Fundamentals of Index Coding

Other titles in Foundations and Trends® in Communications and Information Theory

Multi-way Communications: An Information Theoretic Perspective
Anas Chaaban and Aydin Sezgin
ISBN: 978-1-60198-788-4

An Approximation Approach to Network Information Theory
A. Salman Avestimehr, Suhas N. Diggavi, Chao Tian and David N. C. Tse
ISBN: 978-1-68083-026-2

Energy Efficiency in Wireless Networks via Fractional Programming Theory
Alessio Zappone and Eduard Jorswieck
ISBN: 978-1-68083-042-2

Asymptotic Estimates in Information Theory with Non-Vanishing Error Probabilities
Vincent Y. F. Tan
978-1-60198-852-2

Fundamentals of Index Coding

Fatemeh Arbabjolfaei
Department of Electrical and Computer Engineering
University of California, San Diego
La Jolla, CA, 92093
farbabjo@ucsd.edu

Young-Han Kim
Department of Electrical and Computer Engineering
University of California, San Diego
La Jolla, CA, 92093
yhk@ucsd.edu

now
the essence of knowledge
Boston — Delft

Foundations and Trends® in Communications and Information Theory

Published, sold and distributed by:
now Publishers Inc.
PO Box 1024
Hanover, MA 02339
United States
Tel. +1-781-985-4510
www.nowpublishers.com
sales@nowpublishers.com

Outside North America:
now Publishers Inc.
PO Box 179
2600 AD Delft
The Netherlands
Tel. +31-6-51115274

The preferred citation for this publication is

F. Arbabjolfaei and Y.-H. Kim. *Fundamentals of Index Coding*. Foundations and Trends® in Communications and Information Theory, vol. 14, no. 3-4, pp. 163–346, 2018.

ISBN: x978-1-68083-493-2

Foundations and Trends® in Communications and Information Theory

Volume 14, Issue 3-4, 2018

Editorial Scope

Topics

Foundations and Trends® in Communications and Information Theory publishes survey and tutorial articles in the following topics:

- Coded modulation
- Coding theory and practice
- Communication complexity
- Communication system design
- Cryptology and data security
- Data compression
- Data networks
- Demodulation and Equalization
- Denoising
- Detection and estimation
- Information theory and statistics
- Information theory and computer science
- Joint source/channel coding
- Modulation and signal design
- Multiuser detection
- Multiuser information theory
- Optical communication channels
- Pattern recognition and learning
- Quantization
- Quantum information processing
- Rate-distortion theory
- Shannon theory
- Signal processing for communications
- Source coding
- Storage and recording codes
- Speech and Image Compression
- Wireless Communications

Information for Librarians

Foundations and Trends® in Communications and Information Theory, 2018, Volume 14, 4 issues. ISSN paper version 1567-2190. ISSN online version 1567-2328 . Also available as a combined paper and online subscription.

Contents

Fundamentals of Index Coding

Fatemeh Arbabjolfaei[1] and Young-Han Kim[2]

[1] *Department of Electrical and Computer Engineering, University of California, San Diego; farbabjo@ucsd.edu*
[2] *Department of Electrical and Computer Engineering, University of California, San Diego; yhk@ucsd.edu*

ABSTRACT

Index coding is a canonical problem in network information theory that studies the fundamental limit and optimal coding schemes for broadcasting multiple messages to receivers with different side information. The index coding problem provides a simple yet rich model for several important engineering tasks such as satellite communication, content broadcasting, distributed caching, device-to-device relaying, and interference management. This monograph aims to provide a broad overview of this fascinating subject, focusing on the simplest form of multiple-unicast index coding. A unified treatment on coding schemes based on graph-theoretic, algebraic, and information-theoretic approaches is presented. Although the problem of characterizing the optimal communication rate is open in general, several bounds and structural properties are established. The relationship to other problems such as network coding and distributed storage is also discussed.

Fatemeh Arbabjolfaei and Young-Han Kim (2018), "Fundamentals of Index Coding", Foundations and Trends® in Communications and Information Theory: Vol. 14, No. 3-4, pp 163–346. DOI: 10.1561/0100000094.

1

Introduction

1.1 Motivation

We open our discussion with a simple example. Consider the wireless communication system consisting of one server and three receivers, as depicted in Figure 1.1. The server has three distinct messages x_1, x_2, and x_3. Receiver $i \in \{1, 2, 3\}$ is interested in message x_i and has some of the other messages as side information. In particular, receiver 1 has message x_2 as side information, receiver 2 has x_1 and x_3, and receiver 3 has x_1. The server wishes to communicate all the messages to their designated receivers using the minimum possible number of broadcast transmissions.

The most naive strategy for the server to achieve this goal is to send one message at a time, which takes overall three transmissions. Alternatively, if the server transmits two coded messages $x_1 + x_2$ and x_3 (assuming that the messages can be represented in an alphabet with well-defined addition), then each receiver can recover its desired message using the received coded messages and its side information. Indeed, receiver 1 can recover x_1 from the received message $x_1 + x_2$ and its side information x_2. Similarly, receiver 2 can recover x_2 from $x_1 + x_2$ and x_1. Receiver 3 can trivially recover x_3. This simple example

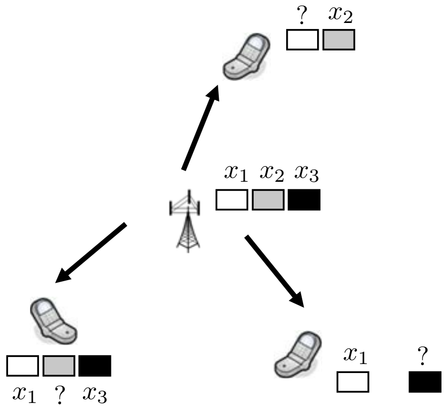

Figure 1.1: An index coding example with three receivers.

illustrates that sending *coded messages* may decrease the number of broadcast transmissions.

Generalizing our initial example, we study the communication problem depicted in Figure 1.2, which is commonly referred to as *index coding*. In this canonical problem in network information theory, a server has a tuple of n messages

$$x^n := (x_1, \ldots, x_n)$$

and broadcasts a coded message y generated from x^n to n receivers. Each receiver $i \in [n] := \{1, 2, \ldots, n\}$ wishes to recover the message x_i from y and a set of other messages

$$x(A_i) := (x_j, j \in A_i), \quad A_i \subseteq [n] \setminus \{i\}$$

it already has as side information. Assuming that the side information sets $A_1, \ldots, A_n$ are known prior to the communication, we are interested in devising a coding scheme that exploits the side information at the receivers and broadcasts the messages reliably with the minimum amount of transmissions.

The index coding problem was introduced by Birk and Kol [26, 27] in the context of satellite communication. Related formulations were studied earlier by Celebiler and Stette [40], Wyner, Wolf, and

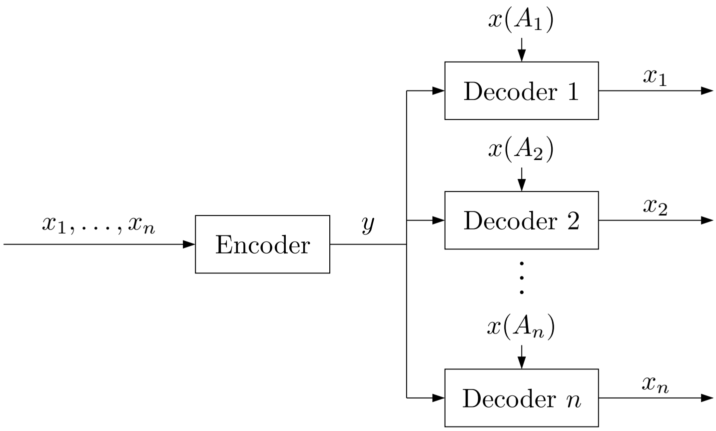

Figure 1.2: The index coding problem.

Willems [157, 161], and Yeung [163]. The term "index coding" is due to Bar-Yossef, Birk, Jayram, and Kol [22], who compared the index coding problem to the problem of zero-error source coding with side information studied by Witsenhausen [158] and contrasted the fact that in the index coding problem the receiver wishes to recover a single component of the source, the *index* of which is unknown to the sender. Hence, as for compound channels [28, 54, 159], the sender can proactively encode its transmission and *broadcast* to the receiver in all possible configurations [49]. In addition to satellite communication, index coding has applications in diverse areas such as multimedia distribution [117], interference management [83], coded caching [106, 84], and distributed computing [96]. This problem is also closely related to many other important problems such as network coding [125, 62, 60], locally recoverable distributed storage [111, 132, 13], guessing games on directed graphs [125, 167, 13], matroid theory [62], and zero-error capacity of channels [135]. Due to this significance, the index coding problem has been broadly studied over the past two decades. Tools from various disciplines including graph theory, coding theory, and information theory have been utilized to propose numerous nontrivial coding schemes [26, 104, 23, 44, 117, 30, 8, 107, 83, 134, 7, 9, 153, 119, 82, 145, 150, 167], as well as performance bounds [165, 23, 56, 29, 18, 145].

