta\hat{c}_B}$$

such that at least a $(1-2\epsilon)$ fraction of source realizations $\hat{\mathbf{x}}_\mathbf{S}$ satisfy $\hat{f}_B(\hat{\mathbf{x}}_\mathbf{S}, \phi(\hat{\mathbf{x}}_\mathbf{S})) = \sigma(\hat{\mathbf{x}}_\mathbf{S})$ for some $\phi(\hat{\mathbf{x}}_\mathbf{S}) \in A_{\hat{\mathbf{x}}_\mathbf{E}}$ and $\sigma(\hat{\mathbf{x}}_\mathbf{S}) \in \Sigma$ dependent only on $\hat{\mathbf{x}}_\mathbf{S}$.

IV. Statement of Results

A. The Edge Addition Statement

We compare the rates achievable on a given network coding problem $\mathcal{I} = (G, S, T, B)$ before and after an edge α of capacity λ is added to network G. The following statement is not known to be true or false.

Statement 1 *(Edge addition):* Let $\mathcal{I} = (G, S, T, B)$ be a network coding problem and $\mathcal{I}^{\alpha(\lambda)}$ be obtained from $\mathcal{I}$ by replacing G with $G^{\alpha(\lambda)}$, for any edge α of capacity λ such that $G^{\alpha(\lambda)}$ is acyclic. The function $k(\lambda, \mathcal{I}, \alpha) = \mathcal{D}(\mathcal{R}(\mathcal{I}), \mathcal{R}(\mathcal{I}^{\alpha(\lambda)}))$ satisfies

$$\lim_{\lambda \to 0^+} k(\lambda, \mathcal{I}, \alpha) = 0.$$

B. The Capacity Equivalence Statement

In [2], the authors ask whether the network coding capacity region can be obtained from the index coding capacity region. The question posed in [2] is captured in the following open statement.

Statement 2 *(Capacity Equivalence):* Let $\mathcal{I}$ be a network coding problem, then $\mathcal{R}(\mathcal{I}) = \hat{\mathcal{R}}(\mathcal{I})$.

C. Results

Our main result shows that the two statements above are equivalent.

Theorem 1*:* Statement 1 is true if and only if Statement 2 is true.

We also show that Statement 2 is true under linear coding.

Theorem 2*:* Let $\mathcal{I} = (G, S, T, B)$ be a network coding problem. Then $\mathcal{R}_l(\mathcal{I}) = \hat{\mathcal{R}}_l(\mathcal{I})$.

The theorems are proven in Sections V and VI.

V. Proof of Theorem 1

A. Statement 1 implies Statement 2

Consider network coding instance $\mathcal{I} = (G, S, T, B)$. Since $\mathcal{R}(\mathcal{I}) \subseteq \hat{\mathcal{R}}(\mathcal{I})$ by [2, Cor. 1], it suffices to show that Statement 1 implies $\hat{\mathcal{R}}(\mathcal{I}) \subseteq \mathcal{R}(\mathcal{I})$.

Let $\hat{\mathcal{I}}$ be the index coding instance corresponding to $\mathcal{I}$. For any $(\mathbf{R}, c_1, ..., c_e) \in \mathcal{R}(\hat{\mathcal{I}})$, $\epsilon > 0$, and $\delta > 0$, there exists an $n \geq 1$ for which $\hat{\mathcal{I}}$ is $(\epsilon, (\mathbf{R}, (c_e)_{e \in E})(1-\delta), n)$-feasible under blocklength n index code $(\hat{\mathcal{F}}, \hat{\mathcal{G}})$. Using notation from Section III-D, let $E(\hat{\mathbf{x}}_\mathbf{S}, \hat{\mathbf{x}}_\mathbf{E})$ be the error indicator function with respect to the code $(\hat{\mathcal{F}}, \hat{\mathcal{G}})$.

