Achievable Schemes and Performance Bounds for Centralized and Distributed Index Coding

Yucheng Liu

January 2021

A Thesis Submitted for the Degree of
Doctor of Philosophy
of The Australian National University

Research School of Electrical,
Energy and Materials Engineering
The Australian National University

Except where otherwise indicated, this thesis is my own original work.

Yucheng Liu
25 January 2021

To my parents, Mei and Chunjie, and my beloved fiancée, Nan.

Acknowledgments

I would like to first express my heartfelt and profound gratitude to my primary supervisor, Prof. Parastoo Sadeghi, for her constant and invaluable guidance, advice, motivation, and inspiration throughout my entire PhD program. She is the one who introduces the beauty and elegance of information theory to me. Parastoo teaches me with great academic enthusiasm and thoroughness how to understand abstract concepts, how to formulate and analyze problems, and how to write and present my findings and results. I particularly appreciate our productive discussions during which we tackled difficult problems together and did mathematical derivations with pens and papers. I am also full of gratitude for her understanding, patience, and encouragement during several tough times, when it is hard for me to concentrate and make progress due to many reasons. Clearly, I would not be able to complete this thesis without her continuous support. I would also like to thank Prof. Rodney A. Kennedy and Dr. Nan Yang from my supervisory panel for their advice and assistance during the last few years.

I would like to sincerely thank Prof. Young-Han Kim and Dr. Fatemeh Arbabjolfaei from University of California, San Diego as I have learnt a lot from the close and fruitful collaboration between our groups. Young-Han's insightful ideas and broad knowledge in information theory have been a source of inspiration throughout my PhD program.

I would like to thank Dr. Neda Aboutorab from the University of New South Wales, Canberra for kindly giving me the chance to work in her group as a casual research assistant during my last year of PhD. Her guidance and advice during our collaborative work have led me to many interesting results. I would also like to thank Dr. Ni Ding and Dr. Thierry Rakotoarivelo from Data61, Commonwealth Scientific and Industrial Research Organisation (CSIRO) for their kind invitation and mentorship during my fruit-bearing internship in their group.

I am grateful to my colleagues at ANU for creating a friendly and creative research environment, among whom include Prof. Thushara Abhayapala, Assoc. Prof. Xiangyun (Sean) Zhou, Assoc. Prof. Salman Durrani, Dr. Prasanga Samarasinghe, Dr. Wen Zhang, Dr. Hanchi Chen, Dr. Jihui Zhang, Dr. Abdullah Fahim, Dr. Mohammad S. Karim, Dr. Mingchao Yu, Dr. Shihao Yan, Dr. Yirui Cong, Dr. Jing Guo, Dr. Norman Akbar, Dr. Wanchun Liu, Dr. Nicole Sawyer, Dr. Abbas Koohian, Dr. Yuting Fang, Dr. Katrina Zhou, Huanyu Zuo, Fei ma, Zhifeng Tang, Chunhui Li, Akram Shafie, Yizhou Yang, and Yuxin Liu. I am also thankful to people I met during my PhD program outside ANU, especially Assoc. Prof. Lawrence Ong, Prof. Sarah Johnson, Prof. Joerg Kliewer, Dr. Phee Lep Yeoh, Dr. Min Li, Dr. Chandra Thapa, Dr. Xiangxiang Xu, Kameliya Kaneva, Arman Sharififar, Bryan Liu, Jin Yeong Tan, and Emelie

Ekenstedt.

Last, and most importantly, I would like to thank my family for their unconditional support and encouragement through all ups and downs. The endless love from my parents have been accompanying me ever since I was born, which was the most fundamental reason behind every success I have achieved. I would also like to thank my grandparents for loving me and taking care of me all the time. I am highly indebted to my beloved fiancée, Nan Wu, for her understanding, support, and even sacrifice throughout my whole PhD journey. I am also grateful for all the wonderful moments we spent together, which lighted up my world. Special thanks go to my lovely fish and cats (especially Zhuzhu, Xiaoma, and Zaza) for making my life more joyful and full of fun.

Abstract

Index coding studies the efficient broadcast problem where a server broadcasts multiple messages to a group of receivers with side information. Through exploiting the receiver side information, the amount of required communication from the server can be significantly reduced. Thanks to its basic yet highly nontrivial model, index coding has been recognized as a canonical problem in network information theory, which is fundamentally connected with many other problems such as network coding, distributed storage, coded computation, and coded caching.

In this thesis, we study the index coding problem both in its classic setting where the messages are stored at a centralized server, and also in a more general and practical setting where different subsets of messages are stored at multiple servers. In both scenarios the ultimate goal is to establish the capacity region, which contains all the communication rates simultaneously achievable for all the messages. While finding the index coding capacity region remains open in general, we characterize it through developing various inner and outer bounds. The inner bounds we propose on the capacity region are achievable rate regions, each associated with a concrete coding scheme. Our proposed coding schemes are built upon a two-layer random coding scheme referred to as composite coding, introduced by Arbabjolfaei et al. in 2013 for the classic centralized index coding problem. We first propose a series of simplifications for the composite coding scheme, and then enhance it through utilizing more flexible fractional allocation of the broadcast channel capacity. We also show that one can strictly improve composite coding by adding one more layer of random coding into the coding scheme. For the multi-server scenario, we generalize composite coding to a distributed version.

The outer bounds characterize the fundamental performance limits enforced by the problem setup that hold generally for any valid coding scheme. The performance bounds we propose are based on Shannon-type inequalities. For the centralized index coding problem, we define a series of interfering message structures based on the receiver side information. Such structures lead to nontrivial generalizations of the alignment chain model in the literature, based upon which we propose a series of novel iterative performance bounds. For the multi-server scenario, our main result is a general outer bound built upon the polymatroidal axioms of the entropy function. This outer bound utilizes general groupings of servers of different levels of granularity, allowing a natural tradeoff between tightness and computational complexity.

The security aspect of the index coding problem is also studied, for which a number of achievability and performance bounds on the optimal secure communication rate are established. To conclude this thesis, we investigate a privacy-preserving data publishing problem, whose model is inspired by index coding, and characterize its optimal privacy-utility tradeoff.

List of Publications

Almost all work in this thesis has been published or is to be submitted for publication as journal articles or conference papers. These papers are:

J1. **Y. Liu**, P. Sadeghi, F. Arbabjolfaei, and Y.-H. Kim, "Capacity theorems for distributed index coding," *IEEE Trans. Inf. Theory*, vol. 66, no. 8, pp. 4653–4680, Mar. 2020.

J2. **Y. Liu** and P. Sadeghi "Performance bounds for index coding based on acyclic chains," to be submitted to *IEEE Trans. Inf. Theory*.

C1. **Y. Liu**, P. Sadeghi, F. Arbabjolfaei, and Y.-H. Kim, "On the capacity for distributed index coding," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Aachen, Germany, June 2017, pp. 3055–3059.

C2. **Y. Liu**, P. Sadeghi, F. Arbabjolfaei, and Y.-H. Kim, "Simplified composite coding for index coding," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Vail, CO, June 2018, pp. 456–460.

C3. **Y. Liu**, P. Sadeghi, and Y.-H. Kim, "Three-layer composite coding for index coding," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Guangzhou, China, Nov. 2018, pp. 170–174.

C4. **Y. Liu** and P. Sadeghi, "Generalized alignment chain: improved converse results for index coding," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Paris, France, Jul. 2019, pp. 1242–1246.

C5. **Y. Liu** and P. Sadeghi, "From alignment to acyclic chains: lexicographic performance bounds for index coding," in *Proc. 57th Ann. Allerton Conf. Comm. Control Comput.*, Monticello, IL, Sep. 2019, pp. 260–267.

C6. **Y. Liu**, N. Ding, P. Sadeghi, and T. Rakotoarivelo, "Privacy-utility tradeoff in a guessing framework inspired by index coding," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Los Angeles, CA, June 2020.[1]

C7. **Y. Liu**, P. Sadeghi, N. Aboutorab, and A. Sharififar, "Secure index coding with security constraints on receivers," in *Proc. Int. Symp. on Inf. Theory and its Applications (ISITA)*, Kapolei, HI, Oct. 2020.

[1] This work was produced during my internship at Data61, Commonwealth Scientific and Industrial Research Organisation (CSIRO).

The following publication was also produced during my PhD program but is not directly related to nor included in this thesis:

C8. A. Sharififar, N. Aboutorab, **Y. Liu**, and P. Sadeghi, “Independent user partition multicast scheme for the groupcast index coding problem,” in *Proc. Int. Symp. on Inf. Theory and its Applications (ISITA)*, Kapolei, HI, Oct. 2020.

Acronyms

CIC	centralized index coding
DIC	distributed index coding
NP	nondeterministic polynomial time
MAIS	maximum acyclic induced subgraph
PM	polymatroidal
AWGN	additive white Gaussian noise
CC	composite coding
ECC	enhanced composite coding
TLCC	three-layer composite coding
DCC	distributed composite coding
GPM	grouping polymatroidal
MDS	maximum distance separable
FLPCC	fractional local partial clique covering
FME	Fourier-Motzkin elimination
LP	linear programming
CCC	cooperative composite coding
S-FLPCC	secure fractional local partial clique covering
S-PM	secure polymatroidal
S-MAIS	secure maximum acyclic induced subgraph

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

Introduced by Birk and Kol in Birk and Kol [1998, 2006] for satellite communication, the index coding problem encapsulates the communication problem where a server broadcasts a number of messages to multiple receivers. Each receiver requests some messages and knows some other messages *a priori* as side information. The server knows which messages are requested and which messages are known by each receiver. Through exploiting the side information availability at the receivers, the server attempts to minimize the overall amount of information to be broadcast through the channel to satisfy each receiver's requests.

Consider a simple motivating toy example as follows. Assume there is one server containing three same-length binary messages, x_1, x_2, x_3. There are three receivers, receiver $1, 2, 3$. Each receiver requests one unique message. In particular, receiver i wants message x_i for any $i \in \{1, 2, 3\}$. Each receiver knows some other messages apart from its own requested message as side information. More specifically, receiver 1 knows nothing, receiver 2 knows message x_3, and receiver 3 knows message x_2. Assume that the broadcast channel is noiseless, and that a codeword of length equal to one message can be transmitted through the channel within each time slot. Such example can be visualized in Figure 1.1.

To satisfy the three receivers, the server could simply send the uncoded messages x_1, x_2, x_3 one-by-one in total 3 time slots. Nevertheless, such naive transmission scheme is suboptimal as the receivers can in fact be satisfied in merely 2 time slots. To achieve this, the server could transmit first the uncoded message x_1 in the first time slot and then the coded symbol $x_2 \oplus x_3$ in the second time slot, where $\oplus$ means bitwise XORing of message bits. Upon receiving such transmitted codewords $(x_1, x_2 \oplus x_3)$, one can easily verify that every receiver can decode its own requested message: receiver 1 can trivially recover x_1; receiver 2 can decode x_2 utilizing the received coded symbol $x_2 \oplus x_3$ and its own side information x_3 as $x_2 = (x_2 \oplus x_3) \oplus x_3$; receiver 3 can decode x_3 in a similar manner as receiver 2 since it knows x_2 as side information and $x_3 = (x_2 \oplus x_3) \oplus x_2$. This example illustrates the key fact that sending coded symbols can result in reduced number of transmissions to fulfill the receivers' requests and thus achieve

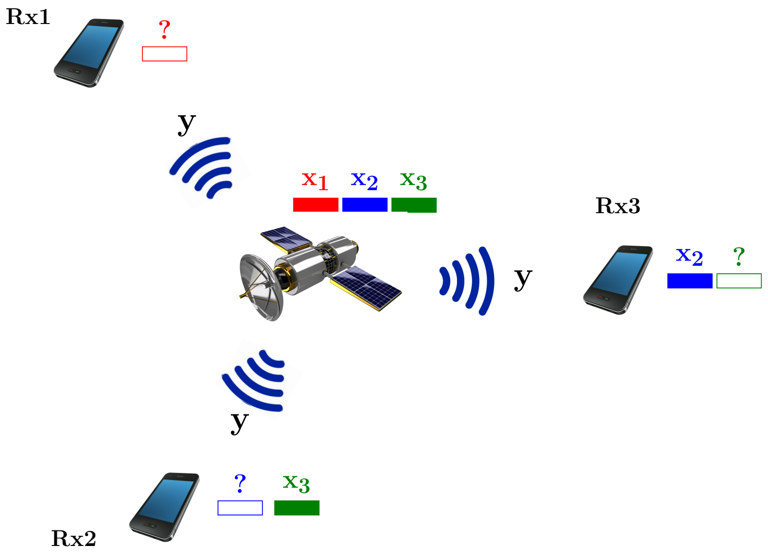


Figure 1.1: An index coding example with 3 messages and 3 receivers. Each receiver wants to know one unique message and may know some other messages as side information.

a better communication efficiency. Note that such savings of the number of transmissions will not be possible if the receivers do not possess any side information.

At first glance, the system model of index coding may seem rather simple as there is only one shared broadcast link between the server and the receivers. However, quite surprisingly, it is indeed rich enough to capture any multiple-unicast communication network. This is because it allows arbitrary side information among the receivers. Such basic yet nontrivial model not only leads to clear formulation of transmission techniques, but also results in strong connections between index coding and many other important problems such as coded caching, distributed computing, distributed storage, topological interference management, as well as the long-lasting network coding problem. Consequently, the index coding problem has been recognized as one of the canonical problems in network information theory.

As the index coding problem captures the essence of the cache-aided network communication, it has various real-world applications in diverse areas such as satellite communication, opportunistic wireless communication, and multimedia distribution. In the classic setting of the index coding problem there is a central server which contains all the messages. However, an even more practical and general setup would allow the messages to be distributed among multiple servers, where each server stores a unique subset of messages and is connected to all the receivers through its own broadcast channel. Such communication model has clear applications for practical scenarios, in which the information is geographically distributed and stored across many locations. We refer to such multi-server index coding problem as the *distributed* index coding (DIC) problem, whereas the classic index coding problem with one central server is referred to as the *centralized* index coding (CIC) problem. Throughout the thesis, unless otherwise specified, we study the CIC and DIC problem while assuming a *multiple-unicast* setting where each message is requested by exactly one receiver.

There are two fundamental objectives in studying the index coding problem, whether centralized or distributed. One is to establish the capacity region, which contains all the communication rates simultaneously achievable for all the messages. The other is to design an optimal coding scheme that achieves the capacity region. In particular, we are typically interested in finding a computable single-letter expression of the capacity region, and we wish to develop an optimal coding scheme that is implementable in practice.

1.2 Literature Review

Since its introduction by Birk and Kol Birk and Kol [1998] in 1998, the centralized index coding (CIC) problem has intrigued various research communities and has been extensively investigated from different perspectives such as algebraic coding theory, graph theory, network coding, Shannon theory, and interference alignment.

While the fundamental objectives of establishing the capacity region and designing optimal coding schemes remain open in general, various achievable index coding schemes and performance bounds for the CIC problem have been established. Each achievable scheme provides an achievable communication rate region serving as an inner bound on the capacity region. And each performance bound serves as a converse result that characterizes the fundamental performance limits enforced by the problem setup holding generally for any valid index coding scheme. In contrast to the CIC problem, the distributed index coding problem (DIC) was recently introduced and thus less literature exists. In the following we first review existing works on coding scheme design for both the CIC and DIC problems. Then we discuss relevant works on performance bounds for both problems. To conclude this section, we also provide a brief overview of some interesting results on the CIC and DIC problems or their variants other than the achievable schemes and performance bounds. In particular, we discuss the previous works in the literature with regard to the *security* aspects of the index coding problem. We also elucidate the relationship between the index coding problem and other related problems such as network coding and locally recoverable distributed storage.

1.2.1 Existing Achievable Coding Schemes

All the index coding schemes can be broadly categorized into two classes: *linear* and *non-linear* coding schemes. Any coding scheme is a linear coding scheme if the encoding and decoding functions are all linear. Otherwise the coding scheme is non-linear. The linear coding schemes are also called the *vector* linear coding schemes, which include a special class of coding schemes referred to as the *scalar* linear coding schemes. In the scalar linear coding schemes, messages are encoded while being treated as an entity, while in the more general vector linear coding schemes, each message is further divided into a number of smaller size sub-messages through rate splitting. It has been shown in Bar-Yossef et al. [2011] that for the CIC problem the optimal binary scalar linear codeword length is equal to a certain algebraic notion called *minrank*, which can be found via rank minimization over a group of *fitting* matrices constructed based on the receiver side information. Such result was generalized to allow coding to be performed over an arbitrary finite field in Lubetzky and Stav [2009], characterizing the optimal performance of scalar linear index code for the CIC problem. As shown in El Rouayheb et al. [2007], computing the minrank even within the binary field is NP-complete. Nevertheless, the minrank method has been further extended into a general vector linear version in Maleki et al. [2014]. Based on the result from Maleki et al. [2014], a multi-letter characterization of the optimal performance of linear index code for the CIC problem has been established in [Arbabjolfaei and Kim, 2018, Section 6.7], whose computability remains unknown in general.

Another line of work which can be dated back to the original paper by Birk and Kol Birk and Kol [1998] applies graph-theoretic tools to the designing of CIC coding schemes. In Birk and Kol [1998], the authors introduced the idea of characterizing the receiver side information through a directed graph, which is later formally named as the *side information graph* in Bar-Yossef et al. [2011]. Most of the graph-theoretic coding schemes involve decomposition of the side information graph based on certain graph-theoretic structures (e.g., cliques or cycles). See Birk and Kol [2006]; Chaudhry et al. [2011]; Neely et al. [2013]; Blasiak et al. [2013]; Shanmugam et al. [2013]; Arbabjolfaei and Kim [2014]; Thapa et al. [2017]; Agarwal and Mazumdar [2016] and references therein. Although such graph-theoretic schemes in general lead to smaller acheivable rate regions (i.e., looser inner bounds on the capacity region) compared to the algebraic minrank approaches, they have the advantage of being more practical because of their lower complexities and simpler encoding and decoding processes.

The *interference alignment* method, introduced from the study of the interference management for wireless interference networks, has been adapted in Maleki et al. [2014]; Jafar [2014] to develop scalar and vector linear coding schemes for the CIC problem. In particular, the optimal *symmetric* communication rate has been established for several special classes of the CIC problem in Maleki et al. [2014], including the scenario where the symmetric rate of $1/2$ is simultaneously achievable for all the messages and the scenario where the side information availability at the receivers follows a certain symmetric structure. Notice that the sufficient and necessary condition for the symmetric rate of $1/2$ being achievable was also independently established in Blasiak et al. [2013], where a polynomial time algorithm to determine whether such condition is satisfied was also proposed.

In general linear index coding schemes are not sufficient to achieve the capacity region for the CIC problem as shown in Lubetzky and Stav [2009]; Blasiak et al. [2011]; Maleki et al. [2014]. Existing non-linear index coding schemes are designed mainly from two perspectives. One is to utilize the classic information-theoretic tool: *random coding*. In Arbabjolfaei et al. [2013], the authors first introduced the *flat* coding scheme and then the more general *composite coding* scheme. The former is a single-layer random coding scheme, while the latter is a superpositioned coding scheme that consists of two layers of random coding. It has been verified in Arbabjolfaei et al. [2013] that the composite coding scheme is able to establish the capacity region for all the CIC problems with five or fewer messages. Despite the general good performance of the composite coding scheme, it is still suboptimal as demonstrated in [Arbabjolfaei et al., 2014, Example 4], [Thapa et al., 2017, Section III].

The other major perspective for designing non-linear CIC coding schemes is based upon *coloring* of the *confusion graphs* for the CIC problem, the notion of which was introduced in Alon et al. [2008]. Roughly speaking, the confusion graph captures the *distinguishability* of any two realizations of the messages at the receivers. In other words, the confusion graph indicates the groups of message realizations that *must not* be mapped to the same codeword to satisfy the decoding requirement at the receivers. Any valid coloring scheme of the confusion graph corresponds to a valid possibly non-linear CIC coding scheme. Indeed, based on the (fractional) chromatic number of the confusion graph a multi-letter characterization of the capacity region for the general CIC problem has been established Arbabjolfaei and Kim [2018], the computability of which is unknown in general.

Thus far we have listed and discussed a number of existing achievable coding schemes for the CIC problem that are developed from several most common directions. More details can be found in Arbabjolfaei and Kim [2018] and references therein. In the following we provide a brief overview of previous works on the coding scheme design for the DIC problem, most of which are based upon extending the existing CIC coding schemes.

The DIC problem with the messages being distributed among multiple servers was first introduced in Ong et al. [2016a] in the special form where each receiver knows only one message a priori as side information. Subsequently, the authors in Thapa et al. [2016] considered another interesting special case with two servers each containing an arbitrary message subset. In both works linear coding schemes based on graph-theoretic tools have been studied.

The most general setup of the DIC problem where for every subset of messages there is one corresponding server with an arbitrary finite broadcast channel link capacity was first introduced and investigated in Sadeghi et al. [2016]. In particular, a distributed version of the composite coding scheme was developed in Sadeghi et al. [2016], achieving the capacity region for all the DIC problems with three or fewer messages and equal broadcast link capacities for all the servers. Improvements upon such DIC composite coding scheme have been proposed in Li et al. [2017, 2018] under a similar system model.

The minrank approach have also been extended to the general DIC problem in Li et al.

[2019]; Kim and No [2019]. In particular, the authors of Li et al. [2019] developed a rank-minimization framework that explicitly takes into account the message availability at the servers and then checks whether the decoding requirements of the receivers are satisfied for each coding matrix candidate. To reduce the search space dimension, a number of optimality-preserving simplification strategies have also been proposed in Li et al. [2019]. In contrast, the authors of Kim and No [2019] proposed an alternative way of constructing the fitting matrices for the DIC problem for which the decodability of the messages at their corresponding receivers are always guaranteed.

1.2.2 Existing Performance Bounds

While each achievability result is induced by a specific coding scheme, performance bounds for index coding provide converse results, characterizing the fundamental limits that any valid index coding scheme cannot exceed. Throughout the thesis, whenever we say performance bounds, we mean converse results. In the context of capacity region, this should be understood as outer bounds (upper bounds) on capacity region (capacity values). For broadcast rate, which is the reciprocal of the symmetric capacity, this should be understood as lower bounds on broadcast rate.

Shannon-type inequalities, also known as the polymatroidal axioms of the entropy function, have been playing a central role in developing performance bounds in network information theory. In particular, an explicit characterization of the minimal set of Shannon-type inequalities for an arbitrary set of random variables was given in Yeung [2008]. Consequently, most of the existing performance bounds for both the CIC and DIC problems are based purely on Shannon-type inequalities. Note that for some specific CIC problem instances, non-Shannon-type inequalities are strictly needed to tightly characterize the capacity region as shown in Sun and Jafar [2015]; Baber et al. [2013]. As the CIC problem can be seen as a special case of the DIC problem, it follows automatically that non-Shannon-type inequalities are also needed for some DIC problem instances.

Within the scope of the CIC performance bounds based on Shannon-type inequalities, one can form the tightest bound through explicitly including *all* the Shannon-type inequalities for the message and codeword random variables according to the minimal set presented in Yeung [2008]. Such bound has been presented in [Arbabjolfaei and Kim, 2018, Section 5.3]. The polymatroidal (PM) bound from Blasiak et al. [2011] is established using only a subset of Shannon-type inequalities. The PM bound is tight on the capacity region for all the CIC problem instances with five or fewer messages (as verified in Arbabjolfaei et al. [2013]), as well as many larger instances with more than five messages. Notably, the PM bound has been proved in Liu et al. [2018a] to be as tight as the aforementioned tightest bound which explicitly includes all Shannon-type inequalities. One common drawback of these two bounds is the high computational complexity, which grows exponentially as the number of the messages in the CIC problem increases.

Despite possible performance loss in terms of tightness, several less computationally intensive performance bounds for the CIC problem have been proposed and proved to be useful in the literature. The arguably most widely-used performance bound is the maximal acyclic induced subgraph (MAIS) bound proposed in Bar-Yossef et al. [2011], which utilizes the *acyclic* structure within the side information graph. Another interesting bound is the internal conflict bound presented in Maleki et al. [2014]; Jafar [2014], which is based on the alignment chain model introduced in Maleki et al. [2014] from the interference alignment perspective. The MAIS bound and the internal conflict can outperform each other for different CIC problem instances, while both of them are implied by the PM bound.

The MAIS bound and the PM bound have been generalized to the DIC problem in Sadeghi et al. [2016]. Noticeably, the bound in Sadeghi et al. [2016] based on the polymatroidal axioms can only effectively capture a subset of Shannon-type inequalities for the DIC problem, and thus is strictly looser than the tightest possible DIC performance bound based on all Shannon-type inequalities. The distributed version of the MAIS bound in Sadeghi et al. [2016] has also been presented in an alternative form in Li et al. [2018].

1.2.3 Other Related Results

Rather than directly characterizing the capacity region for the CIC problem via establishing tight inner and outer bounds, an alternative "divide-and-conquer" approach has been taken in Arbabjolfaei and Kim [2020], which studies the structural properties of the receiver side information availability to decide whether a CIC problem can be decomposed into subproblems such that its capacity region is characterized in terms of those of the subproblems. A different but related question is that whether a CIC problem's capacity region remains unchanged if some side information at a certain receiver is removed. Partial answers to this question have been proposed in Tahmasbi et al. [2015]; Arbabjolfaei and Kim [2015a]. Similar techniques have been investigated for the DIC problem in Thapa et al. [2019]; Arunachala et al. [2019]; Kim and No [2019].

As mentioned earlier the CIC problem has strong connections to many other important research problems in network information theory. In particular, while the CIC problem can be seen as a special case of the multiple-unicast network coding problem Ahlswede et al. [2000]; Yeung [2006, 2008], a code-level equivalence has been established in Effros et al. [2015] between these two problems, providing a mapping from any network coding instance to a corresponding CIC problem and vice versa such that a valid coding scheme for one problem can be translated to a valid (but different) coding scheme for the other problem. A rather interesting observation is that the DIC problem itself is also a special case of the multiple-unicast network coding problem, and thus the code-level equivalence established in Effros et al. [2015] also holds between the CIC and DIC problems. In addition, there exists a duality between the CIC problem and the locally recoverable distributed storage problem, which has been identified in Mazumdar [2014]; Shanmugam and Dimakis [2014] and further investigated in Arbabjolfaei and Kim [2015b]. The authors of Arbabjolfaei and Kim [2015b] also studied the complementarity between the guessing game problem on graphs Riis [2007] and the CIC problem.

Another research direction is to *approximate* the capacity for the index coding problem rather than trying to directly establish it. Such approximation problem has been studied in Blasiak et al. [2013]; Arbabjolfaei and Kim [2016] for the CIC scenario. To the best of our knowledge, there exists no work considering capacity approximation for the DIC problem.

Variants of the CIC and DIC problems, each of different theoretical or practical significance, have been formulated and studied in the literature. We review several interesting directions in the following.

The CIC problem with a *noisy* broadcast channel (e.g., AWGN channel) has been investigated in Dau et al. [2013]; Asadi et al. [2015]; Natarajan et al. [2015a,b]; Huang [2017]; Karat et al. [2018]. Such model has clear value for some practical scenarios where the assumption of a noiseless channel is unrealistic.

Another practical variant of the CIC problem called *pliable* index coding was introduced in Brahma and Fragouli [2015] and subsequently studied in Song and Fragouli [2017]; Song et al. [2019]; Liu and Tuninetti [2019c]; Ong et al. [2019a,b]; Sasi and Rajan [2019], where the receivers are "pliable" such that each receiver will be satisfied as long as it can decode *any* one of its unknown messages from the transmitted codeword.

Information security and privacy have been receiving growing attention in both communication and data science. The security aspect of the CIC problem was first studied in Dau et al. [2012], where in addition to the *legitimate* receivers there is an *eavesdropper* with some side information, and the server must broadcast in a way such that the messages are kept secure against the eavesdropper while at the same time each receiver can decode its requested message. Several extensions of such problem, often collectively termed as *secure* index coding, have been studied in Ong et al. [2016c]; Mojahedian et al. [2017]; Ong et al. [2016b, 2018]; Liu et al. [2018a]. An alternative problem setup where there is no eavesdropper and the security constraints are against the receivers themselves has been briefly discussed in Dau et al. [2012] and later investigated in Narayanan et al. [2020]. The authors of Karmoose et al. [2019] considered the CIC problem from a different privacy-preserving perspective, where instead of trying to protect the *content* of the messages, their goal was to limit the information that a receiver can infer about the *identities* of the requests of other receivers.

A *peer-to-peer* version of the DIC problem has been studied in Porter and Wootters [2019] under the name *embedded* index coding, where each party in the communication system acts as both a server and a receiver. That is, each party broadcasts a codeword as a function of its own side information to other parties, and at the same time, needs to decode some of its unknown messages from the other receivers' codewords.

The key features of some CIC and DIC variants can be combined together. See, for example, Liu and Tuninetti [2020], where a peer-to-peer, secure, and pliable version of the index coding problem was investigated; see also Liu and Tuninetti [2019a], which studies peer-to-peer pliable index coding, and Liu and Tuninetti [2019b], which studies secure pliable index coding. See, for another example, Byrne and Calderini [2017], where the broadcast channel is assumed to be noisy and at the same time, the side information at receivers can be linear combinations of the messages.

The majority of the previous works introduced thus far considered the CIC and DIC problems or their variants in the multiple-unicast setting. Indeed, most of their results have natural extensions to the more general *multiple-groupcast* scenario, where some message can be demanded by two or more receivers. See also Alon et al. [2008]; Blasiak et al. [2011]; Tehrani et al. [2012]; Neely et al. [2013]; Shanmugam et al. [2014]; Unal and Wagner [2016] for existing studies on such model.

For other existing works on the CIC and DIC problems or their variants, see Dai et al. [2014]; Lee et al. [2015]; Kim and No [2017]; Thomas and Rajan [2017]; Esfahanizadeh et al. [2014]; Wan et al. [2017]; Yu and Neely [2014]; Kao et al. [2016]; El Rouayheb et al. [2010]; Haviv and Langberg [2012]; Arbabjolfaei and Kim [2018]; Hsu et al. [2020]; Huang et al. [2017]; Haviv [2020] and references therein.

1.3 Thesis Contributions, Organization, and Notation

In this thesis our contributions are mainly in three domains. First, in Chapter 3, we investigate the achievable coding schemes for both the CIC and DIC problems. Then, in Chapter 4, we develop a number of performance bounds for the CIC and DIC problems. Thirdly, in Chapter 5, we consider the security and privacy issues in index coding. Detailed description and formal definition of the system model and problem setup are presented in Chapter 2, where we also provide some useful mathematical preliminaries and review several pertinent existing results. We conclude the thesis in Chapter 6 with several concluding remarks and a brief discussion on future directions. In the following we introduce our technical contributions in a more detailed manner.

1.3.1 Proposed Achievable Schemes

We study the achievable coding scheme design for the CIC and DIC problems from a random coding perspective in Chapter 3. Generally speaking, we focus on the simplification, improvement, and generalization (to multi-server scenario) of the composite coding (CC) scheme, introduced for the CIC problem in Arbabjolfaei et al. [2013]. We list our contributions as follows.

- As a starting point, in Section 3.1 we address the high computational complexity issue of the CC scheme by proposing a series of simplification techniques. More specifically, we propose a sufficient condition to systematically exclude some decoding choices for the receivers without any performance loss on the achievable rate region. Also, we introduce a pairwise comparison method to safely remove some intermediate variables involved in the expression and computation of the achievable rate region. Based on such techniques a simplified CC scheme is established, leading towards not only reduced complexity, but also a better understanding of the coding scheme.

- We then improve the performance of the CC scheme to achieve larger achievable rate regions from two independent directions. One direction is to employ a more flexible fractional allocation of the broadcast channel capacity to achieve a better convexified achievable rate region compared to that of the original CC scheme. We formalize this idea to introduce the *enhanced* composite coding (ECC) scheme in Section 3.2. The other direction is motivated by a simple observation: the CC scheme, consisting of two layers of random coding, subsumes and strictly outperforms the flat coding scheme Arbabjolfaei et al. [2013], which is a single-layer random coding scheme. Hence, it is natural to ask whether we can further improve the CC scheme by adding one more layer of random coding into it. In Section 3.3, we answer this question positively by introducing the *three-layer* composite coding (TLCC) scheme and showing that the TLCC scheme yields strictly larger achievable rate region than the CC scheme through a concrete example.

- In Sections 3.4-3.5, we generalize the CC scheme to the multi-server DIC problem. We build our *distributed* composite coding (DCC) scheme through combining our enhanced fractional allocation method of the server link capacities and the cooperative compression idea from Li et al. [2017, 2018], and, moreover, adding a new dimension of decoding flexibility. For gentler presentation, in Section 3.4, we first describe the basic form of our DCC scheme with fixed decoding choices for the receivers and provide the achievable rate region and detailed error analysis. Then in Section 3.5, we present our general DCC scheme with varying decoding choices and its associated achievable rate region.

Most results in this chapter have been presented in Liu et al. [2020c, 2017, 2018b,c], which are also listed below for ease of reference:

Liu et al. [2020c] **Y. Liu**, P. Sadeghi, F. Arbabjolfaei, and Y.-H. Kim, "Capacity theorems for distributed index coding," *IEEE Trans. Inf. Theory*, vol. 66, no. 8, pp. 4653–4680, Mar. 2020.
Liu et al. [2017] **Y. Liu**, P. Sadeghi, F. Arbabjolfaei, and Y.-H. Kim, "On the capacity for distributed index coding," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Aachen, Germany, Jun. 2017, pp. 3055–3059.
Liu et al. [2018b] **Y. Liu**, P. Sadeghi, F. Arbabjolfaei, and Y.-H. Kim, "Simplified composite coding for index coding," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Vail, CO, Jun. 2018, pp. 456–460.
Liu et al. [2018c] **Y. Liu**, P. Sadeghi, and Y.-H. Kim, "Three-layer composite coding for index coding," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Guangzhou, China, Nov. 2018, pp. 170–174.

1.3.2 Proposed Performance Bounds

In Chapter 4, we study the information-theoretic converse results and establish a series of performance bounds based on Shannon-type inequalities for the CIC and DIC problems. Our contributions are summarized as follows.

- First, in Section 4.1, we generalize the *alignment chain* model Maleki et al. [2014] for the CIC problem to develop the *basic acyclic chain* model. Based on such model, we propose an iterative performance bound on the symmetric capacity for the CIC problem, namely the basic acyclic chain bound. This new bound subsumes the maximal acyclic induced subgraph (MAIS) bound Bar-Yossef et al. [2011] based on the acyclic structure and the internal conflict bound Maleki et al. [2014]; Jafar [2014] based on the alignment chain model as special cases and can be strictly tighter.

- In Section 4.2, we further generalize the basic acyclic chain model by allowing a more flexible way of interactions among the segments of the chain. We name such generalized acyclic chain model as the *regular acyclic chain*. In a similar manner to the basic acyclic chain bound, we develop the more general regular acyclic chain bound on the symmetric capacity for the CIC problem. Note that both the basic and regular acyclic chain bounds are established purely based on Shannon-type inequalities, which indicates that they are implied by the polymatroidal (PM) bound Blasiak et al. [2011]. Nevertheless, the acyclic chain bounds have the merit over the PM bound of lower computational complexity, especially for large CIC problems where the complexity of computing the PM bound is forbiddingly high.

- We move on to developing performance bounds for the DIC problem in Section 4.4. The definitions of the acyclic chain models do not depend on the server setup and thus also apply to the DIC problem. However, in general it is not clear whether there exist some iterative DIC performance bounds as counterparts to the acyclic chain bounds for the CIC problem introduced above. We reduce the acyclic chain models to certain less general variants, and then propose two associated (non-iterative) DIC performance bounds, which are tight on the symmetrical capacity for some DIC problems.

- In the sequel, we extend the PM bound Blasiak et al. [2011] for the CIC problem to the DIC scenario in Section 4.5. To do this, we incorporate the most flexible use of server groups to derive the necessary conditions for the achievable rates, and subsequently we name our general multi-server performance bound as the *grouping polymatroidal* (PM) bound. Through utilizing different server groupings of varied granularities, the grouping PM bound allows a natural tradeoff between its tightness and computational complexity. In particular, in Sections 4.6-4.7, we specify a number of useful server grouping construction techniques and present the corresponding specialized grouping PM bounds. We also formalize the idea of the hierarchy of server groupings in terms of the tightness and complexity of the resulted bounds in Section 4.8.

Most results in this chapter have been presented in Liu et al. [2020c, 2017]; Liu and Sadeghi [2019b,a], which are also listed below for ease of reference:

Liu et al. [2020c] **Y. Liu**, P. Sadeghi, F. Arbabjolfaei, and Y.-H. Kim, "Capacity theorems for distributed index coding," *IEEE Trans. Inf. Theory*, vol. 66, no. 8, pp. 4653–4680, Mar. 2020.

Liu et al. [2017] **Y. Liu**, P. Sadeghi, F. Arbabjolfaei, and Y.-H. Kim, "On the capacity for distributed index coding," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Aachen, Germany, Jun. 2017, pp. 3055–3059.

Liu and Sadeghi [2019b] **Y. Liu** and P. Sadeghi, "Generalized alignment chain: improved converse results for index coding," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Paris, France, Jul. 2019, pp. 1242–1246.

Liu and Sadeghi [2019a] **Y. Liu** and P. Sadeghi, "From alignment to acyclic chains: lexicographic performance bounds for index coding," in *Proc. 57th Ann. Allerton Conf. Comm. Control Comput.*, Monticello, IL, Sep. 2019, pp. 260–267.

1.3.3 Security and Privacy Investigation

In Chapter 5, we investigate the security and privacy aspects of the index coding problem. We restrict our attention to the single-server CIC problem in this chapter. Our contributions include the following:

- In Section 5.1, we formulate and investigate the *secure* CIC problem with security constraints on the legitimate receivers. That is, there exists a prohibited message set for each receiver. The server must broadcast in a way such that upon receiving the transmitted codeword, each receiver is able to decode its requested message and, at the same time, is not able to decode any single message in its prohibited message set. We first provide a linear coding scheme to fulfill both the decoding and security requirements at the receivers, which is an extended version of the fractional local partial clique covering scheme Arbabjolfaei and Kim [2014] introduced for the normal CIC problem. We then propose two performance bounds for the secure CIC problem together with two associated necessary conditions for a given CIC problem to be *securely feasible* (i.e., to have nonzero rates for every message).

- In Section 5.2, we study the fundamental *privacy-utility tradeoff* within a multi-terminal *guessing* framework inspired by the CIC problem, where the centralized server broadcasts a (possibly stochastically coded) codeword to multiple legitimate receivers, which is also overheard by an adversary. Being a data publishing problem rather than a communication problem, the goal is to balance the privacy leakage and the data utility instead of maximizing the communication rate. We propose two information-theoretic performance bounds on the privacy leakage given the source distribution and utility constraints. As for the achievable scheme (i.e., privacy mechanism), we develop a greedy algorithm based on the agglomerative clustering method that has been used in the information bottleneck and privacy funnel problems Slonim and Tishby [2000]; Makhdoumi et al. [2014]; Ding and Sadeghi [2019].

The results in this chapter have been presented in Liu et al. [2020a,b], which are also listed below for ease of reference:

Liu et al. [2020a] **Y. Liu**, N. Ding, P. Sadeghi, and T. Rakotoarivelo, "Privacy-utility tradeoff in a guessing framework inspired by index coding," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Los Angeles, CA, Jun. 2020.

Liu et al. [2020b] **Y. Liu**, P. Sadeghi, N. Aboutorab, and A. Sharififar, "Secure index coding with security constraints on receivers," in *Proc. Int. Symp. on Inf. Theory and its Applications (ISITA)*, Kapolei, HI, Oct. 2020.

1.3.4 Notation

For non-negative integers a and b, $[a]$ denotes the set $\{1, 2, \dots, a\}$, and $[a : b]$ denotes the set $\{a, a+1, \dots, b\}$. If $a > b$, $[a : b] = \emptyset$.

For a finite set A, $|A|$ denotes its cardinality, and 2^A denotes the set of all subsets of A. For a subset A of a ground set, A^c denotes its complement with respect to the ground set. For two sets A and B, $A \setminus B \doteq \{i \in A : i \notin B\}$. Throughout the thesis, unless otherwise specified, the base of logarithm is 2.

For any discrete random variable Z with probability distribution P_Z, we denote its alphabet by $\mathcal{Z}$ with realizations $z \in \mathcal{Z}$. We denote an estimation of Z by $\hat{Z}$, whose alphabet is also $\mathcal{Z}$. The entropy of the discrete random variable Z is denoted by $H(Z)$. For two discrete random variables X, Y, their joint entropy and mutual information are denoted by $H(X, Y)$ and $I(X; Y)$, respectively. The conditional entropy of X given Y is denoted by $H(X|Y)$. For discrete random variables X, Y, Z, the conditional mutual information between X and Y given Z is denoted by $I(X; Y|Z)$. A review on the aforementioned information measures is presented in Section 2.2.

Chapter 2

System Model and Preliminaries

In this chapter we provide a detailed description for the system model and give necessary mathematical preliminaries. More specifically, in Section 2.1 we describe the system model for the distributed index coding (DIC) problem, which includes the centralized index coding (CIC) problem as a special case. We also formally introduce a number of key definitions for both the DIC and CIC problems, such as the distributed and centralized index codes, the capacity region, and the symmetric capacity. In Section 2.2 we recall some basic graph-theoretic terms and notions. We also define a crucial structure, namely the *touch* structure, which will be used extensively in later chapters, especially for the DIC problem. In Section 2.3, we review several existing results that are pertinent to our work.

2.1 System Model and Problem Setup

We describe the system model for the distributed index coding (DIC) problem, noting that the centralized index coding (CIC) problem is a special case of the DIC problem where there is only one server that contains every message.

Consider the DIC problem with n messages, $x_i \in \{0,1\}^{t_i}$, $i \in [n]$, each of length t_i bits. For brevity, when we say message i, we mean message x_i. Let X_i be the random variable corresponding to x_i. For any set $K \subseteq [n]$, we use the shorthand notation x_K and X_K to denote the collection of messages and the collection of message random variables, whose index is in K, respectively. We use the convention $x_\emptyset = X_\emptyset = \emptyset$. We assume that $X_1, \ldots, X_n$ are uniformly distributed and independent of each other.

There are 2^n servers, one per each subset of the messages. The server indexed by $J \in N \doteq 2^{[n]}$ has access to messages x_J. For brevity, when we say server J, we mean the server indexed by J. Server J is connected to all receivers via a noiseless broadcast link of finite capacity $C_J \geq 0$. Note that this model allows for all possible servers containing any subset of messages to be present in the system. If $C_J = 1$ for $J = [n]$ and is zero otherwise, we recover the CIC problem. Let y_J be the output of server J, which is a deterministic function of x_J, and Y_J be the random variable corresponding to y_J. Note also that for simplicity of notation, we include a *dummy* server indexed by $J = \emptyset$, which does nothing ($C_\emptyset = 0$ and $Y_\emptyset = \emptyset$). For a collection $P \subseteq N$ of servers, we use the shorthand notation y_P and Y_P to denote the corresponding collection of outputs and the collection of output random variables from servers in P, respectively. In particular, for $P = N$, Y_N denotes the collection of output random variables from all the servers.

There are n receivers, where receiver $i \in [n]$ wishes to obtain x_i and knows x_{A_i} as side information for some $A_i \subseteq [n] \setminus \{i\}$. The set of messages which receiver i does not know or want is referred to as the *interfering message set* of receiver i and denoted by $B_i \doteq [n] \setminus A_i \setminus \{i\} = (A_i \cup \{i\})^c$. For any set of receivers $K \subseteq [n]$, define B_K as

$$B_K \doteq \bigcap_{i \in K} B_i, \tag{2.1}$$

denoting the *common* interfering message set of the receivers in K. One can easily verify that

$$B_{K'} \subseteq B_K, \qquad \forall K \subseteq K' \subseteq [n]. \tag{2.2}$$

Any instance of such DIC problem can be represented by a sequence $(i|j \in A_i)$, $i \in [n]$ describing the side information availability at receivers, together with a tuple $\mathbf{C} \doteq (C_J, J \in N)$ describing the server broadcast link capacities. The side information availability can also be represented by a directed graph with n vertices, namely the *side information graph*, denoted by $\mathcal{G}$. In the side information graph $\mathcal{G}$, vertex $i \in [n]$ represents message/receiver i, and a directed edge from i to j, denoted by (i, j), means that $i \in A_j$[1]. Therefore, any DIC problem instance can also be represented by the tuple $(\mathcal{G}, \mathbf{C})$. For the CIC problem where there is only one central server $J = [n]$ with unit link capacity $C_{[n]} = 1$, we denote any given instance by its corresponding $\mathcal{G}$ or the sequence $(i|j \in A_i)$, $i \in [n]$ only.

The central question of the DIC problem is to find the maximum amount of information that can be communicated to the receivers and the optimal coding scheme that achieves this maximum. To answer this question formally, we define a $(\mathbf{t}, \mathbf{r}) = ((t_i, i \in [n]), (r_J, J \in N))$ *distributed index code* by

- 2^n encoders, one for each server $J \in N$, such that

$$\phi_J : \prod_{j \in J} \{0,1\}^{t_j} \to \{0,1\}^{r_J}$$

maps the messages x_J in server J to an r_J-bit sequence y_J, and

[1] Note that in some work (e.g., Bar-Yossef et al. [2011]; Ong [2014]) edge (i, j) in the side information graph means that $j \in A_i$ instead of $i \in A_j$.

- n decoders, one for each receiver $i \in [n]$, such that

$$\psi_i : \prod_{J \in N} \{0,1\}^{r_J} \times \prod_{k \in A_i} \{0,1\}^{t_k} \to \{0,1\}^{t_i}$$

maps the sequences y_N and the side information x_{A_i} to $\hat{x}_i$.

Given a link capacity tuple $\mathbf{C}$, we say a rate tuple $\mathbf{R} \doteq (R_i, i \in [n])$ is achievable if for every $\epsilon > 0$, there exists a $(\mathbf{t}, \mathbf{r})$ code and a common normalization positive integer r such that the message rates $\frac{t_i}{r}$, $i \in [n]$ and codeword rates $\frac{r_J}{r}$, $J \in N$ satisfy

$$R_i \leq \frac{t_i}{r}, \quad i \in [n], \qquad C_J \geq \frac{r_J}{r}, \quad J \in N, \tag{2.3}$$

and the average probability of error satisfies

$$\mathrm{P}\{(\hat{X}_1, \ldots, \hat{X}_n) \neq (X_1, \ldots, X_n)\} \leq \epsilon. \tag{2.4}$$

For the DIC problem $\mathcal{G}$: $(i|j \in A_i)$, $i \in [n]$ with link capacity tuple $\mathbf{C}$, its *capacity region* $\mathscr{C}(\mathcal{G}, \mathbf{C})$ is the closure of the set of all achievable rate tuples $\mathbf{R}$. Sometimes we consider the *symmetric rate* R such that $\mathbf{R} = (R, R, \ldots, R)$ is achievable. The *symmetric capacity* is defined as

$$C_{\text{sym}}(\mathcal{G}, \mathbf{C}) \doteq \max\{R : (R, \ldots, R) \in \mathscr{C}(\mathcal{G}, \mathbf{C})\}. \tag{2.5}$$

Sometimes we may be interested in the *sum capacity* of the DIC problem, defined as

$$C_{\text{sum}}(\mathcal{G}, \mathbf{C}) \doteq \max\{C_{\text{sum}} : \exists \mathbf{R} \in \mathscr{C}(\mathcal{G}, \mathbf{C}) \text{ s.t. } C_{\text{sum}} = \sum_{i \in [n]} R_i\}. \tag{2.6}$$

For the CIC problem $\mathcal{G}$: $(i|j \in A_i)$, $i \in [n]$, a $(\mathbf{t}, r)$ centralized index code, an achievable rate tuple $\mathbf{R}$, the capacity region $\mathscr{C}(\mathcal{G})$, the symmetric capacity $C_{\text{sym}}(\mathcal{G})$, and the sum capacity $C_{\text{sum}}(\mathcal{G})$ can be defined accordingly. The broadcast rate $\beta(\mathcal{G})$ characterizes the minimum number of transmissions from the server to satisfy all the receivers when the messages are of the same length, and is defined as the reciprocal of the symmetric capacity as

$$\beta(\mathcal{G}) \doteq \frac{1}{C_{\text{sym}}(\mathcal{G})}. \tag{2.7}$$

The broadcast rate can be alternatively defined as

$$\beta(\mathcal{G}) \doteq \inf_{t} \inf_{((t,t,\ldots,t),r) \text{ code}} \frac{r}{t} = \lim_{t \to \infty} \inf_{((t,t,\ldots,t),r) \text{ code}} \frac{r}{t}, \tag{2.8}$$

where the equality between the infimum and the limit follows by Fekete's lemma Fekete [1923] and the subadditivity

$$\inf_{((t_1+t_2,\ldots,t_1+t_2),r)\text{ codes}} r \le \inf_{((t_1,\ldots,t_1),r_1)\text{ codes}} r_1 + \inf_{((t_2,\ldots,t_{,2}),r_2)\text{ codes}} r_2.$$

Note that the broadcast rate is only defined and to be studied for the CIC problem. For the DIC problem, we sometimes study the symmetric capacity as defined in (2.5).

Remark 2.1. *Zero-error* capacity region can be defined for CIC and DIC problems similar to our definition of the capacity region, which allows *vanishing error*, by simply requesting that $\mathrm{P}\{(\hat{X}_1,\ldots,\hat{X}_n) \neq (X_1,\ldots,X_n)\} = 0$. For more details please refer to [Arbabjolfaei and Kim, 2018, Section 1.2]. For the CIC problem it has been shown in Langberg and Effros [2011] that the capacity region is equal to the zero-error capacity region. See also Chan and Grant [2010]. However, it is not known whether these two capacity regions are equal for the more general DIC problem.

We illustrate the system models for the CIC and DIC problems in the following example.

Example 2.1. Figures 2.1(a) and 2.1(b) respectively show the system models for the CIC problem and the DIC problem with $n = 3$ messages. Recall that the server indexed by $J \in N$ contains messages $x_J = (x_j, j \in J)$. The multi-server nature of the DIC problem can lead to fundamentally different properties compared with the CIC problem. For example, consider the CIC and DIC problems, both with the same receiver side information $(1|3), (2|3), (3|2)$, i.e., $A_1 = \{3\}$, $A_2 = \{3\}$, $A_3 = \{2\}$. Figure 2.2 shows the corresponding side information graph $\mathcal{G}$. For the centralized problem, the capacity region remains unchanged if the side information at receiver 1 is removed Tahmasbi et al. [2015], i.e., if A_1 becomes $\emptyset$. However, in its distributed counterpart, as long as the broadcast link capacities from the servers $J_1 = \{1,2\}$ and $J_2 = \{1,3\}$ are positive, removing the side information at receiver 1 does result in a strictly smaller capacity region Sadeghi et al. [2016].

2.2 Mathematical Preliminaries

In this section we introduce a few definitions that will be useful in later sections and chapters.

2.2.1 Graph-Theoretic Terms and Notions

Throughout the thesis, unless otherwise specified, we use $\mathcal{G}$ to denote a directed, finite, and simple graph, and $V(\mathcal{G})$ and $E(\mathcal{G})$ to denote the vertex set and the edge set of the graph, respectively. Similarly, let $\mathcal{U}$ denote an undirected, finite, and simple graph, and $V(\mathcal{U})$ and $E(\mathcal{U})$ denote its corresponding vertex set and edge set, respectively. Recall that the side

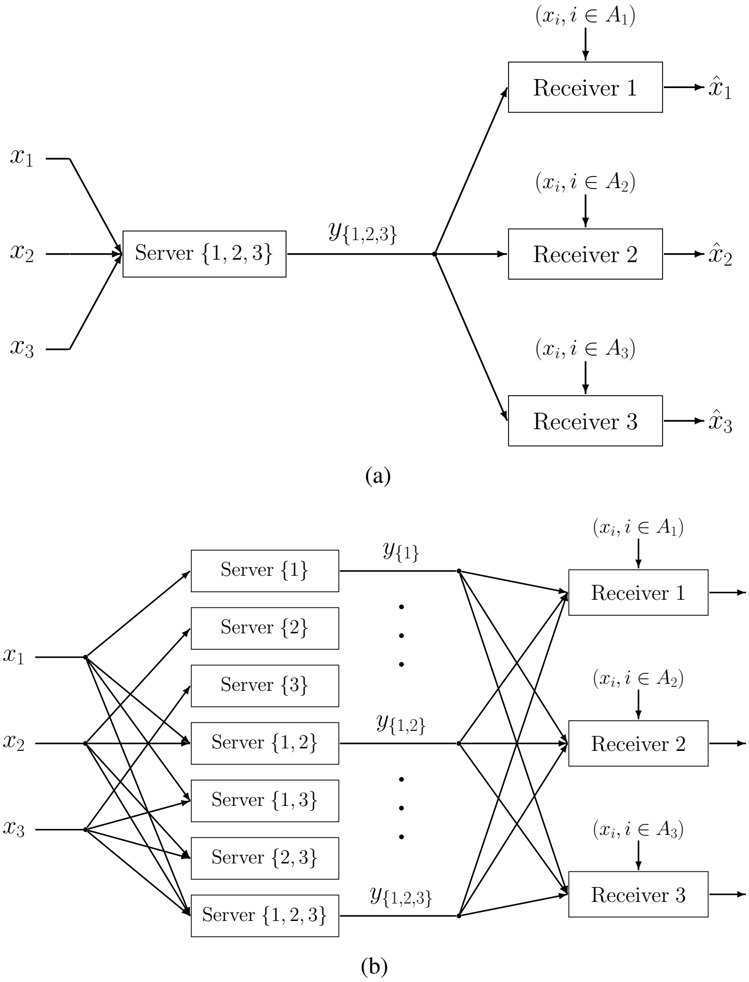


Figure 2.1: Centralized and distributed index coding problems with $n = 3$ messages.

information graph $\mathcal{G}$ for a CIC or DIC problem is a directed graph with $V(\mathcal{G}) = [n]$ and $E(\mathcal{G}) = \{(i,j) : i,j \in [n], i \in A_j\}$.

Given a directed graph $\mathcal{G}$, for any set of vertices $S \subseteq V(\mathcal{G})$, the vertex induced subgraph $\mathcal{G}|_S$ is generated from $\mathcal{G}$ as $V(\mathcal{G}|_S) = \{i \in V(\mathcal{G}) : i \in S\}$ and $E(\mathcal{G}|_S) = \{(i,j) \in E(\mathcal{G}) : i \in S, j \in S\}$. For the CIC problem $\mathcal{G}$ (whose side information graph is $\mathcal{G}$) and a message/receiver set $S \subseteq [n]$, $\mathcal{G}|_S$ denotes the subproblem whose side information graph is $\mathcal{G}|_S$. When the context is clear, we simply use S to denote the subproblem $\mathcal{G}|_S$. Similarly, for the DIC problem $(\mathcal{G}, \mathbf{C})$, $(\mathcal{G}_S, \mathbf{C}_S)$ denotes the subproblem induced by set S, whose side information graph is $\mathcal{G}|_S$ and link capacity tuple $\mathbf{C}_S$ is defined as $\mathbf{C}_S \doteq \{C_J : J \subseteq S\}$.

A group of vertices $v_{i_1}, v_{i_2}, \ldots, v_{i_k}$ form a *cycle* if there is a directed edge $(v_{i_j}, v_{i_{j+1}}) \in E(\mathcal{G})$ for any $j \in [k-1]$, and there is a directed edge $(v_{i_k}, v_{i_1}) \in E(\mathcal{G})$. If $\mathcal{G}|_S$ is *acyclic* (i.e., $\mathcal{G}|_S$ does not contain any cycle), then we say the set S forms an acyclic structure or S

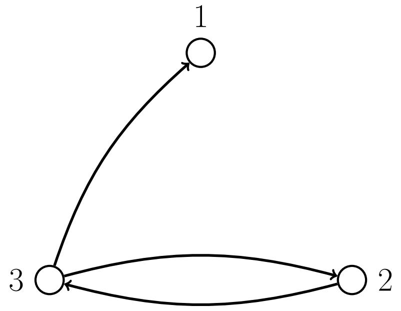


Figure 2.2: The side information graph $\mathcal{G}$ corresponds to the side information availability sequence $(1|3), (2|3), (3|2)$.

is acyclic. It can be verified by contradiction that for any directed graph $\mathcal{G}$, set $S \subseteq V(\mathcal{G})$ is acyclic if and only if there exists an ordering of the vertices in S, say $i_1, i_2, \ldots, i_{|S|}$, such that

$$\begin{aligned} (i_j, i_{|S|}) &\notin E(\mathcal{G}), & \forall j &\in [|S|-1], \\ (i_j, i_{|S|-1}) &\notin E(\mathcal{G}), & \forall j &\in [|S|-2], \\ (i_j, i_{|S|-2}) &\notin E(\mathcal{G}), & \forall j &\in [|S|-3], \\ &\cdots, & &\cdots, \\ (i_j, i_3) &\notin E(\mathcal{G}), & \forall j &\in [2], \\ (i_j, i_2) &\notin E(\mathcal{G}), & j &= 1. \end{aligned} \tag{2.9}$$

Note that (2.9) can be compactly represented as

$$(i_j, i_\ell) \notin E(\mathcal{G}), \qquad \forall j \in [\ell - 1], \ell \in [2 : |S|]. \tag{2.10}$$

If $\mathcal{G}$ is a side information graph for a CIC or DIC problem, then as every directed edge $(i, j) \in E(\mathcal{G})$ denotes $i \in A_j$, for the acyclic set $S = \{i_1, i_2, \ldots, i_{|S|}\} \subseteq [n]$, (2.10) simplifies to

$$i_j \notin A_{i_\ell}, \qquad \forall j \in [\ell - 1], \ell \in [2 : |S|], \tag{2.11}$$

which can be equivalently represented using the interfering message sets as

$$\{i_1, i_2, \ldots, i_{\ell-1}\} \subseteq B_{i_\ell}, \qquad \forall \ell \in [2 : |S|]. \tag{2.12}$$

So far we discussed the sufficient and necessary condition for a set of vertices to be acyclic. In a similar manner, we can define the notion of being *set-level* acyclic for a set of vertex sets. Consider a directed graph $\mathcal{G}$ and a group of disjoint vertex sets $S_1, S_2, \ldots, S_h \subseteq V(\mathcal{G})$, $S_k \cap S_\ell = \emptyset$, $\forall k \neq \ell \in [h]$. We say that these vertex sets are acyclic at the set level if and only if we have

$$(i, j) \notin E(\mathcal{G}), \qquad \forall i \in S_k, j \in S_\ell, k \in [\ell - 1], \ell \in [2 : h]. \tag{2.13}$$

If $\mathcal{G}$ is a side information graph for a CIC or DIC problem, (2.13) simplifies to

$$i \notin A_j, \qquad \forall i \in S_k, j \in S_\ell, k \in [\ell - 1], \ell \in [2 : h], \tag{2.14}$$

which can be equivalently represented using the common interfering message sets defined in (2.1) as

$$S_1 \cup S_2 \cup \cdots \cup S_{\ell-1} \subseteq B_{S_\ell}, \qquad \forall \ell \in [2 : h]. \tag{2.15}$$

Example 2.2. Consider the 4-message CIC problem

$$(1|4), (2|1), (3|1, 2), (4|1),$$

whose side information graph is shown in Figure 2.3(a). The message set $\{1, 2, 3\}$ induces an acyclic subgraph, and thus is acyclic. One can verify that (2.12) holds as

$$\emptyset \subseteq B_3, \quad \{3\} \subseteq B_2, \quad \{2, 3\} \subseteq B_1.$$

Next consider the 6-message CIC problem

$$(1|2), (2|1), (3|1, 4), (4|3), (5|6), (6|4, 5),$$

whose side information graph is shown in Figure 2.3(b). Message sets $\{5, 6\}$, $\{3, 4\}$, $\{1, 2\}$ are acyclic at the set level as we have

$$\emptyset \subseteq B_{\{5,6\}}, \quad \{5, 6\} \subseteq B_{\{3,4\}}, \quad \{3, 4, 5, 6\} \subseteq B_{\{1,2\}}.$$

Then consider the 8-message CIC problem

$$(1|2, 3, 7, 8), (2|1, 5, 6), (3|5, 7, 8), (4|6, 8), (5|4, 6), (6|5, 7), (7|8), (8|1, 3, 4),$$

whose side information graph is shown in Figure 2.3(c). Message sets $\{1, 2\}$, $\{3\}$, $\{4, 5, 6\}$, $\{7\}$ are acyclic at the set level as we have

$$\emptyset \subseteq B_{\{1,2\}}, \quad \{1, 2\} \subseteq B_{\{3\}}, \quad \{1, 2, 3\} \subseteq B_{\{4,5,6\}}, \quad \{1, 2, 3, 4, 5, 6\} \subseteq B_{\{7\}}.$$

In the following we review the notion of lexicographic product of two directed graphs

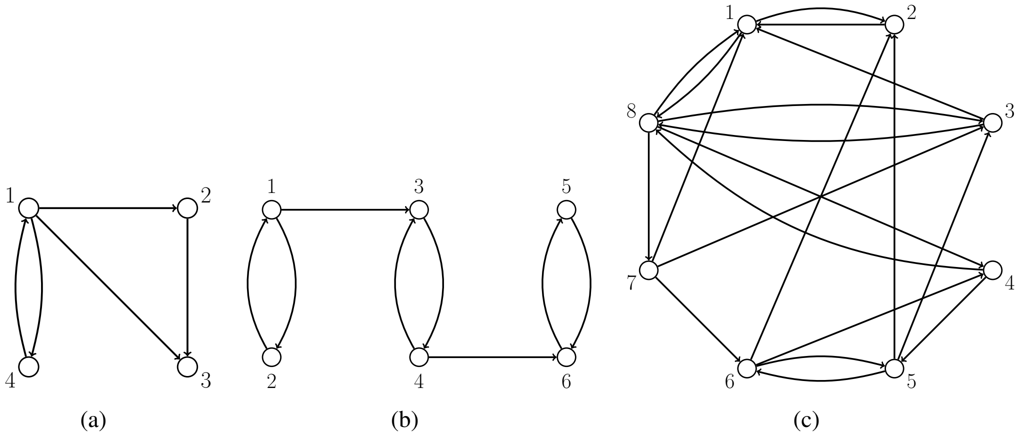


Figure 2.3: Three CIC problems where (a) message set $\{1,2,3\}$ is acyclic, (b) message sets $\{5,6\}$, $\{3,4\}$, $\{1,2\}$ are acyclic at the set level, and (c) message sets $\{1,2\}$, $\{3\}$, $\{4,5,6\}$, $\{7\}$ are acyclic at the set level.

Hammack et al. [2011]; Arbabjolfaei and Kim [2018].

Definition 2.1 (Lexicographic product, Hammack et al. [2011]; Arbabjolfaei and Kim [2018])**.** For any two directed graphs $\mathcal{G}_0$ and $\mathcal{G}_1$, the directed graph

$$\mathcal{G} = \mathcal{G}_0 \circ \mathcal{G}_1$$

is called the lexicographic product of $\mathcal{G}_0$ and $\mathcal{G}_1$ if and only if

$$V(\mathcal{G}) = V(\mathcal{G}_0) \times V(\mathcal{G}_1) = \{(i_1, i_2) : i_1 \in V(\mathcal{G}_0), i_2 \in V(\mathcal{G}_1)\},$$

and

$$E(\mathcal{G}) = \{((i_1, i_2), (j_1, j_2)) : (i_1, j_1) \in E(\mathcal{G}_0) \text{ or } i_1 = j_1, (i_2, j_2) \in E(\mathcal{G}_1)\}.$$

That is, to produce $\mathcal{G}$ as $\mathcal{G}_0 \circ \mathcal{G}_1$, every vertex in $\mathcal{G}_0$ is replaced by a copy of $\mathcal{G}_1$, and all the vertices of one copy of $\mathcal{G}_1$ are connected with all the vertices of another copy according to the edge connectivity given by $E(\mathcal{G}_0)$. If $\mathcal{G}_0$ and $\mathcal{G}_1$ both represent some CIC problems (as their side information graphs), then by computing the lexicographic product of them we generate a new CIC problem with side information graph $\mathcal{G} = \mathcal{G}_0 \circ \mathcal{G}_1$. For an example see Figure 2.4.

For a given directed graph $\mathcal{G}$ and any positive integer k, define the shorthand notation

$$\mathcal{G}^{\circ k} \doteq \underbrace{\mathcal{G} \circ \mathcal{G} \circ \cdots \circ \mathcal{G}}_{k \text{ times}}.$$

In this thesis, we study the structural properties of the lexicographic product of CIC problems (i.e., their side information graphs) only, and do not consider that of DIC problems.

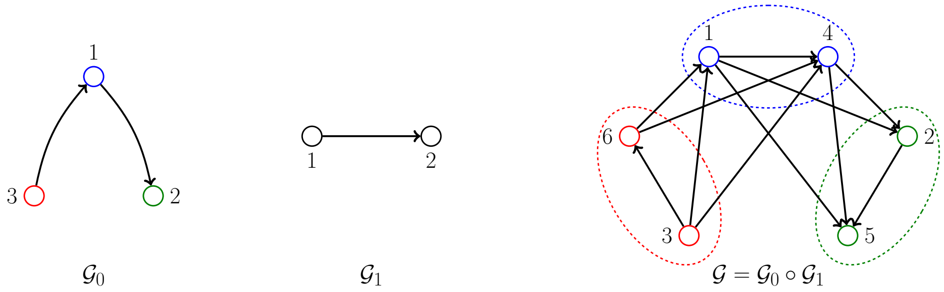


Figure 2.4: An example of the lexicographic product of two directed graphs. If $\mathcal{G}_0$ represents the 3-message CIC problem $(1|3),(2|1),(3|-)$ and $\mathcal{G}_1$ represents the 2-message CIC problem $(1|-),(2|1)$, then $\mathcal{G} = \mathcal{G}_0 \circ \mathcal{G}_1$ represents the 6-message CIC problem $(1|3,6),(2|1,4),(3|-),(4|1,3,6),(5|1,2,4),(6|3)$, which is the lexicographic product of $\mathcal{G}_0$ and $\mathcal{G}_1$. Note that according to Definition 2.1, the vertex set of $\mathcal{G}$ is $\{(1,1),(1,2),(2,1),(2,2),(3,1),(3,2)\}$. We re-label these vertices as $(1,1)=1$, $(1,2)=4$, $(2,1)=2$, $(2,2)=5$, $(3,1)=3$, $(3,2)=6$.

2.2.2 Information Measures

We briefly review some of the most basic and classic information measures in the field of information theory, which were introduced in the seminal paper Shannon [1948] by Claude Shannon, and thus sometimes referred to as Shannon's information measures. For more details, see [Cover, 2006, Chapter 2].

For a discrete random variable Z with alphabet $\mathcal{Z}$ and probability distribution P_Z, the *entropy* of Z is defined as

$$H(Z) = \sum_{z \in \mathcal{Z}} P_Z(z) \log \frac{1}{P_Z(z)}.$$

Note that in the above definition, we set by convention that $0 \log \frac{1}{0} = 0$, which can be justified by continuity as $x \log \frac{1}{x} \to 0$ when $x \to 0$.

Roughly speaking, entropy is average of the uncertainty of the random variable. The entropy is non-negative and upper-bounded by the logarithm of the alphabet size. That is,

$$0 \leq H(Z) \leq \log |\mathcal{Z}|, \tag{2.16}$$

where the lower bound is tight if and only if Z is deterministic, and the upper bound is tight if and only if Z has a uniform distribution.

For two discrete random variables X and Y with alphabet $\mathcal{X}$ and $\mathcal{Y}$, respectively, and joint distribution $P_{X,Y}$, their *joint entropy* is defined as

$$H(X,Y) = \sum_{x \in \mathcal{X}, y \in \mathcal{Y}} P_{X,Y}(x,y) \log \frac{1}{P_{X,Y}(x,y)}.$$

Note that the random variables X, Y can be collectively viewed as a single vector-valued random variable $Z = (X, Y)$. Subsequently, the joint entropy $H(X, Y)$ is simply the entropy of $Z = (X, Y)$.

As a measure of the average remaining uncertainty of a random variable Y given the knowledge of another random variable X, we define the *conditional entropy* of Y given X as

$$H(Y|X) = \sum_{x\in\mathcal{X}} P_X(x) \sum_{y\in\mathcal{Y}} P_{Y|X}(y|x) \log \frac{1}{P_{Y|X}(y|x)} = H(X, Y) - H(X),$$

where the second equality characterizing the relationship between entropy, joint entropy, and conditional entropy is often referred to as the *chain rule*. The conditional entropy $H(Y|X)$ is non-negative and no larger than the entropy $H(Y)$. That is,

$$0 \leq H(Y|X) \leq H(Y),$$

where the lower bound is tight if and only if Y is a deterministic function of X, and the upper bound is tight if and only if X and Y are independent.

The average reduction of uncertainty of Y due to the knowledge of X is defined as the *mutual information* between Y and X, denoted as $I(Y; X)$. Clearly, we have

$$I(Y; X) = H(Y) - H(Y|X) = \sum_{x\in\mathcal{X}, y\in\mathcal{Y}} P_{X,Y}(x, y) \log \frac{P_{X,Y}(x, y)}{P_X(x)P_Y(y)}.$$

The mutual information is symmetric in X and Y, non-negative, and upper-bounded by the entropies of X and Y. That is,

$$\begin{aligned}
&I(Y; X) = I(X; Y) = H(X) + H(Y) - H(X, Y),\\
&0 \leq I(X; Y) \leq \min\{H(X), H(Y)\},
\end{aligned}$$

where the lower bound is tight if and only if X and Y are independent, and the upper bound is tight if and only if one of X or Y is a deterministic function of the other.

The *conditional mutual information* between X and Y given Z is defined as

$$\begin{aligned}
I(X; Y|Z) &= H(X|Z) - H(X|Y, Z)\\
&= \sum_{x\in\mathcal{X}, y\in\mathcal{Y}, z\in\mathcal{Z}} P_{X,Y,Z}(x, y, z) \log \frac{P_{X,Y|Z}(x, y|z)}{P_{X|Z}(x|z)P_{Y|Z}(y|z)}\\
&= I(X; Y, Z) - I(X; Z),
\end{aligned}$$

where the last equality is the chain rule for mutual information and $I(X; Y, Z)$ is simply the mutual information between X and the vector-valued random variable (Y, Z).

2.2.3 Touch Structure

Now we define the following *touch* structure, which will be used extensively when dealing with multiple servers for the DIC problem.

Definition 2.2 (Touch structure)**.** For any set of messages $K \subseteq [n]$ and a given server $J \in N$, we say that server J *touches* K if $J \cap K \neq \emptyset$ and that J *does not touch* K if $J \cap K = \emptyset$. We denote by T_K the collection of servers that touch K and denote by $T_{\overline{K}}$ the collection of servers that do not touch K, that is,

$$T_K = \{J \in N : J \cap K \neq \emptyset\},$$
$$T_{\overline{K}} = \{J \in N : J \cap K = \emptyset\} = \{J \in N : J \subseteq [n] \setminus K\} = N \setminus T_K = 2^{K^c}.$$

And furthermore, for two sets of messages K and L,

$$T_{K,L} = T_{L,K} = \{J \in N : J \cap K \neq \emptyset, J \cap L \neq \emptyset\},$$
$$T_{K,\overline{L}} = T_{\overline{L},K} = \{J \in N : J \cap K \neq \emptyset, J \cap L = \emptyset\}.$$

Example 2.3. If $n = 4$, then

$$T_{\{1\}} = \{\{1\}, \{1,2\}, \{1,3\}, \{1,4\}, \{1,2,3\}, \{1,2,4\}, \{1,3,4\}, \{1,2,3,4\}\},$$
$$T_{\{1\},\overline{\{2\}}} = \{\{1\}, \{1,3\}, \{1,4\}, \{1,3,4\}\}.$$

Remark 2.2. Any set T_K can be broken into two disjoint subsets $T_{K,L}$ and $T_{K,\overline{L}}$. Thus, $T_{K,L} \subseteq T_K, T_{K,\overline{L}} \subseteq T_K$ in general and $T_{K,L} = T_K$ if $K \subseteq L$. It is also easy to verify that $T_K \cup T_L = T_{K \cup L}$ and $T_{K \cap L} \subseteq T_{K,L} = T_K \cap T_L$. Definition 2.2 can be naturally extended to three or more message sets. For example, if $n = 5$, then

$$T_{\{1,4\},\overline{\{3\}},\{4,5\}} = \{\{4\}, \{1,4\}, \{1,5\}, \{1,4,5\}, \{2,4\}, \{1,2,4\}, \{1,2,5\}, \{1,2,4,5\}\}.$$

In our usual notation, Y_{T_K} thus denotes the output random variables from servers that have at least one message from the message set K, e.g., $Y_{T_{[n]}} = Y_N$. When the context is clear, we shall use the shorthand notation T_K for Y_{T_K}, e.g., $H(T_K)$ means $H(Y_{T_K})$.

2.3 Review of Some Existing Results

We briefly review some existing results on both the achievable coding schemes and the performance bounds for the CIC and DIC problems that are pertinent to this thesis.

2.3.1 Existing Coding Schemes for the CIC problem

We begin with several well-known CIC coding schemes, which can be broadly categorized into two classes: linear CIC coding schemes and non-linear CIC coding schemes. Every coding scheme leads to an inner bound on the capacity region $\mathscr{C}$ and/or an upper bound on the broadcast rate β.

2.3.1.1 Existing Linear CIC Coding Schemes

Consider the CIC problem $\mathcal{G}$: $(i|j \in A_i)$, $i \in [n]$. Let $a_{\min} \doteq \min_{i\in[n]} |A_i|$ denote the minimum number of side information messages across the receivers. Obviously $0 \leq a_{\min} < n$. Over sufficiently large fields, there always exists some systematic $(n + n - a_{\min}, n - a_{\min})$ MDS code with total $n + n - a_{\min}$ code symbols, including n message symbols and $n - a_{\min}$ parity symbols, such that given any n of these total $n + n - a_{\min}$ symbols, one can recover the rest of them. Therefore by broadcasting the $n - a_{\min}$ parity symbols from the central server to the receivers, receiver j has $|A_j| + n - a_{\min} = |A_j| + n - \min_{i\in[n]} |A_i| \geq n$ MDS code symbols and hence is capable of recovering all the remaining symbols. That is, broadcasting the $n - a_{\min}$ parity symbols will enable every receiver to decode every message not in its side information. Such coding schemes leads to the following achievability bound on β.

Proposition 2.1 (MDS bound, Birk and Kol [1998]). Consider the CIC problem $\mathcal{G}$: $(i|j \in A_i)$, $i \in [n]$. The broadcast rate $\beta(\mathcal{G})$ is upper bounded by $\beta_{\text{MDS}}(\mathcal{G})$ as

$$\beta(\mathcal{G}) \leq \beta_{\text{MDS}}(\mathcal{G}) \doteq n - a_{\min} = n - \min_{i\in[n]} |A_i|. \tag{2.17}$$

Note that for the CIC problem $\mathcal{G}$ the MDS coding scheme requires $\beta_{\text{MDS}}(\mathcal{G})$ transmissions and thus a symmetric rate of $\frac{1}{\beta_{\text{MDS}}(\mathcal{G})}$ is achieved.

To generalize the above MDS coding scheme for the CIC problem, consider time sharing among a number of induced subproblems $\mathcal{G}|_L$ for some $L \subseteq [n]$. Each subproblem $\mathcal{G}|_L$ is assigned with a certain fraction of the unit channel link capacity λ_L. Any message x_i that appears in multiple subproblems is split into sub-messages $x_{i,L}, i \in L, \lambda_L > 0$. For each subproblem $\mathcal{G}|_L$ we apply the MDS coding scheme. That is, for each subproblem $\mathcal{G}|_L$, a systematic $(|L| + \beta_{\text{MDS}}(\mathcal{G}|_L), |L|)$ MDS code is used such that every receiver $i \in L$ can decode sub-message $x_{i,L}$ at rate $\frac{\lambda_L}{\beta_{\text{MDS}}(\mathcal{G}|_L)}$. Moreover, to exploit the fact that each receiver can recover some MDS code symbols from its own side information, an MDS code is applied at the subproblem level to the MDS code symbols for subproblems. In this way, the channel capacity is shared only among the MDS code symbols not available locally at each receiver.

The above coding scheme is called the fractional local partial clique covering (FLPCC) scheme Arbabjolfaei and Kim [2014]. The following theorem presents the achievable rate region given by the FLPCC scheme.

Proposition 2.2 (Fractional local partial clique covering (FLPCC) bound, Arbabjolfaei and Kim [2014])**.** Consider the CIC problem $\mathcal{G}$: $(i|j \in A_i)$, $i \in [n]$. The capacity region $\mathscr{C}(\mathcal{G})$ is inner bounded by the rate region $\mathscr{R}_{\text{FLPCC}}(\mathcal{G})$ that consists of all rate tuples $\mathbf{R}$ satisfying

$$R_i \leq \sum_{L \subseteq [n]: i \in L} \frac{\lambda_L}{\beta_{\text{MDS}}(\mathcal{G}|_L)}, \qquad \forall i \in [n] \tag{2.18}$$

for some $\lambda_L, L \subseteq [n]$ such that

$$\lambda_L \in [0,1], \qquad \forall L \subseteq [n], \tag{2.19}$$

$$\sum_{L \subseteq [n]: L \not\subseteq A_i} \lambda_L \leq 1, \qquad \forall i \in [n]. \tag{2.20}$$

In particular, the broadcast rate $\beta(\mathcal{G})$ is upper bounded by $\beta_{\text{FLPCC}}(\mathcal{G})$, which is the solution to the optimization problem

$$\text{minimize} \quad \max_{i \in [n]} \sum_{L \subseteq [n]: L \not\subseteq A_i} \kappa_L \cdot \beta_{\text{MDS}}(\mathcal{G}|_L) \tag{2.21}$$

$$\text{subject to} \quad \kappa_L \in [0,1], \qquad \forall L \subseteq [n], \tag{2.22}$$

$$\sum_{L \subseteq [n]: i \in L} \kappa_L \geq 1, \qquad \forall i \in [n]. \tag{2.23}$$

To compute the FLPCC bound, one can use an optimization tool such as Fourier Motzkin Elimination (FME) [El Gamal and Kim, 2011, Appendix D] or linear programming (LP) to eliminate the intermediate variables in Proposition 2.2. The FLPCC scheme leads to tight inner bounds for a number of CIC problems. Nevertheless, it is always outperformed by (i.e., no tighter than) the recursive coding scheme in Arbabjolfaei and Kim [2014]. We do not go through the recursive coding scheme or other existing linear CIC coding schemes in detail as we are not going to study or use them in later chapters.

2.3.1.2 Existing Non-linear CIC Coding Schemes

Next we introduce several existing non-linear coding schemes, which are based on the classic random coding method.

We begin with the *flat* coding scheme Arbabjolfaei et al. [2013], a single-layer random coding scheme for the CIC problem[2]. The codebook generation is as follows. For each realization of messages $x_{[n]}$, a codeword $y(x_{[n]})$ is drawn uniformly at random from $[2^r]$. Such codebook is revealed to all parties. To communicate messages $x_{[n]}$, the server transmits $y(x_{[n]})$. Since the

[2] The codebook generation for flat coding is "flat" over all messages, in contrast to any "layered" superposition coding scheme. In other words, the codeword y is directly generated according to the messages $x_1, \ldots, x_n$ without any intermediate steps.

encoding is flat, receiver i must decode all messages not in its side information. Such coding scheme yields the following achievable rate region.

Proposition 2.3 (Flat coding bound, Arbabjolfaei et al. [2013]). Consider the CIC problem $\mathcal{G}$: $(i|j \in A_i)$, $i \in [n]$. The capacity region $\mathscr{C}(\mathcal{G})$ is inner bounded by the rate region $\mathscr{R}_{\text{flat}}(\mathcal{G})$ that consists of all rate tuples $\mathbf{R}$ satisfying

$$R_i + \sum_{j \in B_i} R_j < 1, \qquad \forall i \in [n]. \tag{2.24}$$

In particular, the broadcast rate $\beta(\mathcal{G})$ is upper bounded by $\beta_{\text{FLAT}}(\mathcal{G})$ as

$$\beta(\mathcal{G}) \leq \beta_{\text{FLAT}}(\mathcal{G}) \doteq 1 + \max_{i \in [n]} |B_i| = n - \min_{i \in [n]} |A_i|. \tag{2.25}$$

Note that the flat coding upper bound $\beta_{\text{flat}}(\mathcal{G})$ is identical to the MDS upper bound $\beta_{\text{MDS}}(\mathcal{G})$. As pointed out in [Arbabjolfaei and Kim, 2018, Remark 6.11], the flat coding scheme can be seen as a random construction of an approximate MDS code, which achieves $\beta_{\text{MDS}}(\mathcal{G})$ asymptotically.

Composite Coding Arbabjolfaei et al. [2013]: For the CIC problem $(1|4), (2|3,4), (3|1,2), (4|2,3)$, flat coding (or more generally time sharing of flat coding over subproblems) gives a suboptimal achievable rate region as $R_1 + R_2 + R_3 < 1$, $R_1 + R_4 < 1$, and $R_3 + R_4 < 1$. However, using the two-layer random coding scheme described below, namely, the *composite coding* (CC) scheme Arbabjolfaei et al. [2013], the capacity region as $R_1 + R_2 < 1$, $R_1 + R_3 < 1$, $R_1 + R_4 < 1$, and $R_3 + R_4 < 1$ can be achieved.

We review the CC scheme in detail as follows.

Codebook generation.

Step 1. For each message set $K \subseteq [n]$ and each realization of messages x_K, generate a composite index $w_K = w_K(x_K)$ drawn uniformly at random from $[2^{s_K}]$, where $s_K = \lceil rS_K \rceil$ and S_K is the rate of the composite index w_K. That is, the composite index w_K is generated according to the random mapping w_K as

$$w_K : \prod_{i \in K} [2^{t_i}] \to [2^{s_K}].$$

Step 2. For each realization of the composite index tuple $(w_K, K \subseteq [n])$, generate a codeword $y = y((w_K, K \subseteq [n]))$ drawn uniformly at random from $[2^r]$. That is, the codeword y is generated according to the random mapping y as

$$y : \prod_{K \subseteq [n]} [2^{s_K}] \to [2^r].$$

The codebook is revealed to all parties.[3]

Encoding. To communicate messages $x_{[n]}$, the server first encodes the composite index $w_K = w_K(x_K)$ for each $K \subseteq [n]$, and then encodes and transmits the codeword $y = y((w_K, K \subseteq [n]))$.

Decoding. Decoding takes place in the reverse order of stages as follows.

Step 1. Receiver i finds the unique composite index tuple $(\hat{w}_K, K \subseteq [n])$ such that $y = y(\hat{w}_K, K \subseteq [n])$. If there is more than one such tuple it declares an error.