1.2 Formal Definition of the Problem

We formulate the problem more precisely. A $(t_1, \dots, t_n, r)$ *index code* is defined by

- n messages, where the i-th message $x_i = (x_{i1}, \dots, x_{it_i})$ takes values from $\mathcal{X}^{t_i}$ for some common finite alphabet $\mathcal{X}$,
- an encoder $\phi : \prod_{i=1}^n \mathcal{X}^{t_i} \to \mathcal{X}^r$ that maps the message n-tuple x^n to an index $y = (y_1, \dots, y_r) \in \mathcal{X}^r$, and
- n decoders, where the decoder at receiver $i \in [n]$, $\psi_i : \mathcal{X}^r \times \prod_{j \in A_i} \mathcal{X}^{t_j} \to \mathcal{X}^{t_i}$, maps the received index $y = \phi(x^n)$ and the side information $x(A_i)$ back to x_i.

Thus, for every $x^n \in \prod_{i=1}^n \mathcal{X}^{t_i}$,

$$\psi_i(\phi(x^n), x(A_i)) = x_i, \quad i \in [n].$$

A $(t, \dots, t, r)$ code will be written simply as a (t, r) code.

Suppose that $\mathcal{X}$ is an alphabet over which linear operations are well-defined, for example, a finite field $\mathbb{F}_q$ or a ring (see, for example, [48]). If the encoder of a code is a linear function of $x^n = (x_1^{t_1}, \dots, x_n^{t_n}) = (x_{11}, \dots, x_{1t_1}, \dots, x_{n1}, \dots, x_{nt_n})$ and the decoders are also linear functions of x^n (and $y = (y_1, \dots, y_r)$), then the code is referred to as a *linear* index code. If $t_i = 1$ for all $i \in [n]$, then the linear index code is said to be a *scalar* linear index code. Otherwise, the code is a *vector* linear index code.

As an example, for the 3-message index coding problem in Figure 1.1, a $(1, 1, 1, 2)$ index code with $\mathcal{X} = \mathbb{F}_2$, the encoder defined by

$$y_1 = x_1 + x_2 \quad \text{and} \quad y_2 = x_3,$$

and the decoders defined by

$$x_1 = y_1 + x_2, \quad x_2 = y_1 + x_3, \quad \text{and} \quad x_3 = y_2,$$

where in both cases the addition operations are in $\mathbb{F}_2$, is a scalar linear code.

A tuple $(R_1, \ldots, R_n) \in \mathbb{R}^n_{\geq 0}$ of nonnegative real numbers is said to be an *achievable* rate tuple for the index coding problem if there exists a $(t_1, \ldots, t_n, r)$ index code such that

$$R_i \leq \frac{t_i}{r}, \quad i \in [n].$$

The *capacity region* $\mathscr{C}$ of the index coding problem is defined as the closure of the set of all achievable rate tuples. The ultimate goal of studying the index coding problem is to characterize the capacity region of a general index coding problem and develop a simple coding scheme that achieves or approximates the capacity region.

Remark 1.1. Our definition of capacity region, with the stringent requirement of perfect, zero-error recovery of the messages, should be distinguished from the more common definition of *vanishing-error capacity region* in network information theory that allows for arbitrarily small probability of error. It can be shown, however, that these two capacity regions coincide; see Appendix 1.A for details.

The definition of the capacity region depends on the alphabet $\mathcal{X}$ on which the messages are defined and one may well denote it by $\mathscr{C}_{\mathcal{X}}$ to emphasize this dependence. As we will prove in Appendix 1.B, however, the choice of $\mathcal{X}$ is irrelevant to the actual capacity region itself.

Lemma 1.1. For any two finite alphabets $\mathcal{X}$ and $\mathcal{X}'$,

$$\mathscr{C}_{\mathcal{X}} = \mathscr{C}_{\mathcal{X}'}.$$

Consequently, we assume without loss of generality that $\mathcal{X} = \mathbb{F}_2$ for a general index code and consequently that the base of logarithm is 2 throughout, unless specified otherwise.

By limiting our attention to *linear* codes, we can similarly define *linearly achievable* rate tuple and *linear capacity region* $\mathscr{L}$. In contrast to the capacity region, the linear capacity region of the index coding problem may depend on the chosen alphabet $\mathcal{X}$ and indeed does so [104] (see Section 6.7).

The capacity region, linear or nonlinear, is closed by definition. Based on the standard time-sharing argument (see, for example, [61, Sec. 4.4]), it can be readily checked to be *convex*.

In many cases, it is convenient to focus on a single performance metric instead of a multidimensional region. Let $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n) \in \mathbb{R}^n_{\geq 0}$ be a tuple of nonnegative real numbers. The $\boldsymbol{\mu}$-*directed capacity* of the index coding problem is defined as

$$C(\boldsymbol{\mu}) = \max\{R\colon R\boldsymbol{\mu} \in \mathscr{C}\}.$$

Remark 1.2. The capacity region can be written in terms of $\boldsymbol{\mu}$-directed capacities as

$$\mathscr{C} = \bigcup_{\boldsymbol{\mu}} \{(R_1, \ldots, R_n)\colon R_i \leq C(\boldsymbol{\mu})\mu_i,\, i \in [n]\}. \tag{1.1}$$

Note that if $\boldsymbol{\mu} = c\boldsymbol{\mu}'$ for some constant c, then $C(\boldsymbol{\mu})\boldsymbol{\mu} = C(\boldsymbol{\mu}')\boldsymbol{\mu}'$ and thus in (1.1), it suffices to take the union only over normalized vectors, e.g., over $\boldsymbol{\mu}$ such that $\sum_{i=1}^n \mu_i = n$.

The **1**-directed capacity of the index coding problem is referred to as the *symmetric capacity* (or the *capacity* in short), that is,

$$C_{\text{sym}} = C(\mathbf{1}) = \max\{R\colon (R, \ldots, R) \in \mathscr{C}\}.$$

The symmetric capacity can be equivalently defined as

$$C_{\text{sym}} = \sup_r \sup_{(t,r)\text{ codes}} \frac{t}{r} = \lim_{r\to\infty} \sup_{(t,r)\text{ codes}} \frac{t}{r},$$

where the equality between the supremum and the limit follows by Fekete's lemma [69] and the superadditivity

$$\sup_{(t,r_1+r_2)\text{ codes}} t \geq \sup_{(t_1,r_1)\text{ codes}} t_1 + \sup_{(t_2,r_2)\text{ codes}} t_2.$$

The reciprocal of the symmetric capacity,

$$\beta = \frac{1}{C_{\text{sym}}},$$

is referred to as the *broadcast rate*, which can be alternatively defined as

$$\beta = \inf_t \inf_{(t,r)\text{ codes}} \frac{r}{t} = \lim_{t\to\infty} \inf_{(t,r)\text{ codes}} \frac{r}{t}. \tag{1.2}$$

The *linear broadcast rate* is similarly defined as

$$\lambda = \inf_{t} \inf_{(t,r) \text{ linear codes}} \frac{r}{t} = \lim_{t\to\infty} \inf_{(t,r) \text{ linear codes}} \frac{r}{t}.$$

As with the linear capacity region $\mathscr{L}$, the linear broadcast rate depends on the underlying alphabet $\mathcal{X}$ and will be sometimes denoted $\lambda_{\mathcal{X}}$ to emphasize this dependence. For any $\mathcal{X}$, $\beta \le \lambda_{\mathcal{X}}$.

Note that the capacity region $\mathscr{C}$ of the index coding problem includes the simplex $\{(R_1, \ldots, R_n) \colon R_1 + \cdots + R_n = 1, R_i \ge 0,\ i \in [n]\}$ and is included in the hypercube $\{(R_1, \ldots, R_n) \colon 0 \le R_i \le 1,\ i \in [n]\}$. Consequently, the capacity and the broadcast rate are bounded as $1/n \le C_{\text{sym}} \le 1$ and $1 \le \beta \le n$, respectively. Similar bounds hold for $\mathscr{L}$ and λ as well.