We follow the construction in [5]. Define a new network $\mathcal{I}'$ by replacing the source nodes in S by $|S|$ new source nodes $s_1',...,s_{|S|}'$ and a supernode s_c' with access to all the new sources $s_1',...,s_{|S|}'$ through edges (s_i', s_c'). For each old source node $s_i \in S$, there is a capacity-R_s edge (s_i', s_i) from new source s_i' to the old source s_i. Supernode s_c' is connected to an intermediate node $\tilde{s}_c'$ through a bottleneck link $\alpha' = (s_c', \tilde{s}_c')$ of capacity γ. Intermediate node $\tilde{s}_c'$ is in turn connected to all nodes V of network G. By [2, Claim 3], for each source realization $\mathbf{x_S}$, supernode s_c' can compute $\sigma(\mathbf{x_S})$ such that almost all source realizations $\mathbf{x_S}$ satisfy $f_B(\mathbf{x_S}, \phi(\mathbf{x_S})) = \sigma(\mathbf{x_S})$ for some $\phi(\mathbf{x_S}) \in A_{\hat{\mathbf{x}}_\mathbf{S}}$.

Following the analysis and construction in [2], we define a network code that operates on $\mathcal{I}'$ in two phases. During the first phase, each new source s_i' transmits its source information over all outgoing edges; the supernode takes in $\mathbf{x_S}$ and broadcasts an overhead message $X_{\alpha'}(\mathbf{x_S}) = \hat{f}_B(\mathbf{x_S}, \phi(\mathbf{x_S})) = \sigma(\mathbf{x_S})$ (or an "error" if not found) to all the nodes in the network G. In the second phase, the sources $\mathbf{x_S}$ are transmitted through the network G by sending on edge e the message $X_e = \hat{g}_{\hat{t}_e}(X_{\alpha'}, (X_{e'} : e' \in \mathrm{In}_G(\mathrm{Tail}(e))))$; each terminal t_i implements the decoding function $\hat{g}_{\hat{t}_i}(X_{\alpha'}, (X_{e'} : e' \in \mathrm{In}_G(t_i)))$. It is shown in [2, Cor. 1] that this code is indeed a $((2\epsilon, \mathbf{R}(1-\delta), n)$ code for $\mathcal{I}'$.

Since [2, Claim 3] guarantees existence of a set Σ of small cardinality, supernode s_c' can broadcast the message $\sigma(\mathbf{x_S})$ to each node in the network with broadcast rate at most γ bits, where $\gamma = \log|\Sigma|/n$.

By the above argument, network $\mathcal{I}'$ is $(2\epsilon, \mathbf{R}(1-\delta), n)$-feasible. Let $\mathcal{I}^*$ be the network obtained from $\mathcal{I}'$ by removing bottleneck link α'. Note that $\mathcal{R}(\mathcal{I}) = \mathcal{R}(\mathcal{I}^*)$. Now if Statement 1 is true then $\mathcal{I}^*$ is $(2\epsilon, (\mathbf{R}(1-\delta) - k(\gamma, \mathcal{I}^*, \alpha')\mathbf{1})_+, n)$-feasible. Since γ and $k(\gamma, \mathcal{I}^*, \alpha')$ tend to zero as δ tends to zero, we conclude that $\mathbf{R} \in \mathcal{R}(\mathcal{I})$. ■.

B. Statement 2 implies Statement 1

We shall proceed by contradiction, showing that if edge addition is false, then capacity equivalence is false.

First note that in general, $\mathcal{R}(\mathcal{I}) \subseteq \mathcal{R}(\mathcal{I}^{\alpha(\lambda')}) \subseteq \mathcal{R}(\mathcal{I}^{\alpha(\lambda)})$, for $0 < \lambda' < \lambda$. Suppose that Statement 1 is false, then there exist network coding instances $\mathcal{I}$ and $\mathcal{I}^{\alpha(\lambda)}$ and a constant $\Delta > 0$ such that $\lim_{\lambda \to 0^+} k(\lambda, \mathcal{I}, \alpha) > \Delta$. This implies that $\mathcal{R}(\mathcal{I}) \subsetneq \lim_{\lambda \to 0^+} \mathcal{R}(\mathcal{I}^{\alpha(\lambda)})$. We can therefore find a rate vector $\mathbf{R}^\Delta$ such that $\min_{\mathbf{R} \in \mathcal{R}(\mathcal{I})} \mathrm{d}_\infty(\mathbf{R}, \mathbf{R}^\Delta) > \Delta$; and for any $\lambda > 0$, $\mathbf{R}^\Delta \in \mathcal{R}(\mathcal{I}^{\alpha(\lambda)})$ but $\mathbf{R}^\Delta \notin \mathcal{R}(\mathcal{I})$. (Roughly speaking, if edge addition is false, then there exist network coding instances where adding a link of vanishing capacity leads to a discontinuity in the capacity region.)