Step 2. Assuming $(\hat{w}_K, K \subseteq [n])$ is correct, receiver i decodes for a subset of messages indexed by $D_i \subseteq [n] \setminus A_i$, such that $i \in D_i$. That is, receiver i finds the unique message tuple $\hat{x}_{D_i}$ such that $\hat{w}_K = w_K(\hat{x}_K)$ for all $K \subseteq D_i \cup A_i$. If there is more than one such tuple it declares an error.

The tuple of decoding message sets is denoted by $\mathbf{D} \doteq (D_i, i \in [n])$ and referred to as the *decoding configuration* for composite coding. Let $\mathcal{D}_i(\mathcal{G}) \doteq \{D_i \subseteq [n] \setminus A_i : i \in D_i\}$ denote the set of all possible decoding message sets at receiver i, and thus $\mathcal{D}(\mathcal{G}) \doteq \prod_{i \in [n]} \mathcal{D}_i(\mathcal{G})$ denotes the collection of all possible decoding configurations for the given CIC problem.

The achievable rate region of the CC scheme is summarized below. Note that time sharing is applied over all possible decoding configurations.

Proposition 2.4 (Composite coding (CC) bound, Arbabjolfaei et al. [2013]; Arbabjolfaei and Kim [2018])**.** Consider the CIC problem $\mathcal{G}$: $(i|j \in A_i), i \in [n]$. The capacity region $\mathscr{C}(\mathcal{G})$ is inner bounded by the rate region $\mathscr{R}_{\mathrm{CC}}(\mathcal{G})$ as

$$\mathscr{R}_{\mathrm{CC}}(\mathcal{G}) \doteq \mathrm{co}\Big(\bigcup_{\mathbf{D} \in \mathcal{D}(\mathcal{G})} \mathscr{R}_{\mathrm{CC}}(\mathcal{G}, \mathbf{D})\Big), \tag{2.26}$$

where $\mathrm{co}(\cdot)$ denotes the convex hull, and $\mathscr{R}_{\mathrm{CC}}(\mathcal{G}, \mathbf{D})$ consists of all the rate tuples $\mathbf{R}$ satisfying

$$\sum_{K \subseteq [n], K \nsubseteq A_i} S_K(\mathbf{D}) < 1, \qquad \forall i \in [n], \tag{2.27}$$

$$\sum_{j \in L} R_j < \sum_{\substack{K \subseteq D_i \cup A_i, \\ K \cap L \neq \emptyset}} S_K(\mathbf{D}), \quad \forall L \subseteq D_i,\, i \in [n], \tag{2.28}$$

for some $S_K(\mathbf{D}) \geq 0$, $K \subseteq [n]$.

Here the collection of inequalities in (2.27) signify the first-step decoding constraints for the composite indices to be recovered by each receiver. The collection of inequalities in (2.28)

[3] It is worth contrasting single-layer *flat* coding of messages into random codewords versus two-layer *composite* coding of messages into random composite indices and then composite indices into random codewords. In other words, composite coding can be viewed as random construction of an approximate MDS code for composite indices, rather than messages themselves. This adds flexibility in decoding conditions and can enhance the achievable rate region. For more details see Arbabjolfaei et al. [2013]; Arbabjolfaei and Kim [2018].

signify the second-step decoding constraints for recovering messages in each receiver's decoding message set given the decoding configuration $\mathbf{D}$.

To compute an explicit achievable rate region or rate tuple from (2.27) and (2.28), one has to eliminate the intermediate variables $(S_K(\mathbf{D}), K \subseteq [n])$ using optimization tools such as FME or LP. It has been verified that the CC scheme can give tight inner bounds on the capacity region for all CIC problems with five or fewer messages Arbabjolfaei et al. [2013], as well as many larger problems with six or more messages.

2.3.2 Existing Coding Schemes for the DIC problem

In this section, we briefly discuss the existing coding schemes and their corresponding achievability results for the DIC problem.

There is a series of useful graph-theoretic coding schemes for the CIC problem such as the FLPCC scheme (for more details see [Arbabjolfaei and Kim, 2018, Sections 6.2-6.5]). In contrast, to the best of our knowledge, there has been no existing graph-theoretic coding schemes holding generally for any DIC problem beyond trivial extensions of the CIC graph-theoretic coding schemes (e.g., every server employs its own coding scheme with no cooperation between each other). Several graph-theoretic results have been proposed under certain constraints, such as the cyclic codes for the single-uniprior multi-sender index coding problem where each receiver knows one message as side information Ong et al. [2016a], and the graph theoretic approaches for the two-sender index coding problem Thapa et al. [2016].

The minrank approach for the CIC problem has been extended to the DIC problem in Kim and No [2019]; Li et al. [2019], providing nontrivial coding techniques. However, despite simplification techniques proposed, the code construction can still be computationally expensive. Also the suboptimality of minrank based codes compared to non-linear codes naturally carries over from the CIC to the DIC problem.

As for the non-linear coding schemes for the DIC problem, the CC scheme has been extended to the multi-server scenario in Sadeghi et al. [2016], and then further generalized in Li et al. [2017, 2018]. In Chapter 3, we will propose a new coding scheme, namely the distributed composite coding (DCC) scheme, which is more general than all those existing schemes based on composite coding. More specifically, the DCC scheme adopts the key technique introduced by the cooperative composite coding (CCC) scheme in Li et al. [2018] as a baseline, which is then combined with our own improvements to jointly lead to a more advanced coding scheme. The achievable rate region of the CCC scheme is rather involved and bear a resemblance to that of our DCC scheme. Given such fact, we postpone the details about the CCC scheme's achievability results till after the introduction of the DCC scheme in Sections 3.4 and 3.5.

2.3.3 Existing Performance Bounds for the CIC problem

In the following we review a series of performance bounds for the CIC problem, namely, the maximal acyclic induced subgraph (MAIS) bound Bar-Yossef et al. [2011], the internal conflict bound Maleki et al. [2014]; Jafar [2014], and the polymatroidal (PM) bound Arbabjolfaei et al. [2013]. Every performance bound essentially serves as an outer bound on the capacity region $\mathscr{C}$, which can be specified to a lower bound on the broadcast rate β.

Proposition 2.5 (Maximal acyclic induced subgraph (MAIS) bound, Bar-Yossef et al. [2011])**.** Consider the CIC problem $\mathcal{G}$: $(i|j \in A_i), i \in [n]$. The capacity region $\mathscr{C}(\mathcal{G})$ is outer bounded by the rate region $\mathscr{R}_{\text{MAIS}}(\mathcal{G})$ that consists of all rate tuples $\mathbf{R}$ satisfying

$$\sum_{i \in S:\ S \subseteq [n] \text{ is acyclic}} R_i \leq 1. \tag{2.29}$$

In particular, the broadcast rate $\beta(\mathcal{G})$ is lower bounded by $\beta_{\text{MAIS}}(\mathcal{G})$ as

$$\beta(\mathcal{G}) \geq \beta_{\text{MAIS}}(\mathcal{G}) \doteq \max_{S \subseteq [n] \text{ is acyclic}} |S|. \tag{2.30}$$

It has been verified in [Arbabjolfaei and Kim, 2018, Section 8.6] that for every CIC problem instance with $n = 4$ or fewer messages, the MAIS outer bound is tight as it coincides with the fractional local partial clique covering (FLPCC) inner bound Arbabjolfaei and Kim [2014] (Proposition 2.2 in Section 2.3.1), thus establishing the capacity region. In general however, the MAIS bound is loose.

Example 2.4. Consider the 5-message CIC problem $\mathcal{G}$:

$$(1|2,5), (2|1,3), (3|2,4), (4|3,5), (5|1,4).$$

The MAIS bound yields $\beta(\mathcal{G}) \geq \beta_{\text{MAIS}}(\mathcal{G}) = 2$. Yet the actual broadcast rate $\beta(\mathcal{G})$ for this problem is 2.5 (see Example 2.5 to be presented shortly).

For the internal conflict bound, we first recall the definition of the *alignment chain* model from Maleki et al. [2014].

Definition 2.3 (Alignment chain, Maleki et al. [2014])**.** Consider the CIC problem $\mathcal{G}$: $(i|j \in A_i), i \in [n]$. Messages $i(1), i(2), \cdots, i(m), i(m+1)$ and $k(1), k(2), \cdots, k(m)$ constitute an alignment chain, Ch_{AC}, of length m as

$$Ch_{\text{AC}}: \quad \underline{i(1)} \xleftrightarrow{k(1)} i(2) \xleftrightarrow{k(2)} i(3) \cdots \xleftrightarrow{k(m)} \underline{i(m+1)}, \tag{2.31}$$

if the conditions listed below are satisfied:

1. $i(1) \in B_{i(m+1)}$ or $i(m+1) \in B_{i(1)}$;

2. for any $j \in [m]$, we have $\{i(j), i(j+1)\} \subseteq B_{k(j)}$.

For any alignment chain we call the edge between $i(j)$ and $i(j+1)$ edge j. In Definition 2.3, the two terminals $i(1)$ and $i(m+1)$ of the chain are underlined to indicate that the message set $\{i(1), i(m+1)\}$ is acyclic. We call the i-labeled and k-labeled messages the *components* of the alignment chain Ch_{AC}. The chain becomes trivial when $m = 0$ as it contains only one component $i(1)$. To avoid such trivial case, we always require that $m \geq 1$.

Remark 2.3. For the CIC problem $\mathcal{G}$, let $\mathfrak{C}_{\text{AC}}(\mathcal{G})$ denote the collection of its alignment chains. If there exits at least one alignment chain, then we say that there is an *internal conflict* between the two terminal messages $i(1)$ and $i(m+1)$ of the alignment chain and that the problem is *internally conflicted*. The broadcast rate for the CIC problems that are not internally conflicted (i.e., $\mathfrak{C}_{\text{AC}}(\mathcal{G}) = \emptyset$) is known to be either $\beta(\mathcal{G}) = 1$ or $\beta(\mathcal{G}) = 2$ Maleki et al. [2014]; Blasiak et al. [2013]. Moreover, the broadcast rate is strictly larger than 2 if the problem is internally conflicted. In other words, a problem is not internally conflicted if and only if a symmetric rate $R = 1/\beta(\mathcal{G})$ of at least 0.5 can be achieved, and thus we also call such problem a *half-rate-feasible* problem. The internally conflicted problems with $\mathfrak{C}_{\text{AC}}(\mathcal{G}) \neq \emptyset$ are also referred to as *half-rate-infeasible* problems.

Proposition 2.6 (Internal conflict bound, Maleki et al. [2014]; Jafar [2014])**.** Consider the internally conflicted (i.e., half-rate-infeasible) CIC problem $\mathcal{G}$: $(i|j \in A_i)$, $i \in [n]$. The capacity region $\mathscr{C}(\mathcal{G})$ is outer bounded by the rate region $\mathscr{R}_{\text{AC}}(\mathcal{G})$ that consists of all rate tuples $\mathbf{R}$ such that for every alignment chain $Ch_{\text{AC}} \in \mathfrak{C}_{\text{AC}}(\mathcal{G})$ as shown in (2.31),

$$\sum_{j \in [m+1]} R_{i(j)} + \sum_{j \in [m]} R_{k(j)} \leq m. \tag{2.32}$$

In particular, the broadcast rate $\beta(\mathcal{G})$ is lower bounded by $\beta_{\text{AC}}(\mathcal{G})$ as

$$\beta(\mathcal{G}) \geq \beta_{\text{AC}}(\mathcal{G}) \doteq \max_{m \in \mathbb{Z}^+\colon \exists Ch_{\text{AC}} \in \mathfrak{C}_{\text{AC}}(\mathcal{G}) \text{ of length } m} \frac{1+2m}{m}. \tag{2.33}$$

The MAIS bound in Proposition 2.5 and the internal conflict bound in Proposition 2.6 can outperform each other for some CIC problem instances.

Example 2.5. Consider the 5-message CIC problem $\mathcal{G}$ in Example 2.4, for which we have the following alignment chain of length $m = 2$,

$$Ch_{\text{AC}}: \quad \underline{1} \xleftrightarrow{4} 2 \xleftrightarrow{5} \underline{3}. \tag{2.34}$$

According to the internal conflict bound in Proposition 2.6, we have

$$\beta(\mathcal{G}) \geq \beta_{\mathrm{AC}}(\mathcal{G}) \geq \frac{1+2*2}{2} = 2.5.$$

This upper bound is tight as it coincides with the FLPCC bound Arbabjolfaei and Kim [2014] (Proposition 2.2 in Section 2.3.1) on the broadcast rate. More specifically, the optimal linear code given by the FLPCC scheme can be written as

$$y = (x_{1,1} \oplus x_{2,1}, x_{2,2} \oplus x_{3,1}, x_{3,2} \oplus x_{4,1}, x_{4,2} \oplus x_{5,1}, x_{5,2} \oplus x_{1,2}),$$

where the codeword length $r = 5$ and message length $t_i = 2$ for every $i \in [5]$. With such linear code, every receiver can decode 2 bits of its wanted message in every 5 channel uses, and thus $\beta(\mathcal{G}) = 5/2 = 2.5$ is achieved. Recall that the MAIS bound gives a loose result as $\beta_{\mathrm{MAIS}}(\mathcal{G}) = 2 < 2.5$ for this problem.

Remark 2.4. One may notice that the MAIS bound is loose for any CIC problem whose broadcast rate is not an integer as $\beta_{\mathrm{MAIS}}(\mathcal{G})$ is always an integer. As for the internal conflict bound, by (2.33) one can see that $\beta_{\mathrm{AC}}(\mathcal{G})$ is always no larger than $\frac{1+2m}{m}|_{m=1} = 3$. Hence for those CIC problems whose broadcast rate is larger than 3, the internal conflict bound is always loose. For a simple example, consider the CIC problem $\mathcal{G}$: $(1|-), (2|-), (3|-), (4|-)$ where there are 4 receivers and no one knows anything as side information. Quite obviously for this problem $\beta(\mathcal{G}) = \beta_{\mathrm{MAIS}}(\mathcal{G}) = 4$. However, one can verify that the internal conflict bound gives only $\beta_{\mathrm{AC}}(\mathcal{G}) = 3$.

In the following we review the more general PM bound Blasiak et al. [2011], which implies both the MAIS bound and the internal conflict bound.

Proposition 2.7 (Polymatroidal (PM) bound, Blasiak et al. [2011])**.** Consider the CIC problem $\mathcal{G}$: $(i|j \in A_i), i \in [n]$. The capacity region $\mathscr{C}(\mathcal{G})$ is outer bounded by the rate region $\mathscr{R}_{\mathrm{PM}}(\mathcal{G})$ that consists of all rate tuples $\mathbf{R}$ satisfying

$$R_i \leq g(B_i \cup \{i\}) - g(B_i), \qquad \forall i \in [n], \tag{2.35}$$

for at least one set function $g : 2^{[n]} \to [0,1]$ such that

$$g(\emptyset) = 0, \tag{2.36}$$

$$g([n]) \leq 1, \tag{2.37}$$

$$g(K) \leq g(K'), \qquad \text{if } K \subseteq K', \tag{2.38}$$

$$g(K \cap K') + g(K \cup K') \leq g(K) + g(K'), \tag{2.39}$$

$$g(B_i \cup \{i\}) - g(B_i) = g(\{i\}), \qquad \forall i \in [n]. \tag{2.40}$$

In particular, the broadcast rate $\beta(\mathcal{G})$ is lower bounded by $\beta_{\text{PM}}(\mathcal{G})$ as

$$\beta(\mathcal{G}) \geq \beta_{\text{PM}}(\mathcal{G}) \doteq \min_{g\,:\,2^{[n]} \to [0,1] \text{ satisfying (2.36)-(2.40)}} \max_{i\in[n]} \frac{1}{g(B_i \cup \{i\}) - g(B_i)} \tag{2.41}$$

Remark 2.5. Properties (2.36), (2.38), and (2.39) capture standard PM axioms of the entropy function. Property (2.37) captures the unit channel capacity constraint. The inequality (2.35) and property (2.40) are based on the system assumptions and conditions such as message independence and decoding requirement at receivers. For more details, also see Appendix B.1.1.

It has been shown in Liu et al. [2018a] that the PM bound in Proposition 2.7 is the tightest bound one can get using *all* Shannon-type inequalities. The PM bound is tight on the capacity region for all CIC problems with less than or equal to 5 messages Arbabjolfaei et al. [2013], as well as many larger problems. The smallest CIC problem known so far, for which non-Shannon-type inequalities are strictly needed for capacity region characterization, is of $n = 9$ messages [Liu and Sadeghi, 2019b, Section IV-C]. Several other larger problems for which non-Shannon-type inequalities are necessary have been identified in Sun and Jafar [2015]; Baber et al. [2013].

2.3.4 Existing Performance Bounds for the DIC problem

Now we move on to the existing performance bounds for the DIC problem. To the best of our knowledge, the only two information-theoretical performance bounds in the literature holding generally for any DIC problem are those proposed in Sadeghi et al. [2016], which are presented in the following.

Proposition 2.8 (Distributed maximal acyclic induced subgraph (MAIS) bound, Sadeghi et al. [2016])**.** Consider the DIC problem $(i|j \in A_i)$, $i \in [n]$ with link capacity tuple $\mathbf{C}$. The capacity region $\mathscr{C}(\mathcal{G}, \mathbf{C})$ is outer bounded by the rate region $\mathscr{R}_{\text{DMAIS}}(\mathcal{G}, \mathbf{C})$ that consists of all rate tuples $\mathbf{R}$ satisfying

$$\sum_{i\,\in\, S\colon\, S \subseteq [n] \text{ is acyclic}} R_i \leq \sum_{J\in T_S} C_J, \tag{2.42}$$

where $T_S = \{J \in N : J \cap S \neq \emptyset\}$ as defined in Definition 2.2. In particular, if any symmetric rate R is achievable, then

$$R \leq \min_{S \subseteq [n] \text{ is acyclic}} \frac{1}{|S|} \sum_{J\in T_S} C_J. \tag{2.43}$$

Proposition 2.9 (All-server distributed polymatroidal (PM) bound, Sadeghi et al. [2016])**.** Consider the DIC problem $\mathcal{G}$: $(i|j \in A_i)$, $i \in [n]$ with link capacity tuple $\mathbf{C}$. The capacity region $\mathscr{C}(\mathcal{G}, \mathbf{C})$ is outer bounded by the rate region $\mathscr{R}_{\text{ADPM}}(\mathcal{G}, \mathbf{C})$ that consists of all rate tuples $\mathbf{R}$ such that for any $L \subseteq [n]$,

$$R_i \leq f_L((B_i \cup \{i\}) \cap L) - f_L(B_i \cap L), \qquad \forall i \in L, \tag{2.44}$$

for at least one set function $f_L : 2^L \to [0,1]$ such that

$$f_L(\emptyset) = 0, \tag{2.45}$$

$$f_L(L) \leq \sum_{J \in T_L} C_J, \tag{2.46}$$

$$f_L(K) \leq f_L(K'), \qquad \text{if } K \subseteq K', \tag{2.47}$$

$$f_L(K \cap K') + f_L(K \cup K') \leq f_L(K) + f_L(K'). \tag{2.48}$$

In particular, if any symmetric rate R is achievable, then

$$R \leq \min_{L \subseteq [n]} \max_{f_L \,:\, 2^L \to [0,1] \text{ satisfying (2.45)-(2.48)}} \min_{i \in L} \big(f_L((B_i \cup \{i\}) \cap L) - f_L(B_i \cap L) \big). \tag{2.49}$$

The distributed MAIS bound in Proposition 2.8 can be seen as a generalized version of the MAIS bound for the CIC problem, and is implied by the all-server distributed PM bound in Proposition 2.9.

Proposition 2.9 is capable of giving tight results on the sum capacity C_{sum} for 145 of 218 non-isomorphic 4-message DIC problems with equal link capacities $C_J = 1, J \in N \setminus \{\emptyset\}$. Nevertheless, for all these 218 problems tight upper bounds on the sum capacity C_{sum} can indeed be established using Shannon-type inequalities. That is to say, unlike the PM bound in Proposition 2.7 which is the tightest bound for the CIC problem one can get using Shannon-type inequalities, the all-server distributed PM bound in Proposition 2.9 only captures a subset of Shannon-type inequalities for the DIC problem. In Chapter 4, we will propose a more general performance bound also based on the PM axioms of the entropy function for the DIC problem. Our bound strictly outperforms the all-server distributed PM bound in Proposition 2.9 and indeed captures all Shannon-type inequalities.

2.4 Chapter Summary

In this chapter, we formulated and rigorously defined the centralized index coding (CIC) and the distributed index coding (DIC) problems, as well as their capacity region and other related notions. We illustrated that the CIC problem is a special case of the more general DIC problem, and that the multi-server nature of the DIC problem can lead to fundamentally different properties compared with the CIC scenario. We provided mathematical preliminaries for the whole thesis. In particular, we reviewed and defined a number of graph-theoretic terms including the notions of a vertex set being acyclic and a set of vertex sets being set-level acyclic. We also reviewed the concept of lexicographic product of directed graphs, and defined the touch structure for the multi-server scenario. We reviewed a number of existing linear and non-linear coding schemes, as well as several performance bounds for the CIC and DIC problems. Our results to be presented in the next three chapters will be nontrivially built upon these existing works.

Chapter 3

Achievability and Coding Schemes

In this chapter, we investigate the coding schemes and corresponding achievability results for index coding. Our results are built upon the two-layer random coding scheme, composite coding (CC) Arbabjolfaei et al. [2013] (cf. Proposition 2.4), where the messages are mapped to composite indices and composite indices are then mapped to the codeword to be transmitted.

In Sections 3.1-3.3 we focus on the CIC problem. First in Section 3.1, we simplify the original CC scheme through a number of simplification techniques, leading to a great reduction in the computational complexity. We then improve the CC scheme in two independent directions. In Section 3.2, we improve the CC scheme by employing a more flexible enhanced fractional allocation of the centralized channel capacity. In Section 3.3, we extend the CC scheme by adding one more layer of random coding into it. That is, instead of directly mapping the composite indices to the codeword, we map them to a number of *doubly* composite indices, which are then mapped to the codeword for transmission. We name such coding scheme as *three-layer* composite coding (TLCC) and show that it can strictly outperform the CC scheme.

In Sections 3.4-3.5, we propose a distributed version of the CC scheme, namely *distributed composite coding* (DCC). The DCC scheme is built upon combining the idea of cooperative compression of composite indices across all servers introduced in Li et al. [2018] and our aforementioned enhanced fractional allocation of server link capacities, together with a new dimension of decoding flexibility: each receiver can choose their own group of server outputs for decoding. We also present the resulting achievable rate region of the DCC scheme in a series of equivalent or simplified forms that can help better understand the coding scheme or reduce the computational complexity.

Detailed examples are provided to showcase the use and efficacy of our results throughout the chapter. Some technical proofs are presented in Appendices A.1-A.2.

3.1 Simplified Composite Coding

The composite coding (CC) scheme is a two-layer coding scheme for the CIC problem, which is based on the classic idea of random coding by Claude Shannon. See Section 2.3.1 for a detailed description of the CC scheme and its corresponding achievability bound, the CC bound in Proposition 2.4.

One practical concern about the CC scheme is its computational complexity. As the problem size n grows, the number of composite indices $W_K, K \subseteq [n]$ grows exponentially, and the number of possible decoding configurations grows super exponentially. These two facts together lead to considerably high computational complexity.

In this section we propose several simplification techniques to address both complexity issues mentioned above. Removing composite indices is achieved by pairwise comparison of any two indices and removing one if its corresponding rate can be *transferred* without performance (tightness) loss to the other in the expressions of the achievable rate region. A heuristic method is also proposed for reducing the number of composite indices even further, but possibly with some performance loss. Decoding configurations are reduced by establishing a sufficient condition for a specific decoding configuration to be safely excluded without affecting the performance of the scheme.

We first show in Subsection 3.1.1 how the CC scheme can be simplified for a fixed decoding configuration via removing some unnecessary composite indices. Then we show in Subsection 3.1.2 how one can ignore some decoding configurations without affecting the achievable rate region.

3.1.1 Reducing Composite Indices for a Fixed Decoding Configuration

Consider the CIC problem $\mathcal{G}$: $(i|j \in A_i)$, $i \in [n]$. For easier reference, we repeat the CC bound in Proposition 2.4 as

$$\mathscr{R}_{\mathrm{CC}}(\mathcal{G}) = \mathrm{co}\Big(\bigcup_{\mathbf{D} \in \mathcal{D}(\mathcal{G})} \mathscr{R}_{\mathrm{CC}}(\mathcal{G}, \mathbf{D})\Big), \tag{3.1}$$

where for each $\mathbf{D} \in \mathcal{D}(\mathcal{G})$, $\mathscr{R}_{\mathrm{CC}}(\mathcal{G}, \mathbf{D})$ consists of all the rate tuples $\mathbf{R}$ satisfying

$$\sum_{K \subseteq [n], K \nsubseteq A_i} S_K(\mathbf{D}) < 1, \qquad \forall i \in [n], \tag{3.2}$$

$$\sum_{j \in L} R_j < \sum_{\substack{K \subseteq D_i \cup A_i, \\ K \cap L \neq \emptyset}} S_K(\mathbf{D}), \quad \forall L \subseteq D_i,\, i \in [n], \tag{3.3}$$

for some $S_K(\mathbf{D}) \geq 0$, $K \subseteq [n]$. Recall that the inequalities in (3.2) signify the first-step decoding constraints for the composite indices to be recovered for each receiver, and the inequalities in (3.3) signify the second-step decoding constraints for recovering messages in each receiver's decoding message set specified by the decoding configuration $\mathbf{D}$. When the context is clear, we simply call the inequalities in (3.2) and (3.3) the first-step inequalities and second-step inequalities, respectively.

The goal is to determine whether it is possible to remove $S_K(\mathbf{D})$ for some $K \subseteq [n]$ in (3.2) and (3.3) without affecting the rate region $\mathscr{R}_{\mathrm{CC}}(\mathcal{G}, \mathbf{D})$ for a given decoding configuration $\mathbf{D}$. Our proposed simplification involves pairwise comparison of any two composite index rates, say $S_K(\mathbf{D})$ and $S_{K'}(\mathbf{D})$, $K, K' \subseteq [n]$. Roughly speaking, if $S_K(\mathbf{D})$ appears in every first-step inequality in (3.2) in which $S_{K'}(\mathbf{D})$ appears, and at the same time, $S_K(\mathbf{D})$ does not appear in any second-step inequality in (3.3) in which $S_{K'}(\mathbf{D})$ does not appear, then $S_K(\mathbf{D})$ can be safely removed from computations.

More formally, fix an ordering for all inequalities identified in (3.2) and (3.3). Enumerate all inequalities in (3.2) by indices $\ell_1 \in [m_1]$ and all inequalities in (3.3) by indices $\ell_2 \in [m_2]$, where m_1, m_2 denote the number of first-step and second-step inequalities, respectively. Note that including possibly inactive inequalities, there is one first-step inequality and $2^{|D_i|} - 1$ second-step inequalities for each receiver. Now, assume $S_K(\mathbf{D})$ respectively appears in first-step and second-step inequalities that are identified by indices $Q_1(K) \subseteq [m_1]$ and $Q_2(K) \subseteq [m_2]$. Similarly, assume $S_{K'}(\mathbf{D})$ respectively appears in first-step and second-step inequalities identified by $Q_1(K') \subseteq [m_1]$ and $Q_2(K') \subseteq [m_2]$. We establish a sufficient condition for removing $S_K(\mathbf{D})$ without performance loss as follows.

Theorem 3.1 (Composite index removing condition)**.** If $Q_2(K) \subseteq Q_2(K')$ and $Q_1(K') \subseteq Q_1(K)$, then $S_K(\mathbf{D})$ can be removed from the inequalities in (3.2) and (3.3) without affecting the resulting rate region $\mathscr{R}_{\mathrm{CC}}(\mathcal{G}, \mathbf{D})$.

Proof. Assume that $S_K(\mathbf{D}) = a$ and $S_{K'}(\mathbf{D}) = b$ in the full set of expressions in (3.2) and (3.3). Since $Q_2(K) \subseteq Q_2(K')$, whenever $S_K(\mathbf{D})$ appears in any second-step inequality, so does $S_{K'}(\mathbf{D})$. Therefore, transferring the rate of $S_K(\mathbf{D})$ to $S_{K'}(\mathbf{D})$, that is setting $S_K(\mathbf{D}) = 0$ and $S_{K'}(\mathbf{D}) = a + b$, cannot decrease message rates. Since $Q_1(K') \subseteq Q_1(K)$, whenever $S_{K'}(\mathbf{D})$ appears in any first-step inequality, so does $S_K(\mathbf{D})$. Hence, transferring the rate of $S_K(\mathbf{D})$ to $S_{K'}(\mathbf{D})$ cannot result in an invalid composite index rate assignment in (3.2). Therefore, one can remove $S_K(\mathbf{D})$ from (3.2) and (3.3) without affecting $\mathscr{R}_{\mathrm{CC}}(\mathcal{G}, \mathbf{D})$. □

Note that for the CIC problem $\mathcal{G}$ with n messages, simplification via Theorem 3.1 requires no more than $\binom{2^n-1}{2}$ pairwise comparisons.

Example 3.1. Consider the 4-message CIC problem $(1|4), (2|3,4), (3|1,2), (4|2,3)$. Set $\mathbf{D}$ as $D_i = [n] \setminus A_i$, $i = 2,3,4$, and $D_1 = \{1\}$. For brevity, for any $S_K(\mathbf{D})$, $K \subseteq [n]$ we simply write S_K. We compare the relative presence of $S_{\{1,3\}}$ and $S_{\{1,2,3\}}$ in the decoding inequalities in (3.2) and (3.3). Since for any $i \in [n]$, $\{1,3\} \not\subseteq A_i$ and $\{1,2,3\} \not\subseteq A_i$, $S_{\{1,3\}}$ and $S_{\{1,2,3\}}$ are present in all first-step inequalities. Writing second-step decoding inequalities in (3.3) yields

$$\begin{aligned}
R_1 &< S_{\{1\}} + S_{\{1,4\}}, \\
R_1 &< S_{\{1\}} + S_{\{1,2\}} + S_{\{1,3\}} + S_{\{1,2,3\}} + S_{\{1,4\}} + S_{\{1,2,4\}} + S_{\{1,3,4\}} + S_{[n]}, \\
R_2 &\overset{(a)}{<} S_{\{2\}} + S_{\{1,2\}} + S_{\{2,3\}} + S_{\{1,2,3\}} + S_{\{2,4\}} + S_{\{1,2,4\}} + S_{\{2,3,4\}} + S_{[n]}, \\
R_1 + R_2 &< S_{\{1\}} + S_{\{2\}} + S_{\{1,2\}} + S_{\{1,3\}} + S_{\{2,3\}} \\
&\quad + S_{\{1,2,3\}} + S_{\{1,4\}} + S_{\{2,4\}} + S_{\{1,2,4\}} + S_{\{1,3,4\}} + S_{\{2,3,4\}} + S_{[n]}, \\
R_3 &< S_{\{3\}} + S_{\{1,3\}} + S_{\{2,3\}} + S_{\{1,2,3\}} + S_{\{3,4\}} + S_{\{1,3,4\}} + S_{\{2,3,4\}} + S_{[n]}, \\
R_4 &< S_{\{4\}} + S_{\{1,4\}} + S_{\{2,4\}} + S_{\{1,2,4\}} + S_{\{3,4\}} + S_{\{1,3,4\}} + S_{\{2,3,4\}} + S_{[n]}, \\
R_3 + R_4 &< S_{\{3\}} + S_{\{1,3\}} + S_{\{2,3\}} + S_{\{1,2,3\}} + S_{\{4\}} \\
&\quad + S_{\{1,4\}} + S_{\{2,4\}} + S_{\{1,2,4\}} + S_{\{3,4\}} + S_{\{1,3,4\}} + S_{\{2,3,4\}} + S_{[n]}, \\
R_1 + R_4 &< S_{\{1\}} + S_{\{1,2\}} + S_{\{1,3\}} + S_{\{1,2,3\}} + S_{\{4\}} \\
&\quad + S_{\{1,4\}} + S_{\{2,4\}} + S_{\{1,2,4\}} + S_{\{3,4\}} + S_{\{1,3,4\}} + S_{\{2,3,4\}} + S_{[n]},
\end{aligned}$$

Now we observe that $S_{\{1,2,3\}}$ is present in one more second-step decoding inequality compared with $S_{\{1,3\}}$ (in the inequality (a) above, $S_{\{1,2,3\}}$ is present, but $S_{\{1,3\}}$ is not). Hence, $S_{\{1,3\}}$ can be removed without affecting the achievable rate performance $\mathscr{R}_{\text{CC}}(\mathcal{G}, \mathbf{D})$. Continuing this procedure for all distinct $K, K' \subseteq [n]$, the only remaining composite index rates are S_K, $K \in \{\{1\}, \{2\}, \{1,2\}, \{3\}, \{2,3\}, \{4\}, \{1,4\}, \{3,4\}, \{1,2,3,4\}\}$, reducing the number of rate variables from 15 to 9. Note that in general further reductions may be possible if we first remove inactive inequalities from the inequalities given by (3.2) and (3.3), and then apply Theorem 3.1. For the problem in this example, the second inequality of first-step decoding (for $i = 2$) is inactive. If this inequality is removed first, one can further remove $S_{\{3,4\}}$ relative to $S_{[n]}$ using Theorem 3.1.

In the following we present a heuristic algorithm that can result in further reductions in the number of composite index rates, albeit with possible performance loss. It is based upon Theorem 3.1 and allows reduction of some composite index rates even when some of the conditions are violated in a controlled manner.

Specifically, to remove $S_K(\mathbf{D})$ given the decoding configuration $\mathbf{D}$, we still require that $Q_2(K) \subseteq Q_2(K')$, which will ensure that the corresponding composite index $w_{K'}(\mathbf{D})$ is at least as useful as $w_K(\mathbf{D})$ in the second-step decoding for all receivers. Now, note that $Q_1(K') \subseteq Q_1(K)$ in Theorem 3.1 implies that there does not exist any receiver for which $S_{K'}(\mathbf{D})$ appears in their first-step decoding inequalities while $S_K(\mathbf{D})$ does not. We define the subset of receivers who know all messages in x_K, but not all messages in $x_{K'}$ as follows,

$$M(K,K') = \{j \in [n] : K \subseteq A_j, K' \nsubseteq A_j\}. \tag{3.4}$$

The condition $Q_1(K') \subseteq Q_1(K)$ of Theorem 3.1 implies $M(K,K')$ is empty. In the heuristic, we allow $S_K(\mathbf{D})$ to be removed in comparison to $S_{K'}(\mathbf{D})$ even when $M(K,K') \neq \emptyset$, provided that $w_{K'}(\mathbf{D})$ is useful in the second-step decoding for all receivers in $M(K,K')$. Intuitively, although $w_{K'}(\mathbf{D})$ has to be decoded in the first-step decoding at those receivers, it is not an interference and is useful in their second-step decoding.

Simplification via Algorithm 1 can be implemented in such a way that no more than $\binom{2^n-1}{2}$ pairwise comparisons (same as Theorem 3.1) are required for any n-message CIC problem. The algorithm is shown below.

```
Algorithm 1: Heuristic composite index rate reduction
Input  : Two composite index rates $S_K(\mathbf{D})$ and $S_{K'}(\mathbf{D})$ that appear respectively in
         (3.2) and (3.3) indexed by $Q_1(K), Q_2(K)$ and by $Q_1(K'), Q_2(K')$. Receiver
         subset $M(K,K')$ in (3.4).
Output: A flag indicating whether $S_K(\mathbf{D})$ can be removed from the inequalities in
         (3.2) and (3.3).
If Theorem 3.1 holds (i.e., $Q_2(K) \subseteq Q_2(K')$ and $M(K,K') = \emptyset$), then flag = TRUE.
Else if $Q_2(K) \subseteq Q_2(K')$, $M(K,K') \neq \emptyset$, and $K' \subseteq D_j \cup A_j, \forall j \in M(K,K')$, then
  flag = TRUE.
Else flag = FALSE.
```

Applying Algorithm 1 to the problem in Example 3.1 with the same $\mathbf{D}$, all composite index rates, but $S_{\{1,4\}}(\mathbf{D})$ and $S_{\{1,2,3,4\}}(\mathbf{D})$ can be eliminated. The achievable rate region $\mathscr{R}_{\mathrm{CC}}(\mathcal{G},\mathbf{D})$ using only $S_{\{1,4\}}(\mathbf{D})$ and $S_{\{1,2,3,4\}}(\mathbf{D})$ coincides with the MAIS bound $\beta_{\mathrm{MAIS}}(\mathcal{G})$, which establishes the capacity region and confirms efficacy of Algorithm 1 for this problem.

3.1.2 Reducing Decoding Configurations

The main idea behind ruling out some decoding configurations is as follows. Given the CIC problem $\mathcal{G} : (i|j \in A_i)$, $i \in [n]$ and a decoding configuration $\mathbf{D}$, if $A_j \subseteq A_i \cup D_i$ for some $j \neq i$, then receiver i already knows or will know more than receiver j does. Therefore, receiver i can mimic the behavior of receiver j to decode whichever messages receiver j decodes without leading to any extra constraints on the achievable rate region. That is, one can update $D_i \leftarrow D_i \cup D_j$ at no cost to the achievable rate region. Based on such idea, we can establish the following sufficient condition for excluding some decoding configurations without performance loss.

Theorem 3.2 (Decoding configuration removing condition)**.** Let $\mathbf{D}$ be a decoding configuration satisfying the condition that there exists some receivers $i, j \in [n]$, $A_j \subseteq A_i \cup D_i$, $D_j \setminus (A_i \cup D_i) \neq \emptyset$. Then $\mathbf{D}$ can be removed from (3.1) without affecting the resulting rate region $\mathscr{R}_{\mathrm{CC}}(\mathcal{G})$.

To prove the above theorem, we first show the following lemma.

Lemma 3.1. Let $\mathbf{D}$ be a decoding configuration satisfying the condition in Theorem 3.2 (i.e., there exists some receivers $i, j \in [n]$, $A_j \subseteq A_i \cup D_i$, $D_j \setminus (A_i \cup D_i) \neq \emptyset$). Then there always exists some $\mathbf{D}' = (D'_k, k \in [n])$ such that $\mathscr{R}_{\mathrm{CC}}(\mathcal{G}, \mathbf{D}) \subseteq \mathscr{R}_{\mathrm{CC}}(\mathcal{G}, \mathbf{D}')$.

Proof. Set $\mathbf{D}' = (D'_k, k \in [n])$ as

$$D'_k = \begin{cases} D_k, & k \neq i, \\ D_i \cup (D_j \setminus (A_i \cup D_i)) = D_i \cup D_j \setminus A_i, & k = i. \end{cases} \tag{3.5}$$

Consider an arbitrary achievable rate tuple $\mathbf{R} \in \mathscr{R}_{\mathrm{CC}}(\mathcal{G}, \mathbf{D})$. There exists some $(S_K, K \subseteq [n])$ satisfying (3.2) and (3.3) under the decoding configuration $\mathbf{D}$. That is, we have

$$\sum_{K \subseteq [n], K \not\subseteq A_k} S_K < 1, \qquad \forall k \in [n], \tag{3.6}$$

$$\sum_{\ell \in L} R_\ell < \sum_{K \subseteq D_k \cup A_k, K \cap L \neq \emptyset} S_K, \quad \forall L \subseteq D_k, \, k \in [n]. \tag{3.7}$$

In the following we show that $\mathbf{R} \in \mathscr{R}_{\mathrm{CC}}(\mathcal{G}, \mathbf{D}')$ by proving that $\mathbf{R}$ and $(S_K, K \subseteq [n])$ satisfy (3.2) and (3.3) under the decoding configuration $\mathbf{D}'$.

Note that the first-step inequalities in (3.2) do not depend on the decoding configuration. Also, since $D'_k = D_k$ for any $k \in \{i\}^c$, the second-step inequalities in (3.3) for receivers other than i are the same for $\mathbf{D}$ and $\mathbf{D}'$. Therefore, it suffices to show that for receiver i,

$$\sum_{\ell \in L} R_\ell < \sum_{K \subseteq D'_i \cup A_i, K \cap L \neq \emptyset} S_K, \quad \forall L \subseteq D'_i. \tag{3.8}$$

Consider any $L \subseteq D'_i$. Note that D'_i as constructed in (3.5) can be partitioned into two disjoint parts D_i and $D_j \setminus (A_i \cup D_i)$. Partition L as $L = L_1 \cup L_2$ where $L_1 \subseteq D_i$ and $L_2 \subseteq D_j \setminus (A_i \cup D_i)$. Hence, $L_1 \cap L_2 = \emptyset$, $L_2 \cap (A_i \cup D_i) = \emptyset$, and $L_2 \subseteq D_j$.

Given $L_1 \subseteq D_i$, $L_2 \subseteq D_j$, by (3.7), we have

$$\sum_{\ell \in L_1} R_\ell < \sum_{K \subseteq D_i \cup A_i, K \cap L_1 \neq \emptyset} S_K, \tag{3.9}$$

$$\sum_{\ell \in L_2} R_\ell < \sum_{K \subseteq D_j \cup A_j, K \cap L_2 \neq \emptyset} S_K. \tag{3.10}$$

Therefore, we have

$$\begin{aligned}
\sum_{\ell\in L} R_\ell = \sum_{\ell\in L_1} R_\ell + \sum_{\ell\in L_2} R_\ell &< \sum_{\substack{K\subseteq D_i\cup A_i,\\ K\cap L_1\neq\emptyset}} S_K + \sum_{\substack{K\subseteq D_j\cup A_j,\\ K\cap L_2\neq\emptyset}} S_K && (3.11)\\
&\leq \sum_{\substack{K\subseteq D_i\cup A_i,\\ K\cap L_1\neq\emptyset}} S_K + \sum_{\substack{K\subseteq D_j\cup(A_i\cup D_i),\\ K\cap L_2\neq\emptyset}} S_K && (3.12)\\
&= \sum_{\substack{K\subseteq D_i\cup A_i,\\ K\cap L_1\neq\emptyset}} S_K + \sum_{\substack{K\subseteq D_j\cup(A_i\cup D_i),\\ K\nsubseteq A_i\cup D_i,\\ K\cap L_2\neq\emptyset}} S_K && (3.13)\\
&\leq \sum_{\substack{K\subseteq D_i\cup A_i,\\ K\cap L\neq\emptyset}} S_K + \sum_{\substack{K\subseteq D_j\cup(A_i\cup D_i),\\ K\nsubseteq A_i\cup D_i,\\ K\cap L\neq\emptyset}} S_K && (3.14)\\
&= \sum_{K\subseteq D_i'\cup A_i, K\cap L\neq\emptyset} S_K, && (3.15)
\end{aligned}$$

where (3.11) is due to (3.9) and (3.10), (3.12) is due to the fact that $A_j \subseteq A_i \cup D_i$, (3.13) is due to the fact that $L_2 \cap (A_i \cup D_i) = \emptyset$ and thus any $K \cap L_2 \neq \emptyset$ must not be a subset of $A_i \cup D_i$, (3.14) simply follows from the fact that $L_1 \subseteq L$, $L_2 \subseteq L$, and (3.15) is due to the fact that $D_i' = D_i \cup D_j \setminus A_i$ as constructed in (3.5) and thus $D_i' \cup A_i = D_j \cup (A_i \cup D_i)$. This concludes the proof of (3.8).

So far we have shown that for any $\mathbf{R} \in \mathscr{R}_{\mathrm{CC}}(\mathcal{G}, \mathbf{D})$, we have $\mathbf{R} \in \mathscr{R}_{\mathrm{CC}}(\mathcal{G}, \mathbf{D}')$. Therefore, $\mathscr{R}_{\mathrm{CC}}(\mathcal{G}, \mathbf{D}) \subseteq \mathscr{R}_{\mathrm{CC}}(\mathcal{G}, \mathbf{D}')$. □

With the help of Lemma 3.1, Theorem 3.2 can be shown as follows.

Proof. We prove Theorem 3.2 by showing that for any $\mathbf{D}$ satisfying the condition in Theorem 3.2 (i.e., there exists some receivers $i, j \in [n]$, $A_j \subseteq A_i \cup D_i'$, $D_j' \setminus (A_i \cup D_i') \neq \emptyset$), there exists some $\mathbf{D}'$, which does not satisfy that condition, such that $\mathscr{R}_{\mathrm{CC}}(\mathcal{G}, \mathbf{D}) \subseteq \mathscr{R}_{\mathrm{CC}}(\mathcal{G}, \mathbf{D}')$.

Given any $\mathbf{D}$ satisfying the condition in Theorem 3.2, we can generate $\mathbf{D}'$ according to (3.5) and by Lemma 3.1 we have $\mathscr{R}_{\mathrm{CC}}(\mathcal{G}, \mathbf{D}) \subseteq \mathscr{R}_{\mathrm{CC}}(\mathcal{G}, \mathbf{D}')$. If $\mathbf{D}'$ still satisfies that condition, then we treat $\mathbf{D}'$ as $\mathbf{D}$ and generate a new $\mathbf{D}'$ based on this new $\mathbf{D}$ according to (3.5) again. Repeating such procedure until we reach some $\mathbf{D}'$ for which there exists no two receivers $i, j \in [n]$ such that $A_j \subseteq A_i \cup D_i'$, $D_j' \setminus (A_i \cup D_i') \neq \emptyset$. It remains to show that such terminating point $\mathbf{D}'$ always exists.

Notice that every time we generate a new $\mathbf{D}' = (D_k', k \in [n])$ from a $\mathbf{D} = (D_i, i \in [n])$ according to (3.5), the sum of the size of the decoding message set for each receiver is increased by at least 1. That is, $\sum_{k\in[n]} |D_k'| - \sum_{i\in[n]} |D_i| \geq 1$. Recall the minimal decoding message set for each receiver $i \in [n]$ is the singleton message set $\{i\}$, and the maximal decoding message set for each receiver $i \in [n]$ is the complement of its side information, A_i^c. Also notice that the maximal decoding configuration $\mathbf{D}^* = (D_i^* = A_i^c, i \in [n])$ itself does not satisfy the condition in Theorem 3.2. Therefore, by repeating the aforementioned procedure at most $\sum_{i\in[n]} |A_i^c| - \sum_{i\in[n]} |\{i\}| = \sum_{i\in[n]} |B_i|$ times, we can always reach some $\mathbf{D}'$ which does not satisfy the condition in Theorem 3.2. □

Example 3.2. Consider the 5-message CIC problem $(1|5), (2|4,5), (3|1,4), (4|1,2,3), (5|1,3)$. For this problem, there exists in total $|\mathcal{D}(\mathcal{G})| = 2^{10} = 1024$ decoding configurations. However, after excluding all $\mathbf{D}$ satisfying the condition in Theorem 3.2, only the following 5 decoding configurations remain:

$$\mathbf{D}_1 : D_1 = \{1\}, D_2 = \{1,2,3\}, D_3 = \{2,3,5\}, D_4 = \{4,5\}, D_5 = \{5\};$$
$$\mathbf{D}_2 : D_1 = \{1,3\}, D_2 = \{1,2,3\}, D_3 = \{2,3,5\}, D_4 = \{4,5\}, D_5 = \{5\};$$
$$\mathbf{D}_3 : D_1 = \{1\}, D_2 = \{1,2,3\}, D_3 = \{2,3,5\}, D_4 = \{4,5\}, D_5 = \{5\};$$
$$\mathbf{D}_4 : D_1 = \{1,2\}, D_2 = \{1,2,3\}, D_3 = \{2,3,5\}, D_4 = \{4,5\}, D_5 = \{5\};$$
$$\mathbf{D}_5 : D_1 = \{1,2,3,4\}, D_2 = \{1,2,3\}, D_3 = \{2,3,5\}, D_4 = \{4,5\}, D_5 = \{5\}.$$

One can verify that for any decoding configuration $\mathbf{D}_i$, $i = 1,2,3,4,5$ presented above, there exists no two receivers i, j such that $A_j \subseteq A_i \cup D_i$, $D_j \setminus (A_i \cup D_i) \neq \emptyset$.

Despite the usefulness of the decoding configuration removing condition in Theorem 3.2, sometimes even checking whether this condition holds or not for every $\mathbf{D} \in \mathcal{D}(\mathcal{G})$ is rather computationally expensive due to the considerably large number of all possible decoding configurations in $\mathcal{D}(\mathcal{G})$. Consider the CIC problem $\mathcal{G}$: $(i|j \in A_i)$, $i \in [n]$. The number of decoding configurations to consider for this problem grows on average super-exponentially in n as

$$|\mathcal{D}(\mathcal{G})| = |\prod_{i\in[n]} \mathcal{D}_i(\mathcal{G})| = 2^{n^2 - \sum_{i\in[n]} |A_i| - n} = 2^{\sum_{i\in[n]} |B_i|}. \tag{3.16}$$

As a partial solution to such issue, we iteratively build a decoding configuration in Algorithm 2, namely the *natural* decoding configuration, denoted by $\underline{\mathbf{D}}$. The natural decoding configuration $\underline{\mathbf{D}}$ serves as a baseline such that it suffices to consider only the decoding configurations that are element-wisely no smaller than it.[1]

Excluding any $\mathbf{D}$ that is not element-wisely no smaller than $\underline{\mathbf{D}}$ from (3.1) does not affect the achievable rate region $\mathscr{R}_{\mathrm{CC}}(\mathcal{G})$, as illustrated by the following proposition.

Proposition 3.1. Consider the CIC problem $\mathcal{G}$: $(i|j \in A_i)$, $i \in [n]$ with natural decoding configuration $\underline{\mathbf{D}} = (\underline{D}_i, i \in [n])$ given by Algorithm 2. Let $\mathbf{D} = (D_i, i \in [n])$ be a decoding

[1] The notion of natural decoding configuration has also been presented in another PhD thesis Arbabjolfaei [2017], as this result is an outcome of collaboration Liu et al. [2017]. We have chosen to present Algorithm 2 and Proposition 3.1 in this chapter because it helps with application of our original Theorem 3.2.

Algorithm 2: Natural decoding configuration

```
Input  : CIC problem $\mathcal{G}$: $(i|j \in A_i), i \in [n]$.
Output: Natural decoding configuration $\underline{\mathbf{D}} = (\underline{D}_i, i \in [n])$.
Initialize $\underline{D}_i = \{i\}, i \in [n]$.
For as long as there exists $i, j \in [n]$ such that $A_j \subseteq A_i \cup \underline{D}_i$, $\underline{D}_j \setminus (A_i \cup \underline{D}_i) \neq \emptyset$,
  update $\underline{D}_i \leftarrow \underline{D}_i \cup (\underline{D}_j \setminus (A_i \cup \underline{D}_i))$. If no such $i, j$ exist, terminate the algorithm.
```

configuration such that $\underline{D}_k \nsubseteq D_k$ for some $k \in [n]$. Then there exists another decoding configuration $\mathbf{D}' = (D'_i, i \in [n])$, for which $\underline{D}_i \subseteq D'_i$ for all $i \in [n]$, such that $\mathscr{R}_{\text{CC}}(\mathcal{G}, \mathbf{D}) \subseteq \mathscr{R}_{\text{CC}}(\mathcal{G}, \mathbf{D}')$.

Proposition 3.1 can be proved using similar techniques to Theorem 3.2.

For any CIC problem with a large number of decoding configurations, our proposed approach is to first compute the natural decoding configuration $\underline{\mathbf{D}}$ according to Algorithm 2, and then apply Theorem 3.2 to the set of decoding configurations that are element-wisely no smaller than $\underline{\mathbf{D}}$. See the following example.

Example 3.3. Consider the 6-message CIC problem

$$(1|3,4), (2|4,6), (3|5,6), (4|2,6), (5|1,2,6), (6|1,2,3).$$

For this problem, there exists in total $|\mathcal{D}(\mathcal{G})| = 2^{16} = 65536$. The natural decoding configuration $\underline{\mathbf{D}}$ can be computed via Algorithm 2 as $\underline{D}_i = \{i\}$, $i = 1,2,3,4$, $\underline{D}_5 = \{3,4,5\}$, and $\underline{D}_6 = \{4,5,6\}$. Among the 65536 decoding configurations, there are only $2^{12} = 4096$ of them that are element-wisely no smaller than $\underline{\mathbf{D}}$. Hence according to Proposition 3.1, it suffices to consider only these 4096 decoding configurations. Applying Theorem 3.2 to these 4096 decoding configurations, only 18 remain. That is, it suffices to consider only these 18 remaining decoding configurations to compute the achievable rate region $\mathscr{R}_{\text{CC}}(\mathcal{G})$ of the CC scheme for this problem $\mathcal{G}$.

3.1.3 Numerical Results for the Simplified Composite Coding

To show the efficacy of our proposed simplifications, we first consider all 218 and 9,608 non-isomorphic CIC problems with $n = 4$ and 5 messages, respectively. For each problem, we only use the natural decoding configuration $\underline{\mathbf{D}}$ obtained through Algorithm 2. Hence for each problem, the reduction in the number of decoding configurations is $(100 - \frac{100}{|\mathcal{D}(\mathcal{G})|})\%$, where $|\mathcal{D}(\mathcal{G})|$ was given in (3.16). We then apply Algorithm 1 to all possible rate pairs S_K and $S_{K'}$ in Proposition 3.7. If for a given problem, Algorithm 1 only retains m out of $2^n - 1$ such rates, the reduction rate for that problem is $(100 - \frac{100m}{2^n-1})\%$. Finally, we use FME to eliminate the only m remaining $S_{K'}$ variables and compute the rate region of Proposition 2.4. For comparison with the obtained achievable rate region for each problem, we compute corresponding outer bound on the capacity region using the polymatroidal (PM) bound in Proposition 2.7 introduced from Blasiak et al. [2011]. The tests for all the CIC problems with $n = 4$ and 5 messages were run on an Apple iMac 4GHz Intel Core i7 with 16 GB memory and using Matlab® R2017b and the FME software Gattegno et al. [2015].

According to the test results, the natural decoding configuration $\underline{\mathbf{D}}$ together with Algorithm 1 are sufficient to match the PM outer bound and thus establish the capacity region for all problems with $n = 4$ and 5 messages. Table 4.1 indicates the average computational savings. Although not strictly needed, the average reduction using the more conservative Theorem 3.1 is shown for comparison. As shown in the last column, it takes on average 5.6 and 11.63 times longer to eliminate all $(2^n - 1)$ S_K variables in the non-reduced problems with $n = 4$ and 5 messages, respectively, as compared to eliminating only the few rates retained by Algorithm 1. Finally, the distribution of composite rate reduction among problems is quite good. For $n = 4$, only in 7 of 218 problems a maximum of $m = 3$ rates were retained, where for all other problems, $m \leq 2$. For $n = 5$, in only 31 of 9,608 problems, $m = 7$ to at most $m = 10$ composite rates were used.

Table 3.1: Results for all CIC problems with $n = 4$ and 5 messages.

Problem Size	Ave. dec. config. reduction	Ave. S_K reduction by Alg. 1	Ave. S_K reduction by Thm. 3.1	$\frac{\text{Not-reduced time}}{\text{Alg. 1 time}}$
$n = 4$	96.13%	89.6%	63.06%	560%
$n = 5$	99.64%	92.36%	71.42%	1163%

Next, we consider the 259 β-critical CIC problems with $n = 6$ messages as described in [Arbabjolfaei and Kim, 2018, Section 8.6]. Note that among all 1,540,944 non-isomorphic 6-message CIC problems, any problem $\mathcal{G}$ that is not β-critical can be transformed to some β-critical problem $\mathcal{G}'$ by removing certain receiver side information, such that the transformed problem shares the same broadcast rate as the original problem (i.e., $\beta(\mathcal{G}) = \beta(\mathcal{G}')$). For these 259 β-critical problems, lower bounds on β are computed using the PM bound for comparison with the achievability results. Note that the PM lower bound on β is tight for any 6-message CIC problem [Arbabjolfaei and Kim, 2018, Section 8.6].

As shown by Table 3.2, for a vast majority of problems (230 out of 259), the natural decoding configuration $\underline{\mathbf{D}}$ given by Algorithm 2 together with the heuristic composite index reduction method in Algorithm 1 are sufficient to obtain β. In 2 problems, $\underline{\mathbf{D}}$ is still sufficient yet Theorem 3.1 was required. For the remaining 27 problems, $\underline{\mathbf{D}}$ is not sufficient to achieve the broadcast rate β. Instead, we consider all the decoding configurations that do not satisfy the condition specified in Theorem 3.2. The decoding configuration reduction for each problem $\mathcal{G}$

is thus computed as

$$(100 - \frac{100 \times |\{\mathbf{D} \in \mathcal{D}(\mathcal{G}) : \mathbf{D} \text{ does not satisfy the condition in Thm. 3.2}\}|}{|\mathcal{D}(\mathcal{G})|})\%.$$

According to the numerical results, for these 27 problems, the composite index reduction methods in Algorithm 1 and Theorem 3.1 make no difference in terms of the tightness of the resulting upper bounds on β. Therefore, Algorithm 1 is sufficient. For 14 out of 27 problems, the composite coding upper bound matches the PM lower bound, thus establishing β. For the other 13 problems, we are not able to achieve β using composite coding.

Table 3.2: Results for 259 β-critical CIC problems with $n = 6$ messages.

Number of Problems	Ave. dec. config. reduction	Ave. S_K reduction by Alg. 1	Ave. S_K reduction by Thm. 3.1	How β can be obtained
230	99.55%	92.32%	76.66%	$\underline{\mathbf{D}}$ & Alg.1
2	99.99%	90.48%	81.75%	$\underline{\mathbf{D}}$ & Thm. 3.1
14	99.93%	91.59%	75.43%	Thm. 3.2 & Alg.1
13	-	-	-	Unable to achieve β using composite coding

Remark 3.1. We list the 13 β-critical 6-message CIC problems for which composite coding cannot achieve the broadcast rate. For these 13 problems, the PM lower bound β_{PM} matches the upper bound given by the minrank approach Bar-Yossef et al. [2011], thus establishing the broadcast rate. These 13 problems are listed below, where each line represents one problem:

$$(1|4,5,6),(2|4,5),(3|4,6),(4|1,2,3),(5|1,3),(6|1,2);$$
$$(1|4,5,6),(2|4,6),(3|5,6),(4|1,3),(5|1,2),(6|2,4);$$
$$(1|4,5,6),(2|4,6),(3|5,6),(4|1,3),(5|1,2),(6|1,4,5);$$
$$(1|3,4),(2|4,6),(3|5,6),(4|2,6),(5|1,2),(6|1,2,3);$$
$$(1|4,6),(2|4,5),(3|1,5),(4|1,6),(5|1,3),(6|2,3);$$
$$(1|3,4),(2|4,6),(3|5,6),(4|2,6),(5|1,2),(6|1,4,5);$$
$$(1|4,5,6),(2|5,6),(3|1,5),(4|1,2),(5|2,6),(6|3,4);$$
$$(1|3,4),(2|4,5),(3|5,6),(4|1,5,6),(5|1,4,6),(6|1,2);$$
$$(1|3,4),(2|4,5),(3|5,6),(4|2,3,6),(5|1,4,6),(6|1,2);$$
$$(1|3,4,5),(2|3,4),(3|5,6),(4|1,6),(5|1,2),(6|1,2,3);$$
$$(1|3,4,6),(2|3,4,5),(3|2,6),(4|1,2),(5|1,3),(6|4,5);$$
$$(1|3,5),(2|3,4),(3|1,5),(4|1,6),(5|2,6),(6|1,2,3);$$
$$(1|3,5),(2|4,5),(3|2,6),(4|1,6),(5|3,4,6),(6|1,2,5).$$

3.2 Enhanced Composite Coding

In this section, we extend the original CC scheme for the CIC problem by employing a more flexible enhanced fractional allocation of the centralized channel capacity.

Consider the CIC problem $\mathcal{G}$: $(i|j \in A_i)$, $i \in [n]$. Splitting the rate of each message, we represent the message x_i by independent parts $x_i(\mathbf{D})$ at rate $R_i(\mathbf{D})$, where each sub-message is associated with a decoding configuration $\mathbf{D} \in \mathcal{D}(\mathcal{G})$. Thus,

$$R_i = \sum_{\mathbf{D} \in \mathcal{D}(\mathcal{G})} R_i(\mathbf{D}), \qquad \forall i \in [n]. \tag{3.17}$$

To send $(x_i(\mathbf{D}), i \in [n])$, the server generates composite indices $w_K(\mathbf{D})$ at composite rate $S_K(\mathbf{D})$ for any $K \subseteq [n]$ via random coding. Then it encodes all the composite indices $(w_K(\mathbf{D}), K \subseteq [n], \mathbf{D} \in \mathcal{D}(\mathcal{G}))$ together to a codeword y again via random coding. Upon receiving y, each receiver first recovers all composite indices $(w_K(\mathbf{D}), K \subseteq [n], \mathbf{D} \in \mathcal{D}(\mathcal{G}))$. This is successful with vanishing probability of error if

$$\sum_{\mathbf{D} \in \mathcal{D}(\mathcal{G})} \sum_{K \subseteq [n], K \not\subseteq A_i} S_K(\mathbf{D}) < 1, \qquad \forall i \in [n]. \tag{3.18}$$

For a given $\mathbf{D} \in \mathcal{D}(\mathcal{G})$, receiver i can successfully recover messages $x_{D_i}(\mathbf{D})$, with vanishing probability of error if

$$\sum_{j \in L} R_j(\mathbf{D}) < \sum_{K \subseteq D_i \cup A_i, K \cap L \neq \emptyset} S_K(\mathbf{D}), \qquad \forall L \subseteq D_i. \tag{3.19}$$

We refer to the above coding scheme as the *enhanced* composite coding (ECC) scheme, whose achievable rate region $\mathscr{R}_{\text{ECC}}(\mathcal{G})$ is summarized by the following theorem.

Theorem 3.3 (Enhanced composite coding (ECC) bound)**.** Consider the CIC problem $\mathcal{G}$: $(i|j \in A_i)$, $i \in [n]$. The capacity region $\mathscr{C}(\mathcal{G})$ is inner bounded by the rate region $\mathscr{R}_{\text{ECC}}(\mathcal{G})$ that consists of all rate tuple $\mathbf{R}$ such that

$$R_i = \sum_{\mathbf{D} \in \mathcal{D}(\mathcal{G})} R_i(\mathbf{D}), \qquad \forall i \in [n], \tag{3.20}$$

for some $R_i(\mathbf{D}) \geq 0$, $i \in [n]$, $\mathbf{D} \in \mathcal{D}(\mathcal{G})$ and some $S_K(\mathbf{D}) \geq 0$, $K \subseteq [n]$, $\mathbf{D} \in \mathcal{D}(\mathcal{G})$

satisfying

$$\sum_{\mathbf{D}\in\mathcal{D}(\mathcal{G})}\ \sum_{K\subseteq[n],K\not\subseteq A_i} S_K(\mathbf{D}) < 1, \qquad \forall i\in[n], \tag{3.21}$$

$$\sum_{j\in L} R_j(\mathbf{D}) < \sum_{\substack{K\subseteq D_i\cup A_i,\\ K\cap L\neq\emptyset}} S_K(\mathbf{D}), \qquad \forall L\subseteq D_i, i\in[n], \mathbf{D}\in\mathcal{D}(\mathcal{G}). \tag{3.22}$$

It can be verified that the ECC achievable rate region $\mathscr{R}_{\mathrm{ECC}}(\mathcal{G})$ is convex. One can compute $\mathscr{R}_{\mathrm{ECC}}(\mathcal{G})$ by projecting out $(S_K(\mathbf{D}), K\subseteq[n], \mathbf{D}\in\mathcal{D}(\mathcal{G}))$ and $(R_i(\mathbf{D}), i\in[n], \mathbf{D}\in\mathcal{D}(\mathcal{G}))$ in (3.20)-(3.22) through certain optimization tools such as Fourier-Motzkin elimination (FME). Note that linear programing (LP) can be used to solve for a desired weighted sum-rate subject to (3.20)-(3.22) if one is not interested in the whole rate region, which results in a lower computational complexity.

The ECC bound in Theorem 3.3 is always no looser than the original CC bound in Proposition 2.4 due to a larger degrees of freedom in choosing the composite index rates for different decoding choices $\mathbf{D}$. That is, for any CIC problem $\mathcal{G}$,

$$\mathscr{R}_{\mathrm{CC}}(\mathcal{G}) \subseteq \mathscr{R}_{\mathrm{ECC}}(\mathcal{G}). \tag{3.23}$$

Such relationship can be strict, as illustrated by the following example.

Example 3.4. Consider the 6-message CIC problem $\mathcal{G}$:

$$(1|3,4), (2|4,5), (3|5,6), (4|2,3,6), (5|1,4,6), (6|1,2).$$

The largest symmetric rate in $\mathscr{R}_{\mathrm{CC}}(\mathcal{G})$ is $8/27$, while the largest symmetric rate in $\mathscr{R}_{\mathrm{ECC}}(\mathcal{G})$ is $23/77$, and thus $\mathscr{R}_{\mathrm{CC}}(\mathcal{G}) \subset \mathscr{R}_{\mathrm{ECC}}(\mathcal{G})$. Note that a strictly larger symmetric rate of $R = 1/3$ can be achieved by the following scalar linear coding scheme (where $r = 1$ and $t_i = 1$ for every $i \in [n]$):

$$y = (x_1\oplus x_3\oplus x_4, x_2\oplus x_4\oplus x_5, x_1\oplus x_2\oplus x_6),$$

which matches the MAIS bound in Proposition 2.5, thus establishing the symmetric capacity $C_{\mathrm{sym}}(\mathcal{G})$ to be $1/3$.

The simplification techniques presented in Section 3.1 can be readily extended to the ECC scheme, but will not be detailed here.

3.3 Three-Layer Composite Coding

In this section we take a different approach to extend the CC scheme. Recall that the flat coding Arbabjolfaei et al. [2013] (cf. Proposition 2.3) is a single-layer random coding scheme while the CC scheme is also based on the random coding technique, but is two-layer. A natural question that arises is whether one can benefit from adding more layers of random coding into the CC scheme. Towards answering this question, in this section we design a three-layer composite coding scheme or the TLCC scheme by adding one more layer of random coding into the CC scheme.

3.3.1 The Three-Layer Composite Coding Scheme and Achievability Bound

We describe the TLCC scheme in a detailed manner, present its corresponding achievable rate region, and then show that it can strictly outperform the CC scheme through a concrete example.

Codebook generation.

Step 1. For each $K \subseteq [n]$ and each realization of messages x_K, generate a composite index $w_K = w_K(x_K)$ drawn uniformly at random from $[2^{s_K}]$, where $s_K = \lceil rS_K \rceil$ and S_K is the rate of the composite index w_K. That is, the composite index w_K is generated according to the random mapping w_K as

$$w_K : \prod_{i \in K} [2^{t_i}] \to [2^{s_K}].$$

Step 2. For each $M \subseteq N$ and each realization of composite indices $(w_K, K \in M)$, generate a *doubly* composite index $v_M = v_M(w_K, K \in M)$ drawn uniformly at random from $[2^{z_M}]$, where $z_M = \lceil rZ_M \rceil$ and Z_M is the rate for the doubly composite index v_M. That is, the doubly composite index v_M is generated according to the random mapping V_M as

$$v_M : \prod_{K \in M} [2^{s_K}] \to [2^{z_M}].$$

Step 3. For each realization of the doubly composite index tuple $(v_M, M \subseteq N)$, generate a codeword $y = y((v_M, M \subseteq N))$ drawn uniformly at random from $[2^r]$. That is, the codeword y is generated according to the random mapping y as

$$y : \prod_{M \subseteq N} [2^{z_M}] \to [2^r].$$

The codebook is revealed to all parties.

Encoding. To communicate messages $x_{[n]}$, the server first encodes the composite index $w_K = w_K(x_K)$ for each $K \subseteq [n]$, then encodes the doubly composite index $v_M = v_M(w_K, K \in$

$M)$ for each $M \subseteq N$, and finally encodes and transmits the codeword $y = y((v_M, M \subseteq N))$.

Decoding. Decoding takes place in the reverse order of stages.

Step 1. Receiver i finds the unique doubly composite index tuple $(\hat{v}_M, M \subseteq N)$ such that $y = y((\hat{v}_M, M \subseteq N))$. If there is more than one such tuple, it declares an error.

Step 2. Assuming that $(\hat{v}_M, M \subseteq N)$ is correct, receiver i decodes for a subset of composite indices indexed by $P_i \subseteq N \setminus 2^{A_i}$. That is, receiver i finds the unique composite index tuple $(\hat{w}_K, K \in P_i)$ such that $\hat{v}_M = v_M(\hat{w}_K, K \in M)$ for every $M \subseteq P_i \cup 2^{A_i}$. If there is more than one such tuple, it declares an error.

Step 3. Assuming that $(\hat{w}_K, K \in P_i)$ is correct, receiver i decodes for a subset of messages indexed by $D_i \subseteq [n] \setminus A_i$. That is, receiver i finds the unique message tuple $\hat{x}_{D_i}$ such that $\hat{w}_K = w_K(\hat{x}_K)$ for every $K \subseteq D_i \cup A_i$, $K \in P_i$. If there is more than one such tuple, it declares an error.

Remark 3.2. Recall that for the CC scheme, each receiver i can choose a subset of messages to decode. The tuple of decoding message sets is denoted by $\mathbf{D} = (D_i, i \in [n])$ and referred to as the decoding configuration for the CC scheme. For the TLCC scheme, after successfully decoding all the doubly composite indices, each receiver i can choose both the subset of composite indices and the subset of messages to decode. Hence, the decoding configuration for the TLCC scheme is the 2-tuple $(\mathbf{P}, \mathbf{D})$, where $\mathbf{P} \doteq (P_i, i \in [n])$ denotes the tuple of decoding composite index sets. Let $\mathcal{P}_i(\mathcal{G}) \doteq \{P_i \subseteq N \setminus 2^{A_i} : i \in \cup_{K \in P_i} K\}$ denote the set of all possible decoding composite index set at receiver i, and thus $\mathcal{P}(\mathcal{G}) \doteq \prod_{i \in [n]} \mathcal{P}_i(\mathcal{G})$ denotes the collection of all possible tuples of decoding composite index sets for the CIC problem $\mathcal{G}$.

Remark 3.3. Note that in the third step of decoding, receiver i can only possibly recover messages in set $\cup_{K \in P_i} K$, because only composite indices $(\hat{w}_K, K \in P_i) = (w_K, K \in P_i)$ are available at receiver i as the result of the second step of decoding. Hence, without loss of generality, we always require that $D_i \subseteq \cup_{K \in P_i} K$ and do not consider any $(\mathbf{P}, \mathbf{D})$ violating this condition.

We summarize the achievable rate region of the TLCC scheme below. Time sharing over all possible decoding configurations is applied. Note that we can also apply the more flexible *enhanced* fractional allocation of the broadcast channel capacity as in the ECC scheme, the details of which are not given here.

Theorem 3.4 (Three-layer composite coding (TLCC) bound)**.** Consider the CIC problem $\mathcal{G}$: $(i|j \in A_i), i \in [n]$. The capacity region $\mathscr{C}(\mathcal{G})$ is inner bounded by the rate region $\mathscr{R}_{\text{TLCC}}(\mathcal{G})$ as

$$\mathscr{R}_{\text{TLCC}}(\mathcal{G}) \doteq \text{co}\big(\bigcup_{\mathbf{P} \in \mathcal{P}(\mathcal{G}), \mathbf{D} \in \mathcal{D}(\mathcal{G})} \mathscr{R}_{\text{TLCC}}(\mathcal{G}, \mathbf{P}, \mathbf{D}) \big), \tag{3.24}$$

where $\mathrm{co}(\cdot)$ denotes the convex hull, and $\mathscr{R}_{\mathrm{TLCC}}(\mathcal{G}, \mathbf{P}, \mathbf{D})$ consists of all the rate tuples $\mathbf{R}$ satisfying

$$\sum_{M \subseteq N, M \not\subseteq 2^{A_i}} Z_M(\mathbf{P}, \mathbf{D}) < 1, \qquad \forall i \in [n], \tag{3.25}$$

$$\sum_{K \in Q} S_K(\mathbf{P}, \mathbf{D}) < \sum_{\substack{M \subseteq P_i \cup 2^{A_i}, \\ M \cap Q \neq \emptyset}} Z_M(\mathbf{P}, \mathbf{D}), \qquad \forall Q \subseteq P_i, i \in [n], \tag{3.26}$$

$$\sum_{j \in L} R_j < \sum_{\substack{K \subseteq D_i \cup A_i, \\ K \cap L \neq \emptyset, \\ K \in P_i}} S_K(\mathbf{P}, \mathbf{D}), \qquad \forall L \subseteq D_i,\ i \in [n], \tag{3.27}$$

for some $Z_M(\mathbf{P}, \mathbf{D}) \geq 0, M \subseteq N$ and $S_K(\mathbf{P}, \mathbf{D}) \geq 0, K \subseteq [n]$.

Proof. We only present the error analysis for the second step decoding of the TLCC scheme, as the analysis for other steps is quite similar to the error analysis for the CC scheme (which can be found in Arbabjolfaei et al. [2014], [Arbabjolfaei and Kim, 2018, Section 6.10]).

Towards this end, assume that the doubly composite indices $(\hat{v}_M, M \subseteq N)$ have been correctly decoded. For receiver i, we partition the error event according to the collection $Q \subseteq P_i$ for erroneous composite indices. That is, $\hat{w}_K \neq w_K$ iff $K \in Q$. Hence, for the second step decoding error probability P_e, we have

$$\begin{aligned} P_e =& P\{\hat{v}_M = v_M(\hat{w}_K, K \in M) \text{ for all } M \subseteq 2^{A_i} \cup P_i \text{ for some } \hat{w}_K \neq w_K, K \in P_i\} \\ \leq& \sum_{Q \subseteq P_i} \sum_{\substack{(\hat{w}_K, K \in P_i): \\ \hat{w}_K \neq w_K, K \in Q \\ \hat{w}_K = w_K, K \notin Q}} P\left\{ \bigcap_{\substack{M \subseteq P_i \cup 2^{A_i} \\ M \cap Q \neq \emptyset}} \{\hat{v}_M = v_M(\hat{w}_K, K \in M)\} \right\} \\ <& \sum_{Q \subseteq P_i} 2^{\sum_{K \in Q} s_K} / 2^{\sum_{M \subseteq P \cup 2^{A_i}, M \cap Q \neq \emptyset} z_M} \\ <& \sum_{Q \subseteq P_i} 2^{r(\sum_{K \in Q} S_K - \sum_{M \subseteq P \cup 2^{A_i}, M \cap Q \neq \emptyset} Z_M) + |Q|}, \end{aligned} \tag{3.28}$$

where the first inequality is due to the union bound, the second inequality holds since for each Q, the number of erroneous tuples is $\prod_{K \in Q}(2^{s_K} - 1) < 2^{\sum_{K \in Q} s_K}$, and the probability for any two distinct composite index tuples being mapped to the same doubly composite index v_M for all $M \subseteq P \cup 2^{A_i}, M \cap Q \neq \emptyset$ is $2^{-\sum_{M \subseteq P \cup 2^{A_i}, M \cap Q \neq \emptyset} z_M}$, and the last inequality simply follows from that $s_K = \lceil r S_K \rceil < r S_K + 1$ and $z_M = \lceil r Z_M \rceil \geq r Z_M$.

Given (3.28), one can see that the second step decoding error probability P_e tends to zero as $r \to \infty$ if (3.26) is satisfied. □

The inequalities in (3.25), (3.26), and (3.27) signify the first-step, second-step, and third-step decoding constraints for the TLCC scheme.

When there is no ambiguity, we simply refer to the doubly composite index rate variables and the composite index rate variables in (3.25)-(3.27) as Z_M or Z variables and S_K or S variables, respectively.

The TLCC scheme subsumes the CC scheme in general. A rigorous proof is presented in Section 3.3.2.3. For some problems, it can give strictly better results as demonstrated below.

Example 3.5. Consider the 7-message CIC problem $\mathcal{G}$:

$$(1|5),(2|3,5,6),(3|4,6,7),(4|1,2,7),(5|2,3,4,7),(6|3,4,7),(7|1,2,4).$$

The ECC bound in Theorem 3.3 indicates that any symmetric rate $R < \frac{1}{3.25}$ is achievable. Nevertheless, a strictly better result $R < \frac{1}{3}$ can be obtained according to the TLCC scheme in Theorem 3.4 as follows. Set $(\mathbf{P},\mathbf{D})$ such that $D_i = \{i\}, \forall i \in [n]$, and

$$\begin{aligned} &P_1 = \{\{1\}\}, P_2 = \{\{2\},\{4,7\}\}, P_5 = \{\{5\}\},\\ &P_3 = P_6 = \{\{3,6\}\}, P_4 = P_7 = \{\{5\},\{4,7\}\}. \end{aligned}$$

Then set $Z_M = 0, \forall M \subseteq N$ except for $Z_{\{\{1\},\{5\}\}}$, $Z_{\{\{2\},\{4,7\},\{5\}\}}$ and $Z_{\{\{3,6\},\{4,7\}\}}$. Also set $S_K = 0, \forall K \subseteq [n]$ except for $S_{\{1\}}, S_{\{2\}}, S_{\{5\}}, S_{\{3,6\}}$, and $S_{\{4,7\}}$. Writing all the active decoding inequalities of Theorem 3.4 yields

$$Z_{\{\{1\},\{5\}\}} + Z_{\{\{2\},\{4,7\},\{5\}\}} + Z_{\{\{3,6\},\{4,7\}\}} < 1,$$

and

$$\begin{aligned} S_{\{1\}} &< Z_{\{\{1\},\{5\}\}},\\ S_{\{2\}} + S_{\{4,7\}} &< Z_{\{\{2\},\{4,7\},\{5\}\}} + Z_{\{\{3,6\},\{4,7\}\}},\\ S_{\{2\}} &< Z_{\{\{2\},\{4,7\},\{5\}\}},\\ S_{\{3,6\}} &< Z_{\{\{3,6\},\{4,7\}\}},\\ S_{\{5\}} + S_{\{4,7\}} &< Z_{\{\{1\},\{5\}\}} + Z_{\{\{2\},\{4,7\},\{5\}\}},\\ S_{\{4,7\}} &< Z_{\{\{2\},\{4,7\},\{5\}\}},\\ S_{\{5\}} &< Z_{\{\{2\},\{4,7\},\{5\}\}}, \end{aligned}$$

and

$$R_{\mathrm{sym}} < S_K, \qquad \forall K \in \{\{1\},\{2\},\{5\},\{3,6\},\{4,7\}\}.$$

For an arbitrary $\epsilon \in (0, \frac{1}{3}]$, assigning $R = \frac{1}{3} - \epsilon$, $S_{\{1\}} = S_{\{2\}} = S_{\{5\}} = S_{\{3,6\}} = S_{\{4,7\}} =$

$\frac{1}{3} - \frac{\epsilon}{2}$, $Z_{\{\{1\},\{5\}\}} = Z_{\{\{2\},\{4,7\},\{5\}\}} = Z_{\{\{3,6\},\{4,7\}\}} = \frac{1}{3} - \frac{\epsilon}{4}$ satisfies all the inequalities above. Hence any symmetric rate $R < \frac{1}{3}$ is achievable by the TLCC scheme, which matches the MAIS bound in Proposition 2.5, thus establishing the symmetric capacity $C_{\text{sym}}(\mathcal{G})$ to be $1/3$. Note that the optimal symmetric rate $R = 1/3$ can also be achieved by the following scalar linear coding scheme:

$$y = (x_1 \oplus x_5, x_1 \oplus x_2 \oplus x_3 \oplus x_6, x_1 \oplus x_2 \oplus x_4 \oplus x_7).$$

3.3.2 Simplifications for the Three-Layer Composite Coding

The main challenges for the TLCC's computation are the overwhelming number of Z_M variables, which grows doubly exponentially with n, and the choice of decoding configuration $(\mathbf{P}, \mathbf{D})$. To circumvent these, we now present a series of simplifications.

3.3.2.1 Limiting the Choice of Decoding Configuration

First, it can be shown that Theorem 3.2 and Proposition 3.1 also apply to the TLCC scheme. That is, we have

$$\begin{aligned} \mathscr{R}_{\text{TLCC}}(\mathcal{G}) &= \text{co}\Big(\bigcup_{\mathbf{P} \in \mathcal{P}(\mathcal{G}), \mathbf{D} \in \mathcal{D}(\mathcal{G})} \mathscr{R}_{\text{TLCC}}(\mathcal{G}, \mathbf{P}, \mathbf{D}) \Big) \\ &= \text{co}\Big(\bigcup_{\mathbf{P} \in \mathcal{P}(\mathcal{G}), \mathbf{D} \in \Delta(\mathcal{G})} \mathscr{R}_{\text{TLCC}}(\mathcal{G}, \mathbf{P}, \mathbf{D}) \Big) \qquad (3.29) \\ &= \text{co}\Big(\bigcup_{\mathbf{P} \in \mathcal{P}(\mathcal{G}), \mathbf{D} \in \Delta_{\text{natural}}(\mathcal{G})} \mathscr{R}_{\text{TLCC}}(\mathcal{G}, \mathbf{P}, \mathbf{D}) \Big), \qquad (3.30) \end{aligned}$$

where $\Delta(\mathcal{G})$ denotes the collection of $\mathbf{D}$ for which the condition in Theorem 3.2 is not satisfied (i.e., there exists no two receivers $i, j \in [n]$ such that $A_j \subseteq A_i \cup D_i$, $D_j \setminus (A_i \cup D_i) \neq \emptyset$), and $\Delta_{\text{natural}}(\mathcal{G})$ denotes the collection of $\mathbf{D}$ that is element-wisely no smaller than the natural decoding message set tuple $\underline{\mathbf{D}}$ as generated in Algorithm 2.

In the following we propose a *heuristic* baseline decoding composite index set tuple $\underline{\mathbf{P}}^{\mathbf{D}}$ for a given $\mathbf{D}$. Note that removing some $\mathbf{P}$ that is not element-wisely no smaller than $\underline{\mathbf{P}}^{\mathbf{D}}$ may affect the performance of the TLCC scheme. The idea is as follows. For a given $\mathbf{D}$, the starting set $P_i^{\mathbf{D}}$ for receiver i contains all subsets $K \subseteq [n]$ that intersects with D_i and does not intersect with the interfering messages $B_i = (A_i \cup D_i)^c$. Then, following similar lines of thought as in Algorithm 2, we iteratively add missing elements from $P_j^{\mathbf{D}}$ to $P_i^{\mathbf{D}}$ if $2^{A_j} \subseteq 2^{A_i} \cup P_i^{\mathbf{D}}$ and $P_j^{\mathbf{D}} \nsubseteq 2^{A_i} \cup P_i^{\mathbf{D}}$. This is summarized in Algorithm 3, which makes use of the touch structure defined in Definition 2.2.