Any instance of the index coding problem is fully determined by the side information sets $A_1, \ldots, A_n$, and is represented compactly by a sequence $(i|A_i), i \in [n]$. For example, the 3-message index coding problem with $A_1 = \{2\}, A_2 = \{1, 3\}$, and $A_3 = \{1\}$ in Figure 1.1 is represented as

$$(1|2), (2|1, 3), (3|1).$$

Each instance of the index coding problem can be equivalently specified by a directed graph $G = (V, E)$ with n vertices, referred to as the *side information graph* [27, 44]. Each vertex of G corresponds to a receiver (and its associated message) and there is a directed edge $j \to i$ if and only if (iff) receiver i knows message x_j as side information, i.e., $j \in A_i$ (see Figure 1.3). The reader is cautioned that in the literature the opposite convention is sometimes used to describe the availability of side information, in which there is a directed edge $i \to j$ if $j \in A_i$.

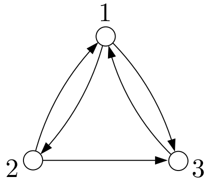

Figure 1.3: The graph representation for the index coding problem with $A_1 = \{2, 3\}, A_2 = \{1\}$, and $A_3 = \{1, 2\}$.

Either way, the number of index coding problems with n messages is $2^{n(n-1)}$, which blows up quickly with n. Even when we remove *isomorphic* (i.e., symmetric up to vertex relabeling) instances and concentrate on nonisomorphic instances, the number of such problems is equal to that of nonisomorphic directed graphs with n vertices [151, Seq. A000273], which grows superexponentially. Throughout the monograph, we identify an instance of the index coding problem with its side information graph G and often write "index coding problem G." We also denote the broadcast rate and the capacity region of problem G with $\beta(G)$ and $\mathscr{C}(G)$, respectively, when this dependence is to be emphasized.

The index coding problem can be also formulated as a special case of the multiple-unicast network coding problem [2, 164]; see Section 10 for a self-contained description of the latter. For example, the index coding problem with $A_1 = \{2\}, A_2 = \{1, 3\}$, and $A_3 = \{1\}$ can be represented as a network coding problem depicted in Figure 1.4. In this network coding graph, each edge (solid line) represents a link of unit capacity and each vertex represents a node. There are three messages communicated from source nodes on the left to destination nodes on the right (depicted by dashed lines). Each source node is connected to the top left node, which encodes the messages into an index (coded message) and communicates it to the top right node (under the capacity constraint). The top right node is connected to each destination node and forwards (broadcasts) the index to all of them. The remaining edges connect source nodes and destination nodes directly according to the availability pattern of side information. The essence of each problem

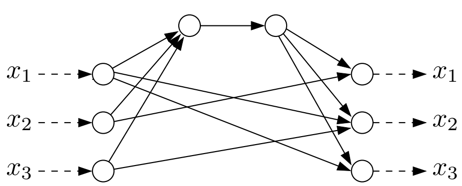

Figure 1.4: The network coding representation of the index coding problem with $A_1 = \{2\}, A_2 = \{1, 3\}$, and $A_3 = \{1\}$.

instance is captured by such direct connections between the source and destination nodes. If we consider the subgraph consisting of only source and destination nodes and overlap each source–destination pair, then we recover the side information graph in Figure 1.3.

1.3 Objectives and the Organization

As mentioned earlier, our main objectives in studying the index coding problem are to characterize the capacity region for a general index coding instance in a computable expression and to develop the coding scheme that can achieve it. Despite their simplicity, these two closely related questions are extremely difficult and precise answers to them, after twenty years of vigorous investigation, are still in terra incognita. There are, nonetheless, many elegant results that shed light on the fundamental challenges in multiple-unicast network communication and expose intriguing interplay between coding theory, graph theory, and information theory. This monograph thus aims to provide a concise survey of these results in a unified framework.

To facilitate the development of this framework, the rest of the monograph is organized as follows. Section 2 reviews some known results in graph theory that will be recalled frequently throughout. Our main story starts with Section 3, which presents a few noncomputable characterizations of the capacity of a general index coding problem in graph-theoretic and information-theoretic expressions. As a main application of these characterizations, we present basic structural properties of index coding capacity in Section 4. The next two sections develop upper and lower bounds on the capacity. In Section 5, we establish a few capacity upper bounds and the relationships among them. In Section 6, we develop several coding schemes based on algebraic, graph-theoretic, and information-theoretic tools along with the corresponding lower bounds on the capacity. Section 7 is devoted to the notion of criticality, namely, whether the side information graph cannot be reduced without lowering the capacity, and presents necessary and sufficient conditions for an index coding problem to be critical. In Section 8, we combine the results in Sections 4 through 7 to characterize the capacity of several classes of the index coding problem. The capacity approximation results beyond these classes of problems are presented in Section 9. The next two sections explore the connection between index coding and other related problems. In Section 10, we relate index coding to the well-known multiple-unicast network coding problem. In Section 11, we present the intriguing duality between index coding, locally recoverable distributed storage, and guessing games. We conclude our discussion in Section 12 by pointing out numerous variations and extensions of the basic index coding problem presented thus far. Some technical proofs are relegated to the end of each section.

1.A Capacity Region Under Average Error Probability Criterion

Let X_i and $\hat{X}_i$ be random variables representing the i-th message and its estimate, respectively. Suppose that $(X_1, \ldots, X_n)$ is uniformly distributed over $\mathcal{X}^{t_1} \times \cdots \times \mathcal{X}^{t_n}$, i.e., the messages are uniformly distributed and independent of each other. A rate tuple $(R_1, \ldots, R_n)$ is said to be *vanishing-error achievable* if for every $\epsilon > 0$, there exists a $(t_1, \ldots, t_n, r)$ code with $R_i \leq t_i/r$, $i \in [n]$, such that the average probability of error

$$P_e = \mathsf{P}\{(\hat{X}_1, \ldots, \hat{X}_n) \neq (X_1, \ldots, X_n)\} \leq \epsilon. \tag{1.3}$$

Equivalently, a rate tuple $(R_1, \ldots, R_n)$ is vanishing-error achievable if there exists a sequence of $(\lceil rR_1 \rceil, \ldots, \lceil rR_n \rceil, r)$ index codes such that the average probability of error P_e converges to 0 as $r \to \infty$. The *vanishing-error capacity region* $\mathscr{C}^*$ of the index coding problem is the closure of the set of all vanishing-error achievable rate tuples $(R_1, \ldots, R_n)$.

For a general network communication problem, the vanishing-error capacity region and the (zero-error) capacity region are not the same [59, 154]. For a network with a single sender that broadcasts multiple messages, however, these two regions can be shown to be identical [156] (see also [61, Problem 8.11]), which was rediscovered by Chan and Grant [42], and Langberg and Effros [92] in the context of index coding and single-sender network coding problems.

Lemma 1.2. For any index coding problem, $\mathscr{C} = \mathscr{C}^*$.

The proof is delegated to Appendix 1.C. One can similarly define *vanishing-error linearly achievability.* The *vanishing-error linear capacity region* $\mathscr{L}^*$ is then defined to be the closure of the set of all vanishing-error achievable rate tuples, which is also the same as the (zero-error) linear capacity region.

Lemma 1.3. For any $\mathcal{X} = \mathbb{F}_q$, $\mathscr{L} = \mathscr{L}^*$.

To prove Lemma 1.3, we establish the following stronger result.

Lemma 1.4. For any linear index code, if the probability of error $P_e < 1/2$, then $P_e = 0$.