Let $\hat{\mathcal{I}}$ and $\widehat{\mathcal{I}^{\alpha(\lambda)}}$ be the transformation of $\mathcal{I}$ and $\mathcal{I}^{\alpha(\lambda)}$ to index coding instances, respectively (as described in Section III-A). Recall that the edges in $\mathcal{I}$ have capacities $\{c_e\}_{e \in E}$ while the edges in $\mathcal{I}^{\alpha(\lambda)}$ have capacities $\{c_e\}_{e \in E} \cup \{\lambda\}$. The corresponding broadcast capacity for $\hat{\mathcal{I}}$ and $\widehat{\mathcal{I}^{\alpha(\lambda)}}$ is thus $\hat{c}_B = \sum_{e \in E} c_e$ and $\hat{c}_B + \lambda$, respectively. By assumption, $\widehat{\mathcal{I}^{\alpha(\lambda)}}$ is $(\mathbf{R}^\Delta, (c_e)_{e \in E}, \lambda)$–feasible. We now state a connection between the capacity region of $\hat{\mathcal{I}}$ and $\widehat{\mathcal{I}^{\alpha(\lambda)}}$, proven at the end of this section.

Lemma 1: If $\widehat{\mathcal{I}^{\alpha(\lambda)}}$ is $(\mathbf{R}, (c_e)_{e \in E}, \lambda)$-feasible, $\hat{\mathcal{I}}$ is $(\mathbf{R}, (c_e)_{e \in E})(1 - \frac{\lambda}{\hat{c}_B + \lambda})$-feasible.

By Lemma 1, $\hat{\mathcal{I}}$ is $(\mathbf{R}^\Delta, (c_e)_{e \in E})(1 - \frac{\lambda}{\hat{c}_B + \lambda})$-feasible for any $\lambda > 0$. This implies that for any $\lambda > 0$, $\epsilon > 0$, and $\delta > 0$, there exists a blocklength n such that $\hat{\mathcal{I}}$ is $(\epsilon, (\mathbf{R}^\Delta, (c_e)_{e \in E})(1 - \frac{\lambda}{\hat{c}_B + \lambda})(1-\delta), n)$-feasible. Since $(1 - \frac{\lambda}{\hat{c}_B + \lambda})(1-\delta)$ tends to 1 as $\lambda \to 0^+$ and $\delta \to 0^+$, we conclude that $\hat{\mathcal{I}}$ is $(\mathbf{R}^\Delta, (c_e)_{e \in E})$–feasible.

If Statement 2 is true, then $\mathcal{I}$ is $\mathbf{R}^\Delta$–feasible, which is a contradiction of the assumption above that $\mathbf{R}^\Delta \notin \mathcal{R}(\mathcal{I})$. Hence if Statement 1 is false, then Statement 2 is false. ■

Proof of Lemma 1: Due to space limitations, we summarize the proof. To prove the lemma we must prove two claims. The first is that removing the source corresponding to edge α from the index coding instance $\widehat{I^{\alpha(\lambda)}}$ does not shrink the capacity region of the remaining sources. Intuitively, this follows by using the code for $\widehat{I^{\alpha(\lambda)}}$ while fixing the value of the removed source to a suitable constant value. The second claim, proven in [6, Theorem 4.1], states that reducing the capacity of any edge in a network by a multiplicative factor of $\kappa < 1$ shrinks the capacity region by at most factor κ. We apply this second claim to the broadcast capacity of $\widehat{I^{\alpha(\lambda)}}$.