Algorithm 3: Heuristic baseline decoding index set.

```
Input  : CIC problem $\mathcal{G}$: $(i|j \in A_i), i \in [n]$, and decoding message set tuple
         $\mathbf{D} = (D_i, i \in [n])$
Output: Heuristic baseline decoding composite index set tuple $\mathbf{P^D} = (P_i^{\mathbf{D}}, i \in [n])$.
Initialize $P_i^{\mathbf{D}} = T_{D_i,\overline{(A_i \cup D_i)^c}}, i \in [n]$.
If there exists $i, j \in [n]$ such that $2^{A_j} \subseteq 2^{A_i} \cup P_i^{\mathbf{D}}$ and $P_j^{\mathbf{D}} \nsubseteq 2^{A_i} \cup P_i^{\mathbf{D}}$, update
  $P_i^{\mathbf{D}} \leftarrow P_i^{\mathbf{D}} \cup (P_j^{\mathbf{D}} \setminus 2^{A_i})$. If no such $i, j$ exist, terminate the algorithm.
```

One may only consider the supersets of $\mathbf{P^D}$ computed by Algorithm 3 in the second-step decoding of the TLCC scheme. This can lead to great reduction in the number of possible decoding configurations, albeit with possible performance loss.

For any collection of composite indices $M \subseteq N$, we set

$$\Gamma_*(M) = \bigcup_{K \in M} \{L \in N : L \subseteq K\} = \bigcup_{K \in M} 2^K \tag{3.31}$$

to be the subset completion of M. Note $2^K = \Gamma_*(\{K\})$. The set $M \subseteq N$ is subset complete if $M = \Gamma_*(M)$.

For any $\mathbf{P}$, if $P_i \cup 2^{A_i}$ is subset complete for any $i \in [n]$, we simply say that $\mathbf{P}$ is subset complete. Then we have the following lemma.

Lemma 3.2. For any $\mathbf{D}$, its corresponding $\mathbf{P^D}$ given by Algorithm 3 is subset complete.

Proof. For any $i \in [n]$, consider the initial set $P_i^{\mathbf{D}} = T_{D_i,\overline{(A_i \cup D_i)^c}}$. As $2^{A_i} \cup P_i^{\mathbf{D}} = 2^{A_i \cup D_i}$, we have $\Gamma_*(2^{A_i} \cup P_i^{\mathbf{D}}) = \Gamma_*(2^{A_i \cup D_i}) = 2^{A_i \cup D_i} = 2^{A_i} \cup P_i^{\mathbf{D}}$. Since the union of any two subset complete sets is also subset complete, for the final $P_i^{\mathbf{D}}$, $2^{A_i} \cup P_i^{\mathbf{D}}$ must be subset complete as well. □

For a given $\mathbf{D}$, despite possible performance loss, we may further narrow down the range of $\mathbf{P}$ to consider by enforcing that $\mathbf{P}$ is element-wisely no smaller than $\mathbf{P^D}$ and is subset complete.

3.3.2.2 Reducing Doubly Composite Indices

In the following theorem, we adopt and modify the composite index rate removal techniques from Section 3.1.1 to exclude some doubly composite index rates $Z_M, M \subseteq N$.

Theorem 3.5 (Doubly composite index removing condition)**.** For a given decoding configuration and arbitrary $M \neq M' \subseteq N$, compare the relative presence for $Z_M, Z_{M'}$ in the inequalities identified by Theorem 3.4.

1. If $Z_{M'}$ appears in any first-step decoding inequality in (3.25) then so does Z_M, AND

2. If Z_M appears in any second-step decoding inequality in (3.26) then so does $Z_{M'}$,

then Z_M can be removed from the rate expressions without affecting the resulting rate region.

The proof is similar to that of Theorem 3.1 and thus omitted.

The number of Z_M variables remaining after applying Theorem 3.5 to every M, M' pairs can be much smaller than its original amount, leading to considerably lower computational complexity. However, as the original number of Z_M variables is extremely large for large n, even applying Theorem 3.5 for every possible M, M' can be computationally unaffordable. Hence, we propose to systematically exclude some Z_M variables even before applying Theorem 3.5. This can be done if $\mathbf{P}$ is subset-complete, such as $\mathbf{P}^{\mathbf{D}}$ at the output of Algorithm 3.

Corollary 3.1. For any decoding configuration $(\mathbf{P}, \mathbf{D})$ such that $\mathbf{P}$ is subset complete, it suffices to only consider Z_M such that M is subset complete.

Proof. Consider an arbitrary $M \subseteq N$ such that $M \neq \Gamma_*(M)$, set $M' = \Gamma_*(M)$. Then M' is subset complete and $M \subset M'$. Since $\bigcup_{K \in M} K = \bigcup_{K \in M'} K$, whenever Z_M appears in a first-step decoding inequality in (3.25), so does $Z_{M'}$ and vice versa. Consider the relative presence for $Z_M, Z_{M'}$ in the second-stage inequalities. For any $i \in [n]$, since $2^{A_i} \cup P_i$ is subset-complete, if $M \subseteq 2^{A_i} \cup P_i$, we must have $M' \subseteq 2^{A_i} \cup P_i$. Also, as $M \subset M'$, if $M \cap L \neq \emptyset$ for any $L \subseteq D_i$, we must also have $M' \cap L \neq \emptyset$. Therefore, whenever Z_M appears in a second-step decoding inequality in (3.26) so does $Z_{M'}$. According to Theorem 3.5 any such Z_M can be removed. □

For the CIC problem $\mathcal{G}$ and decoding message set tuple $\mathbf{D}$, use $N_{\mathcal{K}}(\mathcal{G}, \mathbf{D})$ and $N'_{\mathcal{K}}(\mathcal{G}, \mathbf{D})$ to denote the collection of K such that S_K remains after applying Theorem 3.1 and Algorithm 1, respectively. When the context is clear, we simply use the shorthand notation $N_{\mathcal{K}}$ and $N'_{\mathcal{K}}$. Note that $N'_{\mathcal{K}} \subseteq N_{\mathcal{K}} \subseteq N$.

Example 3.6. Consider the 7-message problem in Example 3.5, setting $\mathbf{D}$ to be the natural decoding configuration $\underline{\mathbf{D}}$ and applying Theorem 3.1 and Algorithm 1, we find $|N_{\mathcal{K}}| = 16$ and $|N'_{\mathcal{K}}| = 7$, respectively. Note that the original number of S_K variables (excluding the dummy $S_\emptyset$) is $2^7 - 1 = 127$.

Remark 3.4. For the TLCC scheme, one can remove any composite index w_K, $K \notin N_{\mathcal{K}}$ or $K \notin N'_{\mathcal{K}}$ from the coding scheme. Since doubly composite indices v_M are generated from composite indices w_K, this will naturally narrow down the range of v_M from $M \subseteq N$ to $M \subseteq N_{\mathcal{K}}$ or $M \subseteq N'_{\mathcal{K}}$, which can lead to a huge reduction in the number of Z_M variables from $2^{|N|} - 1 = 2^{2^n - 1} - 1$ to $2^{|N_{\mathcal{K}}|} - 1$ or $2^{|N'_{\mathcal{K}}|} - 1$. Note that such reduction may lead to performance loss.

The simplification techniques for the TLCC scheme discussed so far are summarized in Table 3.3.

Table 3.3: Simplification techniques and whether they retain optimality

Simplification	References	Optimality Preservation
Limiting $\mathbf{D}$ such that $\mathbf{D} \in \Delta(\mathcal{G})$ or $\mathbf{D} \in \Delta_{\text{natural}}(\mathcal{G})$	(3.29) and (3.30)	Yes
Limiting $\mathbf{P}$ to be supersets of $\mathbf{P}^{\mathbf{D}}$ such that $\mathbf{P}$ is subset complete	Alg. 3, Lem. 3.2	Unknown
Removing Z variables by pairwise comparison	Thm. 3.5	Yes
Removing $Z_M, M \neq \Gamma_*(M)$ when $\mathbf{P}$ is subset complete	Cor. 3.1	Yes
Removing S_K and Z_M not fully embedded in $N_{\mathcal{K}}$ or $N'_{\mathcal{K}}$	Rmk. 3.4	Unknown

3.3.2.3 Simplified TLCC Subsumes Composite Coding

In this section, we prove that even with some of the simplifications discussed so far, the TLCC scheme is guaranteed to perform at least as well as the CC scheme.

Consider the CIC problem $\mathcal{G}$: $(i|j \in A_i), i \in [n]$ and an arbitrary $\mathbf{D} \in \mathcal{D}(\mathcal{G})$.

Within this section we slightly abuse the notation as follows. For any $K \in N_{\mathcal{K}}$ and $M \subseteq N_{\mathcal{K}}$, let 2^K denote the collection of subsets of K with respect to $N_{\mathcal{K}}$, $2^K = \{L \in N_{\mathcal{K}} : L \subseteq K\}$, and let $\Gamma_*(M)$ denote the subset completion of M with respect to $N_{\mathcal{K}}$, $\Gamma_*(M) = \bigcup_{K \in M}\{L \in N_{\mathcal{K}} : L \subseteq K\} = \bigcup_{K \in M} 2^K$. Also, set $T_{K,\overline{L}} = \{J \in N_{\mathcal{K}} : J \cap K \neq \emptyset, J \cap L = \emptyset\}$, and hence the heuristic baseline decoding index set $\mathbf{P}^{\mathbf{D}} = (P_i^{\mathbf{D}}, i \in [n])$ given by Algorithm 3 is within the range of $N_{\mathcal{K}}$. Let $\mathscr{R}_{\text{TLCC}}^{\text{simplified}}(\mathcal{G}, \mathbf{P}, \mathbf{D})$ denote the achievable rate region of the TLCC with decoding configuration $(\mathbf{P}, \mathbf{D})$ with the following simplifications:

1. any $S_K, K \notin N_{\mathcal{K}}$ and $Z_M, M \nsubseteq N_{\mathcal{K}}$ are removed from the decoding inequalities (3.25)-(3.27), and

2. $\mathbf{P} = (P_i, i \in [n])$ is a decoding composite index set such that for any $i \in [n]$, $P_i^{\mathbf{D}} \subseteq P_i \subseteq N_{\mathcal{K}} \setminus 2^{A_i}$ and that $2^{A_i} \cup P_i$ is subset complete with respect to $N_{\mathcal{K}}$.

Theorem 3.6. $\mathscr{R}_{\text{CC}}(\mathcal{G}, \mathbf{D}) \subseteq \mathscr{R}_{\text{TLCC}}^{\text{simplified}}(\mathcal{G}, \mathbf{P}, \mathbf{D})$.

Proof. Let $\mathbf{R} \in \mathscr{R}_{\text{CC}}(\mathcal{G}, \mathbf{D})$ be an achievable rate tuple of the CC scheme. According to Theorem 3.1, removing any S_K that $K \notin N_{\mathcal{K}}$ does not affect the achievable rate region of the

CC, hence there exists some $(S_K, K \in N_{\mathcal{K}})$ such that $\mathbf{R}$ and $(S_K, K \in N_{\mathcal{K}})$ satisfy (3.2) and (3.3) with $\mathbf{D}$. Set $Z_M, M \subseteq N_{\mathcal{K}}$ as

$$Z_M = \begin{cases} S_K, & \text{if } M = 2^K \text{ for some } K \in N_{\mathcal{K}}, \\ 0, & \text{otherwise.} \end{cases}$$

Now we show that $\mathbf{R}$, $(S_K, K \in N_{\mathcal{K}})$ and $(Z_M, M \subseteq N_{\mathcal{K}})$ satisfy the decoding inequalities of the TLCC scheme, (3.25)-(3.27), with $(\mathbf{P}, \mathbf{D})$ defined above.

First, for any $i \in [n]$, we have

$$\sum_{M \subseteq N_{\mathcal{K}} : M \not\subseteq 2^{A_i}} Z_M = \sum_{2^K \subseteq N_{\mathcal{K}} : 2^K \not\subseteq 2^{A_i}} Z_{2^K} = \sum_{K \in N_{\mathcal{K}} : K \not\subseteq A_i} S_K < 1.$$

Second, for any $Q \subseteq P_i, i \in [n]$, as $2^{A_i} \cup P_i$ is subset complete with respect to $N_{\mathcal{K}}$, we know that for any $K \in Q$, $K \in 2^{A_i} \cup P_i$ and thus $2^K = \Gamma_*(\{K\}) \subseteq 2^{A_i} \cup P_i$. Also, for any $K \in Q$, $2^K \cap Q \neq \emptyset$. Hence, we have

$$\sum_{\substack{M \subseteq N_{\mathcal{K}}: \\ M \subseteq 2^{A_i} \cup P_i \\ M \cap Q \neq \emptyset}} Z_M = \sum_{\substack{2^K \subseteq N_{\mathcal{K}}: \\ 2^K \subseteq 2^{A_i} \cup P_i \\ 2^K \cap Q \neq \emptyset}} Z_{2^K} \geq \sum_{\substack{K \in N_{\mathcal{K}}: \\ K \in Q}} S_K. \tag{3.32}$$

Third, for any $L \subseteq D_i, i \in [n]$, note that $T_{L, \overline{(A_i \cup D_i)^c}} \subseteq T_{D_i, \overline{(A_i \cup D_i)^c}} \subseteq P_i^{\mathbf{D}} \subseteq P_i$. Therefore, if $K \in N_{\mathcal{K}}$, $K \subseteq A_i \cup D_i$ and $K \cap L \neq \emptyset$ then K must be in set P_i. Hence,

$$\begin{aligned} \sum_{j \in L} R_j &< \sum_{K \in N_{\mathcal{K}} : K \subseteq A_i \cup D_i, K \cap L \neq \emptyset} S_K \qquad (3.33) \\ &= \sum_{K \in N_{\mathcal{K}} : K \subseteq A_i \cup D_i, K \cap L \neq \emptyset, K \in P_i} S_K. \qquad (3.34) \end{aligned}$$

So far we have proven that $\mathbf{R}$, $(S_K, K \in N_{\mathcal{K}})$ and $(Z_M, M \subseteq N_{\mathcal{K}})$ satisfy (3.25)-(3.27) with $(\mathbf{P}, \mathbf{D})$, which means that $\mathbf{R} \in \mathscr{R}_{\text{TLCC}}^{\text{simplified}}(\mathcal{G}, \mathbf{P}, \mathbf{D})$.

In summary, $\mathscr{R}_{\text{CC}}(\mathcal{G}, \mathbf{D}) \subseteq \mathscr{R}_{\text{TLCC}}^{\text{simplified}}(\mathcal{G}, \mathbf{P}, \mathbf{D})$. □

The simplified TLCC bound $\mathscr{R}_{\text{TLCC}}^{\text{simplified}}(\mathcal{G}, \mathbf{P}, \mathbf{D})$ can be strictly larger than $\mathscr{R}_{\text{CC}}(\mathcal{G}, \mathbf{D})$. In fact, $\mathscr{R}_{\text{TLCC}}^{\text{simplified}}(\mathcal{G}, \mathbf{P}, \mathbf{D})$ can even strictly outperform the ECC scheme for some cases, as shown by the following example.

Example 3.7. Consider the 7-message CIC problem from Example 3.5 again. Time sharing of the ECC scheme over subproblems gives that any symmetric achievable rate $R < 1/3.25$.

For the TLCC scheme, we apply the following simplifications. Fix $\mathbf{D}$ to be $\underline{\mathbf{D}}$. Apply Theorem 3.1 to obtain $N_{\mathcal{K}}$, where $|N_{\mathcal{K}}| = 16$. Remove any $S_K, K \notin N_{\mathcal{K}}$ and $Z_M, M \not\subseteq N_{\mathcal{K}}$. Thus the number of Z_M variables is hugely reduced from $2^{127} - 1$ to $2^{16} - 1$. Find the output of Algorithm 3, $\mathbf{P}^{\mathbf{D}}$, and then we set $\mathbf{P}$ as $P_i = P_i^{\mathbf{D}} \cap N_{\mathcal{K}}, \forall i \in \{1,3,5,6\}$, $P_2 = (P_2^{\mathbf{D}} \cup 2^{\{4,7\}}) \cap N_{\mathcal{K}}$ and $P_i = (P_i^{\mathbf{D}} \cup 2^{\{5\}}) \cap N_{\mathcal{K}}, \forall i \in \{4,7\}$. Use Theorem 3.5 to further remove unnecessary Z_M variables. Applying Theorem 3.4 we obtain that any symmetric rate $R < 1/3$ can be achieved, strictly outperforming the ECC result.

Moreover, $R < 1/3$ can be obtained with even much lower computational complexity via using $N'_{\mathcal{K}}$ given by Algorithm 1 instead of $N_{\mathcal{K}}$ for the entire simplifying process. Note that $|N'_{\mathcal{K}}| = 7$, and thus the number of Z_M variables is only $2^7 - 1$. We can simply set $\mathbf{P}$ such that $P_i = P_i^{\mathbf{D}} \cap N'_{\mathcal{K}}, \forall i \in [n]$. Applying Theorem 3.5 to further remove unnecessary Z_M variables, the final number of remaining Z_M variables is merely 6.

3.4 Distributed Composite Coding for a Fixed Decoding Configuration

In this and the next section, we extend the composite coding (CC) scheme to the DIC problem through combining and generalizing the key ideas of cooperative composite coding (CCC) from Li et al. [2018] and our enhanced fractional allocation in Section 3.2. We also add another degree of freedom into the coding scheme by allowing receiver-dependent decoding choices. The extended coding scheme is referred to as the distributed composite coding (DCC) scheme. For gentler presentation, in this section we focus on the DCC scheme for a fixed decoding configuration (to be defined shortly), while the general DCC scheme incorporating varying decoding configurations is to be presented in Section 3.5.

Let $t_i = \lceil rR_i \rceil$, $i \in [n]$. Recall that t_i is the length and R_i is the rate of message i. The DCC scheme is a two-stage nonlinear coding scheme based on random coding. In the first stage of encoding, each nonempty subset of messages $K \subseteq [n]$ is mapped via random coding to a corresponding composite index, denoted by $w_K \in \{0,1\}^{s_K}$, with rate S_K and length $s_K = \lceil rS_K \rceil$. By convention, $s_\emptyset = S_\emptyset = 0$. In the second stage, composite indices $w_K, K \subseteq J$ that are available at server J are mapped together to a codeword $y_J \in \{0,1\}^{r_J}$, again via random coding, where $r_J = \lfloor rC_J \rfloor$, $J \in N$.

Decoding takes place in the reverse order of stages. For each receiver $i \in [n]$, fix a decoding server group $P_i \subseteq N \setminus 2^{A_i}$ and a decoding message set $D_i \subseteq [n] \setminus A_i$ such that $i \in D_i$. In the first stage of decoding, receiver i decodes all the composite indices to which some servers in P_i have access. In the second stage, receiver i decodes the messages in D_i according to the decoded composite indices from the first stage. The tuples $\mathbf{P} = (P_i, i \in [n])$ and $\mathbf{D} = (D_i, i \in [n])$ are collectively referred to as the decoding configuration for the DCC scheme.

Remark 3.5. Note that the same notation $\mathbf{P} = (P_i, i \in [n])$ that appears in both the decoding configuration definitions for the TLCC scheme (cf. Remark 3.2) and the DCC scheme has different meanings. For the TLCC scheme, **P** denotes the tuple of decoding composite index sets such that each receiver i decodes all the composite indices $w_K, K \in P_i$. In contrast for the DCC scheme, **P** denotes the tuple of decoding server groups, and each receiver i decodes all the composite indices w_K such that there exists some $J \in P_i$, $K \subseteq J$.

Remark 3.6. Note that $\bigcup_{J \in P_i} J$ is the set of message indices that is collectively available to servers in P_i. In the second step of decoding receiver i can only possibly recover messages contained by the servers J in the decoding server group P_i. Hence, we always require that $D_i \subseteq \cup_{J \in P_i} J$, which is similar to the TLCC scheme (cf. Remark 3.3).

As stated above, in this section we provide an achievable rate region for the DCC scheme for a fixed $(\mathbf{P}, \mathbf{D})$. For any $P \subseteq N$, $\Gamma_*(P) = \bigcup_{J \in P}\{K : K \subseteq J\} = \bigcup_{J \in P} 2^J$ is the subset completion of P (cf. (3.31)). Similarly, for any $M \subseteq N$, $\Gamma^*(M) = \bigcup_{K \in M}\{J \in N : K \subseteq J\}$ is the superset completion of M (with respect to N). One can think of $\Gamma_*(P)$ as denoting the set of all composite indices that the servers in P can collectively access. One can think of $\Gamma^*(M)$ as the set of all servers that have access to at least one composite index in M.

Theorem 3.7 (Distributed composite coding (DCC) bound for a fixed decoding configuration)**.** Consider the DIC problem $\mathcal{G}$: $(i|j \in A_i)$, $i \in [n]$ with link capacity tuple **C** and a given decoding configuration $(\mathbf{P}, \mathbf{D})$. The capacity region $\mathscr{C}(\mathcal{G}, \mathbf{C})$ is inner bounded by the rate region $\mathscr{R}_{\text{DCC}}(\mathcal{G}, \mathbf{C}, \mathbf{P}, \mathbf{D})$ that consists of all rate tuples **R** satisfying

$$\sum_{j \in L} R_j < \sum_{\substack{K \subseteq D_i \cup A_i, \\ K \in \Gamma_*(P_i), \\ K \cap L \neq \emptyset}} S_K, \qquad \forall L \subseteq D_i, i \in [n], \tag{3.35}$$

$$\sum_{K \in M} S_K < \sum_{J \in \Gamma^*(M) \cap P_i} C_J, \qquad \forall M \subseteq \Gamma_*(P_i) \setminus 2^{A_i}, i \in [n], \tag{3.36}$$

for some $S_K \geq 0, K \subseteq [n]$.

Before outlining the coding scheme corresponding to Theorem 3.7, we showcase its use via the following example.

Example 3.8. Consider the DIC problem $(1|-), (2|3), (3|2)$ with $n = 3$ messages and non-negative, but otherwise arbitrary link capacities $C_J \geq 0, J \in N \setminus \{\emptyset\}$ and $C_\emptyset = 0$. Note that the set $N = \{\emptyset, \{1\}, \{2\}, \{1,2\}, \{3\}, \{1,3\}, \{2,3\}, \{1,2,3\}\}$ is subset complete and hence, $\Gamma_*(N) = N$.

We fix $P_1 = \{\{1\}\}$, $P_2 = \{\{2\}, \{1,2\}, \{2,3\}\}$, and $P_3 = \{\{3\}, \{1,3\}, \{1,2,3\}\}$. We fix $D_1 = \{1\}$, $D_2 = \{1,2\}$, and $D_3 = \{1,3\}$. Hence, $D_1 \cup A_1 = \{1\}$, and $D_i \cup A_i = \{1,2,3\}, i = 2,3$. Note that $\Gamma_*(P_1) = \{\emptyset, \{1\}\}, \Gamma_*(P_2) = \{\emptyset, \{1\}, \{2\}, \{1,2\}, \{3\}, \{2,3\}\}$ and $\Gamma_*(P_3) = N$. Inequality (3.35) (including active and inactive inequalities) gives (3.37). Inequality (3.36) (excluding inactive inequalities) yields (3.38).

$$\begin{aligned}
&R_1 < S_{\{1\}}, && i=1, L=D_1,\\
&R_1 < S_{\{1\}} + S_{\{1,2\}}, && i=2, L=\{1\}\subset D_2,\\
&R_2 < S_{\{2\}} + S_{\{1,2\}} + S_{\{2,3\}}, && i=2, L=\{2\}\subset D_2,\\
&R_1 + R_2 < S_{\{1\}} + S_{\{2\}} + S_{\{1,2\}} + S_{\{2,3\}}, && i=2, L=D_2,\\
&R_1 < S_{\{1\}} + S_{\{1,2\}} + S_{\{1,3\}} + S_{\{1,2,3\}}, && i=3, L=\{1\}\subset D_3,\\
&R_3 < S_{\{3\}} + S_{\{1,3\}} + S_{\{2,3\}} + S_{\{1,2,3\}}, && i=3, L=\{3\}\subset D_3,\\
&R_1 + R_3 < S_{\{1\}} + S_{\{1,2\}} + S_{\{3\}} + S_{\{1,3\}} + S_{\{2,3\}} + S_{\{1,2,3\}}, && i=3, L=D_3.
\end{aligned} \tag{3.37}$$

$$\begin{aligned}
&S_{\{1\}} < C_{\{1\}}, && i=1, M_1=\Gamma_*(P_1), \Gamma^*(M_1)\cap P_1 = P_1,\\
&S_{\{1\}} + S_{\{1,2\}} < C_{\{1,2\}}, && i=2, M_2=\{\{1\},\{1,2\}\}, \Gamma^*(M_2)\cap P_2=\{\{1,2\}\},\\
&S_{\{2,3\}} < C_{\{2,3\}}, && i=2, M_3=\{\{2,3\}\}, \Gamma^*(M_3)\cap P_2=\{\{2,3\}\},\\
&S_{\{1\}} + S_{\{1,2\}} + S_{\{2,3\}} < C_{\{1,2\}} + C_{\{2,3\}}, && i=2, M_4=\{\{1\},\{1,2\},\{2,3\}\},\\
&S_{\{1\}} + S_{\{2\}} + S_{\{1,2\}} + S_{\{2,3\}} && \\
&\qquad\qquad < C_{\{2\}} + C_{\{1,2\}} + C_{\{2,3\}}, && i=2, M_5=\{\{1\},\{2\},\{1,2\},\{2,3\}\},\\
&S_{\{1,2\}} + S_{\{2,3\}} + S_{\{1,2,3\}} < C_{\{1,2,3\}}, && i=3, M_6=\{\{1,2\},\{2,3\}\,\{1,2,3\}\},\\
&S_{\{1\}} + S_{\{1,2\}} + S_{\{1,3\}} + S_{\{2,3\}} + S_{\{1,2,3\}} && \\
&\qquad\qquad < C_{\{1,3\}} + C_{\{1,2,3\}}, && i=3, M_7=\{\{1\},\{1,2\},\{1,3\},\{2,3\}\,\{1,2,3\}\},\\
&\textstyle\sum_{K\in M_8} S_K < C_{\{3\}} + C_{\{1,3\}} + C_{\{1,2,3\}}, && i=3, M_8=\Gamma_*(P_3)\setminus\{\{2\}\}, \Gamma^*(M_8)\cap P_3 = P_3.
\end{aligned} \tag{3.38}$$

We apply FME to eliminate all present S_K and find the achievable rate region as

$$\begin{aligned}
R_1 &< \min\{C_{\{1\}}, C_{\{1,2\}}, C_{\{1,3\}} + C_{\{1,2,3\}}\},\\
R_1 + R_2 &< C_{\{2\}} + C_{\{1,2\}} + C_{\{2,3\}},\\
R_1 + R_3 &< C_{\{3\}} + C_{\{1,3\}} + C_{\{1,2,3\}}.
\end{aligned}$$

Note that we are not claiming that the choice of $(\mathbf{P}, \mathbf{D})$ is optimal for this toy example. The chosen $\mathbf{D}$ is indeed optimal, but for a more compact exposition, we chose different and smaller suboptimal sets $P_i \subset N$, instead of the optimal $P_i = N$, $i \in [n]$.

Outline of Coding Scheme for Theorem 3.7. **Codebook generation.** *Step 1.* For each message set $K \subseteq [n]$ and each realization of x_K, generate a composite index $w_K = w_K(x_K)$ drawn independently and uniformly at random from $[2^{s_K}]$. That is, the composite index w_K is generated according to the random mapping w_K as

$$w_K : \prod_{j\in K}[2^{t_j}] \to [2^{s_K}].$$

For brevity, when we say composite index K, we mean composite index w_K.

Step 2. For each server $J \in N$ and each realization of composite index tuple $(w_K, K \in 2^J)$, generate a codeword $y_J = y_J((w_K, K \in 2^J))$ drawn independently and uniformly at random from $[2^{r_J}]$. That is, the codeword y_J is generated according to the random mapping y_J as

$$y_J : \prod_{K \in 2^J} [2^{s_K}] \to [2^{r_J}].$$

The codebook $\{(w_K = w_K(x_K), K \subseteq [n]), (y_J = y_J((w_K, K \in 2^J)), J \in N)\}$ is revealed to all *corresponding* parties.[2]

Encoding. To communicate messages $x_{[n]}$, each server $J \in N$ computes $w_K = w_K(x_K)$ for each $K \in 2^J$ and transmits $y_J = y_J((w_K, K \in 2^J))$.

Decoding. *Step 1.* For $i \in [n]$, receiver i finds the unique composite index tuple $(\hat{w}_K, K \in \Gamma_*(P_i))$ such that $y_J = y_J((\hat{w}_K, K \in 2^J))$ for every $J \in P_i$. If there is more than one such tuple, it declares an error.

Step 2. Assuming that $(\hat{w}_K, K \in \Gamma_*(P_i))$ is correct, receiver i finds the unique message tuple $\hat{x}_{D_i}$ such that $\hat{w}_K = w_K(\hat{x}_K)$ for every $K \in \Gamma_*(P_i)$ with $K \subseteq D_i \cup A_i$. If there is more than one such tuple, it declares an error.

The inequalities in (3.35) signify the second-step decoding constraints for the messages in D_i to be recovered with vanishingly small probability of error from all composite indices K in $\Gamma_*(P_i)$, with the help of side information A_i. The inequalities in (3.36) signify the first-step decoding constraints for the composite indices that the servers in P_i have access to (except those that can be generated from side information) to be recovered with vanishingly small probability of error from the outputs y_J from the servers J in P_i. The details of error analysis of Theorem 3.7 is provided in Appendix A.1. □

To help with understanding of Theorem 3.7, we also present the error analysis for a specific example as follows.

Example 3.9. Let us revisit Example 3.8 and consider decoding for receiver 2 with $P_2 = \{\{2\}, \{1,2\}, \{2,3\}\}$, $D_2 = \{1,2\}$.

In the first step of decoding, receiver 2 tries to decode composite indices $\hat{w}_K, K \in \Gamma_*(P_2)$ from the received codewords $y_J, J \in P_2$ and its side information x_{A_2}. The decoding error probability P_e is the probability that there exists some composite index tuple $(\hat{w}_K, K \in \Gamma_*(P_2)) = (\hat{w}_{\{1\}}, \hat{w}_{\{2\}}, \hat{w}_{\{1,2\}}, \hat{w}_{\{3\}}, \hat{w}_{\{2,3\}})$ other than the correct (actually transmitted) tuple $(w_{\{1\}}, w_{\{2\}}, w_{\{1,2\}}, w_{\{3\}}, w_{\{2,3\}})$ such that they are mapped to the same codeword y_J for every $J \in P_2 = \{\{2\}, \{1,2\}, \{2,3\}\}$. To utilize the union bound for upper bounding P_e, we

[2] One of the servers must act as a representative or a central processing unit to generate the codebook and reveal the codebook to *all corresponding* servers and all users. This is because the random mapping of x_K to composite index w_K should be identical among all servers that can generate w_K. For composite index K, the corresponding servers are indexed by the superset of K with respect to the set of all servers N, $\Gamma^*(K)$.

partition the error event according to the erroneous composite index set $M \subseteq \Gamma_*(P_2) \setminus 2^{A_2} = \{\{1\},\{2\},\{1,2\},\{2,3\}\}$. That is, $\hat{w}_K \neq w_K$ iff $K \in M$. Note that 2^{A_2} is always excluded from M for the reason that $\hat{w}_{\{3\}}$ can be generated by receiver 2 from $x_{A_2} = x_{\{3\}}$ and thus will never be erroneous. Therefore, by the union bound, we have $P_e \leq \sum_{M \subseteq \{\{1\},\{2\},\{1,2\},\{2,3\}\}} P_e^M$, where

$$P_e^M \doteq \sum_{\substack{(\hat{w}_{\{1\}},\hat{w}_{\{2\}},\hat{w}_{\{1,2\}},w_{\{3\}},\hat{w}_{\{2,3\}}): \\ \hat{w}_K \neq w_K, K \in M, \\ \hat{w}_K = w_K, K \notin M}} \mathrm{P}\left\{ \bigcap_{\substack{J \in \{\{2\},\{1,2\},\{2,3\}\}, \\ J \in \Gamma^*(M)}} \left\{ y_J = y_J(\hat{w}_K, K \in 2^J) \right\} \right\}.$$

To ensure vanishingly small decoding error probability P_e, each P_e^M has to be vanishingly small. In particular, we present detailed analysis for P_e^M with $M = M_4 = \{\{1\},\{1,2\},\{2,3\}\}$ as specified in Example 3.8, while other P_e^M can be analyzed similarly. Since $P_2 \cap \Gamma^*(M) = \{\{2\},\{1,2\},\{2,3\}\} \cap \Gamma^*(\{\{1\},\{1,2\},\{2,3\}\}) = \{\{1,2\},\{2,3\}\}$, we have

$$\begin{aligned} P_e^{M_4} &= \sum_{\substack{(\hat{w}_{\{1\}},w_{\{2\}},\hat{w}_{\{1,2\}},w_{\{3\}},\hat{w}_{\{2,3\}}): \\ \hat{w}_K \neq w_K, K \in \{\{1\},\{1,2\},\{2,3\}\}}} \mathrm{P}\left\{ \bigcap_{J \in \{\{1,2\},\{2,3\}\}} \left\{ y_J = y_J(\hat{w}_K, K \in 2^J) \right\} \right\} \\ &= \frac{(2^{s_{\{1\}}} - 1) \cdot (2^{s_{\{1,2\}}} - 1) \cdot (2^{s_{\{2,3\}}} - 1)}{2^{r_{\{1,2\}}} \cdot 2^{r_{\{2,3\}}}} \qquad (3.39) \\ &< 2^{s_{\{1\}}} \cdot 2^{s_{\{1,2\}}} \cdot 2^{s_{\{2,3\}}} \cdot 2^{-r_{\{1,2\}}} \cdot 2^{-r_{\{2,3\}}} \\ &< 2^{rS_{\{1\}}+1+rS_{\{1,2\}}+1+rS_{\{2,3\}}+1-(rC_{\{1,2\}}-1)-(rC_{\{2,3\}}-1)} \\ &= 2^5 \cdot 2^{r(S_{\{1\}}+S_{\{1,2\}}+S_{\{2,3\}}-C_{\{1,2\}}-C_{\{2,3\}})}, \qquad (3.40) \end{aligned}$$

where (3.39) holds since there are $(2^{s_{\{1\}}} - 1) \cdot (2^{s_{\{1,2\}}} - 1) \cdot (2^{s_{\{2,3\}}} - 1)$ erroneous tuples $(\hat{w}_{\{1\}}, w_{\{2\}}, \hat{w}_{\{1,2\}}, w_{\{3\}}, \hat{w}_{\{2,3\}})$ where $\hat{w}_K \neq w_K, K \in M_4$, and for each erroneous composite index tuple, it is mapped to the same codeword y_J as the correct tuple for all $J \in \{\{1,2\},\{2,3\}\}$ with probability $1/(2^{r_{\{1,2\}}} \cdot 2^{r_{\{2,3\}}})$ due to the uniform random codebook generation. According to (3.40), $P_e^{M_4}$ tends to 0 as $r \to \infty$, provided that

$$S_{\{1\}} + S_{\{1,2\}} + S_{\{2,3\}} < C_{\{1,2\}} + C_{\{2,3\}}.$$

Note that the above constraint has appeared in the system of inequalities given by (3.36) in Example 3.8. Other inequalities given by (3.36) with $i = 2$ in Example 3.8 are required to ensure vanishingly small P_e^M for other $M \subseteq \{\{1\},\{2\},\{1,2\},\{2,3\}\}$. All these inequalities are to be satisfied to ensure a vanishingly small first-step decoding error probability P_e for receiver 2.

The error analysis for the second step of decoding can be done in a similar way. Assume that all the composite indices $\hat{w}_K, K \in \Gamma_*(P_2)$ have been correctly decoded. In the second step of decoding, receiver 2 tries to decode messages $\hat{x}_{D_2} = \hat{x}_{\{1,2\}}$ from the decoded composite indices $\hat{w}_K, K \in \Gamma_*(P_2)$ and its side information $x_{A_2} = x_{\{3\}}$. The decoding error probability P_e is the probability that there exists some message tuple $(\hat{x}_1, \hat{x}_2, x_3)$ other than the correct tuple (x_1, x_2, x_3) such that they are mapped to the same composite index w_K for every $K \in 2^{D_2 \cup A_2} \cap \Gamma_*(P_2) = 2^{\{1,2,3\}} \cap \{\{1\},\{2\},\{1,2\},\{3\},\{2,3\}\} = \{\{1\},\{2\},\{1,2\},\{3\},\{2,3\}\}$. We partition this error event according to the erroneous message set $L \subseteq D_2 = \{1,2\}$. That is, $\hat{x}_j \neq x_j$ iff $j \in L$. Therefore, by the union bound, we have $P_e \leq \sum_{L \subseteq \{1,2\}} P_e^L$, where

$$P_e^L \doteq \sum_{\substack{(\hat{x}_1,\hat{x}_2): \\ \hat{x}_j \neq x_j, j \in L, \\ \hat{x}_j = x_j, j \notin L}} \mathrm{P}\left\{ \bigcap_{\substack{K \in \{\{1\},\{2\},\{1,2\},\{3\},\{2,3\}\}, \\ K \cap L \neq \emptyset}} \{\hat{w}_K = w_K(\hat{x}_K)\} \right\}.$$

To ensure vanishingly small P_e, each P_e^L has to be vanishingly small. In particular, we present detailed analysis for P_e^L with $L = D_2 = \{1,2\}$, while other P_e^L can be analyzed similarly. We have

$$\begin{aligned} P_e^L &= \sum_{\substack{(\hat{x}_1,\hat{x}_2): \\ \hat{x}_1 \neq x_1, \hat{x}_2 \neq x_2}} \mathrm{P}\left\{ \bigcap_{K \in \{\{1\},\{2\},\{1,2\},\{2,3\}\}} \{\hat{w}_K = w_K(\hat{x}_K)\} \right\} \\ &= \frac{(2^{t_1}-1)\cdot(2^{t_2}-1)}{2^{s_{\{1\}}} \cdot 2^{s_{\{2\}}} \cdot 2^{s_{\{1,2\}}} \cdot 2^{s_{\{2,3\}}}} \qquad (3.41) \\ &< 2^{t_1} \cdot 2^{t_2} \cdot 2^{-s_{\{1\}}} \cdot 2^{-s_{\{2\}}} \cdot 2^{-s_{\{1,2\}}} \cdot 2^{-s_{\{2,3\}}} \\ &< 2^{rR_1+1+rR_2+1-rS_{\{1\}}-rS_{\{2\}}-rS_{\{1,2\}}-rS_{\{2,3\}}} \\ &= 2^2 \cdot 2^{r(R_1+R_2-S_{\{1\}}-S_{\{2\}}-S_{\{1,2\}}-S_{\{2,3\}})}, \qquad (3.42) \end{aligned}$$

where (3.41) holds since there are $(2^{t_1}-1)\cdot(2^{t_2}-1)$ erroneous message tuples, and for each erroneous tuple, it is mapped to the same composite index w_K as the correct tuple for all $K \in \{\{1\},\{2\},\{1,2\},\{2,3\}\}$ with probability $1/(2^{s_{\{1\}}} \cdot 2^{s_{\{2\}}} \cdot 2^{s_{\{1,2\}}} \cdot 2^{s_{\{2,3\}}})$ due to the uniform random codebook generation. According to (3.42), P_e^L tends to 0 as $r \to \infty$, provided that

$$R_1 + R_2 < S_{\{1\}} + S_{\{2\}} + S_{\{1,2\}} + S_{\{2,3\}}.$$

Note that the above constraint appears in the system of inequalities given by (3.35) in Example 3.8. Other inequalities given by (3.35) with $i = 2$ in Example 3.8 are required to ensure vanishingly small P_e^L for other $L \subseteq \{1,2\}$. All these inequalities are to be satisfied to have a vanishingly small second-step decoding error probability P_e for receiver 2.

The achievable rate region $\mathscr{R}_{\mathrm{DCC}}(\mathcal{G}, \mathbf{C}, \mathbf{P}, \mathbf{D})$ in Theorem 3.7 can be represented equivalently as follows.

Proposition 3.2. Consider the DIC problem $\mathcal{G}$: $(i|j \in A_i), i \in [n]$ with link capacity tuple $\mathbf{C}$ and a given decoding configuration $(\mathbf{P}, \mathbf{D})$. The capacity region $\mathscr{C}(\mathcal{G}, \mathbf{C})$ is inner bounded by the rate region $\mathscr{R}_{\mathrm{DCC}}(\mathcal{G}, \mathbf{C}, \mathbf{P}, \mathbf{D})$ that consists of all rate tuples $\mathbf{R}$ satisfying (3.35) and

$$\sum_{K \in \Gamma_*(Q) \setminus \Gamma_*(P_i \setminus Q) \setminus 2^{A_i}} S_K < \sum_{J \in Q} C_J, \qquad \forall Q \subseteq P_i, i \in [n]. \tag{3.43}$$

Here, the summand on the LHS of (3.43) signifies the set of composite indices that can be accessed only by the servers in Q (and not by the servers in $P_i \setminus Q$), and that are not generated freely from the side information A_i. The proof of Proposition 3.2 is provided in Appendix A.2.

Remark 3.7. In Sadeghi et al. [2016], composite index rates had been split across servers, where server-specific rates, $S_{K,J}$, $K \subseteq J$, were limited by the corresponding server capacity, C_J. However, as demonstrated by Li et al. [2017, 2018] such rate splitting can be suboptimal, and cooperative composite coding (CCC) (see Proposition 3.3 in Section 3.5) can generally achieve tighter inner bounds on the capacity region of the DIC problems, where the same subset of messages are mapped to the same composite index at different servers. Subsequently, composite indices w_K and their corresponding rates S_K are not server-specific. In this chapter, we have adopted cooperative compression of composite indices as baseline.

Remark 3.8. Compared to the previous works Sadeghi et al. [2016]; Li et al. [2017, 2018], we have introduced user-specific decoding server groups, $P_i \subseteq N$, $i \in [n]$. Compared to Li et al. [2017, 2018] we use a more flexible enhanced fractional allocation of link capacities over decoding configurations (see Section 3.5). See Remark3.11, as well as Examples 3.13 and 3.15 for more details on how these improvements can lead to generally tighter inner bounds on the capacity region.

Remark 3.9. If $C_J = 0$ for some $J \in N$, we can limit our attention to the set of *active* servers J with positive capacity, denoted by $N_A \doteq \{J \in N : C_J > 0\}$. Our results in Theorem 3.7 and Proposition 3.2 can easily incorporate the set of active servers N_A, which can reduce the computational complexity of characterizing the rate region. Example 3.10 in the following illustrates how the rate region is easily specialized when active servers N_A is a strict subset of N. Example 3.10 also shows an instance of equivalence of Theorem 3.7 and Proposition 3.2. Note that the results in Li et al. [2017, 2018] are presented based on the set of active servers.

Example 3.10. Consider the 4-message DIC problem $\mathcal{G}$: $(1|4), (2|1,3), (3|1,2), (4|2,3)$ with two active servers $N_A = \{\{1,2,3\}, \{2,3,4\}\}$ with positive link capacities $C_J > 0, J \in N_A$

and $C_J = 0$, $J \in N \setminus N_A$. Note that

$$\begin{aligned}
\Gamma_*(N_A) &= \bigcup_{J \in N_A} 2^J \\
&= \{\emptyset, \{1\}, \{2\}, \{1,2\}, \{3\}, \{1,3\}, \{2,3\}, \{1,2,3\}\} \\
&\quad \cup \{\emptyset, \{2\}, \{3\}, \{2,3\}, \{4\}, \{2,4\}, \{3,4\}, \{2,3,4\}\} \\
&= \{\emptyset, \{1\}, \{2\}, \{1,2\}, \{3\}, \{1,3\}, \{2,3\}, \{1,2,3\}, \\
&\quad \{4\}, \{2,4\}, \{3,4\}, \{2,3,4\}\}.
\end{aligned}$$

We choose $P_i = N_A$, $i \in [n]$, $D_1 = \{1\}$, $D_2 = \{2,4\}$, $D_3 = \{3,4\}$ and $D_4 = \{1,4\}$. Hence, $D_1 \cup A_1 = \{1,4\}$, and $D_i \cup A_i = [n]$, $i = 2,3,4$. Active inequalities from (3.35) are given in (3.44).

$$\begin{aligned}
& R_1 < S_{\{1\}}, && i = 1, L = D_1, \\
& R_2 < S_{\{2\}} + S_{\{1,2\}} + S_{\{2,3\}} + S_{\{1,2,3\}} && \\
& \qquad + S_{\{2,4\}} + S_{\{2,3,4\}}, && i = 2, L = \{2\} \subset D_2, \\
& R_4 < S_{\{4\}} + S_{\{2,4\}} + S_{\{3,4\}} + S_{\{2,3,4\}}, && i = 2,3,4, L = \{4\} \subset D_i, \\
& R_2 + R_4 < S_{\{2\}} + S_{\{1,2\}} + S_{\{1,3\}} + S_{\{2,3\}} + S_{\{1,2,3\}} && \\
& \qquad + S_{\{4\}} + S_{\{2,4\}} + S_{\{2,3,4\}}, && i = 2, L = D_2, \\
& R_3 < S_{\{3\}} + S_{\{1,3\}} + S_{\{2,3\}} + S_{\{1,2,3\}} && \\
& \qquad + S_{\{3,4\}} + S_{\{2,3,4\}}, && i = 3, L = \{3\} \subset D_3, \\
& R_3 + R_4 < S_{\{3\}} + S_{\{1,3\}} + S_{\{2,3\}} + S_{\{1,2,3\}} + S_{\{4\}} && \\
& \qquad + S_{\{2,4\}} + S_{\{3,4\}} + S_{\{2,3,4\}}, && i = 3, L = D_3 \\
& R_1 + R_4 < S_{\{1\}} + S_{\{1,2\}} + S_{\{1,3\}} + S_{\{1,2,3\}} + S_{\{4\}} && \\
& \qquad + S_{\{2,4\}} + S_{\{3,4\}} + S_{\{2,3,4\}}, && i = 4, L = D_4.
\end{aligned} \tag{3.44}$$

$$\begin{aligned}
& S_{\{1\}} + S_{\{1,2\}} + S_{\{1,3\}} + S_{\{1,2,3\}} < C_{\{1,2,3\}}, && i = 1,4, M = M_1, \\
& S_{\{4\}} + S_{\{2,4\}} + S_{\{3,4\}} + S_{\{2,3,4\}} < C_{\{2,3,4\}}, && i = 2,3,4, M = M_2, \\
& \textstyle\sum_{K \in N, K \neq \{4\}} S_K < C_{\{1,2,3\}} + C_{\{2,3,4\}}, && i = 1, M = M_3 \setminus \{\{4\}\}, \\
& \textstyle\sum_{K \in \{\{2\},\{1,2\},\{2,3\},\{1,2,3\},\{4\},\{2,4\},\{3,4\},\{2,3,4\}\}} S_K && \\
& \qquad < C_{\{1,2,3\}} + C_{\{2,3,4\}}, && i = 2, M = M_3 \setminus 2^{\{1,3\}}, \\
& \textstyle\sum_{K \in \{\{3\},\{1,3\},\{2,3\},\{1,2,3\},\{4\},\{2,4\},\{3,4\},\{2,3,4\}\}} S_K && \\
& \qquad < C_{\{1,2,3\}} + C_{\{2,3,4\}}, && i = 3, M = M_3 \setminus 2^{\{1,2\}}, \\
& \textstyle\sum_{K \in \{\{1\},\{1,2\},\{1,3\},\{1,2,3\},\{4\},\{2,4\},\{3,4\},\{2,3,4\}\}} S_K && \\
& \qquad < C_{\{1,2,3\}} + C_{\{2,3,4\}}, && i = 4, M = M_3 \setminus 2^{\{2,3\}}.
\end{aligned} \tag{3.45}$$

Note that with respect to N_A, for $M_1 = \{\{1\},\{1,2\},\{1,3\},\{1,2,3\}\}$, we have $\Gamma^*(M_1) = \{\{1,2,3\}\}$, for $M_2 = \{\{4\},\{2,4\},\{3,4\},\{2,3,4\}\}$, we have $\Gamma^*(M_2) = \{\{2,3,4\}\}$, and for $M_3 = \Gamma_*(N_A)$, we have $\Gamma^*(M_3) = N_A$. With this in mind, we write the active inequalities (3.36) as (3.45).

We apply FME to eliminate all present S_K variables in (3.44) and (3.45) and find the achievable rate region as

$$\begin{aligned} R_1 &< C_{\{1,2,3\}}, \\ R_4 &< C_{\{2,3,4\}}, \\ R_1 + R_2 &< C_{\{1,2,3\}} + C_{\{2,3,4\}}, \\ R_1 + R_3 &< C_{\{1,2,3\}} + C_{\{2,3,4\}}, \\ R_2 + R_4 &< C_{\{1,2,3\}} + C_{\{2,3,4\}}, \\ R_3 + R_4 &< C_{\{1,2,3\}} + C_{\{2,3,4\}}. \end{aligned}$$

It can be verified that the above rate region matches the all-server grouping PM bound $\mathscr{R}_*(\mathcal{G})$ in Corollary 4.5 to be presented in Section 4.8, and thus establishes the capacity region for the problem.

We can rewrite active first-step decoding inequalities in the form of (3.43) in Proposition 3.2. Recall that $P_i = N_A = \{\{1,2,3\},\{2,3,4\}\}$, $i \in [n]$. For $Q_1 = \{\{1,2,3\}\}$, we have $\Gamma_*(Q_1) \setminus \Gamma_*(P_i \setminus Q_1) = \{\{1\},\{1,2\},\{1,3\},\{1,2,3\}\}$, for $Q_2 = \{\{2,3,4\}\}$, we have $\Gamma_*(Q_2) \setminus \Gamma_*(P_i \setminus Q_2) = \{\{4\},\{2,4\},\{3,4\},\{2,3,4\}\}$, and for $Q_3 = P_i = N_A$, we have

$$\begin{aligned} &\Gamma_*(Q_3) \setminus \Gamma_*(P_i \setminus Q_3) \\ &= \{\{1\},\{2\},\{1,2\},\{3\},\{1,3\},\{2,3\},\{1,2,3\},\{4\},\{2,4\},\{3,4\},\{2,3,4\}\}. \end{aligned}$$

Inequality (3.43) (excluding inactive inequalities) gives (3.45), which showcases the equivalence of (3.36) and (3.43) as claimed by Proposition 3.2.

We now present a few simplifications of Theorem 3.7. First, setting $P_i = N$, $i \in [n]$ (that is, all receivers access all server outputs for decoding) yields the following.

Corollary 3.2. Consider the DIC problem $\mathcal{G}$: $(i|j \in A_i)$, $i \in [n]$ with link capacity tuple $\mathbf{C}$ and a given decoding message set tuple $\mathbf{D}$. The capacity region $\mathscr{C}(\mathcal{G},\mathbf{C})$ is inner bounded by the rate region $\mathscr{R}_{\mathrm{DCC}}(\mathcal{G},\mathbf{C},(N,N,\dots,N),\mathbf{D})$ that consists of all rate tuples $\mathbf{R}$ satisfying

$$\sum_{j\in L} R_j < \sum_{\substack{K\subseteq D_i\cup A_i,\\ K\cap L\neq\emptyset}} S_K, \qquad \forall L \subseteq D_i,\ i\in[n], \tag{3.46}$$

$$\sum_{K\in M} S_K < \sum_{J\in\Gamma^*(M)} C_J, \qquad \forall M \subseteq N\setminus 2^{A_i},\ i\in[n]. \tag{3.47}$$

Corollary 3.2 can still result in a tight bound on the sum capacity.

Example 3.11. Consider the 4-message DIC problem $(1|-), (2|4), (3|4), (4|3)$ with equal link capacities $C_J = 1$ for all $J \in N \setminus \{\emptyset\}$. Choose $P_i = N$, $i \in [n]$. Choose $D_1 = \{1\}$ and $D_i = [n] \setminus A_i$ for $i = 2, 3, 4$. Maximizing the sum-rate under the constraints (3.46) and (3.47) results in $R_1 + R_2 + R_3 + R_4 < 21$, which is the sum capacity of this DIC problem; see Example 4.10 for the matching upper bound. Note that the sum capacity of 21 can also be achieved by some linear coding scheme as follows. Set the codeword length $r_J = 1$ for every server $J \in N \setminus \{\emptyset\}$ and message length $t_1 = 4$, $t_2 = 3$, $t_3 = t_4 = 7$. Thus message x_1 can be written as $x_1 = (x_{1,1}, x_{1,2}, x_{1,3}, x_{1,4})$. Similarly, $x_2 = (x_{2,1}, x_{2,2}, x_{2,3})$, $x_3 = (x_{3,1}, \dots, x_{3,7})$, and $x_4 = (x_{4,1}, \dots, x_{4,7})$. Then we generate the codewords y_J, $J \in N \setminus \{\emptyset\}$ as

$$\begin{aligned}
&y_{\{1\}} = (x_{1,1}), \quad y_{\{2\}} = (x_{2,1}), \quad y_{\{3\}} = (x_{3,1}), \quad y_{\{4\}} = (x_{4,1}),\\
&y_{\{1,2\}} = (x_{1,2}), \quad y_{\{1,3\}} = (x_{1,3}), \quad y_{\{1,4\}} = (x_{1,4}),\\
&y_{\{2,3\}} = (x_{2,2} \oplus x_{3,2}), \quad y_{\{2,4\}} = (x_{2,2} \oplus x_{4,2}),\\
&y_{\{1,2,3\}} = (x_{2,3} \oplus x_{3,3}), \quad y_{\{1,2,4\}} = (x_{2,3} \oplus x_{4,3}),\\
&y_{\{3,4\}} = (x_{3,4} \oplus x_{4,4}), \quad y_{\{1,3,4\}} = (x_{3,5} \oplus x_{4,5}),\\
&y_{\{2,3,4\}} = (x_{3,6} \oplus x_{4,6}), \quad y_{\{1,2,3,4\}} = (x_{3,7} \oplus x_{4,7}).
\end{aligned}$$

It can be verified that upon receiving the above codewords, every receiver i can decode x_i with the help of its side information x_{A_i}. For the above linear coding scheme, rate tuple $\mathbf{R} = \mathbf{t} = (4, 3, 7, 7)$ is achievable, thus achieving a sum-rate of $4 + 3 + 7 + 7 = 21$.

We now further simplify Corollary 3.2 by choosing the composite rates explicitly (and potentially suboptimally) as $S_K = C_K$, $K \in N$, which essentially prevents cooperation among the servers and forces server J to transmit the composite index w_J by a one-to-one mapping $y_J(w_J(x_J))$.

Corollary 3.3. Consider the DIC problem $\mathcal{G}$: $(i|j \in A_i)$, $i \in [n]$ with link capacity tuple $\mathbf{C}$ and a given decoding message set tuple $\mathbf{D}$. A rate tuple $\mathbf{R}$ is achievable if

$$\sum_{j \in L} R_j < \sum_{\substack{J \subseteq D_i \cup A_i, \\ J \cap L \neq \emptyset}} C_J, \qquad \forall L \subseteq D_i,\ i \in [n]. \tag{3.48}$$

With no need for Fourier–Motzkin elimination of the composite index rates, the rate region in Corollary 3.3 can be easily evaluated.

Example 3.12. Consider the 4-message DIC problem $(1|4), (2|4), (3|2), (4|3)$ with equal link capacities $C_J = 1$ for all $J \in N \setminus \{\emptyset\}$. Choose $D_i = [n] \setminus A_i$, $i \in [n]$. Corollary 3.3 then simplifies to the set of inequalities $R_i < 8$, $i \in [n]$, $R_i + R_j < 12$, $i \neq j \in [n]$,

$R_1 + R_2 + R_3 < 14$, $R_1 + R_2 + R_4 < 14$, and $R_1 + R_3 + R_4 < 14$. For this problem, Theorem 3.7 (with no pre-specified restriction on S_K) yields the same rate region under equal link capacities, which indeed matches the touch grouping PM bound in Corollary 4.2 in Section 4.6 and thus establishes the capacity region.

3.5 Distributed Composite Coding with Varying Decoding Configurations

In this section we allow message rates R_i, $i \in [n]$, and composite index rates S_K, $K \subseteq [n]$, to be a function of the decoding configuration $(\mathbf{P}, \mathbf{D})$ at the receivers.

More formally, consider the DIC problem $\mathcal{G}$ with link capacity tuple $\mathbf{C}$. Let $\mathcal{P}_i(\mathcal{G}) \doteq \{P_i \subseteq N \setminus 2^{A_i} : i \in \cup_{J \in P_i} J\}$ denote the set of all possible decoding server groups at receiver i, and thus $\mathcal{P}(\mathcal{G}) \doteq \prod_{i \in [n]} \mathcal{P}_i(\mathcal{G})$ denotes the collection of all possible decoding server group tuples among all receivers. Note that such decoding server group tuples does not depend on the link capacity tuple $\mathbf{C}$. Similarly, $\mathcal{D}_i(\mathcal{G}) = \{D_i \subseteq [n] \setminus A_i : i \in D_i\}$ denotes the set of all possible decoding message sets at receiver i, and thus $\mathcal{D}(\mathcal{G}) \doteq \prod_{i \in [n]} \mathcal{D}_i(\mathcal{G})$ denotes the collection of all possible decoding message set tuples among all receivers for the DIC problem. Whenever we refer to a decoding configuration $(\mathbf{P}, \mathbf{D})$, we refer to a decoding server group tuple $\mathbf{P} \in \mathcal{P}(\mathcal{G}) = \prod_{i \in [n]} \{P_i \subseteq N \setminus 2^{A_i} : i \in \cup_{J \in P_i} J\}$ and a decoding message set tuple $\mathbf{D} \in \mathcal{D}(\mathcal{G}) = \prod_{i \in [n]} \{D_i \subseteq [n] \setminus A_i : i \in D_i\}$, such that for any receiver $i \in [n]$, $D_i \subseteq \cup_{J \in P_i} J$.

Let for each $(\mathbf{P}, \mathbf{D})$ and $i \in [n]$, $t_i(\mathbf{P}, \mathbf{D}) = \lceil r R_i(\mathbf{P}, \mathbf{D}) \rceil$, where $R_i(\mathbf{P}, \mathbf{D})$ is the rate of message i communicated via decoding configuration $(\mathbf{P}, \mathbf{D})$. Let $x_i(\mathbf{P}, \mathbf{D}) \in [2^{t_i(\mathbf{P}, \mathbf{D})}]$ be the part of message i communicated via decoding configuration $(\mathbf{P}, \mathbf{D})$. Denote $s_K(\mathbf{P}, \mathbf{D}) = \lceil r S_K(\mathbf{P}, \mathbf{D}) \rceil$, $K \subseteq [n]$, where $S_K(\mathbf{P}, \mathbf{D})$ is the rate of composite index K and configuration $(\mathbf{P}, \mathbf{D})$. Denote $r_J(\mathbf{P}) = \lfloor r C_J(\mathbf{P}) \rfloor$, where $C_J(\mathbf{P})$ is the fractional capacity of server J for decoding server group $\mathbf{P}$. By convention, $s_\emptyset(\mathbf{P}, \mathbf{D}) = S_\emptyset(\mathbf{P}, \mathbf{D}) = 0$ for each $(\mathbf{P}, \mathbf{D})$.

Theorem 3.8 (Distributed composite coding (DCC) bound)**.** Consider the DIC problem $\mathcal{G}$: $(i|j \in A_i)$, $i \in [n]$ with link capacity tuple $\mathbf{C}$. The capacity region $\mathscr{C}(\mathcal{G}, \mathbf{C})$ is inner bounded by the rate region $\mathscr{R}_{\text{DCC}}(\mathcal{G}, \mathbf{C})$ that consists of all rate tuples $\mathbf{R}$ such that

$$R_i = \sum_{\mathbf{P} \in \mathcal{P}(\mathcal{G}), \mathbf{D} \in \mathcal{D}(\mathcal{G})} R_i(\mathbf{P}, \mathbf{D}), \qquad \forall i \in [n], \tag{3.49}$$

$$C_J = \sum_{\mathbf{P} \in \mathcal{P}(\mathcal{G})} C_J(\mathbf{P}), \qquad \forall J \in N, \tag{3.50}$$

for some $R_i(\mathbf{P},\mathbf{D})$, $C_J(\mathbf{P})$, and $S_K(\mathbf{P},\mathbf{D}) \geq 0$ that satisfy

$$\sum_{j\in L} R_j(\mathbf{P},\mathbf{D}) < \sum_{\substack{K\subseteq D_i\cup A_i,\\ K\in\Gamma_*(P_i),\\ K\cap L\neq\emptyset}} S_K(\mathbf{P},\mathbf{D}), \qquad \forall L\subseteq D_i, i\in[n], \tag{3.51}$$

and

$$\sum_{\mathbf{D}}\sum_{K\in M} S_K(\mathbf{P},\mathbf{D}) < \sum_{J\in\Gamma^*(M)\cap P_i} C_J(\mathbf{P}), \qquad \forall M\subseteq \Gamma_*(P_i)\setminus 2^{A_i}, i\in[n]. \tag{3.52}$$

We now outline the coding scheme corresponding to the achievable rate region $\mathscr{R}_{\text{DCC}}(\mathcal{G},\mathbf{C})$ in Theorem 3.8. The corresponding error analysis follows similar steps as in the proof of Theorem 3.7 and thus is omitted for brevity.

Codebook generation: *Step 1.* For each message set $K\subseteq[n]$, decoding configuration $(\mathbf{P},\mathbf{D})$, and each realization of $x_K(\mathbf{P},\mathbf{D})$, a corresponding composite index $w_{K,\mathbf{P},\mathbf{D}} = w_{K,\mathbf{P},\mathbf{D}}(x_K(\mathbf{P},\mathbf{D}))$ is drawn independently and uniformly at random from $[2^{s_K(\mathbf{P},\mathbf{D})}]$. That is, the composite index $w_{K,\mathbf{P},\mathbf{D}}$ is generated according to the random mapping $w_{K,\mathbf{P},\mathbf{D}}$ as

$$w_{K,\mathbf{P},\mathbf{D}} : \prod_{j\in K}[2^{t_j(\mathbf{P},\mathbf{D})}] \to [2^{s_K(\mathbf{P},\mathbf{D})}].$$

Step 2. For each server $J\in N$, decoding server group tuple $\mathbf{P}\in\mathcal{P}(\mathcal{G})$, and each realization of composite index tuple $(w_{K,\mathbf{P},\mathbf{D}},(K,\mathbf{D})\in 2^J\times\mathcal{D}(\mathcal{G}))$, a fractional server index $y_{J,\mathbf{P}}((w_{K,\mathbf{P},\mathbf{D}},(K,\mathbf{D})\in 2^J\times\mathcal{D}(\mathcal{G})))$ is drawn independently and uniformly at random from $[2^{r_J(\mathbf{P})}]$. That is, the fractional server index $y_{J,\mathbf{P}}$ is generated according to the random mapping function $y_{J,\mathbf{P}}$ as

$$y_{J,\mathbf{P}} : \prod_{K\in 2^J}\prod_{\mathbf{D}\in\mathcal{D}(\mathcal{G})}[2^{s_K(\mathbf{P},\mathbf{D})}] \to [2^{r_J(\mathbf{P})}].$$

For each $J\in N$, the final codeword y_J is the deterministic concatenation of the fractional server index tuples, $(y_{J,\mathbf{P}},\mathbf{P}\in\mathcal{P}(\mathcal{G}))$. The random codebook

$$\begin{aligned}&\{(w_{K,\mathbf{P},\mathbf{D}} = w_{K,\mathbf{P},\mathbf{D}}(x_K(\mathbf{P},\mathbf{D})), K\subseteq[n], \mathbf{P}\in\mathcal{P}(\mathcal{G}), \mathbf{D}\in\mathcal{D}(\mathcal{G})),\\ &(y_{J,\mathbf{P}} = y_{J,\mathbf{P}}((w_{K,\mathbf{P},\mathbf{D}},(K,\mathbf{D})\in 2^J\times\mathcal{D}(\mathcal{G}))), J\in N, \mathbf{P}\in\mathcal{P}(\mathcal{G}))\}\end{aligned}$$

is revealed to all corresponding parties. See Footnote 2.

Encoding: To communicate messages $x_{[n]}$, each server $J\in N$ computes the composite index $w_{K,\mathbf{P},\mathbf{D}} = w_{K,\mathbf{P},\mathbf{D}}(x_K(\mathbf{P},\mathbf{D}))$ for each $K\in 2^J$ and $(\mathbf{P},\mathbf{D})$, as well as the server index $y_{J,\mathbf{P}} = y_{J,\mathbf{P}}((w_{K,\mathbf{P},\mathbf{D}},(K,\mathbf{D})\in 2^J\times\mathcal{D}(\mathcal{G})))$ for each $\mathbf{P}$, and then transmits the final codeword $y_J = (y_{J,\mathbf{P}},\mathbf{P}\in\mathcal{P}(\mathcal{G}))$.

Decoding: *Step 1.* For each $i \in [n]$ and each $\mathbf{P} \in \mathcal{P}(\mathcal{G})$, receiver i finds the unique tuple $(\hat{w}_{K,\mathbf{P},\mathbf{D}}, (K, \mathbf{D}) \in \Gamma_*(P_i) \times \mathcal{D}(\mathcal{G}))$ such that $y_{J,\mathbf{P}} = y_{J,\mathbf{P}}((\hat{w}_{K,\mathbf{P},\mathbf{D}}, (K, \mathbf{D}) \in 2^J \times \mathcal{D}(\mathcal{G})))$ for every $J \in P_i$. If there is more than one such tuple, it declares an error.

Step 2. Assuming Step 1 is correctly executed and for each $i \in [n]$ and each $(\mathbf{P}, \mathbf{D})$, receiver i finds the unique message tuple $\hat{x}_{D_i}(\mathbf{P}, \mathbf{D})$ such that $\hat{w}_{K,\mathbf{P},\mathbf{D}} = w_{K,\mathbf{P},\mathbf{D}}(\hat{x}_K(\mathbf{P}, \mathbf{D}))$ for every $K \in \Gamma_*(P_i)$ with $K \subseteq D_i \cup A_i$. If there is more than one such tuple, it declares an error.

Remark 3.10. Computing the DCC bound $\mathscr{R}_{\text{DCC}}(\mathcal{G}, \mathbf{C})$ in Theorem 3.8 over all decoding configurations $(\mathbf{P}, \mathbf{D})$ can be rather expensive, but a few simplifications are possible. First, if $J \notin \cup_{i \in [n]} P_i$, then we can simply set $C_J(\mathbf{P}) = 0$ without loss of generality. Second, as mentioned in Remark 3.9, we can focus on active servers and consider $P_i \subseteq N_A = \{J \in N : C_J > 0\}$. Third, it suffices to consider subset complete decoding server groups, such that $P_i \cup 2^{A_i} = \Gamma_*(P_i \cup 2^{A_i})$. This is because, all subsets of composite indices that can be generated by the servers in P_i (except those that are already known) appear on the LHS of (3.52). That is, composite indices K in $M \subseteq \Gamma_*(P_i) \setminus 2^{A_i}$ appear on the LHS of (3.52). However, on the RHS of (3.52), the contributing server capacities are "cut" by P_i. Therefore, the region becomes no smaller if we use subset completion of P_i instead of P_i itself. Finally, it can be shown that the reduction method for decoding message set tuple $\mathbf{D}$ in Theorem 3.2 and Proposition 3.1 hold even for the DCC scheme. Thus we can safely limit the choices of $\mathbf{D}$ to be the tuples that are element-wisely no smaller than the natural decoding configuration $\underline{\mathbf{D}}$ (cf. Algorithm 2) and, at the same time, not satisfying the removing condition in Theorem 3.2.

As mentioned in Section 2.3.2, a less general version of the fractional coding over different decoding configurations was proposed in the cooperative composite coding (CCC) scheme Li et al. [2018]. We present the achievable rate region given by the CCC scheme in our notation as follows. Note that the CCC scheme uses the same decoding server group $P_i = P \subseteq N$ for all the receivers. For a given $(P, \mathbf{D})$, define $\Delta_i = D_i \cap (\cup_{J \in P} J)$, $\forall i \in [n]$.

Proposition 3.3 (Cooperative composite coding (CCC) bound, Li et al. [2018]). Consider the DIC problem $\mathcal{G}$: $(i|j \in A_i)$, $i \in [n]$ with link capacity tuple $\mathbf{C}$. The capacity region $\mathscr{C}(\mathcal{G}, \mathbf{C})$ is inner bounded by the rate region $\mathscr{R}_{\text{CCC}}(\mathcal{G}, \mathbf{C})$ as

$$\mathscr{R}_{\text{CCC}}(\mathcal{G}, \mathbf{C}) = \text{co}\Big(\bigcup_{\mathbf{D} \in \mathcal{D}(\mathcal{G})} \mathscr{R}_{\text{CCC}}(\mathcal{G}, \mathbf{C}, \mathbf{D}) \Big) \tag{3.53}$$

where $\text{co}(\cdot)$ denotes the convex hull, and $\mathscr{R}_{\text{CCC}}(\mathcal{G}, \mathbf{C}, \mathbf{D})$ consists of all the rate tuples $\mathbf{R}$ such that

$$R_i = \sum_{P \subseteq N} R_i(P), \qquad \forall i \in [n], \tag{3.54}$$

$$C_J = \sum_{P \subseteq N} C_J(P), \qquad \forall J \in N, \tag{3.55}$$

for some $R_i(P)$, $S_K(P)$, and $C_J(P) \geq 0$ such that

$$\sum_{j \in L} R_j(P) < \sum_{\substack{K \subseteq \Delta_i \cup A_i, \\ K \in \Gamma_*(P), \\ K \cap L \neq \emptyset}} S_K(P), \qquad \forall L \subseteq \Delta_i, \tag{3.56}$$

$$\sum_{K \in M} S_K(P) < \sum_{J \in \Gamma^*(M) \cap P} C_J(P), \qquad \forall M \subseteq \Gamma_*(P) \setminus 2^{A_i}, \tag{3.57}$$

for $i \in [n]$ such that $i \in \Delta_i$ and for $J \in N$ such that $J \in P$. Otherwise, set $R_i(P) = 0$ for $i \notin \Delta_i$ and set $C_J(P) = 0$ for $J \notin P$.

Remark 3.11. Compared with the CCC bound in Proposition 3.3, the DCC bound in Theorem 3.8 is more general in two aspects. First, as mentioned earlier in Remark 3.8, our coding scheme allows more degrees of freedom in choosing the decoding server groups $P_i \subseteq N$ that are receiver-dependent. Second, more subtly, our coding scheme requires the fractional link capacity constraints to be satisfied on average over $\mathbf{D}$ (cf. (3.52)), whereas CCC in Li et al. [2018] requires the link capacity constraints to be satisfied for each $\mathbf{D}$ (cf. (3.57)). As illustrated by Examples 3.13 and 3.15, respectively, flexibility in choosing different decoding server groups or averaging fractional link capacities over different decoding configurations can strictly increase the achievable rates.

Example 3.13. Consider the 5-message DIC problem $\mathcal{G}$:

$$(1|2,5), (2|3,4), (3|-), (4|2,5), (5|1,2,4)$$

with $C_J = 1$ for $J = J_1 = \{1,2,3\}$ and $J = J_2 = \{1,4\}$, $C_J = 2$ for $J = J_3 = \{1,3,4,5\}$, and $C_J = 0$ otherwise. Hence, the set of active servers is $N_A = \{J_1, J_2, J_3\}$. We use a single decoding message set tuple $\mathbf{D}$ with $D_2 = \{2\}$, $D_3 = \{3\}$, and $D_i = [n] \setminus A_i$ for $i = 1,4,5$. The sum-rate achievable by the CCC bound in Proposition 3.3, which is computed using (3.54)-(3.57) across all $2^3 - 1 = 7$ decoding server group tuples $\mathbf{P}$, $P_i = P \subseteq N_A$, $i \in [n]$, satisfies $R_1 + R_2 + R_3 + R_4 + R_5 < 6$. The sum-rate achievable by the DCC bound in Theorem 3.8, which is computed with variable P_i as a function of $i \in [n]$ using the following 7 randomly found decoding server group tuples, satisfies $R_1 + R_2 + R_3 + R_4 + R_5 < 7$. Therefore, there can be a benefit in allowing P_i to vary across i.

In $\mathbf{P}_1$, we set $P_1 = \{J_1, J_2\}$, $P_2 = \{J_3\}$, $P_3 = \{J_2\}$, $P_4 = \{J_2, J_3\}$, and $P_5 = N_A$.

In $\mathbf{P}_2$, we set $P_1 = P_2 = P_4 = \{J_1, J_2\}$, $P_3 = \{J_2\}$, and $P_5 = \{J_2, J_3\}$.

In $\mathbf{P}_3$, we set $P_1 = \{J_2, J_3\}$, $P_2 = P_5 = \{J_3\}$, $P_3 = \{J_1, J_2\}$, and $P_4 = \{J_2\}$.

In $\mathbf{P}_4$, we set $P_1 = \{J_1, J_3\}$, $P_2 = \{J_1, J_2\}$, $P_3 = \{J_2\}$, $P_4 = \{J_3\}$, and $P_5 = \{J_1\}$.

In $\mathbf{P}_5$, we set $P_1 = \{J_2, J_3\}$, $P_2 = N_A$, $P_3 = \{J_2\}$, $P_4 = \{J_1, J_3\}$, and $P_5 = \{J_2\}$.

In $\mathbf{P}_6$, we set $P_1 = \{J_2\}$, $P_2 = \{J_1, J_2\}$, $P_3 = \{J_2, J_3\}$, $P_4 = \{J_1, J_3\}$, and $P_5 = N_A$.

In $\mathbf{P}_7$, we set $P_1 = \{J_1, J_3\}$, $P_2 = \{J_2\}$, $P_3 = P_4 = \{J_1, J_3\}$, and $P_5 = N_A$.

We note that even with slight variations in the above 7 decoding server group tuples, one can still obtain $R_1 + R_2 + R_3 + R_4 + R_5 < 7$ with the DCC bound $\mathscr{R}_{\text{DCC}}(\mathcal{G})$. For example, if we keep $\mathbf{P}_2$ to $\mathbf{P}_7$ unchanged and in $\mathbf{P}_1$ we set $P_1 = \{J_3\}$, $P_2 = N_A$, $P_3 = \{J_1\}$, $P_4 = \{J_2, J_3\}$, and $P_5 = \{J_1, J_2\}$, we still obtain the same sum-rate lower bound. Applying the touch grouping PM bound in Corollary 4.2 (to be presented in Section 4.6) with the touch grouping $\mathcal{P}_{\text{t}} = \{T_{\{2,5\}} \cap N_A, T_{\{1,3,4\}} \cap N_A\}$ gives a matching upper bound on the sum capacity $\mathscr{C}_{\text{sum}}(\mathcal{G})$, thus establishing the sum capacity to be 7. Also note that the sum-rate of 7 can indeed be achieved by the following linear coding scheme. Set the codeword length $r_{J_1} = r_{J_2} = 1$, $r_{J_3} = 2$ and message length $t_1 = t_2 = 1$, $t_3 = 0$ (i.e., R_3 will be zero in this linear coding scheme), $t_4 = 3$, and $t_5 = 2$. We generate the codewords y_J, $J \in N_A$ as

$$y_{J_1} = (x_1 \oplus x_2), \quad y_{J_2} = (x_1 \oplus x_{4,1}) \quad y_{J_3} = (x_{4,2} \oplus x_{5,1}, x_{4,3} \oplus x_{5,2}).$$

It can be verified that upon receiving the above codewords, every receiver i can decode x_i with the help of x_{A_i}. For the above coding scheme, rate tuple $\mathbf{R} = \mathbf{t} = (1, 1, 0, 3, 2)$ is achievable, thus achieving a sum-rate of $1 + 1 + 0 + 3 + 2 = 7$.

Example 3.14. In Example 3.13, we can compute the whole achievable rate region of DCC using FME. In our programs, there were 181 variables to eliminate, which was completed in a few minutes on an Apple iMac 4GHz Intel Core i7 with 16 GB memory and using Matlab® R2017b and the FME software Gattegno et al. [2015]. The achievable rate region is

$$\begin{aligned} R_2 &< 1, & R_4 &< 3, \quad R_5 < 2, \\ R_1 + R_2 + R_3 &< 4, & R_1 + R_3 + R_4 &< 4, \\ R_2 + R_3 + R_4 &< 4, & R_2 + R_3 + R_5 &< 3. \end{aligned}$$

Comparison of the DCC achievable rate region with that obtained using the single-server grouping PM bound in Corollary 4.4 (see Section 4.8 for details) shows the region is tight, thus establishing the capacity region for this problem.

Example 3.15. Consider the 5-message DIC problem $\mathcal{G}$:

$$(1|4), (2|1,3,4), (3|1,2,4), (4|1,3), (5|3)$$

with $C_J = 1$ for $J = \{1,2,5\}$, $\{1,2,3,5\}$, and $\{2,4,5\}$, and $C_J = 0$ otherwise. The sum-rate achievable by the CCC bound in Proposition 3.3, after taking the convex hull over all possible decoding server groups $P \subseteq N_A = \{\{1,2,5\},\{1,2,3,5\},\{2,4,5\}\}$ ($2^3 - 1 = 7$ possibilities) and $\mathbf{D} \in \mathcal{D}(\mathcal{G})$ ($2^{\sum_{i\in[n]}|B_i|} = 2^{10} = 1024$ possibilities according to (3.16)) satisfies $R_1 + R_2 + R_3 + R_4 + R_5 < 4$.

In contrast, if we set $R_i(\mathbf{P},\mathbf{D})$, $C_J(\mathbf{P},\mathbf{D})$, and $S_K(\mathbf{P},\mathbf{D})$ to zero in the DCC bound in Theorem 3.8 except one decoding server group tuple $\mathbf{P}$ with $P_i = N_A$, $i \in [n]$ and two decoding message set tuples $\mathbf{D}$ with $D_1 = \{1\}$, $D_5 = \{5\}$, and $D_i = [n] \setminus A_i$ for $i = 2,3,4$, and $\mathbf{D}'$ with $D_1 = \{1,2\}$, $D_5 = \{5\}$, and $D_i = [n] \setminus A_i$ for $i = 2,3,4$, then the achievable sum-rate satisfies $R_1 + R_2 + R_3 + R_4 + R_5 < 5$. This is strictly larger than that of the CCC bound. Applying the touch grouping PM bound in Corollary 4.2 (see Section 4.6) with the touch grouping $\mathcal{P}_\text{t} = \{T_{\{1,3\}} \cap N_A, T_{\{2,4,5\}} \cap N_A\}$ gives a matching upper bound on the sum capacity $\mathscr{C}(\mathcal{G})$, thus establishing the sum capacity to be 5. Also note that the sum-rate of 5 can indeed be achieved by the following linear coding scheme. Set the codeword length $r_J = 1$ for every $J \in N_A$ and message length $t_1 = t_3 = t_4 = 1$, $t_2 = 2$, $t_5 = 0$ (i.e., R_5 will be zero in this linear coding scheme). We generate the codewords y_J, $J \in N_A$ as

$$y_{\{1,2,5\}} = (x_1 \oplus x_{2,1}), \quad y_{\{1,2,3,5\}} = (x_{2,2} \oplus x_3), \quad y_{\{2,4,5\}} = (x_{2,1} \oplus x_4).$$

It can be verified that upon receiving the above codewords, every receiver i can decode x_i with the help of x_{A_i}. For the above linear coding scheme, rate tuple $\mathbf{R} = \mathbf{t} = (1,2,1,1,0)$ is achievable, thus achieving a sum-rate of $1+2+1+1+0 = 5$.

Just like Theorem 3.7, Proposition 3.2 and Corollaries 3.2 and 3.3 can be extended by fractional allocation of rates over decoding configurations. In the following, we present the extension of Corollary 3.3 as an illustration. Fix a single decoding server group tuple $\mathbf{P}$ with $P_i = N$ for all $i \in [n]$. Let $S_K(\mathbf{P},\mathbf{D}) = S_K(\mathbf{D}) = C_K(\mathbf{D})$, $K \in N$, where $C_K(\mathbf{D})$ is to be determined. This essentially prevents cooperation among the servers and turns the coding scheme to

$$w_{J,\mathbf{D}} : \prod_{j\in J}[2^{t_j(\mathbf{D})}] \to [2^{C_J(\mathbf{D})}], \qquad \forall J \in N, \mathbf{D} \in \mathcal{D}(\mathcal{G}),$$

where the final codeword y_J is the deterministic concatenation of composite index tuples, $(w_{J,\mathbf{D}}, \mathbf{D} \in \prod_{i=1}^n \mathcal{D}_i)$.

Corollary 3.4. Consider the DIC problem $\mathcal{G}$: $(i|j \in A_i)$, $i \in [n]$ with link capacity tuple $\mathbf{C}$.

A rate tuple $\mathbf{R}$ is achievable if

$$R_i = \sum_{\mathbf{D} \in \mathcal{D}(\mathcal{G})} R_i(\mathbf{D}), \qquad \forall i \in [n], \tag{3.58}$$

$$C_J = \sum_{\mathbf{D} \in \mathcal{D}(\mathcal{G})} C_J(\mathbf{D}), \qquad \forall J \in N, \tag{3.59}$$

for some $R_j(\mathbf{D})$ and $C_J(\mathbf{D}) \geq 0$ such that

$$\sum_{j \in L} R_j(\mathbf{D}) < \sum_{\substack{J \subseteq D_i \cup A_i, \\ J \cap L \neq \emptyset}} C_J(\mathbf{D}), \qquad \forall L \subseteq D_i,\ i \in [n]. \tag{3.60}$$

3.6 Chapter Summary

In this chapter, we investigated the coding schemes based on the two-layer random-coding-based composite coding scheme Arbabjolfaei et al. [2013] and characterized their corresponding achievable rate regions. For the CIC problem, we started with simplifying the composite coding scheme from two perspectives: reducing the number of composite indices for the each decoding configuration, and reducing the number of decoding configurations. Next, we moved on to extending the composite coding scheme to more general versions. Again we took two approaches. First, we introduced a more flexible fractional rate splitting method compared to standard time sharing, based on which we proposed the enhanced composite coding scheme. Second, we developed the three-layer composite coding scheme by adding one more layer of random coding into the original two-layer composite coding scheme. For the DIC problem, we generalized the centralized composite coding scheme to its distributed version. In particular, we combined our enhanced fractional rate splitting method with the key idea of cooperatively generating composite indices among all the servers introduced in Li et al. [2018]. We even added another degree of freedom into the distributed composite coding scheme by allowing the receivers to choose different groups of server outputs used for decoding.

Chapter 4

Performance Bounds

In this chapter, we study the information-theoretic performance bounds for index coding. Such bounds serve as converse results for the index coding problem, characterizing the fundamental limits on its broadcast rate and capacity region.

In Sections 4.1-4.3, we generalize the alignment chain model Maleki et al. [2014] (Definition 2.3) to derive a series of new performance bounds for the CIC problem, which are strictly tighter than the MAIS bound Bar-Yossef et al. [2011] (Proposition 2.5) and the internal conflict bound Maleki et al. [2014]; Jafar [2014] (Proposition 2.6), and at the same time, less computationally intensive than the PM bound Blasiak et al. [2011]; Arbabjolfaei et al. [2013] (Proposition 2.7). See Figure 4.1 for a visualized summary.