Proof. Let $t_\Sigma = \sum_{i=1}^n t_i$. Every linear encoder ϕ is specified by a matrix $\Phi \in \mathbb{F}_q^{r \times t_\Sigma}$ such that $y = \phi(x^n) = \Phi x^n$. Assume by contradiction that $0 < P_e < 1/2$. Since the probability of error is nonzero, there exist distinct $\tilde{x}^n, \tilde{z}^n \in \mathbb{F}_q^{t_\Sigma}$ such that $\Phi\tilde{x}^n = \Phi\tilde{z}^n$ and $\tilde{x}(A_i) = \tilde{z}(A_i)$ for some $i \in [n]$, or equivalently, $\Phi e^n = 0$ and $e(A_i) = 0$, where $e^n = \tilde{x}^n - \tilde{z}^n$. Now for every x^n, since the code is linear, there exists $z^n = x^n + e^n$ for which $\Phi x^n = \Phi z^n$ and $x(A_i) = z(A_i)$ for some i. Therefore, at most one of x^n and z^n can be recovered correctly and consequently the average probability of error $P_e \geq 1/2$, which contradicts the assumption. □

We are now ready to prove Lemma 1.3. Clearly $\mathscr{L} \subseteq \mathscr{L}^*$. Thus, it suffices to show that $\mathscr{L}^* \subseteq \mathscr{L}$. Let $(R_1, \ldots, R_n)$ be a vanishing error linearly achievable rate tuple. Then, by definition, there exists a sequence of $(\lceil rR_1 \rceil, \ldots, \lceil rR_n \rceil, r)$ index codes for which (1.3) is satisfied. Therefore, there exists a sufficiently large r such that the error probability of the index code $(\lceil rR_1 \rceil, \ldots, \lceil rR_n \rceil, r)$ is less than 1/2. By Lemma 1.4, the error probability of this code is zero and thus, $(R_1, \ldots, R_n)$ is also a (zero-error) linearly achievable rate tuple. Hence, we have $\mathscr{L}^* \subseteq \mathscr{L}$, which completes the proof.

1.B Proof of Lemma 1.1

Let I and I' be index coding instances defined over finite alphabets $\mathcal{X}$ and $\mathcal{X}'$, respectively, and let $\mathscr{A}$ and $\mathscr{A}'$ be the associated sets of achievable rate tuples. We consider two cases.

Case 1. $\log_{|\mathcal{X}|}|\mathcal{X}'|$ is a rational number, i.e., $\log_{|\mathcal{X}|}|\mathcal{X}'| = a/b$ for some $a, b \in \mathbb{N}$. To show that the capacity regions are equal, it suffices to show $\mathscr{A} = \mathscr{A}'$. Suppose that $\mathbf{R} = (R_1, \ldots, R_n) \in \mathscr{A}$. Then, by definition, there exists a $(\mathbf{t}, r) = (t_1, \ldots, t_n, r)$ code for problem I such that $R_i \le t_i/r$, $i \in [n]$. Repeat the $(\mathbf{t}, r)$ code a times to construct an $(a\mathbf{t}, ar)$ index code for problem I. Since the two instances are both defined on the same set of side information, and $|\mathcal{X}|^a = |\mathcal{X}'|^b$, this leads to a $(b\mathbf{t}, br)$ code for problem I'. Therefore, $\mathbf{R} \in \mathscr{A}'$, and thus $\mathscr{A} \subseteq \mathscr{A}'$. By similar steps we can show $\mathscr{A}' \subseteq \mathscr{A}$, which completes the proof.

Case 2. $\log_{|\mathcal{X}|}|\mathcal{X}'|$ is irrational. First, we show that $\mathscr{A}' \subseteq \mathscr{C}_{\mathcal{X}}$. Suppose that $\mathbf{R} \in \mathscr{A}'$. Then, by definition, there exists a $(\mathbf{t}, r)$ index code for problem I' such that $R_i \le t_i/r$, $i \in [n]$. For any $b \in \mathbb{N}$ sufficiently large, there exists $a \in \mathbb{N}$ such that $a/b < \log_{|\mathcal{X}|}|\mathcal{X}'| < (a+1)/b$. Construct a $(b\mathbf{t}, br)$ index code for problem I' by repeating the $(\mathbf{t}, r)$ code b times. Since $|\mathcal{X}|^a < |\mathcal{X}'|^b < |\mathcal{X}|^{a+1}$ and the two problems are defined on the same set of side information, a $(a\mathbf{t}, (a+1)r)$ code for problem I can be constructed from the $(b\mathbf{t}, br)$ code for problem I'. Letting $b \to \infty$ (and hence $a \to \infty$) proves that $\mathbf{R} \in \mathscr{C}_{\mathcal{X}}$, and thus $\mathscr{A}' \subseteq \mathscr{C}_{\mathcal{X}}$. Since $\mathscr{C}_{\mathcal{X}}$ is closed, we have $\mathscr{C}_{\mathcal{X}'} \subseteq \mathscr{C}_{\mathcal{X}}$. By similar steps we can show $\mathscr{C}_{\mathcal{X}} \subseteq \mathscr{C}_{\mathcal{X}'}$, which completes the proof.

1.C Proof of Lemma 1.2

We adapt Telatar's simplification of the classical proof by Willems [156] on the invariance of the broadcast channel capacity region under the average and maximal error probability criteria that appeared in [61, Problem 8.11].

It is trivial to see that $\mathscr{C} \subseteq \mathscr{C}^*$. We thus prove the other direction. Let $(R_1, \ldots, R_n) \subseteq \mathscr{C}^*$. Then for every $\epsilon > 0$, there exists a sequence of $(t_1, \ldots, t_n, r)$ codes with $t_i = \lceil rR_i \rceil$, $i \in [n]$, such that the average probability of error $P_e \le \epsilon$ for r sufficiently large. Assume without loss of generality that $R_i > 0$, $i \in [n]$. (Otherwise, the message x_i of zero rate $R_i = 0$ is fixed and can be ignored.) We will identify the set $\{0,1\}^t$ of all t-bit sequences with the set $[2^t] = \{1, \ldots, 2^t\}$ of integers throughout. Then the set of codewords of the $(t_1, \ldots, t_n, r)$ index code can be expressed as

$$\mathcal{C} = \{\phi(x^n) \in [2^r] : x^n \in [2^{t_1}] \times \cdots \times [2^{t_n}]\},$$

which will be referred to as the *codebook*. For each message tuple x^n, we define its probability of error as

$$P_e(x^n) = \mathsf{P}\{X^n \neq \hat{X}^n \mid X^n = x^n\}. \tag{1.4}$$

Note that $P_e(x^n)$ is either 0 or 1 for any index code. We say that a codeword $\phi(x^n)$ is said to be "bad" if the corresponding $P_e(x^n) = 1$. Since the average probability of error is

$$\epsilon = \frac{1}{2^{t_\Sigma}} \sum_{x^n} P_e(x^n),$$

there are $2^{t_\Sigma}\epsilon$ "bad" codewords $\phi(x^n)$, where $t_\Sigma = \sum_{i=1}^n t_i$. Randomly and independently permute the messages $x_1, \ldots, x_n$ to generate a new codebook $\overline{\mathcal{C}}$ that consists of codewords $\phi(\pi_1(x_1), \ldots, \pi(x_n))$, where $\pi_1, \ldots, \pi_n$ denote the independent random permutations.

We now proceed with the *multicoding* technique by Marton [108] originally developed for broadcast channels. We partition the codebook $\overline{\mathcal{C}}$ into subcodebooks $\overline{\mathcal{C}}(x_1', \ldots, x_n')$ for a new set of message tuples $(x_1', \ldots, x_n') \in [2^{t_1}/r^2] \times \cdots \times [2^{t_n}/r^2]$, each subcodebook consisting of $r^2 \times \cdots \times r^2 = r^{2n}$ codewords of length r. We will show that there exists a new encoder $\phi'(x_1', \ldots, x_n')$ that maps each message tuple $(x_1', \ldots, x_n')$ to some codeword in the corresponding subcodebook $\overline{\mathcal{C}}(x_1', \ldots, x_n')$, so that every codeword $\phi'(x_1', \ldots, x_n')$ is "good" (= not "bad") and hence distinguishable from the rest with zero error. Since the rate of the new code is $R_i' = (t_i - 2\log r)/r = (\lceil rR_i \rceil - 2\log r)/r$, which converges to the original rate R_i as $r \to \infty$, there is no asymptotic rate loss for achieving the zero error probability. Details on the existence of $\phi'(x_1', \ldots, x_n')$ are as follows.