VI. Proof of Theorem 2

Consider network coding instance $\mathcal{I} = (G, S, T, B)$. Since $\mathcal{R}_l(\mathcal{I}) \subseteq \hat{\mathcal{R}}_l(\mathcal{I})$ [2, Theorem 1], it suffices to show that $\hat{\mathcal{R}}_l(\mathcal{I}) \subseteq \mathcal{R}_l(\mathcal{I})$. The proof is similar to that of Theorem 1, except in this case all codes are linear. For this proof, we use the result that edge removal is true for linear network codes [3]. We present Lemma 2 below, which is a linear version of [2, Claim 3]. Using notation of Section III-D, we require function $\sigma(\mathbf{x_S})$ in [2, Claim 3] to be linear in $\mathbf{x_S}$.

We begin with definitions. Let $\hat{\mathcal{I}}$ be the index coding instance corresponding to $\mathcal{I}$ and $(\mathbf{R}, (c_e)_{e \in E}) \in \mathcal{R}_l(\hat{\mathcal{I}})$. Fix some $0 < \epsilon < \frac{1}{2}$ and $\delta > 0$. Then $\hat{\mathcal{I}}$ is $(\epsilon, (\mathbf{R}, (c_e)_{e \in E})(1-\delta), n)$-feasible, under a certain blocklength n linear index code $(\hat{f}_B^l, \hat{\mathcal{G}}^l)$. Since the code is linear, the message $\hat{f}_B^l(\hat{\mathbf{x}}_\mathbf{S}, \hat{\mathbf{x}}_\mathbf{E})$ sent on the bottleneck link is a linear combination of the sources $\hat{\mathbf{x}}_\mathbf{S} = (\hat{x}_s)_{s \in S}$ and $\hat{\mathbf{x}}_\mathbf{E} = (\hat{x}_e)_{e \in E}$. Namely,

$$\hat{f}_B^l(\hat{\mathbf{x}}_\mathbf{S}, \hat{\mathbf{x}}_\mathbf{E}) = \hat{\mathbf{x}}_\mathbf{S}\hat{A}_S + \hat{\mathbf{x}}_\mathbf{E}\hat{A}_E.$$

We treat all source messages as binary vectors. Here $\hat{\mathbf{x}}_\mathbf{S}$ and $\hat{\mathbf{x}}_\mathbf{E}$ are binary vectors of dimension $\log_2|\hat{\mathcal{X}}_S| = \sum_{s \in S} \log_2|\hat{\mathcal{X}}_s|$ and $\log_2|\hat{\mathcal{X}}_E| = \sum_{s \in E} \log_2|\hat{\mathcal{X}}_e|$ respectively. Binary matrices $\hat{A}_S$ and $\hat{A}_E$ have dimensions $\log_2|\hat{\mathcal{X}}_S| \times n\hat{c}_B$ and $\log_2|\hat{\mathcal{X}}_E| \times n\hat{c}_B$ respectively. All additions are over $\mathbb{F}_2$. Let $E(\hat{\mathbf{x}}_\mathbf{S}, \hat{\mathbf{x}}_\mathbf{E})$ be the error indicator function under the linear code $(\hat{\mathcal{F}}^l, \hat{\mathcal{G}}^l)$; then $E(\hat{\mathbf{x}}_\mathbf{S}, \hat{\mathbf{x}}_\mathbf{E})$ is equal to 1 if any of the terminals decode incorrectly when $(\hat{\mathbf{X}}_\mathbf{S}, \hat{\mathbf{X}}_\mathbf{E}) = (\hat{\mathbf{x}}_\mathbf{S}, \hat{\mathbf{x}}_\mathbf{E})$ and 0 otherwise. We are now ready to state Lemma 2.

Lemma 2: For $0 < \epsilon < \frac{1}{2}$, there exists a $\log_2 |\hat{\mathcal{X}}_S| \times n\hat{c}_B$ binary matrix $\hat{A}$ with $rank(\hat{A}) \le n\delta\hat{c}_B$ and some fixed $k^* \in \mathbb{F}_2^{n\hat{c}_B}$ such that all source realizations $\hat{\mathbf{x}}_\mathbf{S}$ satisfy $\hat{f}^l_B(\hat{\mathbf{x}}_\mathbf{S}, \phi(\hat{\mathbf{x}}_\mathbf{S})) = \sigma^l(\hat{\mathbf{x}}_\mathbf{S})$ for some arbitrary function $\phi(\hat{\mathbf{x}}_\mathbf{S}) \in \hat{\mathcal{X}}_\mathbf{E}$ and linear function $\sigma^l(\hat{\mathbf{x}}_\mathbf{S}) = \hat{\mathbf{x}}_\mathbf{S}\hat{A} + k^*$ with

$$\Pr[E(\hat{\mathbf{X}}_\mathbf{S}, \phi(\hat{\mathbf{X}}_\mathbf{S})) = 1] \le \epsilon,$$

when $\hat{\mathbf{X}}_\mathbf{S}$ is uniformly distributed in $\mathbb{F}_2^{\log_2 |\hat{\mathcal{X}}_S|}$.