In Sections 4.4-4.9, we extend the acyclic chain bounds and the PM bound for the CIC problem to the multi-server scenario. We pay particular attention to the computational complexity of such extended performance bounds since they can grow very quickly as the number of servers involved in the system increases. The DIC performance bounds are summarized in Figure 4.2.

Some of the technical proofs are presented in Appendices B.1-B.8.

4.1 The Basic Acyclic Chain Bound

To motivate the acyclic chain bounds to be developed henceforth, consider the following observation on the MAIS bound β_{MAIS}.

As can be seen from Proposition 2.5, the MAIS bound β_{MAIS} depends on the acyclic structure among receivers' side information (cf. (2.11)), or equivalently, receivers' interfering messages (cf. (2.12)). For a given CIC problem $\mathcal{G}$ whose side information graph $\mathcal{G}$ is acyclic, $\beta_{\text{MAIS}}(\mathcal{G})$ coincides with the broadcast rate $\beta(\mathcal{G})$ and thus is tight. Otherwise, when there are multiple maximal acyclic sets within $\mathcal{G}$, $\beta_{\text{MAIS}}(\mathcal{G})$ depends only on one such acyclic structure with the largest size and may be loose for the problem.

Hence, a natural question arises: is it possible to develop certain performance bound that not only depends on the largest acyclic set, but can take the internal structure among multiple

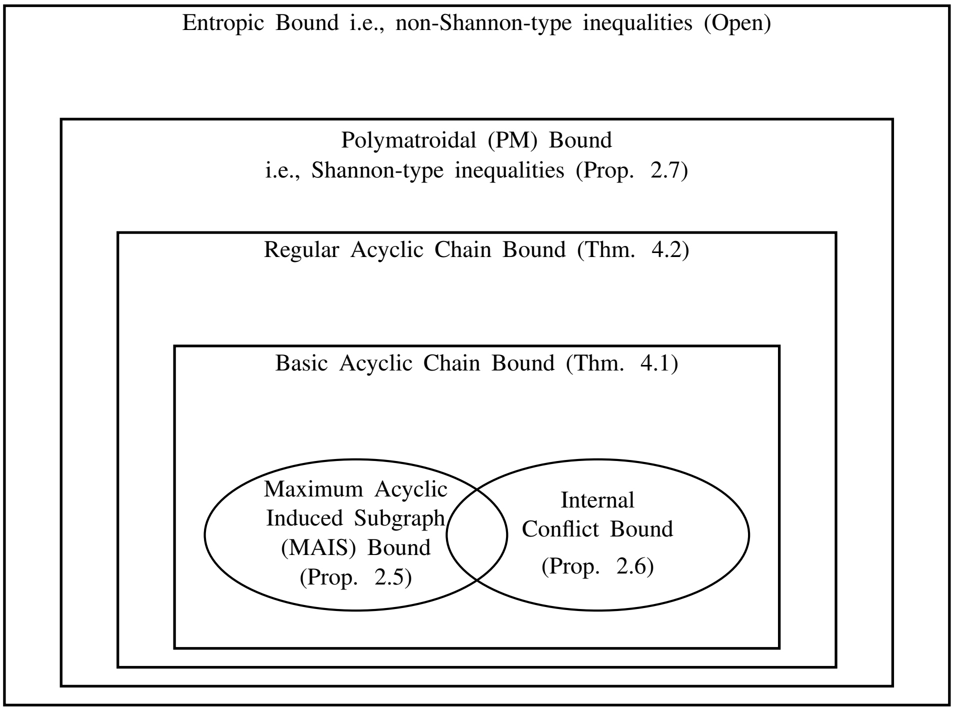

Entropic Bound i.e., non-Shannon-type inequalities (Open)
Polymatroidal (PM) Bound
i.e., Shannon-type inequalities (Prop. 2.7)
Regular Acyclic Chain Bound (Thm. 4.2)
Basic Acyclic Chain Bound (Thm. 4.1)
Maximum Acyclic Induced Subgraph (MAIS) Bound (Prop. 2.5)
Internal Conflict Bound (Prop. 2.6)

Figure 4.1: Summary of the performance bounds for the CIC problem. Bound A being contained within bound B means that bound A is a special case of bound B.

acyclic structures into account. In the following, we answer such question positively by showing that converse results more stringent than the MAIS bound can be achieved if the multiple acyclic structures jointly satisfy certain concatenative relationship. That is, we define a series of chain models, namely acyclic chains, formed via the concatenation of multiple acyclic structures.

We start by defining the basic building block of the acyclic chains, the *basic tower*.

Definition 4.1 (Basic tower)**.** For the CIC problem $\mathcal{G} : (i|j \in A_i), i \in [n]$, nonempty message sets $I, I', K_1, K_2, \dots, K_h \subseteq [n]$, constitute the following basic tower $\mathcal{B}$,

$$\mathcal{B}: \quad I \overset{\substack{K_h \\ \vdots \\ K_2 \\ K_1}}{\longleftrightarrow}_{\mathrm{b}} I',$$

if $I \cup I' \cup K_1 \cup K_2 \cup \cdots \cup K_{\ell-1} \subseteq B_{K_\ell}$ for any $\ell \in [h]$.

See Fig. 4.3(a) for a visualization of Definition 4.1. In the basic tower $\mathcal{B}$, message sets I and I' are placed horizontally at the *ground level* of the tower, and message set K_ℓ is placed

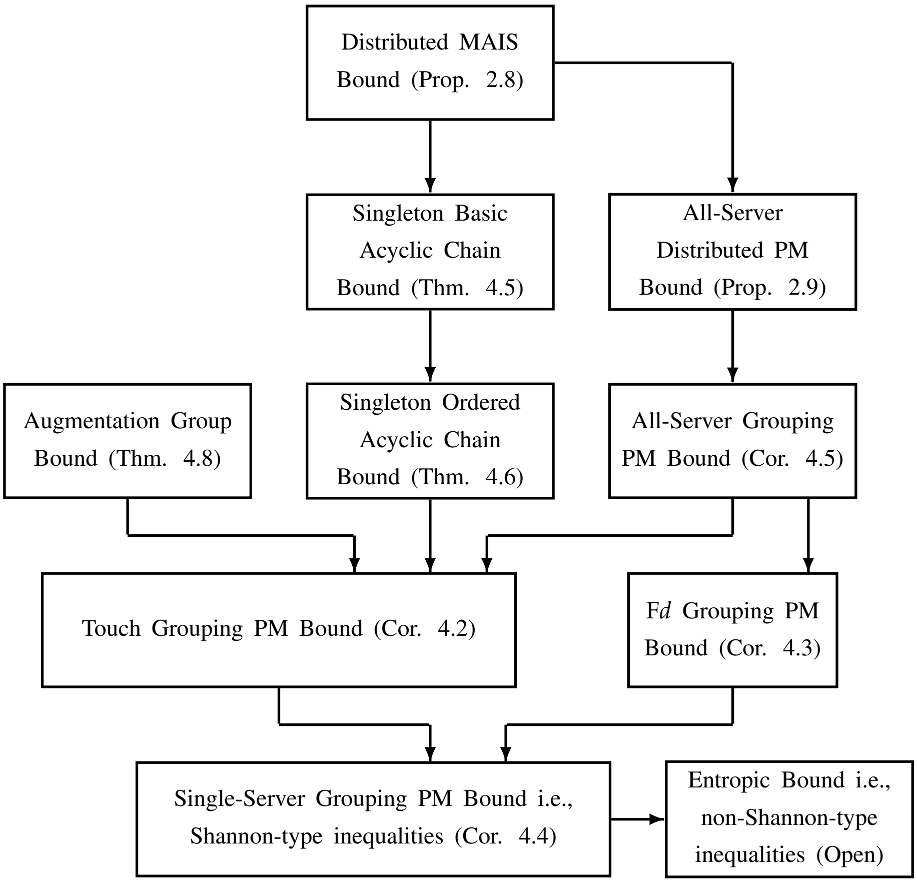

Distributed MAIS Bound (Prop. 2.8)
Singleton Basic Acyclic Chain Bound (Thm. 4.5)
All-Server Distributed PM Bound (Prop. 2.9)
Augmentation Group Bound (Thm. 4.8)
Singleton Ordered Acyclic Chain Bound (Thm. 4.6)
All-Server Grouping PM Bound (Cor. 4.5)
Touch Grouping PM Bound (Cor. 4.2)
Fd Grouping PM Bound (Cor. 4.3)
Single-Server Grouping PM Bound i.e., Shannon-type inequalities (Cor. 4.4)
Entropic Bound i.e., non-Shannon-type inequalities (Open)

Figure 4.2: Summary of the performance bounds for the DIC problem. A directed path from bound A to B means that bound A is a special case of bound B. Note that the bounds in Corollary 4.2-4.5 are all special forms of the grouping PM bound in Theorem 4.7.

on the ℓ-th *floor* for any $\ell \in [h]$, where h is called the *height* of the tower. The receivers corresponding to the message set on a higher floor must have all messages in the message sets located on lower floors, including the ground level, in their common interfering message set (cf. (2.1)). As a result, the message sets $I, K_1, K_2, \ldots, K_h$ form an acyclic structure *at the set level* as defined in (2.15), and so do the message sets $I', K_1, K_2, \ldots, K_h$.

For some nonempty message sets $I, I', K_1, K_2, \ldots, K_h$, when there is no ambiguity, we simply say that they form a basic tower if for any $\ell \in [h]$, $I \cup I' \cup K_1 \cup \cdots \cup K_{\ell-1} \subseteq B_{K_\ell}$. For any basic tower $\mathcal{B}$, we also use $\mathcal{B}$ to denote the union of all the message sets within it.

Example 4.1. Consider the 5-message CIC problem $\mathcal{G}$:

$$(1|2,5),(2|1,3),(3|2,4),(4|3,5),(5|1,4),$$

which has previously been discussed in Examples 2.4 and 2.5. The problem can be equivalently

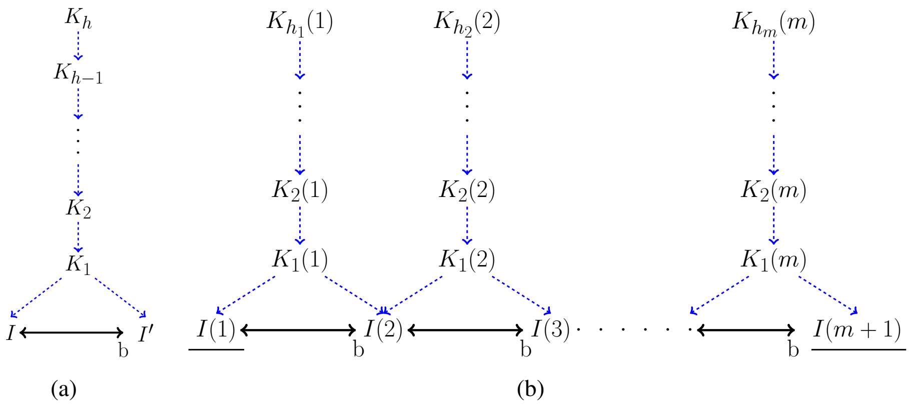


Figure 4.3: Schematic graphs for (a) Definition 4.1 and (b) Definition 4.2. To help with understanding, we draw blue dashed arrows such that if there is a directed path formed by dashed arrows of the same color from message set Q to P, then $P \subseteq B_Q$.

represented in the format of $(i\|j \in B_i)$[1], $i \in [n]$ based on the receivers' interfering message sets as

$$(1\|3,4), (2\|4,5), (3\|1,5), (4\|1,2), (5\|2,3).$$

According to Definition 4.1, we have the following five basic towers,

$$\{3\} \overset{\{1\}}{\longleftrightarrow}_{\mathrm{b}} \{4\}, \qquad \{4\} \overset{\{2\}}{\longleftrightarrow}_{\mathrm{b}} \{5\}, \qquad \{5\} \overset{\{3\}}{\longleftrightarrow}_{\mathrm{b}} \{1\},$$
$$\{1\} \overset{\{4\}}{\longleftrightarrow}_{\mathrm{b}} \{2\}, \qquad \{2\} \overset{\{5\}}{\longleftrightarrow}_{\mathrm{b}} \{3\}.$$

In the following we present the *basic acyclic chain* model built upon the basic tower in Definition 4.1.

Definition 4.2 (Basic acyclic chain)**.** For the CIC problem $\mathcal{G} : (i|j \in A_i)$, $i \in [n]$, we have the basic acyclic chain, Ch_{b}, of length m, constituted by nonempty message sets as

$$Ch_{\mathrm{b}}: \quad \underline{I(1)} \overset{\substack{K_{h_1}(1)\\ \cdots \\ K_2(1)\\ K_1(1)}}{\longleftrightarrow}_{\mathrm{b}} I(2) \overset{\substack{K_{h_2}(2)\\ \cdots \\ K_2(2)\\ K_1(2)}}{\longleftrightarrow}_{\mathrm{b}} \cdots \overset{\substack{K_{h_m}(m)\\ \cdots \\ K_2(m)\\ K_1(m)}}{\longleftrightarrow}_{\mathrm{b}} \underline{I(m+1)},$$

if the conditions listed below are satisfied:

1. $I(1) \subseteq B_{I(m+1)}$ or $I(m+1) \subseteq B_{I(1)}$;
2. For any $j \in [m]$, message sets $I(j), I(j+1), K_1(j), K_2(j), \ldots, K_{h_j}(j)$ form a basic tower $\mathcal{B}_j$, i.e., for any $\ell \in [h_j]$, $I(j) \cup I(j+1) \cup K_1(j) \cup K_2(j) \cup \cdots \cup K_{\ell-1}(j) \subseteq B_{K_\ell(j)}$.

[1] Note the difference between the two notations $(i|j \in A_i)$ and $(i\|j \in B_i)$ to avoid ambiguity.

In the above definition, the two terminal message sets $I(1)$ and $I(m+1)$ of the horizontal chain are underlined to indicate that they are acyclic at the set level as specified in Condition 1. We call the edge between message sets $I(j)$ and $I(j+1)$ edge j. A basic acyclic chain of length m has m edges, and can be seen as a horizontal concatenation (with overlapping) of m basic towers, $\mathcal{B}_1, \cdots, \mathcal{B}_m$, such that the terminal message sets $I(1)$ and $I(m+1)$ are acyclic at the set level. For the basic acyclic chain and any other acyclic chain models to be developed later, to avoid trivial cases, we always require that the length m is no less than 1.

For the CIC problem $\mathcal{G}$, let $\mathfrak{C}_{\mathrm{b}}(\mathcal{G})$ denote the collection of its basic acyclic chains. Obviously any alignment chain as defined in Definition 2.3 can be viewed as a special basic acyclic chain constituted by basic towers of height 1, each of which is formed by singleton message sets. It can be easily shown that for a given problem $\mathcal{G}$ there exists no valid basic acyclic chain if and only if there exists no alignment chain. That is, $\mathfrak{C}_{\mathrm{b}}(\mathcal{G}) = \emptyset$ if and only if $\mathfrak{C}_{\mathrm{AC}}(\mathcal{G}) = \emptyset$, or equivalently, the problem $\mathcal{G}$ is half-rate-feasible (cf. Remark 2.3).

We present the following iterative lower bound on the broadcast rate for the CIC problem.

Theorem 4.1 (Basic acyclic chain bound)**.** For the CIC problem $\mathcal{G} : (i|j \in A_i)$, $i \in [n]$, we have

$$\beta(\mathcal{G}) \geq \Gamma_{\mathrm{b}}(\mathcal{G}) \doteq \max_{Ch_{\mathrm{b}} \in \mathfrak{C}_{\mathrm{b}}(\mathcal{G})} \Gamma_{\mathrm{b}}(Ch_{\mathrm{b}}),$$

where for any basic acyclic chain $Ch_{\mathrm{b}} \in \mathfrak{C}_{\mathrm{b}}(\mathcal{G})$,

$$\Gamma_{\mathrm{b}}(Ch_{\mathrm{b}}) \doteq \frac{1}{m}\Big(\sum_{j\in[m]}\sum_{\ell\in[h_j]} \Gamma_{\mathrm{b}}(\mathcal{G}|_{K_\ell(j)}) + \sum_{j\in[m+1]} \Gamma_{\mathrm{b}}(\mathcal{G}|_{I(j)})\Big),$$

with the following recursion termination condition:

$$\Gamma_{\mathrm{b}}(\mathcal{G}|_S) \doteq \beta_{\mathrm{MAIS}}(\mathcal{G}|_S), \quad \text{if } S \text{ is acyclic or half-rate-feasible.}$$

We omit Theorem 4.1's proof since it can be seen a special case of Theorem 4.2 to be presented in Section 4.2 and proved in Appendix B.1.

It can be verified that the basic acyclic chain bound on the broadcast rate β is always no looser than the MAIS bound β_{MAIS} in Proposition 2.5 or the internal conflict bound on β_{AC} in Proposition 2.6. We formalize such relationships in the following proposition.

Proposition 4.1. For the CIC problem $\mathcal{G}$, we have

$$\Gamma_{\mathrm{b}}(\mathcal{G}) \geq \max\{\beta_{\mathrm{MAIS}}(\mathcal{G}), \beta_{\mathrm{AC}}(\mathcal{G})\}.$$

Proof. Since any alignment chain is a special basic acyclic chain, it is clear that $\Gamma_{\mathrm{b}}(\mathcal{G}) \geq \beta_{\mathrm{AC}}(\mathcal{G})$. It remains to show that $\Gamma_{\mathrm{b}}(\mathcal{G}) \geq \beta_{\mathrm{MAIS}}(\mathcal{G})$. Set $s = \beta_{\mathrm{MAIS}}(\mathcal{G})$, then there exists an acyclic message set with cardinality of s, say $\{i_1, i_2, \cdots, i_s\}$. In other words (cf. (2.12)),

$$\{i_1, \cdots, i_{\ell-1}\} \subseteq B_{i_\ell}, \qquad \forall \ell \in [s]. \tag{4.1}$$

Therefore, we have the following one-edge basic acyclic chain,

$$Ch_{\mathrm{b}}: \quad \underline{\{i_1\}} \xleftrightarrow{\substack{\{i_s\} \\ \{i_{s-1}\} \\ \cdots \\ \{i_3\}}}_{\mathrm{b}} \underline{\{i_2\}}, \tag{4.2}$$

and thus by Theorem 4.1, we have $\Gamma_{\mathrm{b}}(\mathcal{G}) \geq \Gamma_{\mathrm{b}}(Ch_{\mathrm{b}}) = s = \beta_{\mathrm{MAIS}}(\mathcal{G})$. Indeed, any acyclic set forms a one-edge basic acyclic chain, from which it is clear that the basic acyclic chain bound on the broadcast rate is always no looser than the MAIS bound. □

The relationships in Proposition 4.1 can be strict.

Example 4.2. Consider the 6-message CIC problem $\mathcal{G}$:

$$\begin{gathered}(1|2,3,4,6), \quad (2|4,5,6), \quad (3|1,2,4,5,6) \\ (4|1,2,6), \quad (5|2,3,4,6), \quad (6|-).\end{gathered}$$

For this problem, $\beta_{\mathrm{AC}}(\mathcal{G}) = \beta_{\mathrm{MAIS}}(\mathcal{G}) = 3$. However, we have the following basic acyclic chain of length $m = 2$,

$$Ch_{\mathrm{b}}: \quad \underline{\{1\}} \xleftrightarrow{\substack{\{6\} \\ \{2\}}}_{\mathrm{b}} \{3\} \xleftrightarrow{\substack{\{6\} \\ \{4\}}}_{\mathrm{b}} \underline{\{5\}},$$

and thus by Theorem 4.1 we have $\beta(\mathcal{G}) \geq \Gamma_{\mathrm{b}}(\mathcal{G}) \geq \Gamma(Ch_{\mathrm{b}})$ where

$$\begin{aligned}\Gamma(Ch_{\mathrm{b}}) &= \frac{1}{2}\Big(\sum_{j\in[2]}\sum_{\ell\in[h_j]} \Gamma_{\mathrm{b}}(\mathcal{G}|_{K_\ell(j)}) + \sum_{j\in[3]} \Gamma_{\mathrm{b}}(\mathcal{G}|_{I_\ell(j)})\Big) \\ &= \frac{1}{2}(\Gamma_{\mathrm{b}}(\{2\}) + \Gamma_{\mathrm{b}}(\{6\}) + \Gamma_{\mathrm{b}}(\{4\}) + \Gamma_{\mathrm{b}}(\{6\}) + \Gamma_{\mathrm{b}}(\{1\}) + \Gamma_{\mathrm{b}}(\{3\}) + \Gamma_{\mathrm{b}}(\{5\})) \\ &= \frac{1}{2}(1+1+1+1+1+1+1) = 3.5,\end{aligned}$$

which matches the composite coding (CC) bound in Proposition 2.4. Therefore, for this problem, we have

$$\beta(\mathcal{G}) = \Gamma_{\mathrm{b}}(\mathcal{G}) = 3.5 > \beta_{\mathrm{AC}}(\mathcal{G}) = \beta_{\mathrm{MAIS}}(\mathcal{G}) = 3.$$

Note that in Ch_{b}, message set $\{6\}$ appears twice in two different basic towers, which is allowed by Definition 4.2.

Example 4.3. First consider the following 5-message CIC problem $\mathcal{G}_0$:

$$(1|2,5),(2|1,3),(3|2,4),(4|3,5),(5|1,4), \tag{4.3}$$

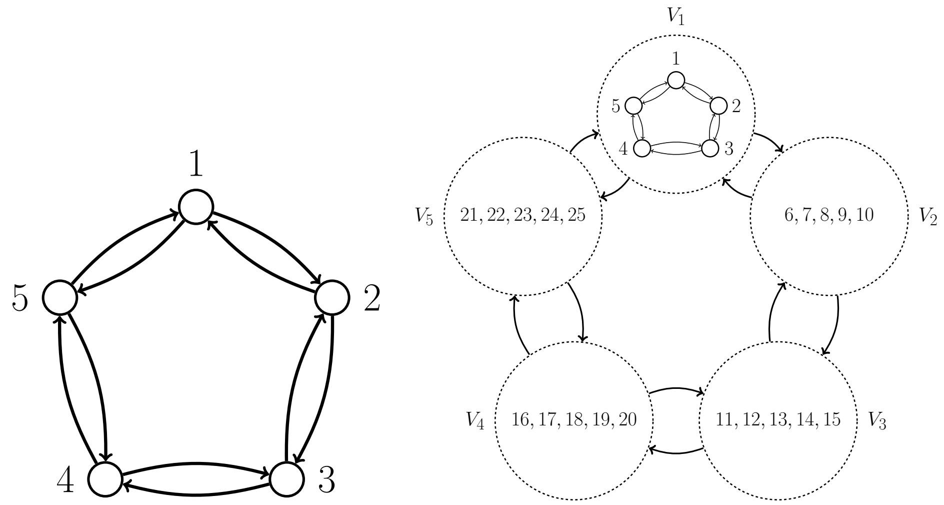
whose side information graph is shown in Figure 4.4(a). Note that this problem has been studied in Examples 2.4, 2.5, and 4.1. We have $\beta_{\mathrm{MAIS}}(\mathcal{G}_0) = 2$, which is strictly looser than $\beta(\mathcal{G}_0) = \beta_{\mathrm{AC}}(\mathcal{G}_0) = \Gamma_{\mathrm{b}}(\mathcal{G}_0) = 2.5$ given by the basic acyclic chain

$$Ch_{\mathrm{b}}: \quad \underline{\{1\}} \xleftrightarrow{\{4\}}_{\mathrm{b}} \{2\} \xleftrightarrow{\{5\}}_{\mathrm{b}} \underline{\{3\}}. \tag{4.4}$$

Now consider the 25-message CIC problem $\mathcal{G}$ whose side information graph is the lexicographic square of $\mathcal{G}_0$, i.e., $\mathcal{G} = \mathcal{G}_0^{\circ 2} = \mathcal{G} \circ \mathcal{G}$ (cf. Definition 2.1). For easier reference, set

$$V_i = \{5i-4, 5i-3, 5i-2, 5i-1, 5i,\}, \qquad \forall i \in [5].$$

and hence for $\mathcal{G}$, $[n] = \bigcup_{j\in[i]} V_i$. Note that for any $i \in [5]$, the subproblem $\mathcal{G}|_{V_i}$ can be seen as just a copy of the 5-message problem $\mathcal{G}_0$. A simplified version of $\mathcal{G}$ is shown in Figure 4.4(b). Since both the MAIS bound and the broadcast rate are multiplicative under the lexicographic product of CIC side information graphs (see Blasiak et al. [2011],[Arbabjolfaei and Kim, 2018, Sections 4.3 and 5.1]), we have $\beta_{\mathrm{MAIS}}(\mathcal{G}) = (\beta_{\mathrm{MAIS}}(\mathcal{G}_0))^2 = 4$ strictly smaller than $\beta(\mathcal{G}) = (\beta(\mathcal{G}_0))^2 = 6.25$. It can also be verified via exhaustive search that the



(a) (b)

Figure 4.4: (a) The side information graph of $\mathcal{G}_0$ in (4.3). (b) A simplified version of the side information graph of $\mathcal{G} = \mathcal{G}_0^{\circ 2}$. We only draw $\mathcal{G}|_{V_1}$ in detail. An edge from a dashed circle V_i to another dashed circle V_j means that there is an edge from every node in V_i to every node in V_j.

internal conflict bound gives $\beta_{\mathrm{AC}}(\mathcal{G}) = 4.5$ which is also loose. Nevertheless, the broadcast rate of 6.25 can indeed be obtained using the basic acyclic chain bound given that we have

$$Ch_{\mathrm{b}}^{\circ}: \quad \underline{\{V_1\}} \xleftrightarrow{\{V_4\}}_{\mathrm{b}} \{V_2\} \xleftrightarrow{\{V_5\}}_{\mathrm{b}} \underline{\{V_3\}}, \tag{4.5}$$

and thus we have

$$\Gamma_{\mathrm{b}}(\mathcal{G}) \geq \Gamma_{\mathrm{b}}(Ch_{\mathrm{b}}^{\circ}) = \frac{1}{2} \sum_{i \in [5]} \Gamma_{\mathrm{b}}(V_i) = \frac{1}{2} * 5 * \Gamma_{\mathrm{b}}(\mathcal{G}_0) = 6.25.$$

4.2 The Regular Acyclic Chain Bound

To present a more general acyclic chain model and its corresponding performance bound, we first build the *regular tower* as follows.

Definition 4.3 (Regular tower)**.** For the CIC problem $\mathcal{G} : (i|j \in A_i)$, $i \in [n]$, we have the following regular tower, $\mathcal{X}_j$, constituted by nonempty message sets as

$$\mathcal{X}_j : \quad I(1) \xleftrightarrow{\substack{K_{h_1}(1) \\ \cdots \\ K_2(1) \\ K_1(1)}}_{\mathrm{b}} \cdots \xleftrightarrow{\substack{K_{h_{j-1}}(j-1) \\ \cdots \\ K_2(j-1) \\ K_1(j-1)}}_{\mathrm{b}} I(j) \xleftrightarrow{\substack{K_{h_j}(j) \\ \cdots \\ K_2(j) \\ K_1(j)}}_{\mathrm{c}} I(j+1) \xleftrightarrow{\substack{K_{h_{j+1}}(j+1) \\ \cdots \\ K_2(j+1) \\ K_1(j+1)}}_{\mathrm{b}} \cdots \xleftrightarrow{\substack{K_{h_q}(q) \\ \cdots \\ K_2(q) \\ K_1(q)}}_{\mathrm{b}} I(q+1),$$

if the conditions listed below are satisfied:

1. for any $j' \in [q] \setminus \{j\}$, message sets $I(j'), I(j'+1), K_1(j'), K_2(j'), \ldots, K_{h_{j'}}(j')$ form a basic tower $\mathcal{B}_{j'}$.
2. for any $\ell \in [h_j]$ there exist two integers $s_{\ell,j} \in [j]$ and $t_{\ell,j} \in [j+1 : q+1]$ such that
   (a) $K_1(j) \cup K_2(j) \cup \cdots \cup K_{\ell-1}(j) \cup I(s_{\ell,j}) \cup I(t_{\ell,j}) \subseteq B_{K_\ell(j)}$;
   (b) for any $\ell_1 \neq \ell_2 \in [h_j]$, if $s_{\ell_1,j} \leq s_{\ell_2,j}$ then $t_{\ell_1,j} \geq t_{\ell_2,j}$;
   (c) there exists some $\theta_j, \iota_j \in [h_j]$ such that $s_{\theta_j,j} = 1$, $t_{\theta_j,j} = q+1$, and $s_{\iota_j,j} = j$, $t_{\iota_j,j} = j+1$.

For the regular tower $\mathcal{X}_j$ defined above, it has $q = q+1-1 = t_{\theta_j,j} - s_{\theta_j,j}$ edges. We call edge j the *central* edge, and the collection of message sets $I(j), I(j+1), K_1(j)$, $K_2(j), \ldots, K_{h_j}(j)$ the *core*. Every other edge $j' \in [q] \setminus \{j\}$ corresponds to a collection of message sets $I(j'), I(j'+1), K_1(j'), K_2(j'), \ldots, K_{h_{j'}}(j')$ that form a basic tower $\mathcal{B}_{j'}$. Note that we use different subscripts for the edges (c for the central edge and b for the other edges) in the horizontal chain in Definition 4.3 to distinguish the two different types of edges.

Conditions 2b and 2c in Definition 4.3 are described as follows. In the core, message set $K_\ell(j)$ on floor ℓ has message sets $I(s_{\ell,j})$ and $I(t_{\ell,j})$ to *start* and *terminate* its *coverage*

$G_{\ell,j} \doteq [s_{\ell,j} : t_{\ell,j} - 1]$, respectively. For any two message sets in the core on different floors $\ell_1 \neq \ell_2 \in [h_j]$, we must have $G_{\ell_1,j} \subseteq G_{\ell_2,j}$ or $G_{\ell_2,j} \subseteq G_{\ell_1,j}$, i.e., one coverage is contained within the other. For any regular tower $\mathcal{X}_j$ with central edge j, all of its edges are within its *total coverage* $G_j \doteq G_{\theta_j,j} = [s_{\theta_j,j} : t_{\theta_j,j} - 1]$.[2]

Note that any basic tower $\mathcal{B}_{j'}$ can be seen as a special regular tower with only the central edge j' and the core, for which $s_{\ell,j'} = j'$, $t_{\ell,j'} = j' + 1$ for any $\ell \in [h_{j'}]$. For any basic tower $\mathcal{B}_{j'}$, $|G_{j'}| = 1$, and for any regular tower that is not a basic tower $\mathcal{X}_j$, $|G_j| \geq 2$. In the rest of the thesis, unless otherwise specified, whenever we say regular tower we assume that it is not a basic tower.

Similar to basic tower, for any regular tower $\mathcal{X}_j$ we simply use $\mathcal{X}_j$ to denote the union of all the message sets within it.

For visualization of Definition 4.3, see Figure 4.5. To avoid clutter, we only draw $\mathcal{B}_1$ and the core of central edge j.

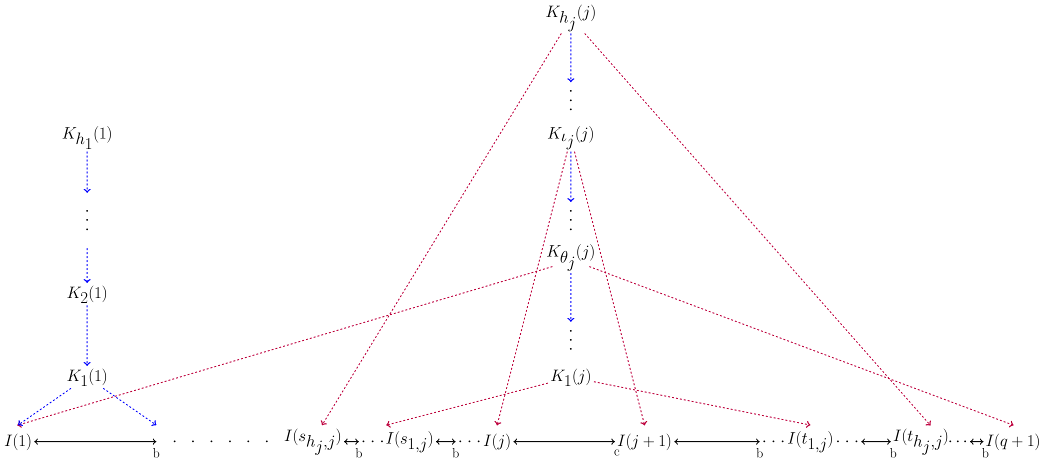


Figure 4.5: A visualization example for the regular tower in Definition 4.3. To help with understanding, we draw blue and purple dashed arrows. If there is a directed path formed by dashed arrows of the same color from message set Q to P, then $P \subseteq B_Q$. Note that for the positions of the $I(s_{\ell,j})$ and $I(t_{\ell,j})$ message sets on the horizontal chain for any $\ell \in [h_j]$, we do not impose any symmetric distance requirements such as $j - s_{\ell,j} = t_{\ell,j} - (j+1)$. They only need to satisfy the regularity conditions, i.e., Conditions 2b and 2c, in Definition 4.3.

Below we present our most general acyclic chain model, namely, the *regular acyclic chain*, which is built upon the basic and regular towers.

Definition 4.4 (Regular acyclic chain)**.** For the CIC problem $\mathcal{G} : (i|j \in A_i), i \in [n]$, we have

[2] We call $\mathcal{X}_j$ a "regular" (in contrast to "irregular") tower as it has to satisfy the balanced requirements on the relative positions of $I(s_{\ell,j})$ and $I(t_{\ell,j})$ in the horizontal chain as specified in Conditions 2b and 2c in Definition 4.3

the regular acyclic chain, Ch, of length m, constituted by nonempty message sets as

$$Ch: \quad \underline{I(1)} \xleftrightarrow{\substack{K_{h_1}(1)\\ \cdots \\ K_2(1)\\ K_1(1)}} I(2) \xleftrightarrow{\substack{K_{h_2}(2)\\ \cdots \\ K_2(2)\\ K_1(2)}} \cdots \xleftrightarrow{\substack{K_{h_m}(m)\\ \cdots \\ K_2(m)\\ K_1(m)}} \underline{I(m+1)}, \tag{4.6}$$

if the conditions listed below are satisfied:

1. $I(1) \subseteq B_{I(m+1)}$ or $I(m+1) \subseteq B_{I(1)}$;
2. For every $j \in [m]$, message sets $I(j), I(j+1), K_1(j), K_2(j), \ldots, K_{[h_j]}(j)$ constitute either a basic tower $\mathcal{B}_j$ or the core of a regular tower $\mathcal{X}_j$;
3. For any $j_1 \neq j_2 \in M$, $G_{j_1} \cap G_{j_2} = \emptyset$, where $M \doteq \{j \in [m] : |G_j| \geq 2\}$ denotes the set of central edges of the regular towers within the chain.

We remove subscripts for the edges in the horizontal chain in the above definition as the positions of the basic and regular towers are flexible.

Define

$$M' \doteq [m] \setminus (\bigcup_{j \in M} G_j) \tag{4.7}$$

as the set of edges located outside the coverage of any regular tower. Then the regular acyclic chain can be seen as a horizontal concatenation of the non-overlapping regular towers $\mathcal{X}_j$, $j \in M$ and the basic towers $\mathcal{B}_{j'}$, $j' \in M'$, such that the terminals of the chain $I(1)$ and $I(m+1)$ are acyclic at the set level. For a visualization of Definition 4.4, see Fig. 4.6.

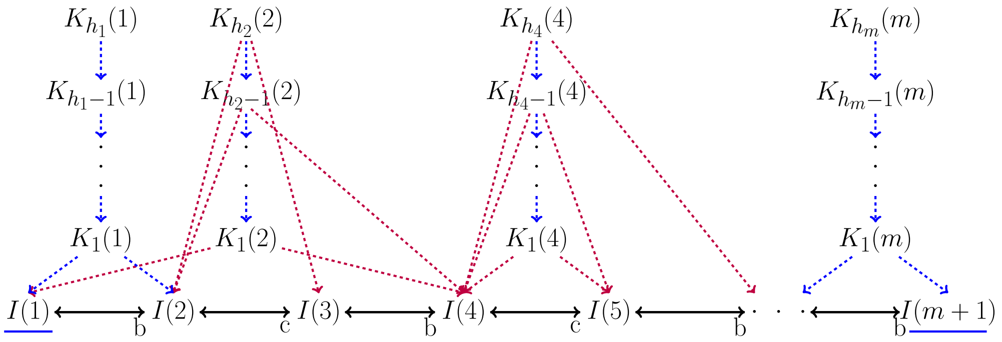


Figure 4.6: A visualization example for the regular acyclic chain in Definition 4.4. To help with understanding, we draw blue and purple dashed arrows. If there is a directed path formed by dashed arrows of the same color from message set Q to P, then $P \subseteq B_Q$.

For any basic tower, regular tower, basic acyclic chain, or regular acyclic chain, we call the I-labeled and K-labeled sets within it its *components*.

Given a problem $\mathcal{G}$, let $\mathfrak{C}(\mathcal{G})$ denote the collection of its regular acyclic chains. It can be verified that $\mathfrak{C}(\mathcal{G}) = \emptyset$ if and only if $\mathcal{G}$ is half-rate-feasible, in which case we also have $\mathfrak{C}_{\mathrm{b}}(\mathcal{G}) = \mathfrak{C}_{\mathrm{AC}}(\mathcal{G}) = \emptyset$. In other words, $\mathfrak{C}(\mathcal{G}) = \emptyset$, $\mathfrak{C}_{\mathrm{b}}(\mathcal{G}) = \emptyset$, $\mathfrak{C}_{\mathrm{AC}}(\mathcal{G}) = \emptyset$, and $\mathcal{G}$ being half-rate-feasible imply each other.

Theorem 4.2 (Regular acyclic chain bound)**.** For the CIC problem $\mathcal{G} : (i|j \in A_i), i \in [n]$, we have

$$\beta(\mathcal{G}) \geq \Gamma(\mathcal{G}) \doteq \max_{Ch \in \mathfrak{C}(\mathcal{G})} \Gamma(Ch),$$

where for any regular acyclic chain $Ch \in \mathfrak{C}(\mathcal{G})$,

$$\Gamma(Ch) \doteq \frac{1}{m}\Big(\sum_{j\in[m]}\sum_{\ell\in[h_j]} \Gamma(\mathcal{G}|_{K_\ell(j)}) + \sum_{j\in[m+1]} \Gamma(\mathcal{G}|_{I(j)})\Big),$$

with the following recursion termination condition:

$$\Gamma(\mathcal{G}|_S) \doteq \beta_{\mathrm{MAIS}}(\mathcal{G}|_S), \quad \text{if } S \text{ is acyclic or half-rate-feasible.}$$

We prove Theorem 4.2 in Appendix B.1 using mathematical induction.

Note that any basic acyclic chain is a special regular acyclic chain constituted by basic towers only with $M = \{j \in [m] : |G_j| \geq 2\} = \emptyset$. Consequently, the basic acyclic chain bound in Theorem 4.1 is implied by the regular acyclic chain bound in Theorem 4.2. Recall that the MAIS bound and the internal conflict bound are outperformed by the basic acyclic chain bound as shown in Proposition 4.1, and hence we have the following result.

Proposition 4.2. For the CIC problem $\mathcal{G}$, we have

$$\Gamma(\mathcal{G}) \geq \Gamma_{\mathrm{b}}(\mathcal{G}) \geq \max\{\beta_{\mathrm{MAIS}}(\mathcal{G}), \beta_{\mathrm{AC}}(\mathcal{G})\}.$$

The regular acyclic chain bound $\Gamma(\mathcal{G})$ in Theorem 4.2 can strictly outperform the basic acyclic chain bound $\Gamma_{\mathrm{b}}(\mathcal{G})$ in Theorem 4.1.

The following examples demonstrate the efficacy of the regular acyclic chain bound $\Gamma(\mathcal{G})$.

Example 4.4. Consider the 10-message CIC problem $\mathcal{G}$:

$$\begin{array}{cc}
(1|3,4,5,6,7,8,9,10), & (2|3,4,5,6,7,8,9,10) \\
(3|1,2,4,5,6,7,8,9,10), & (4|1,2,3,5,6,7,8,9,10), \\
(5|1,3,6,7,8,9,10), & (6|2,4,5,7,8,9,10), \\
(7|1,2,5,6,8,9,10), & (8|1,3,6,7,9,10), \\
(9|2,3,5,7,8,10), & (10|1,2,5,6,8,9).
\end{array}$$

For this problem, $\beta_{\mathrm{MAIS}}(\mathcal{G}) = \beta_{\mathrm{AC}}(\mathcal{G}) = \Gamma_{\mathrm{b}}(\mathcal{G}) = 3$. However, we have the following regular acyclic chain of length $m = 3$,

$$Ch: \quad \underline{\{1\}} \xleftrightarrow{\substack{\{9\}\\ \{6\}}}_{\mathrm{c}} \{3\} \xleftrightarrow{\substack{\{10\}\\ \{7\}}}_{\mathrm{b}} \{4\} \xleftrightarrow{\substack{\{8\}\\ \{5\}}}_{\mathrm{b}} \underline{\{2\}}.$$

Note that the above regular acyclic chain is not a basic acyclic chain due to the existence of the regular tower $\mathcal{X}_1$ whose central edge is edge 1, i.e., the edge between message sets $\{1\}$ and $\{3\}$. For $\mathcal{X}_1$, message set $\{6\}$ on the first floor of the core has message sets $\{1\}$ and $\{3\}$ to start and terminate its coverage, respectively, and message set $\{9\}$ on the second floor has message sets $\{1\}$ and $\{4\}$ to start and terminate its coverage, respectively, and thus $|G_1| = 2$. Given Ch, by Theorem 4.2 we have $\beta(\mathcal{G}) \geq \Gamma(\mathcal{G}) \geq \Gamma(Ch)$ where

$$\Gamma(Ch) = \frac{1}{3}\Big(\sum_{j\in[3]} \sum_{\ell\in[h_j]} \Gamma(\mathcal{G}|_{K_\ell(j)}) + \sum_{j\in[4]} \Gamma(\mathcal{G}|_{I(j)})\Big) = \frac{6+4}{3} = \frac{10}{3},$$

which matches the CC bound in Proposition 2.4. Therefore, for this problem, we have

$$\beta(\mathcal{G}) = \Gamma(\mathcal{G}) = \frac{10}{3} > \beta_{\mathrm{MAIS}}(\mathcal{G}) = \beta_{\mathrm{AC}}(\mathcal{G}) = \Gamma_{\mathrm{b}}(\mathcal{G}) = 3.$$

Example 4.5. Consider the 17-message CIC problem $\mathcal{G}$ as follows, which is denoted by $(i\|j \in B_i), i \in [n]$ rather than $(i|A_i), i \in [n]$ for brevity,

$$\begin{aligned}
&(1\|6), (2\|7,8), (3\|8,11,17), (4\|-), (5\|-), (6\|1), (7\|1,2),\\
&(8\|1,2,3,4,7), (9\|2,3), (10\|1,4,9), (11\|3,4,8), (12\|5,6),\\
&(13\|4,5), (14\|4,6,7,13,15,17), (15\|-), (16\|5,6,12), (17\|8).
\end{aligned}$$

For this problem, $\beta_{\mathrm{MAIS}}(\mathcal{G}) = \beta_{\mathrm{AC}}(\mathcal{G}) = \Gamma_{\mathrm{b}}(\mathcal{G}) = 3$. However, we have the following regular acyclic chain of length $m = 5$,

$$Ch: \quad \underline{\{1\}} \xleftrightarrow{\substack{\{8\}\\ \{7\}}}_{\mathrm{b}} \{2\} \xleftrightarrow{\substack{\{10\}\\ \{9\}}}_{\mathrm{c}} \{3\} \xleftrightarrow{\substack{\{11\}\\ \{8\}}}_{\mathrm{b}} \{4\} \xleftrightarrow{\substack{\{14\}\\ \{13\}}}_{\mathrm{c}} \{5\} \xleftrightarrow{\substack{\{16\}\\ \{12\}}}_{\mathrm{b}} \underline{\{6\}}.$$

Note that the above chain is not a basic acyclic chain due to the existences of the two regular towers $\mathcal{X}_2$ and $\mathcal{X}_4$, whose central edges are edge 2 that is between message sets $\{2\}$ and $\{3\}$, and edge 4 that is between message sets $\{4\}$ and $\{5\}$, respectively. For $\mathcal{X}_2$, its total coverage starts at message set $\{1\}$ and terminates at message set $\{4\}$, and thus $|G_2| = 3$. For $\mathcal{X}_4$, its total coverage starts at message set $\{4\}$ and terminates at message set $\{6\}$, and thus $|G_4| = 2$.

Given Ch, by Theorem 4.2 we have

$$\beta(\mathcal{G}) \geq \Gamma(\mathcal{G}) \geq \Gamma(Ch) = \frac{10+6}{5} = \frac{16}{5},$$

which matches the CC bound in Proposition 2.4. Therefore, for this problem, we have

$$\beta(\mathcal{G}) = \Gamma(\mathcal{G}) = \frac{16}{5} > \beta_{\mathrm{MAIS}}(\mathcal{G}) = \beta_{\mathrm{AC}}(\mathcal{G}) = \Gamma_{\mathrm{b}}(\mathcal{G}) = 3.$$

Example 4.6. Consider the 34-message CIC problem $\mathcal{G}$ as follows, which is denoted by $(i\|j \in B_i)$, $i \in [n]$ rather than $(i|j \in A_i)$, $i \in [n]$ for brevity,

$$\begin{aligned}
&(1\|3,23,24,31,32,33), (2\|4,5,23,24,31,32,33), (3\|4,5,23,24,31,32,33),\\
&(4\|1,2,23,24,31,32,33), (5\|2,3,23,24,31,32,33), (6\|8,9,16,17,29,30),\\
&(7\|9,10,16,17,29,30), (8\|6,10,16,17,29,30), (9\|6,7,16,17,29,30),\\
&(10\|7,8,16,17,29,30), (11\|13,14,17,18,29,30), (12\|14,15,17,18,29,30),\\
&(13\|11,15,17,18,29,30), (14\|11,12,17,18,29,30), (15\|12,13,17,18,29,30),\\
&(16\|18,29,30), (17\|29,30), (18\|16,29,30), (19\|20,21,30,31,32),\\
&(20\|19,21,30,31,32), (21\|22,30,31,32), (22\|19,21,30,31,32), (23\|24,30,34),\\
&(24\|23,30,34), (25\|33,34), (26\|25,33,34), (27\|25,26,33,34), (28\|25,26,27,33,34),\\
&(29\|34), (30\|-), (31\|32), (32\|31), (33\|-), (34\|29).
\end{aligned}$$

For this problem, $\beta_{\mathrm{AC}}(\mathcal{G}) = 3$ (c.f. Remark 2.4), and $\beta_{\mathrm{MAIS}}(\mathcal{G}) = 5$. However, we have the regular acyclic chain Ch of length $m = 4$ as:

$$\underline{\{29\}} \xleftrightarrow{\{6,7,8,\ldots,18\}}_{\mathrm{b}} \{30\} \xleftrightarrow{\{19,20,21,22\}}_{\mathrm{b}} \{31,32\} \xleftrightarrow{\overset{\{1,2,3,4,5\}}{\{23,24\}}}_{\mathrm{c}} \{33\} \xleftrightarrow{\{25,26,27,28\}}_{\mathrm{b}} \underline{\{34\}}.$$

Note that the above chain is not a basic acyclic chain due to the existences of the regular tower $\mathcal{X}_3$, whose central edges is edge 3 that is between message sets $\{31,32\}$ and $\{33\}$. The total coverage of $\mathcal{X}_3$ starts at message set $\{30\}$ and terminates at message set $\{34\}$, and thus $|G_3| = 3$. Given Ch, by Theorem 4.2 we have $\beta(\mathcal{G}) \geq \Gamma(\mathcal{G}) \geq \Gamma(Ch)$ where

$$\begin{aligned}
\Gamma(Ch) = &\frac{1}{4}(\Gamma(\{29\}) + \Gamma(\{30\}) + \Gamma(\{31,32\}) + \Gamma(\{33\}) + \Gamma(\{34\}) + \Gamma(\{6,7,\ldots,18\})\\
&+ \Gamma(\{19,20,21,22\}) + \Gamma(\{23,24\}) + \Gamma(\{1,2,3,4,5\}) + \Gamma(\{25,26,27,28\}))\\
= &\frac{1}{4}(1+1+2+1+1+4+3+2+4+2.5) \qquad (4.8)\\
= &\,5.375, \qquad (4.9)
\end{aligned}$$

where (4.8) follows from the following facts:

- For any subproblem $S \in \{\{29\},\{30\},\{31,32\},\{33\},\{34\},\{23,24\},\{25,26,27,28\}\}$, as S is acyclic, we simply have $\Gamma(S) = \beta_{\text{MAIS}}(S)$.
- For the subproblem $S = \{6,7,\dots,18\}$, by Definition 4.2 we have the following basic acyclic chain

$$\underline{\{16\}} \xleftrightarrow{\{6,7,8,9,10\}}_{\text{b}} \{17\} \xleftrightarrow{\{11,12,13,14,15\}}_{\text{b}} \underline{\{18\}}.$$

One can verify that the two subproblems $\{6,7,8,9,10\}$ and $\{11,12,13,14,15\}$ are both isomorphic to the CIC problem in (4.3) in Example 4.3. Hence, according to the results in Example 4.3, we have $\Gamma(\{6,7,8,9,10\}) = \Gamma(\{11,12,13,14,15\}) = 2.5$. Therefore, by Theorem 4.1, we have

$$\Gamma(S) = \frac{1}{2}(1+1+1+2.5+2.5) = 4.$$

It is worth noticing that for this subproblem $S = \{6,7,\dots,18\}$, we have $\beta_{\text{MAIS}}(S) = \beta_{\text{AC}}(S) = 3$, i.e., both the MAIS bound and the internal conflict bound are loose.

- For the subproblem $S = \{19,20,21,22\}$, one can verify that

$$\beta(S) = \Gamma(S) = \beta_{\text{MAIS}}(S) = 3.$$

- For the subproblem $S = \{1,2,3,4,5\}$, by Definition 4.2 we have the following basic acyclic chain

$$\underline{\{1\}} \xleftrightarrow{\{4\}}_{\text{b}} \{2\} \xleftrightarrow{\{5\}}_{\text{b}} \underline{\{3\}}.$$

Therefore, by Theorem 4.2, we have

$$\Gamma(S) = \frac{1}{2}(1+1+1+1+1) = 2.5.$$

By (4.9) we know that $\beta(\mathcal{G}) \geq 5.375$. It can be verified that this bound is indeed tight according to [Arbabjolfaei and Kim, 2018, Theorem 4.1, Remark 4.5]. See also [Arbabjolfaei and Kim, 2018, Sections 1.2 and 4.1] for more details.

Remark 4.1. Note that the basic and regular acyclic chain bounds are derived purely based on Shannon-type inequalities, and hence the PM bound in Proposition 2.7 is always no looser than them. Nevertheless, due to its high computational complexity[3], the PM bound is unfeasible for

[3] The LP for computing the PM bound involves a large number of variables and linear constraints, both growing exponentially in the problem size n.

most large CIC problems. In comparison, the complexities of computing $\Gamma_{\mathrm{b}}(\mathcal{G})$ and $\Gamma(\mathcal{G})$ are lower and thus affordable even for relatively large problems. Generally speaking, $\Gamma_{\mathrm{b}}(\mathcal{G})$ and $\Gamma(\mathcal{G})$ can be computed through exhaustive search over all possible chains in $\mathfrak{C}_{\mathrm{b}}(\mathcal{G})$ and $\mathfrak{C}(\mathcal{G})$, respectively. Efficient heuristic algorithms for computing or approximating $\Gamma_{\mathrm{b}}(\mathcal{G})$ and $\Gamma(\mathcal{G})$ remain to be designed in future work.

Remark 4.2. Another merit of the acyclic chain bounds over the PM bound is that the acyclic chain bounds provide a new way of deriving converse results for a given CIC problem through observing and exploiting the side information structure at receivers. In contrast, to obtain converse results using the PM bound one has to solve an LP.

4.3 Properties of the Acyclic Chain Bounds

We identify several important properties of the regular acyclic chain bound $\Gamma(\mathcal{G})$. Same properties hold for the less general basic acyclic bound $\Gamma_{\mathrm{b}}(\mathcal{G})$. We start with the following theorem.

Theorem 4.3. For the CIC problem $\mathcal{G} : (i|j \in A_i), i \in [n]$, for any sets $P, Q \subseteq [n]$ such that $P \subseteq B_Q$, we have

$$\Gamma(P \cup Q) \geq \Gamma(P) + \Gamma(Q). \tag{4.10}$$

Proof. Assume that for the subproblem P, the regular acyclic chain $Ch \in \mathfrak{C}(P)$:

$$\underline{I(1)} \xleftrightarrow{\substack{K_{h_1}(1)\\ \cdots \\ K_2(1)\\ K_1(1)}} I(2) \xleftrightarrow{\substack{K_{h_2}(2)\\ \cdots \\ K_2(2)\\ K_1(2)}} \cdots \xleftrightarrow{\substack{K_{h_m}(m)\\ \cdots \\ K_2(m)\\ K_1(m)}} \underline{I(m+1)},$$

satisfies that $\Gamma(P) = \Gamma(Ch)$. Since $P \subseteq B_Q$, we can construct a regular acyclic chain Ch^Q based on Ch as

$$Ch^Q : \quad \underline{I(1)} \xleftrightarrow{\substack{Q\\ K_{h_1}(1)\\ \cdots \\ K_1(1)}} I(2) \xleftrightarrow{\substack{Q\\ K_{h_2}(2)\\ \cdots \\ K_1(2)}} \cdots \xleftrightarrow{\substack{Q\\ K_{h_m}(m)\\ \cdots \\ K_1(m)}} \underline{I(m+1)},$$

such that $Ch^Q \in \mathfrak{C}(P \cup Q)$. By Theorem 4.2, we have

$$\begin{aligned}
\Gamma(Ch^Q) =& \frac{1}{m}\Big(\sum_{j\in[m]} \Gamma(Q) + \sum_{j\in[m]} \sum_{\ell\in[h_j]} \Gamma(K_\ell(j)) + \sum_{j\in[m+1]} \Gamma(I(j))\Big)\\
=& \Gamma(Q) + \Gamma(Ch)\\
=& \Gamma(Q) + \Gamma(P).
\end{aligned} \tag{4.11}$$

Since $Ch^{Q} \in \mathfrak{C}(P \cup Q)$, we have

$$\Gamma(P \cup Q) = \max_{Ch' \in \mathfrak{C}(P \cup Q)} \Gamma(Ch') \geq \Gamma(Ch^{Q}). \tag{4.12}$$

Combining (4.11) and (4.12) leads to (4.10). □

Recall that the MAIS bound and the broadcast rate are multiplicative under the lexicographic product of CIC side information graphs Blasiak et al. [2011]; Arbabjolfaei and Kim [2018]. That is, for two given CIC problems $\mathcal{G}_0$ and $\mathcal{G}_1$, we have

$$\beta_{\text{MAIS}}(\mathcal{G}_0 \circ \mathcal{G}_1) = \beta_{\text{MAIS}}(\mathcal{G}_0) \cdot \beta_{\text{MAIS}}(\mathcal{G}_1), \tag{4.13}$$

$$\beta(\mathcal{G}_0 \circ \mathcal{G}_1) = \beta(\mathcal{G}_0) \cdot \beta(\mathcal{G}_1). \tag{4.14}$$

A similar albeit weaker property has been shown for the PM bound [Arbabjolfaei and Kim, 2018, Proposition 5.4]: for two given CIC problems $\mathcal{G}_0$ and $\mathcal{G}_1$, we have

$$\beta_{\text{PM}}(\mathcal{G}_0 \circ \mathcal{G}_1) \geq \beta_{\text{PM}}(\mathcal{G}_0) \cdot \beta_{\text{PM}}(\mathcal{G}_1). \tag{4.15}$$

Similar to the PM bound $\beta_{\text{PM}}(\mathcal{G})$, the regular acyclic chain bound $\Gamma(\mathcal{G})$ has the following structural property.

Theorem 4.4. For two given problems $\mathcal{G}_0$ and $\mathcal{G}_1$, we have

$$\Gamma(\mathcal{G}_0 \circ \mathcal{G}_1) \geq \Gamma(\mathcal{G}_0) \cdot \Gamma(\mathcal{G}_1). \tag{4.16}$$

The proof of Theorem 4.4 is presented in Appendix B.2.

The corollary below states that the gap between the regular acyclic chain bound $\Gamma(\mathcal{G})$ and the MAIS bound $\beta_{\text{MAIS}}(\mathcal{G})$ can be magnified to a multiplicative factor, which grows polynomially in the problem size n.

Corollary 4.1. Let $\mathcal{G}_0$ be a CIC problem with $n_0 = |V(\mathcal{G}_0)|$ messages for which $\frac{\Gamma(\mathcal{G}_0)}{\beta_{\text{MAIS}}(\mathcal{G}_0)} = \rho > 1$. Consider the CIC problem $\mathcal{G} = \mathcal{G}^{\circ k}$ with $n = |V(\mathcal{G})| = n_0^k$ messages for any positive integer k, we have

$$\frac{\Gamma(\mathcal{G})}{\beta_{\text{MAIS}}(\mathcal{G})} \geq n^{\log_{n_0}(\rho)}.$$

Proof. We have

$$\frac{\Gamma(\mathcal{G})}{\beta_{\text{MAIS}}(\mathcal{G})} \geq \frac{(\Gamma(\mathcal{G}_0))^k}{(\beta_{\text{MAIS}}(\mathcal{G}_0))^k} = \rho^k = (n_0^{\log_{n_0}(\rho)})^k = n^{\log_{n_0}(\rho)},$$

where the first inequality follows from Theorem 4.4 and (4.13) □

4.4 Acyclic Chain Bounds for DIC

Now we move on to the distributed scenario and develop a series of performance bounds through generalizing the performance bounds for the CIC problem. Recall that the MAIS bound in Proposition 2.5 has already been extended to the multi-server case as the distributed MAIS bound in Proposition 2.8. In this section we investigate the converse results based on the acyclic chain models for the DIC problem.

Note that Definitions 4.1-4.4 do not depend on the server setup and thus also apply to the DIC problem. However, whether there exist some iterative DIC performance bounds as counterparts to Theorems 4.1-4.2 is unknown and remains to be investigated. In the following, we reduce the tower and acyclic chain models in Definitions 4.1-4.4 to their less general variants. Then we propose two DIC performance bounds associated to these less general acyclic chain models. It should be noted that these two bounds are non-iterative as the elements of the less general acyclic chain models are individual messages rather than sets of messages.

Definition 4.5 (Singleton basic tower)**.** For the CIC or DIC problem $\mathcal{G} : (i|j \in A_i), i \in [n]$, messages $i, i', k_1, k_2, \dots, k_h \subseteq [n]$, constitute the following singleton basic tower $\mathcal{B}^{\text{s}}$,

$$\mathcal{B}^{\text{s}}: \quad i \overset{\substack{k_h \\ \cdots \\ k_2 \\ k_1}}{\longleftrightarrow}_{\text{b}} i',$$

if $\{i, i', k_1, k_2, \dots, k_{\ell-1}\} \subseteq B_{k_\ell}$ for any $\ell \in [h]$.

Definition 4.6 (Singleton basic acyclic chain)**.** For the CIC or DIC problem $\mathcal{G} : (i|j \in A_i)$, $i \in [n]$, we have the singleton basic acyclic chain, Ch^{s}_{b}, of length m, constituted by individual messages as

$$Ch^{\text{s}}_{\text{b}}: \quad \underline{i(1)} \overset{\substack{k_{h_1}(1) \\ \cdots \\ k_2(1) \\ k_1(1)}}{\longleftrightarrow}_{\text{b}} i(2) \overset{\substack{k_{h_2}(2) \\ \cdots \\ k_2(2) \\ k_1(2)}}{\longleftrightarrow}_{\text{b}} \cdots \overset{\substack{k_{h_m}(m) \\ \cdots \\ k_2(m) \\ k_1(m)}}{\longleftrightarrow}_{\text{b}} \underline{i(m+1)},$$

if the conditions listed below are satisfied:

1. $i(1) \in B_{i(m+1)}$ or $i(m+1) \in B_{i(1)}$;

2. For any $j \in [m]$, messages $i(j), i(j+1), k_1(j), k_2(j), \dots, k_{h_j}(j)$ form a singleton basic tower $\mathcal{B}^{\text{s}}_j$, i.e., for any $\ell \in [h_j]$, $\{i(j), i(j+1), k_1(j), k_2(j), \dots, k_{\ell-1}(j)\} \subseteq B_{k_\ell(j)}$.

Definition 4.7 (Singleton ordered tower)**.** For the CIC or DIC problem $\mathcal{G} : (i|j \in A_i), i \in [n]$, we have the following singleton ordered tower, $\mathcal{X}_j^{\mathrm{s,o}}$, constituted by individual messages as

$$\mathcal{X}_j^{\mathrm{s}}: \quad i(1) \xleftrightarrow{\substack{k_{h_1}(1) \\ \cdots \\ k_2(1) \\ k_1(1)}}_{\mathrm{b}} \cdots \xleftrightarrow{\substack{k_{h_{j-1}}(j-1) \\ \cdots \\ k_2(j-1) \\ k_1(j-1)}}_{\mathrm{b}} i(j) \xleftrightarrow{\substack{k_{h_j}(j) \\ \cdots \\ k_2(j) \\ k_1(j)}}_{\mathrm{c}} i(j+1) \xleftrightarrow{\substack{k_{h_{j+1}}(j+1) \\ \cdots \\ k_2(j+1) \\ k_1(j+1)}}_{\mathrm{b}} \cdots \xleftrightarrow{\substack{k_{h_q}(q) \\ \cdots \\ k_2(q) \\ k_1(q)}}_{\mathrm{b}} i(q+1),$$

if the conditions listed below are satisfied:

1. for any $j' \in [q] \setminus \{j\}$, messages $i(j'), i(j'+1), k_1(j'), k_2(j'), \ldots, k_{h_{j'}}(j')$ form a singleton basic tower $\mathcal{B}_{j'}^{\mathrm{s}}$.
2. for any $\ell \in [h_j]$ there exist two integers $s_{\ell,j} \in [j]$ and $t_{\ell,j} \in [j+1 : q+1]$ such that
   (a) $\{k_1(j), k_2(j), \ldots, k_{\ell-1}(j), i(s_{\ell,j}), i(t_{\ell,j})\} \subseteq B_{k_\ell(j)}$;
   (b) for any $\ell_1 < \ell_2 \in [h_j]$, we have $j = s_{1,j} \geq s_{\ell_1,j} \geq s_{\ell_2,j} \geq s_{h_j,j} = 1$, and $j+1 = t_{1,j} \leq t_{\ell_1,j} \leq t_{\ell_2,j} \leq t_{h_j,j} = q+1$.

For the singleton ordered tower $\mathcal{X}_j^{\mathrm{s,o}}$ defined above, it has $q = q+1-1 = t_{h_j,j} - s_{h_j,j}$ edges. We call edge j the central edge, and the message set $\{i(j), i(j+1), k_1(j), \ldots, k_{h_j}(j)\}$ the core of the singleton ordered tower. Every other edge $j' \in [q] \setminus \{j\}$ corresponds to a group of messages $i(j'), i(j'+1), k_1(j')$, $k_2(j'), \ldots, k_{h_{j'}}(j')$ that form a singleton basic tower $\mathcal{B}_{j'}^{\mathrm{s}}$. Different subscripts for the edges in the horizontal chain in Definition 4.7 are used.

Condition 2 in Definition 4.7 is described as follows. In the core, message $k_\ell(j)$ on the ℓ-th floor has messages $i(s_{\ell,j})$ and $i(t_{\ell,j})$ to start and terminate its coverage, respectively. In particular, for message $k_1(j)$ on the first floor, we have $i(s_{\ell,j}) = i(j)$, and $i(t_{\ell,j}) = i(j+1)$. The coverage of a message on a lower floor is within the range of the coverage of any message on a higher floor. We call the coverage of the message $k_{h_j}(j)$ on the top floor the total coverage of the singleton ordered tower, which is defined as $G_j \doteq [s_{h_j,j} : t_{h_j,j} - 1]$.[4]

Note that any singleton basic tower $\mathcal{B}_{j'}^{\mathrm{s}}$ can be seen as a special singleton ordered tower with $s_{\ell,j'} = j'$ and $t_{\ell,j'} = j'+1$ for any $\ell \in [h_{j'}]$, and hence $G_{j'} = \{j'\}$, and $|G_{j'}| = 1$. Unless otherwise stated, when we say a singleton ordered tower we assume that it is not a singleton basic tower.

For visualization of Definition 4.3, see Figure 4.7. To avoid clutter, we only draw $\mathcal{B}_1^{\mathrm{s}}$ and the core of central edge j.

Definition 4.8 (Singleton ordered acyclic chain)**.** For the CIC or DIC problem $\mathcal{G} : (i|j \in A_i), i \in [n]$, we have the singleton ordered acyclic chain, $Ch^{\mathrm{s,o}}$, of length m, constituted by

[4] We can see that the regular tower in Definition 4.3 is more general compared to the singleton ordered tower in Definition 4.7 in two facets: the regular tower is constructed by message sets while the singleton ordered tower is constructed by individual messages; the relationship among the coverages of different floors in the core of the singleton ordered tower can be seen as a special case of that of the regular tower.

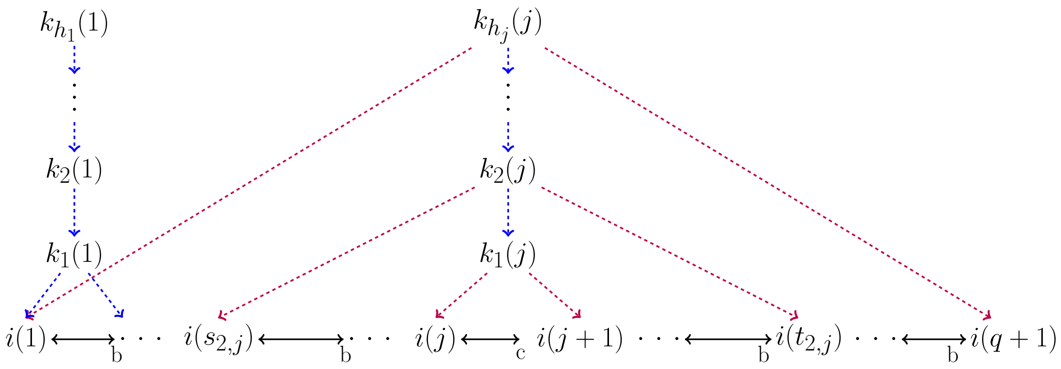


Figure 4.7: A schematic graph for the singleton ordered tower in Definition 4.7. A directed path that contains arrows of only one color (either blue or purple) from message a to b indicates that $b \in B_a$. According to Condition 2b of Definition 4.7, two purple arrows do not criss-cross.

individual messages as

$$Ch^{\mathrm{s,o}}: \quad \underline{i(1)} \xleftrightarrow{\substack{k_{h_1}(1)\\ \cdots \\ k_2(1)\\ k_1(1)}} i(2) \xleftrightarrow{\substack{k_{h_2}(2)\\ \cdots \\ k_2(2)\\ k_1(2)}} \cdots \xleftrightarrow{\substack{k_{h_m}(m)\\ \cdots \\ k_2(m)\\ k_1(m)}} \underline{i(m+1)},$$

if the conditions listed below are satisfied:

1. $i(1) \in B_{i(m+1)}$ or $i(m+1) \in B_{i(1)}$;

2. For every $j \in [m]$, messages $i(j), i(j+1), k_1(j), k_2(j), \ldots, k_{h_j}(j)$ constitute either a singleton basic tower $\mathcal{B}_j^{\mathrm{s}}$ or the core of a singleton ordered tower $\mathcal{X}_j^{\mathrm{s,o}}$;

3. For any $j_1 \neq j_2 \in M$, $G_{j_1} \cap G_{j_2} = \emptyset$, where $M \doteq \{j \in [m] : |G_j| \geq 2\}$ denotes the set of central edges of the singleton ordered towers within the chain.

We remove subscripts for the edges in the horizontal chain in the above definition as the positions of the singleton basic and singleton ordered towers are flexible.

Similar to the regular acyclic chain in Section 4.2, given a singleton ordered acyclic chain $Ch^{\mathrm{s,o}}$, let $M' = [m] \setminus (\bigcup_{j \in M} G_j)$ denote the set of edges located outside the coverage of any singleton ordered tower (cf. (4.7)). Then the singleton ordered acyclic chain can be seen as a horizontal concatenation of the non-overlapping singleton ordered towers $\mathcal{X}_j^{\mathrm{s,o}}$, $j \in M$ and the singleton basic towers $\mathcal{B}_{j'}^{\mathrm{s}}$, $j' \in M'$, such that the terminals of the chain $i(1)$ and $i(m+1)$ form an acyclic set. For a visualization of Definition 4.8, see Fig. 4.8.

For any singleton basic tower $\mathcal{B}_j^{\mathrm{s}}$, $\mathcal{B}_j^{\mathrm{s}}$ also denotes the collection of all the messages within it. Similarly, for any singleton ordered tower $\mathcal{X}_j^{\mathrm{s,o}}$, $\mathcal{X}_j^{\mathrm{s,o}}$ also denotes the collection of all the messages within it. For any singleton basic tower, singleton ordered tower, singleton basic

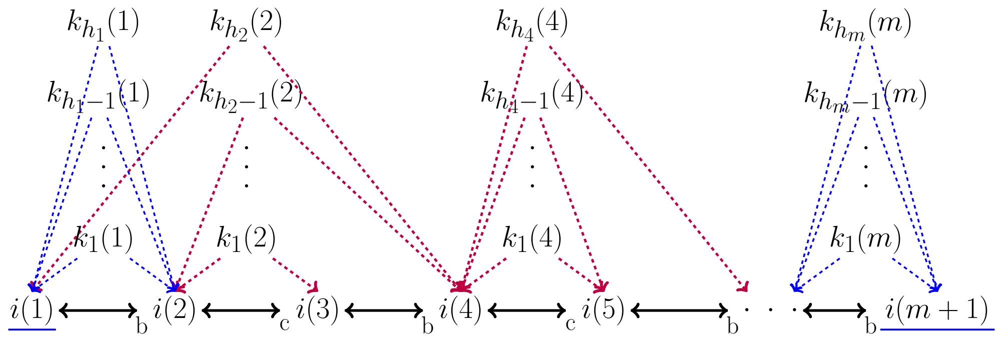


Figure 4.8: A visualization example for the singleton ordered acyclic chain in Definition 4.8. To help with understanding, we draw blue and purple dashed arrows. If there is a directed path formed by dashed arrows of the same color from message a to b, then $b \in B_a$. Definitions 4.7 and 4.8 jointly ensure that purple arrows can never criss-cross.

acyclic chain, or singleton ordered acyclic chain, we call the i-labeled and k-labeled messages within it its components.

Given a problem $\mathcal{G}$, let $\mathfrak{C}^{\mathrm{s,o}}(\mathcal{G})$ and $\mathfrak{C}^{\mathrm{s}}_{\mathrm{b}}(\mathcal{G})$ denote the collection of its singleton ordered acyclic chains and the collection of its singleton basic acyclic chains, respectively. Note that any alignment chain can be viewed as a special singleton basic acyclic chain, and any singleton basic acyclic chain can be viewed as a special singleton ordered acyclic chain. It can be verified that $\mathfrak{C}^{\mathrm{s}}_{\mathrm{b}}(\mathcal{G}) = \emptyset$ if and only if $\mathcal{G}$ is half-rate-feasible. It can also be verified that $\mathfrak{C}^{\mathrm{s,o}}(\mathcal{G}) = \emptyset$ if and only if $\mathcal{G}$ is half-rate-feasible.

Theorem 4.5 (Singleton basic acyclic chain bound)**.** Consider the DIC problem $\mathcal{G}$: $(i|j \in A_i)$, $i \in [n]$ with link capacity tuple $\mathbf{C}$. Let R be any achievable symmetric rate. Then for any of its singleton basic acyclic chain $Ch^{\mathrm{s}}_{\mathrm{b}} \in \mathfrak{C}^{\mathrm{s}}_{\mathrm{b}}(\mathcal{G})$, we have

$$R \leq \frac{1}{1+m+\sum_{j\in[m]} h_j} \sum_{j\in[m]} \sum_{\substack{J\in N: J\cap\{i(j),i(j+1),k_1(j),\ldots,k_{h_j}(j)\}\neq\emptyset,\\ J\cap\{i(j),i(m+1),k_1(j),\ldots,k_{h_j}(j)\}\neq\emptyset}} C_J. \tag{4.17}$$

We omit Theorem 4.5's proof since it can be seen a special case of Theorem 4.6 to be presented below and proved in Appendix B.3.

Recall that any singleton basic tower $\mathcal{B}^{\mathrm{s}}_{j'}$ can be seen as a special singleton ordered tower with $s_{\ell,j'} = j'$ and $t_{\ell,j'} = j'+1$ for any $\ell \in [h_{j'}]$. Also recall that given a singleton ordered acyclic chain, $M' = [m] \setminus (\bigcup_{j\in M} G_j)$ denotes the set of edges located outside the coverage of any singleton ordered tower.

Theorem 4.6 (Singleton ordered acyclic chain bound)**.** Consider the DIC problem $\mathcal{G}$: $(i|j \in A_i)$, $i \in [n]$ with link capacity tuple $\mathbf{C}$. Let R be any achievable symmetric rate. Then for any of its singleton ordered acyclic chain $Ch^{\mathrm{s,o}} \in \mathfrak{C}^{\mathrm{s,o}}(\mathcal{G})$ we have

$$
\begin{aligned}
R \leq \frac{1}{1+m+\sum_{j\in[m]} h_j} (\sum_{j\in M\cup M'} \sum_{J\in N: J\cap T_1(j)\neq\emptyset, J\cap T_2(j)\neq\emptyset} C_J \\
+ \sum_{j\in M} \sum_{\ell\in[2:h_j]} (\sum_{j'\in[s_{\ell,j}:s_{\ell-1,j}-1]} \sum_{\substack{J\in N: J\cap T_3(j')\neq\emptyset,\\ J\cap T_4(j,\ell,j')\neq\emptyset}} C_J + \sum_{j'\in[t_{\ell-1,j}:t_{\ell,j}-1]} \sum_{\substack{J\in N: J\cap T_3(j')\neq\emptyset,\\ J\cap T_5(j,\ell,j')\neq\emptyset}} C_J)),
\end{aligned}
\tag{4.18}
$$

where

$$
\begin{aligned}
T_1(j) &= \{i(s_{h_j,j}), i(t_{h_j,j}), k_1(j), \ldots, k_{h_j}(j)\},\\
T_2(j) &= \{i(s_{h_j,j}), i(m+1), k_1(j), \ldots, k_{h_j}(j)\},\\
T_3(j') &= \{i(j'), i(j'+1), k_1(j'), \ldots, k_{h_{j'}}(j')\},\\
T_4(j,\ell,j') &= \{i(j'), i(s_{\ell-1,j}), k_1(j'), \ldots, k_{h_{j'}}(j')\},\\
T_5(j,\ell,j') &= \{i(j'), i(t_{\ell,j}), k_1(j'), \ldots, k_{h_{j'}}(j')\}.
\end{aligned}
$$

The proof for Theorem 4.6 is presented in Appendix B.3.

Remark 4.3. The singleton basic acyclic chain bound in Theorem 4.5 is always no looser than the distributed MAIS bound in Proposition 2.8, which can be proved in a similar way to Proposition 4.1.

Theorems 4.5 and 4.6 can give tight bounds for some DIC problems.

Example 4.7. Consider the following 5-message DIC problem $\mathcal{G}$:

$$(1|2,3,4,5), (2|1,3,4,5), (3|2,4,5), (4|3,5), (5|1,4),$$

with equal link capacities $C_J = 1, J \in N \setminus \{\emptyset\}$. For this problem, there exists a singleton basic acyclic chain as

$$Ch^{\mathrm{s}}_{\mathrm{b}} : \underline{1} \xleftrightarrow{4}_{\mathrm{b}} 2 \xleftrightarrow{5}_{\mathrm{b}} \underline{3}.$$

Therefore, by the singleton basic acyclic chain bound in Theorem 4.5, for any symmetric achievable rate R, we have

$$R \leq \frac{1}{5}(\sum_{\substack{J\in N: J\cap\{1,2,4\}\neq\emptyset,\\ J\cap\{1,3,4\}\neq\emptyset}} 1 + \sum_{J\in N: J\cap\{2,3,5\}\neq\emptyset} 1) = \frac{54}{5}.$$

It can be verified that the above upper bound on the symmetric capacity matches the lower bound given by the distributed composite coding (DCC) scheme in Theorem 3.8. Thus, we establish the symmetric capacity of this problem to be $54/5$.

4.5 Grouping Polymatroidal Bound for DIC

Thus far, we have been focusing on the performance bounds based on various acyclic chain structures for the CIC and DIC problems. Despite of their usefulness as we demonstrated through a number of examples, they only provide bounds on the broadcast rate or symmetric capacity, and can be loose in general. For the CIC problem we know that the PM bound in Proposition 2.7 introduced in Blasiak et al. [2011] is the tightest bound on the entire capacity region one can get based on Shannon-type inequalities. There is no existing counterpart of such a powerful PM bound for the DIC problem in the literature.

In this section we develop a general performance bound for the DIC problem, which can potentially capture all the Shannon-type inequalities. The new bound is based on the PM axioms of the entropy function and server groupings, the latter of which is defined as follows.

Definition 4.9 (Valid server grouping)**.** A *server grouping* $\mathcal{P} = \{P_1, P_2, \cdots, P_m\}$, consisting of m server groups $P_i \subseteq N, i \in [m]$, is said to be *valid* if $\bigcup_{i\in[m]} P_i = N$. Given a server grouping $\mathcal{P}$, we denote by $P_G = \bigcup_{i\in G} P_i$ the collection of servers in the server groups identified by $G \subseteq [m]$. By convention, $P_\emptyset = \emptyset$.

Note that we allow overlaps between different server groups in a grouping. Also note that $P_{[m]} = \bigcup_{i\in[m]} P_i = N$. Let Y_{P_G} denote the output random variables from the server collection P_G, e.g., $Y_{P_{[m]}} = Y_N$. When the context is clear, we shall use the shorthand notation P_G for Y_{P_G}, e.g., $H(P_G|X_{K^c})$ means $H(Y_{P_G}|X_{K^c})$.

We briefly review the standing assumptions and conditions of $(\mathbf{t}, \mathbf{r})$ distributed index codes and achievable rate–capacity tuples $(\mathbf{R}, \mathbf{C})$. Since the messages are assumed to be independent and uniformly distributed, for any two disjoint sets $K, K' \subseteq [n]$ we have

$$H(X_K|X_{K'}) = H(X_K) = \sum_{i\in K} t_i. \tag{4.19}$$

The *encoding condition* at server $J \in N$ is

$$H(Y_J|X_J) = 0. \tag{4.20}$$

The *decoding condition* at receiver i stipulates

$$H(X_i|Y_N, X_{A_i}) \leq t_i \cdot \delta(\epsilon) \tag{4.21}$$

with $\lim_{\epsilon\to 0} \delta(\epsilon) = 0$ by Fano's inequality [Cover, 2006, Theorem 2.10.1].

Also recall the touch structure defined in Definition 2.2, which will be extensively used in this and upcoming sections.

As stated in Remark 2.1, it is not known whether the capacity region and the zero-error capacity region of the DIC problem are equal to each other. Hence, we use the general vanishing error decoding condition in (4.21) rather than assuming zero-error decoding.

In the following we state the main result of this section, namely, the *grouping polymatroidal (PM) bound*.

Theorem 4.7 (Grouping ploymatroidal (PM) bound)**.** Consider the DIC problem $\mathcal{G}$: $(i|j \in A_i)$, $i \in [n]$ with link capacity tuple $\mathbf{C}$. Its capacity region $\mathscr{C}(\mathcal{G}, \mathbf{C})$ is outer bounded by the rate region $\mathscr{R}_{\text{GPM}}(\mathcal{G}, \mathbf{C})$ that consists of all rate tuples $\mathbf{R}$ such that for any valid server grouping $\mathcal{P} = \{P_1, P_2, \cdots, P_m\}$,

$$R_i \leq f([m], B_i \cup \{i\}) - f([m], B_i), \qquad i \in [n], \tag{4.22}$$

for at least one set function $f(G,K)$, for all $G, G' \subseteq [m]$, $K, K' \subseteq [n]$, such that

$$f(G,K) = f(G',K), \qquad \text{if } (P_G \cup P_{G'}) \setminus (P_G \cap P_{G'}) \subseteq T_{\overline{K}}, \tag{4.23}$$

$$f(\emptyset, K) = f(G, \emptyset) = 0, \tag{4.24}$$

$$f(G,K) \leq \sum_{J: J \in P_G, J \in T_K} C_J, \tag{4.25}$$

$$f(G,K) \leq f(G',K'), \qquad \text{if } K \subseteq K', G \subseteq G', \tag{4.26}$$

$$f(G \cup G', K \cap K') + f(G \cap G', K \cup K') \leq f(G,K) + f(G',K'), \tag{4.27}$$

$$f([m], B_i \cup \{i\}) - f([m], B_i) = f([m], \{i\}), \;\; i \in [n], \tag{4.28}$$

$$f(G,K) + f(G,K') = f(G, K \cup K'), \qquad \text{if } K \cap K' = \emptyset, P_G \subseteq (N \setminus T_{K,K'}). \tag{4.29}$$

Proof. If a rate tuple $\mathbf{R}$ is achievable, then for every $\epsilon > 0$ there exists a $(\mathbf{t}, \mathbf{r})$ distributed index code satisfying (2.3). For any $i \in [n]$, we have

$$\begin{aligned} t_i &= H(X_i | X_{A_i}) && (4.30)\\ &\leq H(X_i|X_{A_i}) - H(X_i|Y_N, X_{A_i}) + t_i \cdot \delta(\epsilon) && (4.31)\\ &= I(X_i; Y_N | X_{A_i}) + t_i \cdot \delta(\epsilon) && (4.32)\\ &= H(Y_N|X_{A_i}) - H(Y_N|X_{A_i \cup \{i\}}) + t_i \cdot \delta(\epsilon), && (4.33) \end{aligned}$$

where (4.30) follows from the fact that the messages are independent and uniformly distributed as specified in (4.19), and (4.31) is due to the decoding condition in (4.21). Now, given the server grouping $\mathcal{P} = \{P_1, P_2, \cdots, P_m\}$, define

$$f_\epsilon(G,K) \doteq \frac{1}{r} H(P_G | X_{K^c}) = \frac{1}{r} H(Y_{\{J: J \in \bigcup_{i \in G} P_i\}} | X_{K^c}), \tag{4.34}$$

for $G \subseteq [m]$ and $K \subseteq [n]$. Recall that as specified in Section 2.1, the $(\mathbf{t}, \mathbf{r})$ distributed index code depends on ϵ and thus so does the set function $f_\epsilon(G,K)$. Then,

$$R_i \leq \frac{t_i}{r} \leq \frac{H(Y_N|X_{A_i}) - H(Y_N|X_{A_i\cup\{i\}})}{r\cdot(1-\delta(\epsilon))} = \frac{f_\epsilon([m],B_i\cup\{i\}) - f_\epsilon([m],B_i)}{1-\delta(\epsilon)}, \tag{4.35}$$

where the second inequality follows from (4.33) and the equality from the definition of $f_\epsilon(G,K)$.