First note that every subcodebook has the same distribution as the set

$$\{\phi(\pi_1(x_1), \ldots, \pi_n(x_n))\colon x_1, \ldots, x_n \in [r^2]\}.$$

The probability that all r^{2n} codewords in this set are "bad" is upper bounded by the probability that all r^2 "diagonal" codewords, that is, all codewords in

$$\{\phi(\pi_1(x), \ldots, \pi_n(x)) \colon x \in [r^2]\},$$

are "bad." Since the permutations are independent and there are $2^{t_\Sigma}\epsilon$ "bad" codewords, the probability that all "diagonal" codewords are "bad" is upper bounded by

$$\frac{2^{t_\Sigma}\epsilon}{\prod_{i=1}^n 2^{t_i}} \cdot \frac{2^{t_\Sigma}\epsilon - 1}{\prod_{i=1}^n (2^{t_i} - 1)} \cdot \frac{2^{t_\Sigma}\epsilon - 2}{\prod_{i=1}^n (2^{t_i} - 2)} \cdots \frac{2^{t_\Sigma}\epsilon - (r^2 - 1)}{\prod_{i=1}^n (2^{t_i} - (r^2 - 1))}$$
$$\leq \left(\prod_{i=1}^n \frac{2^{t_i}}{2^{t_i} - (r^2 - 1)}\right)^{r^2} \epsilon^{r^2},$$

which is further upper bounded by $(2\epsilon)^{r^2}$ for r sufficiently large (since $2^{t_i}/(2^{t_i} - (r^2 - 1)) \to 1$ as $r \to \infty$).

Next, since every subcodebook has the same distribution, the expected number of subcodebooks for which all of their constituent codewords are "bad" is upper bounded by

$$\frac{2^{t_\Sigma}}{r^{2n}}(2\epsilon)^{r^2},$$

which is the product of the number of all subcodebooks and the probability bound of $(2\epsilon)^{r^2}$ we computed above. Since this bound tends to zero as $r \to \infty$, there exists at least one permutation tuple $(\pi_1, \ldots, \pi_n)$ such that every subcodebook has at least one codeword that is "good." Hence, we can define the new encoder $\phi'(x'_1, \ldots, x'_n)$ that maps each message tuple $(x'_1, \ldots, x'_n) \in [2^{t_1}/r^2] \times \cdots \times [2^{t_n}/r^2]$ to a "good" codeword in the subcodebook $\overline{\mathcal{C}}(x'_1, \ldots, x'_n)$. This completes the proof of Lemma 1.2.

References

[1] Agarwal, A., L. Flodin, and A. Mazumdar (2018). "Linear programming approximations for index coding". URL: http://arxiv.org/abs/1807.07193.

[2] Ahlswede, R., N. Cai, S.-Y. R. Li, and R. W. Yeung (2000). "Network information flow". *IEEE Trans. Inf. Theory*. 46(4): 1204–1216.

[3] Alon, N., A. Hassidim, E. Lubetzky, U. Stav, and A. Weinstein (2008). "Broadcasting with side information". In: *Proc. 49th Ann. IEEE Symp. Found. Comput. Sci.* Philadelphia, PA. 823–832.

[4] Alon, N. and N. Kahale (1998). "Approximating the independence number via the Θ-function". *Math. Prog.* 80(3): 253–264.

[5] Alon, N. and A. Orlitsky (1996). "Source coding and graph entropies". *IEEE Trans. Inf. Theory*. 42(5): 1329–1339.

[6] Appel, K., W. Haken, and J. Koch (1977). "Every planar map is four colorable—II: Reducibility". *Illinois J. Math.* 21(3): 491–567.

[7] Arbabjolfaei, F., B. Bandemer, and Y.-H. Kim (2014). "Index coding via random coding". In: *Proc. 2nd Iran Workshop Commun. Inf. Theory*. Tehran, Iran. 1–7.

[8] Arbabjolfaei, F., B. Bandemer, Y.-H. Kim, E. Sasoglu, and L. Wang (2013). "On the capacity region for index coding". In: *Proc. IEEE Int. Symp. Inf. Theory*. Istanbul, Turkey. 962–966.

[9] Arbabjolfaei, F. and Y.-H. Kim (2014). "Local time sharing for index coding". In: *Proc. IEEE Int. Symp. Inf. Theory*. Honolulu, HI. 286–290.

[10] Arbabjolfaei, F. (2017). "Index coding: Fundamental limits, coding schemes, and structural properties". *Ph.D. Thesis*. La Jolla, CA: University of California, San Diego.

[11] Arbabjolfaei, F. and Y.-H. Kim (2015a). "On critical index coding problems". In: *Proc. IEEE Inf. Theory Workshop*. Jeju Island, Korea. 9–13.

[12] Arbabjolfaei, F. and Y.-H. Kim (2015b). "Structural properties of index coding capacity using fractional graph theory". In: *Proc. IEEE Int. Symp. Inf. Theory*. Hong Kong. 1034–1038.

[13] Arbabjolfaei, F. and Y.-H. Kim (2015c). "Three stories on a two-sided coin: Index coding, locally recoverable distributed storage, and guessing games on graphs". In: *Proc. 53rd Ann. Allerton Conf. Comm. Control Comput.* Monticello, IL. 843–850.

[14] Arbabjolfaei, F. and Y.-H. Kim (2016). "Approximate capacity of index coding for some classes of graphs". In: *Proc. IEEE Int. Symp. Inf. Theory*. Barcelona, Spain. 2154–2158.

[15] Arbabjolfaei, F. and Y.-H. Kim (2017). "Generalized lexicographic products and the index coding capacity". URL: http://arxiv.org/abs/1608.03689.

[16] Arbabjolfaei, F., Y.-H. Kim, and P. Sadeghi (2018). "GitHub index coding repository". URL: http://github.com/index-coding.

[17] Asadi, B., L. Ong, and S. J. Johnson (2015). "Optimal coding schemes for the three-receiver AWGN broadcast channel with receiver message side information". *IEEE Trans. Inf. Theory*. 61(10): 5490–5503.

[18] Baber, R., D. Christofides, A. N. Dang, S. Riis, and E. R. Vaughan (2013). "Multiple unicasts, graph guessing games, and non-Shannon inequalities". In: *Proc. Int. Symp. Netw. Coding*. Calgary, AB. 1–6.

[19] Bachem, A. and W. Kern (1992). *Linear Programming Duality: An Introduction to Oriented Matroids*. Berlin: Springer.

[20] Banawan, K. and S. Ulukus (2018). "The capacity of private information retrieval from coded databases". *IEEE Trans. Inf. Theory*. 64(3): 1945–1956.

[21] Bandemer, B., A. El Gamal, and Y.-H. Kim (2015). "Optimal achievable rates for interference networks with random codes". *IEEE Trans. Inf. Theory*. 61(12): 6536–6549.

[22] Bar-Yossef, Z., Y. Birk, T. S. Jayram, and T. Kol (2006). "Index coding with side information". In: *Proc. 47th Ann. IEEE Symp. Found. Comput. Sci.* Berkeley, CA. 197–206.

[23] Bar-Yossef, Z., Y. Birk, T. S. Jayram, and T. Kol (2011). "Index coding with side information". *IEEE Trans. Inf. Theory*. 57(3): 1479–1494.

[24] Belmonte, R., P. Heggernes, P. van't Hof, A. Rafiey, and R. Saei (2014). "Graph classes and Ramsey numbers". *Discrete Appl. Math.* 173: 16–27.

[25] Berliner, Y. and M. Langberg (2011). "Index coding with outerplanar side information". In: *Proc. IEEE Int. Symp. Inf. Theory*. Saint Petersburg, Russia. 806–810.

[26] Birk, Y. and T. Kol (1998). "Informed-source coding-on-demand (ISCOD) over broadcast channels". In: *Proc. 17th Ann. IEEE Int. Conf. Comput. Commun. (INFOCOM)*. San Francisco, CA. 1257–1264.

[27] Birk, Y. and T. Kol (2006). "Coding on demand by an informed source (ISCOD) for efficient broadcast of different supplemental data to caching clients". *IEEE Trans. Inf. Theory*. 52(6): 2825–2830.

[28] Blackwell, D., L. Breiman, and A. J. Thomasian (1959). "The capacity of a class of channels". *Ann. Math. Statist.* 30(4): 1229–1241.

[29] Blasiak, A., R. Kleinberg, and E. Lubetzky (2011). "Lexicographic products and the power of non-linear network coding". In: *Proc. 52nd Ann. IEEE Symp. Found. Comput. Sci.* Palm Springs, CA. 609–618.

[30] Blasiak, A., R. Kleinberg, and E. Lubetzky (2013). "Broadcasting with side information: Bounding and approximating the broadcast rate". *IEEE Trans. Inf. Theory*. 59(9): 5811–5823.

[31] Blasiak, A. (2013). "A graph-theoretic approach to network coding". *PhD thesis*. Cornell University.

[32] Brahma, S. and C. Fragouli (2015). "Pliable index coding". *IEEE Trans. Inf. Theory*. 61(11): 6192–6203.

[33] Byrne, E. and M. Calderini (2017). "Error correction for index coding with coded side information". *IEEE Trans. Inf. Theory*. 63(6): 3712–3728.