Our proof follows the lines of that given for Theorem 1 in Section V-A. Namely, we modify $\mathcal{I}$ to $\mathcal{I}'$ by replacing the old sources by a set of new sources and adding a supernode. By Lemma 2, for each source realization $\mathbf{x}_\mathbf{S}$, the supernode can compute the linear function $\sigma^l(\mathbf{x}_\mathbf{S}) = \mathbf{x}_\mathbf{S}\hat{A} + k^*$ such that all source realizations $\mathbf{x}_\mathbf{S}$ satisfy $\hat{f}^l_B(\mathbf{x}_\mathbf{S}, \phi(\mathbf{x}_\mathbf{S})) = \sigma^l(\mathbf{x}_\mathbf{S})$, for some $\phi(\mathbf{x}_\mathbf{S}) \in \hat{\mathcal{X}}_E$.

As in the proof of Theorem 1, we define a linear network code that operates on $\mathcal{I}'$ in two phases. During the first phase, the supernode takes in $\mathbf{x}_\mathbf{S}$, broadcasts an overhead message $X_{\alpha'}(\mathbf{x}_\mathbf{S}) = \hat{f}^l_B(\hat{\mathbf{x}}_\mathbf{S}, \phi(\hat{\mathbf{x}}_\mathbf{S})) = \mathbf{x}_\mathbf{S}\hat{A} + k^*$ to all nodes in the network. The source $\mathbf{x}_\mathbf{S}$ is then transmitted through the network using the modified function $(\mathcal{F}^{l*}, \mathcal{G}^{l*}) = (\{f^{l*}_e\}, \{g^{l*}_t\})$, where $f^{l*}_e = \hat{g}^l_{\hat{t}_e}(X_{\alpha'}, (X_{e'} : e' \in \text{In}_G(\text{Tail}(e))))$ and $g^{l*}_{t_i} = \hat{g}^l_{\hat{t}_i}(X_{\alpha'}, (X_{e'} : e' \in \text{In}_G(t_i)))$. Since each $\hat{g}^l_{\hat{t}_i}$ and $\hat{g}^l_{\hat{t}_e}$ are linear, it is not hard to verify that $g^{l*}_{t_i}$ and f^{l*}_e are also linear.

Since Lemma 2 guarantees $rank(\hat{A}) \le n\delta\hat{c}_B$, supernode s'_c can broadcast message $\hat{\mathbf{x}}_\mathbf{S}\hat{A}$ to each node in the network using a broadcast rate of at most γ, where $\gamma = \delta\hat{c}_B$. Our selection of matrix $\hat{A}$ satisfies $\Pr[E(\hat{\mathbf{X}}_\mathbf{s}, \phi(\hat{\mathbf{X}}_\mathbf{s})) = 1] \le \epsilon$. Following the proof of Theorem 1, we conclude that the above scheme gives us a blocklength n linear network code for $\mathcal{I}'$ that achieves rate $\mathbf{R}(1-\delta)$ with error probability at most ϵ.