We now show that the set function $f_\epsilon(G,K)$ is bounded from above for any $G \subseteq [m]$, $K \subseteq [n]$ and any $\epsilon > 0$. We have

$$f_\epsilon(G,K) = \frac{1}{r}H(P_G|X_{K^c}) \leq \frac{1}{r}H(P_G) \leq \sum_{J\in P_G}\frac{1}{r}H(Y_J) \leq \sum_{J\in P_G}\frac{r_J}{r} \leq \sum_{J\in P_G} C_J. \tag{4.36}$$

Also, $f_\epsilon(G,K)$ is bounded from below as $f_\epsilon(G,K) \geq 0$ due to the nonnegativity of the entropy function.

To characterize the achievable rate tuple $\mathbf{R}$, we need to define a set function $f(G,K)$ that does not depend on the decoding error threshold ϵ. We define such set function $f(G,K)$ as the limit infimum of $f_\epsilon(G,K)$:

$$f(G,K) \doteq \liminf_{\epsilon\to 0} f_\epsilon(G,K), \tag{4.37}$$

which is real and bounded given the boundedness of $f_\epsilon(G,K)$. Now taking the limit infimum as ϵ approaches zero on both sides of (4.35) yields

$$\begin{aligned} R_i &\leq \liminf_{\epsilon\to 0} \frac{f_\epsilon([m],B_i\cup\{i\}) - f_\epsilon([m],B_i)}{1-\delta(\epsilon)} && (4.38)\\ &= \liminf_{\epsilon\to 0}(f_\epsilon([m],B_i\cup\{i\}) - f_\epsilon([m],B_i)) && (4.39)\\ &= \liminf_{\epsilon\to 0} f_\epsilon([m],B_i\cup\{i\}) - \limsup_{\epsilon\to 0} f_\epsilon([m],B_i) && (4.40)\\ &\leq \liminf_{\epsilon\to 0} f_\epsilon([m],B_i\cup\{i\}) - \liminf_{\epsilon\to 0} f_\epsilon([m],B_i) && (4.41)\\ &= f([m],B_i\cup\{i\}) - f([m],B_i). && (4.42) \end{aligned}$$

We have thus far established (4.22). It is also checked in Section B.4 that $f(G,K)$ satisfies the conditions in (4.23)–(4.29), which establishes Theorem 4.7. □

Remark 4.4. The conditions (4.23)–(4.29) in Theorem 4.7 will be referred to as the *Axioms* of the grouping PM bound. Axioms (4.24), (4.26), and (4.27) capture standard polymatroidal properties of the entropy function. Axioms (4.23) and (4.25) capture the encoding conditions at servers, as well as the link capacity constraints. Axiom (4.29) captures the conditional message independence given by the fd-separation Kramer [1998]; Thakor et al. [2016], and thus is referred to as the fd-separation axiom. Refer to Appendix B.5 for a brief treatment on how fd-separation applies to the DIC problem. The inequality (4.22) will be referred to as the *rate constraint inequality* jointly satisfied by $\mathbf{R}$ and $f(G,K)$, which is based on the message independence, as well as the decoding conditions at receivers.

Remark 4.5. Axiom (4.28) captures the additional decoding conditions at the receivers (cf. property (2.40) in Proposition 2.7). It has been found in Liu et al. [2018a] that this axiom is strictly needed to obtain tight outer bounds on the capacity region for the *secure* CIC problem. Currently, we are not aware of any instance of *non-secure* CIC or DIC for which Axiom (4.28) can tighten the outer bound. However, we include this axiom for the following reason. Through including the additional decoding conditions in the CIC problem, it was shown in Liu et al. [2018a] that the PM bound in Proposition 2.7 is as tight as the apparently stronger bound in which *all* Shannon-type inequalities are used. A similar relation exists between the most *refined* version of the grouping PM bound for the DIC problem and the one obtained based on all Shannon-type inequalities of the entropy function. See Remark 4.11 in Section 4.8.

The tightness and computational complexity of the grouping PM bound in Theorem 4.7 depends pivotally on the specific server grouping $\mathcal{P}$. For a fixed problem size n, the number of variables in the theorem is exponential in m, the size of $\mathcal{P}$. To fully compute the rate region $\mathscr{R}_{\text{GPM}}$ satisfying (4.22)–(4.29) for a given DIC problem $(\mathcal{G},\mathbf{C})$ and a given server grouping $\mathcal{P}$, one should use FME to remove all the 2^{m+n} intermediate variables $f(G,K)$, $G \subseteq [m]$, $K \subseteq [n]$. In general, this operation is prohibitively complex even for small m and n. For a given $\mathbf{C}$, however, it is typically tractable to establish an upper bound on the (weighted) sum-capacity using LP subject to (4.22)–(4.29). In the next three sections, we specialize the outer bound using a number of explicit constructions for server grouping. In general, the optimal tightness-complexity tradeoff in choosing a server grouping remains open.

Throughout the rest of this chapter, when there is no ambiguity, we denote the outer bound given by the grouping PM bound in Theorem 4.7 with a specific server grouping $\mathcal{P}$ as $\mathscr{R}_{\mathcal{P}}(\mathcal{G},\mathbf{C})$ or simply $\mathscr{R}_{\mathcal{P}}$. For brevity, whenever we say that one server grouping $\mathcal{P}$ is tighter (looser) than another grouping $\mathcal{P}'$, we mean that the corresponding outer bound $\mathscr{R}_{\mathcal{P}}$ is tighter (looser) than the outer bound $\mathscr{R}_{\mathcal{P}'}$.

4.6 Outer Bounds Based on Server Groupings Utilizing the Touch Structure

In this section we explicitly construct server groupings based on the touch structure of Definition 2.2. Let us motivate the construction through a series of examples and definitions. Together, they will lead to a closed-form upper bound on the sum capacity, which is implied by Theorem 4.7 with server groupings based on a specific touch structure.

Example 4.8. Consider the DIC problem $(1|-),(2|4),(3|4),(4|3)$ with equal unit link capacities $C_J = 1$ for all $J \in N \setminus \{\emptyset\}$. Consider two sets $L = \{4,2\}$ and $K = \{1,3\}$, where set L is ordered as $L = \{i_1 = 4, i_2 = 2\}$. For this problem, we can verify that

$$A_{i_1} = A_4 = \{3\} \subset K = \{1,3\}, \qquad A_{i_2} = A_2 = \{4\} \subset K \cup \{i_1\} = \{1,3,4\}. \tag{4.43}$$

We say that L is an augmentation set of K. Similarly, $L = \{2,3\}$ is an augmentation set of $K = \{1,4\}$. When $K = \emptyset$, we find a unique *maximal* augmentation set, called the peripheral. The set $U = \{1\}$ is the peripheral for this problem as $A_1 = \emptyset \subseteq \emptyset$, and $A_i \nsubseteq \{1\}$, $i \in \{2,3,4\}$.

Generally, the idea is to find two disjoint subsets $L, K \subseteq [n]$, such that any valid index code *augments* the singular message decoding condition (4.21) to

$$H(X_L | Y_N, X_K) \leq \sum_{i \in L} t_i \cdot \delta(\epsilon).$$

For the peripheral set U, the decoding condition gives

$$H(X_U | Y_N) \leq \sum_{i \in U} t_i \cdot \delta(\epsilon).$$

We formalize this through the following definitions.

Definition 4.10 (Augmentation set)**.** For the DIC problem $\mathcal{G}$: $(i|j \in A_i)$, $i \in [n]$ and any two disjoint sets $L, K \subseteq [n]$, we say L is an *augmentation set* of K if there exists an ordering $i_1, i_2, \cdots, i_{|L|}$ of the elements in L such that $A_{i_j} \subseteq \{i_1, \cdots, i_{j-1}\} \cup K, j \in [|L|]$. The empty set $\emptyset$ is an augmentation set of any set $K \subseteq [n]$.

Definition 4.11 (Peripheral)**.** For the DIC problem $\mathcal{G}$: $(i|j \in A_i)$, $i \in [n]$, we say that set $U \subseteq [n]$ is a *peripheral* if U is an augmentation set of the empty set $\emptyset$, and that for any $i \in U^c$, we have $A_i \nsubseteq U$.

For a given problem $(i|A_i), i \in [n]$, peripheral U is unique. This can be verified by contradiction as follows. Assume that there exist two different peripherals U, U', and $U \setminus U' \neq \emptyset$. Define $u \doteq |U|$, $U_1 \doteq U \setminus U'$ and $U_0 \doteq U \cap U'$. Then, $U = U_0 \cup U_1$ and $U_1 \cap U_0 = \emptyset$. By Definition 4.11, there exists an ordering $i_1, i_2, \cdots, i_u$ of the elements in U such that $A_{i_j} \subseteq \{i_1, \cdots, i_{j-1}\}$, $j \in [u]$. There always exists some $s \in [u]$ such that $i_s \in U_1$ and $\{i_1, \cdots, i_{s-1}\} \subseteq U_0$. Hence, we have $A_{i_s} \subseteq \{i_1, \cdots, i_{s-1}\} \subseteq U_0 \subseteq U'$. We also have $i_s \in U'^c$ since $i_s \in U_1$ and $U_1 \cap U' = \emptyset$. Combining that $A_{i_s} \subseteq U'$ and that $i_s \in U'^c$ leads to a contradiction against Definition 4.11, and thus completes the proof.

Remark 4.6. In Liu and Tuninetti [2018], a decoding chain is established based on the idea that receiver j can mimic another receiver i and decode message i at no cost to the achievable rates if receiver j knows everything receiver i does. This is similar to the notion of augmentation set in Definition 4.10. A chaining procedure, starting from a receiver with an empty side information set, is used in Neely et al. [2013] to prove a lower bound on the broadcast rate of the CIC problem. This is similar to the procedure of building the peripheral in Definition 4.11.

Based on Definitions 4.10 and 4.11, we define the augmentation group as follows.

Definition 4.12 (Augmentation group)**.** For the DIC problem $\mathcal{G}$: $(i|j \in A_i), i \in [n]$ with peripheral U, we use $\mathbf{V} = (V_1, V_2, \cdots, V_k)$, referred to as an *augmentation group*, to denote the tuple of k disjoint nonempty sets $V_1, \cdots, V_k \subseteq U^c$ for some $k \geq 1$ such that the following conditions are satisfied:

1. for any $j \in [k]$, V_j is an augmentation set of its complement set V_j^c;
2. set $W \doteq [n] \setminus U \setminus (\bigcup_{j\in[k]} V_j)$ is an augmentation set of its complement set W^c;
3. there does not exist another tuple of disjoint nonempty sets $\mathbf{V}' = (V_1', V_2', \cdots, V_{k'}')$ such that it satisfies the first two conditions, and that $\bigcup_{j'\in[k']} V_{j'}' \subset \bigcup_{j\in[k]} V_j$.
4. there does not exist another tuple of disjoint nonempty sets $\mathbf{V}' = (V_1', V_2', \cdots, V_{k'}')$ such that it satisfies the first two conditions, $\bigcup_{j'\in[k']} V_{j'}' = \bigcup_{j\in[k]} V_j$, and that $k' < k$.

Note that there can be multiple augmentation groups for a given problem.

Example 4.9. Consider the DIC problem $(1|-), (2|4), (3|4), (4|3)$ discussed in Example 4.8. We have $U = \{1\}$ and there are in total 2 augmentation groups $\mathbf{V} = (\{3\})$ and $\mathbf{V}' = (\{4\})$. Note that $\mathbf{V}'' = (\{3\}, \{4\})$ does satisfy the first two conditions of Definition 4.12, yet given the existence of $\mathbf{V}$ and $\mathbf{V}'$, according to the third condition, $\mathbf{V}''$ is not a valid augmentation group. For another example, consider the six-message problem $(1|4), (2|3), (3|2), (4|1), (5|-), (6|5)$. We have $U = \{5, 6\}$ and in total 4 augmentation groups shown as follows,

$$\mathbf{V}^1 = (\{1,2\}), \qquad \mathbf{V}^2 = (\{1,3\}), \qquad \mathbf{V}^3 = (\{2,4\}), \qquad \mathbf{V}^4 = (\{3,4\}). \tag{4.44}$$

Note that $\mathbf{V}' = (\{1\}, \{2\})$ does satisfy the first two conditions of Definition 4.12, yet given the existence of $\mathbf{V}^1$, according to the fourth condition, $\mathbf{V}'$ is not a valid augmentation group.

A closed-form upper bound on the sum capacity, namely the *augmentation group bound*, is given by the following theorem.

Theorem 4.8 (Augmentation group bound)**.** Consider the DIC problem $\mathcal{G}$: $(i|j \in A_i), i \in [n]$ with the link capacity tuple $\mathbf{C}$ and the peripheral U. Let $\mathbf{R}$ be any achievable rate tuple. For any of its valid augmentation group $\mathbf{V} = (V_1, V_2, \cdots, V_k)$, we have

$$\sum_{i\in[n]} R_i \leq \sum_{J\in N} C_J + \sum_{\ell\in[k]} \sum_{J\in\mathcal{T}_\ell} C_J, \tag{4.45}$$

where $\mathcal{T}_\ell \doteq T_{V_\ell,(\bigcup_{j\in[\ell+1:k]} V_j)\cup W} = \{J \in N : J \cap V_\ell \neq \emptyset, J \cap ((\bigcup_{j\in[\ell+1:k]} V_j) \cup W) \neq \emptyset\}$ and $W = [n] \setminus U \setminus (\bigcup_{j\in[k]} V_j)$.

Remark 4.7. For a given DIC problem, every augmentation group yields one performance bound as specified in (4.45) on its sum capacity. To compute the *best* augmentation group bound, one can simply compute the bound for all augmentation groups through exhaustive search. The computational complexity of such search depends only on the problem size n and receiver side information structure rather than the server setup.

We prove Theorem 4.8 by showing that (4.45) is implied by Theorem 4.7 with a specific server grouping $\mathcal{P}_\mathbf{V}$ defined below. The proof details are given in Appendix B.6.

Definition 4.13 ($\mathcal{P}_\mathbf{V}$ grouping)**.** For the DIC problem $\mathcal{G}$: $(i|j \in A_i), i \in [n]$ with an augmentation group $\mathbf{V} = (V_1, V_2, \cdots, V_k)$, the server grouping $\mathcal{P}_\mathbf{V}$ is defined as follows

$$\mathcal{P}_\mathbf{V} = \{T_{V_1}, T_{V_2}, \cdots, T_{V_k}, T_{(\bigcup_{j\in[k]} V_j)^c}\}. \tag{4.46}$$

In some cases, Theorem 4.8 gives a tight bound on the sum capacity, as illustrated below.

Example 4.10. Recall the 4-message DIC problem $(1|-), (2|4), (3|4), (4|3)$ discussed in Examples 4.8 and 4.9. The all-server distributed PM bound in Proposition 2.9 yields that $R_1 + R_2 + R_3 + R_4 \leq 22$. In comparison, given the peripheral $U = \{1\}$ and the augmentation group $\mathbf{V} = \{\{3\}\}$, we have $W = [n] \setminus \{1\} \setminus \{3\} = \{2, 4\}$, and hence Theorem 4.8 tightens the sum-capacity upper bound to

$$R_1 + R_2 + R_3 + R_4 \leq \sum_{J\in N} C_J + \sum_{J\in T_{\{3\},\{2,4\}}} C_J = 15 + 6 = 21, \tag{4.47}$$

which matches the lower bound presented in Example 3.11 in Chapter 3. Note that with another augmentation group $\mathbf{V}' = \{\{4\}\}$, Theorem 4.8 yields the same tight upper bound of 21 on the sum capacity.

Note that $\mathcal{P}_\mathbf{V}$ is a server grouping whose server groups are in the form of touch structure. To generalize this further, we introduce the touch grouping and its resulting outer bound as follows.

Definition 4.14 (Touch grouping)**.** For a given $m \leq n$ and disjoint nonempty sets $L_i \subseteq [n]$,

$i \in [m]$, such that $\bigcup_{i\in[m]} L_i = [n]$, the *touch grouping* $\mathcal{P}_{\mathrm{t}}$ is defined as

$$\mathcal{P}_{\mathrm{t}} = \{T_{L_1}, T_{L_2}, \cdots, T_{L_m}\}. \tag{4.48}$$

For the special case $m = n$, $L_i = \{i\}$, and $P_i = T_{\{i\}}$, the touch grouping is called the *individual touch grouping* and is denoted by

$$\mathcal{P}_{\mathrm{t}}^* = \{T_{\{1\}}, T_{\{2\}}, \ldots, T_{\{n\}}\}. \tag{4.49}$$

Note that we have $P_G = \bigcup_{i\in G} T_{L_i} = T_{L_G}$, where $L_G \doteq \bigcup_{i\in G} L_i$. With the touch grouping $\mathcal{P}_{\mathrm{t}}$, we obtain a special case of the general grouping PM bound, namely the *touch grouping PM bound*. The expression of the touch grouping PM bound's axioms are somewhat simplified compared with the original axioms in Theorem 4.7.

Corollary 4.2 (Touch grouping PM bound)**.** Consider the DIC problem $\mathcal{G}$: $(i|j \in A_i), i \in [n]$ with link capacity tuple $\mathbf{C}$. For a given touch grouping $\mathcal{P}_{\mathrm{t}} = \{T_{L_1}, T_{L_2}, \cdots, T_{L_m}\}$, the capacity region $\mathscr{C}(\mathcal{G}, \mathbf{C})$ is outer bounded by the rate region $\mathscr{R}_{\mathcal{P}_{\mathrm{t}}}(\mathcal{G}, \mathbf{C})$ that consists of all rate tuples $\mathbf{R}$ such that

$$R_i \le f([m], B_i \cup \{i\}) - f([m], B_i), \qquad i \in [n], \tag{4.50}$$

for at least one set function $f(G, K)$ for all $G, G' \subseteq [m], K, K' \subseteq [n]$ satisfying

$$f(G,K) = f(G',K), \qquad \text{if } K \subseteq (L_G \cap L_{G'}), \tag{4.51}$$

$$f(\emptyset, K) = f(G, \emptyset) = 0, \tag{4.52}$$

$$f(G,K) \le \sum_{J:J\in T_{L_G,K}} C_J, \tag{4.53}$$

$$f(G,K) \le f(G',K'), \qquad \text{if } K \subseteq K', G \subseteq G', \tag{4.54}$$

$$f(G \cup G', K \cap K') + f(G \cap G', K \cup K') \le f(G,K) + f(G',K'), \tag{4.55}$$

$$f([m], B_i \cup \{i\}) - f([m], B_i) = f([m], \{i\}), \qquad \forall i \in [n]. \tag{4.56}$$

Proof. It is obvious that Axioms (4.24), (4.26), (4.27), and (4.28) in Theorem 4.7 and Axioms (4.52), (4.54), (4.55), and (4.56) above do not depend on the underlying server grouping, and thus remain unchanged. Since $P_G = \bigcup_{j\in G} T_{L_j} = T_{L_G}$, Axioms (4.25) and (4.53) are the same.

Note that $[n] \in T_L$ for any nonempty $L \subseteq [n]$, which indicates that the server $J = [n]$ containing all messages is common among all server groups in the touch grouping $\mathcal{P}_{\mathrm{t}}$. Hence, with $\mathcal{P}_{\mathrm{t}}$, there never exists two disjoint nonempty sets $K, K' \subseteq [n]$ such that $T_{L_G} \subseteq (N \setminus T_{K,K'})$ for any nonempty $G \subseteq [m]$. This implies that the fd-separation axiom, Axiom (4.29), in Theorem 4.7, can only give trivial inequalities (e.g., $f(\emptyset, K) + f(\emptyset, K') = f(\emptyset, K \cup K')$).

Hence, in Corollary 4.2, there is no axiom corresponding to the fd-separation axiom.

It remains to prove that Axiom (4.23) in Theorem 4.7 simplifies to Axiom (4.51), which is relegated to Appendix B.7. □

Remark 4.8. The grouping PM bound of Theorem 4.7 can easily incorporate the set of active servers $N_A = \{J \in N : C_J > 0\}$ (cf. Remark 3.9). One can simply replace N with N_A and $\mathcal{P} = \{P_1, \cdots, P_m\}$ with $\mathcal{P}_A = \{P_1 \cap N_A, \cdots, P_m \cap N_A\}$. However, notice that the axioms of Corollary 4.2 (and those of Corollary 4.5 to be introduced in Section 4.8) are expressed in simplified forms based on the assumption that all the servers $J \in N$ are active. When $N_A \subset N$, using these simplified axioms might result in looser outer bounds. One can avoid this issue by using the axioms in their original unsimplified forms of Theorem 4.7 with the desired server grouping.

Remark 4.9. Within the general class of touch grouping, it is unclear which touch grouping can give the tightest capacity outer bound with the lowest possible computational cost. Since $\mathcal{P}_{\mathrm{V}}$ grouping has an explicit construction, one may first try this grouping and compare the obtained performance bound with an achievable coding scheme. If the results match, no further action is required. Otherwise, the finest touch grouping $\mathcal{P}_{\mathrm{t}}^* = \{T_{\{1\}}, T_{\{2\}}, \ldots, T_{\{n\}}\}$ can be tried, which results in the tightest outer bound on the capacity region among all possible touch groupings. See Section 4.8 for the hierarchy of server groupings in terms of their tightness.

4.7 Outer Bounds Based on Server Groupings Utilizing fd-separation

One limitation of the touch grouping PM bound is the missing fd-separation axiom (also see the discussion in the proof of Corollary 4.2). To show the usefulness of the fd-separation axiom, we present a list of problems with discussion, leading to a construction of server grouping based on fd-separation.

Definition 4.15 (Isolated vertex and disjoint cycles)**.** For the DIC problem $(i|j \in A_i)$, $i \in [n]$ with side information graph $\mathcal{G}$, a vertex $v \in V(\mathcal{G})$ is said to be *isolated* if it has no incoming edges. That is, there does not exist any edge $e = (v', v) \in E(\mathcal{G})$ for some $v' \in V(\mathcal{G})$. Two cycles $K, K' \subseteq V(\mathcal{G})$ in $\mathcal{G}$ are said to be *disjoint* if $K \cap K' = \emptyset$.

Example 4.11. Consider the following six DIC problems with $n = 4$ and equal link capacities

$C_J = 1, J \in N \setminus \{\emptyset\}$,

$$\begin{gathered}(1|-),(2|4),(3|2),(4|3);\\(1|-),(2|4),(3|2),(4|1,3);\\(1|-),(2|1,4),(3|1,2),(4|1,3);\\(1|4),(2|3),(3|2),(4|1,3);\\(1|4),(2|3),(3|2),(4|1,2,3);\\(1|4),(2|3),(3|1,2),(4|1,2),\end{gathered} \tag{4.57}$$

whose side information graphs are shown in Figure 4.9. For the problems shown in Figures 4.9(a), 4.9(b), and 4.9(c), the touch grouping PM bound $\mathscr{R}_{\mathcal{P}_t}$ yields $\sum_{i\in[4]} R_i \leq 19.5$. With $\mathcal{P} = \{P_1, N \setminus P_1\}$, where $P_1 = \{J \in N : |J \setminus \{1\}| \leq 1\}$, a tighter upper bound of 19 can be obtained by the grouping PM bound $\mathscr{R}_{\text{GPM}}$ for these three problems, matching their sum capacity. For the problems shown in Figures 4.9(d), 4.9(e), and 4.9(f), the touch grouping PM bound $\mathscr{R}_{\mathcal{P}_t}$ yields $\sum_{i\in[4]} R_i \leq 24$. With $\mathcal{P}' = \{P'_1, P'_2, N \setminus P'_1 \setminus P'_2\}$, where $P'_1 = \{\{1\},\{2\},\{3\},\{4\}\}, P'_2 = \{J \in N : |J| = 2, J \in T_{\{1,4\},\{2,3\}}\}$, a tighter upper bound of 23.5 can be obtained by the grouping PM bound for these three problems, matching their sum capacity. As we can see, there is a common pattern among the problems of Figures 4.9(a), 4.9(b), and 4.9(c). That is, there is an isolated vertex, vertex 1. Also, in their capacity-achieving server grouping $\mathcal{P}$, P_1 only contains servers that have no more than one message apart from message 1. There also exists a common pattern among the problems of Figures 4.9(d), 4.9(e), and 4.9(f). That is, there are two disjoint cycles, cycle $\{1,4\}$ and cycle $\{2,3\}$. Also, in their capacity-achieving server grouping $\mathcal{P}'$, P'_2 only contains servers that have one message from each disjoint cycle.

Remark 4.10. For the problems in Figures 4.9(a), 4.9(b), and 4.9(c) with the capacity-achieving $\mathcal{P}$, the following constraints are given by Axiom (4.29) (the fd-separation axiom)

$$f(\{1\},\{j\}) + f(\{1\},\{2,3,4\} \setminus \{j\}) = f(\{1\},\{2,3,4\}), \qquad j \in \{2,3,4\}.$$

If the constraints above were to be removed from Theorem 4.7, then the upper bounds would become looser than 19. Similarly, for the problems in Figures 4.9(d), 4.9(e), and 4.9(f) with the capacity-achieving $\mathcal{P}'$, the following two constraints are given by Axiom (4.29)

$$\begin{aligned}f(\{1,2\},\{1\}) + f(\{1,2\},\{4\}) &= f(\{1,2\},\{1,4\}),\\f(\{1,2\},\{2\}) + f(\{1,2\},\{3\}) &= f(\{1,2\},\{2,3\}).\end{aligned}$$

If the constraints above were to be removed from Theorem 4.7, then the upper bounds would become looser than 23.5.

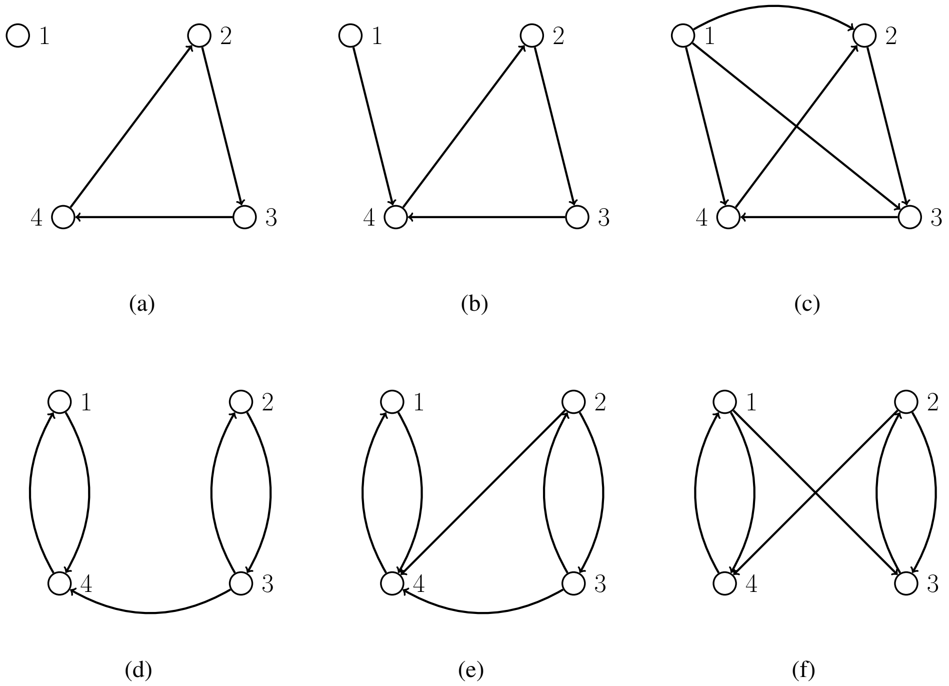


Figure 4.9: The side information graphs for the six 4-message problems given in (4.57). In Figures (a), (b) and (c), there is one isolated vertex, vertex 1, which has no incoming edges. In Figures (d), (e), and (f), there is a pair of disjoint cycles, $\{1,4\}$ and $\{2,3\}$.

The following definition generalizes the server grouping construction discussed in Example 4.11 to exploit fd-separation.

Definition 4.16 (fd grouping)**.** Consider the DIC problem $\mathcal{G}$: $(i|j \in A_i), i \in [n]$, whose side information graph $\mathcal{G}$ contains $k \geq 0$ mutually disjoint cycles, denoted by $K_j \subseteq [n], j \in [k]$, $K_j \cap K_{j'} = \emptyset$ for any $j, j' \in [k], j \neq j'$, and $|K_0| \geq 0$ isolated vertices, $v \in K_0 \subseteq [n]$. The fd grouping, denoted by $\mathcal{P}_{\text{fd}}$, with $m = (2^k - 1 - k) + 2 = 2^k - k + 1$ groups is defined as follows. The first server group is given by

$$P_1 = \{J \in N : |J \setminus K_0| \leq 1\}. \tag{4.58}$$

For $\ell = 2, \cdots, m-1$, the server groups are

$$P_{\tilde{G}_\ell} = \{J \in N : |J \setminus K_0| = |\tilde{G}_\ell|, J \in \bigcap_{j \in \tilde{G}_\ell} T_{K_j}\}, \tag{4.59}$$

where $\tilde{G}_\ell \subseteq [k], |\tilde{G}_\ell| \geq 2$. Note that there are $(2^k - 1 - k)$ such groups. Finally, the last server group P_m is defined as

$$P_m = N \setminus (\bigcup_{\tilde{G}_\ell \subseteq [k] : |\tilde{G}_\ell| \geq 2} P_{\tilde{G}}) \setminus P_1. \tag{4.60}$$

With $\mathcal{P}_{\text{fd}}$, we have the following corollary, namely the fd *grouping PM bound*, from Theorem 4.7. As the fd grouping does not result in any simplified expression compared to the grouping PM bound in Theorem 4.7, we do not repeat the rate constraint inequality (4.22) and Axioms (4.23)-(4.29) below.

Corollary 4.3 (fd grouping PM bound)**.** Consider the DIC problem $\mathcal{G}$: $(i|j \in A_i)$, $i \in [n]$ with link capacity tuple $\mathbf{C}$. Its capacity region $\mathscr{C}(\mathcal{G}, \mathbf{C})$ satisfies

$$\mathscr{C}(\mathcal{G}, \mathbf{C}) \subseteq \mathscr{R}_{\mathcal{P}_{\text{fd}}}(\mathcal{G}, \mathbf{C}),$$

where $\mathscr{R}_{\mathcal{P}_{\text{fd}}}(\mathcal{G}, \mathbf{C})$ denotes the outer bound given by the grouping PM bound with any valid fd grouping $\mathcal{P}_{\text{fd}}$.

The proof is trivial and omitted.

Note that the fd grouping and the fd grouping PM bound can both easily incorporate the set of active servers N_A.

We give an example showing the efficacy of the fd grouping PM bound when there are both disjoint cycles and isolated vertices in the side information graph.

Example 4.12. Consider the 5-message DIC problem $(1|-), (2|3), (3|2), (4|5), (5|4)$ with equal link capacities $C_J = 1, J \in N \setminus \{\emptyset\}$. The side information graph of the problem consists of two disjoint cycles $K_1 = \{2,3\}$ and $K_2 = \{4,5\}$, as well as one isolated vertex 1, and thus $K_0 = \{1\}$. For easier notation, set

$$Q_1 = \{\{1\},\{2\},\{3\},\{4\},\{5\},\{1,2\},\{1,3\},\{1,4\},\{1,5\}\}, \tag{4.61}$$

$$Q_2 = \{\{2,4\},\{2,5\},\{3,4\},\{3,5\},\{1,2,4\},\{1,2,5\},\{1,3,4\},\{1,3,5\}\}, \tag{4.62}$$

and $Q_3 = N \setminus Q_1 \setminus Q_2$. We have $m = 2^2 - 2 + 1 = 3$, and the fd grouping as $\mathcal{P}_{\text{fd}} = \{P_1, P_{\{1,2\}}, P_3\}$, where $P_1 = Q_1$, $P_{\{1,2\}} = Q_2$, and $P_3 = Q_3$. The fd grouping PM bound with $\mathcal{P}_{\text{fd}}$ yields $R_1 + R_2 + R_3 + R_4 + R_5 \leq 47\frac{2}{3}$, which is tight and matches the DCC lower bound in Theorem 3.8 on the sum capacity. For the latter, we use seven decoding message set tuples $\mathbf{D}_1, \mathbf{D}_2, \cdots, \mathbf{D}_7$ as follows.

In $\mathbf{D}_1$, we set $D_1 = \{1\}$ and $D_i = \{1, i\}$, $i \in \{2,3,4,5\}$.

In $\mathbf{D}_2$, we set $D_1 = \{1\}$, $D_i = [n] \setminus A_i$, $i \in \{2,3\}$, and $D_i = \{1, i\}$, $i \in \{4,5\}$.

In $\mathbf{D}_3$, we set $D_1 = \{1,4,5\}$, $D_i = [n] \setminus A_i$, $i \in \{2,3\}$, and $D_i = \{1, i\}$, $i \in \{4,5\}$.

In $\mathbf{D}_4$, we set $D_1 = \{1\}$, $D_i = \{1, i\}$, $i \in \{2,3\}$, and $D_i = [n] \setminus A_i$, $i \in \{4,5\}$.

In $\mathbf{D}_5$, we set $D_1 = \{1,2,3\}$, $D_i = \{1,i\}$, $i \in \{2,3\}$, and $D_i = [n] \setminus A_i$, $i \in \{4,5\}$.

In $\mathbf{D}_6$, we set $D_1 = \{1\}$, $D_i = [n] \setminus A_i$, $i \in \{2,3,4,5\}$.

In $\mathbf{D}_7$, we set $D_i = [n] \setminus A_i$, $i \in [n]$.

We also use the following three decoding server groups $\mathbf{P}_1, \mathbf{P}_2, \mathbf{P}_3$. In $\mathbf{P}_1$, we set $P_i = Q_1$, $i \in [n]$. In $\mathbf{P}_2$, we set $P_i = Q_2$, $i \in [n]$, and in $\mathbf{P}_3$, we set $P_i = Q_3$, $i \in [n]$. Hence, there are in total $7 * 3 = 21$ decoding configurations, $(\mathbf{P}_j, \mathbf{D}_k)$, $j \in [3], k \in [7]$. We set $R_i(\mathbf{P}, \mathbf{D}), i \in [n]$, $C_J(\mathbf{P}, \mathbf{D}), J \in N$, and $S_K(\mathbf{P}, \mathbf{D})$, $K \subseteq [n]$, to zero for all other $(\mathbf{P}, \mathbf{D})$ configurations. Notice that there is an interesting correspondence between the server groups in $\mathcal{P}_{\text{fd}}$ used in the outer bound and the decoding server groups used in the inner bound. Whether such correspondence has its roots in some deeper structural properties of the problem remains to be studied in future.

4.8 A Hierarchy of Server Groupings

In some DIC problems, it is more advantageous to use "finer" server groupings than what we have introduced so far. As alluded before in Remark 4.9, there is a natural hierarchy of server groupings in terms of tightness of the resulting outer bound. We need the following definition to formalize this.

Definition 4.17 (Grouping refinement and aggregation)**.** For any two valid server groupings $\mathcal{Q} = \{Q_1, Q_2, \cdots, Q_\ell\}$ and $\mathcal{P} = \{P_1, \cdots, P_m\}$. We say that $\mathcal{P}$ is a *refinement* of $\mathcal{Q}$ and that $\mathcal{Q}$ is an *aggregation* of $\mathcal{P}$, if for every $i \in [\ell]$, $Q_i = P_G$ for some $G \subseteq [m]$.

In words, every server group in $\mathcal{Q}$ is the union of some server groups in $\mathcal{P}$. We have the following relationship between the outer bounds $\mathscr{R}_{\mathcal{P}}$ and $\mathscr{R}_{\mathcal{Q}}$.

Proposition 4.3. If $\mathcal{P}$ is a refinement of $\mathcal{Q}$, or equivalently, $\mathcal{Q}$ is an aggregation of $\mathcal{P}$, then $\mathscr{R}_{\mathcal{P}}$ is no looser than $\mathscr{R}_{\mathcal{Q}}$, i.e.,

$$\mathscr{R}_{\mathcal{P}} \subseteq \mathscr{R}_{\mathcal{Q}}.$$

The proof is presented in Appendix B.8. Note that Proposition 4.3 clarifies the relationship between the individual touch grouping and any other touch grouping.

Definition 4.18 (Intersecting refinement of groupings)**.** For two valid server groupings $\mathcal{P}$ and $\mathcal{Q}$,

$$\mathcal{P} \wedge \mathcal{Q} = \{P \cap Q : P \in \mathcal{P}, Q \in \mathcal{Q}\} \tag{4.63}$$

is the *intersecting* refinement of both groupings.

Example 4.13. Consider the 5-message DIC problem $(1|2), (2|1), (3|5), (4|3), (5|4)$ with link capacities

$$C_J = 1, J \in N_A = \{\{1,3\}, \{1,4\}, \{1,5\}, \{2,3\}, \{2,4\}, \{2,5\},\\ \{3,4,5\}, \{1,3,4,5\}, \{2,3,4,5\}, \{1,2,3,4,5\}\}$$

and $C_J = 0$ otherwise. The touch grouping PM bound with

$$\mathcal{P}_{\mathrm{t}} = \{T_{\{1\}} \cap N_A, T_{\{2,3,4,5\}} \cap N_A\}, \tag{4.64}$$

and the fd grouping PM bound $\mathcal{P}_{\mathrm{fd}} = \{P_1, P_2 = N_A \setminus P_1\}$, where

$$P_1 = \{\{1,3\}, \{1,4\}, \{1,5\}, \{2,3\}, \{2,4\}, \{2,5\}\}, \tag{4.65}$$

both yield an upper bound of 14.5 on the sum capacity. With the intersecting refinement grouping,

$$\mathcal{P}_{\mathrm{t}} \wedge \mathcal{P}_{\mathrm{fd}} = \{T_{\{1\}} \cap P_1, T_{\{1\}} \cap P_2, T_{\{2,3,4,5\}} \cap P_1, T_{\{2,3,4,5\}} \cap P_2\}, \tag{4.66}$$

a tighter upper bound of 14 is obtained, which indeed matches the sum capacity of this problem. For the latter, we use Theorem 3.7 with $(\mathbf{P}, \underline{\mathbf{D}})$, where $P_i = N_A$, $i \in [n]$, and $\underline{\mathbf{D}}$ denotes the natural decoding message set tuples generated according to Algorithm 2.

Based on Proposition 4.3, we can establish the tightest grouping PM bound by using the "finest" server grouping $\mathcal{P}^* = \{\{J\} : J \in N \setminus \{\emptyset\}\}$ with $m = 2^n - 1$, referred to as the *single-server grouping*, which consists of all single nonempty servers and is a refinement of every other valid server grouping.

We present the following corollary, namely the *single-server grouping PM bound*, without repeating (4.22)-(4.29).

Corollary 4.4 (Single-server grouping PM bound)**.** Consider the DIC problem $\mathcal{G}$: $(i|j \in A_i)$, $i \in [n]$ with link capacity tuple $\mathbf{C}$. Its capacity region $\mathscr{C}(\mathcal{G}, \mathbf{C})$ satisfies

$$\mathscr{C}(\mathcal{G}, \mathbf{C}) \subseteq \mathscr{R}_{\mathcal{P}^*}(\mathcal{G}, \mathbf{C}),$$

where $\mathscr{R}_{\mathcal{P}^*}(\mathcal{G}, \mathbf{C})$ denotes the outer bound given by the grouping PM bound with the single-server grouping $\mathcal{P}^*$.

Remark 4.11. In a similar fashion as in Liu et al. [2018a], it can be shown that the single-server grouping PM bound $\mathscr{R}_{\mathcal{P}^*}$ is as tight as the apparently stronger outer bound in which all Shannon-type inequalities of the entropy function for the DIC problem are used.

If all servers are active, the computational complexity of $\mathscr{R}_{\mathcal{P}^*}$ is prohibitive even for small n as the number of the intermediate variables $f(G,K), G \subseteq [m], K \subseteq [n]$ in Corollary 4.4 is $2^{|N\setminus\{\emptyset\}|+n} = 2^{2^n-1+n}$, which is doubly exponential to n.

Finally, based on Proposition 4.3 we can establish the loosest grouping PM bound by using the "coarsest" server grouping $\mathcal{P}_* = \{N\}$ with $m = 1$, referred to as the *all-server grouping*, which consists of a single all-server group and is an aggregation of every other valid server grouping. With $\mathcal{P}_*$, the grouping PM bound in Theorem 4.7 simplifies to $\mathscr{R}_{\mathcal{P}_*}$, namely the *all-server grouping PM bound.*

Corollary 4.5 (All-server grouping PM bound)**.** Consider the DIC problem $\mathcal{G}$: $(i|j \in A_i)$, $i \in [n]$ with link capacity tuple $\mathbf{C}$. Its capacity region $\mathscr{C}(\mathcal{G},\mathbf{C})$ satisfies

$$\mathscr{C}(\mathcal{G},\mathbf{C}) \subseteq \mathscr{R}_{\mathcal{P}_*}(\mathcal{G},\mathbf{C}),$$

where $\mathscr{R}_{\mathcal{P}_*}(\mathcal{G},\mathbf{C})$ consists of all rate tuples $\mathbf{R}$ such that

$$R_i \leq f_*(B_i \cup \{i\}) - f_*(B_i), \qquad i \in [n], \tag{4.67}$$

for at least one set function $f_*(K), K \subseteq [n]$, satisfying

$$\begin{aligned} f_*(\emptyset) &= 0, && && (4.68)\\ f_*(K) &\leq \sum_{J:J\in T_K} C_J, && && (4.69)\\ f_*(K) &\leq f_*(K'), && K \subseteq K', && (4.70)\\ f_*(K\cap K') + f_*(K\cup K') &\leq f_*(K) + f_*(K'), && && (4.71)\\ f_*(B_i\cup\{i\}) - f_*(B_i) &= f_*(\{i\}), && i \in [n]. && (4.72) \end{aligned}$$

Proof. As $G \subseteq [1]$, G can be either $\emptyset$ or $\{1\}$. Also, $f(\emptyset,K) = 0$ for any $K \subseteq [n]$. Therefore, it suffices to use a single set function $f_*(K) = f(\{1\},K), K \subseteq [n]$ in the axioms and the rate constraint inequality in Corollary 4.5. With all-server grouping $\mathcal{P}_* = \{N\}, m = 1$, we have $(P_G \cup P_{G'}) \setminus (P_G \cap P_{G'}) \subseteq T_{\overline{K}}$ only for $G = G'$ or $K = \emptyset$, and thus Axiom (4.23) in Theorem 4.7 becomes trivial. Also, there never exists two disjoint nonempty sets $K, K' \subseteq [n]$ such that $P_{\{1\}} = N \subseteq (N \setminus T_{K,K'})$. This means that Axiom (4.29) can only give trivial inequalities with the all-server grouping. In summary, in Corollary 4.5, there are no constraints corresponding to Axioms (4.23) and (4.29) in Theorem 4.7. □

Even though the all-server grouping PM bound $\mathscr{R}_{\mathcal{P}_*}$ on the capacity region is the loosest bound one can get from the grouping PM bound, it is already no looser than the all-server distributed PM bound in Proposition 2.9 introduced in Sadeghi et al. [2016].

Proposition 4.4. For any DIC problem $\mathcal{G}$: $(i|j \in A_i), i \in [n]$, with link capacity tuple $\mathbf{C}$, it holds that $\mathscr{R}_{\mathcal{P}_*}(\mathcal{G}, \mathbf{C}) \subseteq \mathscr{R}_{\text{ADPM}}(\mathcal{G}, \mathbf{C})$.

The proof is presented in Appendix B.9.

The computational complexity of the all-server grouping PM bound $\mathscr{R}_{\mathcal{P}_*}$ will be the lowest among all the bounds from the grouping PM bound. And even performing FME to compute the outer bound on the entire capacity region for a general $\mathbf{C}$ is possible for small to moderate n as the total number of variables is only $2^n + 2^n - 1 + n$ in Corollary 4.5, accounting for 2^n $f_*(K), K \subseteq [n]$ variables, $2^n - 1$ link capacity variables $C_J, J \in N \setminus \{\emptyset\}$, and n rate variables $R_i, i \in [n]$.

Example 4.14. We first revisit the problem $(1|-), (2|4), (3|4), (4|3)$ discussed in Examples 4.10. A looser upper bound of 22 on the sum capacity is given by the all-server grouping PM bound $\mathscr{R}_{\mathcal{P}_*}$ in comparison to the tight bound established earlier. However, $\mathscr{R}_{\mathcal{P}_*}$ can sometimes yield tight bounds. For example, consider the problem $(1|4), (2|1,4), (3|1,2,4), (4|1,2,3)$ with equal link capacities $C_J = 1, J \in N \setminus \{\emptyset\}$. The all-server grouping PM bound $\mathscr{R}_{\mathcal{P}_*}$ yields the tight upper bound of 22 on the sum capacity, which matches the lower bound in Corollary 3.2 with $P_i = N$ and $D_i = [n] \setminus A_i$, $i \in [n]$.

4.9 Summary of the Speical Forms of the Grouping PM Bound

Different server groupings and their corresponding performance bounds presented in Sections 4.6-4.8 are summarized in Table 4.1.

4.10 Numerical Results for the Grouping PM Bound

We numerically evaluate lower and upper bounds on the sum capacity for all 218 non-isomorphic 4-message DIC problems with equal link capacities, $C_J = 1$, $J \in N \setminus \{\emptyset\}$. For brevity, each problem in this section is represented with a problem number and the corresponding receiver side information can be found in [Liu et al., 2020c, Appendix J]. The upper bounds on the sum capacity are computed using the special cases of the grouping PM bound proposed in Sections 4.6-4.8. The lower bound are given by the DCC scheme, computed using a fixed decoding configuration in Theorem 3.7. For both upper and lower bounds, we use LP to maximize the sum-rate $R_1 + R_2 + R_3 + R_4$. It turns out that the lower bounds match the upper bounds, thus establishing the sum capacity, for all 218 problems.

The results are summarized in Table 4.2. On the right column, each tuple denotes a list of problem numbers, followed by their sum-capacity in bold face. For example, (16, 30, 60, 102; **19**) means that problems 16, $\cdots$, 102 have the same sum-capacity of 19. The left column shows special cases of the grouping PM bound that are used to yield the tight upper bounds on

Table 4.1: Special cases of the grouping PM bound and their indicative computational complexities.

Server groupings	Entropic variables	References	Total # of variables
The touch grouping $\mathcal{P}_{\text{t}} = \{T_{L_1}, \ldots, T_{L_m}\}$ with $m \leq n$ groups: $f(G,K), G \subseteq [m], K \subseteq [n]$	$\frac{1}{r}H(Y_{T_{L_G}}\|X_{K^c})$	Cor. 4.2	$2^{m+n} + 2^n - 1 + n$
The individual touch grouping $\mathcal{P}_{\text{t}}^* = \{T_{\{1\}}, \ldots, T_{\{n\}}\}$ with $m = n$ groups: $f(G,K), G, K \subseteq [n]$	$\frac{1}{r}H(Y_{T_G}\|X_{K^c})$	(4.49), Rmk. 4.9	$2^{2n} + 2^n - 1 + n$
The fd grouping $\mathcal{P}_{\text{fd}}$ constructed according to (4.58)-(4.60) with $m = 2^k - k + 1$ groups, where k is the number of pairwisely disjoint cycles: $f(G,K), G \subseteq [m], K \subseteq [n]$	$\frac{1}{r}H(Y_{P_G}\|X_{K^c})$	Cor. 4.3, Exm. 4.11, 4.12	$2^{2^k-k+1+n} + 2^n - 1 + n$
The intersecting refined grouping $\mathcal{P} \wedge \mathcal{Q} = \{P \cap Q : P \in \mathcal{P}, Q \in \mathcal{Q}\}$ with $m\ell$ groups, where $m = \|\mathcal{P}\|, \ell = \|\mathcal{Q}\|$: $f(G,K), G \subseteq [m\ell], K \subseteq [n]$	$\frac{1}{r}H(Y_{(P \cap Q)_G}\|X_{K^c})$	Def. 4.17, 4.18, Prop. 4.3, Exm. 4.13	$2^{m\ell+n} + 2^n - 1 + n$
The single-server grouping $\mathcal{P}^* = \{\{J\} : J \in N \setminus \{\emptyset\}\}$ with $m = 2^n - 1$: $f(G,K), G \subseteq [2^n - 1], K \subseteq [n]$	$\frac{1}{r}H(Y_{P_G}\|X_{K^c})$	Cor. 4.4	$2^{2^n-1+n} + 2^n - 1 + n$
The all-server grouping $\mathcal{P}_* = \{N\}$ with $m = 1$ server group: $g(K), K \subseteq [n]$	$\frac{1}{r}H(Y_N\|X_{K^c})$	Cor. 4.5, Exm. 4.14	$2^n + 2^n - 1 + n$

the sum-capacity. It turns out that the all-server grouping PM bound $\mathscr{R}_{\mathcal{P}_*}$ in Corollary 4.5 with $m = 1$ group can solve 145 out of 218 problems with minimum computational complexity. For the 63 problems shown in the middle row in Table 4.2, tight upper bounds on the sum capacity can be obtained by the touch grouping PM bound $\mathscr{R}_{\mathcal{P}_{\text{t}}}$ with $\mathcal{P}_{\text{V}}$ defined in (4.46). Notice that for these aforementioned $63 + 145 = 208$ problems, except for the 6 problems (149, 176, 179, 200, 203, 212; **26**) with sum capacity of 26, they are also solvable by the augmentation group bound in Theorem 4.8. For the remaining 10 problems, the touch grouping PM bound $\mathscr{R}_{\mathcal{P}_{\text{t}}}$ gives loose results, and the fd grouping PM bound $\mathscr{R}_{\mathcal{P}_{\text{fd}}}$ is necessary to yield tight upper bounds. A subset of these problems were discussed in Example 4.11 and shown in Figure 4.9 and they all involve either isolated vertices or disjoint cycles in their side information graph.

For the lower bounds on the sum capacity for all 218 problems, we used $(\mathbf{P}, \mathbf{D})$ in Theorem 3.7 where $P_i = N$, $i \in [n]$, and $\mathbf{D} = \underline{\mathbf{D}}$, the natural decoding configuration generated according to Algorithm 2.

Table 4.2: Sum capacity for all 218 non-isomorphic 4-message DIC problems with equal link capacities $C_J = 1, J \in N \setminus \{\emptyset\}$.

Upper bounds	(Problem numbers; **sum capacity**)
The all-server grouping PM bound $\mathscr{R}_{\mathcal{P}_*}$ in Corollary 4.5	(1, 2, 3, 5, 6, 7, 8, 10, 11, 12, 13, 15, 17, 19, 20, 22, 25, 26, 33, 35, 38, 39, 40, 41, 49, 63, 65, 67, 69, 70, 100; **15**), (47; **18.6667**), (43, 78, 83, 85, 130, 132; **20**), (42, 44, 45, 71, 72, 73, 74, 75, 76, 77, 79, 80, 82, 84, 103, 104, 105, 106, 107, 108, 109, 110, 111, 113, 116, 117, 118, 120, 122, 123, 124, 125, 126, 127, 128, 131, 133, 142, 143, 144, 145, 147, 151, 152, 153, 154, 158, 159, 161, 162, 163, 164, 165, 166, 167, 168, 169, 174, 177, 182, 183, 184, 185, 186, 187, 201; **22**), (114, 121, 129, 146, 150, 155, 156, 157, 160, 170, 171, 175, 178, 180, 181, 188, 189, 190, 191, 192, 194, 195, 196, 197, 198, 202, 204, 206, 208, 210, 216; **24**), (207; **26**), (193, 205, 209, 211, 213, 214, 215, 217; **28**), (218; **32**)
The touch grouping PM bound $\mathscr{R}_{\mathcal{P}_{\mathrm{t}}}$ in Corollary 4.2 with $\mathcal{P}_{\mathbf{V}}$ in (4.46)	(4, 9, 18, 21, 23, 24, 34, 36, 48, 55, 64, 66, 68, 86, 95, 99, 138; **19**), (14, 27, 28, 29, 31, 32, 37, 50, 51, 52, 53, 54, 56, 57, 58, 59, 61, 62, 87, 88, 89, 90, 91, 92, 94, 96, 97, 98, 101, 134, 136, 137, 139, 140, 141, 173; **21**), (93, 135, 172, 199; **25**), (149, 176, 179, 200, 203, 212; **26**)
The fd grouping PM bound $\mathscr{R}_{\mathcal{P}_{\mathrm{fd}}}$ in Corollary 4.3	(16, 30, 60, 102; **19**), (46; **23.3333**), (81, 112, 115, 119, 148; **23.5**)

4.11 Chapter Summary

In this chapter, we proposed a number of information-theoretic performance bounds for the CIC and DIC problems. For the CIC problem, we generalized the alignment chain model Maleki et al. [2014] to acyclic chain models, which capture concatenated set-level acyclic structures implied by the interfering message sets at receivers. Based on the acyclic chains, we developed iterative performance bounds on the broadcast rate of the CIC problem and illustrated their usefulness via a number of concrete examples. For the DIC problem, we began with deriving performance bounds based on certain simplified versions of the acyclic chain models. Whether there exist some stronger (tighter) DIC performance bounds based on the non-reduced form of the acyclic chain models remains unclear at the current stage. We provided nontrivial generalization of the polymatroidal (PM) bound Blasiak et al. [2011] for the CIC problem to the DIC scenario. More specifically, we proposed the grouping PM bound based on the PM axioms of the entropy function and flexible utilization of server groupings. Through employing server groupings of different granularity, a natural tradeoff between the tightness and computational complexity of the bound can be exploited. In particular, we introduced a number of useful explicit constructions for server grouping and specialized the grouping PM bound based on these groupings. We also rigorously proved the hierarchy of server groupings in terms of the tightness of their corresponding bounds.

Chapter 5

Security and Privacy

In this chapter, we study secure and private variants of the centralized index coding (CIC) problem.

In Section 5.1, we consider the CIC problem with *security constraints on the legitimate receivers themselves*. Instead of assuming the existence of an eavesdropper as in Dau et al. [2012]; Ong et al. [2016c, 2018]; Mojahedian et al. [2017]; Liu et al. [2018a], we impose security requirements on the legitimate receivers. That is, each receiver must decode the legitimate message it requests and, at the same time, cannot learn any single message from a certain subset of *prohibited* messages. Similar problem setup was first investigated in [Dau et al., 2012, Section IV-E] and later studied in Narayanan et al. [2020]. On the achievability side, we extend the fractional local partial clique covering scheme Arbabjolfaei and Kim [2014] (cf. Proposition 2.2) to meet such security constraints. On the converse side, we develop two information-theoretic performance bounds for the secure CIC problem. The structure of the performance bounds enables us to further develop two necessary conditions for a given CIC problem to be *securely feasible* (i.e., to have nonzero rates for every message).

In Section 5.2, we consider a *data publishing* problem under a multi-terminal guessing framework with side information, which is inspired by the CIC problem, but with a significant twist to place emphasis on privacy. Instead of trying to maximize the communication rate, in this new framework the server's goal is to balance the data privacy and utility performance in the broadcast, both of which are measured based on the success rate of correctly guessing either the messages themselves or an arbitrary random function thereof. Such framework has applications in various real-world scenarios where sensitive data needs to be broadcasted/published in the presence of an adversary, e.g., field data broadcasting from a paddock aggregator in the presence of a malicious agent in precision agriculture. We first derive two lower bounds on the privacy leakage given the message distribution and utility constraints. We then propose a greedy algorithm as the privacy-preserving mechanism, which is inspired by the agglomerative clustering method used in the information bottleneck Slonim and Tishby [2000] and privacy funnel problems Makhdoumi et al. [2014].

5.1 Index Coding with Security Constraints on Receivers

We refer to the centralized index coding problem with security constraints on legitimate receivers simply as the secure index coding problem. The system model of such problem is the same as that of the CIC problem as defined in Section 2.1 with an extra security constraint defined as follows. Throughout the section, we assume zero-error decoding instead of vanishing-error decoding.

We assume *weak security* constraints against the receivers. That is, for each receiver $i \in [n]$, there is a set of *prohibited* messages $P_i \subseteq B_i$, which the receiver is prohibited from learning. More specifically, receiver i should not be able to decode any information about each individual message $j \in P_i$ given the side information x_{A_i} and the received codeword y.[1] A $(\mathbf{t}, r) = ((t_i, i \in [n]), r)$ *secure index code* is defined by

- An encoder at the centralized server, $\phi : \prod_{i \in [n]} \{0,1\}^{t_i} \to \{0,1\}^r$, which maps the messages $x_{[n]}$ to an r-bit sequence y;
- n decoders, one for each receiver $i \in [n]$, such that $\psi_i : \{0,1\}^r \times \prod_{k \in A_i} \{0,1\}^{t_k} \to \{0,1\}^{t_i}$ maps the received sequence y and the side information x_{A_i} to $\hat{x}_i$.

To summarize, we say a rate tuple $\mathbf{R} = (R_i, i \in [n])$ is *securely* achievable if there exists a $(\mathbf{t}, r)$ secure index code satisfying

$$\textbf{Rate:} \quad R_i \le \frac{t_i}{r}, \qquad \forall i \in [n], \tag{5.1}$$

$$\textbf{Message:} \quad H(X_S | X_{S'}) = H(X_S) = \sum_{i \in S} t_i, \qquad \forall S, S' \subseteq [n], S \cap S' = \emptyset, \tag{5.2}$$

$$\textbf{Codeword:} \quad H(Y) \le r, \tag{5.3}$$

$$\textbf{Encoding:} \quad H(Y | X_{[n]}) = 0, \tag{5.4}$$

$$\textbf{Decoding:} \quad H(X_i | Y, X_{A_i}) = 0, \qquad \forall i \in [n], \tag{5.5}$$

$$\textbf{Security:} \quad I(X_j; Y | X_{A_i}) = 0, \qquad \forall j \in P_i, i \in [n], \tag{5.6}$$

where (5.1) is the definition of $R_i, i \in [n]$, (5.2) follows from the assumption that the messages are independent and uniformly distributed, (5.3) is due to the length of the codeword being r, (5.4) follows from the fact that y is a function of $x_{[n]}$, (5.5) is stipulated by the decoding requirement at receivers: $\mathrm{P}\{(\hat{X}_1, \ldots, \hat{X}_n) \neq (X_1, \ldots, X_n)\} = 0$, together with Fano's inequality, and (5.6) is stipulated by the security constraints on the receivers.

[1] In contrast, a *strong security* constraint would require that receiver i cannot learn any information about the prohibited message set x_{P_i}, rather just individual messages in that set, given the side information and received codeword. That is, $I(X_{P_i}; Y | X_{A_i}) = 0$. Such constraint is overly strong when imposed on legitimate receivers as it enforces that $R_j = H(X_j) = 0$ for any $j \in P_i$, $i \in [n]$.

Any instance of the secure index coding problem can be represented by a sequence $(i|A_i, P_i)$, $i \in [n]$ specifying the side information availability and security constraints at receivers. For example, for a 3-message secure index coding problem with $A_1 = \emptyset, A_2 = \{3\}, A_3 = \{2\}$, and $P_1 = \{2,3\}, P_2 = P_3 = \emptyset$, we write

$$(1|\emptyset, \{2,3\}), (2|\{3\}, \emptyset), (3|\{2\}, \emptyset). \tag{5.7}$$

Recall that the side information availability at receivers can also be represented by the side information graph $\mathcal{G}$ as defined in Section 2.1. The secure index coding problem $(i|A_i, P_i)$, $i \in [n]$ can also be represented by the tuple $(\mathbf{A}, \mathbf{P})$, where $\mathbf{A} \doteq (A_i, i \in [n])$ and $\mathbf{P} \doteq (P_i, i \in [n])$, or equivalently the tuple $(\mathcal{G}, \mathbf{P})$.

For the secure index coding problem $(\mathcal{G}, \mathbf{P})$, its capacity region $\mathscr{C}(\mathcal{G}, \mathbf{P})$ is the closure of the set of all rate tuples $\mathbf{R}$ that are securely achievable. The symmetric capacity is defined as

$$C_{\text{sym}}(\mathcal{G}, \mathbf{P}) \doteq \max\{R : (R, \cdots, R) \in \mathscr{C}(\mathcal{G}, \mathbf{P})\}. \tag{5.8}$$

The broadcast rate $\beta(\mathcal{G}, \mathbf{P})$ is defined as the reciprocal of the symmetric capacity as

$$\beta(\mathcal{G}, \mathbf{P}) \doteq \frac{1}{C_{\text{sym}}(\mathcal{G}, \mathbf{P})}. \tag{5.9}$$

5.1.1 A Secure Linear Coding Scheme

We first investigate the achievability aspect of the secure index coding problem. Note that a secure variant of the composite coding scheme has been studied in Liu et al. [2018a], establishing an inner bound on the secure capacity region. Like other random-coding based coding schemes, the secure composite coding scheme yields achievability results (i.e., achievable rate regions), yet does not directly lead to specific code design. To design a practical linear coding scheme for secure index coding, we extend the fractional local partial clique covering (FLPCC) coding scheme from Arbabjolfaei and Kim [2014], originally developed for the non-secure CIC problem. The FLPCC scheme was reviewed in Section 2.3.1, and its corresponding achievability bound $\mathscr{R}_{\text{FLPCC}}(\mathcal{G})$ was presented in Proposition 2.2.

In the following, we describe our extended secure fractional local partial clique covering (S-FLPCC) scheme for the secure index coding problem $(\mathcal{G}, \mathbf{P})$. Similar to the FLPCC scheme, the S-FLPCC scheme utilizes time sharing and rate splitting among a number of subproblems of $\mathcal{G}$, and applies an MDS code at the subproblem level, while each subproblem also uses an MDS code. The main difference with the FLPCC scheme is that for the S-FLPCC scheme, we consider only the subproblems that satisfy relevant security constraints.

More specifically, for each subproblem $\mathcal{G}|_L$ for some message set $L \subseteq [n]$ assigned with a nonzero fraction of the channel capacity $0 < \lambda_L \leq 1$, we use a systematic $(|L| +$

$\beta_{\text{MDS}}(\mathcal{G}|_L), |L|)$ MDS code such that every receiver $i \in L$ can decode sub-message $x_{i,L}$ at rate $\frac{\lambda_L}{\beta_{\text{MDS}}(\mathcal{G}|_L)}$ (cf. Proposition 2.1).

For each subproblem $\mathcal{G}|_L$, due to the nature of MDS codes, every receiver $i \in L$ will be able to decode all the sub-messages in L at rate $\frac{\lambda_L}{\beta_{\text{MDS}}(\mathcal{G}|_L)}$ from the corresponding parity symbols. Hence, we require that $L \cap P_i = \emptyset$, since otherwise receiver i will be able to obtain some information about the messages in $L \cap P_i$, which violates the security constraint (5.6).

On the other hand, any receiver $i \notin L$ acts like an eavesdropper to the subproblem $\mathcal{G}|_L$ if $L \cap P_i \neq \emptyset$. It has been shown in Ong et al. [2016c] that for $\mathcal{G}|_L$, there exists some systematic $(|L| + \beta_{\text{MDS}}(\mathcal{G}|_L), |L|)$ MDS code over a large enough finite field that is secure against an eavesdropper who knows less than $|L| - \beta_{\text{MDS}}(\mathcal{G}|_L) - 1$ messages within $\mathcal{G}|_L$ as its side information (see [Ong et al., 2016c, Theorem 1] and its proof for more details). Therefore, to make sure that receiver $i \notin L$ with $L \cap P_i \neq \emptyset$ cannot learn any single message from the parity symbols of $\mathcal{G}|_L$, we simply require that $|A_i \cap L| < |L| - \beta_{\text{MDS}}(\mathcal{G}|_L)$.

Referring to our system model, sub-messages $x_{i,L}, i \in L, \lambda_L > 0$ are independent of each other. Hence, by combining MDS code symbols from different subproblems a receiver cannot gain any extra information than considering the MDS code symbols for each subproblem separately. Therefore, the general security constraint in (5.6) can be satisfied as long as the aforementioned security constraints for each subproblem $\mathcal{G}|_L$ are satisfied.

We present the achievability bound corresponding to the S-FLPCC scheme below.

Theorem 5.1 (Secure fractional local partial clique covering (S-FLPCC) bound)**.** Consider the secure index coding problem $(\mathcal{G}, \mathbf{P})$. Its capacity region $\mathscr{C}(\mathcal{G}, \mathbf{P})$ is inner bounded by the rate region $\mathscr{R}_{\text{S-FLPCC}}(\mathcal{G}, \mathbf{P})$ that consists of all rate tuples $\mathbf{R}$ such that

$$R_i \leq \sum_{L \subseteq [n]: i \in L} \frac{\lambda_L}{\beta_{\text{MDS}}(\mathcal{G}|_L)}, \tag{5.10}$$

for some $\lambda_L, L \subseteq [n]$ satisfying

$$\lambda_L \in [0,1], \qquad \forall L \subseteq [n], \tag{5.11}$$

$$\sum_{L \subseteq [n]: L \not\subseteq A_i} \lambda_L \leq 1, \qquad \forall i \in [n], \tag{5.12}$$

$$P_i \cap L = \emptyset, \qquad \forall L \subseteq [n], \lambda_L > 0, i \in L, \tag{5.13}$$

$$P_i \cap L = \emptyset \text{ or } |A_i \cap L| < |L| - \beta_{\text{MDS}}(\mathcal{G}|_L), \qquad \forall L \subseteq [n], \lambda_L > 0, i \notin L. \tag{5.14}$$

Remark 5.1. Note that (5.10)-(5.12) together form the same achievable rate region for the FLPCC scheme for non-secure index coding in Proposition 2.2. The security constraints against receivers are enforced by (5.13) and (5.14).

The S-FLPCC can give tight result for some problem, as illustrated by the following example. For simplicity, we compute the securely achievable symmetric rate rather than the whole rate region.

Example 5.1. Consider the following 9-message secure index coding problem $(i|A_i, P_i)$, $i \in [n]$ with $P_i = B_i$ for any $i \in [9]$,

$$\begin{array}{ccc} (1|\{1\}^c, \emptyset), & (2|\{2\}^c, \emptyset), & (3|\{4,5,6,8,9\}, \{1,2,7\}), \\ (4|\{5,6,7,8\}, \{1,2,3,9\}), & (5|\{3,4,7,8,9\}, \{1,2,6\}), & (6|\{2,3,4,5,7,9\}, \{1,8\}), \\ (7|\{7\}^c, \emptyset), & (8|\{8\}^c, \emptyset), & (9|\{9\}^c, \emptyset). \end{array}$$

The symmetric rate $R = \frac{1}{4}$ can be securely achieved by assigning $\lambda_L = \frac{1}{4}$ to the subproblems $\mathcal{G}|_L$ for $L \in \{\{1,2,8\}, \{2,6,7,9\}, \{3,9\}, \{4,5\}\}$, which is optimal for this problem (see Example 5.3 in Section 5.1.2 for the matching converse result). For each subproblem $\mathcal{G}|_L$, we have $\beta_{\text{MDS}}(\mathcal{G}|_L) = 1$ (i.e., the induced subgraph $\mathcal{G}|_L$ is actually a clique). One can check that the security constraints (5.13) and (5.14) are met for each subproblem. For example, consider $L = \{2,6,7,9\}$. For any $i \in L$, $P_i \cap L = \emptyset$ and thus (5.13) is satisfied for this L. The receivers $i \in L^c = \{1,3,4,5,8\}$ can be divided into two groups. For $i \in \{1,8\}$, we have $P_i = \emptyset$ and thus $P_i \cap L = \emptyset$. For $i \in \{3,4,5\}$, we have $|A_i \cap L| = 2 < |L| - \beta_{\text{MDS}}(\mathcal{G}|_L) = 3$. Therefore, (5.14) is also satisfied for this L.

5.1.2 Performance Bounds and Necessary Conditions for Feasibility

In this section, we introduce two performance bounds for the secure index coding problem. We also investigate the feasibility of the secure index coding problem. By saying a secure index coding problem to be feasible we mean that at least one rate tuple that is nonzero for every message is securely achievable. Subsequently, a secure index coding problem is infeasible if every securely achievable rate tuple has at least one zero rate element. A simple sufficient condition for a secure index coding problem to be feasible is that the achievable rate region $\mathscr{R}_{\text{S}-\text{FLPCC}}$ in Theorem 5.1 contains at least one rate tuple that is nonzero for every message. In this section, we focus on developing necessary conditions for feasibility.

5.1.2.1 An Outer Bound on the Secure Capacity Region

Theorem 5.2 (Secure polymatroidal (S-PM) bound)**.** Consider the secure index coding problem $(\mathcal{G}, \mathbf{P})$. Its capacity region $\mathscr{C}(\mathcal{G}, \mathbf{P})$ is outer bounded by the rate region $\mathscr{R}_{\text{S}-\text{PM}}(\mathcal{G}, \mathbf{P})$ that consists of all rate tuples $\mathbf{R}$ such that

$$R_i \leq g(B_i \cup \{i\}) - g(B_i), \qquad \forall i \in [n], \tag{5.15}$$

for at least one set function $g : 2^{[n]} \to [0,1]$ such that

$$
\begin{aligned}
& g(\emptyset) = 0, && && (5.16)\\
& g([n]) \le 1, && && (5.17)\\
& g(K) \le g(K'), && \text{if } K \subseteq K', && (5.18)\\
& g(K \cap K') + g(K \cup K') \le g(K) + g(K'), && && (5.19)\\
& g(B_i \cup \{i\}) - g(B_i) = g(\{i\}), && \forall i \in [n], && (5.20)\\
& g(B_i) = g(B_i \setminus \{j\}), && \forall j \in P_i, i \in [n]. && (5.21)
\end{aligned}
$$

Proof. Define the set function $g : 2^{[n]} \to [0,1]$ as in (B.4) in Section B.1, which is repeated below as

$$g(S) \doteq \frac{1}{r} H(Y|X_{S^c}), \quad \forall S \subseteq [n]. \tag{5.22}$$

Properties (5.15)-(5.20) can be derived in the same manner as (2.35)-(2.40) in Proposition 2.7. Proof details can be found in [Arbabjolfaei and Kim, 2018, Section 5.2]. To show (5.21), for any $i \in [n]$, $j \in P_i$, we use (5.2), (5.5), (5.22), as well as the security constraints specified in (5.6), to obtain

$$
\begin{aligned}
rg(B_i) &= H(Y|X_{A_i \cup \{i\}})\\
&= H(X_i|Y, X_{A_i}) + H(Y|X_{A_i}) - H(X_i|X_{A_i})\\
&= H(Y|X_{A_i}) - I(X_j; Y|X_{A_i}) - H(X_i|X_{A_i})\\
&= H(Y|X_{A_i \cup \{j\}}) - H(X_i|X_{A_i \cup \{j\}})\\
&= H(Y|X_{A_i \cup \{i\} \cup \{j\}}) - H(X_i|Y, X_{A_i \cup \{j\}})\\
&= rg(B_i \setminus \{j\}),
\end{aligned} \tag{5.23}
$$

where the second and second last equalities are simply due to the chain rule. □

Similar to the PM bound Blasiak et al. [2011] in Proposition 2.7 and the grouping PM bound in Theorem 4.7, the S-PM bound presented above can be solved using optimization tools such as Fourier-Motzkin Elimination (FME) and Linear Programming (LP). While the S-PM is useful by itself, in below, we will further develop a series of converse results utilizing the set function g defined aboved in (5.22). Indeed, the results to be presented in the next two subsections hold for any set function g satisfying the properties (5.15)-(5.21).

5.1.2.2 A Partition of N Based on the Security Constraints

The security property (5.21) is the only difference between the S-PM bound and the normal PM bound for the non-secure CIC problem, which can be seen as a direct consequence of the security constraint (5.6). It enforces the value of the set function g defined in (5.22) to be equal for certain arguments. For the toy example in (5.7), $B_1 = P_1 = \{2,3\}$. Thus, by (5.21), we have $g(\{2,3\}) = g(\{2\}) = g(\{3\})$.

Moreover, combining properties (5.18) and (5.19) of g with (5.21) may result in g to be equal for even more arguments. In the above example, since $g(\{2,3\}) = g(\{2\})$, by (5.19) we have $g(\{1,2,3\}) \leq g(\{1,2\}) + g(\{2,3\}) - g(\{2\}) = g(\{1,2\})$, and by (5.18) we have $g(\{1,2,3\}) \geq g(\{1,2\})$. Thus $g(\{1,2,3\}) = g(\{1,2\})$. Similarly, $g(\{1,2,3\}) = g(\{1,3\})$.

Based on the above ideas, we formally define a partition on the set $N = 2^{[n]}$, namely the g-partition, denoted by $\mathcal{N} = \{N_1, N_2, \cdots, N_\gamma\}$.

Definition 5.1 (g-partition)**.** Given a secure index coding problem $(\mathcal{G}, \mathbf{P})$, its g-partition $\mathcal{N}$ can be constructed using the following steps:

1. For any receiver $i \in [n]$ whose P_i is nonempty, for any $T \subseteq B_i^c$, form $N(i,T) \in \mathcal{N}$ as

$$N(i,T) = \{T \cup B_i \setminus \{j\} : j \in P_i\} \cup \{T \cup B_i\}.$$

   Note that $N(i,T)$ is a subset of N. All elements $S \in N$ that are not in any subset $N(i,T)$ are placed in $N_{\text{remaining}}$, i.e.,

$$N_{\text{remaining}} = N \setminus (\cup_{T \subseteq B_i^c, i \in [n]: P_i \neq \emptyset} N(i,T)).$$

2. As long as there exist two subsets $N(i,T), N(i',T')$ such that $N(i,T) \cap N(i',T') \neq \emptyset$, we merge these two subsets into one new subset. We keep merging overlapping subsets until all subsets in $\mathcal{N}$ are disjoint, then we index the elements of $\mathcal{N}$ as $N_1, N_2, \cdots, N_\gamma$ in an arbitrary order, except for $N_\gamma = N_{\text{remaining}}$.

For a given secure index coding problem, one can verify that its g-partition is unique. We call any component within the g-partition except for the last one a g-subset. The values of the set function g with arguments from one g-subset are always equal, enforced by (5.18), (5.19), and (5.21).

Lemma 5.1. Consider the secure index coding problem $(\mathcal{G}, \mathbf{P})$ with g-partition $\mathcal{N} = \{N_1, N_2, \cdots, N_\gamma\}$. For any g-subset N_k, $k \in [\gamma - 1]$, we have

$$g(S) = g(S'), \qquad \forall S, S' \in N_k.$$

Proof. Step 2 in Definition 5.1 is simply merging subsets that have at least one common element. Therefore, it suffices to show that for any receiver $i \in [n]$ whose P_i is nonempty, any $T \subseteq B_i^c$, the initialized $N(i,T)$ in Step 1 satisfies that

$$g(S) = g(S'), \qquad \forall S, S' \in N(i,T). \tag{5.24}$$

Consider any such $N(i,T)$. For any $j \in P_i$, we have

$$\begin{aligned} g(T \cup B_i) &\leq g(T \cup B_i \setminus \{j\}) + g(B_i) - g(B_i \setminus \{j\}) \\ &= g(T \cup B_i \setminus \{j\}), \end{aligned} \tag{5.25}$$

where the inequality follows from the submodularity condition in (5.19) and the equality follows from (5.21). On the other hand, by the monotonicity condition in (5.18), we have

$$g(T \cup B_i) \geq g(T \cup B_i \setminus \{j\}). \tag{5.26}$$

Combining (5.25) and (5.26) yields $g(T \cup B_i) = g(T \cup B_i \setminus \{j\})$, which implies (5.24) and thus concludes the proof. □

Let g_{N_k} denote the value of the set function g of any set S that belongs to the g-subset N_k for some $k \in [\gamma - 1]$, within a given $\mathcal{N} = \{N_1, N_2, \cdots, N_\gamma\}$. That is,

$$g_{N_k} \doteq g(S), \qquad \forall S \in N_k. \tag{5.27}$$

We state our first necessary condition for a given secure index coding problem to be feasible based on its g-partition.

Theorem 5.3. Consider the secure index coding problem $(\mathcal{G}, \mathbf{P})$ with g-partition as $\mathcal{N} = \{N_1, N_2 \cdots, N_\gamma\}$. For any $k \in [\gamma - 1]$, if there exist some message sets $S, S' \in N_k$ and receiver $i \in [n]$ such that $S' \cup \{i\} \subseteq S$ and $S' \subseteq B_i$, then the problem is infeasible.

Proof. Consider any valid $(\mathbf{t}, r)$ secure index code and any securely achievable rate tuple $\mathbf{R}$. For any $k \in [\gamma - 1]$, suppose that there exist some $S, S' \in N_k$ and $i \in [n]$ such that $S' \cup \{i\} \subseteq S$ and $S' \subseteq B_i$. We have

$$g(S') = g(S) \geq g(S', i) \geq g(B_i \cup \{i\}) - g(B_i) + g(S') \geq R_i + g(S'), \tag{5.28}$$

where the equality follows from Lemma 5.1 with the fact that S and S' belong to the same g-subset N_k, the first inequality follows from the fact that $S' \cup \{i\} \subseteq S$ and (5.18), the second inequality follows from the fact that $S' \subseteq B_i$ and (5.19), and the third inequality follows from (5.15). Clearly, (5.28) implies $R_i = 0$ and thus the problem is infeasible. □

Remark 5.2. One common scenario where a problem is infeasible is that there exit two receivers $i \neq j \in [n]$ such that $A_i \subseteq A_j \cup \{j\}$ and $i \in P_j$. For example, see [Narayanan et al., 2018, Proposition 2]. In this case, receiver j, after decoding its requested message j, knows more messages than receiver i and thus can always mimic the behaviour of receiver i to decode message i. This violates the security constraint $i \in P_j$ and thus, the problem is infeasible. We can show that such scenario is captured by Theorem 5.3 as a special case. Since $i \in P_j \subseteq B_j$, there exists a g-subset N_k for some $k \in [\gamma - 1]$ within $\mathcal{N}$ such that $S = B_j \in N_k$ and $S' = B_j \setminus \{i\} \in N_k$. First, as $i \in B_j$, we have

$$S' \cup \{i\} = (B_j \setminus \{i\}) \cup \{i\} = B_j \subseteq B_j = S.$$

Second, since $A_i \subseteq A_j \cup \{j\}$, we have $B_i \cup \{i\} = A_i^c \supseteq (A_j \cup \{j\})^c = B_j$, which indicates that

$$S' = B_j \setminus \{i\} \subseteq B_i.$$

Therefore, according to Theorem 5.3, the problem is infeasible.

Example 5.2. Consider the 5-message secure index coding problem

$$(1|\{2,4,5\},\emptyset),(2|\{1,5\},\{4\}),(3|\emptyset,\{1,2,5\}),(4|\{2\},\{1\}),(5|\{1,2\},\emptyset).$$

By Definition 5.1, the g-partition of the problem can be written as $\mathcal{N} = \{N_1, \cdots, N_\gamma\}$ where $\gamma = 6$, and

$$\begin{aligned} &N_1 = \{\{3\},\{3,4\}\},\ N_2 = \{\{1,2,4,5\} \setminus \{j\} : j \in \{1,2,5\}\},\\ &N_3 = \{\{1,3\},\{1,3,4\}\},\ N_4 = \{\{2,3\},\{2,3,4\}\}\\ &N_5 = \{\{1,2,3\},\{1,2,3,4\},\{3,5\},\{1,3,5\},\{2,3,5\},\\ &\quad \{1,2,3,5\},\{3,4,5\},\{1,3,4,5\},\{2,3,4,5\},[5]\}, \end{aligned}$$

and $N_6 = N \setminus (\cup_{k \in [5]} N_k)$. Consider message sets $\{1,3,5\},\{1,3,4,5\} \in N_5$ and receiver $4 \in [5]$, we have $\{1,3,5\} \cup \{4\} \subseteq \{1,3,4,5\}$ and $\{1,3,5\} \subseteq B_4$. Thus, by Theorem 5.3 the problem is infeasible.

5.1.2.3 A Lower Bound on the Secure Broadcast Rate

The MAIS lower bound on the broadcast rate $\beta_{\text{MAIS}}(\mathcal{G})$ Bar-Yossef et al. [2011] in Proposition 2.5 can be trivially extended to the secure index coding problem as $\beta_{\text{MAIS}}(\mathcal{G}, \mathbf{P}) = \beta_{\text{MAIS}}(\mathcal{G})$. We present a slightly stronger statement in the following lemma.

Lemma 5.2. Consider the secure index coding problem $(\mathcal{G}, \mathbf{P})$. For any set function $g : 2^{[n]} \to [0,1]$ satisfying the properties (5.15)-(5.21), any message set $S \subseteq [n]$, and any securely achievable symmetric rate R, we have

$$g(S) \geq R \cdot \beta_{\text{MAIS}}(\mathcal{G}|_S).$$

Proof. Assume $u = \beta_{\text{MAIS}}(\mathcal{G}|_S)$. Then there exists some set $U = \{i_1, i_2, \cdots, i_u\} \subseteq S$ whose induced subgraph $\mathcal{G}|_U$ is acyclic. Hence, without loss of generality (see (2.12)), we have

$$\{i_1, i_2, \cdots, i_{p-1}\} \subseteq B_{i_p}, \qquad \forall p \in [u]. \tag{5.29}$$

Therefore, we have

$$\begin{aligned} g(S) &\geq g(\{i_1, i_2, \cdots, i_u\}) \\ &\geq g(i_1, i_2, \cdots, i_{u-1}) + R_{i_u} \\ &\geq g(i_1, i_2, \cdots, i_{u-2}) + R_{i_{u-1}} + R_{i_u} \\ &\vdots \\ &\geq \sum_{p \in [u]} R_{i_p} = R \cdot \beta_{\text{MAIS}}(\mathcal{G}|_S), \end{aligned}$$

where the first inequality follows from (5.18), the other inequalities follow from (5.15) and (5.19) together with (5.29), and the equality simply follows from the definition of u. □

Based upon the above lemma, we propose a new performance bound, namely, the secure maximum acyclic induced subgraph (S-MAIS) lower bound as follows.