[34] Byrne, E. and M. Calderini (2018). "Index coding, network coding and broadcast with side-information". In: *Network Coding and Subspace Designs*. Ed. by M. Greferath, M. O. Pavčević, N. Silberstein, and M. Á. Vázquez-Castro. Cham: Springer. 247–293.

[35] Cadambe, V. R. and S. A. Jafar (2008). "Interference alignment and degrees of freedom of the K-user interference channel". *IEEE Trans. Inf. Theory*. 54(8): 3425–3441.

[36] Cadambe, V. R., S. A. Jafar, H. Maleki, K. Ramchandran, and C. Suh (2013). "Asymptotic interference alignment for optimal repair of MDS codes in distributed storage". *IEEE Trans. Inf. Theory*. 59(5): 2974–2987.

[37] Cai, N. and T. Chan (2011). "Theory of secure network coding". *Proc. IEEE*. 99(3): 421–437.

[38] Cai, N. and R. W. Yeung (2002). "Secure network coding". In: *Proc. IEEE Int. Symp. Inf. Theory*. Lausanne, Switzerland. 323.

[39] Cai, N. and R. W. Yeung (2011). "Secure network coding on a wiretap network". *IEEE Trans. Inf. Theory*. 57(1): 424–435.

[40] Celebiler, M. and G. Stette (1978). "On increasing the downlink capacity of a regenerative satellite repeater in point-to-point communications". *Proc. IEEE*. 66(1): 98–100.

[41] Chan, T. and A. Grant (2008). "Dualities between entropy functions and network codes". *IEEE Trans. Inf. Theory*. 54(10): 4470–4487.

[42] Chan, T. and A. Grant (2010). “On capacity regions of non-multicast networks”. In: *Proc. IEEE Int. Symp. Inf. Theory*. Austin, TX. 2378–2382.

[43] Chartrand, G. and F. Harary (1967). “Planar permutation graphs”. *Ann. Inst. H. Poincaré B*. 3: 433–438.

[44] Chaudhry, M. A. R., Z. Asad, A. Sprintson, and M. Langberg (2011). “On the complementary index coding problem”. In: *Proc. IEEE Int. Symp. Inf. Theory*. Saint Petersburg, Russia. 244–248.

[45] Chor, B., O. Goldreich, E. Kushilevitz, and M. Sudan (1995). “Private information retrieval”. In: *36th Ann. IEEE Symp. Found. Comp. Sci.* Milwaukee, WI. 41–50.

[46] Chudnovsky, M., N. Robertson, P. Seymour, and R. Thomas (2006). “The strong perfect graph theorem”. *Ann. Math.* 164(1): 51–229.

[47] Chudnovsky, M. and P. Seymour (2005). “The structure of claw-free graphs”. In: *Surv. Combin.* Ed. by B. S. Webb. Cambridge: Cambridge University Press. 153–171.

[48] Connelly, J. M. (2018). “Linear network coding over ring alphabets”. *Ph.D. Thesis*. La Jolla, CA: University of California, San Diego.

[49] Cover, T. M. (1972). “Broadcast channels”. *IEEE Trans. Inf. Theory*. 18(1): 2–14.

[50] Cover, T. M. and J. A. Thomas (2006). *Elements of Information Theory*. Second. New York: Wiley.

[51] Dai, M., K. W. Shum, and C. W. Sung (2014). “Data dissemination with side information and feedback”. *IEEE Trans. Wireless Commun.* 13(9): 1536–1276.

[52] Dau, S. H., V. Skachek, and Y. M. Chee (2012). “On the security of index coding with side information”. *IEEE Trans. Inf. Theory*. 58(6): 3975–3988.

[53] Dau, S. H., V. Skachek, and Y. M. Chee (2013). “Error correction for index coding with side information”. *IEEE Trans. Inf. Theory*. 59(3): 1517–1531.

[54] Dobrushin, R. L. (1959). "Optimum information transmission through a channel with unknown parameters". *Radio Eng. Electron.* 4(12): 1–8.

[55] Dougherty, R., C. Freiling, and K. Zeger (2006). "Unachievability of network coding capacity". *IEEE Trans. Inf. Theory.* 52(6): 2365–2372.

[56] Dougherty, R., C. Freiling, and K. Zeger (2011a). "Network coding and matroid theory". *Proc. IEEE.* 99(3): 388–405.

[57] Dougherty, R., C. Freiling, and K. Zeger (2011b). "Non-Shannon information inequalities in four random variables". URL: http://arxiv.org/abs/1104.3602.

[58] Dougherty, R., C. Freiling, and K. Zeger (2005). "Insufficiency of linear coding in network information flow". *IEEE Trans. Inf. Theory.* 51(8): 2745–2759.

[59] Dueck, G. (1978). "Maximal error capacity regions are smaller than average error capacity regions for multi-user channels". *Probl. Control Inf. Theory.* 7(1): 11–19.

[60] Effros, M., S. El Rouayheb, and M. Langberg (2015). "An equivalence between network coding and index coding". *IEEE Trans. Inf. Theory.* 61(5): 2478–2487.

[61] El Gamal, A. and Y.-H. Kim (2011). *Network Information Theory.* Cambridge: Cambridge University Press.

[62] El Rouayheb, S., A. Sprintson, and C. Georghiades (2010). "On the index coding problem and its relation to network coding and matroid theory". *IEEE Trans. Inf. Theory.* 56(7): 3187–3195.

[63] El Rouayheb, S., M. Chaudhry, and A. Sprintson (2007). "On the minimum number of transmissions in single-hop wireless coding networks". In: *Proc. IEEE Inf. Theory Workshop.* Tahoe City, CA. 120–125.

[64] El Rouayheb, S. Y. (2009). "Network and Index Coding with Application to Robust and Secure Communications". *PhD thesis.* Texas A & M University.

[65] El Rouayheb, S., E. Soljanin, and A. Sprintson (2012). "Secure network coding for wiretap networks of type II". *IEEE Trans. Inf. Theory.* 58(3): 1361–1371.

[66] Erdös, P., Z. Füredi, A. Hajnal, P. Komjáth, V. Rödl, and Á. Seress (1986). "Coloring graphs with locally few colors". *Discrete Math.* 59(1): 21–34.

[67] Erdös, P. and L. Moser (1964). "On the representation of directed graphs as unions of orderings". *Publ. Math. Inst. Hungar. Acad. Sci.* 9: 125–132.

[68] Erdös, P. and G. Szekeres (1935). "A combinatorial problem in geometry". *Compositio Math.* 2: 463–470.

[69] Fekete, M. (1923). "Uber die Verteilung der Wurzeln bei gewissen algebraischen Gleichungen mit ganzzahligen Koeffizienten". *Math. Z.* 17(1): 228–249.

[70] Fragouli, C. and E. Soljanin (Dec. 2007a). "Network coding applications". *Found. Trends Netw.* 2(2): 135–269.

[71] Fragouli, C. and E. Soljanin (Jan. 2007b). "Network coding fundamentals". *Found. Trends Netw.* 2(1): 1–133.

[72] Fujishige, S. (2005). *Submodular Functions and Optimization.* Second. Amsterdam: Elsevier.

[73] Gadouleau, M. and S. Riis (2011). "Graph-theoretical constructions for graph entropy and network coding based communications". *IEEE Trans. Inf. Theory.* 57(10): 6703–6717.

[74] Godsil, C. and B. McKay (1978). "A new graph product and its spectrum". *Bull. Austral. Math. Soc.* 18: 21–28.

[75] Golovnev, A., O. Regev, and O. Weinstein (2018). "The minrank of random graphs". *IEEE Trans. Inf. Theory.* 64.

[76] Haemers, W. (1978). "An upper bound for the Shannon capacity of a graph". *Colloq. Math. Soc. Janos Bolyai.* 25: 267–272.

[77] Haemers, W. (1979). "On some problems of Lovász concerning the Shannon capacity of a graph". *IEEE Trans. Inf. Theory.* 25(2): 231–232.

[78] Hammack, R., W. Imrich, and S. Klavzar (2011). *Handbook of Product Graphs.* Second. Boca Raton, FL: CRC Press.

[79] Haviv, I. and M. Langberg (2012). "On linear index coding for random graphs". In: *Proc. IEEE Int. Symp. Inf. Theory.* Cambridge, MA. 2231–2235.

[80] Ho, T. and D. Lun (2008). *Network Coding: An Introduction.* Cambridge: Cambridge University Press.