By the above argument, network $\mathcal{I}'$ is $(\epsilon, \mathbf{R}(1-\delta), n)$-feasible. Let $\mathcal{I}^*$ be the network obtained from $\mathcal{I}'$ by removing bottleneck link α'. Note that $\mathcal{R}_l(\mathcal{I}) = \mathcal{R}_l(\mathcal{I}^*)$. Since the edge removal statement is true for linear codes [3], network $\mathcal{I}^*$ is $(\epsilon, (\mathbf{R}(1-\delta) - k(\gamma, \mathcal{I}^*, \alpha')\mathbf{1})_+, n)$-feasible. Since γ and $k(\gamma, \mathcal{I}^*, \alpha')$ tends to zero as δ tends to zero, $\mathbf{R}$ is in the linear coding capacity region of $\mathcal{I}$. ■

Proof of Lemma 2: Consider a blocklength n, linear index code $(\hat{\mathcal{F}}^l, \hat{\mathcal{G}}^l)$ over a binary field such that $\hat{\mathcal{I}}$ is $(\epsilon, \hat{\mathbf{R}}(1-\delta), n)$-feasible. Since $\hat{t}_{all}$ is able to decode $\hat{\mathbf{x}}_\mathbf{E}$ from $(\hat{\mathbf{x}}_\mathbf{S}, \hat{\mathbf{x}}_\mathbf{S}\hat{A}_S + \hat{\mathbf{x}}_\mathbf{E}\hat{A}_E)$ with error probability less than $\frac{1}{2}$, $\hat{A}_E$ must be full rank (i.e. $rank(\hat{A}_E) = \log_2 |\hat{\mathcal{X}}_E| = (1-\delta)n\hat{c}_B$).

Next, we construct the low rank matrix $\hat{A}$ to satisfy this lemma. For any matrix M, denote by $M(i)$ row i of M and by $RS(M)$ the rowspace of M. For vector spaces V and W, let $V \oplus W = \{v + w | v \in V, w \in W\}$. Let $B_E \subseteq \{1, ..., \log_2 |\hat{\mathcal{X}}_E|\}$ such that $\{\hat{A}_E(i)\}_{i \in B_E}$ forms a basis for $RS(\hat{A}_E)$. Let $B_{S1} \subseteq \{1, ..., \log_2 |\hat{\mathcal{X}}_S|\}$ such that $\{\hat{A}_S(i)\}_{i \in B_{S1}} \cup \{\hat{A}_E(i)\}_{i \in B_E}$ forms a basis for $RS(\hat{A}_S) \oplus RS(\hat{A}_E)$. Since $\hat{A}_E$ is full rank, we must have $|B_{S1}| \le \delta n\hat{c}_B$ or we would have more than $n\hat{c}_B$ independent vectors in $\mathbb{F}_2^{n\hat{c}_B}$. Therefore, $\forall i \in \{1, ..., \log_2 |\hat{\mathcal{X}}_S|\}$, $\hat{A}_S(i)$ can be expressed as a linear combination of the basis $\{\hat{A}_S(i)\}_{i \in B_{S1}}, \{\hat{A}_E(i)\}_{i \in B_E}$:

$$\hat{A}_S(i) = \sum_{j \in B_{S1}} \beta_{i,j}\hat{A}_S(j) + \sum_{j \in B_E} \gamma_{i,j}\hat{A}_E(j) = v_{Si} + v_{Ei},$$

where $v_{Si} \in RS(\hat{A}_S)$ and $v_{Ei} \in RS(\hat{A}_E)$. Now define $\log_2 |\hat{\mathcal{X}}_S| \times n\hat{c}_B$ matrices $\hat{A}_{S1}$ and $\hat{A}_{S2}$ with respect to this decomposition. We let $\hat{A}_{S1}(i) = v_{Si}$ and $\hat{A}_{S2}(i) = v_{Ei}$. We observe that $\hat{A}_S = \hat{A}_{S1} + \hat{A}_{S2}$ and $rank(\hat{A}_{S1}) \le \delta n\hat{c}_B$. Since $\hat{\mathbf{x}}_\mathbf{S}\hat{A}_{S2} \in RS(\hat{A}_E)$, we can define a function $\phi'(\hat{\mathbf{x}}_\mathbf{S}) : \mathbb{F}_2^{n\sum \hat{R}_s} \to \mathbb{F}_2^{n\sum \hat{R}_e}$ such that $\phi'(\hat{\mathbf{x}}_\mathbf{S})\hat{A}_E = \hat{\mathbf{x}}_\mathbf{S}\hat{A}_{S2}$. Therefore, for all $\hat{\mathbf{x}}_\mathbf{S} \in \hat{\mathcal{X}}_\mathbf{s}$ and $\hat{\mathbf{x}}_\mathbf{E} \in \hat{\mathcal{X}}_\mathbf{e}$, we have