Theorem 5.4 (Secure maximum acyclic induced subgraph (S-MAIS) bound)**.** Consider the secure index coding problem $(\mathcal{G}, \mathbf{P})$ with g-partition $\mathcal{N} = \{N_1, N_2, \cdots, N_\gamma\}$. Its broadcast rate $\beta(\mathcal{G}, \mathbf{P})$ is lower bounded by $\beta_{\text{S-MAIS}}(\mathcal{G}, \mathbf{P})$, which is constructed by the following steps:

1. For any subset N_k, $k \in [\gamma]$, initialize ρ_{N_k} as

$$\rho_{N_k} = \max_{S \in N_k} \beta_{\text{MAIS}}(\mathcal{G}|_S).$$

2. As long as there exist two g-subsets N_k, N_ℓ, $k \neq \ell \in [\gamma - 1]$ such that there exist some sets $S \in N_k$, $S' \in N_\ell$ satisfying that $S' \subseteq S$, and that

$$\beta_{\text{MAIS}}(\mathcal{G}|_{\{j \in S \setminus S' : S' \subseteq B_j\}}) + \rho_{N_\ell} > \rho_{N_k},$$

update $\rho_{N_k} \leftarrow \beta_{\text{MAIS}}(\mathcal{G}|_{\{j \in S \setminus S' : S' \subseteq B_j\}}) + \rho_{N_\ell}$.

3. If no such N_k, N_ℓ exist, set $\beta_{\text{S-MAIS}}(\mathcal{G}, \mathbf{P}) = \max_{k \in [\gamma]} \rho_{N_k}$ and terminate the algorithm.

Proof. We show that $\beta_{\mathrm{S-MAIS}}(\mathcal{G},\mathbf{P}) \le \beta(\mathcal{G},\mathbf{P})$, which is equivalent to showing that $1 \ge R \cdot \beta_{\mathrm{S-MAIS}}(\mathcal{G},\mathbf{P}) = \max_{k\in[\gamma]} \rho_{N_k}$ for any valid $(\mathbf{t},r)$ secure index code and any securely achievable symmetric rate R.

If $\max_{k\in[\gamma]} \rho_{N_k} = \rho_{N_\gamma}$, as ρ_{N_γ} remains unchanged since its initialization, we have

$$\begin{aligned} R \cdot \beta_{\mathrm{S-MAIS}}(\mathcal{G},\mathbf{P}) = R \cdot \rho_{N_\gamma} &= R \cdot \max_{S\in N_\gamma} \beta_{\mathrm{MAIS}}(\mathcal{G}|_S) \\ &\le R \cdot \beta_{\mathrm{MAIS}}(\mathcal{G}) \le 1. \end{aligned}$$

It remains to show that $1 \ge R \cdot \beta_{\mathrm{S-MAIS}}(\mathcal{G},\mathbf{P})$ when $\beta_{\mathrm{S-MAIS}}(\mathcal{G},\mathbf{P}) = \max_{k\in[\gamma]} \rho_{N_k} = \rho_{N_k}$ for some $k \in [\gamma-1]$. Note that $g(S) \le 1, \forall S \subseteq [n]$ according to (5.17) and (5.18). We show that $1 \ge R \cdot \beta_{\mathrm{S-MAIS}}(\mathcal{G},\mathbf{P})$ by showing a slightly stronger statement that

$$g_{N_k} = g(S) \ge R \cdot \rho_{N_k}, \qquad \forall S \in N_k, k \in [\gamma-1]. \tag{5.30}$$

Recall that g_{N_k} has been defined in (5.27).

By induction, it suffices to show that

1. for the initialized $\rho_{N_k} = \max_{S\in N_k} \beta_{\mathrm{MAIS}}(\mathcal{G}|_S), k \in [\gamma-1]$, (5.30) holds;

2. for any N_k, N_ℓ, $k \ne \ell \in [\gamma-1]$ satisfying the conditions in Step 2 in Theorem 5.4, the updated $\rho_{N_k} = \beta_{\mathrm{MAIS}}(\mathcal{G}|_{\{j\in S\setminus S' : S'\subseteq B_j\}}) + \rho_{N_\ell}$ still satisfies (5.30), provided that $g_{N_\ell} \ge R \cdot \rho_{N_\ell}$.

Consider the initialized $\rho_{N_k} = \max_{S\in N_k} \beta_{\mathrm{MAIS}}(\mathcal{G}|_S), k \in [\gamma-1]$. By Lemmas 5.1 and 5.2, we have

$$g_{N_k} \ge R \cdot \beta_{\mathrm{MAIS}}(\mathcal{G}|_S), \qquad \forall S \in N_k,$$

which directly leads to (5.30).

Consider any $N_k, N_\ell, k \ne \ell \in [\gamma-1]$ and $S \in N_k, S' \in N_\ell$ satisfying the conditions in Step 2 in Theorem 5.4. We have the updated ρ_{N_k} as $\rho_{N_k} = \beta_{\mathrm{MAIS}}(\mathcal{G}|_{\{j\in S\setminus S' : S'\subseteq B_j\}}) + \rho_{N_\ell}$. Set

$$s = \beta_{\mathrm{MAIS}}(\mathcal{G}|_{\{j\in S\setminus S' : S'\subseteq B_j\}}). \tag{5.31}$$

Then, by (2.12), there exists some set $\{j_1, j_2, \cdots, j_s\} \subseteq \{j \in S \setminus S' : S' \subseteq B_j\}$ whose induced subgraph is acyclic satisfying that

$$\{j_1, \cdots, j_{p-1}\} \subseteq B_{j_p}, \qquad \forall p \in [s]. \tag{5.32}$$

Note that we also have

$$S' \subseteq B_{j_p}, \qquad \forall p \in [s], \tag{5.33}$$

since any j_p is an element of the set $\{j \in S \setminus S' : S' \subseteq B_j\}$. Combining (5.32) and (5.33) we have

$$S' \cup \{j_1, \cdots, j_{p-1}\} \subseteq B_{j_p}, \qquad \forall p \in [s]. \tag{5.34}$$

Hence, we have

$$\begin{aligned} g_{N_k} &\geq g(S' \cup \{j \in S \setminus S' : S' \in B_j\}) \\ &\geq g(S' \cup \{j_1, j_2, \cdots, j_s\}) \\ &\geq g(S' \cup \{j_1, j_2, \cdots, j_{s-1}\}) + R_{j_s} \\ &\geq g(S' \cup \{j_1, j_2, \cdots, j_{s-2}\}) + R_{j_{s-1}} + R_{j_s} \\ &\vdots \\ &\geq g(S') + \sum_{p \in [s]} R_{j_p} \\ &= g_{N_\ell} + R \cdot \beta_{\mathrm{MAIS}}(\mathcal{G}|_{\{j \in S \setminus S' : S' \subseteq B_j\}}) \\ &\geq R \cdot (\rho_{N_\ell} + \beta_{\mathrm{MAIS}}(\mathcal{G}|_{\{j \in S \setminus S' : S' \subseteq B_j\}})) = R \cdot \rho_{N_k}, \end{aligned}$$

where the first and second inequalities follow from (5.18), the inequalities except for the first two and the last one follow from (5.15) and (5.19) together with (5.34), the first equality follows from (5.31), the last inequality follows from the assumption that $g_{N_\ell} \geq R \cdot \rho_{N_\ell}$, and, finally, the last equality follows from the fact that the updated $\rho_{N_k} = \rho_{N_\ell} + \beta_{\mathrm{MAIS}}(\mathcal{G}|_{\{j \in S \setminus S' : S' \subseteq B_j\}})$. □

Based on the g-partition in Definition 5.1 and the S-MAIS bound in Theorem 5.4, we state our second necessary condition for feasibility of secure index coding below.

Theorem 5.5. Consider the secure index coding problem $(\mathcal{G}, \mathbf{P})$ with g-partition as $\mathcal{N} = \{N_1, N_2, \cdots, N_\gamma\}$. The problem is infeasible if there exists some $k \in [\gamma - 1]$ such that

$$\rho_{N_k} > \min_{S \in N_k} |S|,$$

where $\rho_{N_k}, k \in [\gamma - 1]$ are iteratively defined by Steps 1-3 in Theorem 5.4.

Proof. Suppose for some $k \in [\gamma - 1]$, we have $\rho_{N_k} > \min_{S \in N_k} |S|$. We show that the problem is infeasible by contradiction.

Assume that the problem is feasible. Then there exists some $((t, t, \ldots, t), r)$ secure index code such that some symmetric rate $R > 0$ is securely achievable. For any message $j \in [n]$

and any subset of its interfering message set $B \subseteq B_j$, setting $A = (B \cup \{j\})^c$, we have

$$\begin{aligned} \frac{t}{r} &= \frac{H(X_j)}{r} \\ &= \frac{1}{r}(H(X_j|X_A) - H(X_j|Y, X_A)) \\ &= \frac{1}{r}(I(X_j; Y|X_A)) \\ &= \frac{1}{r}(H(Y|X_A) - H(Y|X_{A\cup\{j\}})) \\ &= g(B \cup \{j\}) - g(B), \end{aligned} \tag{5.35}$$

where the first equality follows from that message x_j is uniformly distributed and of length t specified in (5.2), the second equality follows from the message independence also specified in (5.2) and the decoding condition in (5.5) together with $A_j = (B_j \cup \{j\})^c \subseteq (B \cup \{j\})^c = A$, the third and fourth equalities simply follow from the definition of mutual information, and the last equality follows from the definition of the set function g in (5.22). If we set $B = \emptyset$ in (5.35), by (5.16) we have

$$\frac{t}{r} = g(\{j\}). \tag{5.36}$$

Recall that for some $k \in [\gamma - 1]$, we have $\rho_{N_k} > \min_{S\in N_k} |S|$. There exists some $S_0 \in N_k$ such that $|S_0| = \min_{S\in N_k} |S|$. Then, we have

$$g(S_0) \leq \sum_{i\in S_0} g(\{i\}) = \sum_{i\in S_0} \frac{t}{r} = \frac{t}{r} \cdot \min_{s\in N_k} |S|, \tag{5.37}$$

where the inequality follows from repeated application of (5.19), and the first equality follows from (5.36).

One can also prove that

$$g_{N_k} \geq \frac{t}{r} \cdot \rho_{N_k}, \qquad \forall S \in N_k, k \in [\gamma - 1], \tag{5.38}$$

in the similar manner as proving (5.30) in the proof of Theorem 5.4. Therefore,

$$\begin{aligned} \frac{t}{r} \cdot \min_{S\in N_k} |S| &\geq g(S_0) \\ &= g_{N_k} \geq \frac{t}{r} \cdot \rho_{N_k} > \frac{t}{r} \cdot \min_{S\in N_k} |S|, \end{aligned} \tag{5.39}$$

where the first, second, and the last inequality follow from (5.37), (5.38), and the assumption that $\rho_{N_k} > \min_{S\in N_k} |S|$, respectively. Clearly, (5.39) leads to a contradiction, and therefore the problem must be infeasible. □

The following two examples demonstrate the efficacy of Theorems 5.4 and 5.5 in providing tight bounds as well as identifying infeasible secure index coding problems.

Example 5.3. Revisit the 9-message secure index coding problem in Example 5.1. While the normal MAIS lower bound gives $\beta(\mathcal{G},\mathbf{P}) \geq \beta_{\text{MAIS}}(\mathcal{G},\mathbf{P}) = \beta_{\text{MAIS}}(\mathcal{G}) = 3$, the S-MAIS lower bound in Theorem 5.4 could give a strictly tighter result. More specifically, consider the two g-subsets N_k, N_ℓ, $k \neq \ell \in [\gamma]$ such that $S = \{1,2,3,4\} \in N_k$ and $S' = \{1,2\} \in N_\ell$. Since we have $\{1,2,6\} = B_i = P_i$ for receiver $i = 5$, by (5.21), we have

$$g(\{1,2\}) = g(\{1,2,6\}) = g(\{1,6\}),$$

which indicates that the message subset $\{1,6\}$ is also in the g-subset N_ℓ. It can be verified that $\max_{\tilde{S}\in N_k} \beta_{\text{MAIS}}(\mathcal{G}|_{\tilde{S}}) = \beta_{\text{MAIS}}(\mathcal{G}|_{\{1,2,3,4\}}) = 3$, and thus we initialize ρ_{N_k} as $\rho_{N_k} = 3$. It can also be verified that $\max_{\tilde{S}\in N_\ell} \beta_{\text{MAIS}}(\mathcal{G}|_{\tilde{S}}) = \beta_{\text{MAIS}}(\mathcal{G}|_{\{1,6\}}) = 2$, and thus we initialize ρ_{N_ℓ} as $\rho_{N_\ell} = 2$. However, we have

$$\begin{aligned}\beta_{\text{MAIS}}(\mathcal{G}|_{\{j\in S\setminus S': S'\subseteq B_j\}}) + \rho_{N_\ell} &= \beta_{\text{MAIS}}(\mathcal{G}|_{\{j\in\{3,4\}:\{1,2\}\subseteq B_j\}}) + \rho_{N_\ell} \\ &= \beta_{\text{MAIS}}(\mathcal{G}|_{\{3,4\}}) + \rho_{N_\ell} = 2 + 2 = 4 > 3 = \rho_{N_k}.\end{aligned}$$

Therefore, according to Step 2 in Theorem 5.4, we update ρ_{N_k} from 3 to 4. We keep repeating such search-and-update procedure untill we reach the termination condition in Step 3. It can be verified that we cannot obtain any result larger than 4, and thus we have

$$\beta(\mathcal{G},\mathbf{P}) \geq \beta_{\text{S-MAIS}}(\mathcal{G},\mathbf{P}) = 4 > 3 = \beta_{\text{MAIS}}(\mathcal{G},\mathbf{P}).$$

Therefore, any securely achievable symmetric rate satisfies that $R \leq 1/4$, which matches the result in Example 5.1 and thus establishes the symmetric capacity $C_{\text{sym}}(\mathcal{G},\mathbf{P})$ to be $1/4$.

Example 5.4. Revisit the 5-message secure index coding problem in Example 5.2, which is infeasible according to Theorem 5.3. One can also see that the problem is infeasible by Theorem 5.5 since for $N_5 \in \mathcal{N}$, $\rho_{N_5} = 4$, and thus

$$\min_{S\in N_5} |S| = |\{3,5\}| = 2 < 4 = \rho_{N_5}.$$

Remark 5.3. While all our examples discussed in this section have a small or moderate problem size n, Theorems 5.3-5.5 can also be relatively easily applied to large secure index coding problems (e.g., $n \geq 12$). In contrast, the S-PM bound in Theorem 5.2 has a relatively high computational complexity similar to the PM bound for the CIC problem, and thus can be computationally impractical when n is large. Nevertheless, the S-PM bound is more general than

Theorems 5.3-5.5: it always gives no looser converse result than the S-MAIS bound in Theorem 5.4, and when the necessary condition in Theorem 5.3 or 5.5 implies infeasibility for the problem, the S-PM bound will directly enforce the corresponding rate R_i to be zero.

5.2 Privacy-Utility Tradeoff in a Guessing Framework Inspired by Index Coding

In this section, we study the tradeoff in privacy and utility in a single-trial multi-terminal guessing (estimation) framework using a system model that is inspired by index coding. Instead of maximizing the communication rate, in this new framework the sender's goal is to balance the privacy and utility performance in the broadcast, both measured based on the success rate of correctly guessing either the messages themselves or an arbitrary random function thereof. More specifically, we consider multiple independent sources (messages) available at a data curator[2] and assume that there are multiple legitimate users (receivers), as well as one adversary in the system. Each party (a user or the adversary) knows some sources a priori as side information. The data curator broadcasts (discloses) a distorted function of the sources to the users, which is also overheard by the adversary.

For the adversary, we adopt the maximal leakage introduced in Issa et al. [2020] as the privacy metric. It measures the worst-case information leakage in terms of the gain of the adversary in maximum a posteriori estimation of any target function of the unknown sources *after and before* observing the disclosed data. For the legitimate users, we define our utility metric such that it also reflects the improvement in users' guessing ability of the sources themselves. That is, as opposed to the adversary, we assume that the legitimate users are interested in the sources themselves. Quite often in practice different sources are of different levels of priority to a user. We capture this by dividing the unknown sources of each legitimate user into two subsets. Some essential sources are required to be perfectly reconstructed by the user. That is, the correct guessing probability of such sources after observing the disclosed data is non-negotiable and must be 1. The remaining sources are less critical, for which the user requires the guessing gain about each such source to be larger than a certain negotiable threshold. The privacy-utility tradeoff is cast as a constrained optimization problem, where the objective is to minimize the privacy leakage to the adversary, conditioned that the requirements on the utility of the unknown sources for the legitimate users are satisfied.

Given such system setting, we derive fundamental performance lower bounds on the maximal leakage to the adversary, which are inspired by the notion of confusion graph Alon et al. [2008] and the polymatroidal (PM) bound Blasiak et al. [2011]; Arbabjolfaei et al. [2013] (cf.

[2] Note that in this section we assume a general discrete finite-alphabet distribution for each source, rather than the uniform distribution normally assumed for the index coding problem. We still require the sources to be independent of each other.

Proposition 2.7) for the centralized index coding (CIC) problem. We also detail a greedy privacy enhancing mechanism, which is inspired by the agglomerative clustering algorithms in the information bottleneck Slonim and Tishby [2000] and privacy funnel Makhdoumi et al. [2014] problems.

Recall that for any discrete random variable Z with probability distribution P_Z, we denote its alphabet by $\mathcal{Z}$ with realizations $z \in \mathcal{Z}$. We denote an estimation of Z by $\hat{Z}$, whose alphabet is also $\mathcal{Z}$.

5.2.1 System Model and Problem Formulation

In this section, we provide a more detailed description on our multi-terminal guessing framework.

Assume that a data curator observes n independent discrete random sources $X_1, X_2, \ldots, X_n$. We assume a general distribution P_{X_i} for each source. Without loss of generality, we assume every source has full support. For brevity, when we say source i, we mean source X_i. For any $S \subseteq [n]$, set $\mathcal{X}_S = \prod_{i \in S} \mathcal{X}_i$. The data curator broadcasts a distorted version of $X_{[n]}$, denoted by Y, generated according to the privacy mechanism $P_{Y|X_{[n]}}$, to a number of legitimate users, which is also overheard by a single adversary.

Both the legitimate users and the adversary attempt to guess (estimate) a certain parameter V of interest about the sources in a single trial. Both the privacy and utility are defined in terms of a guessing gain r defined as follows. Consider any party (a user or the adversary) that wishes to guess V with the side information Z. Note that V and Z are independent of each other due to source independence. The party aims to maximize the correct guessing probability of V upon observing Y (i.e., the party employs the maximum *a posteriori* estimator). For each $(y, z) \in \mathcal{Y} \times \mathcal{Z}$, we define the ratio between such maximized guessing probability after and before observing a $y \in \mathcal{Y}$ given $z \in \mathcal{Z}$ as

$$r(V \to y|z) \doteq \frac{\max\limits_{P_{\hat{V}|Y=y,Z=z}} \mathbb{E}\left[P_{\hat{V}|Y=y,Z=z}(V|y,z)\right]}{\max\limits_{P_{\hat{V}|Z=z}} \mathbb{E}\left[P_{\hat{V}|Z=z}(V|z)\right]} \tag{5.40}$$

$$= \frac{\max\limits_{v} P_{V|Y,Z}(v|y,z)}{\max\limits_{v} P_{V|Z}(v|z)}. \tag{5.41}$$

Adversary: We assume the adversary has side information X_P for some $P \subseteq [n]$, and is interested in a (possibly randomized) discrete function U of the sources X_Q it does not know, where $Q \doteq P^c$. That is, we have the Markov chain $U - X_Q - (Y, X_P)$. Quite often in practice, this function is chosen by the adversary, and is unknown to the data curator. Therefore, we consider a worst-case privacy leakage measure, a conditional version of the maximal leakage from Issa et al. [2020], as our privacy metric.

Definition 5.2 (Maximal leakage, Issa et al. [2020])**.** Given a finite discrete joint distribution $P_{X_{[n]},Y}$, the maximal leakage from X_Q to Y given X_P is defined as

$$\mathcal{L}_{\max}(X_Q \to Y|X_P) \doteq \sup_{U:U-X_Q-(Y,X_P)} \mathcal{L}(U \to Y|X_P), \tag{5.42}$$

where

$$\mathcal{L}(U \to Y|X_P) \doteq \log \mathbb{E}_{P_{Y,X_P}}\big[r(U \to Y|X_P)\big]. \tag{5.43}$$

Remark 5.4. Note that the maximal leakage in Definition 5.2 assumes a different Markov chain model from [Issa et al., 2020, Section III-E]: for our problem, the Markov chain model studied in Issa et al. [2020] always reduces to $U - X_{[n]} - Y$ regardless of Q.

When there is no ambiguity, we refer to $\mathcal{L}_{\max}(X_Q \to Y|X_P)$ simply as $\mathcal{L}_{\max}$. A computable expression of the maximal leakage in Definition 5.2 is presented as

$$\mathcal{L}_{\max} = \log \sum_{y,x_P} \max_{x_Q} P_{Y,X_P|X_Q}(y, x_P|x_Q), \tag{5.44}$$

which can be obtained following a similar approach to Issa et al. [2020]. Note that the right hand side of (5.44) is equal to the Sibson mutual information of order ∞ Sibson [1969]; Verdú [2015].

Legitimate Users: There are m legitimate users. User $i \in [m]$ knows some sources X_{A_i} a priori as side information for some $A_i \subseteq [n]$, and is interested in all the remaining sources $X_{A_i^c}$. More specifically, for each user i, the sources $X_{A_i^c}$ are divided into two groups of different levels of priority:

- Source X_{W_i} for some $W_i \subseteq A_i^c$ are indispensable to the user, and thus must be correctly guessed by user i with probability of 1 (i.e., perfect decoding).
- The rest of the sources, X_{G_i}, where $G_i \doteq A_i^c \setminus W_i$, are less essential, yet still useful/interesting. Thus, user i requires the guessing ability gain upon observing Y to be larger than a certain threshold d_i.

These result in the following two kinds of utility constraints. For any $i \in [m]$, we have

$$H(X_{W_i}|Y, X_{A_i}) = 0, \qquad \forall i \in [m], \tag{5.45}$$

$$D(X_{G_i} \to Y|X_{A_i}) \geq d_i, \qquad \forall i \in [m], \tag{5.46}$$

where $D(X_{G_i} \to Y|X_{A_i})$ is defined as

$$D(X_{G_i} \to Y|X_{A_i}) \doteq \mathbb{E}_{P_{Y,X_{A_i}}}\big[\log r(X_{G_i} \to Y|X_{A_i})\big]. \tag{5.47}$$

Note that for constraint (5.46), each legitimate user i is interested in obtaining the source X_{G_i} rather than a function/feature of X_{G_i}. We simplify the notation $D(X_{G_i} \to Y|X_{A_i})$ to D_i when there is no ambiguity.

Remark 5.5. Note the subtle difference between D_i and

$$\mathcal{L}(X_{G_i} \to Y|X_{A_i}) = \log \mathbb{E}_{P_{Y,X_{A_i}}} \left[r(X_{G_i} \to Y|X_{A_i}) \right],$$

$\mathcal{L}$ defined in (5.43). The latter is lower bounded by the former due to Jensen's inequality. From the data curator's viewpoint, requesting D_i to be above a certain threshold is more stringent than requesting $\mathcal{L}(X_{G_i} \to Y|X_{A_i})$ to be above that threshold. We use D_i rather than $\mathcal{L}(X_{G_i} \to Y|X_{A_i})$ as our utility measure as it leads to a simple closed-form result characterizing $\mathcal{L}_{\max}$ in terms of D_i (cf. Lemma 5.4).

Remark 5.6. The two types of utility constraints in (5.45) and (5.46) can be unified to be represented in terms of the same function D: The perfect decoding constraint (5.45) is equivalent to requiring that

$$\begin{aligned} D(X_{W_i} \to Y|X_{A_i}) &= \mathbb{E}_{P_{Y,X_{A_i}}} \left[\log \frac{1}{\max\limits_{x_{W_i}} P_{X_{W_i}}(x_{W_i})} \right] \\ &= H_\infty(X_{W_i}), \end{aligned} \tag{5.48}$$

where $H_\infty(X_{W_i})$ denotes the min-entropy (i.e., Rényi entropy of order ∞ Rényi [1961]) of X_{W_i}. One can show that for any $i \in [m]$, $D_i \leq H_\infty(X_{G_i})$. Consequently, to avoid an invalid system model, we always require that $0 \leq d_i \leq H_\infty(X_{G_i})$.

Privacy-Utility Tradeoff: We denote such system by the 5-tuple $(P_{X_{[n]}}, \mathbf{A}, \mathbf{W}, \mathbf{d}, P)$, where $\mathbf{A} \doteq (A_i, i \in [m])$, $\mathbf{W} \doteq (W_i, i \in [m])$, and $\mathbf{d} \doteq (d_i, i \in [m])$. Note that G_i is determined by W_i and A_i, and Q is determined by P.

To design the privacy mechanism $P_{Y|X_{[n]}}$, we need to consider the fundamental tradeoff between the privacy and utility. Any data distortion that reduces the information leakage to the adversary can decrease the utility obtained by the users. Such tradeoff is formulated by the following constrained optimization problem.

$$\inf_{\substack{P_{Y|X_{[n]}} \in \\ \mathcal{P}_{Y|X_{[n]}}(P_{X_{[n]}}, \mathbf{A}, \mathbf{W}, \mathbf{d}, P)}} \quad \mathcal{L}_{\max}(X_Q \to Y|X_P), \tag{5.49}$$

where $\mathcal{P}_{Y|X_{[n]}}(P_{X_{[n]}}, \mathbf{A}, \mathbf{W}, \mathbf{d}, P)$ denotes the collection of randomized mappings $P_{Y|X_{[n]}}$ that satisfy (5.45) and (5.46) for the system $(P_{X_{[n]}}, \mathbf{A}, \mathbf{W}, \mathbf{d}, P)$.

Due to the non-convexity of (5.49), instead of providing an explicit solution, we derive performance bounds and achievability results to characterize $\mathcal{L}_{\max}$ by taking inspiration from index coding.

5.2.2 Performance Bounds on the Privacy Leakage

We derive two information-theoretic lower bounds on the privacy leakage. One is based on the utility constraint (5.45) only, while the other is obtained based on both (5.45) and (5.46).

5.2.2.1 Lower Bound Based on the Confusion Graph

The utility constraint (5.45) indicates that for user i, any different realizations $x_{W_i} \neq x'_{W_i} \in \mathcal{X}_{W_i}$ must be distinguishable based on the released $y \in \mathcal{Y}$, as well as the user's side information $x_{A_i} \in \mathcal{X}_{A_i}$. To describe such distinguishability, we recall the notion of confusion graph for index coding Alon et al. [2008].

Definition 5.3 (Confusion graph, Alon et al. [2008])**.** Any two realizations of the n sources $x^1_{[n]}, x^2_{[n]} \in \mathcal{X}_{[n]}$ are *confusable* if there exists some user $i \in [m]$ such that $x^1_{W_i} \neq x^2_{W_i}$ and $x^1_{A_i} = x^2_{A_i}$. A confusion graph Γ is an undirected graph with $|\mathcal{X}_{[n]}|$ vertices such that every vertex corresponds to a realization $x_{[n]} \in \mathcal{X}_{[n]}$ and an edge connects two vertices if and only if their corresponding realizations are confusable.

See Figure 5.1 for an example of the confusion graph.

To ensure (5.45), a group of realizations of $X_{[n]}$ can be mapped to the same y with nonzero probability only if they are pairwisely not confusable. To show such statement, consider any two confusable realizations $x^1_{[n]}, x^2_{[n]}$ such that for some user i, $x^1_{W_i} \neq x^2_{W_i}$ and $x^1_{A_i} = x^2_{A_i}$. If $x^1_{[n]}, x^2_{[n]}$ are mapped to the same y with nonzero probability, then upon observing this y, user i will not be able to tell whether its requested sources X_{W_i} give $x^1_{W_i}$ or $x^2_{W_i}$ according to its side information (as $x^1_{A_i}$ and $x^2_{A_i}$ are simply the same). More rigorously, for any source set $S \subseteq [n]$, define

$$\begin{aligned} \mathcal{Y}(x_S) &\doteq \{y \in \mathcal{Y} : P_{Y|X_S}(y|x_S) > 0\}, \qquad && \forall x_S \in \mathcal{X}_S, \\ \mathcal{X}_S(y) &\doteq \{x_S \in \mathcal{X}_S : P_{Y|X_S}(y|x_S) > 0\}, && \forall y \in \mathcal{Y}. \end{aligned}$$

Then, we have the following lemma. We omit the proof as it can be simply done by contradiction.

Lemma 5.3. Given a $P_{Y|X_{[n]}}$ satisfying (5.45), for any two confusable $x^1_{[n]}, x^2_{[n]} \in \mathcal{X}_{[n]}$, we have $\mathcal{Y}(x^1_{[n]}) \cap \mathcal{Y}(x^2_{[n]}) = \emptyset$.

Given a source set $S \subseteq [n]$ and a specific realization $x_S \in \mathcal{X}_S$, we define $\Gamma(x_S)$ as the subgraph of Γ induced by all the vertices $x_{[n]}$ such that $x_{[n]} = (x_S, x_{S^c})$ for some $x_{S^c} \in \mathcal{X}_{S^c}$.

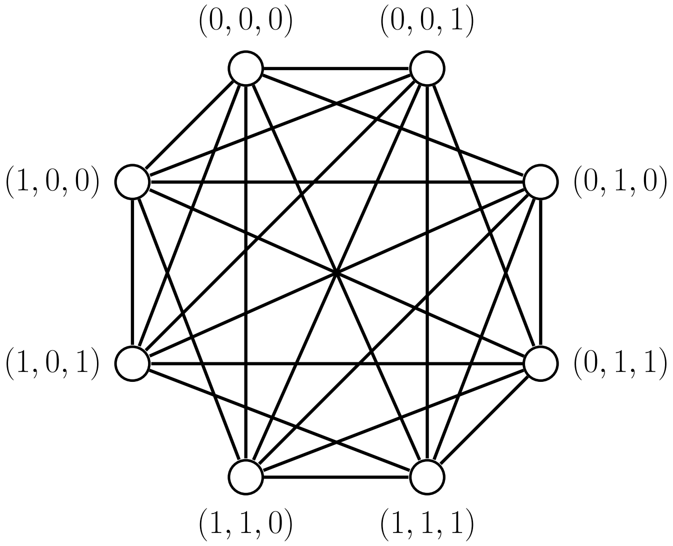


Figure 5.1: The confusion graph Γ when there are three binary sources, $\mathbf{W} = (\{1\}, \{2\}, \{3\})$, and $\mathbf{A} = (\emptyset, \{3\}, \{2\})$. For example, the two vertices corresponding to realizations $x^1_{[n]} = (0, 0, 0)$ and $x^2_{[n]} = (0, 0, 1)$ are confusable and thus connected in Γ, because at user $i = 3$, we have $x^1_{W_3} = 0 \neq x^2_{W_3} = 1$ and $x^1_{A_3} = x^2_{A_3} = 0$.

Notice that for any $x^1_S \neq x^2_S \in \mathcal{X}_S$ and $x^1_{S^c} \neq x^2_{S^c} \in \mathcal{X}_{S^c}$, $(x^1_S, x^1_{S^c})$ and $(x^1_S, x^2_{S^c})$ are confusable if and only if $(x^2_S, x^1_{S^c})$ and $(x^2_S, x^2_{S^c})$ are confusable. Hence, given the source set $S \subseteq [n]$, the subgraphs $\Gamma(x_S)$, $x_S \in \mathcal{X}_S$ are isomorphic to each other, and thus we simply denote any such subgraph by $\Gamma(S)$.

We present our first lower bound on the privacy leakage as follows.

Theorem 5.6. For the problem (5.49) with confusion graph Γ:

$$\mathcal{L}_{\max} \geq \log \omega(\Gamma(P)), \tag{5.50}$$

where $\omega(\cdot)$ is the clique number (size of the largest clique) of a graph.

Proof. Consider any $x_P \in \mathcal{X}_P$. There exists some realizations $x^1_Q, x^2_Q, \ldots, x_Q^{\omega(\Gamma(x_P))} \in \mathcal{X}_Q$ whose corresponding vertices in the subgraph $\Gamma(x_P)$ form a clique, which indicates that the realizations $(x_P, x^1_Q), (x_P, x^2_Q), \ldots, (x_P, x_Q^{\omega(\Gamma(x_P))})$ also form a clique in Γ. That is, the realizations $(x_P, x^1_Q), \ldots, (x_P, x_Q^{\omega(\Gamma(x_P))})$ are pairwisely confusable. Then by Lemma 5.3, for any $k \neq k' \in [\omega(\Gamma(x_P))]$, we have

$$\mathcal{Y}((x_P, x^k_Q)) \cap \mathcal{Y}((x_P, x^{k'}_Q)) = \emptyset. \tag{5.51}$$

Therefore, we have

$$\begin{aligned}\sum_{y}\max_{x_Q} P_{Y|X_{[n]}}(y|x_P,x_Q) &\geq \sum_{k\in[\omega(\Gamma(x_P))]}\ \sum_{y\in\mathcal{Y}(x_P,x_Q^k)} \max_{x_Q} P_{Y|X_{[n]}}(y|x_P,x_Q)\\ &\geq \sum_{k\in[\omega(\Gamma(x_P))]}\ \sum_{y\in\mathcal{Y}(x_P,x_Q^k)} P_{Y|X_{[n]}}(y|x_P,x_Q^k)\\ &= \sum_{k\in[\omega(\Gamma(x_P))]} 1 = \omega(\Gamma(x_P)) = \omega(\Gamma(P)),\end{aligned} \tag{5.52}$$

where the first inequality follows from (5.51). Hence, we have

$$\begin{aligned}\mathcal{L}_{\max} &= \log\sum_{x_P} P_{X_P}(x_P)\sum_{y\in\mathcal{Y}}\max_{x_Q} P_{Y|X_{[n]}}(y|x_P,x_Q)\\ &\geq \log\sum_{x_P} P_{X_P}(x_P)\cdot\omega(\Gamma(P)) = \log\omega(\Gamma(P)),\end{aligned}$$

where the first equality follows from the source independence, and the inequality follows from (5.52). □

5.2.2.2 Lower Bound Based on Polymatroidal Functions

We introduce a key lemma that serves as the baseline in the lower bound to be developed in this subsection.

Lemma 5.4. For the problem (5.49), we have

$$\mathcal{L}_{\max} \geq \max\{I(X_Q;Y|X_P), \max_{i\in[m]:G_i\subseteq Q}\Delta_i\}, \tag{5.53}$$

where $\Delta_i \doteq D_i + H(X_{W_i\cap Q}) + I(X_{A_i\cap Q};Y|X_P)$.

Proof. Recall that the maximal leakage $\mathcal{L}_{\max}$ is equal to the Sibson mutual information of order ∞ Sibson [1969]; Verdú [2015], denoted by I_∞^{S}, and hence we have

$$\mathcal{L}_{\max}(X_Q\to Y|X_P) = I_\infty^{\mathrm{S}}(X_Q;Y,X_P) \geq I(X_Q;Y,X_P) = I(X_Q;Y|X_P),$$

where the inequality follows from [Verdú, 2015, Theorem 2], and the last equality follows from the source independence.

It remains to show $\mathcal{L}_{\max}\geq\Delta_i$ for any user $i\in[m]$ with $G_i\subseteq Q$. For brevity, we drop the subscript i remembering that W, A, G stand for W_i, A_i, G_i, respectively. Set

$$\begin{aligned}W_P &= P\cap W, & W_Q &= Q\cap W,\\ A_P &= P\cap A, & A_Q &= Q\cap A.\end{aligned}$$

Since $G \subseteq Q$, we have $P \cap G = \emptyset$, $Q \cap G = G$, and thus $P = W_P \cup A_P$, and $Q = W_Q \cup A_Q \cup G$. We have

$$\begin{aligned}
&\mathcal{L}_{\max}(X_Q \to Y|X_P)\\
&\geq \sum_{y,x_P} P_{Y,X_P}(y,x_P) \log \max_{x_Q} \frac{P_{Y,X_P|X_Q}(y,x_P|x_Q)}{P_{Y,X_P}(y,x_P)} && (5.54)\\
&= \sum_{y,x_P} P_{Y,X_P}(y,x_P) \log \Big(\max_{x_{W_Q},x_{A_Q}} \Big(\frac{P_{Y,X_P|X_{W_Q},X_{A_Q}}(y,x_P|x_{W_Q},x_{A_Q})}{P_{Y,X_P}(y,x_P)}\\
&\quad \cdot \max_{x_G} \frac{P_{Y,X_P|X_Q}(y,x_P|x_Q)}{P_{Y,X_P|X_{W_Q},X_{A_Q}}(y,x_P|x_{W_Q},x_{A_Q})} \Big)\Big)\\
&= \sum_{y,x_P} P_{Y,X_P}(y,x_P) \log \Big(\max_{x_{W_Q},x_{A_Q}} \Big(\frac{P_{Y,X_P|X_{W_Q},X_{A_Q}}(y,x_P|x_{W_Q},x_{A_Q})}{P_{Y,X_P}(y,x_P)}\\
&\quad \cdot \max_{x_G} \frac{P_{Y,X_A|X_G}(y,x_A|x_G)}{P_{Y,X_A}(y,x_A)} \Big)\Big) && (5.55)\\
&\geq \sum_{y,x_P,x_{A_Q}} P_{Y,X_P,X_{A_Q}}(y,x_P,x_{A_Q}) \Big(\log \max_{x_{W_Q}} \frac{P_{Y,X_P|X_{W_Q},X_{A_Q}}(y,x_P|x_{W_Q},x_{A_Q})}{P_{Y,X_P}(y,x_P)}\\
&\quad + \log \max_{x_G} \frac{P_{Y,X_A|X_G}(y,x_A|x_G)}{P_{Y,X_A}(y,x_A)} \Big) && (5.56)\\
&= I(Y,X_P;X_{W_Q},X_{A_Q}) + \sum_{y,x_A} P_{Y,X_A}(y,x_A) \log \max_{x_G} \frac{P_{Y,X_A|X_G}(y,x_A|x_G)}{P_{Y,X_A}(y,x_A)} && (5.57)\\
&\geq I(Y,X_P;X_{W_Q},X_{A_Q}) + D_i && (5.58)\\
&= I(Y;X_{A_Q}|X_P) + H(X_{W_Q}) + D_i = \Delta_i, && (5.59)
\end{aligned}$$

where

- (5.54) follows from Jensen's inequality;

- (5.55) follows from the fact that for any $y \in \mathcal{Y}$, $x_{W\cup A} \in \mathcal{X}_{W\cup A}$,

$$\begin{aligned}
&\max_{x_G} \frac{P_{Y,X_P|X_Q}(y,x_P|x_Q)}{P_{Y,X_P|X_{W_Q},X_{A_Q}}(y,x_P|x_{W_Q},x_{A_Q})}\\
&= \max_{x_G} \frac{P_{Y,X_{A\cup G}}(y,x_{A\cup G}) \cdot P_{X_W|Y,X_{A\cup G}}(x_W|y,x_{A\cup G})}{P_{X_G}(x_G) \cdot P_{Y,X_A}(y,x_A) \cdot P_{X_W|Y,X_A}(x_W|y,x_A)}\\
&\overset{(a)}{=} \max_{x_G} \frac{P_{Y,X_A|X_G}(y,x_A|x_G)}{P_{Y,X_A}(y,x_A)},
\end{aligned}$$

where (a) follows from the Markov chain $X_W - (Y,X_A) - X_G$ as a result of the utility constraint (5.45);

- (5.56) follows from replacing maximum over x_{A_Q} with expectation over $P_{X_{A_Q}|Y,X_P}$, and Jensen's inequality;

- (5.57) follows from the fact that X_{W_Q} is a deterministic function of (Y, X_P, X_{A_Q}) according to (5.45);

- (5.58) follows from the fact that

$$\begin{aligned}
&\sum_{y,x_A} P_{Y,X_A}(y,x_A) \log \max_{x_G} \frac{P_{Y,X_A|X_G}(y,x_A|x_G)}{P_{Y,X_A}(y,x_A)} \\
&\geq \sum_{y,x_A} P_{Y,X_A}(y,x_A) \log \frac{\max_{x_G} P_{Y,X_A,X_G}(y,x_A,x_G)}{\max_{x_G} P_{X_G}(x_G) \cdot P_{Y,X_A}(y,x_A)} \\
&= \mathbb{E}_{P_{Y,X_{A_i}}} \left[\log r(X_{G_i} \to Y|X_{A_i}) \right] = D_i;
\end{aligned}$$

- (5.59) follows from the source independence, as well as (5.45).

This concludes the proof. □

For a given system $(P_{X_{[n]}}, \mathbf{A}, \mathbf{W}, \mathbf{d}, P)$, $H(X_{W_i \cap Q})$ has a fixed value and D_i is lower bounded by d_i according to (5.46). Hence, the only terms in (5.53) that still depend on $P_{Y|X_{[n]}}$ are the mutual information $I(X_P;Y|X_Q)$ and $I(X_{A_i \cap Q};Y|X_P)$. We further bound these mutual information terms below.

We draw inspiration from the polymatroidal (PM) bound Blasiak et al. [2011] in Proposition 2.7 for the CIC problem.

Lemma 5.5. Consider the system $(P_{X_{[n]}}, \mathbf{A}, \mathbf{W}, \mathbf{d}, P)$. For any disjoint $V, Z \subseteq [n]$, we have

$$I(X_V;Y|X_Z) = g(Z^c) - g(Z^c \cap V^c), \tag{5.60}$$

for some polymatroidal set function $g(S), S \subseteq [n]$ such that for any $i \in [n]$, $W \subseteq W_i$, $G \subseteq W^c \cap A_i^c$,

$$H(X_W) = g(G \cup W) - g(G), \tag{5.61}$$

and

$$g(\emptyset) = 0, \tag{5.62}$$

$$g(S') \geq g(S), \qquad \text{if } S \subseteq S', \tag{5.63}$$

$$g(S') + g(S) \geq g(S' \cup S) + g(S' \cap S). \tag{5.64}$$

Proof. Define $g(S) \doteq H(Y|X_{S^c}) - H(Y|X_{[n]}), \forall S \subseteq [n]$. We have $I(X_V;Y|X_Z) = g(Z^c) - g(Z^c \cap V^c)$. It remains to show that this $g(S)$ satisfies (5.61)-(5.64).

For (5.61), consider any $i \in [n]$, $W \subseteq W_i$, $G \subseteq W^c \cap A_i^c$. Set $A = [n] \setminus W \setminus G$, and one can verify that $A_i \subseteq A$. Hence,

$$\begin{aligned} H(X_W) &= H(X_W|X_A) - H(X_W|Y,X_A) \qquad (5.65)\\ &= H(Y|X_A) - H(Y|X_W,X_A)\\ &= g(W,G) - g(G), \end{aligned}$$

where (5.65) is due to source independence, (5.45), and $A_i \subseteq A$.

For (5.62), we have $g(\emptyset) = H(Y|X_{[n]}) - H(Y|X_{[n]}) = 0$.

For (5.63), for any $S \subseteq S' \subseteq [n]$, $S'^c \subseteq S^c$, and thus

$$g(S') = H(Y|X_{S'^c}) \geq H(Y|X_{S^c}) = g(S).$$

For (5.64), consider any $S, S' \subseteq [n]$. Set $S^c \cap S'^c = S_0$, $S^c \setminus S_0 = S_1$, and $S'^c \setminus S_0 = S_2$. We have

$$\begin{aligned} &g(S') + g(S)\\ &= H(Y,X_{S_0 \cup S_2}) + H(Y,X_{S_0 \cup S_1}) - H(X_{S_0 \cup S_2}) - H(X_{S_0 \cup S_1})\\ &\geq H(Y,X_{S_0}) + H(Y,X_{S_0 \cup S_1 \cup S_2}) - H(X_{S_0}) - H(X_{S_0 \cup S_1 \cup S_2})\\ &= g(S \cup S') + g(S \cap S'), \end{aligned}$$

where the inequality follows from the submodularity of the entropy function, as well as source independence. □

The above bound can be solved using optimization tools such as LP or FME similar to the PM bound in Proposition 2.7 and the grouping polymatroidal (GPM) bound in Theorem 4.7.

In the system $(P_{X_{[n]}}, \mathbf{A}, \mathbf{W}, \mathbf{d}, P)$, for any disjoint $V, Z \subseteq [n]$, let

$$\Lambda(V,Z) \doteq \min_{g \text{ satisfying (5.61)-(5.64)}} \{g(Z^c) - g(Z^c \cap V^c)\}.$$

Combining Lemmas 5.4 and 5.5 gives the following result.

Theorem 5.7. For the problem (5.49), we have

$$\mathcal{L}_{\max} \geq \max\{\Lambda(Q,P), \max_{i \in [m]: G_i \subseteq Q} \big(d_i + H(X_{W_i \cap Q}) + \Lambda(A_i \cap Q, P)\big)\}.$$

Remark 5.7. In the index coding problem, we usually assume uniformly distributed independent sources and a deterministic mapping $P_{Y|X_{[n]}}$. In contrast, Theorems 5.6 and 5.7 generally hold for any discrete independent source distribution and make no assumption on the privacy mechanism.

In general, Theorems 5.6 and 5.7 can outperform each other. See the numerical tests in Section 5.2.4 for more discussions about the comparison between the two performance bounds.

5.2.3 Privacy-Preserving Mechanism Design

We develop a greedy algorithm to provide a solution for the problem (5.49). The algorithm is based on the agglomerative clustering method, which has been used in the information bottleneck Slonim and Tishby [2000] and the privacy funnel problems Makhdoumi et al. [2014].

Consider a given system $(P_{X_{[n]}}, \mathbf{A}, \mathbf{W}, \mathbf{d}, P)$. To design the privacy-preserving mechanism $P_{Y|X_{[n]}}$, we start from the one-to-one deterministic mapping with $\mathcal{Y} = \mathcal{X}_{[n]}$ and $P_{Y|X_{[n]}}(y|x_{[n]}) = 1$ if and only if $y = x_{[n]}$[3] and then iteratively merge some elements of $\mathcal{Y}$ to make the privacy leakage smaller (in other words, we "blur" the revealed information), while still ensuring the utility for the users at an acceptable level, i.e., satisfying (5.45) and (5.46). In particular, to ensure (5.45), we again utilize the notion of confusion graph Alon et al. [2008] introduced in Section 5.2.2.1. See Figure 5.2 for an example about how the merging operation is restricted in order to satisfy (5.45).

Based on the merging idea discussed above, we propose an agglomerative clustering algorithm in Algorithm 4. Let Y^{y_1,y_2} be the resulting Y from merging any $y_1, y_2 \in \mathcal{Y}$. Let Θ denote the collection of $\{y_1, y_2\}$ such that merging them does not violate (5.45) and (5.46) and strictly reduces the privacy leakage to the adversary:

$$\begin{aligned}\Theta \doteq \{\{y_1,y_2\} \in \mathcal{Y}\times\mathcal{Y}\colon y_1 \neq y_2, &\text{ any two}\\ x_{[n]}, x'_{[n]} \in \mathcal{X}_{[n]}(y_1) \cup \mathcal{X}_{[n]}(y_2) &\text{ are not confusable},\\ D(X_{G_i} \to Y^{y_1,y_2}|X_{A_i}) \geq d_i, &\forall i \in [m],\\ \mathcal{L}_{\max}(X_Q \to Y^{y_1,y_2}|X_P) < \mathcal{L}_{\max}&(X_Q \to Y|X_P)\}.\end{aligned} \tag{5.66}$$

The algorithm terminates if Θ becomes an empty set.

Remark 5.8. In order to compute Algorithm 4 more efficiently, notice that in step 3 finding $\arg\min_{\{y_1,y_2\}\in\Theta} L_{\max}(X_Q \to Y^{y_1,y_2}|X_P)$ is equivalent to finding

$$\arg\max_{\{y_1,y_2\}\in\Theta} \left(2^{\mathcal{L}_{\max}(X_Q\to Y|X_P)} - 2^{\mathcal{L}_{\max}(X_Q\to Y^{y_1,y_2}|X_P)}\right),$$

[3] Note that such one-to-one mapping allows every user to perfectly reconstruct every source and thus definitely satisfies (5.45) and (5.46). Nevertheless, it also leads to the largest privacy leakage as the adversary can also perfectly reconstruct X_Q and subsequently any function U thereof it is interested in.

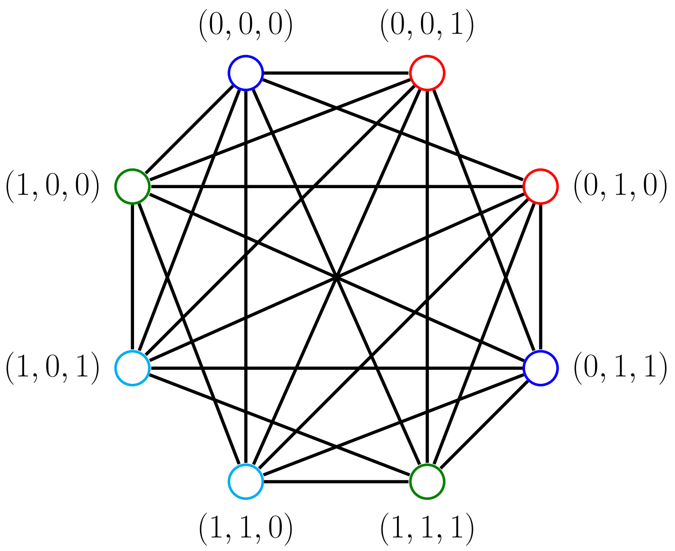


Figure 5.2: The confusion graph Γ when there are three binary sources, $\mathbf{W} = (\{1\}, \{2\}, \{3\})$, and $\mathbf{A} = (\emptyset, \{3\}, \{2\})$ (repeated from Figure 5.1). To help with understanding, we color-code the vertices of Γ such that to ensure (5.45), two codewords $y \neq y' \in \mathcal{Y}$ can be possibly merged if and only if their corresponding groups of source realizations $\mathcal{X}_{[n]}(y)$ and $\mathcal{X}_{[n]}(y')$ all belong to the same color class.

which can be computed as

$$\begin{aligned}
&2^{\mathcal{L}_{\max}(X_Q \to Y|X_P)} - 2^{\mathcal{L}_{\max}(X_Q \to Y^{y_1,y_2}|X_P)} \\
&= \sum_{x_P} \max_{x_Q} P_{X_P}(x_P) \cdot P_{Y|X_{[n]}}(y_1|x_P, x_Q) + \sum_{x_P} \max_{x_Q} P_{X_P}(x_P) \cdot P_{Y|X_{[n]}}(y_2|x_P, x_Q) \\
&\quad - \sum_{x_P} \max_{x_Q} P_{X_P}(x_P) \cdot P_{Y|X_{[n]}}(\bar{y}|x_P, x_Q) \\
&= \sum_{x_P \in \mathcal{X}_P(y_1)} P_{X_P}(x_P) \cdot 1 + \sum_{x_P \in \mathcal{X}_P(y_2)} P_{X_P}(x_P) \cdot 1 - \sum_{x_P \in \mathcal{X}_P(\bar{y})} P_{X_P}(x_P) \cdot 1 \\
&= P_{X_P}(\mathcal{X}_P(y_1)) + P_{X_P}(\mathcal{X}_P(y_2)) - P_{X_P}(\mathcal{X}_P(\bar{y})).
\end{aligned}$$

5.2.4 Numerical Test Results

To evaluate the performance of our proposed performance bounds and privacy mechanism, we consider 500 systems $(P_{X_{[n]}}, \mathbf{A}, \mathbf{W}, \mathbf{d}, P)$ randomly generated according to the following conditions:

- $n = m = 5$, $W_i = \{i\}$ for any user $i \in [5]$, and $\mathbf{A}$ is generated based on a randomly chosen graph $\mathcal{G}$ from the 9608 non-isomorphic 5-vertex directed graphs Arbabjolfaei and Kim [2018] such that $A_i = \{j \in [5] : (j, i) \in \mathcal{G}\}$.
- For any $i \in [5]$, $\mathcal{X}_i = \{0, 1\}$ and $X_i \sim \mathrm{Bern}(p_i)$, where p_i is uniformly randomly

Algorithm 4: Agglomerative clustering algorithm for problem (5.49)

```
input : The system $(P_{X_{[n]}}, \mathbf{W}, \mathbf{A}, \mathbf{d}, P)$.
output: Privacy-preserving mechanism $P_{Y|X_{[n]}}$.
Initialization: $\mathcal{Y} \leftarrow \mathcal{X}_{[n]}$, $P_{Y|X_{[n]}}(y|x_{[n]}) \leftarrow 1$ iff $y = x_{[n]}$ and obtain $\Theta$ based on $\mathcal{Y}$ by (5.66);
repeat
    $\{y_1^*, y_2^*\} \leftarrow \arg\min_{\{y_1,y_2\}\in\Theta} \mathcal{L}_{\max}(X_Q \to Y^{y,y'}|X_P)$;
    Merge $y_1^*$ and $y_2^*$ into $\bar{y}$: $\bar{y} \leftarrow \{y_1^*, y_2^*\}$;
    Obtain the new $Y$ by letting $\mathcal{Y} \leftarrow \mathcal{Y} \setminus \{y_1^*, y_2^*\} \cup \{\bar{y}\}$ and
      $P_{Y|X_{[n]}}(\bar{y}|x_{[n]}) \leftarrow P_{Y|X_{[n]}}(y_1^*|x_{[n]}) + P_{Y|X_{[n]}}(y_2^*|x_{[n]})$ for any $x_{[n]} \in \mathcal{X}_{[n]}$ while keeping
      the rest of $P_{Y|X_{[n]}}$ unchanged;
    Obtain the new $\Theta$ by (5.66) based on updated $Y$;
until $\Theta = \emptyset$;
return $P_{Y|X_{[n]}}$;
```

chosen from the interval $(0,1)$;

- For any $i \in [5]$, $d_i = \tilde{d}_i \cdot H_\infty(X_{G_i})$, where $\tilde{d}_i$ is uniformly randomly chosen from the interval $(0,1)$;
- $P \subseteq [5]$ is randomly generated assuring that $|P| \leq 2$. That is, we consider a relatively *weak* adversary that does not possess too many sources as side information.

For each system, we compute the lower bounds $\mathcal{L}_{\max}^{\text{Thm.1}}$ and $\mathcal{L}_{\max}^{\text{Thm.2}}$ by Theorems 5.6 and 5.7, respectively. An interesting observation is that we have $\mathcal{L}_{\max}^{\text{Thm.2}} > \mathcal{L}_{\max}^{\text{Thm.1}}$ for only 2 among the 500 tested systems, while $\mathcal{L}_{\max}^{\text{Thm.1}} > \mathcal{L}_{\max}^{\text{Thm.2}}$ for all the 498 remaining systems.

We also compute the maximal leakage according to the privacy-preserving mechanism $P_{Y|X_{[n]}}$ given by Algorithm 4, denoted as $\mathcal{L}_{\max}^{\text{Alg.1}}$. Then we compute the ratio

$$R = \mathcal{L}_{\max}^{\text{Alg.1}} / \max(\mathcal{L}_{\max}^{\text{Thm.1}}, \mathcal{L}_{\max}^{\text{Thm.2}}).$$

A lower ratio R (close to 1) means that our converse and achievable results perform well and are quite close to the optimal $\mathcal{L}_{\max}^*$, while a higher ratio indicates bad performance of the agglomerative privacy-preserving mechanism or loose converse bounds. We summarize the values of R from 500 tests in Table 5.1, from which we can see that the proposed techniques achieves a satisfactory level of performance for the majority of tested problems. Notably, we have $\mathcal{L}_{\max}^{\text{Thm.1}} = \mathcal{L}_{\max}^{\text{Alg.1}}$ and thus $R = 1$ for 162 among the 500 tests.

5.3 Chapter Summary

In this chapter, we studied the security and privacy aspects of (centralized) index coding. We formulated our secure index coding model by introducing security constraints on each receiver

Table 5.1: Performance of the Lower Bounds versus Algorithm 4.

Ratio, R	$= 1$	< 1.05	< 1.1	< 1.2	≥ 1.2
Number of Systems	162	401	429	460	40

that prevent the receiver to decode any messages in a certain prohibited message set. A linear secure coding scheme and two performance bounds on the secure capacity have been proposed. In addition, two necessary conditions for a given secure index coding problem to be feasible (i.e., have at least one securely achievable rate tuple that is nonzero for every message) were established. As for the privacy problem, we formulated a privacy-preserving data publishing problem in a multi-terminal guessing framework inspired by index coding, where a server (data curator) broadcasts messages to multiple legitimate receivers in the presence of an adversary. Instead of trying to maximize the communication rate, the goal of the server is to balance the privacy and utility performance in the system. Performance bounds on the optimal privacy leakage were derived, and a greedy agglomerative clustering algorithm was introduced as a practical privacy-preserving mechanism.

Chapter 6

Conclusion

In this chapter, we conclude this thesis with a brief summary on our contributions, together with a number of open questions and intriguing future directions.

In this thesis, we studied the index coding problem in both the classical centralized setup and the more general distributed scenario. In the centralized index coding (CIC) problem, a central server containing n messages broadcasts to n receivers with side information over a noiseless channel with unit capacity. In the distributed index coding (DIC) problem, all 2^n servers, each containing a different subset of the n messages in the system were taken into account. Towards the ultimate goal of characterizing the problems' capacity region, series of achievable coding schemes and performance bounds have been developed. The security and privacy aspects of the CIC problem have also been investigated.

For the achievability side, we took an information-theoretic perspective and built coding schemes based on the classic tool of random coding. For the CIC problem, we started from simplifying the existing random-coding-based composite coding scheme Arbabjolfaei et al. [2013] through excluding redundant composite index rates and decoding configurations from the expression and computation of its achievable rate region. We then extended the composite coding scheme by employing an enhanced fractional rate splitting technique, which strictly outperforms the standard convexification techniques based on time sharing. Following a different approach, the composite coding scheme was also generalized by allowing the composite indices to be further layered. In other words, we added one more layer of random coding into the original two-layer coding scheme. Subsequently, we developed the three-layer composite coding scheme, which leads to a strictly enlarged achievable rate region. Motivated by such advancement due to further layering of random coding, we ask the following fundamental question on composite coding.

Open Problem I: Can we always strictly enlarge the achievable rate region by adding more layers of random coding into the multi-layer composite coding scheme? What is the fundamental limit on such multi-layer composite coding?

One practical drawback of the multi-layer composite coding scheme would be the high computational complexity that increases rapidly as the number of random coding layers increases. Current simplification techniques for the composite coding and three-layer composite coding would be not efficient enough to address the forbiddingly high complexity when the number of layers is large. Therefore, it is of importance to extend and further improve the simplification methods in a systematic manner.

Open Problem II: Find the minimal set of necessary composite indices and decoding configurations for composite coding for a general CIC problem, and extend such result to the multi-layer composite coding schemes.

An interesting observation is that thus far no concrete CIC instance has been identified in the literature for which the composite coding scheme (even with enhanced convexification method and multiple layers of random coding) can strictly outperform the general vector linear coding schemes.

Open Problem III: Show that the linear capacity region (i.e., the closure of the set of all achievable rate tuples given by linear coding schemes) always subsumes the achievable rate region of composite coding (or its extensions), or provide a counterexample.

For the DIC problem, we generalized the composite coding scheme to the multi-server scenario. More specifically, we combined the key ideas of our enhanced fractional rate splitting with the cooperative compression method introduced in Li et al. [2018]. We also showed that the distributed composite coding can be presented in a series of equivalent or simplified forms.

The open problems for the centralized composite coding carries over to its distributed variant. Beyond the scope of composite coding, practical (linear) code design is an important open problem in the DIC problem.

Open Problem IV: Design linear coding schemes that are applicable to all DIC problems as a general purpose coding scheme and are implementable in practice with low complexity.

For the converse side of the capacity region, we introduced a number of performance bounds that characterize the fundamental limits on the communication rates imposed by the system model. Our performance bounds for the CIC problem were developed based on the acyclic chain model, which are generalized versions of the alignment chain model Maleki et al. [2014] that can capture concatenated acyclic structures among the interfering messages at the receivers. The acyclic chain bounds are strictly tighter than the maximum acyclic induced subgraph bound Bar-Yossef et al. [2011] and the internal conflict bound Maleki et al. [2014]; Jafar [2014], and, at the same time, less computationally intensive compared with the more general polymatroidal bounds Blasiak et al. [2011]. The proposed bounds have also been shown to satisfy several desirable properties.

Open Problem V: Characterize the time complexity of computing the acyclic chain bounds for a general CIC problem with n messages. And design algorithms for the computation or approximation of the acyclic chain bounds with low complexity (e.g., polynomial-time algorithms).

As for the DIC problem, we first proposed two performance bounds based on some less general variants of the acyclic chain model. Whether there exist more general iterative DIC performances as counterparts to the CIC acyclic chain bounds remains unknown at the moment.

Open Problem VI: Show the general iterative performance bounds based on the acyclic chain models for the DIC problem, which allow message sets rather than individual messages to be components of the chains, or provide counterexamples.

Next, we proposed a general performance bound based on the polymatroidal (PM) axioms of the entropy function, which can essentially capture all Shannon-type inequalities. The bound is referred to as the grouping PM bound as it utilizes general groupings of servers in the system with different levels of granularity. Such flexibility allows a natural tradeoff between computational complexity and tightness of the bound. We also specified a number of construction techniques for grouping servers and presented the corresponding specialized forms of the grouping PM bound.

Open Problem VII: Find the minimal set of inequalities needed to characterize the grouping PM bound with a given server grouping.

Open Problem VIII: The correspondence between the server groups used in the matching achievable and performance bounds for the problem in Example 4.12 is interesting and worth of future investigation. With deeper understanding of the possible reasons behind such correspondence, can we establish the capacity region for certain class of DIC problems by proving that the grouping PM bound matches the distributed composite coding bound? Or, can we prove that there is a constant (with respect to the system parameters) multiplicative gap between these two bounds for a general DIC problem?

In general, Shannon-type inequalities are not sufficient to obtain tight performance bounds on the capacity region even for the CIC problem Sun and Jafar [2015]; Baber et al. [2013]; Liu and Sadeghi [2019b]. As the CIC problem is a special case of the DIC problem, it follows automatically that non-Shannon-type inequalities are needed for some DIC instances.

Open Problem IX: Can we construct a class of CIC or DIC problems for which non-Shannon-type inequalities are needed? How much would non-Shannon-type inequalities improve upon Shannon-type inequalities (i.e., the grouping PM bound) for these problems?

In the last part of the thesis, we studied the secure and private (centralized) index coding problems in two different settings. In our secure index coding model, each receiver stipulates not only a decoding requirement for its wanted message, but also a set of security constraints prohibiting the receiver from learning any individual message that belongs to a certain prohibited message set. A linear coding scheme to ensure such secure communication was proposed. On the converse side, performance bounds on the secure capacity region were developed, based on which we could also establish two necessary conditions for a given problem to have at least one securely achievable rate tuple that is nonzero for every message. Further improvements on the secure coding schemes and performance bounds remain to be investigated. Beyond that, we would also like to list the following challenging research directions.

Open Problem X: Characterize the sufficient and necessary conditions for a given secure index coding to be feasible (i.e., to have at least one securely achievable rate tuple that is nonzero for every message).

Open Problem XI: Characterize the secure capacity region under the stronger security constraint that each receiver can learn no information about its entire prohibited message set, where secret keys shared among subsets of receivers and the server might be needed.

We formulated a privacy-preserving data publishing problem inspired by the index coding problem, but with a significant twist to place emphasis on privacy. We considered a multi-terminal guessing framework where the published (broadcasted) codeword from the server are observed by the legitimate receivers and a malicious adversary. The privacy and utility were measured according to the success rates of correctly guessing certain parameters of interest by the adversary and legitimate receivers, respectively. Performance bounds and privacy-preserving mechanism were proposed towards characterizing the fundamental privacy-utility tradeoff. For the performance bounds we drew inspiration from the confusion graph technique Alon et al. [2008] and the polymatroidal bound Blasiak et al. [2011]. For the privacy-preserving mechanism we adopted the method of agglomerative clustering from Slonim and Tishby [2000]; Makhdoumi et al. [2014].

Open Problem XII: Characterize the privacy-utility tradeoff with a communication rate constraint for the data publishing problem in the multi-terminal guessing framework.

Open Problem XIII: Characterize the privacy-utility tradeoff with or without a communication rate constraint for the data publishing problem when the privacy and utility are measured using different metrics other than the correctly guessing probability.

Appendix A

In Appendix A, we prove certain results presented in Chapter 3.

A.1 Proof of Theorem 3.7

Analysis of error for the first-step decoding is as follows. We partition the error event according to the collection $M \subseteq \Gamma_*(P_i) \setminus 2^{A_i}$ for erroneous composite indices. That is, $\hat{w}_K \neq w_K$ iff $K \in M$. Therefore, by the union bound, we have

$$\begin{aligned}
P_e &= \mathrm{P}\{y_J = y_J(\hat{w}_K, K \in 2^J) \text{ for all } J \in P_i \text{ for some} \\
&\qquad (\hat{w}_K, K \in \Gamma_*(P_i)) \neq (w_K, K \in \Gamma_*(P_i))\} \\
&\leq \sum_{M \subseteq \Gamma_*(P_i) \setminus 2^{A_i}} \sum_{\substack{(\hat{w}_K, K \in \Gamma_*(P_i)): \\ \hat{w}_K \neq w_K, K \in M, \\ \hat{w}_K = w_K, K \notin M}} \mathrm{P}\left\{ \bigcap_{J \in P_i, J \in \Gamma^*(M)} \left\{ y_J = y_J(\hat{w}_K, K \in 2^J) \right\} \right\} \\
&< \sum_{M \subseteq \Gamma_*(P_i) \setminus 2^{A_i}} 2^{\sum_{K \in M} s_K - \sum_{J \in \Gamma^*(M) \cap P_i} r_J} \qquad \text{(A.1)} \\
&< \sum_{M \subseteq \Gamma_*(P_i) \setminus 2^{A_i}} 2^{\sum_{K \in M} (rS_K + 1) - \sum_{J \in \Gamma^*(M) \cap P_i} (rC_J - 1)} \\
&= \sum_{M \subseteq \Gamma_*(P_i) \setminus 2^{A_i}} \frac{2^{r \sum_{K \in M} S_K} \cdot 2^{|M| + |\Gamma^*(M) \cap P_i|}}{2^{r \sum_{J \in \Gamma^*(M) \cap P_i} C_J}},
\end{aligned}$$

where (A.1) holds since for each composite index collection M, the number of erroneous tuples is $\prod_{K \in M} (2^{s_K} - 1) < 2^{\sum_{K \in M} s_K}$, and for each erroneous composite index tuple with $\hat{w}_K \neq w_K$, iff $K \in M$, the probability that it is mapped to the same codeword y_J as the correct composite index tuple for all $J \in \Gamma^*(M) \cap P_i$ is $2^{-\sum_{J \in \Gamma^*(M) \cap P_i} r_J}$. Note that only servers in $\Gamma^*(M)$ can generate composite index (indices) in the collection M and the intersection with P_i is necessary due to receiver i's choice of server group P_i.

Therefore, the error probability P_e tends to zero as $r \to \infty$, provided that

$$\sum_{K \in M} S_K < \sum_{J \in \Gamma^*(M) \cap P_i} C_J, \qquad \forall M \subseteq \Gamma_*(P_i) \setminus 2^{A_i}, i \in [n].$$

Analysis of error for the second-step decoding is as follows. Recall that we always require that $D_i \subseteq \cup_{J \in P_i} J, \forall i \in [n]$. We partition the error event according to message index subsets $L \subseteq D_i$. That is, $\hat{x}_j \neq x_j$ iff $j \in L$. Therefore, by the union bound, we have

$$\begin{aligned}
P_e &= \mathrm{P}\{w_K(\hat{x}_K) = \hat{w}_K \text{ for all } K \in \Gamma_*(P_i),\, K \subseteq D_i \cup A_i \text{ for some } \hat{x}_j \neq x_j,\, j \in D_i\} \\
&\leq \sum_{L \subseteq D_i} \sum_{\substack{\hat{x}_{D_i}: \\ \hat{x}_j \neq x_j, j \in L \\ \hat{x}_j = x_j, j \notin L}} \mathrm{P}\left\{ \bigcap_{\substack{K \subseteq D_i \cup A_i, \\ K \in \Gamma_*(P_i), \\ K \cap L \neq \emptyset}} \{w_K(\hat{x}_K) = \hat{w}_K\} \right\} \\
&\leq \sum_{L \subseteq D_i} 2^{\sum_{j \in L} t_j - \sum_{K \subseteq D_i \cup A_i, K \in \Gamma_*(P_i), K \cap L \neq \emptyset} s_K} \qquad \text{(A.2)} \\
&< \sum_{L \subseteq D_i} 2^{\sum_{j \in L} (rR_j + 1) - \sum_{K \subseteq D_i \cup A_i, K \in \Gamma_*(P_i), K \cap L \neq \emptyset} rS_K} \\
&= \sum_{L \subseteq D_i} 2^{r(\sum_{j \in L} R_j - \sum_{K \subseteq D_i \cup A_i, K \in \Gamma_*(P_i), K \cap L \neq \emptyset} S_K) + |L|},
\end{aligned}$$

where (A.2) holds since for each message index subset L, the number of erroneous messages is $\prod_{j \in L}(2^{t_j} - 1) < 2^{\sum_{j \in L} t_j}$ and for each erroneous message tuple with $\hat{x}_j \neq x_j$, iff $j \in L$, the probability that it is mapped to the same composite index w_K as the correct message tuple for all K such that $K \subseteq D_i \cup A_i, K \in \Gamma_*(P_i), K \cap L \neq \emptyset$ is $2^{-\sum_{K \subseteq D_i \cup A_i, K \in \Gamma_*(P_i), K \cap L \neq \emptyset} s_K}$.

Therefore, the error probability P_e tends to zero as $r \to \infty$, provided that

$$\sum_{j \in L} R_j < \sum_{K \subseteq \Delta_i \cup A_i, K \in \Gamma_*(P_i), K \cap L \neq \emptyset} S_K, \qquad \forall L \subseteq D_i, i \in [n].$$

A.2 Proof of Proposition 3.2

We first prove the following lemma.

Lemma A.1. If $P' \subseteq P$, then

$$\Gamma_*(P') \setminus \Gamma_*(P \setminus P') = \Gamma_*(P) \setminus \Gamma_*(P \setminus P'). \qquad \text{(A.3)}$$

Proof. As $P' \subseteq P$, we have that $\Gamma_*(P') \subseteq \Gamma_*(P)$, and therefore, $\Gamma_*(P') \setminus \Gamma_*(P \setminus P') \subseteq \Gamma_*(P) \setminus \Gamma_*(P \setminus P')$.

Now consider an arbitrary $J \in \Gamma_*(P) \setminus \Gamma_*(P \setminus P')$. As $J \in \Gamma_*(P)$, there must exist some

$J_1 \in P$ such that $J \subseteq J_1$. Since $J \notin \Gamma_*(P \setminus P')$, we know that $J_1 \notin P \setminus P'$. Hence, $J_1 \in P'$ and thus $J \in \Gamma_*(P')$. Therefore, we have $J \in \Gamma_*(P') \setminus \Gamma_*(P \setminus P')$ and thus $\Gamma_*(P) \setminus \Gamma_*(P \setminus P') \subseteq \Gamma_*(P') \setminus \Gamma_*(P \setminus P')$.

In summary, we have $\Gamma_*(P') \setminus \Gamma_*(P \setminus P') = \Gamma_*(P) \setminus \Gamma_*(P \setminus P')$. □

Now we prove Proposition 3.2 as follows.

For easier reference, we repeat (3.36) here for a given $(P_i, i \in [n])$,

$$\sum_{K \in M} S_K < \sum_{J \in \Gamma^*(M) \cap P_i} C_J, \qquad \forall M \subseteq \Gamma_*(P_i) \setminus 2^{A_i}, i \in [n]. \tag{A.4}$$

According to Lemma A.1, for any $Q \subseteq P_i$ we have

$$\Gamma_*(Q) \setminus \Gamma_*(P_i \setminus Q) \setminus 2^{A_i} = \Gamma_*(P_i) \setminus \Gamma_*(P_i \setminus Q) \setminus 2^{A_i}.$$

Therefore, for the same P_i as chosen above, (3.43) can be written as

$$\sum_{K \in \Gamma_*(P_i) \setminus \Gamma_*(P_i \setminus Q) \setminus 2^{A_i}} S_K < \sum_{J \in Q} C_J, \qquad \forall Q \subseteq P_i, i \in [n]. \tag{A.5}$$

We prove that for any given inequality from the system of inequalities (A.4) there exists an inequality in the system of inequalities (A.5) that is no looser and vice versa.

First, for any $M \subseteq \Gamma_*(P_i) \setminus 2^{A_i}$ in (A.4) we construct $Q = \Gamma^*(M) \cap P_i$. Therefore, the RHS of (A.4) and (A.5) become identical. Our claim is that $M \subseteq \Gamma_*(P_i) \setminus \Gamma_*(P_i \setminus Q) \setminus 2^{A_i}$ or that if $K \in M$ then $K \in \Gamma_*(P_i) \setminus \Gamma_*(P_i \setminus Q) \setminus 2^{A_i}$. Note that we have $M \subseteq \Gamma_*(P_i) \setminus 2^{A_i}$, therefore, if $K \in M$ it is automatic that $K \in \Gamma_*(P_i) \setminus 2^{A_i}$. So it remains to show that $K \notin \Gamma_*(P_i \setminus Q)$, which can be proven by contradiction as follows. For any $K \in M$, assume that $K \in \Gamma_*(P_i \setminus Q) = \Gamma_*(P_i \setminus \Gamma^*(M))$, which indicates that there exists some $J \in P_i \setminus \Gamma^*(M)$ such that $K \subseteq J$. However, as $K \in M$ and $K \subseteq J$, we have $J \in \Gamma^*(M)$, which contradicts with $J \in P_i \setminus \Gamma^*(M)$. Therefore, for any $K \in M$, we must have $K \notin \Gamma_*(P_i \setminus Q)$. In summary, for any given $M \subseteq \Gamma_*(P_i) \setminus 2^{A_i}$ and the corresponding inequality from (A.4), we have proved that there exists an inequality in (A.5) that is no looser.

To prove the other direction, for any $Q \subseteq P_i$ we construct $M = \Gamma_*(P_i) \setminus \Gamma_*(P_i \setminus Q) \setminus 2^{A_i}$. Therefore, the LHS of (A.4) and (A.5) become identical. Our claim is that if $J \in \Gamma^*(M) \cap P_i$ then $J \in Q$ or that $\Gamma^*(M) \cap P_i \subseteq Q$. Before proving our claim, we show that with the choice of $M = \Gamma_*(P_i) \setminus \Gamma_*(P_i \setminus Q) \setminus 2^{A_i}$, we have

$$M = \Gamma^*(M) \cap \Gamma_*(P_i). \tag{A.6}$$

The direction $M \subseteq \Gamma^*(M) \cap \Gamma_*(P_i)$ is easy, as $M \subseteq \Gamma^*(M)$ and $M \subseteq \Gamma_*(P_i)$ by construction. To show $\Gamma^*(M) \cap \Gamma_*(P_i) \subseteq M$, for all $J \in \Gamma^*(M) \cap \Gamma_*(P_i)$, we have $J \in \Gamma^*(M)$ and $J \in \Gamma_*(P_i)$. One can show a contradiction in assuming $J \in \Gamma_*(P_i \setminus Q)$ or $J \in 2^{A_i}$. Therefore, $J \in \Gamma_*(P_i)$, $J \notin \Gamma_*(P_i \setminus Q)$ and $J \notin 2^{A_i}$, hence $J \in M$, which completes the proof of (A.6).

Now we go back to proving the claim $\Gamma^*(M) \cap P_i \subseteq Q$ or equivalently, proving $P_i \setminus Q \subseteq P_i \setminus (\Gamma^*(M) \cap P_i)$. For all $J \in P_i \setminus Q$, we have $J \in P_i$, which means $J \in \Gamma_*(P_i)$. Also, $J \in P_i \setminus Q$ means $J \in \Gamma_*(P_i \setminus Q)$. Since M was constructed as $M = \Gamma_*(P_i) \setminus \Gamma_*(P_i \setminus Q) \setminus 2^{A_i}$, then $J \in \Gamma_*(P_i \setminus Q)$ means that $J \notin M$. However, from $J \notin M$, $J \in \Gamma_*(P_i)$ and (A.6), we conclude that $J \notin \Gamma^*(M)$, which means $J \notin \Gamma^*(M) \cap P_i$. Therefore, $J \in P_i \setminus (\Gamma^*(M) \cap P_i)$, which completes the proof of our claim. In summary, for any given $Q \subseteq P_i$ and the corresponding inequality from (A.5), we have proved that there exists an inequality in (A.4) that is no looser.

In conclusion, we have proved that the system of inequalities (A.4) and (A.5) are identical for a given $(P_i, i \in [n])$.

Appendix B

In Appendix B, we prove certain results presented in Chapter 4.

B.1 Proof of Theorem 4.2

To prove Theorem 4.2, we take the following steps:

- We first introduce several definitions which will extensively be used throughout the proof. In particular, we introduce a set function $g : 2^{[n]} \to [0, 1]$ satisfying the properties (2.35)-(2.40). Indeed, most derivations involved in the proof of Theorem 4.2 are based on this set function g and its properties.

- Then we propose a series of lemmas establishing useful intermediate results. The lemmas study different modules within a general regular acyclic chain: basic tower, horizontal chain, and regular tower. Roughly speaking, the lemmas provide upper bounds on the achievable symmetric rate in terms of the set function g and the regular acyclic chain bound Γ of the subproblems induced by the components of the module under consideration;

- Combining these lemmas for different modules of a regular acyclic chain, and utilizing the concatenation structure of the chain, we show a key lemma through mathematical induction.

- Finally, we show that this key lemma implies Theorem 4.2.

B.1.1 A Number of Useful Definitions

First, we define an *s-layer* regular acyclic chain and an *s-layer* CIC problem.

Definition B.1 (An s-layer regular acyclic chain and an *s-layer* CIC problem)**.** For any integer $s \geq 0$, an s-layer regular acyclic chain Ch and an *s-layer* CIC problem $\mathcal{G}$ are defined as follows.

1. We say a problem $\mathcal{G}$ is 0-layer if $\mathfrak{C}(\mathcal{G}) = \emptyset$, i.e., $\mathcal{G}$ is half-rate-feasible.

2. We say a chain Ch is 1-layer, if its components are all acyclic or half-rate-feasible. We say a problem $\mathcal{G}$ is 1-layer if it has at least one valid chain and all its valid chains $Ch \in \mathfrak{C}(\mathcal{G})$ are 1-layer.

3. We say a chain Ch is 2-layer, if its components are all at most 1-layer, and there is at least one component that is neither half-rate-feasible nor acyclic. We say a problem $\mathcal{G}$ is 2-layer if it has at least one 2-layer chain and all its valid chains $Ch \in \mathfrak{C}(\mathcal{G})$ are at most 2-layer.

4. We say a chain Ch is 3-layer, if its components are all at most 2-layer, and there is at least one 2-layer component. We say a problem $\mathcal{G}$ is 3-layer if it has at least one 3-layer chain and all its valid chains $Ch \in \mathfrak{C}(\mathcal{G})$ are at most 3-layer.

5. A chain or a problem that is of s-layer, for any integer $s > 3$, can be similarly defined.

Example B.1. Consider the 25-message CIC problem discussed in Example 4.3. The chain in (4.5) is a 2-layer acyclic chain. This is because that its each component V_i, $i \in [5]$ is a 1-layer subproblem which is neither half-rate-feasible nor acyclic.

To prove Theorem 4.2, consider an arbitrary CIC problem $\mathcal{G} : (i|A_i), i \in [n]$.

Recall that for the CIC problem the zero-error capacity region coincides with the capacity region (cf. Remark 2.1) Langberg and Effros [2011]. Therefore, throughout this section, without loss of generality we assume zero-error decoding as

$$H(X_i|Y, X_{A_i}) = 0, \qquad \forall i \in [n]. \tag{B.1}$$

Consider an arbitrary achievable rate tuple $\mathbf{R}$ and some $(\mathbf{t}, r)$ centralized index code satisfying

$$R_i \leq \frac{t_i}{r}, \qquad \forall i \in [n]. \tag{B.2}$$

and the decoding condition in (B.1). Without loss of generality, assume (B.2) holds with equality. That is,

$$R_i = \frac{t_i}{r}, \qquad \forall i \in [n]. \tag{B.3}$$

Define set function $g : 2^{[n]} \to [0, 1]$ as

$$g(S) \doteq \frac{1}{r} H(Y|X_{S^c}), \qquad \forall S \subseteq [n]. \tag{B.4}$$

For simplicity, we use i to denote $\{i\}$ when there is no ambiguity, and thus $g(i)$ simply means $g(\{i\})$. Also, for any sets $S_1, S_2, \cdots \subseteq [n]$, we use $g(S_1, S_2, \cdots)$ to denote $g(S_1 \cup S_2 \cup$

$\cdots$). For example, for $n = 3$, $g(\{1,2\} \cup \{3\})$, $g(\{1,2\},\{3\})$, $g(\{1,2\},3)$ and $g(1,2,3)$ mean the same thing. For another example, for $n = 4$, $S_1 = \{1,2\}$, $S_2 = \{3\}$, and $S_3 = \{4\}$, the following terms mean the same thing

$$\begin{array}{cccc} g(S_1,S_2,S_3), & g(\{1,2\},\{3\},\{4\}), & g(\{1,2\},3,4), & g(1,2,3,4), \\ g(S_1 \cup S_2 \cup S_3), & g(\{1,2\} \cup \{3\} \cup \{4\}), & g(\{1,2\} \cup 3 \cup 4), & g(\{1,2,3,4\}). \end{array}$$

In the following, we list several properties of the set function $g(S)$, which will be extensively used later in the proofs.

We have

$$g(\emptyset) = 0, \tag{B.5}$$

$$g([n]) \leq 1, \tag{B.6}$$

$$g(S) \leq g(S'), \qquad \forall S \subseteq S' \subseteq [n], \tag{B.7}$$

$$g(S \cap S') + g(S \cup S') \leq g(S) + g(S'), \quad \forall S, S' \subseteq [n]. \tag{B.8}$$

We refer to (B.7) and (B.8) as the *monotonicity* and *submodularity* of the set function $g(S)$. Notice that (B.5). (B.6), and (B.7) jointly ensure that $0 \leq g(S) \leq 1$ for any $S \subseteq [n]$.

We recall the following lemma introduced in Liu et al. [2018a], namely, the *single-receiver decoding lemma*, which characterizes the key relationship between the rate tuple $\mathbf{R}$ and the set function $g(S)$.

Lemma B.1 (Single-receiver decoding lemma, Liu et al. [2018a]). For any receiver $i \in [n]$, we have

$$R_i + g(B) = g(B \cup \{i\}), \qquad \forall B \subseteq B_i. \tag{B.9}$$

Particularly, when $B = \emptyset$, $g(B) = 0$, and thus $R_i = g(i)$.

So far we have introduced the set function $g(S)$ in (B.4) and its properties in (B.5)-(B.8) and Lemma B.1[1].