[81] Huang, Y.-C. (2017). "Lattice index codes from algebraic number fields". *IEEE Trans. Inf. Theory.* 63(4): 2098–2112.

[82] Huang, X. and S. El Rouayheb (2015). "Index coding and network coding via rank minimization". In: *Proc. IEEE Inf. Theory Workshop.* Jeju Island, Korea. 14–18.

[83] Jafar, S. A. (2014). "Topological interference management through index coding". *IEEE Trans. Inf. Theory.* 60(1): 529–568.

[84] Ji, M., G. Caire, and A. F. Molisch (2016). "Fundamental limits of caching in wireless D2D networks". *IEEE Trans. Inf. Theory.* 62(2): 849–869.

[85] Kao, D. T. H., M. A. Maddah-Ali, and A. S. Avestimehr (2017). "Blind index coding". *IEEE Trans. Inf. Theory.* 63(4): 2076–2097.

[86] Karmoose, M., L. Song, M. Cardone, and C. Fragouli (2017). "Private broadcasting: An index coding approach". In: *Proc. IEEE Int. Symp. Inf. Theory.* Aachen, Germany. 2543–2547.

[87] Kim, J.-W. and J.-S. No (2017). "Index coding with erroneous side information". *IEEE Trans. Inf. Theory.* 63(12): 7687–7697.

[88] Körner, J. and A. Orlitsky (1998). "Zero-error information theory". *IEEE Trans. Inf. Theory.* 44(6): 2207–2229.

[89] Körner, J., C. Pilotto, and G. Simonyi (2005). "Local chromatic number and Sperner capacity". *J. Combin. Theory Ser. B.* 95(1): 101–117.

[90] Kramer, G. (1998). *Directed Information for Channels with Feedback.* Dr. sc. thchn. Dissertation, Swiss Federal Institute of Technology (ETH) Zurich. Konstanz: Hartung-Gorre Verlag.

[91] Kramer, G. and S. Shamai (2007). "Capacity for classes of broadcast channels with receiver side information". In: *Proc. IEEE Inf. Theory Workshop.* Tahoe City, CA. 313–318.

[92] Langberg, M. and M. Effros (2011). "Network coding: Is zero error always possible?" In: *Proc. 49th Ann. Allerton Conf. Comm. Control Comput.* Monticello, IL. 1478–1485.

[93] Langberg, M. and A. Sprintson (2011). "On the hardness of approximating the network coding capacity". *IEEE Trans. Inf. Theory*. 57(2): 1008–1014.

[94] Lee, N., A. G. Dimakis, and R. W. Heath (2015). "Index coding with coded side-information". *IEEE Comm. Lett.* 19(3): 319–322.

[95] Li, M., L. Ong, and S. J. Johnson (2017). "Improved bounds for multi-sender index coding". In: *Proc. IEEE Int. Symp. Inf. Theory*. Aachen, Germany. 3060–3064.

[96] Li, S., M. A. Maddah-Ali, Q. Yu, and A. S. Avestimehr (2018). "A fundamental tradeoff between computation and communication in distributed computing". *IEEE Trans. Inf. Theory*. 64(1): 109–128.

[97] Liu, Y., P. Sadeghi, F. Arbabjolfaei, and Y.-H. Kim (2017). "On the capacity for distributed index coding". In: *Proc. IEEE Int. Symp. Inf. Theory*. Aachen, Germany. 3065–3069.

[98] Liu, Y., P. Sadeghi, F. Arbabjolfaei, and Y.-H. Kim (2018a). "Capacity theorems for distributed index coding". URL: http://arxiv.org/abs/1801.09063.

[99] Liu, Y., P. Sadeghi, F. Arbabjolfaei, and Y.-H. Kim (2018b). "Simplified composite coding for index coding". In: *Proc. IEEE Int. Symp. Inf. Theory*. Vail, CO. 456–460.

[100] Liu, Y., P. Sadeghi, and Y.-H. Kim (2018c). "Three-layer composite coding for index coding". In: *Proc. IEEE Inf. Theory Workshop*. Guangzhou, China.

[101] Liu, Y., B. N. Vellambi, Y.-H. Kim, and P. Sadeghi (2018d). "On the capacity region for secure index coding". In: *Proc. IEEE Inf. Theory Workshop*. Guangzhou, China.

[102] Lovász, L. (1972). "Normal hypergraphs and the perfect graph conjecture". *Discrete Math.* 2(3): 253–267.

[103] Lovász, L. (1979). "On the Shannon capacity of a graph". *IEEE Trans. Inf. Theory*. 25(1): 1–7.

[104] Lubetzky, E. and U. Stav (2009). "Nonlinear index coding outperforming the linear optimum". *IEEE Trans. Inf. Theory*. 55(8): 3544–3551.

[105] Maddah-Ali, M. A., A. S. Motahari, and A. K. Khandani (2008). "Communication over MIMO X channels: Interference alignment, decomposition, and performance analysis". *IEEE Trans. Inf. Theory*. 54(8): 3457–3470.

[106] Maddah-Ali, M. A. and U. Niesen (2014). "Fundamental limits of caching". *IEEE Trans. Inf. Theory*. 60(5): 2856–2867.

[107] Maleki, H., V. R. Cadambe, and S. A. Jafar (2014). "Index coding: An interference alignment perspective". *IEEE Trans. Inf. Theory*. 60(9): 5402–5432.

[108] Marton, K. (1979). "A coding theorem for the discrete memoryless broadcast channel". *IEEE Trans. Inf. Theory*. 25(3): 306–311. ISSN: 0018-9448.

[109] Matthews, M. M. and D. P. Sumner (1985). "Longest paths and cycles in $K_{1,3}$-free graphs". *J. Graph Theory*. 9(2): 269–277.

[110] Matúš, F. (2007). "Infinitely many information inequalities". In: *Proc. IEEE Int. Symp. Inf. Theory*. Nice, France. 41–44.

[111] Mazumdar, A. (2015). "Storage capacity of repairable networks". *IEEE Trans. Inf. Theory*. 61(11): 5810–5821.

[112] Mazumdar, A., A. McGregor, and S. Vorotnikova (2017). "Storage capacity as an information-theoretic analogue of vertex cover". URL: http://arxiv.org/abs/1706.09197.

[113] Mojahedian, M. M., M. R. Aref, and A. Gohari (2017). "Perfectly secure index coding". *IEEE Trans. Inf. Theory*. 63(11): 7382–7395.

[114] Narayanan, V., V. M. Prabhakaran, J. Ravi, V. K. Mishra, B. K. Dey, and N. Karamchandani (2018). "Private index coding". In: *Proc. IEEE Int. Symp. Inf. Theory*. Vail, CO. 596–600.

[115] Natarajan, L., Y. Hong, and E. Viterbo (2015a). "Index codes for the Gaussian broadcast channel using quadrature amplitude modulation". *IEEE Comm. Lett.* 19(8): 1291–1294.

[116] Natarajan, L., Y. Hong, and E. Viterbo (2015b). "Lattice index coding". *IEEE Trans. Inf. Theory*. 61(12): 6505–6525.

[117] Neely, M. J., A. S. Tehrani, and Z. Zhang (2013). "Dynamic index coding for wireless broadcast networks". *IEEE Trans. Inf. Theory*. 59(11): 7525–7540.

[118] Ong, L. (2017). "Optimal finite-length and asymptotic index codes for five or fewer receivers". *IEEE Trans. Inf. Theory*. 63(11): 7116–7130.

[119] Ong, L., C. K. Ho, and F. Lim (2016a). "The single uniprior index coding problem: The single sender case and the multi sender extension". *IEEE Trans. Inf. Theory*. 62(6): 3165–3182.

[120] Ong, L., J. Kliewer, and B. N. Vellambi (2018). "Secure network-index code equivalence: Extension to non-zero error and leakage". In: *Proc. IEEE Int. Symp. Inf. Theory*. Vail, CO. 841–845.

[121] Ong, L., B. N. Vellambi, P. L. Yeoh, J. Kliewer, and J. Yuan (2016b). "Secure index coding: Existence and construction". In: *Proc. IEEE Int. Symp. Inf. Theory*. Barcelona, Spain. 2834–2838.

[122] Ore, O. (1962). *Theory of Graphs*. Providence, RI: American Mathematical Society.

[123] Papailiopoulos, D., A. G. Dimakis, and V. Cadambe (2013). "Repair optimal erasure codes through Hadamard designs". *IEEE Trans. Inf. Theory*. 59(5): 3021–3037.