$$\hat{f}^l_B(\hat{\mathbf{x}}_\mathbf{S}, \hat{\mathbf{x}}_\mathbf{E}) = \hat{\mathbf{x}}_\mathbf{S}\hat{A}_{S1} + \hat{\mathbf{x}}_\mathbf{S}\hat{A}_{S2} + \hat{\mathbf{x}}_\mathbf{E}\hat{A}_E,$$
$$\hat{f}^l_B(\hat{\mathbf{x}}_\mathbf{S}, \hat{\mathbf{x}}_\mathbf{E} - \phi'(\hat{\mathbf{x}}_\mathbf{S})) = \hat{\mathbf{x}}_\mathbf{S}\hat{A}_{S1} + \hat{\mathbf{x}}_\mathbf{E}\hat{A}_E.$$

Finally, to satisfy the lemma, we set $\hat{A}$ to be $\hat{A}_{S1}$, and for a suitable $\hat{\mathbf{x}}^*_\mathbf{E} \in \hat{\mathcal{X}}_\mathbf{E}$ that satisfies the error requirement, we set $\phi(\hat{\mathbf{x}}_\mathbf{S}) = \hat{\mathbf{x}}^*_\mathbf{E} - \phi'(\hat{\mathbf{x}}_\mathbf{S})$ and $k^* = \hat{\mathbf{x}}^*_\mathbf{E}\hat{A}_E$.

We now pick a suitable $\hat{\mathbf{x}}^*_\mathbf{E}$ by way of an averaging argument. Namely, we would like to find $\hat{\mathbf{x}}^*_\mathbf{E}$ such that

$$\Pr[E(\hat{\mathbf{X}}_\mathbf{S}, \hat{\mathbf{x}}^*_\mathbf{E} - \phi'(\hat{\mathbf{X}}_\mathbf{S})) = 1] \le \epsilon.$$

For $\hat{\mathbf{X}}^*_\mathbf{E}$ uniformly distributed and independent of $\hat{\mathbf{X}}_\mathbf{S}$, it holds that $\hat{\mathbf{X}}^*_\mathbf{E} - \phi'(\hat{\mathbf{X}}_\mathbf{S})$ is also uniform, and thus

$$\Pr_{\hat{\mathbf{X}}_\mathbf{S}, \hat{\mathbf{X}}^*_\mathbf{E}}[E(\hat{\mathbf{X}}_\mathbf{S}, \hat{\mathbf{X}}^*_\mathbf{E} - \phi'(\hat{\mathbf{X}}_\mathbf{S})) = 1] \le \epsilon.$$

We conclude the existence of $\hat{\mathbf{x}}^*_\mathbf{E}$ as needed. ■

VII. Conclusion

This work addressed the capacity equivalence problem left open in [2]. We connect this problem to the edge removal problem [3], [4], and resolve it for the case of linear coding. The question whether one can determine the capacity region of network coding instances by those of index coding instances obtained through a reduction different than that used in [2] remains an intriguing open problem.

Acknowledgment

This material is based upon work supported by the National Science Foundation under Grant No. CCF-1018741 and ISF grant 480/08. This work was done while M. Langberg was at the California Institute of Technology, he would like to thank S. Jafar for suggesting to study Theorem 2, and S. El Rouayheb for helpful discussions.

References

[1] Z. Bar-Yossef, Y. Birk, T. S. Jayram, and T. Kol, "Index coding with side information," *Proceedings of IEEE Symposium on Foundations of Computer Science*, 2006.

[2] M. Effros, S. E. Rouayheb, and M. Langberg, "An equivalence between network coding and index coding," *arXiv*, arxiv.org/abs/1211.6660, 2012.

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

[4] T. Ho, M. Effros, and S. Jalali, "On equivalence between network topologies," *Allerton Conference on Communication, Control and Computing*, 2010.

[5] M. Langberg and M. Effros, "Network coding: Is zero error always possible?," *49th Annual Allerton Conference on IEEE*, 2011.

[6] M. Effros, "On capacity outer bounds for a simple family of wireless networks," *Information Theory and Applications Workshop*, 2010.