Based on the set function $g(S)$, we also define the shorthand notation

$$g(A|B) \doteq g(A,B) - g(B), \tag{B.10}$$

for any $A, B \subseteq [n]$. Based on the properties of $g(S)$, one can verify that $g(A|B)$ has the

[1] Note that this set function $g(S)$ in (B.4) indeed satisfies the properties (2.35)-(2.40) in the PM bound in Proposition 2.7, which are essentially the same as the properties in (B.5)-(B.9).

following properties,

$$g(A|B) \leq g(A'|B'), \quad A \subseteq A' \subseteq [n], B' \subseteq B \subseteq [n], \tag{B.11}$$

$$g(A|B,C) + g(B|C) = g(A,B|C), \quad A,B,C \subseteq [n]. \tag{B.12}$$

B.1.2 Lemmas and Intermediate Results

The lemma below is the key to showing Theorem 4.2, which can be proved using mathematical induction on the number of layers, as defined in Definition B.1, of subproblem $\mathcal{G}|_Q$, or simply subproblem Q. Recall that the regular acyclic chain lower bound of any subproblem Q is denoted by $\Gamma(Q)$ as defined in Theorem 4.2.

Lemma B.2. Let R be any achievable symmetric rate. For any nonempty message/receiver set $Q \subseteq [n]$, we have

$$g(Q,P) \geq g(P) + R \cdot \Gamma(Q), \qquad \forall P \subseteq B_Q. \tag{B.13}$$

To show Lemma B.2, we first show a more general version of the single-receiver decoding lemma (Lemma B.1), and refer to it as the *multi-receiver decoding lemma*.

Lemma B.3 (Multi-receiver decoding lemma)**.** Let R be any achievable symmetric rate. Consider any two message/receiver sets $K, K' \subseteq [n]$ such that $K' \subseteq B_K$. If K is acyclic or subproblem K is half-rate-feasible, then

$$g(K,K') = R \cdot \Gamma(K) + g(K'). \tag{B.14}$$

Particularly, when $K' = \emptyset$, $g(K') = 0$, and thus $g(K) = R \cdot \Gamma(K)$.

Proof. We first show (B.14) when subproblem K is half-rate-feasible, where according to Maleki et al. [2014]; Blasiak et al. [2013] we must have either that $\mathcal{G}|_K$ is a complete graph and $\Gamma(K) = 1$, or that $\Gamma(K) = \beta_{\text{MAIS}}(K) = 2$. For the former where $\Gamma(K) = 1$, we have

$$g(K,K') \geq g(k,K') = R_k + g(K') = R \cdot \Gamma(K) + g(K'),$$

where k is an arbitrary element of set K and the inequality follows from the monotonicity of $g(S)$ in (B.7), and the first equality follows from Lemma B.1 with $K' \subseteq B_K \subseteq B_k$. For the latter where $\Gamma(K) = \beta_{\text{MAIS}}(K) = 2$, there always exist two elements, $k_1, k_2 \in K$, such that $k_2 \in B_{k_1}$. Since that $K' \subseteq B_K \subseteq B_{k_1}$, we have $\{k_2\} \cup K' \subseteq B_{k_1}$. Hence,

$$g(K,K') \geq g(k_1,k_2,K') = R_{k_1} + R_{k_2} + g(K') = R \cdot \Gamma(K) + g(K'),$$

where the inequality follows from the monotonicity of $g(S)$ in (B.7), and the first equality follows from Lemma B.1 with $\{k_2\} \cup K' \subseteq B_{k_1}$ and $K \subseteq B_K \subseteq B_{k_2}$.

Now consider the case when K is acyclic. Given that K is acyclic and that $K' \subseteq B_K$, there exists an ordering of elements in K as $k_1, k_2, \cdots, k_{|K|}$ such that $\{k_j : j \in [\ell - 1]\} \cup K' \subseteq B_{k_\ell}$ for any $\ell \in [|K|]$ (cf. (2.12)). Hence, by Lemma B.1, we have

$$\begin{aligned} g(K, K') &= g(k_1, k_2, \cdots, k_{|K|-2}, k_{|K|-1}, k_{|K|}, K') \\ &= R_{k_{|K|}} + g(k_1, k_2, \cdots, k_{|K|-2}, k_{|K|-1}, K') \\ &= R_{k_{|K|}} + R_{k_{|K|-1}} + g(k_1, k_2, \cdots, k_{|K|-2}, K') \\ &= \cdots \\ &= \sum_{i \in K} R_i + g(K') = R \cdot \Gamma(K) + g(K'), \end{aligned}$$

where the last equality is due to the fact that $\Gamma(K) = |K|$ for any acyclic K. □

With the help of the multi-receiver decoding lemma (Lemma B.3), we develop a series of lemmas about the regular acyclic chain Ch as defined in Definition 4.4, which will prove useful in showing Lemma B.2. We start with the following lemma for the basic tower.

Lemma B.4. Consider any basic tower $\mathcal{B}_j \subseteq [n]$. Let R be any achievable symmetric rate. If every component of $\mathcal{B}_j$ is acyclic or half-rate-feasible, we have

$$g(\mathcal{B}_j, P) = g(I(j), I(j+1), P) + R \sum_{\ell \in [h_j]} \Gamma(K_\ell(j)), \qquad \forall P \subseteq B_{\mathcal{B}_j}. \tag{B.15}$$

Proof. Consider any $P \subseteq B_{\mathcal{B}_j}$. For any $\ell \in [h_j]$, we have $I(j) \cup I(j+1) \cup K_1(j) \cup \cdots \cup K_{\ell-1}(j) \subseteq B_{K_\ell(j)}$ by Definition 4.1. Recall that for any $K \subseteq K' \subseteq [n]$, $B_{K'} \subseteq B_K \subseteq [n]$ (cf. (2.2)). Thus we have $P \subseteq B_{\mathcal{B}_j} \subseteq B_{K_\ell(j)}$ since $K_\ell(j) \subseteq \mathcal{B}_j$. Combining these two results yields

$$I(j) \cup I(j+1) \cup P \cup K_1(j) \cup \cdots \cup K_{\ell-1}(j) \subseteq B_{K_\ell(j)}, \qquad \forall \ell \in [h_j]. \tag{B.16}$$

Given (B.16) and that every component of $\mathcal{B}_j$ is acyclic or half-rate-feasible, by Lemma B.3,

$$\begin{aligned} g(\mathcal{B}_j, P) &= g(I(j), I(j+1), P, K_1(j), K_2(j), \ldots, K_{h_j}(j)) \\ &= R\Gamma(K_{h_j}(j)) + g(I(j), I(j+1), P, K_1(j), K_2(j), \ldots, K_{h_j-1}(j)) \\ &= R\Gamma(K_{h_j}(j)) + R\Gamma(K_{h_j-1}(j)) + g(I(j), I(j+1), P, K_1(j), K_2(j), \ldots, K_{h_j-2}(j)) \\ &\cdots \\ &= R \sum_{\ell \in [h_j]} \Gamma(K_\ell(j)) + g(I(j), I(j+1), P), \end{aligned}$$

which completes the proof. □

The following lemma considers a group of message sets that are concatenatively located on the horizontal chain of a regular acyclic chain.

Lemma B.5. Consider a group of message sets $I(a), I(a+1), \dots, I(b), I(b+1) \subseteq [n]$ for some positive integers $a < b$, concatenatively located on the horizontal chain of some regular acyclic chain $Ch \in \mathfrak{C}(\mathcal{G})$, which are all acyclic or half-rate-feasible. Let R be any achievable symmetric rate. For any set P such that $P \subseteq B_{I(a)\cup I(a+1)\cup\cdots\cup I(b)\cup I(b+1)}$, we have

$$\begin{aligned}&\sum_{j\in[a:b]} g(I(j), I(j+1), P)\\&\geq (b-a)g(P) + R\sum_{j\in[a+1:b]}\Gamma(I(j)) + g(I(a), I(b+1), P).\end{aligned}\tag{B.17}$$

Proof. Consider any $P \subseteq B_{I(a)\cup I(a+1)\cup\cdots\cup I(b)\cup I(b+1)}$. Then for any $j \in [a:b+1]$, we have $I(j) \subseteq I(a+1) \cup \cdots \cup I(b) \cup I(b+1)$, and thus $P \subseteq B_{I(j)}$ by (2.2). Hence, we have

$$\begin{aligned}&\sum_{j\in[a:b]} g(I(j), I(j+1), P)\\
&= g(I(a), I(a+1), P) + g(I(a+1), I(a+2), P)\\
&\quad + g(I(a+2), I(a+3), P) + \cdots + g(I(b), I(b+1), P)\\
&\geq g(P) + R\Gamma(I(a+1)) + g(I(a), I(a+1), I(a+2), P)\\
&\quad + g(I(a+2), I(a+3), P) + \cdots + g(I(b), I(b+1), P) &\text{(B.18)}\\
&\geq 2g(P) + R\Gamma(I(a+1)) + R\Gamma(I(a+2)) + g(I(a), I(a+1), I(a+2), I(a+3), P)\\
&\quad + \cdots + g(I(b), I(b+1), P) &\text{(B.19)}\\
&\cdots &\text{(B.20)}\\
&\geq (b-a)g(P) + R\Gamma(I(a+1)) + R\Gamma(I(a+2)) + \cdots + R\Gamma(I(b))\\
&\quad + g(I(a), I(a+1), I(a+2), \cdots, I(b), I(b+1), P) &\text{(B.21)}\\
&\geq (b-a)g(P) + R\sum_{j\in[a+1:b]}\Gamma(I(j)) + g(I(a), I(b+1), P), &\text{(B.22)}\end{aligned}$$

where (B.18) follows from the following derivation:

$$\begin{aligned}&g(I(a), I(a+1), P) + g(I(a+1), I(a+2), P)\\
&\geq g(I(a+1), P) + g(I(a), I(a+1), I(a+2), P)\\
&= g(P) + R\Gamma(I(a+1)) + g(I(a), I(a+1), I(a+2), P),\end{aligned}$$

where the inequality is due to the submodularity of $g(S)$ in (B.8) and the equality is due to

Lemma B.3 given that for any $j \in [a : b+1]$, $P \subseteq B_{I(j)}$ and $I(j)$ is acyclic or half-rate-feasible. Inequalities (B.19)-(B.21) follow similarly. The last inequality, (B.22), is due to the monotonicity of $g(S)$ in (B.7). □

The following lemma holds for any regular tower, whose proof involves Lemmas B.3-B.5 presented above.

Lemma B.6. Consider any regular tower $\mathcal{X}_j \subseteq [n]$ with central edge j whose components are all acyclic or half-rate-feasible. Let R be any achievable symmetric rate. For any $P \subseteq B_{\mathcal{X}_j}$, we have

$$\begin{aligned} |G_j| \cdot g(\mathcal{X}_j, P) \geq g(I(s_{\theta_j,j}), I(t_{\theta_j,j}), P) &+ R \sum_{j' \in G_j} \sum_{\ell \in [h_{j'}]} \Gamma(K_\ell(j')) \\ &+ R \sum_{j' \in [s_{\theta_j,j}+1 : t_{\theta_j,j}-1]} \Gamma(I(j')) + (|G_j|-1) g(P). \end{aligned} \tag{B.23}$$

Proof of Lemma B.6 for a specific example: To help with understanding, before presenting the proof of Lemma B.6, we illustrate our proof techniques by applying it to a specific example. Consider the following regular tower $\mathcal{X}_j$ with the central edge $j = 3$,

$$\mathcal{X}_j : \quad I(1) \xleftrightarrow{K_1(1)}_{\mathrm{b}} I(2) \xleftrightarrow{K_1(2)}_{\mathrm{b}} I(3) \overset{\substack{K_3(3)\\ K_2(3)}}{\xleftrightarrow{K_1(3)}}_{\mathrm{c}} I(4) \xleftrightarrow{K_1(4)}_{\mathrm{b}} I(5) \xleftrightarrow{K_1(5)}_{\mathrm{b}} I(6), \tag{B.24}$$

where

- for brevity, we denote the height h_j of the central edge simply as h, and for any $\ell \in [h]$, the message set $K_\ell(j)$ as K_ℓ, $s_{\ell,j}$ as s_ℓ, $t_{\ell,j}$ as t_ℓ, and $s_{\iota_j,j}$, $s_{\theta_j,j}$, $t_{\iota_j,j}$, $t_{\theta_j,j}$ as s_ι, s_θ, t_ι, t_θ, respectively;
- again for brevity, for any edge $j' \in [5] \setminus \{3\} = \{1,2,4,5\}$ other than the central edge, we denote $K_1(j')$ as $K(j')$;
- $s_1 = 2$, $t_1 = 5$, $s_2 = s_\iota = 3$, $t_2 = t_\iota = 4$, $s_3 = s_\theta = 1$, and $t_3 = t_\theta = 6$ (one can verifiy that Conditions 2b and 2c in Definition 4.3 are satisfied);
- every component of $\mathcal{X}_j$ is acyclic or half-rate-feasible.

Given the above setup, the regular tower $\mathcal{X}_j$ in (B.24) can be equivalently represented as

$$\mathcal{X}_j : \quad I(s_3) \xleftrightarrow{K(1)}_{\mathrm{b}} I(s_1) \xleftrightarrow{K(2)}_{\mathrm{b}} I(s_2) \overset{\substack{K_3\\ K_2}}{\xleftrightarrow{K_1}}_{\mathrm{c}} I(t_2) \xleftrightarrow{K(4)}_{\mathrm{b}} I(t_1) \xleftrightarrow{K(5)}_{\mathrm{b}} I(t_3). \tag{B.25}$$

See Figure B.1 as a visualization of this regular tower $\mathcal{X}_j$.

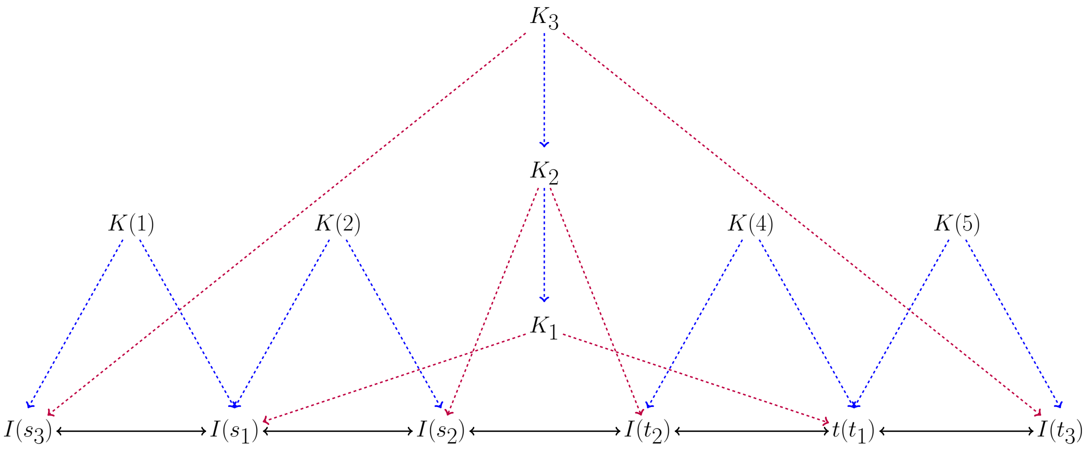


Figure B.1: A visualization for the regular tower $\mathcal{X}_j$ in (B.25). To help with understanding, we draw blue and purple dashed arrows. If there is a directed path formed by dashed arrows of the same color from message set Q to P, then $P \subseteq B_Q$.

For this specific $\mathcal{X}_j$, (B.23) in Lemma B.6 reduces to

$$\begin{aligned}5g(\mathcal{X}_j, P) \geq g(I(s_3), I(t_3), P) + R \sum_{j' \in [5] \setminus \{3\}} \Gamma(K(j')) + R \sum_{\ell \in [3]} \Gamma(K_\ell) \\ + R\Gamma(I(s_1)) + R\Gamma(I(s_2)) + R\Gamma(I(t_2)) + R\Gamma(I(t_1)) + 4g(P). \qquad \text{(B.26)}\end{aligned}$$

We show (B.26) in the following.

Define a one-to-one mapping function $\mu : [3] \to [3]$ as

$$1 = \mu(3), \qquad 2 = \mu(1), \qquad 3 = \mu(2). \qquad \text{(B.27)}$$

Let $\eta = \mu^{-1}$ denote the inverse function of μ. Then we have

$$\eta(1) = 3, \qquad \eta(2) = 1, \qquad \eta(3) = 2. \qquad \text{(B.28)}$$

With such mapping function μ and η, we have

$$\begin{aligned} s_{\eta(1)} &= s_3, \quad t_{\eta(1)} = t_3, \\ s_{\eta(2)} &= s_1, \quad t_{\eta(2)} = t_1, \\ s_{\eta(3)} &= s_2, \quad t_{\eta(3)} = t_2, \end{aligned} \qquad \text{(B.29)}$$

Hence, for any $i_1 < i_2 \in [3]$, $s_{\eta(i_1)} < s_{\eta(i_2)}$ and thus $I(s_{\eta(i_1)})$ is placed on the left to $I(s_{\eta(i_2)})$ in the horizontal chain of $\mathcal{X}_j$ as shown in (B.25). Similarly, for any $i_1 < i_2 \in [3]$, $t_{\eta(i_1)} > t_{\eta(i_2)}$

and thus $I(t_{\eta(i_1)})$ is placed on the right to $I(t_{\eta(i_2)})$ in the horizontal chain of $\mathcal{X}_j$ in (B.25).

For any $i \in [3]$, according to Condition 2a in Definition 4.3, $P \subseteq B_{\mathcal{X}_j}$, and (2.2), we have

$$I(s_{\eta(i)}) \cup I(t_{\eta(i)}) \cup P \cup K_1 \cup \cdots \cup K_{\eta(i)-1} \subseteq B_{K_{\eta(i)}}. \tag{B.30}$$

Hence, we have

$$\begin{aligned} R\Gamma(K_{\eta(i)}) &\overset{(a)}{=} g(K_{\eta(i)}, I(s_{\eta(i)}), I(t_{\eta(i)}), P, K_1, \ldots, K_{\eta(i)-1}) \\ &\quad - g(I(s_{\eta(i)}), I(t_{\eta(i)}), P, K_1, \ldots, K_{\eta(i)-1}) \\ &\overset{(b)}{=} g(K_{\eta(i)} | I(s_{\eta(i)}), I(t_{\eta(i)}), P, K_1, \ldots, K_{\eta(i)-1}), \end{aligned} \tag{B.31}$$

where

- (a) follows from the multi-receiver decoding lemma (Lemma B.3) given (B.30) and the fact that every component of $\mathcal{X}_j$ is acyclic or half-rate-feasible;

- (b) follows from (B.10).

We have

$$\begin{aligned} &\sum_{\ell \in [3]} g(I(s_\ell), I(t_\ell), K_1, K_2, K_3, P) \\ &\overset{(c)}{=} \sum_{i \in \{1,2,3\}} g(I(s_{\eta(i)}), I(t_{\eta(i)}), K_1, K_2, K_3, P) \\ &\overset{(d)}{=} \sum_{i \in \{1,2,3\}} \big(\rho_i + \lambda_i + R\Gamma(K_{\eta(i)})\big), \end{aligned} \tag{B.32}$$

where

- for any $i \in [3]$,

$$\rho_i \doteq g(I(s_{\eta(i)}), I(t_{\eta(i)}), P) + \sum_{\ell \in \{\eta(i'): i' \in [i-1]\}} g(K_\ell | I(s_{\eta(i)}), I(t_{\eta(i)}), P, K_1, \ldots, K_{\ell-1}), \tag{B.33}$$

$$\lambda_i \doteq \sum_{\ell \in [3] \setminus \{\eta(i'): i' \in [i]\}} g(K_\ell | I(s_{\eta(i)}), I(t_{\eta(i)}), P, K_1, \ldots, K_{\ell-1}); \tag{B.34}$$

- (c) follows from the fact that η is a one-to-one mapping function and thus $\{\eta(i) : i \in [3]\} = [3]$;

- (d) follows from the fact that for any $i \in [3]$,

$$\begin{aligned}
&\rho_i + \lambda_i + R\Gamma(K_{\eta(i)}) \\
&\overset{(e)}{=} g(I(s_{\eta(i)}), I(t_{\eta(i)}), P) + \sum_{\ell \in [3] \setminus \{\eta(i)\}} g(K_\ell | I(v_i), I(t_{\eta(i)}), P, K_1, \ldots, K_{\ell-1}) \\
&\quad + g(K_{\eta(i)} | I(s_{\eta(i)}), I(t_{\eta(i)}), P, K_1, \ldots, K_{\eta(i)-1}) \\
&= g(I(s_{\eta(i)}), I(t_{\eta(i)}), P) + \sum_{\ell \in [3]} g(K_\ell | I(v_i), I(t_{\eta(i)}), P, K_1, \ldots, K_{\ell-1}) \\
&\overset{(f)}{=} g(I(s_{\eta(i)}), I(t_{\eta(i)}), P, K_1, K_2, K_3),
\end{aligned}$$

where (e) follows from (B.33) and (B.34), as well as (B.31), and (f) follows from repeated application of (B.12).

According to the definitions of ρ_i, λ_i in (B.33) and (B.34), as well as (B.29), we have

$$\begin{aligned}
&\rho_2 + \lambda_1 + g(I(s_{\eta(1)}), I(s_{\eta(2)}), P) + g(I(t_{\eta(2)}), I(t_{\eta(1)}), P) \\
&= g(I(s_{\eta(2)}), I(s_{\eta(2)}), P) + g(K_{\eta(1)} | I(s_{\eta(2)}), I(t_{\eta(2)}), P, K_1, \ldots, K_{\eta(1)-1}) \\
&\quad + \sum_{\ell \in \{\eta(2), \eta(3)\}} g(K_\ell | I(s_{\eta(1)}), I(t_{\eta(1)}), P, K_1, \ldots, K_{\ell-1}) \\
&\quad + g(I(s_{\eta(1)}), I(s_{\eta(2)}), P) + g(I(t_{\eta(2)}), I(t_{\eta(1)}), P) \\
&= g(I(s_1), I(t_1), P) + g(K_3 | I(s_1), I(t_1), P, K_1, K_2) \\
&\quad + g(K_1 | I(s_3), I(t_3), P) + g(K_2 | I(s_3), I(t_3), P, K_1) \\
&\quad + g(I(s_3), I(s_1), P) + g(I(t_1), I(t_3), P) \\
&\geq \sum_{\ell \in [3]} g(K_\ell | I(s_3), I(s_1), I(t_1), I(t_3), P, K_1, \ldots, K_{\ell-1}) \\
&\quad + g(I(s_1), I(t_1), P) + g(I(s_3), I(s_1), P) + g(I(t_1), I(t_3), P) && \text{(B.35)} \\
&\geq \sum_{\ell \in [3]} g(K_\ell | I(s_3), I(s_1), I(t_1), I(t_3), P, K_1, \ldots, K_{\ell-1}) \\
&\quad + g(I(s_3), I(s_1), I(t_1), I(t_3), P) + g(I(s_1), P) + g(I(t_1), P) && \text{(B.36)} \\
&= g(I(s_3), I(s_1), I(t_1), I(t_3), P, K_1, K_2, K_3) + g(I(s_1), P) + g(I(t_1), P) && \text{(B.37)} \\
&= g(I(s_3), I(s_1), I(t_1), I(t_3), P, K_1, K_2, K_3) + 2g(P) + R\Gamma(I(s_1)) + R\Gamma(I(t_1)), && \text{(B.38)}
\end{aligned}$$

where

- (B.35) follows from the property of $g(A|B)$ in (B.11);
- (B.36) follows from applying the submodularity of $g(S)$ in (B.8) twice;
- (B.37) follows from the property of $g(A|B)$ in (B.12);

- (B.38) follows from the multi-receiver decoding lemma (Lemma B.3) with the fact that $P \subseteq B_{\mathcal{X}_j} \subseteq B_{I(s_1)}$ and $P \subseteq B_{\mathcal{X}_j} \subseteq B_{I(t_1)}$ (given by (2.2)) and that every component of $\mathcal{X}_j$ is acyclic or half-rate-feasible.

Rearranging (B.38), we have

$$\begin{aligned}\rho_2 + \lambda_1 \geq\; & g(I(s_3), I(s_1), I(t_1), I(t_3), P, K_1, K_2, K_3) + 2g(P) + R\Gamma(I(s_1)) + R\Gamma(I(t_1)) \\ & - g(I(s_{\eta(1)}), I(s_{\eta(2)}), P) - g(I(t_{\eta(2)}), I(t_{\eta(1)}), P). \end{aligned} \tag{B.39}$$

Similarly, one can verify that

$$\begin{aligned}\rho_3 + \lambda_2 \geq\; & g(I(s_1), I(s_2), I(t_2), I(t_1), P, K_1, K_2, K_3) + 2g(P) + R\Gamma(I(s_2)) + R\Gamma(I(t_2)) \\ & - g(I(s_{\eta(2)}), I(s_{\eta(3)}), P) - g(I(t_{\eta(3)}), I(t_{\eta(2)}), P). \end{aligned} \tag{B.40}$$

Combining (B.32), (B.39), and (B.40), we have

$$\begin{aligned}
& \sum_{\ell \in [3]} g(I(s_\ell), I(t_\ell), K_1, K_2, K_3, P) \\
& = \sum_{\ell \in [3]} R\Gamma(K_\ell) + \rho_1 + \lambda_3 + (\rho_2 + \lambda_1) + (\rho_3 + \lambda_2) \\
& \geq \sum_{\ell \in [3]} R\Gamma(K_\ell) + \rho_1 + \lambda_3 + R\Gamma(I(s_1)) + R\Gamma(I(s_2)) + R\Gamma(I(t_2)) + R\Gamma(I(t_1)) + 4g(P) \\
& \quad + g(I(s_3), I(s_1), I(t_1), I(t_3), P, K_1, K_2, K_3) + g(I(s_1), I(s_2), I(t_2), I(t_1), P, K_1, K_2, K_3) \\
& \quad - g(I(s_3), I(s_1), P) - g(I(t_1), I(t_3), P) - g(I(s_1), I(s_2), P) - g(I(t_2), I(t_1), P) \\
& \overset{(g)}{\geq} \sum_{\ell \in [3]} R\Gamma(K_\ell) + \rho_1 + \lambda_3 + R\Gamma(I(s_1)) + R\Gamma(I(s_2)) + R\Gamma(I(t_2)) + R\Gamma(I(t_1)) + 4g(P) \\
& \quad + g(I(s_1), I(t_1), P, K_1, K_2, K_3) + g(I(s_2), I(t_2), P, K_1, K_2, K_3) \\
& \quad - g(I(s_3), I(s_1), P) - g(I(t_1), I(t_3), P) - g(I(s_1), I(s_2), P) - g(I(t_2), I(t_1), P) \\
& \overset{(h)}{=} \sum_{\ell \in [3]} R\Gamma(K_\ell) + g(I(s_3), I(t_3), P) + R\Gamma(I(s_1)) + R\Gamma(I(s_2)) + R\Gamma(I(t_2)) + R\Gamma(I(t_1)) \\
& \quad + 4g(P) + g(I(s_1), I(t_1), P, K_1, K_2, K_3) + g(I(s_2), I(t_2), P, K_1, K_2, K_3) \\
& \quad - g(I(s_3), I(s_1), P) - g(I(t_1), I(t_3), P) - g(I(s_1), I(s_2), P) - g(I(t_2), I(t_1), P),
\end{aligned} \tag{B.41}$$

where (g) follows from the monotonicity of $g(S)$ in (B.7), and (h) follows from the fact that $\rho_1 = g(I(s_3), I(t_3), P)$ and $\lambda_3 = 0$.

Subtracting $g(I(s_1), I(t_1), P, K_1, K_2, K_3) + g(I(s_2), I(t_2), P, K_1, K_2, K_3)$ from both sides of (B.41), we obtain

$$\begin{aligned}
&g(I(s_3), I(t_3), P, K_1, K_2, K_3) \\
&\geq \sum_{\ell\in[3]} R\Gamma(K_\ell) + g(I(s_3), I(t_3), P) + R\Gamma(I(s_1)) + R\Gamma(I(s_2)) + R\Gamma(I(t_2)) + R\Gamma(I(t_1)) \\
&\quad + 4g(P) - g(I(s_3), I(s_1), P) - g(I(t_1), I(t_3), P) - g(I(s_1), I(s_2), P) - g(I(t_2), I(t_1), P),
\end{aligned}$$

which, together with the fact that $g(\mathcal{X}_j, P) \geq g(I(s_3), I(t_3), P, K_1, K_2, K_3)$ because of the monotonicity of $g(S)$ in (B.7), lead to

$$\begin{aligned}
g(\mathcal{X}_j, P) \geq & \sum_{\ell\in[3]} R\Gamma(K_\ell) + g(I(s_3), I(t_3), P) + R\Gamma(I(s_1)) + R\Gamma(I(s_2)) + R\Gamma(I(t_2)) \\
& + R\Gamma(I(t_1)) + 4g(P) - g(I(s_3), I(s_1), P) - g(I(t_1), I(t_3), P) \\
& - g(I(s_1), I(s_2), P) - g(I(t_2), I(t_1), P). \qquad \text{(B.42)}
\end{aligned}$$

Consider the leftmost basic tower $\mathcal{B}_1$ within the regular tower $\mathcal{X}_j$, which is constitued by the message sets $I(s_3), I(s_1), K(1)$. We have

$$g(\mathcal{X}_j, P) \geq g(\mathcal{B}_1, P) = R\Gamma(K(1)) + g(I(s_3), I(s_1), P), \qquad \text{(B.43)}$$

where the inequality follows from the monotonicity of $g(S)$ in (B.7), and the equality follows from Lemma B.4 with the fact that every component of $\mathcal{X}_j$ is acyclic or half-rate-feasible.

Similarly, considereing the basic towers $\mathcal{B}_2, \mathcal{B}_4, \mathcal{B}_5$ (note that $j = 3$ is the central edge so there is no basic tower $\mathcal{B}_3$), we have

$$\begin{aligned}
g(\mathcal{X}_j, P) &\geq R\Gamma(K(2)) + g(I(s_1), I(s_2), P), && \text{(B.44)} \\
g(\mathcal{X}_j, P) &\geq R\Gamma(K(4)) + g(I(t_2), I(t_1), P), && \text{(B.45)} \\
g(\mathcal{X}_j, P) &\geq R\Gamma(K(5)) + g(I(t_1), I(t_3), P). && \text{(B.46)}
\end{aligned}$$

Summing up (B.42)-(B.46) yield

$$\begin{aligned}
5g(\mathcal{X}_j, P) \geq & \sum_{\ell\in[3]} R\Gamma(K_\ell) + g(I(s_3), I(t_3), P) + R\Gamma(I(s_1)) + R\Gamma(I(s_2)) + R\Gamma(I(t_2)) \\
& + R\Gamma(I(t_1)) + 4g(P) + \sum_{j'\in[5]\setminus\{3\}} R\Gamma(K(j')).
\end{aligned}$$

which completes the proof of (B.26).

So far we have shown that Lemma B.6 holds for the specific $\mathcal{X}_j$ in (B.24).

Proof of Lemma B.6: In the following we prove Lemma B.6 in general.

Proof. For brevity, for the regular tower $\mathcal{X}_j$ with central edge j, we denote the height h_j simply as h, and for any $\ell \in [h]$, the message set $K_\ell(j)$ as K_ℓ, and $s_{\ell,j}$, $t_{\ell,j}$ as s_ℓ, t_ℓ, respectively. Also denote $s_{\iota_j,j}$, $s_{\theta_j,j}$, $t_{\iota_j,j}$, $t_{\theta_j,j}$ as s_ι, s_θ, t_ι, t_θ, respectively. By Definition 4.3, we know that for any $\ell_1, \ell_2 \in [h]$, if $s_{\ell_1} \leq s_{\ell_2}$, then $t_{\ell_1} \geq t_{\ell_2}$.

According to Condition 2b in Defintion 4.3, there exists a bijective function $\mu : [h] \to [h]$ such that for any $\ell_1 \neq \ell_2 \in [h]$, if $\mu(\ell_1) < \mu(\ell_2)$, then $s_{\ell_1} \leq s_{\ell_2}$ and $t_{\ell_1} \geq t_{\ell_2}$.

For example, for the regular tower in (B.24), the funtion μ in (B.27) satisfies the above conditon. For another example, consider the regular tower $\mathcal{X}_1$ in Example 4.4. We have the following bijective function $\mu : [2] \to [2]$,

$$1 = \mu(2), \qquad 2 = \mu(1),$$

which satisfies the condition mentioned above.

Let $\eta = \mu^{-1}$ denote the inverse function of μ. It can be verified that

$$\bigcup_{i \in [h-1]} [s_{\eta(i)} : s_{\eta(i+1)} - 1] = [s_\theta : s_\iota - 1],$$
$$\bigcup_{i \in [h-1]} [t_{\eta(i+1)} : t_{\eta(i)} - 1] = [t_\iota : t_\theta - 1].$$

Given the bijective functions μ and $\eta = \mu^{-1}$, we have

$$\begin{aligned} &\sum_{\ell \in [h]} g(I(s_\ell), I(t_\ell), K_1, K_2, \ldots, K_h, P) \\ &= \sum_{i \in [h]} g(I(s_{\eta(i)}), I(t_{\eta(i)}), K_1, K_2, \ldots, K_h, P) \\ &\overset{(a)}{=} \sum_{i \in [h]} (\rho_i + \lambda_i + R\Gamma(K_\ell)), \end{aligned} \tag{B.47}$$

where for any $i \in [h]$,

$$\begin{aligned} \rho_i &\doteq g(I(s_{\eta(i)}), I(t_{\eta(i)}), P) \\ &\quad + \sum_{\ell \in \{\eta(i') : i' \in [i-1]\}} g(K_\ell | I(s_{\eta(i)}), I(t_{\eta(i)}), P, K_1, \ldots, K_{\ell-1}), \\ \lambda_i &\doteq \sum_{\ell \in [h] \setminus \{\eta(i') : i' \in [i]\}} g(K_\ell | I(s_{\eta(i)}), I(t_{\eta(i)}), P, K_1, \ldots, K_{\ell-1}), \end{aligned}$$

and (a) can be verified as for any $i \in [h]$ we have

$$\begin{aligned}
&\rho_i + \lambda_i \\
&= g(I(s_{\eta(i)}), I(t_{\eta(i)}), P) + \sum_{\ell \in [h]} g(K_\ell | I(s_{\eta(i)}), I(t_{\eta(i)}), P, K_1, \dots, K_{\ell-1}) \\
&\quad - g(K_{\eta(i)} | I(s_{\eta(i)}), I(t_{\eta(i)}), P, K_1, \dots, K_{\eta(i)-1}) \\
&= g(I(s_{\eta(i)}), I(t_{\eta(i)}), P, K_1, \dots, K_h) - g(K_{\eta(i)} | I(s_{\eta(i)}), I(t_{\eta(i)}), P, K_1, \dots, K_{\eta(i)-1}) \\
&= g(I(s_{\eta(i)}), I(t_{\eta(i)}), P, K_1, \dots, K_h) - R\Gamma(K_{\eta(i)}),
\end{aligned}$$

where the second equality follows from repeated application of (B.12), and the third equality follows from (B.10) and Lemma B.3.

Now consider any $i \in [2:h]$, we have

$$\begin{aligned}
&\rho_i + \lambda_{i-1} + g(I(s_{\eta(i-1)}), I(s_{\eta(i)}), P) + g(I(t_{\eta(i)}), I(t_{\eta(i-1)}), P) \\
&= g(I(s_{\eta(i-1)}), I(s_{\eta(i)}), P) + g(I(s_{\eta(i)}), I(t_{\eta(i)}), P) + g(I(t_{\eta(i)}), I(t_{\eta(i-1)}), P) \\
&\quad + \sum_{\ell \in \{\eta(i') : i' \in [i-1]\}} g(K_\ell | I(s_{\eta(i)}), I(t_{\eta(i)}), P, K_1, \dots, K_{\ell-1}) \\
&\quad + \sum_{\ell \in [h] \setminus \{\eta(i') : i' \in [i-1]\}} g(K_\ell | I(s_{\eta(i-1)}), I(t_{\eta(i-1)}), P, K_1, \dots, K_{\ell-1}) \\
&\geq g(I(s_{\eta(i-1)}), I(s_{\eta(i)}), P) + g(I(s_{\eta(i)}), I(t_{\eta(i)}), P) + g(I(t_{\eta(i)}), I(t_{\eta(i-1)}), P) \\
&\quad + \sum_{\ell \in \{\eta(i') : i' \in [i-1]\}} g(K_\ell | I(s_{\eta(i-1)}), I(s_{\eta(i)}), I(t_{\eta(i)}), I(t_{\eta(i-1)}), P, K_1, \dots, K_{\ell-1}) \\
&\quad + \sum_{\ell \in [h] \setminus \{\eta(i') : i' \in [i-1]\}} g(K_\ell | I(s_{\eta(i-1)}), I(s_{\eta(i)}), I(t_{\eta(i)}), I(t_{\eta(i-1)}), P, K_1, \dots, K_{\ell-1}) \\
&\geq 2g(P) + R\Gamma(I(s_{\eta(i)})) + R\Gamma(I(t_{\eta(i)})) + g(I(s_{\eta(i-1)}), I(s_{\eta(i)}), I(t_{\eta(i)}), I(t_{\eta(i-1)}), P) \\
&\quad + \sum_{\ell \in [h]} g(K_\ell | I(s_{\eta(i-1)}), I(s_{\eta(i)}), I(t_{\eta(i)}), I(t_{\eta(i-1)}), P, K_1, \dots, K_{\ell-1}) \\
&= 2g(P) + R\Gamma(I(s_{\eta(i)})) + R\Gamma(I(t_{\eta(i)})) \\
&\quad + g(I(s_{\eta(i-1)}), I(s_{\eta(i)}), I(t_{\eta(i)}), I(t_{\eta(i-1)}), P, K_1, \dots, K_h), \qquad \text{(B.48)}
\end{aligned}$$

where the first inequality follows from (B.11), the second inequality follows from the submodularity of $g(S)$ as well as Lemma B.3, and the last equality follows from repeated application of (B.12).

By (B.47) and (B.48), we have

$$
\begin{aligned}
&\sum_{\ell\in[h]} g(I(s_\ell), I(t_\ell), P, K_1, K_2, \dots, K_h) \\
&\quad + \sum_{i\in[2:h]} \big(g(I(s_{\eta(i-1)}), I(s_{\eta(i)}), P) + g(I(t_{\eta(i)}), I(t_{\eta(i-1)}), P)\big) \\
&= R\sum_{\ell\in[h]} \Gamma(K_\ell) + \sum_{i\in[h]} (\rho_i + \lambda_i) \\
&\quad + \sum_{i\in[2:h]} \big(g(I(s_{\eta(i-1)}), I(s_{\eta(i)}), P) + g(I(t_{\eta(i)}), I(t_{\eta(i-1)}), P)\big) \\
&\geq R\sum_{\ell\in[h]} \Gamma(K_\ell) + \rho_1 + \lambda_h + \sum_{i\in[2:h]} \big(2g(P) + R\Gamma(I(s_{\eta(i)})) + R\Gamma(I(t_{\eta(i)}))\big) \\
&\quad + \sum_{i\in[2:h]} g(I(s_{\eta(i-1)}), I(s_{\eta(i)}), I(t_{\eta(i)}), I(t_{\eta(i-1)}), P, K_1, \dots, K_h) \\
&= R\sum_{\ell\in[h]} \Gamma(K_\ell) + g(I(s_{\eta(1)}), I(t_{\eta(1)}), P) + 0 \\
&\quad + \sum_{i\in[2:h]} \big(2g(P) + R\Gamma(I(s_{\eta(i)})) + R\Gamma(I(t_{\eta(i)}))\big) \\
&\quad + \sum_{i\in[2:h]} g(I(s_{\eta(i-1)}), I(s_{\eta(i)}), I(t_{\eta(i)}), I(t_{\eta(i-1)}), P, K_1, \dots, K_h) \\
&\geq R\sum_{\ell\in[h]} \Gamma(K_\ell) + g(I(s_{\eta(1)}), I(t_{\eta(1)}), P) + \sum_{i\in[2:h]} \big(2g(P) + R\Gamma(I(s_{\eta(i)})) + R\Gamma(I(t_{\eta(i)}))\big) \\
&\quad + \sum_{\ell\in[h-1]} g(I(s_\ell), I(t_\ell), P, K_1, \dots, K_h), \qquad \text{(B.49)}
\end{aligned}
$$

where (B.49) can be verified as

$$
\begin{aligned}
&\sum_{i\in[2:h]} g(I(s_{\eta(i-1)}), I(s_{\eta(i)}), I(t_{\eta(i)}), I(t_{\eta(i-1)}), P, K_1, \dots, K_h) \\
&\geq \sum_{i\in[2:\mu(h)]} g(I(s_{\eta(i-1)}), I(t_{\eta(i-1)}), P, K_1, \dots, K_h) + \sum_{i\in[\mu(h)+1:h]} g(I(s_{\eta(i)}), I(t_{\eta(i)}), P, K_1, \dots, K_h) \\
&= \sum_{i\in[1:\mu(h)-1]} g(I(s_{\eta(i)}), I(t_{\eta(i)}), P, K_1, \dots, K_h) + \sum_{i\in[\mu(h)+1:h]} g(I(s_{\eta(i)}), I(t_{\eta(i)}), P, K_1, \dots, K_h) \\
&= \sum_{i\in[h]\setminus\{\mu(h)\}} g(I(s_{\eta(i)}), I(t_{\eta(i)}), P, K_1, \dots, K_h) \\
&= \sum_{\ell\in[h-1]} g(I(s_\ell), I(t_\ell), P, K_1, \dots, K_h),
\end{aligned}
$$

where the inequality follows from (B.11).

Subtracting $\sum_{\ell\in[h-1]} g(I(s_\ell), I(t_\ell), P, K_1, \dots, K_h)$ on both sides of (B.49), we have

$$\begin{aligned}
&g(I(s_h), I(t_h), P, K_1, \dots, K_h) + \sum_{i\in[2:h]} \big(g(I(s_{\eta(i-1)}), I(s_{\eta(i)}), P) + g(I(t_{\eta(i)}), I(t_{\eta(i-1)}), P)\big) \\
&\geq R \sum_{\ell\in[h]} \Gamma(K_\ell) + g(I(s_{\eta(1)}), I(t_{\eta(1)}), P) + \sum_{i\in[2:h]} \big(2g(P) + R\Gamma(I(s_{\eta(i)})) + R\Gamma(I(t_{\eta(i)}))\big),
\end{aligned}$$

which, together with that $g(\mathcal{X}_j, P) \geq g(I(s_h), I(t_h), P, K_1, K_2, \dots, K_h)$ by the monotonicity of $g(S)$ in (B.7), lead to

$$\begin{aligned}
g(\mathcal{X}_j, P) \geq R &\sum_{\ell\in[h]} \Gamma(K_\ell) + g(I(s_{\eta(1)}), I(t_{\eta(1)}), P) \\
&+ \sum_{i\in[2:h]} \big(2g(P) + R\Gamma(I(s_{\eta(i)})) + R\Gamma(I(t_{\eta(i)}))\big) \\
&- \sum_{i\in[2:h]} \big(g(I(s_{\eta(i-1)}), I(s_{\eta(i)}), P) + g(I(t_{\eta(i)}), I(t_{\eta(i-1)}), P)\big). \qquad \text{(B.50)}
\end{aligned}$$

Consider any $i \in [2:h]$. According to Definition 4.3, for any $j' \in [s_{\eta(i-1)} : s_{\eta(i)} - 1] \cup [t_{\eta(i)} : t_{\eta(i-1)} - 1]$, message sets $I(j'), I(j'+1), K_1(j'), \dots, K_{h_{j'}}(j')$ form a basic tower, denoted as $\mathcal{B}_{j'}$. And thus we have

$$\begin{aligned}
\sum_{j'\in[s_{\eta(i-1)}:s_{\eta(i)}-1]\cup[t_{\eta(i)}:t_{\eta(i-1)}-1]} g(\mathcal{X}_j, P) &\geq \sum_{j'\in[s_{\eta(i-1)}:s_{\eta(i)}-1]\cup[t_{\eta(i)}:t_{\eta(i-1)}-1]} g(\mathcal{B}_{j'}, P) \\
&\geq \sum_{j'\in[s_{\eta(i-1)}:s_{\eta(i)}-1]\cup[t_{\eta(i)}:t_{\eta(i-1)}-1]} g(I(j'), I(j'+1), P) \\
&\quad + R \sum_{j'\in[s_{\eta(i-1)}:s_{\eta(i)}-1]\cup[t_{\eta(i)}:t_{\eta(i-1)}-1]} \sum_{\ell\in[h_{j'}]} \Gamma(K_\ell(j')).
\end{aligned}$$

where the first inequality follows from the monotonicity of $g(S)$ in (B.7), and the second inequality follows from Lemma B.4. Simplifying the above result using Lemma B.5 yields

$$\begin{aligned}
&(s_{\eta(i)} - s_{\eta(i-1)} + t_{\eta(i-1)} - t_{\eta(i)}) \cdot g(\mathcal{X}_j, P) \\
&\geq g(I(s_{\eta(i-1)}), I(s_{\eta(i)}), P) + (s_{\eta(i)} - s_{\eta(i-1)} - 1)g(P) + R \sum_{j'\in[s_{\eta(i-1)}+1:s_{\eta(i)}-1]} \Gamma(I(j')) \\
&\quad + g(I(t_{\eta(i)}), I(t_{\eta(i-1)}), P) + (t_{\eta(i-1)} - t_{\eta(i)} - 1)g(P) + R \sum_{j'\in[t_{\eta(i)}+1:t_{\eta(i-1)}-1]} \Gamma(I(j')) \\
&\quad + R \sum_{j'\in[s_{\eta(i-1)}:s_{\eta(i)}-1]\cup[t_{\eta(i)}:t_{\eta(i-1)}-1]} \sum_{\ell\in[h_{j'}]} \Gamma(K_\ell(j')). \qquad \text{(B.51)}
\end{aligned}$$

Summing up (B.51) for all $i \in [2:h]$ and simplifying yields

$$\begin{aligned}
&(s_{\eta(h)} - s_{\eta(1)} + t_{\eta(1)} - t_{\eta(h)}) \cdot g(\mathcal{X}_j, P) \\
&\geq \sum_{i\in[2:h]} \big(g(I(s_{\eta(i-1)}), I(s_{\eta(i)}), P) + g(I(t_{\eta(i)}), I(t_{\eta(i-1)}), P)\big) \\
&\quad + (s_{\eta(h)} - s_{\eta(1)} + t_{\eta(1)} - t_{\eta(h)} - 2(h-1)) \cdot g(P) \\
&\quad + R \sum_{i\in[2:h]} \Big(\sum_{j'\in[s_{\eta(i-1)}+1:s_{\eta(i)}-1]} \Gamma(I(j')) + \sum_{j'\in[t_{\eta(i)}+1:t_{\eta(i-1)}-1]} \Gamma(I(j'))\Big) \\
&\quad + R \sum_{j'\in[s_{\eta(1)}:s_{\eta(h)}-1]\cup[t_{\eta(h)}:t_{\eta(1)}-1]} \sum_{\ell\in[h_{j'}]} \Gamma(K_\ell(j')). \qquad \text{(B.52)}
\end{aligned}$$

Recall that $s_{\eta(1)} = s_\theta$, $t_{\eta(1)} = t_\theta$, and $s_{\eta(h)} = s_\iota = j$, $t_{\eta(h)} = t_\iota = j+1$. Also recall that $t_\theta - s_\theta = |G_j|$. Simplifying (B.52) yields

$$\begin{aligned}
&(|G_j| - 1) \cdot g(\mathcal{X}_j, P) \\
&\geq \sum_{i\in[2:h]} \big(g(I(s_{\eta(i-1)}), I(s_{\eta(i)}), P) + g(I(t_{\eta(i)}), I(t_{\eta(i-1)}), P)\big) \\
&\quad + R \sum_{i\in[2:h]} \Big(\sum_{j'\in[s_{\eta(i-1)}+1:s_{\eta(i)}-1]} \Gamma(I(j')) + \sum_{j'\in[t_{\eta(i)}+1:t_{\eta(i-1)}-1]} \Gamma(I(j'))\Big) \\
&\quad + R \sum_{j'\in G_j\setminus\{j\}} \sum_{\ell\in[h_{j'}]} \Gamma(K_\ell(j')) + (|G_j| - 1 - 2(h-1)) \cdot g(P). \qquad \text{(B.53)}
\end{aligned}$$

Adding (B.50) and (B.53) and simplifying yields

$$\begin{aligned}
|G_j| \cdot g(\mathcal{X}_j, P) &\geq g(I(s_{\eta(1)}), I(t_{\eta(1)}), P) + R \sum_{\ell\in[h]} \Gamma(K_\ell) + R \sum_{j'\in G_j\setminus\{j\}} \sum_{\ell\in[h_{j'}]} \Gamma(K_\ell(j')) \\
&\quad + (|G_j| - 1) g(P) + R \sum_{i\in[2:h]} \big(\Gamma(I(s_{\eta(i)})) + \Gamma(I(t_{\eta(i)}))\big) \\
&\quad + R \sum_{i\in[2:h]} \Big(\sum_{j'\in[s_{\eta(i-1)}+1:s_{\eta(i)}-1]} \Gamma(I(j')) + \sum_{j'\in[t_{\eta(i)}+1:t_{\eta(i-1)}-1]} \Gamma(I(j'))\Big) \\
&= g(I(s_\theta), I(t_\theta), P) + R \sum_{j'\in G_j} \sum_{\ell\in[h_{j'}]} \Gamma(K_\ell(j')) + (|G_j| - 1) g(P) \\
&\quad + R \sum_{i\in[2:h]} \Big(\sum_{j'\in[s_{\eta(i-1)}+1:s_{\eta(i)}]} \Gamma(I(j')) + \sum_{j'\in[t_{\eta(i)}:t_{\eta(i-1)}-1]} \Gamma(I(j'))\Big) \\
&= g(I(s_\theta), I(t_\theta), P) + R \sum_{j'\in G_j} \sum_{\ell\in[h_{j'}]} \Gamma(K_\ell(j')) + (|G_j| - 1) g(P) \\
&\quad + R \sum_{j'\in[s_\theta+1:t_\theta-1]} \Gamma(I(j')),
\end{aligned}$$

which completes the proof of (B.23) and thus the proof of Lemma B.6. □

B.1.3 Proof of Lemma B.2 utilizing Lemmas B.3-B.6

So far we have introduced Lemmas B.3-B.6, with the help of which we can prove Lemma B.2 using mathematical induction. We first show the induction base that (B.13) holds when the subproblem Q is at most 1-layer in the following.

Proof. When Q is 0-layer, i.e., Q is half-rate-feasible, (B.13) directly follows from Lemma B.3.

When Q is 1-layer, consider any $P \subseteq B_Q$. Assume that the regular acyclic chain $Ch \in \mathfrak{C}(Q)$, as defined in Definition 4.4, satisfies that $\Gamma(Q) = \Gamma(Ch)$. Since Q is a 1-layer problem, Ch must be 1-layer and its components must be acyclic or half-rate-feasible. For easier reference, we repeat (4.6) as

$$Ch: \quad \underline{I(1)} \xleftrightarrow{\substack{K_{h_1}(1)\\ \cdots \\ K_1(1)}} I(2) \xleftrightarrow{\substack{K_{h_2}(2)\\ \cdots \\ K_1(2)}} \cdots \xleftrightarrow{\substack{K_{h_m}(m)\\ \cdots \\ K_1(m)}} \underline{I(m+1)}. \tag{B.54}$$

Recall that Ch can be seen as a concatenation of the regular towers $\mathcal{X}_j$, $j \in M$, and the basic towers $\mathcal{B}_{j'}$, $j' \in M' = [m] \setminus (\bigcup_{j \in M} G_j)$. By Lemmas B.4 and (B.6), as well as the monotonicity of $g(S)$ in (B.7), we have

$$\begin{aligned}
&\sum_{j \in M} \big(R \sum_{j' \in G_j} \sum_{\ell \in [h_{j'}]} \Gamma(K_\ell(j')) + g(I(s_{\theta_j,j}), I(t_{\theta_j,j}), P) \\
&\quad + R \sum_{j' \in [s_{\theta_j,j}+1 : t_{\theta_j,j}-1]} \Gamma(I(j')) + (|G_j| - 1) g(P)\big) \\
&\quad + \sum_{j' \in M'} (R \sum_{\ell \in [h_{j'}]} \Gamma(K_\ell(j')) + g(I(j'), I(j'+1), P)) \\
&\leq \sum_{j \in M} (|G_j| \cdot g(Q,P)) + \sum_{j' \in M'} g(Q,P).
\end{aligned} \tag{B.55}$$

For the RHS of (B.55), we have

$$\begin{aligned}
RHS &= g(Q,P)(\sum_{j \in M} |G_j| + |[m] \setminus (\bigcup_{j \in M} G_j)|) \\
&= m \cdot g(Q,P),
\end{aligned} \tag{B.56}$$

where the last equality is due to the fact that for any $j_1 \neq j_2 \in [m]$, $G_{j_1} \cap G_{j_2} = \emptyset$ by Definition 4.4.

Since that any basic tower $\mathcal{B}_{j'}$, $j' \in M'$ can be seen as a special regular tower with $s_{\theta_{j'},j'} =$

j', $t_{\theta_{j'},j'} = j'+1$, we can rearrange the LHS of (B.55) as follows,

$$LHS = R\sum_{j\in[m]}\sum_{\ell\in[h_j]}\Gamma(K_\ell(j)) + R\sum_{j\in M}\sum_{j'\in[s_{\theta_j,j}+1:t_{\theta_j,j}-1]}\Gamma(I(j')) + g(P)\cdot\sum_{j\in M}(|G_j|-1)$$
$$+\sum_{j\in M\cup M'} g(I(s_{\theta_j,j}),I(t_{\theta_j,j}),P), \tag{B.57}$$

The set $M\cup M'$ denotes the collection of central edges of the regular towers $\mathcal{X}_j$, $j\in M$ and the basic towers $\mathcal{B}_{j'}$, $j'\in M'$. As these towers are concatenated, if we order the elements of $M\cup M'$ as $j_1<j_2<\cdots<j_{|M\cup M'|}$, we have

$$s_{\theta_{j_1},j_1} = 1, \tag{B.58}$$
$$t_{\theta_{j_{|M\cup M'|}},j_{|M\cup M'|}} = m+1, \tag{B.59}$$
$$s_{\theta_{j_v},j_v} = t_{\theta_{j_{v-1}},j_{v-1}}, \qquad \forall v\in[2:|M\cup M'|]. \tag{B.60}$$

By the submodularity and the monotonicity of $g(S)$ in (B.7) and (B.8), and Lemma B.3 with $P\subseteq B_Q\subseteq B_{I(s_{\theta_{j_v},j_v})}$ and $I(s_{\theta_{j_v},j_v})$ being acyclic or half-rate-feasible, for any $v\in[2:|M\cup M'|]$,

$$g(I(s_{\theta_{j_1},j_1}),I(s_{\theta_{j_v},j_v}),P) + g(I(s_{\theta_{j_v},j_v}),I(t_{\theta_{j_v},j_v}),P)$$
$$\geq g(I(s_{\theta_{j_1},j_1}),I(t_{\theta_{j_v},j_v}),P) + R\cdot\Gamma(I(s_{\theta_{j_v},j_v})) + g(P). \tag{B.61}$$

Given (B.58)-(B.60), as well as the fact that j_v, $v\in[|M\cup M'|]$ is a reindexing of j, $j\in M\cup M'$, summing up (B.61) for all $v\in[2:|M\cup M'|]$ and simplifying yields

$$\sum_{j\in M\cup M'} g(I(s_{\theta_j,j}),I(t_{\theta_j,j}),P)$$
$$\geq \sum_{v\in[2:|M\cup M'|]}\big(R\cdot\Gamma(I(s_{\theta_{j_v},j_v}))+g(P)\big) + g(I(s_{\theta_{j_1},j_1}),I(t_{\theta_{j_{|M\cup M'|}},j_{|M\cup M'|}}),P)$$
$$= R\sum_{j'\in(\bigcup_{j\in M\cup M'}\{s_{\theta_j,j},t_{\theta_j,j}\})\backslash\{1,m+1\}}\Gamma(I(j')) + (|M\cup M'|-1)\cdot g(P) + g(I(1),I(m+1),P)$$
$$= R\sum_{j'\in\bigcup_{j\in M\cup M'}\{s_{\theta_j,j},t_{\theta_j,j}\}}\Gamma(I(j')) + |M\cup M'|\cdot g(P), \tag{B.62}$$

where the last equality is due to that

$$g(I(1),I(m+1),P) = g(P) + R\cdot\sum_{j'\in\{1,m+1\}}\Gamma(I(j'))$$

which can be shown by applying Lemma B.3 twice given that $P\subseteq B_Q\subseteq B_{I(j)}$ for any

$j' \in \{1, m+1\}$, that $I(1) \subseteq B_{I(m+1)}$ or $I(m+1) \subseteq B_{I(1)}$ by Definition 4.4, and that every component of Ch is acyclic or half-rate-feasible.

Combining (B.57) and (B.62), we can bound the LHS of (B.55) as

$$\begin{aligned} LHS &\geq R \sum_{j \in [m]} \sum_{\ell \in [h_j]} \Gamma(K_\ell(j)) + R \sum_{j \in M} \sum_{j' \in [s_{\theta_j, j}+1 : t_{\theta_j, j}-1]} \Gamma(I(j') + g(P) \cdot \sum_{j \in M} (|G_j| - 1) \\ &\quad + R \sum_{j' \in \bigcup_{j \in M \cup M'} \{s_{\theta_j, j}, t_{\theta_j, j}\}} \Gamma(I(j')) + |M \cup M'| \cdot g(P) \\ &= R \sum_{j \in [m]} \sum_{\ell \in [h_j]} \Gamma(K_\ell(j)) + R \sum_{j \in [m+1]} \Gamma(I(j))) + m \cdot g(P) \\ &= m \cdot R \cdot \Gamma(Ch) + m \cdot g(P) \\ &= m \cdot R \cdot \Gamma(Q) + m \cdot g(P), \end{aligned} \tag{B.63}$$

where the second equality follows directly from the definition of Γ as defined in Theorem 4.2, and the third equality follows from the assumption that $\Gamma(Q) = \Gamma(Ch)$.

Combining (B.55), (B.56) and (B.63) completes the proof of (B.13) when Q is one-layer, and hence concludes the proof of the induction base. □

It remains to prove the inductive step that (B.13) holds when Q is $s+1$-layer, given the assumption (induction hypothesis) that

$$g(Q, P) \geq g(P) + R \cdot \Gamma(Q), \quad \text{for any } Q \text{ of at most } s\text{-layer, and any } P \subseteq B_Q. \tag{B.64}$$

Such proof can be accomplished following similar steps to the proof of the induction base, utilizing the induction hypothesis (B.64) instead of Lemma B.3, and thus omitted to avoid repetition.

B.1.4 Proof of Theorem 4.2 utilizing Lemma B.2

Given Lemma B.2, Theorem 4.2 can be easily shown as follows.

Proof. It suffices to show $R \cdot \Gamma(\mathcal{G}) \leq 1$. Set $Q = [n]$, $P = \emptyset = B_Q$. Since that $g(S) \leq 1$, $\forall S \subseteq [n]$, and that $g(P) = g(\emptyset) = 0$ and $\mathcal{G}|_Q = \mathcal{G}|_{[n]} = \mathcal{G}$, we have

$$1 \geq g(Q, P) \geq g(P) + R \cdot \Gamma(Q) = R \cdot \Gamma(Q) = R \cdot \Gamma(\mathcal{G}|_{[n]}) = R \cdot \Gamma(\mathcal{G}),$$

where the second inequality follows from Lemma B.2. □

B.2 Proof of Theorem 4.4

We first introduce a few lemmas.

Lemma B.7. For any $\mathcal{G}_0$, $\mathcal{G}_1$, consider any $P, Q \subseteq V(\mathcal{G}_0)$ such that $P \subseteq B_Q$. Then for $\mathcal{G} = \mathcal{G}_0 \circ \mathcal{G}_1$ where $V(P \circ \mathcal{G}_1), V(Q \circ \mathcal{G}_1) \subseteq V(\mathcal{G})$, we have $V(P \circ \mathcal{G}_1) \subseteq B_{V(Q \circ \mathcal{G}_1)}$.

Proof. By the definition of the common interfering message set in (2.1) and Definition 2.1, for any $P, Q \subseteq V(\mathcal{G}_0)$, $P \subseteq B_Q$, we know that there is no edge going from any node in P to any node in Q in the directed graph $\mathcal{G}_0$, and hence there is no edge going from any node in $V(P \circ \mathcal{G}_1)$ to any node in $V(Q \circ \mathcal{G}_1)$ in the directed graph $\mathcal{G} = \mathcal{G}_0 \circ \mathcal{G}_1$, which indicates that $V(P \circ \mathcal{G}_1) \subseteq B_{V(Q \circ \mathcal{G}_1)}$. □

Lemma B.8. For any $\mathcal{G}_0$, $\mathcal{G}_1$, if $\mathcal{G}_0$ is acyclic or half-rate-feasible, then

$$\Gamma(\mathcal{G}_0 \circ \mathcal{G}_1) \geq \Gamma(\mathcal{G}_0) \cdot \Gamma(\mathcal{G}_1). \tag{B.65}$$

Proof. The proof for (B.65) when $\mathcal{G}_0$ is half-rate-feasible is relatively straightforward and thus omitted. Consider the case when $\mathcal{G}_0$ is acyclic. Let $n_0 = |V(\mathcal{G}_0)|$, and there exists an ordering of elements in $V(\mathcal{G}_0)$, denoted as $v_1, v_2, \cdots, v_{n_0}$ such that $v_1 \cup v_2 \cup \cdots \cup v_{i-1} \subseteq B_{v_i}$ for any $i \in [n_0]$ (cf. (2.12)). Hence, by Lemma B.7, we have

$$V(\{v_j : j \in [i-1]\} \circ \mathcal{G}_1) \subseteq B_{V(\{v_i\} \circ \mathcal{G}_1)}, \quad \forall i \in [n_0]. \tag{B.66}$$

Given (B.66), repeatedly applying Theorem 4.3 yields

$$\begin{aligned}
&\Gamma(\mathcal{G}_0 \circ \mathcal{G}_1) \\
&\geq \Gamma(\{v_1, \cdots, v_{n_0-1}\} \circ \mathcal{G}_1) + \Gamma(v_{n_0} \circ \mathcal{G}_1) \\
&\geq \Gamma(\{v_1, \cdots, v_{n_0-2}\} \circ \mathcal{G}_1) + \Gamma(v_{n_0-1} \circ \mathcal{G}_1) + \Gamma(v_{n_0} \circ \mathcal{G}_1) \\
&\cdots \\
&\geq \sum_{i \in [n_0]} \Gamma(\{v_i\} \circ \mathcal{G}_1) = n_0 \cdot \Gamma(\mathcal{G}_1) = \Gamma(\mathcal{G}_0) \cdot \Gamma(\mathcal{G}_1),
\end{aligned}$$

where the last equality follows from that $\Gamma(\mathcal{G}_0) = |V(\mathcal{G}_0)| = n_0$ since $\mathcal{G}_0$ is acyclic. □

With help of Lemmas B.7 and B.8, we prove Theorem 4.4 using mathematical induction on the number of layers of the problem $\mathcal{G}_0$ (cf. Definition B.1). We only provide proof for the induction base that (4.16) holds where $\mathcal{G}_0$ is at most 1-layer. Showing the inductive step can be done via similar steps to the proof of the induction base, utilizing the induction hypothesis instead of Lemma B.8.

Proof. The proof for (4.16) when $\mathcal{G}_0$ is 0-layer is relatively straightforward and thus omitted. When $\mathcal{G}_0$ is 1-layer, assume that the regular acyclic chain $Ch \in \mathfrak{C}(\mathcal{G}_0)$:

$$\begin{tikzcd}
\underline{I(1)} \arrow[r, leftrightarrow, "\substack{K_{h_1}(1)\\ \cdots\\ K_2(1)\\ K_1(1)}"] & I(2) \arrow[r, leftrightarrow, "\substack{K_{h_2}(2)\\ \cdots\\ K_2(2)\\ K_1(2)}"] & \cdots \arrow[r, leftrightarrow, "\substack{K_{h_m}(m)\\ \cdots\\ K_2(m)\\ K_1(m)}"] & \underline{I(m+1)},
\end{tikzcd}$$

satisfies that $\Gamma(\mathcal{G}_0) = \Gamma(Ch)$. As $\mathcal{G}_0$ is a 1-layer problem, Ch must be 1-layer with components being acyclic or half-rate-feasible. By Lemma B.7, we can construct $Ch^{\mathcal{G}_1} \in \mathfrak{C}(\mathcal{G}_0 \circ \mathcal{G}_1)$ as:

$$\begin{tikzcd}
\underline{I(1) \circ \mathcal{G}_1} \arrow[r, leftrightarrow, "\substack{K_{h_1}(1)\circ\mathcal{G}_1\\ \cdots\\ K_1(1)\circ\mathcal{G}_1}"] & {}_{\mathrm{s}}\ \cdots \arrow[r, leftrightarrow, "\substack{K_{h_m}(m)\circ\mathcal{G}_1\\ \cdots\\ K_1(m)\circ\mathcal{G}_1}"] & {}_{\mathrm{s}}\ \underline{I(m+1) \circ \mathcal{G}_1}.
\end{tikzcd}$$

By Theorem 4.2 we have

$$\begin{aligned}
&\Gamma(Ch^{\mathcal{G}_1})\\
&= \frac{1}{m}\Big(\sum_{j\in[m]} \sum_{\ell\in[h_j]} \Gamma(K_\ell(j) \circ \mathcal{G}_1) + \sum_{j\in[m+1]} \Gamma(I(j) \circ \mathcal{G}_1)\Big)\\
&\geq \Gamma(\mathcal{G}_1) \cdot \frac{1}{m}\Big(\sum_{j\in[m]} \sum_{\ell\in[h_j]} \Gamma(K_\ell(j)) + \sum_{j\in[m+1]} \Gamma(I(j))\Big)\\
&= \Gamma(\mathcal{G}_1) \cdot \Gamma(Ch)\\
&= \Gamma(\mathcal{G}_1) \cdot \Gamma(\mathcal{G}_0),
\end{aligned} \tag{B.67}$$

where the inequality is due to Lemma B.8 given the fact that every component of Ch is acyclic or half-rate-feasible. Since $Ch^{\mathcal{G}_1} \in \mathfrak{C}(\mathcal{G}_0 \circ \mathcal{G}_1)$, we have

$$\Gamma(\mathcal{G}_0 \circ \mathcal{G}_1) = \max_{Ch' \in \mathfrak{C}(\mathcal{G}_0\circ\mathcal{G}_1)} \Gamma(Ch') \geq \Gamma(Ch^{\mathcal{G}_1}). \tag{B.68}$$

Combining (B.67) and (B.68) leads to (4.16). □

B.3 Proof of Theorem 4.6

For easier reference, we repeat (4.18) in Theorem 4.6 below,

$$R \leq \frac{1}{1+m+\sum_{j\in[m]} h_j}\Big(\sum_{j\in M\cup M'} \sum_{J\in N: J\cap T_1(j)\neq\emptyset, J\cap T_2(j)\neq\emptyset} C_J \\ + \sum_{j\in M} \sum_{\ell\in[2:h_j]} \Big(\sum_{j'\in[s_{\ell,j}:s_{\ell-1,j}-1]} \sum_{\substack{J\in N: J\cap T_3(j')\neq\emptyset,\\ J\cap T_4(j,\ell,j')\neq\emptyset}} C_J + \sum_{j'\in[t_{\ell-1,j}:t_{\ell,j}-1]} \sum_{\substack{J\in N: J\cap T_3(j')\neq\emptyset,\\ J\cap T_5(j,\ell,j')\neq\emptyset}} C_J\Big)\Big), \tag{B.69}$$

where

$$\begin{aligned}
T_1(j) &= \{i(s_{h_j,j}), i(t_{h_j,j}), k_1(j), \dots, k_{h_j}(j)\},\\
T_2(j) &= \{i(s_{h_j,j}), i(m+1), k_1(j), \dots, k_{h_j}(j)\},\\
T_3(j') &= \{i(j'), i(j'+1), k_1(j'), \dots, k_{h_{j'}}(j')\},\\
T_4(j,\ell,j') &= \{i(j'), i(s_{\ell-1,j}), k_1(j'), \dots, k_{h_{j'}}(j')\},\\
T_5(j,\ell,j') &= \{i(j'), i(t_{\ell,j}), k_1(j'), \dots, k_{h_{j'}}(j')\}.
\end{aligned}$$

Consider an arbitrary DIC problem $\mathcal{G}:(i|A_i), i\in[n]$ with link capacity tuple $\mathbf{C}$. Consider an arbitrary achievable rate tuple $\mathbf{R}$.

It can be verified that there exists some two-argument set function $f: 2^{[n]}\times 2^{[n]} \to [0,1]$ that satisfies the following properties:

$$f(\emptyset;S) = f(L;\emptyset) = 0, \qquad \forall L,S\subseteq[n], \tag{B.70}$$

$$f([n];S) = f(S;S), \qquad \forall S\subseteq[n], \tag{B.71}$$

$$f(L;S) \leq \sum_{J\in N: J\cap L\neq\emptyset, J\cap S\neq\emptyset} C_J, \qquad \forall L,S\subseteq[n] \tag{B.72}$$

$$f(L;S) \leq f(L';S'), \qquad \forall L\subseteq L'\subseteq[n], S\subseteq S'\subseteq[n], \tag{B.73}$$

$$\begin{aligned} &f(L\cap L';S\cup S') + f(L\cup L';S\cap S')\\ &\quad \leq f(L;S)+f(L';S'), \qquad \forall L,L',S,S'\subseteq[n], \end{aligned} \tag{B.74}$$

$$R_i \leq f([n];\{i\}) = f([n];B\cup\{i\}) - f([n];B), \qquad \forall B\subseteq B_i, i\in[n]. \tag{B.75}$$

For more details about the above properties, see Theorem 4.7 and Corollary 4.2 in Sections 4.5 and 4.6, as well as their corresponding proofs.

When there is no ambiguity, we may slightly abuse the notation and write $f(L;S)$ as $f(L;i,i\in S)$, $f(k,k\in L;S)$, or $f(k,k\in L;i,i\in S)$. For example, for $n=3$, $f(\{1,2\};\{2,3\})$, $f(\{1,2\};2,3)$, $f(1,2;\{2,3\})$ and $f(1,2;2,3)$ mean the same thing.

We have the following lemmas.

Lemma B.9. Consider the singleton basic tower $\mathcal{B}_j^{\mathrm{s}}$ constituted by the messages $i(j), i(j+1), k_1(j), k_2(j), \ldots, k_{h_j}(j)$, we have

$$\begin{aligned} f([n]; i(j), i(j+1)) + \sum_{\ell \in [h_j]} f([n]; k_\ell(j)) \\ \leq f([n]; i(j), i(j+1), k_1(j), k_2(j), \ldots, k_{h_j}(j)). \end{aligned} \tag{B.76}$$

Lemma B.9 above can be shown via repeated application of (B.75) given the fact that $\{i(j), i(j+1), k_1(j), \ldots, k_{\ell-1}(j)\} \subseteq B_{k_\ell}, \forall \ell \in [h_j]$ according to Definition 4.5.

Lemma B.10. Consider any singleton ordered acyclic chain $Ch^{\mathrm{s,o}} \in \mathfrak{C}^{\mathrm{s,o}}(\mathcal{G})$ of length m. For any $a, b, c, d \in [m]$, we have

$$\begin{aligned} &f(i(a), i(b), k_1(d), \ldots, k_{h_d}(d); i(a), i(c), k_1(d), \ldots, k_{h_d}(d)) + f([n]; i(a), i(b), i(c)) \\ &\geq f([n]; i(a), i(b), k_1(d), \ldots, k_{h_d}(d)) + f([n]; i(a), i(c)). \end{aligned} \tag{B.77}$$

Proof. We have

$$\begin{aligned} &f(i(a), i(b), k_1(d), \ldots, k_{h_d}(d); i(a), i(c), k_1(d), \ldots, k_{h_d}(d)) + f([n]; i(a), i(b), i(c)) \\ &\geq f(i(a), i(b), k_1(d), \ldots, k_{h_d}(d); i(a), i(b), i(c), k_1(d), \ldots, k_{h_d}(d)) + f([n]; i(a), i(c)) \\ &\geq f(i(a), i(b), k_1(d), \ldots, k_{h_d}(d); i(a), i(b), k_1(d), \ldots, k_{h_d}(d)) + f([n]; i(a), i(c)) \\ &= f([n]; i(a), i(b), k_1(d), \ldots, k_{h_d}(d)) + f([n]; i(a), i(c)), \end{aligned}$$

where the first and second inequalities are due to the submodularity and monotonicity of $f(L; S)$ in (B.74) and (B.73), respectively, and the equality follows from property (B.71). □

Lemma B.11. Consider a group of messages $i(a), i(a+1), \ldots, i(b), i(b+1) \subseteq [n]$ for some positive integers $a < b \in [m]$, concatenatively located on the horizontal chain of some singleton ordered acyclic chain $Ch^{\mathrm{s,o}} \in \mathfrak{C}^{\mathrm{s,o}}(\mathcal{G})$ of length m. We have

$$\begin{aligned} \sum_{j \in [a:b]} f([n]; i(j), i(j+1)) \geq \sum_{j \in [a+1:b]} f([n]; i(j)) + \sum_{j \in [a:b]} f([n]; i(j), i(j+1), i(b+1)) \\ + f([n]; i(a), i(b+1)) - \sum_{j \in [a:b]} f([n]; i(j), i(b+1)). \end{aligned} \tag{B.78}$$

Proof. Due to the fact that when $j = b$, $f([n]; i(j), i(b+1)) = f([n]; i(j), i(j+1), i(b+1))$, it suffices to show that

$$\sum_{j\in[a:b]} f([n];i(j),i(j+1)) \geq \sum_{j\in[a+1:b]} f([n];i(j)) + \sum_{j\in[a:b-1]} f([n];i(j),i(j+1),i(b+1)) \\ + f([n];i(a),i(b+1)) - \sum_{j\in[a:b-1]} f([n];i(j),i(b+1)). \quad \text{(B.79)}$$

For any $j \in [a : b-1]$, by the submodularity of $f(L;S)$ in (B.74) and property (B.75), we have

$$f([n];i(b+1),i(j+1)) + f([n];i(j+1),i(j)) \\ \geq f([n];i(b+1),i(j)) + f([n];i(b+1),i(j+1),i(j)) - f([n];i(b+1),i(j)). \quad \text{(B.80)}$$

Summing up (B.80) for all $j \in [a : b-1]$ and simplifying the result yields (B.79), and thus completes the proof. □

So far we have introduced Lemmas B.9-B.11. Based upon these results, we further introduce the following lemma.

Lemma B.12. Consider the singleton ordered tower $\mathcal{X}_j^{\mathrm{s,o}}$ of central edge j, we have

$$\sum_{j'\in G_j}\sum_{\ell\in[h_{j'}]} f([n];k_\ell(j')) + f([n];i(s_{h_j,j}),i(t_{h_j,j})) + \sum_{\ell\in[s_{h_j,j}+1:t_{h_j,j}-1]} f([n];i(\ell)) \\ \leq f([n];k_1(j),\ldots,k_{h_j}(j),i(s_{h_j,j}),i(t_{h_j,j})) + \sum_{\ell\in[2:h_j]}\sum_{j'\in[s_{\ell,j}:s_{\ell-1,j}-1]}\sum_{\substack{J\in N:J\cap T_3(j')\neq\emptyset,\\ J\cap T_4(j,\ell,j')\neq\emptyset}} C_J \\ + \sum_{\ell\in[2:h_j]}\sum_{j'\in[t_{\ell-1,j}:t_{\ell,j}-1]}\sum_{\substack{J\in N:J\cap T_3(j')\neq\emptyset,\\ J\cap T_5(j,\ell,j')\neq\emptyset}} C_J. \quad \text{(B.81)}$$

Proof. Within the singleton ordered acyclic tower $\mathcal{X}_j^{\mathrm{s,o}}$, any edge $j' \in G_j \setminus \{j\}$ corresponds to a singleton basic tower $\mathcal{B}_{j'}^{\mathrm{s}}$. Thus by Lemma B.9 we have

$$\sum_{j'\in G_j\setminus\{j\}} \Big(\sum_{\ell\in[h_{j'}]} f([n];k_\ell(j')) + f([n];i(j'),i(j'+1)) \Big) \\ \leq \sum_{j'\in G_j\setminus\{j\}} f([n];i(j'),i(j'+1),k_1(j'),\ldots,k_{h_{j'}}(j')). \quad \text{(B.82)}$$

Consider the core of central edge j. Consider any $\ell \in [2 : h_j]$. We have

$$\begin{aligned}
&f([n]; k_1(j), \ldots, k_{\ell-1}(j), i(s_{\ell,j}), i(t_{\ell,j})) - f([n]; i(s_{\ell,j}), i(t_{\ell,j}))\\
&\quad + f([n]; i(s_{\ell,j}), i(t_{\ell,j}), i(s_{\ell-1,j}), i(t_{\ell-1,j}))\\
&\geq f([n]; k_1(j), \ldots, k_{\ell-1}(j), i(s_{\ell-1,j}), i(t_{\ell-1,j}))\\
&= f([n]; k_1(j), \ldots, k_{\ell-2}(j), i(s_{\ell-1,j}), i(t_{\ell-1,j})) + f([n]; k_{\ell-1}(j)),
\end{aligned} \tag{B.83}$$

where the inequality follows from the submodularity and the monotonicity of $f(L; S)$ in (B.74) and (B.73), and the equality follows from property (B.75) together with the fact that $B_{k_{\ell-1}(j)} \supseteq \{k_1(j), \ldots, k_{\ell-2}(j), i(s_{\ell-1,j}), i(t_{\ell-1,j})\}$ by Definition 4.7. Summing up (B.83) for all $\ell \in [2 : h_j]$ and removing redundant terms yields

$$\begin{aligned}
&f([n]; k_1(j), \ldots, k_{h_j-1}(j), i(s_{h_j,j}), i(t_{h_j,j})) + \sum_{\ell \in [2:h_j]} f([n]; i(s_{\ell,j}), i(t_{\ell,j}), i(s_{\ell-1,j}), i(t_{\ell-1,j}))\\
&\geq \sum_{\ell \in [h_j]} f([n]; i(s_{\ell,j}), i(t_{\ell,j})) + \sum_{\ell \in [h_j-1]} f([n]; k_\ell(j)).
\end{aligned} \tag{B.84}$$

Thus, we have

$$\begin{aligned}
&\sum_{\ell \in [h_j]} f([n]; i(s_{\ell,j}), i(t_{\ell,j})) + \sum_{\ell \in [h_j]} f([n]; k_\ell(j))\\
&\leq \sum_{\ell \in [2:h_j]} f([n]; i(s_{\ell,j}), i(t_{\ell,j}), i(s_{\ell-1,j}), i(t_{\ell-1,j}))\\
&\quad + f([n]; k_1(j), \ldots, k_{h_j-1}(j), i(s_{h_j,j}), i(t_{h_j,j})) + f([n]; k_{h_j}(j))\\
&= \sum_{\ell \in [2:h_j]} f([n]; i(s_{\ell,j}), i(t_{\ell,j}), i(s_{\ell-1,j}), i(t_{\ell-1,j}))\\
&\quad + f([n]; k_1(j), \ldots, k_{h_j}(j), i(s_{h_j,j}), i(t_{h_j,j})).
\end{aligned} \tag{B.85}$$

where the inequality follows from (B.84) and the equality follows from property (B.75) with the fact that $B_{k_{h_j}(j)} \supseteq \{k_1(j), \ldots, k_{h_j-1}(j), i(s_{h_j,j}), i(t_{h_j,j})\}$ by Definition 4.7.

Again consider any $\ell \in [2 : h_j]$ and define shorthand notation F_ℓ^s, F_ℓ^t, and F_ℓ as follows,

$$\begin{aligned}
F_\ell^s &\doteq \sum_{j' \in [s_{\ell,j}:s_{\ell-1,j}-1]} \sum_{\ell' \in [h_{j'}]} f([n]; k_{\ell'}(j')) + f([n]; i(s_{\ell,j}), i(s_{\ell-1,j})),\\
F_\ell^t &\doteq \sum_{j' \in [t_{\ell-1,j}:t_{\ell,j}-1]} \sum_{\ell' \in [h_{j'}]} f([n]; k_{\ell'}(j')) + f([n]; i(t_{\ell-1,j}), i(t_{\ell,j})),\\
F_\ell &\doteq f([n]; i(s_{\ell,j}), i(s_{\ell-1,j})) + f([n]; i(t_{\ell-1,j}), i(t_{\ell,j})) + f([n]; i(s_{\ell-1,j}), i(t_{\ell-1,j})).
\end{aligned}$$

Note that the Condition 2b in Definition 4.7 indicates that for any $\ell_1 \neq \ell_2 \in [2 : h_j]$, sets

$[s_{\ell_1,j} : s_{\ell_1-1,j} - 1]$, $[s_{\ell_2,j} : s_{\ell_2-1,j} - 1]$, $[t_{\ell_1-1,j} : t_{\ell_1,j} - 1]$, and $[t_{\ell_2-1,j} : t_{\ell_2,j} - 1]$ are mutually disjoint. Also, recall that $s_{j,1} = j$, and $t_{j,1} = j + 1$. Hence, one can verify that

$$\begin{aligned}
&\sum_{j' \in G_j} \sum_{\ell \in [h_{j'}]} f([n]; k_\ell(j')) + f([n]; i(s_{h_j,j}), i(t_{h_j,j})) \\
&= \sum_{\ell \in [h_j]} f([n]; i(s_{\ell,j}), i(t_{\ell,j})) + \sum_{\ell \in [h_j]} f([n]; k_\ell(j)) + \sum_{\ell \in [2:h_j]} F^s_\ell + \sum_{\ell \in [2:h_j]} F^t_\ell - \sum_{\ell \in [2:h_j]} F_\ell.
\end{aligned}$$

According to the above equation, to show (B.81), it suffices to show that

$$\begin{aligned}
&\sum_{\ell \in [h_j]} f([n]; i(s_{\ell,j}), i(t_{\ell,j})) + \sum_{\ell \in [h_j]} f([n]; k_\ell(j)) \\
&\quad + \sum_{\ell \in [2:h_j]} (F^s_\ell + F^t_\ell - F_\ell) + \sum_{\ell \in [s_{h_j,j}+1 : t_{h_j,j}-1]} f([n]; i(\ell)) \\
&\leq \sum_{\ell \in [2:h_j]} \sum_{j' \in [s_{\ell,j} : s_{\ell-1,j}-1]} \sum_{\substack{J \in N : J \cap T_3(j') \neq \emptyset, \\ J \cap T_4(j,\ell,j') \neq \emptyset}} C_J + \sum_{\ell \in [2:h_j]} \sum_{j' \in [t_{\ell-1,j} : t_{\ell,j}-1]} \sum_{\substack{J \in N : J \cap T_3(j') \neq \emptyset, \\ J \cap T_5(j,\ell,j') \neq \emptyset}} C_J \\
&\quad + f([n]; k_1(j), \ldots, k_{h_j}(j), i(s_{h_j,j}), i(t_{h_j,j})).
\end{aligned} \tag{B.86}$$

We bound F^s_ℓ, F^t_ℓ, and F_ℓ in the following. First, we have

$$\begin{aligned}
F^s_\ell &= \sum_{j' \in [s_{\ell,j} : s_{\ell-1,j}-1]} \Big(\sum_{\ell' \in [h_{j'}]} f([n]; k_{\ell'}(j')) + f([n]; i(j'), i(j'+1)) \Big) \\
&\quad - \sum_{j' \in [s_{\ell,j} : s_{\ell-1,j}-1]} f([n]; i(j'), i(j'+1)) + f([n]; i(s_{\ell,j}), i(s_{\ell-1,j})) \\
&\leq \sum_{j' \in [s_{\ell,j} : s_{\ell-1,j}-1]} f([n]; i(j'), i(j'+1), k_1(j'), \ldots, k_{h_{j'}}(j')) \\
&\quad - \sum_{j' \in [s_{\ell,j} : s_{\ell-1,j}-1]} f([n]; i(j'), i(j'+1)) + f([n]; i(s_{\ell,j}), i(s_{\ell-1,j})) \\
&\leq \sum_{j' \in [s_{\ell,j} : s_{\ell-1,j}-1]} f([n]; i(j'), i(j'+1), k_1(j'), \ldots, k_{h_{j'}}(j')) \\
&\quad + \sum_{j' \in [s_{\ell,j} : s_{\ell-1,j}-1]} f([n]; i(j'), i(s_{\ell-1,j})) - \sum_{j' \in [s_{\ell,j}+1 : s_{\ell-1,j}-1]} f([n]; i(j')) \\
&\quad - \sum_{j' \in [s_{\ell,j} : s_{\ell-1,j}-1]} f([n]; i(j'), i(j'+1), i(s_{\ell-1,j})) \\
&\leq \sum_{j' \in [s_{\ell,j} : s_{\ell-1,j}-1]} f(T_3(j'); T_4(j,\ell,j')) - \sum_{j' \in [s_{\ell,j}+1 : s_{\ell-1,j}-1]} f([n]; i(j')),
\end{aligned} \tag{B.87}$$

where the first inequality follows from Lemma B.9, the second inequality follows from Lemma B.11 with $a = s_{\ell,j}$, $b = s_{\ell-1,j} - 1$, and the third inequality follows from Lemma B.10 with $a = d = j'$, $b = j' + 1$, and $c = s_{\ell-1,j}$ for any $j' \in [s_{\ell,j} + 1 : s_{\ell-1,j} - 1]$.