[124] Peeters, R. (1996). "Orthogonal representations over finite fields and the chromatic number of graphs". *Combinatorica*. 16(3): 417–431.

[125] Riis, S. (2007). "Information flows, graphs and their guessing numbers". *Elec. J. Comb.* 14(R44): 1–17.

[126] Sabidussi, G. (1957). "Graphs with given group and given graph-theoretical properties". *Canad. J. Math.* 9(4): 515–525.

[127] Sadeghi, P., F. Arbabjolfaei, and Y.-H. Kim (2016). "Distributed index coding". In: *Proc. IEEE Inf. Theory Workshop*. Cambridge, UK. 330–334.

[128] Scheinerman, E. R. and D. H. Ullman (2011). *Fractional Graph Theory: A Rational Approach to the Theory of Graphs*. New York: Dover Publications.

[129] Schrijver, A. (2003). *Combinatorial Optimization*. 3 vols. Berlin: Springer-Verlag.

[130] Schwenk, A. J. (1974). "Computing the characteristic polynomial of a graph". In: *Graphs and Combinatorics*. Ed. by R. A. Bari and F. Harary. Berlin: Springer. 153–172.

[131] Shah, N. B., K. V. Rashmi, P. V. Kumar, and K. Ramchandran (2012). "Distributed storage codes with repair-by-transfer and nonachievability of interior points on the storage-bandwidth tradeoff". *IEEE Trans. Inf. Theory*. 58(3): 1837–1852.

[132] Shanmugam, K. and A. G. Dimakis (2014). "Bounding multiple unicasts through index coding and locally repairable codes". In: *Proc. IEEE Int. Symp. Inf. Theory*. Honolulu, HI. 296–300.

[133] Shanmugam, K., A. G. Dimakis, and M. Langberg (2013). "Local graph coloring and index coding". In: *Proc. IEEE Int. Symp. Inf. Theory*. Istanbul, Turkey. 1152–1156.

[134] Shanmugam, K., A. G. Dimakis, and M. Langberg (2014). "Graph theory versus minimum rank for index coding". In: *Proc. IEEE Int. Symp. Inf. Theory*. Honolulu, HI. 291–295.

[135] Shanmugam, K., M. Asteris, and A. G. Dimakis (2015). "On approximating the sum-rate for multiple-unicasts". In: *Proc. IEEE Int. Symp. Inf. Theory*. Hong Kong. 381–385.

[136] Shannon, C. E. (1949). "Communication theory of secrecy systems". *Bell Syst. Tech. J.* 28(4): 656–715. ISSN: 0005-8580.

[137] Shannon, C. E. (1948). "A mathematical theory of communication". *Bell Syst. Tech. J.* 27(3): 379–423, 27(4), 623–656.

[138] Shannon, C. E. (1956). "The zero error capacity of a noisy channel". *IRE Trans. Inf. Theory*. 2(3): 8–19.

[139] Shannon, C. E. (1959). "Coding theorems for a discrete source with a fidelity criterion". In: *IRE Int. Conv. Rec.* Vol. 7, part 4. 142–163.

[140] Silva, D. and F. R. Kschischang (2011). "Universal secure network coding via rank-metric codes". *IEEE Trans. Inf. Theory*. 57(2): 1124–1135.

[141] Sima, J. and W. Chen (2014). "Joint network and Gelfand–Pinsker coding for 3-receiver Gaussian broadcast channels with receiver message side information". In: *Proc. IEEE Int. Symp. Inf. Theory*. Honolulu, HI. 81–85.

[142] Song, L. and C. Fragouli (2018). "A polynomial-time algorithm for pliable index coding". *IEEE Trans. Inf. Theory*. 64(2): 979–999.

[143] Stearns, R. (1959). "The voting problem". *Amer. Math. Monthly*. 66(9): 761–763.

[144] Steinberg, R. and C. A. Tovey (1993). "Planar Ramsey numbers". *J. Combin. Theory Ser. B*. 59(2): 288–296.

[145] Sun, H. and S. A. Jafar (2015). "Index coding capacity: How far can one go with only Shannon inequalities?" *IEEE Trans. Inf. Theory*. 61(6): 3041–3055.

[146] Sun, H. and S. A. Jafar (2017). "The capacity of private information retrieval". *IEEE Trans. Inf. Theory*. 63(7): 4075–4088.

[147] Tahmasbi, M., A. Shahrasbi, and A. Gohari (2015). "Critical graphs in index coding". *IEEE J. Sel. Areas Commun.* 33(2): 225–235.

[148] Thakor, S., A. Grant, and T. Chan (2016). "Cut-set bounds on network information flow". *IEEE Trans. Inf. Theory*. 62(4): 1850–1865.

[149] Thapa, C., L. Ong, and S. J. Johnson (2016). "Graph-theoretic approaches to two-sender index coding". In: *Proc. IEEE Global Telecommun. Conf.* Washington, DC. 1–6.

[150] Thapa, C., L. Ong, and S. J. Johnson (2017). "Interlinked cycles for index coding: Generalizing cycles and cliques". *IEEE Trans. Inf. Theory*. 63(6): 3692–3711.

[151] "The On-Line Encyclopedia of Integer Sequences" (n.d.). URL: http://oeis.org/A000273.

[152] Turán, P. (1941). "Eine extremalaufgabe aus der graphentheorie". *Mat. Fiz. Lapok*. 48: 436–452.

[153] Unal, S. and A. B. Wagner (2016). "A rate–distortion approach to index coding". *IEEE Trans. Inf. Theory*. 62(11): 6359–6378.

[154] Urbanke, R. and Q. Li (1998). "The zero-error capacity region of the 2-user synchronous BAC is strictly smaller than its Shannon capacity region". In: *Proc. IEEE Inf. Theory Workshop*. Killarney, Ireland. 61.

[155] West, D. B. (2001). *Introduction to Graph Theory*. Second. Upper Saddle River, NJ: Prentice Hall.

[156] Willems, F. M. J. (1990). "The maximal-error and average-error capacity region of the broadcast channel are identical: A direct proof". *Probl. Control Inf. Theory*. 19(4): 339–347.

[157] Willems, F. M. J., J. K. Wolf, and A. D. Wyner (1989). "Communicating via a processing broadcast satellite". In: *Proc. IEEE/CAM Inf. Theory Workshop*. Cornell, NY. 3-1.

[158] Witsenhausen, H. S. (1976). "The zero-error side information problem and chromatic numbers". *IEEE Trans. Inf. Theory*. 22(5): 592–593.

[159] Wolfowitz, J. (1960). "Simultaneous channels". *Arch. Rational Mech. Anal.* 4(1): 371–386.

[160] Wu, Y. (2007). "Broadcasting when receivers know some messages a priori". In: *Proc. IEEE Int. Symp. Inf. Theory*. Nice, France. 1141–1145.

[161] Wyner, A. D., J. K. Wolf, and F. M. J. Willems (2002). "Communicating via a processing broadcast satellite". *IEEE Trans. Inf. Theory*. 48(6): 1243–1249.

[162] Xue, F. and S. Sandhu (2007). "PHY-layer network coding for broadcast channel with side information". In: *Proc. IEEE Inf. Theory Workshop*. Lake Tahoe, CA. 108–113.

[163] Yeung, R. W. (1995). "Multilevel diversity coding with distortion". *IEEE Trans. Inf. Theory*. 41(2): 412–422.

[164] Yeung, R. W. (2008). *Information Theory and Network Coding*. New York: Springer.

[165] Yeung, R. W. and Z. Zhang (1999). "Distributed source coding for satellite communications". *IEEE Trans. Inf. Theory*. 45(4): 1111–1120.

[166] Yi, X. and G. Caire (2018). Private communication.

[167] Yi, X., H. Sun, S. A. Jafar, and D. Gesbert (2018). "TDMA is optimal for all-unicast DoF region of TIM if and only if topology is chordal bipartite". *IEEE Trans. Inf. Theory*. 64(3): 2065–2076.

[168] Yoo, J. W., T. Liu, and F. Xue (2009). "Gaussian broadcast channels with receiver message side information". In: *Proc. IEEE Int. Symp. Inf. Theory*. Seoul, Korea. 2472–2476.

[169] Yu, H. and M. J. Neely (2014). "Duality codes and the integrality gap bound for index coding". *IEEE Trans. Inf. Theory*. 60(11): 7256–7268.

[170] Zhang, Z. and R. W. Yeung (1998). "On characterization of entropy function via information inequalities". *IEEE Trans. Inf. Theory*. 44(4): 1440–1452.