Similarly, one can show that

$$F_\ell^t \leq \sum_{j' \in [t_{\ell-1,j}:t_{\ell,j}-1]} f(T_3(j'); T_5(j,\ell,j')) - \sum_{j' \in [t_{\ell-1,j}+1:t_{\ell,j}-1]} f([n]; i(j')). \tag{B.88}$$

By the submodularlity of $f(L;S)$ in (B.74) and property (B.75), we have

$$F_\ell \geq f([n]; i(s_{\ell-1,j})) + f([n]; i(t_{\ell-1,j})) + f([n]; i(s_{\ell,j}), i(s_{\ell-1,j}), i(t_{\ell-1,j}), i(t_{\ell,j})). \tag{B.89}$$

Combining (B.87)-(B.89) and rearranging, we have

$$\begin{aligned} &F_\ell^s + F_\ell^t - F_\ell + \sum_{j' \in [s_{\ell,j}+1:s_{\ell-1,j}]} f([n]; i(j')) + \sum_{j' \in [t_{\ell-1,j}:t_{\ell,j}-1]} f([n]; i(j')) \\ &\leq \sum_{j' \in [s_{\ell,j}:s_{\ell-1,j}-1]} f(T_3(j'); T_4(j,\ell,j')) + \sum_{j' \in [t_{\ell-1,j}:t_{\ell,j}-1]} f(T_3(j'); T_5(j,\ell,j')) \\ &\quad - f([n]; i(s_{\ell,j}), i(s_{\ell-1,j}), i(t_{\ell-1,j}), i(t_{\ell,j})). \end{aligned} \tag{B.90}$$

Summing up (B.90) for all $\ell \in [2:h_j]$ yields

$$\begin{aligned} &\sum_{\ell \in [2:h_j]} (F_\ell^s + F_\ell^t - F_\ell) + \sum_{\ell \in [s_{h_j,j}+1:t_{h_j,j}-1]} f([n]; i(\ell)) \\ &\leq \sum_{\ell \in [2:h_j]} \sum_{j' \in [s_{\ell,j}:s_{\ell-1,j}-1]} f(T_3(j'); T_4(j,\ell,j')) \\ &\quad + \sum_{\ell \in [2:h_j]} \sum_{j' \in [t_{\ell-1,j}:t_{\ell,j}-1]} f(T_3(j'); T_5(j,\ell,j')) \\ &\quad - \sum_{\ell \in [2:h_j]} f([n]; i(s_{\ell,j}), i(s_{\ell-1,j}), i(t_{\ell-1,j}), i(t_{\ell,j})). \end{aligned} \tag{B.91}$$

Finally, (B.86) can be shown as follows,

$$\begin{aligned}
&\sum_{\ell\in[h_j]} f([n];i(s_{\ell,j}),i(t_{\ell,j})) + \sum_{\ell\in[h_j]} f([n];k_\ell(j)) \\
&\quad + \sum_{\ell\in[2:h_j]} (F^s_\ell + F^t_\ell - F_\ell) + \sum_{\ell\in[s_{h_j,j}+1:t_{h_j,j}-1]} f([n];i(\ell)) \\
&\le \sum_{\ell\in[2:h_j]} f([n];i(s_{\ell,j}),i(t_{\ell,j}),i(s_{\ell-1,j}),i(t_{\ell-1,j})) + f([n];k_1(j),\dots,k_{h_j}(j),i(s_{h_j,j}),i(t_{h_j,j})) \\
&\quad + \sum_{\ell\in[2:h_j]}\sum_{j'\in[s_{\ell,j}:s_{\ell-1,j}-1]} f(T_3(j');T_4(j,\ell,j')) + \sum_{\ell\in[2:h_j]}\sum_{j'\in[t_{\ell-1,j}:t_{\ell,j}-1]} f(T_3(j');T_5(j,\ell,j')) \\
&\quad - \sum_{\ell\in[2:h_j]} f([n];i(s_{\ell,j}),i(s_{\ell-1,j}),i(t_{\ell-1,j}),i(t_{\ell,j})) \\
&\le f([n];k_1(j),\dots,k_{h_j}(j),i(s_{h_j,j}),i(t_{h_j,j})) \\
&\quad + \sum_{\ell\in[2:h_j]}\sum_{j'\in[s_{\ell,j}:s_{\ell-1,j}-1]}\sum_{J\in N:J\cap T_3(j')\neq\emptyset,J\cap T_4(j,\ell,j')\neq\emptyset} C_J \\
&\quad + \sum_{\ell\in[2:h_j]}\sum_{j'\in[t_{\ell-1,j}:t_{\ell,j}-1]}\sum_{J\in N:J\cap T_3(j')\neq\emptyset,J\cap T_5(j,\ell,j')\neq\emptyset} C_J,
\end{aligned}$$

where the first inequality follows from (B.85) and (B.91), and the second inequality follows from property (B.72) of $f(L;S)$. □

We show Theorem 4.6, i.e., (B.69), in the following.

Proof. Recall that the singleton ordered acyclic chain $Ch^{\mathrm{s,o}}$ can be seen as a horizontal concatenation of the singleton ordered towers $\mathcal{X}^{\mathrm{s,o}}_j$, $j \in M$ and the singleton basic towers $\mathcal{B}^{\mathrm{s}}_{j'}$, $j' \in M'$, such that the terminals of the chain $i(1)$ and $i(m+1)$ form an acyclic set. By Lemmas B.9 and B.12, we have

$$\begin{aligned}
&\sum_{j\in M}\Big(\sum_{j'\in G_j}\sum_{\ell\in[h_{j'}]} f([n];k_\ell(j')) + f([n];i(s_{h_j,j}),i(t_{h_j,j})) + \sum_{j'\in[s_{h_j,j}+1:t_{h_j,j}-1]} f([n];i(j'))\Big) \\
&\quad + \sum_{j'\in M'}\Big(\sum_{\ell\in[h_{j'}]} f([n];k_\ell(j')) + f([n];i(j'),i(j'+1))\Big)\Big) \\
&\le \sum_{j\in M}\Big(f([n];k_1(j),\dots,k_{h_j}(j),i(s_{h_j,j}),i(t_{h_j,j})) \\
&\quad + \sum_{\ell\in[2:h_j]}\sum_{j'\in[s_{\ell,j}:s_{\ell-1,j}-1]}\sum_{\substack{J\in N:J\cap T_3(j')\neq\emptyset,\\ J\cap T_4(j,\ell,j')\neq\emptyset}} C_J + \sum_{\ell\in[2:h_j]}\sum_{j'\in[t_{\ell-1,j}:t_{\ell,j}-1]}\sum_{\substack{J\in N:J\cap T_3(j')\neq\emptyset,\\ J\cap T_5(j,\ell,j')\neq\emptyset}} C_J\Big) \\
&\quad + \sum_{j'\in M'} f([n];i(j'),i(j'+1),k_1(j'),\dots,k_{h_{j'}}(j')).
\end{aligned} \tag{B.92}$$

Similar to what we have done in the proof of Thoerem 4.2 in Appendex B.1.3, we can reindex the elements of the set $M \cup M'$ as $j_1 < j_2 < \cdots < j_{|M\cup M'|}$, and the following properties hold:

$$s_{h_{j_1},j_1} = 1, \tag{B.93}$$

$$t_{h_{j_{|M\cup M'|}},j_{|M\cup M'|}} = m+1, \tag{B.94}$$

$$s_{h_{j_p},j_p} = t_{h_{j_{p-1}},j_{p-1}}, \qquad \forall p \in [2 : |M \cup M'|]. \tag{B.95}$$

For any $p \in [|M \cup M'| - 1]$, by the submodularity of $f(L;S)$ in (B.74) and property B.75, we have

$$\begin{aligned}
&f([n]; i(m+1), i(t_{h_{j_p},j_p})) + f([n]; i(t_{h_{j_p},j_p}), i(s_{h_{j_p},j_p}))\\
&\geq f([n]; i(t_{h_{j_p},j_p})) + f([n]; i(m+1), i(t_{h_{j_p},j_p}), i(s_{h_{j_p},j_p}))\\
&= f([n]; i(t_{h_{j_p},j_p})) + f([n]; i(m+1), i(s_{h_{j_p},j_p}))\\
&\quad + f([n]; i(m+1), i(t_{h_{j_p},j_p}), i(s_{h_{j_p},j_p})) - f([n]; i(m+1), i(s_{h_{j_p},j_p}))).
\end{aligned} \tag{B.96}$$

Given the fact that j_p, $p \in [|M \cup M'|]$ is a reindexing of j, $j \in M \cup M'$, as well as the properties (B.93)-(B.95), summing up (B.96) for all $p \in [|M \cup M'| - 1]$ yields

$$\begin{aligned}
&\sum_{j\in M\cup M'} f([n]; i(s_{h_j,j}), i(t_{h_j,j}))\\
&\geq \sum_{j'\in(\bigcup_{j\in M\cup M'}\{s_{h_j,j},t_{h_j,j}\})\setminus\{1,m+1\}} f([n]; i(j')) + f([n]; i(m+1), i(1))\\
&\quad + \sum_{p\in[|M\cup M'|-1]} f([n]; i(m+1), i(t_{h_{j_p},j_p}), i(s_{h_{j_p},j_p}))\\
&\quad - \sum_{p\in[|M\cup M'|-1]} f([n]; i(m+1), i(s_{h_{j_p},j_p}))\\
&= \sum_{j'\in\bigcup_{j\in M\cup M'}\{s_{h_j,j},t_{h_j,j}\}} f([n]; i(j'))\\
&\quad + \sum_{p\in[|M\cup M'|-1]} f([n]; i(m+1), i(t_{h_{j_p},j_p}), i(s_{h_{j_p},j_p}))\\
&\quad - \sum_{p\in[|M\cup M'|-1]} f([n]; i(m+1), i(s_{h_{j_p},j_p})),
\end{aligned} \tag{B.97}$$

where the equality follows from property (B.75) with the fact that $i(1) \in B_{i(m+1)}$ or $i(m+1) \in B_{i(1)}$ according to Definition 4.8.

The LHS of (B.92) can be bounded as,

$$
\begin{aligned}
LHS = & \sum_{j\in M}\sum_{j'\in G_j}\sum_{\ell\in[h_{j'}]} f([n];k_\ell(j')) + \sum_{j'\in M'}\sum_{\ell\in[h_{j'}]} f([n];k_\ell(j')) \\
& + \sum_{j\in M}\sum_{j'\in[s_{h_j,j}+1:t_{h_j,j}-1]} f([n];i(j')) + \sum_{j\in M\cup M'} f([n];i(s_{h_j,j}),i(t_{h_j,j})) \\
\geq & \sum_{j\in[m]}\sum_{\ell\in[h_j]} f([n];k_\ell(j)) + \sum_{j\in M}\sum_{j'\in[s_{h_j,j}+1:t_{h_j,j}-1]} f([n];i(j')) \\
& + \sum_{j'\in\bigcup_{j\in M\cup M'}\{s_{h_j,j},t_{h_j,j}\}} f([n];i(j')) + \sum_{p\in[|M\cup M'|-1]} f([n];i(m+1),i(t_{h_{j_p},j_p}),i(s_{h_{j_p},j_p})) \\
& - \sum_{p\in[|M\cup M'|-1]} f([n];i(m+1),i(s_{h_{j_p},j_p})) \\
\geq & \sum_{j\in[m]}\sum_{\ell\in[h_j]} R_{k_\ell(j)} + \sum_{j\in[m+1]} R_{i(j)} + \sum_{p\in[|M\cup M'|-1]} f([n];i(m+1),i(t_{h_{j_p},j_p}),i(s_{h_{j_p},j_p})) \\
& - \sum_{p\in[|M\cup M'|-1]} f([n];i(m+1),i(s_{h_{j_p},j_p})) \\
= & (1+m+\sum_{j\in[m]} h_j)R + \sum_{p\in[|M\cup M'|]} f([n];i(m+1),i(t_{h_{j_p},j_p}),i(s_{h_{j_p},j_p})) \\
& - \sum_{p\in[|M\cup M'|]} f([n];i(m+1),i(s_{h_{j_p},j_p})) \\
= & (1+m+\sum_{j\in[m]} h_j)R + \sum_{j\in M\cup M'} f([n];i(m+1),i(t_{h_j,j}),i(s_{h_j,j})) \\
& - \sum_{j\in M\cup M'} f([n];i(m+1),i(s_{h_j,j})),
\end{aligned} \tag{B.98}
$$

where the first equality follows from simply rearranging the terms of the LHS of (B.92), the first inequality follows from (B.97), the second inequality follows from that fact that $R_i \leq f([n];\{i\}), \forall i \in [n]$ according to property (B.75), the second equality follows from (B.94), and the third equality follows from the fact that $j_p,\ p \in [|M \cup M'|]$ is a reindexing of j, $j \in M \cup M'$.

For the RHS of (B.92), simply rearranging the terms we obtain

$$
\begin{aligned}
RHS = & \sum_{j\in M\cup M'} \big(f([n];k_1(j),\dots,k_{h_j}(j),i(s_{h_j,j}),i(t_{h_j,j}))\big) \\
& + \sum_{j\in M}\Big(\sum_{\ell\in[2:h_j]}\sum_{j'\in[s_{\ell,j}:s_{\ell-1,j}-1]}\sum_{\substack{J\in N:J\cap T_3(j')\neq\emptyset,\\ J\cap T_4(j,\ell,j')\neq\emptyset}} C_J \\
& + \sum_{\ell\in[2:h_j]}\sum_{j'\in[t_{\ell-1,j}:t_{\ell,j}-1]}\sum_{\substack{J\in N:J\cap T_3(j')\neq\emptyset,\\ J\cap T_5(j,\ell,j')\neq\emptyset}} C_J\Big).
\end{aligned} \tag{B.99}
$$

Given (B.92), (B.98), and (B.99), we can conclude that

$$
\begin{aligned}
&(1+m+\sum_{j\in[m]} h_j)R \\
&\quad -\sum_{j\in M}\Big(\sum_{\ell\in[2:h_j]}\ \sum_{j'\in[s_{\ell,j}:s_{\ell-1,j}-1]}\ \sum_{\substack{J\in N:J\cap T_3(j')\neq\emptyset,\\ J\cap T_4(j,\ell,j')\neq\emptyset}} C_J+\sum_{\ell\in[2:h_j]}\ \sum_{j'\in[t_{\ell-1,j}:t_{\ell,j}-1]}\ \sum_{\substack{J\in N:J\cap T_3(j')\neq\emptyset,\\ J\cap T_5(j,\ell,j')\neq\emptyset}} C_J\Big) \\
&\leq \sum_{j\in M\cup M'} \big(f([n];k_1(j),\dots,k_{h_j}(j),i(s_{h_j,j}),i(t_{h_j,j})) \\
&\quad -\sum_{j\in M\cup M'} f([n];i(m+1),i(t_{h_j,j}),i(s_{h_j,j}))+\sum_{j\in M\cup M'} f([n];i(m+1),i(s_{h_j,j})) \\
&\leq \sum_{j\in M\cup M'} f(T_1(j);T_2(j)) \\
&\leq \sum_{j\in M\cup M'}\ \sum_{J\in N:J\cap T_1(j)\neq\emptyset,J\cap T_2(j)\neq\emptyset} C_J, \qquad \text{(B.100)}
\end{aligned}
$$

where the second inequality follows from Lemma B.10 with $a = s_{h_j,j}$, $b = t_{h_j,j}$, $c = m+1$, and $d = j$ for any $j \in M \cup M'$, and the last inequality follows from property (B.72) of $f(L;S)$.

Rearranging and simplifying (B.100) yields (B.69) and thus completes the proof. □

B.4 Remainder of Proof of Theorem 4.7

First we present a lemma based on the encoding condition in (4.20) and the touch structure.

Lemma B.13. For any set $K \subseteq [n]$, $H(T_{\overline{K}}|X_{K^c}) = 0$.

Proof. For any set $K \subseteq [n]$, we have

$$
H(T_{\overline{K}}|X_{K^c}) = H(Y_{\{J:J\in N,J\cap K=\emptyset\}}|X_{K^c}) = H(Y_{\{J:J\subseteq K^c\}}|X_{K^c}) \leq \sum_{J\subseteq K^c} H(Y_J|X_{K^c}) = 0, \qquad \text{(B.101)}
$$

where the last equality follows from the encoding condition in (4.20). □

Now we prove that the set function $f(G,K)$ defined in (4.37) satisfies Axioms (4.23)-(4.29) of Theorem 4.7. Toward that end, we first show that for any $\epsilon > 0$, the set function $f_\epsilon(G,K)$ defined in (4.34) satisfies the following conditions, which are counterparts of (4.23)-(4.27) and

(4.29) (we deal with (4.28) later):

$$f_\epsilon(G,K) = f_\epsilon(G',K), \qquad \text{if } (P_G \cup P_{G'}) \setminus (P_G \cap P_{G'}) \subseteq T_{\overline{K}}, \tag{B.102}$$

$$f_\epsilon(\emptyset,K) = f_\epsilon(G,\emptyset) = 0, \tag{B.103}$$

$$f_\epsilon(G,K) \leq \sum_{J: J \in P_G, J \in T_K} C_J, \tag{B.104}$$

$$f_\epsilon(G,K) \leq f_\epsilon(G',K'), \qquad \text{if } K \subseteq K', G \subseteq G', \tag{B.105}$$

$$f_\epsilon(G \cup G', K \cap K') + f_\epsilon(G \cap G', K \cup K') \leq f_\epsilon(G,K) + f_\epsilon(G',K'), \tag{B.106}$$

$$f_\epsilon(G,K) + f_\epsilon(G,K') = f_\epsilon(G, K \cup K'), \qquad \text{if } K \cap K' = \emptyset, P_G \subseteq (N \setminus T_{K,K'}). \tag{B.107}$$

1. We show that $f_\epsilon(G,K)$ satisfies (B.102) as follows. Note that if $((P_G \cup P_{G'}) \setminus (P_G \cap P_{G'})) \subseteq T_{\overline{K}}$, according to Lemma B.13, we have

$$\begin{aligned} 0 &= H(T_{\overline{K}} | X_{K^c}) \\ &\geq H((P_G \cup P_{G'}) \setminus (P_G \cap P_{G'}) | X_{K^c}) \\ &\geq H(P_G \setminus P_{G'} | X_{K^c}) \\ &\geq H(P_G \setminus P_{G'} | X_{K^c}, P_{G'}), \end{aligned} \tag{B.108}$$

$$\begin{aligned} 0 &= H(T_{\overline{K}} | X_{K^c}) \\ &\geq H((P_G \cup P_{G'}) \setminus (P_G \cap P_{G'}) | X_{K^c}) \\ &\geq H(P_{G'} \setminus P_G | X_{K^c}) \\ &\geq H(P_{G'} \setminus P_G | X_{K^c}, P_G). \end{aligned} \tag{B.109}$$

Given (B.108) and (B.109) as well as the nonnegativity of entropy, we have

$$H(P_G \setminus P_{G'} | X_{K^c}, P_{G'}) = 0, \tag{B.110}$$

$$H(P_{G'} \setminus P_G | X_{K^c}, P_G) = 0. \tag{B.111}$$

Therefore,

$$\begin{aligned} f_\epsilon(G,K) &= \frac{1}{r} H(P_G | X_{K^c}) \\ &= \frac{1}{r} H(P_G, P_{G'} \setminus P_G | X_{K^c}) \\ &= \frac{1}{r} H(P_{G'}, P_G \setminus P_{G'} | X_{K^c}) \\ &= \frac{1}{r} H(P_{G'} | X_{K^c}) = f_\epsilon(G',K), \end{aligned}$$

where the second and fourth equalities follow from (B.111) and (B.110), respectively.

2. We show that $f_\epsilon(G,K)$ satisfies (B.103) as follows. It is obvious that for any $G \subseteq [m]$, $K \subseteq [n]$,

$$f_\epsilon(\emptyset, K) = \frac{1}{r}H(P_\emptyset | X_{K^c}) = \frac{1}{r}H(\emptyset | X_{K^c}) = 0, \tag{B.112}$$

$$f_\epsilon(G, \emptyset) = \frac{1}{r}H(P_G | X_{\emptyset^c}) = \frac{1}{r}H(P_G | X_{[n]}) = 0, \tag{B.113}$$

where the last equality of (B.113) is due to the encoding condition in (4.20).

3. We show that $f_\epsilon(G,K)$ satisfies (B.104) as follows. According to Lemma B.13 and the fact that conditioning cannot increase entropy, we have

$$\begin{aligned} f_\epsilon(G,K) &= \frac{1}{r}H(P_G | X_{K^c}) = \frac{1}{r}H(P_G \cup T_{\overline{K}} | X_{K^c}) \\ &= \frac{1}{r}H(T_{\overline{K}}, P_G \setminus T_{\overline{K}} | X_{K^c}) = \frac{1}{r}H(P_G \setminus T_{\overline{K}} | X_{K^c}) \\ &\leq \frac{1}{r}H(P_G \setminus T_{\overline{K}}) \leq \frac{1}{r}\sum_{J:J\in P_G, J\in T_K} r_J \leq \sum_{J:J\in P_G, J\in T_K} C_J. \end{aligned}$$

4. We show that $f_\epsilon(G,K)$ satisfies (B.105) as follows. Note that if $G \subseteq G' \subseteq [m]$, $K \subseteq K' \subseteq [n]$, then $K'^c \subseteq K^c$. Therefore, we have

$$f_\epsilon(G,K) = \frac{1}{r}H(P_G | X_{K^c}) \leq \frac{1}{r}H(P_G | X_{K'^c}) \leq \frac{1}{r}H(P_{G'} | X_{K'^c}) = f_\epsilon(G',K').$$

5. We show that $f_\epsilon(G,K)$ satisfies (B.106) as follows. Let us define $G_1 = G \setminus G'$, $G_2 = G' \setminus G$ and $G_0 = G \cap G'$, so that $G \cup G' = G_0 \cup G_1 \cup G_2$ is the union of three disjoint sets and $G = G_0 \cup G_1$ and $G' = G_0 \cup G_2$. Similarly, define $K_0 = K^c \cap K'^c$, $K_1 = K^c \setminus K'^c$ and $K_2 = K'^c \setminus K^c$ so that $K^c \cup K'^c = K_0 \cup K_1 \cup K_2$, $K^c = K_0 \cup K_1$ and $K'^c = K_0 \cup K_2$.

Set

$$f_1 = f_\epsilon(G \cup G', K \cap K') + f_\epsilon(G \cap G', K \cup K') + \frac{1}{r}H_1,$$

where $H_1 = H(X_{K_0 \cup K_1 \cup K_2}) + H(X_{K_0})$, and set

$$f_2 = f_\epsilon(G,K) + f_\epsilon(G',K') + \frac{1}{r}H_2,$$

where $H_2 = H(X_{K_0 \cup K_1}) + H(X_{K_0 \cup K_2})$. Due to the message independence in (4.19) and sets K_0, K_1, K_2 being disjoint, we have $H_1 = H_2$.

We can verify that for any server grouping $\mathcal{P} = \{P_1, P_2, \cdots, P_m\}$, $G \subseteq [m]$, we have

$$P_{G \cup G'} = \bigcup_{i \in G_0 \cup G_1 \cup G_2} P_i = P_G \cup P_{G'}, \tag{B.114}$$

$$P_{G \cap G'} = \bigcup_{i \in G_0} P_i \subseteq P_G \cap P_{G'}, \tag{B.115}$$

where (B.115) is due to the possible overlapping between two different server groups, e.g., even for two disjoint sets $G_1, G_2 \subseteq [m]$, $P_{G_1} \cap P_{G_2}$ may not be $\emptyset$. If $P_1, \cdots, P_m$ happen to be disjoint server groups, we will have $P_{G \cap G'} = P_G \cap P_{G'}$.

Therefore, we can write

$$\begin{aligned}
rf_1 &= H(P_{G \cup G'} | X_{K_0 \cup K_1 \cup K_2}) + H(P_{G \cap G'} | X_{K_0}) + H_1 \\
&= H(P_{G \cup G'}, X_{K_0 \cup K_1 \cup K_2}) + H(P_{G \cap G'}, X_{K_0}) \\
&\leq H(P_G \cup P_{G'}, X_{K_0 \cup K_1 \cup K_2}) + H(P_G \cap P_{G'}, X_{K_0}) \\
&\leq H(P_G, X_{K_0 \cup K_1}) + H(P_{G'}, X_{K_0 \cup K_2}) \\
&= H(P_G | X_{K_0 \cup K_1}) + H(P_{G'} | X_{K_0 \cup K_2}) + H_2 = rf_2.
\end{aligned}$$

where the first inequality is due to (B.114) and (B.115) and the second inequality is due to the submodularity of the entropy function. Finally, we have

$$\begin{aligned}
f_\epsilon(G \cup G', K \cap K') + f_\epsilon(G \cap G', K \cup K') = f_1 - \frac{1}{r} H_1 \leq f_2 - \frac{1}{r} H_2 \\
= f_\epsilon(G, K) + f_\epsilon(G', K').
\end{aligned}$$

6. We show that $f_\epsilon(G, K)$ satisfies (B.107) as follows. For any $K, K' \subseteq [n]$ such that $K \cap K' = \emptyset$, set $L = [n] \setminus (K \cup K')$. For $P_G \subseteq (N \setminus T_{K,K'})$, according to Proposition B.1 presented in Section B.5, we have

$$\begin{aligned}
0 &= I(X_K; X_{K'} | P_G, X_L) \\
&= H(X_K | P_G, X_L) + H(X_{K'} | P_G, X_L) - H(X_K, X_{K'} | P_G, X_L) \\
&= H(P_G, X_{K \cup L}) - H(P_G, X_L) + H(P_G, X_{K' \cup L}) \\
&\quad - H(P_G, X_L) - H(P_G, X_{[n]}) + H(P_G, X_L) \\
&= H(P_G, X_{K \cup L}) + H(P_G, X_{K' \cup L}) - H(P_G, X_L) - H(X_{[n]}) \qquad \text{(B.116)} \\
&= H(P_G | X_{K \cup L}) + H(P_G | X_{K' \cup L}) - H(P_G | X_L) \qquad \text{(B.117)} \\
&= r(f_\epsilon(G, K') + f_\epsilon(G, K) - f_\epsilon(G, K \cup K')), \qquad \text{(B.118)}
\end{aligned}$$

where (B.116) follows from the encoding condition in (4.20), and (B.117) follows from the message independence in (4.19). Obviously, given r being positive, by (B.118), we have $f_\epsilon(G,K') + f_\epsilon(G,K) = f_\epsilon(G,K \cup K')$.

Now that we have shown that $f_\epsilon(G,K)$ satisfies (B.102)-(B.107), we use this result to show that $f(G,K)$ satisfies corresponding six axioms (4.23)-(4.27) and (4.29) in the following.

1. For Axiom (4.23), if $((P_G \cup P_{G'}) \setminus (P_G \cap P_{G'})) \subseteq T_{\overline{K}}$, we have

$$f(G,K) = \liminf_{\epsilon\to 0} f_\epsilon(G,K) = \liminf_{\epsilon\to 0} f_\epsilon(G',K) = f(G',K),$$

where the second equality follows from (B.102).

2. For Axiom (4.24), due to (B.103), we have

$$\begin{aligned} f(\emptyset,K) &= \liminf_{\epsilon\to 0} f_\epsilon(\emptyset,K) = 0, \\ f(G,\emptyset) &= \liminf_{\epsilon\to 0} f_\epsilon(G,\emptyset) = 0. \end{aligned}$$

3. For Axiom (4.25), due to (B.104), we have

$$f(G,K) = \liminf_{\epsilon\to 0} f_\epsilon(G,K) \le \sum_{J: J\in P_G, J\in T_K} C_J.$$

4. For Axiom (4.26), consider any $G \subseteq G' \subseteq [m]$, $K \subseteq K' \subseteq [n]$. By (B.105), we have

$$\begin{aligned} 0 &\le \liminf_{\epsilon\to 0}(f_\epsilon(G',K') - f_\epsilon(G,K)) \\ &= \liminf_{\epsilon\to 0} f_\epsilon(G',K') - \limsup_{\epsilon\to 0} f(G,K) \\ &\le \liminf_{\epsilon\to 0} f_\epsilon(G',K') - \liminf_{\epsilon\to 0} f(G,K) \\ &= f(G',K') - f(G,K). \end{aligned}$$

5. For Axiom (4.27), consider any $G, G' \subseteq [m]$, $K, K' \subseteq [n]$. By (B.106), we have

$$\begin{aligned} 0 &\le \liminf_{\epsilon\to 0}(f_\epsilon(G',K') + f_\epsilon(G,K) - f_\epsilon(G\cup G',K\cap K') - f_\epsilon(G\cap G',K\cup K')) \\ &= \liminf_{\epsilon\to 0} f_\epsilon(G',K') + \liminf_{\epsilon\to 0} f_\epsilon(G,K) \\ &\quad - \limsup_{\epsilon\to 0} f(G\cup G',K\cap K') - \limsup_{\epsilon\to 0} f(G\cap G',K\cup K') \\ &\le \liminf_{\epsilon\to 0} f_\epsilon(G',K') + \liminf_{\epsilon\to 0} f_\epsilon(G,K) \\ &\quad - \liminf_{\epsilon\to 0} f(G\cup G',K\cap K') - \liminf_{\epsilon\to 0} f(G\cap G',K\cup K') \\ &= f(G',K') + f(G,K) - f(G\cup G',K\cap K') - f(G\cap G',K\cup K'). \end{aligned}$$

6. For Axiom (4.29), consider any $G \subseteq [m]$, $K, K' \subseteq [n]$ such that $K \cap K' = \emptyset$ and $P_G \subseteq (N \setminus T_{K,K'})$. By (B.107), we have

$$\begin{aligned} f(G,K) + f(G,K') &= \liminf_{\epsilon \to 0} (f_\epsilon(G,K) + f_\epsilon(G,K')) \\ &= \liminf_{\epsilon \to 0} f_\epsilon(G, K \cup K') = f(G, K \cup K'). \end{aligned}$$

Finally for the last remaining axiom, Axiom (4.28), we use an approach similar to the one used in the inequality leading up to (4.42). Consider any $\epsilon > 0$ and $i \in [n]$. We rearrange (4.33) as

$$t_i \leq \frac{H(Y_N|X_{A_i}) - H(Y_N|X_{A_i \cup \{i\}})}{1 - \delta(\epsilon)}. \tag{B.119}$$

We have

$$\begin{aligned} rf_\epsilon([m], \{i\}) &= H(Y_N|X_{\{i\}^c}) - H(Y_N|X_{[n]}) && \text{(B.120)} \\ &= I(X_i; Y_N|X_{A_i \cup B_i}) \\ &= H(X_i|X_{A_i \cup B_i}) - H(X_i|Y_N, X_{A_i \cup B_i}) \\ &\leq H(X_i) \\ &= t_i && \text{(B.121)} \\ &\leq \frac{H(Y_N|X_{A_i}) - H(Y_N|X_{A_i \cup \{i\}})}{1 - \delta(\epsilon)} && \text{(B.122)} \\ &= \frac{rf_\epsilon([m], B_i \cup \{i\}) - rf_\epsilon([m], B_i)}{1 - \delta(\epsilon)} && \text{(B.123)} \end{aligned}$$

where (B.120) follows from the encoding condition in (4.20), (B.121) follows from the fact that the messages are uniformly distributed as specified in (4.19), (B.122) follows from (B.119), and (B.123) follows from the definition of $f_\epsilon(G, K)$. Dividing both sides of (B.123) by r and then taking the limit infimum as ϵ approaches zero, we have

$$\begin{aligned} f([m], \{i\}) &\leq \liminf_{\epsilon \to 0} \frac{f_\epsilon([m], B_i \cup \{i\}) - f_\epsilon([m], B_i)}{1 - \delta(\epsilon)} \\ &= \liminf_{\epsilon \to 0} (f_\epsilon([m], B_i \cup \{i\}) - f_\epsilon([m], B_i)) \\ &= \liminf_{\epsilon \to 0} f_\epsilon([m], B_i \cup \{i\}) - \limsup_{\epsilon \to 0} f_\epsilon([m], B_i) \\ &\leq \liminf_{\epsilon \to 0} f_\epsilon([m], B_i \cup \{i\}) - \liminf_{\epsilon \to 0} f_\epsilon([m], B_i) \\ &= f([m], B_i \cup \{i\}) - f([m], B_i). && \text{(B.124)} \end{aligned}$$

On the other hand, by Axiom (4.27) we have

$$f([m],\{i\}) \geq f([m],B_i \cup \{i\}) - f([m],B_i). \tag{B.125}$$

Combining (B.124) and (B.125) leads to Axiom (4.28).

This concludes the proof of Theorem 4.7.

B.5 Functional Dependence Graph and fd-separation for Distributed Index Coding

We review the functional dependence graph (FDG), which was first introduced in Kramer [1998] and then further developed in Thakor et al. [2016]. We first restate the general definition of FDG and then specialize it to the distributed index coding functional dependence graph (DIC-FDG) based on the DIC problem setup.

Within this section, we use $\mathcal{G} = (\mathcal{V},\mathcal{E})$ to denote a directed graph with vertex set $\mathcal{V} = \{V_1, V_2, \cdots\}$ and directed edge set $\mathcal{E} = \{e_1, e_2, \cdots\}$. We use $\mathsf{tail}(e)$ and $\mathsf{head}(e)$ to denote the tail and the head of the directed edge $e, \forall e \in \mathcal{E}$, respectively. For any $V_j, V_k \in \mathcal{V}$, $j < k$, we say that vertices $V_j, V_k \in \mathcal{V}$ are *connected* if there exist vertices in $\mathcal{V}$, $V_j, V_{j+1}, \ldots, V_k$, and edges in $\mathcal{E}$, $e_j, \ldots, e_{k-1}$, such that for any $i \in [j:k-1]$, we have either $\mathsf{tail}(e_i) = V_i$, $\mathsf{head}(e_i) = V_{i+1}$ or $\mathsf{tail}(e_i) = V_{i+1}$, $\mathsf{head}(e_i) = V_i$. We call such vertex sequence $V_j, V_{j+1}, \ldots, V_k$ a path between V_j and V_k. Correspondingly, we say that V_j and V_k are *disconnected* if such intermediate vertices and edges do not exist (i.e., there is no path between V_j and V_k). Note that we ignore the direction of the edges when determining whether two vertices are connected or not.

Definition B.2 (Functional dependence graph (FDG))**.** Let $\mathcal{V} = \{V_1, V_2, \ldots\}$ be a set of random variables. A directed graph $\mathcal{G} = (\mathcal{V},\mathcal{E})$ is called a functional dependence graph (FDG) for $\mathcal{V}$ if and only if

$$H(V_i | V_j : (V_j, V_i) \in \mathcal{E}) = 0, \quad \forall V_i \in \mathcal{V}. \tag{B.126}$$

Definition B.3 (Distributed index coding functional dependence graph (DIC-FDG))**.** For a given DIC problem, its distributed index coding functional dependence graph (DIC-FDG) is a directed graph $\mathcal{G} = (\mathcal{V},\mathcal{E})$ defined as follows.

- The set of vertices $\mathcal{V} = X_{[n]} \cup Y_N$.
- For any $V, V' \in \mathcal{V}$, $(V, V') \in \mathcal{E}$ if and only if it satisfies one of the following conditions:
  1. $V = X_i, V' = Y_J, J \in N, i \in J$, i.e., (V, V') denotes message availability at server J;

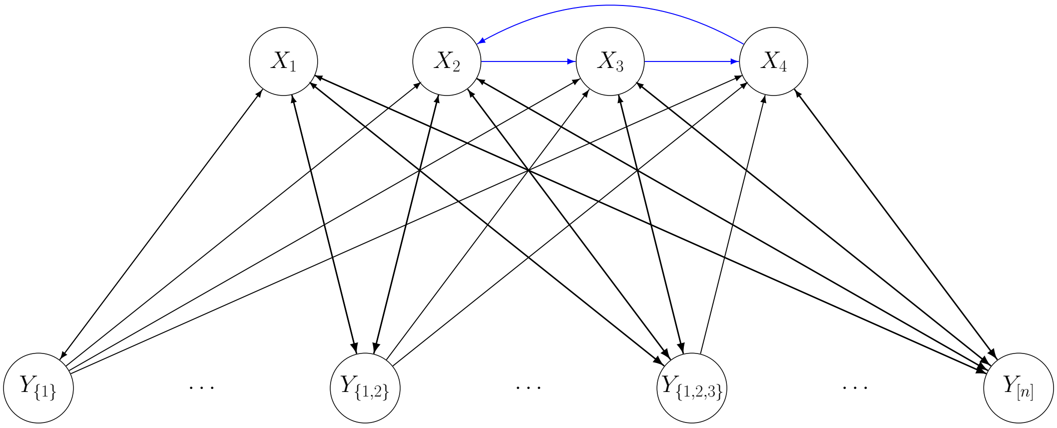


Figure B.2: The DIC-FDG for the 4-message problem: $(1|-), (2|4), (3|2), (4|3)$, with all $2^4 - 1 = 15$ servers. For simplicity, only four output variables, $Y_{\{1\}}$, $Y_{\{1,2\}}$, $Y_{\{1,2,3\}}$, and $Y_{\{1,2,3,4\}}$, and their corresponding links are shown. To avoid clutter, whenever there exist directed edges in both directions between any two vertices, we simply draw an edge with arrows at both ends between the vertices, instead of drawing two separate directed edges. Note that the edges defined in item 2 of Definition B.3 are shown as blue, while all the other edges defined in items 1 and 3 are shown in black.

2. $V = X_i, V' = X_j, j \in [n], i \in A_j$, i.e., (V, V') denotes side information availability at receiver j;
3. $V = Y_J, V' = X_i, J \in N, i \in [n]$, i.e., (V, V') denotes a broadcast link from server J to receiver i;

Note that the encoding conditions $H(Y_J|X_J) = 0$ are captured in the DIC-FDG due to the existence of edges as defined in item 1 above. The decoding conditions $H(X_i|Y_N, X_{A_i}) = 0$ are captured due to the existence of edges as defined in items 2 and 3.[2] Hence, it can be verified that for any DIC-FDG $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ we have

$$H(V_i|V_j : (V_j, V_i) \in \mathcal{E}) = 0, \quad \forall V_i \in \mathcal{V}.$$

Therefore, the DIC-FDG defined in Definition B.3 is indeed an FDG satisfying Definition B.2.

Example B.2. See Figure B.2 for the DIC-FDG for the DIC problem: $(1|-), (2|4), (3|2), (4|3)$.

Now we review the fd-separation criterion, also from Thakor et al. [2016]; Kramer [1998], which leads to the conditional message independence utilized in Axiom (4.29) of Theorem 4.7 in Section 4.5. Similar to the DIC-FDG, the fd-separation presented here has also been specialized to the distributed index coding scenario.

[2] Note that for the DIC-FDG, we assume zero-error decoding conditions at receivers for simplicity. However, as the fd-separation and Proposition B.1 (to be defined shortly) depend only on the message independence and the encoding conditions at servers, they hold in the general case of vanishing decoding error probability.

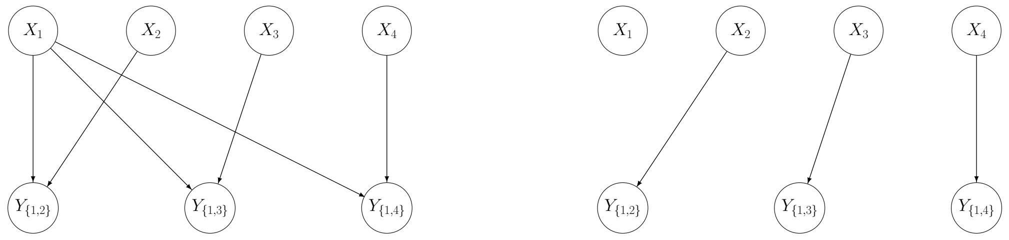


Figure B.3: The ancestral graph $\mathcal{G}_{\mathrm{An}(\mathcal{U}\cup\mathcal{W}\cup\mathcal{Z})}$ is shown on the left. And the remaining of $\mathcal{G}_{\mathrm{An}(\mathcal{U}\cup\mathcal{W}\cup\mathcal{Z})}$ after removing all edges outgoing from vertices in $\mathcal{U}$ is shown on the right, where X_2 becomes disconnected from both X_3 and X_4.

Definition B.4 (Ancestral graph)**.** Consider the DIC-FDG $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ of a given DIC problem. For any subset $\mathcal{A} \subseteq \mathcal{V}$, let $\mathrm{An}(\mathcal{A})$ be the set of all vertices in $\mathcal{V} \setminus \mathcal{A}$ such that for every vertex $V \in \mathrm{An}(\mathcal{A})$, there is a directed path from V to some vertex V' in $\mathcal{A}$ in the subgraph $\tilde{\mathcal{G}} = \mathcal{G} \setminus \{e \in \mathcal{E} : \exists i \in [n], \mathrm{head}(e) = X_i\}$. The ancestral graph with respect to $\mathcal{A}$, denoted by $\mathcal{G}_{\mathrm{An}(\mathcal{A})}$, is a vertex-induced subgraph of $\mathcal{G}$ consisting of vertices $(\mathcal{A} \cup \mathrm{An}(\mathcal{A}))$ and edges $e \in \mathcal{E}$ such that $\mathrm{head(e)}, \mathrm{tail(e)} \in \mathcal{A} \cup \mathrm{An}(\mathcal{A})$.

Definition B.5 (fd-separation)**.** Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be the DIC-FDG of a given DIC problem, and let $\mathcal{U}, \mathcal{W}, \mathcal{Z}$ be three nonempty disjoint subsets of $\mathcal{V}$. Set $\mathcal{U}$ fd-separates sets $\mathcal{W}$ and $\mathcal{Z}$ if every vertex in $\mathcal{W}$ is disconnected from every vertex in $\mathcal{Z}$ in what remains of $\mathcal{G}_{\mathrm{An}(\mathcal{U}\cup\mathcal{W}\cup\mathcal{Z})}$ after removing all edges outgoing from vertices in $\mathcal{U}$.

Example B.3. Recall the 4-message DIC problem whose DIC-FDG is shown in Figure B.2. Set $\mathcal{U} = \{X_{\{1\}}, Y_{\{\{1,2\},\{1,3\},\{1,4\}\}}\}, \mathcal{W} = \{X_{\{2\}}\}, \mathcal{Z} = \{X_{\{3,4\}}\}$. It can be verified that $\mathcal{U}$ fd-separates $\mathcal{W}$ and $\mathcal{Z}$, as illustrated by Figure B.3.

It can be verified that once the subset of vertices $\mathcal{U}$ fd-separates $\mathcal{Z}$ and $\mathcal{W}$ in the DIC-FDG for a given DIC problem, then it also fd-separates $\mathcal{Z}$ and $\mathcal{W}$ in the corresponding network FDG [Thakor et al., 2016, Definition 11]. Therefore, according to Lemma 4 in Thakor et al. [2016], we conclude that the random variables denoted by $\mathcal{Z}$ and $\mathcal{W}$ are conditionally independent given the random variables denoted by $\mathcal{U}$, i.e.,

$$I(\mathcal{Z}; \mathcal{W} | \mathcal{U}) = 0, \quad \text{if } \mathcal{U} \text{ fd-separates } \mathcal{Z} \text{ and } \mathcal{W} \text{ in the DIC-FDG.} \tag{B.127}$$

Now we can state the following proposition.

Proposition B.1. For any DIC problem and two disjoint nonempty subsets $K, K' \subset [n]$, set $L = [n] \setminus (K \cup K')$. Then, we have

$$I(X_K; X_{K'} | X_L, Y_P) = 0, \tag{B.128}$$

for any subset of servers $P \subseteq (N \setminus T_{K,K'})$.

Proof. Set $\mathcal{U} = X_L \cup Y_P$, $\mathcal{Z} = X_K$, $\mathcal{W} = X_{K'}$. Since $P \subseteq N \setminus T_{K,K'} = (T_{K,\overline{K'}} \cup T_{K',\overline{K}} \cup T_{\overline{K},\overline{K'}})$, according to the touch structure in Definition 2.2, we know that in the ancestral graph $\mathcal{G}_{\mathrm{An}(\mathcal{U}\cup\mathcal{W}\cup\mathcal{Z})}$, for any $X_i \in \mathcal{Z}, X_j \in \mathcal{W}$, vertices X_i and X_j are either disconnected, or connected with at least one vertex from $X_L \subseteq \mathcal{U}$ in the path between them.

After removing all edges outgoing from vertices in $\mathcal{U}$, any connected vertices $X_i \in \mathcal{Z}$ and $X_j \in \mathcal{W}$ become disconnected. Hence, we can conclude that the vertex set $\mathcal{U}$ fd-separates $\mathcal{Z}$ and $\mathcal{W}$ in the DIC-FDG. Therefore, from (B.127), we conclude that $I(X_K; X_{K'}|X_L, Y_P) = I(\mathcal{Z}; \mathcal{W}|\mathcal{U}) = 0$. □

B.6 Proof of Theorem 4.8

For easier reference, we repeat (4.45) here for a given DIC problem $\mathcal{G}$: $(i|A_i), i \in [n]$ with peripheral set U and an augmentation group $\mathbf{V} = (V_1, \cdots, V_k)$,

$$\sum_{i\in[n]} R_i \leq \sum_{J\in N} C_J + \sum_{\ell\in[k]} \sum_{J\in\mathcal{T}_\ell} C_J, \tag{B.129}$$

where $\mathcal{T}_\ell = T_{V_\ell,(\bigcup_{j\in[\ell+1:k]} V_j)\cup W}$ and $W = [n] \setminus U \setminus (\bigcup_{j\in[k]} V_j)$. Note that $[k+1:k]$ simply means $\emptyset$.

In the following, we prove the proposition by showing that (B.129) is implied by the rate constraint inequality (4.22) as well as the Axioms (4.23)-(4.29) of Theorem 4.7 with the specific server grouping $\mathcal{P}_{\mathbf{V}} = \{T_{V_1}, T_{V_2}, \cdots, T_{V_k}, T_{(\bigcup_{j\in[k]} V_j)^c}\}$ defined in Definition 4.13.

For any $i \in [n]$, according to (4.22) as well as Axiom (4.27), we have

$$\begin{aligned} R_i &\leq f([m], B_i \cup \{i\}) - f([m], B_i) \\ &\leq f([m], B \cup \{i\}) - f([m], B), \qquad \forall B \subseteq B_i, i \in [n]. \end{aligned} \tag{B.130}$$

Consider any two disjoint sets $L, K \subseteq [n]$. If L is an augmentation set of K, then by Definition 4.10, there exists an ordering $\{i_1, i_2, \cdots, i_{|L|}\}$ of the elements in L such that $A_{i_j} \subseteq \{i_1, \cdots, i_{j-1}\} \cup K, j \in [|L|]$, which indicates that

$$(\{i_1, \cdots, i_{j-1}\} \cup K)^c \subseteq B_{i_j}, \qquad \forall j \in [|L|]. \tag{B.131}$$

Hence, according to (B.130) and (B.131), we have

$$\begin{aligned}\sum_{i\in L} R_i = \sum_{j\in[|L|]} R_{i_j} &\le f([m],K^c) - f([m],K^c\setminus\{i_1\})\\ &\quad + f([m],K^c\setminus\{i_1\}) - f([m],K^c\setminus\{i_1\}\setminus\{i_2\})\\ &\quad + f([m],K^c\setminus\{i_1\}\setminus\{i_2\}) - f([m],K^c\setminus\{i_1\}\setminus\{i_2\}\setminus\{i_3\}) + \cdots\\ &\quad + f([m],K^c\setminus\{i_1\}\setminus\cdots\setminus\{i_{|L|-1}\}) - f([m],K^c\setminus L)\\ &= f([m],K^c) - f([m],K^c\setminus L).\end{aligned}$$

By Definitions 4.11 and 4.12, U is an augmentation set of $\emptyset$, and V_j for any $j\in[k]$ is an augmentation set of V_j^c, and W is an augmentation set of W^c. Therefore, we have

$$\sum_{i\in U} R_i \le f([m],\emptyset^c) - f([m],\emptyset^c\setminus U) = f([m],[n]) - f([m],[n]\setminus U), \tag{B.132}$$

$$\sum_{i\in V_j} R_i \le f([m],(V_j^c)^c) - f([m],(V_j^c)^c\setminus V_j) = f([m],V_j), \qquad \forall j\in[k], \tag{B.133}$$

$$\sum_{i\in W} R_i \le f([m],(W^c)^c) - f([m],(W^c)^c\setminus W) = f([m],W). \tag{B.134}$$

As $U, V_1, \cdots, V_k, W$ are disjoint to each other and $[n] = U \cup (\bigcup_{j\in[k]} V_j) \cup W$, combining (B.132) and (B.134), as well as (B.133) for every $j\in[k]$, we have

$$\begin{aligned}\sum_{i\in[n]} R_i &= \sum_{i\in U} R_i + \sum_{j\in[k]}\sum_{i\in V_j} R_i + \sum_{i\in W} R_i\\ &\le f([m],[n]) - f([m],[n]\setminus U) + \sum_{j\in[k]} f([m],V_j) + f([m],W)\\ &\le f([m],[n]) + \sum_{j\in[k]} f(\{j\},V_j) + f([m],W) - f([m],(\bigcup_{j\in[k]} V_j)\cup W),\end{aligned}$$

where the last inequality is due to Axiom (4.23) of Theorem 4.7. According to Axiom (4.25) of Theorem 4.7, we have

$$f([m],[n]) + \sum_{\ell\in[k]} f(\{\ell\},(\bigcup_{j\in[\ell+1:k]} V_j)\cup W) \le \sum_{J\in N} C_J + \sum_{\ell\in[k]}\sum_{J\in\mathcal{T}_\ell} C_J.$$

Therefore, to complete the proof of the proposition, it suffices to show that

$$\begin{aligned}&f([m],[n]) + \sum_{\ell\in[k]} f(\{\ell\},(\bigcup_{j\in[\ell+1:k]} V_j)\cup W)\\ &\ge f([m],[n]) + \sum_{j\in[k]} f(\{j\},V_j) + f([m],W) - f([m],(\bigcup_{j\in[k]} V_j)\cup W)\end{aligned} \tag{B.135}$$

is implied by the rate constraint inequality (4.22) as well as Axioms (4.23)-(4.29) of Theorem 4.7 with $\mathcal{P}_{\mathbf{V}}$ from Definition 4.13.

By Axioms (4.27) and (4.26), for any $\ell \in [k]$, we have

$$\begin{aligned} &f([m],(\bigcup_{j\in[\ell:k]} V_j)\cup W) + f(\{\ell\},(\bigcup_{j\in[\ell+1:k]} V_j)\cup W) \\ &\geq f([m],(\bigcup_{j\in[\ell+1:k]} V_j)\cup W) + f(\{\ell\},V_\ell). \end{aligned} \tag{B.136}$$

Summing both sides of (B.136) for every $\ell \in [k]$, we have

$$\begin{aligned} &\sum_{\ell\in[k]} f([m],(\bigcup_{j\in[\ell:k]} V_j)\cup W) + \sum_{\ell\in[k]} f(\{\ell\},(\bigcup_{j\in[\ell+1:k]} V_j)\cup W) \\ &\geq \sum_{\ell\in[k]} f([m],(\bigcup_{j\in[\ell+1:k]} V_j)\cup W) + \sum_{\ell\in[k]} f(\{\ell\},V_\ell) \\ &= \sum_{\ell\in[2:k+1]} f([m],(\bigcup_{j\in[\ell:k]} V_j)\cup W) + \sum_{\ell\in[k]} f(\{\ell\},V_\ell). \end{aligned} \tag{B.137}$$

Note that

$$\sum_{\ell\in[1]} f([m],(\bigcup_{j\in[\ell:k]} V_j)\cup W) = f([m],(\bigcup_{j\in[k]} V_j)\cup W)$$

and

$$\sum_{\ell\in[k+1:k+1]} f([m],(\bigcup_{j\in[\ell:k]} V_j)\cup W) = f([m],W).$$

Using the above relations, we have

$$\begin{aligned} &f([m],(\bigcup_{j\in[k]} V_j)\cup W) + \sum_{\ell\in[k]} f(\{\ell\},(\bigcup_{j\in[\ell+1:k]} V_j)\cup W) \\ &= \sum_{\ell\in[k]} f([m],(\bigcup_{j\in[\ell:k]} V_j)\cup W) - \sum_{\ell\in[2:k]} f([m],(\bigcup_{j\in[\ell:k]} V_j)\cup W) \\ &\quad + \sum_{\ell\in[k]} f(\{\ell\},(\bigcup_{j\in[\ell+1:k]} V_j)\cup W) \\ &\geq \sum_{\ell\in[2:k+1]} f([m],(\bigcup_{j\in[\ell:k]} V_j)\cup W) - \sum_{\ell\in[2:k]} f([m],(\bigcup_{j\in[\ell:k]} V_j)\cup W) + \sum_{\ell\in[k]} f(\{\ell\},V_\ell) \\ &= f([m],W) + \sum_{\ell\in[k]} f(\{\ell\},V_\ell), \end{aligned}$$

where the inequality follows from (B.137). This completes the proof of (B.135) being implied by the rate constraint inequality (4.22) as well as Axioms (4.23)-(4.29) of Theorem 4.7 with $\mathcal{P}_{\mathbf{V}}$, and thus completes the proof of this proposition.

B.7 Remainder of Proof of Corollary 4.2

Recall that as specified in Remark 4.8, the following proof is based on the assumption that all the servers $J \in N$ are active.

We finish the proof of Corollary 4.2 by showing that Axiom (4.23) in Theorem 4.7 simplifies to Axiom (4.51). For easier reference, we repeat Axiom (4.23), with $\mathcal{P}_{\text{t}} = \{T_{L_1}, T_{L_2}, \dots, T_{L_m}\}$, as follows.

$$f(G,K) = f(G',K), \qquad \text{if } ((T_{L_G} \cup T_{L_{G'}}) \setminus (T_{L_G} \cap T_{L_{G'}})) \subseteq T_{\overline{K}}.$$

We also repeat Axiom (4.51) as follows.

$$f(G,K) = f(G',K), \qquad K \subseteq (L_G \cap L_{G'}).$$

For brevity, set $L = L_G$, $L' = L_{G'}$, and also set $T = (T_L \cup T_{L'}) \setminus (T_L \cap T_{L'}) = ((T_{L_G} \cup T_{L_{G'}}) \setminus (T_{L_G} \cap T_{L_{G'}}))$. For any $K \subseteq [n]$, $G, G' \subseteq [m]$, we are going to show that $K \subseteq (L \cap L')$ is the sufficient and necessary condition for $T \subseteq T_{\overline{K}}$. If $G = G'$, then both (4.23) and (4.51) becomes trivial. Hence we only consider the case when $G \neq G'$, and since $L_1, \cdots, L_m$ are disjoint to each other, we have $L \neq L'$.

First, to show the sufficiency, we assume that $K \subseteq (L \cap L')$. Consider any $J \in T$, we know that J touches either L or L', but not both. As $K \subseteq (L \cap L')$, we know that $J \cap K = \emptyset$, which means that $J \in T_{\overline{K}}$. Therefore, we have $T \subseteq T_{\overline{K}}$, which proves the sufficient condition.

Second, to show the necessity, we assume that $T \subseteq T_{\overline{K}}$. Since $L \neq L'$, without loss of generality, assume there exists some $j_1 \in L \setminus L'$. Now we show that $K \subseteq (L \cap L')$ by contradiction.

Assume that $K \setminus L' \neq \emptyset$, then there exists some $j_2 \in K \setminus L'$. Note that j_1, j_2 may be the same index. Now set $J = \{j_1\} \cup \{j_2\} \in N$. Then we have

$$J \cap L \neq \emptyset, \quad J \cap L' = \emptyset, \quad J \cap K \neq \emptyset. \tag{B.138}$$

Hence, we have $J \in T$ (since $J \in (T_L \cup T_{L'}), J \notin (T_L \cap T_{L'})$), and also $J \notin T_{\overline{K}}$. This contradicts with the assumption that $T \subseteq T_{\overline{K}}$. And therefore, we must have $K \setminus L' = \emptyset$, i.e., $K \subseteq L'$.

Now assume that $K \setminus L \neq \emptyset$, then there exists some $j_3 \in K \setminus L$. Since $j_3 \in K \setminus L$ and $K \subseteq L'$, for $\{j_3\} \in N$, we have

$$\{j_3\} \cap L = \emptyset, \quad \{j_3\} \cap L' \neq \emptyset, \quad \{j_3\} \cap K \neq \emptyset. \tag{B.139}$$

Hence, we have $\{j_3\} \in T$ (since $\{j_3\} \in (T_L \cup T_{L'}), \{j_3\} \notin (T_L \cap T_{L'})$), and also $\{j_3\} \notin T_{\overline{K}}$. This contradicts with the assumption that $T \subseteq T_{\overline{K}}$. And therefore, we must have $K \setminus L = \emptyset$, i.e., $K \subseteq L$.

Now we have both $K \subseteq L'$ and $K \subseteq L$, which means that $K \subseteq (L \cap L')$ and proves the necessary condition. Therefore, Axioms (4.23) and (4.51) are the same under the touch grouping. This completes the proof.

B.8 Proof of Proposition 4.3

For a given DIC problem, consider two valid server groupings $\mathcal{Q} = \{Q_1, \cdots, Q_\ell\}$ and $\mathcal{P} = \{P_1, \cdots, P_m\}$ such that $\mathcal{P}$ is a refinement of $\mathcal{Q}$. For any $E \subseteq [\ell]$, let $Q_E = \bigcup_{i \in E} Q_i$, and for any $G \subseteq [m]$, let $P_G = \bigcup_{i \in G} P_i$.

According to Definition 4.17, for any $i \in [\ell]$, there exists some set $G \subseteq [m]$ such that $P_G = Q_i$. Define the mapping function G that maps any set $E \subseteq [\ell]$ to a corresponding set $G(E) \subseteq [m]$ as follows.

$$G(E) = \bigcup_{i \in E} \bigcup_{G \subseteq [m]: P_G = Q_i} G. \tag{B.140}$$

Then for any $E \subseteq [\ell]$, we have

$$P_{G(E)} = \bigcup_{i \in E} \bigcup_{G \subseteq [m]: P_G = Q_i} P_G = \bigcup_{i \in E} Q_i = Q_E. \tag{B.141}$$

It can also be verified that the mapping function G has following properties.

$$G(\emptyset) = \emptyset. \tag{B.142}$$

$$G(E) \subseteq G(E'), \qquad \forall E \subseteq E' \subseteq [\ell]. \tag{B.143}$$

$$G(E \cup E') = G(E) \cup G(E'), \qquad \forall E, E' \subseteq [\ell]. \tag{B.144}$$

$$G(E \cap E') \subseteq G(E) \cap G(E'), \qquad \forall E, E' \subseteq [\ell]. \tag{B.145}$$

Consider any rate tuple $\mathbf{R} = (R_i, i \in [n])$ in $\mathscr{R}_\mathcal{P}$. Then there exists some $f(G, K), G \subseteq [m], K \subseteq [n]$ such that $\mathbf{R}$ and $f(G, K)$ satisfy Axioms (4.23)-(4.29), as well as (4.22), with server grouping $\mathcal{P}$. Construct $f_Q(E, K) = f(G(E), K)$, $E \subseteq [\ell], K \subseteq [n]$. We now show that the rate tuple $\mathbf{R}$ is also in $\mathscr{R}_Q$ by showing that $\mathbf{R}$ and $f_Q(E, K)$ satisfy Axioms (4.23)-(4.29), as well as (4.22) with server grouping $\mathcal{Q}$.

For Axiom (4.23), consider any $E, E' \subseteq [\ell], K \subseteq [n]$ such that $(Q_E \cup Q_{E'}) \setminus (Q_E \cap Q_{E'}) \subseteq T_{\overline{K}}$, we have

$$(P_{G(E)} \cup P_{G(E')}) \setminus (P_{G(E)} \cap P_{G(E')}) = (Q_E \cup Q_{E'}) \setminus (Q_E \cap Q_{E'}) \subseteq T_{\overline{K}},$$

where the first equality is due to (B.141). As $f(G, K)$ satisfies Axiom (4.23) with server grouping $\mathcal{P}$, we have $f(G(E),K) = f(G(E'),K)$. Therefore, by the construction of $f_Q(E,K)$, we have $f_Q(E,K) = f(G(E),K) = f(G(E'),K) = f_Q(E',K)$.

For Axiom (4.24), it is clear that for any $E \subseteq [\ell], K \subseteq [n]$, due to (B.142) as well as $f(G,K)$ satisfying Axiom (4.24), we have $f_Q(\emptyset,K) = f(G(\emptyset),K) = f(\emptyset,K) = 0$ and $f_Q(E,\emptyset) = f(G(E),\emptyset) = 0$.

For Axiom (4.25), for any $E \subseteq [\ell], K \subseteq [n]$, due to (B.141) as well as $f(G,K)$ satisfying Axiom (4.25), we have

$$f_Q(E,K) = f(G(E),K) \leq \sum_{J:J\in P_{G(E)}, J\in T_K} C_J = \sum_{J:J\in Q_E, J\in T_K} C_J.$$

For Axiom (4.26), for any $E \subseteq E' \subseteq [\ell], K \subseteq K' \subseteq [n]$, we have

$$f_Q(E,K) = f(G(E),K) \leq f(G(E'),K') = f_Q(E',K'),$$

where the inequality is due to (B.143) and $f(G,K)$ satisfying Axiom (4.26).

For Axiom (4.27), for any $E, E' \subseteq [\ell], K, K' \subseteq [n]$, we have

$$\begin{aligned}
&f_Q(E\cup E', K\cap K') + f_Q(E\cap E', K\cup K') \\
&= f(G(E\cup E'), K\cap K') + f(G(E\cap E'), K\cup K') \\
&\leq f(G(E)\cup G(E'), K\cap K') + f(G(E)\cap G(E'), K\cup K') \\
&\leq f(G(E),K) + f(G(E'),K') \\
&= f_Q(E,K) + f_Q(E',K'),
\end{aligned}$$

where the first inequality is due to (B.144) and (B.145) and $f(G,K)$ satisfying Axiom (4.26). The second inequality is due to $f(G,K)$ satisfying Axiom (4.27).

For Axiom (4.28), for any $K \subseteq [n]$, according to (B.141), we have

$$\begin{aligned}
(P_{[m]} \cup P_{G([\ell])}) \setminus (P_{[m]} \cap P_{G([\ell])}) &= (P_{[m]} \cup Q_{[\ell]}) \setminus (P_{[m]} \cap Q_{[\ell]}) \\
&= (N \cup N) \setminus (N \cap N) = \emptyset \subseteq T_{\overline{K}}.
\end{aligned}$$

Therefore, for any $i \in [n]$, as $f(G,K)$ satisfies Axioms (4.23) and (4.28), we have

$$\begin{aligned}
f_Q([\ell], B_i\cup\{i\}) - f_Q([\ell], B_i) &= f(G([\ell]), B_i\cup\{i\}) - f(G([\ell]), B_i) \\
&= f([m], B_i\cup\{i\}) - f([m], B_i) \qquad \text{(B.146)} \\
&= f([m],\{i\}) = f(G([\ell]),\{i\}) = f_Q([\ell],\{i\}).
\end{aligned}$$

For Axiom (4.29), for any $E \subseteq [\ell], K, K' \subseteq [n]$ such that $Q_E \subseteq (N \setminus T_{K,K'})$, according to (B.141), we have $P_{G(E)} = Q_E \subseteq (N \setminus T_{K,K'})$. As $f(G,K)$ satisfies Axiom (4.29) with server grouping $\mathcal{P}$, we have

$$\begin{aligned} f_Q(E,K) + f_Q(E,K') &= f(G(E),K) + f(G(E),K') \\ &= f(G(E),K \cup K') = f_Q(E,K \cup K'). \end{aligned}$$

Finally, given (B.146) and the fact that $f(G,K)$ and **R** jointly satisfy (4.22), for any $i \in [n]$, we have

$$f_Q([\ell],B_i \cup \{i\}) - f_Q([\ell],B_i) = f([m],B_i \cup \{i\}) - f([m],B_i) \geq R_i,$$

which finishes the proof that $f_Q(E,K)$ and **R** jointly satisfy (4.22).

So far we have shown that for any **R** in rate region $\mathscr{R}_{\mathcal{P}}$, it must be also in $\mathscr{R}_Q$. Therefore, $\mathscr{R}_{\mathcal{P}} \subseteq \mathscr{R}_Q$.

B.9 Proof of Proposition 4.4

We show that $\mathscr{R}_{\mathcal{P}_*}(\mathcal{G},\mathbf{C}) \subseteq \mathscr{R}_{\text{ADPM}}(\mathcal{G},\mathbf{C})$ as follows.

Consider any rate tuple $\mathbf{R} \in \mathscr{R}_{\mathcal{P}_*}(\mathcal{G},\mathbf{C})$. Then, there exists a set function $f_*(K), K \subseteq [n]$ such that $f_*(K)$ satisfies Axioms (4.68)-(4.72) and **R** and $f_*(K)$ jointly satisfy (4.67). Construct $f_L(S), S \subseteq L \subseteq [n]$ as $f_L(S) = f_*(S)$.

It can be verified with relative ease that $f_L(S)$ satisfies all the axioms for Proposition 2.9 (Axioms (2.45)-(2.48)). So it remains to show that **R** and $f_L(S)$ jointly satisfy (2.44) as follows.

For any $L \subseteq [n]$ and $i \in L$, we have

$$\begin{aligned} R_i &\leq f_*(B_i \cup \{i\}) - f_*(B_i) \\ &\leq f_*((B_i \cup \{i\}) \cap L) - f_*(B_i \cap L) \\ &= f_L((B_i \cup \{i\}) \cap L) - f_L(B_i \cap L). \end{aligned}$$

where the first inequality is due to (4.67), and the second inequality is due to the fact that $B_i \cup ((B_i \cup \{i\}) \cap L) = B_i \cup \{i\}, B_i \cap ((B_i \cup \{i\}) \cap L) = B_i \cap L$ and that $f_*(K)$ satisfies the submodularity axiom, Axiom (4.71).

Therefore, we can conclude that $\mathbf{R} \in \mathscr{R}_{\text{ADPM}}(\mathcal{G},\mathbf{C})$ and thus $\mathscr{R}_{\mathcal{P}_*}(\mathcal{G},\mathbf{C}) \subseteq \mathscr{R}_{\text{ADPM}}(\mathcal{G},\mathbf{C})$.

Bibliography

AGARWAL, A. AND MAZUMDAR, A., 2016. Local partial clique and cycle covers for index coding. In *Proc. IEEE Global Commun. Conf. Workshop on Netw. Coding and Appl.*, 1–6. Washington, DC. (cited on page 4)

AHLSWEDE, R.; CAI, N.; LI, S.-Y.; AND YEUNG, R. W., 2000. Network information flow. *IEEE Trans. Inf. Theory*, 46, 4 (2000), 1204–1216. (cited on page 7)

ALON, N.; LUBETZKY, E.; STAV, U.; WEINSTEIN, A.; AND HASSIDIM, A., 2008. Broadcasting with side information. In *Proc. 49th Ann. IEEE Symp. Found. Comput. Sci. (FOCS)*, 823–832. Philadelphia, PA. (cited on pages 5, 9, 131, 135, 141, and 148)

ARBABJOLFAEI, F., 2017. *Index coding: Fundamental limits, coding schemes, and structural properties*. Ph.D. thesis, UC San Diego. (cited on page 44)

ARBABJOLFAEI, F.; BANDEMER, B.; AND KIM, Y.-H., 2014. Index coding via random coding. In *Iran Workshop on Communication and Information Theory (IWCIT)*. Tehran, Iran. (cited on pages 5 and 52)

ARBABJOLFAEI, F.; BANDEMER, B.; KIM, Y.-H.; SASOGLU, E.; AND WANG, L., 2013. On the capacity region for index coding. In *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, 962–966. (cited on pages 5, 6, 10, 27, 28, 29, 30, 31, 34, 37, 50, 75, 77, 131, and 145)

ARBABJOLFAEI, F. AND KIM, Y.-H., 2014. Local time sharing for index coding. In *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, 286–290. Honolulu, HI. (cited on pages 4, 13, 26, 27, 31, 33, 117, and 119)

ARBABJOLFAEI, F. AND KIM, Y.-H., 2015a. On critical index coding problems. In *Proc. IEEE Inf. Theory Workshop (ITW)*, 9–13. Jeju Island, Korea. (cited on page 7)

ARBABJOLFAEI, F. AND KIM, Y.-H., 2015b. Three stories on a two-sided coin: Index coding, locally recoverable distributed storage, and guessing games on graphs. In *Proc. 53th Ann. Allerton Conf. Comm. Control Comput.*, 843–850. (cited on page 8)

ARBABJOLFAEI, F. AND KIM, Y.-H., 2016. Approximate capacity of index coding for some classes of graphs. In *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, 2154–2158. Barcelona, Spain. (cited on page 8)

ARBABJOLFAEI, F. AND KIM, Y.-H., 2018. Fundamentals of index coding. *Foundations and Trends® in Communications and Information Theory*, 14, 3-4 (2018), 163–346. (cited on pages 4, 5, 6, 9, 18, 22, 28, 29, 30, 31, 46, 52, 83, 90, 92, 122, and 142)

ARBABJOLFAEI, F. AND KIM, Y.-H., 2020. Generalized lexicographic products and the index coding capacity. *IEEE Trans. Inf. Theory*, 66, 3 (2020), 1520–1529. (cited on page 7)

ARUNACHALA, C.; AGGARWAL, V.; AND RAJAN, B. S., 2019. On the optimal broadcast rate of the two-sender unicast index coding problem with fully-participated interactions. *IEEE Trans. Commun.*, 67, 12 (2019), 8612–8623. (cited on page 7)

ASADI, B.; ONG, L.; AND JOHNSON, S. J., 2015. Optimal coding schemes for the three-receiver awgn broadcast channel with receiver message side information. *IEEE Trans. Inf. Theory*, 61, 10 (2015), 5490–5503. (cited on page 8)

BABER, R.; CHRISTOFIDES, D.; DANG, A. N.; RIIS, S.; AND VAUGHAN, E. R., 2013. Multiple unicasts, graph guessing games, and non-shannon inequalities. In *Proc. Int. Symp. Netw. Coding.*, 1–6. Calgary, AB. (cited on pages 6, 34, and 147)

BAR-YOSSEF, Z.; BIRK, Y.; JAYRAM, T.; AND KOL, T., 2011. Index coding with side information. *IEEE Trans. Inf. Theory*, 57 (2011), 1479–1494. (cited on pages 4, 7, 11, 16, 31, 47, 77, 125, and 146)

BIRK, Y. AND KOL, T., 1998. Informed-source coding-on-demand (ISCOD) over broadcast channels. In *Proc. 17th Ann. IEEE Int. Conf. Comput. Commun. (INFOCOM)*, 1257–1264. San Francisco, CA. (cited on pages 1, 3, 4, and 26)

BIRK, Y. AND KOL, T., 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 (2006), 2825–2830. (cited on pages 1 and 4)

BLASIAK, A.; KLEINBERG, R.; AND LUBETZKY, E., 2011. Lexicographic products and the power of non-linear network coding. In *Proc. 52th Ann. IEEE Symp. Found. Comput. Sci. (FOCS)*, 609–618. Palm Springs, CA. (cited on pages 5, 6, 9, 11, 12, 33, 46, 77, 83, 92, 98, 116, 122, 131, 139, 146, and 148)

BLASIAK, A.; KLEINBERG, R.; AND LUBETZKY, E., 2013. Broadcasting with side information: Bounding and approximating the broadcast rate. *IEEE Trans. Inf. Theory*, 59, 9 (2013), 5811–5823. (cited on pages 4, 8, 32, and 156)

BRAHMA, S. AND FRAGOULI, C., 2015. Pliable index coding. *IEEE Trans. Inf. Theory*, 61, 11 (2015), 6192–6203. (cited on page 8)

Byrne, E. and Calderini, M., 2017. Error correction for index coding with coded side information. *IEEE Trans. Inf. Theory*, 63, 6 (2017), 3712–3728. (cited on page 9)

Chan, T. and Grant, A., 2010. On capacity regions of non-multicast networks. In *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, 2378–2382. Austin, TX. (cited on page 18)

Chaudhry, M. A. R.; Asad, Z.; Sprintson, A.; and Langberg, M., 2011. On the complementary index coding problem. In *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, 244–248. Saint Petersburg, Russia. (cited on page 4)

Cover, T. M., 2006. *Elements of information theory*. John Wiley & Sons. (cited on pages 23 and 98)

Dai, M.; Shum, K. W.; and Sung, C. W., 2014. Data dissemination with side information and feedback. *IEEE Trans. Wireless Commun.*, 13, 9 (2014), 4708–4720. (cited on page 9)

Dau, S. H.; Skachek, V.; and Chee, Y. M., 2012. On the security of index coding with side information. *IEEE Trans. Inf. Theory*, 58, 6 (2012), 3975–3988. (cited on pages 8 and 117)

Dau, S. H.; Skachek, V.; and Chee, Y. M., 2013. Error correction for index coding with side information. *IEEE Transactions on Information Theory*, 59, 3 (2013), 1517–1531. (cited on page 8)

Ding, N. and Sadeghi, P., 2019. A submodularity-based clustering algorithm for the information bottleneck and privacy funnel. In *Proc. IEEE Inf. Theory Workshop (ITW)*, 1–5. Visby, Sweden. (cited on page 13)

Effros, M.; El Rouayheb, S.; and Langberg, M., 2015. An equivalence between network coding and index coding. *IEEE Trans. Inf. Theory*, 61, 5 (2015), 2478–2487. (cited on pages 7 and 8)

El Gamal, A. and Kim, Y.-H., 2011. *Network Information Theory*. Cambridge: Cambridge University Press. (cited on page 27)

El Rouayheb, S.; Sprintson, A.; and Georghiades, C., 2010. On the index coding problem and its relation to network coding and matroid theory. *IEEE Trans. Inf. Theory*, 56, 7 (2010), 3187–3195. (cited on page 9)

El Rouayheb, S. Y.; Chaudhry, M. A. R.; and Sprintson, A., 2007. On the minimum number of transmissions in single-hop wireless coding networks. In *Proc. IEEE Inf. Theory Workshop (ITW)*, 120–125. Tahoe City, CA. (cited on page 4)

ESFAHANIZADEH, H.; LAHOUTI, F.; AND HASSIBI, B., 2014. A matrix completion approach to linear index coding problem. In *Proc. IEEE Inf. Theory Workshop (ITW)*, 531–535. Hobart, Australia. (cited on page 9)

FEKETE, M., 1923. Uber die verteilung der wurzeln bei gewissen algebraischen gleichungen mit ganzzahligen koeffizienten. *Math. Z.*, 17, 1 (1923), 228–249. (cited on page 17)

GATTEGNO, I. B.; GOLDFELD, Z.; AND PERMUTER, H. H., 2015. FME-IT package for MATLAB. (cited on pages 46 and 73)

HAMMACK, R.; IMRICH, W.; AND KLAVŽAR, S., 2011. *Handbook of product graphs*. CRC press. (cited on page 22)

HAVIV, I., 2020. Task-based solutions to embedded index coding. *IEEE Trans. Inf. Theory*, (2020). (cited on page 9)

HAVIV, I. AND LANGBERG, M., 2012. On linear index coding for random graphs. In *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, 2231–2235. Cambridge, MA. (cited on page 9)

HSU, Y.-P.; HOU, I.-H.; AND SPRINTSON, A., 2020. Joint index coding and incentive design for selfish clients. *arXiv preprint arXiv:2005.08545*, (2020). (cited on page 9)

HUANG, Y.-C., 2017. Lattice index codes from algebraic number fields. *IEEE Trans. Inf. Theory*, 63, 4 (2017), 2098–2112. (cited on page 8)

HUANG, Y.-C.; HONG, Y.; AND VITERBO, E., 2017. Golden-coded index coding. In *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, 2548–2552. Aachen, Germany. (cited on page 9)

ISSA, I.; WAGNER, A. B.; AND KAMATH, S., 2020. An operational approach to information leakage. *IEEE Trans. Inf. Theory*, 66, 3 (2020), 1625–1657. (cited on pages 131, 132, and 133)

JAFAR, S. A., 2014. Topological interference management through index coding. *IEEE Trans. Inf. Theory*, 60, 1 (2014), 529–568. (cited on pages 4, 7, 11, 31, 32, 77, and 146)

KAO, D. T.; MADDAH-ALI, M. A.; AND AVESTIMEHR, A. S., 2016. Blind index coding. *IEEE Trans. Inf. Theory*, 63, 4 (2016), 2076–2097. (cited on page 9)

KARAT, N. S.; SAMUEL, S.; AND RAJAN, B. S., 2018. Optimal error correcting index codes for some generalized index coding problems. *IEEE Trans. Commun.*, 67, 2 (2018), 929–942. (cited on page 8)

KARMOOSE, M.; SONG, L.; CARDONE, M.; AND FRAGOULI, C., 2019. Privacy in index coding: k-limited-access schemes. *IEEE Trans. Inf. Theory*, (2019). (cited on page 8)

KIM, J.-W. AND NO, J.-S., 2017. Index coding with erroneous side information. *IEEE Trans. Inf. Theory*, 63, 12 (2017), 7687–7697. (cited on page 9)

KIM, J.-W. AND NO, J.-S., 2019. Linear index coding with multiple senders and extension to a cellular network. *IEEE Trans. Commun.*, 67, 12 (Dec. 2019), 8666–8677. (cited on pages 6, 7, and 30)

KRAMER, G., 1998. *Directed Information for Channels With Feedback.* Ph.D. thesis, ETH Series in Information Processing. (cited on pages 100, 190, and 191)

LANGBERG, M. AND EFFROS, M., 2011. Network coding: Is zero error always possible? In *Proc. 49th Ann. Allerton Conf. Comm. Control Comput.*, 1478–1485. (cited on pages 18 and 154)

LEE, N.; DIMAKIS, A. G.; AND HEATH, R. W., 2015. Index coding with coded side-information. *IEEE Comm. Lett.*, 19, 3 (2015), 319–322. (cited on page 9)

LI, M.; ONG, L.; AND JOHNSON, S. J., 2017. Improved bounds for multi-sender index coding. In *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, 3060–3064. Aachen, Germany. doi:10.1109/ISIT.2017.8007092. (cited on pages 5, 10, 30, and 65)

LI, M.; ONG, L.; AND JOHNSON, S. J., 2018. Cooperative multi-sender index coding. *IEEE Trans. Inf. Theory*, 65, 3 (2018), 1725–1739. (cited on pages 5, 7, 10, 30, 37, 59, 65, 71, 72, 75, and 146)

LI, M.; ONG, L.; AND JOHNSON, S. J., 2019. Multi-sender index coding for collaborative broadcasting: A rank-minimization approach. *IEEE Trans. Commun.*, 67, 2 (Feb. 2019), 1452–1466. (cited on pages 5, 6, and 30)

LIU, T. AND TUNINETTI, D., 2018. An information theoretic converse for the "consecutive complete–s" picod problem. In *Proc. IEEE Inf. Theory Workshop (ITW)*, 1–5. (cited on page 102)

LIU, T. AND TUNINETTI, D., 2019a. Decentralized pliable index coding. In *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, 532–536. Paris, France. (cited on page 9)

LIU, T. AND TUNINETTI, D., 2019b. Private pliable index coding. In *Proc. IEEE Inf. Theory Workshop (ITW)*. Visby, Sweden. (cited on page 9)

LIU, T. AND TUNINETTI, D., 2019c. Tight information theoretic converse results for some pliable index coding problems. *IEEE Trans. Inf. Theory*, 66, 5 (2019), 2642–2657. (cited on page 8)

LIU, T. AND TUNINETTI, D., 2020. Secure decentralized pliable index coding. *arXiv preprint arXiv:2001.03810*, (2020). (cited on page 9)

LIU, Y.; DING, N.; SADEGHI, P.; AND RAKOTOARIVELO, T., 2020a. Privacy-utility tradeoff in a guessing framework inspired by index coding. *arXiv preprint arXiv:2001.06828*, (2020). (cited on page 13)

LIU, Y.; KIM, Y.-H.; VELLAMBI, B.; AND SADEGHI, P., 2018a. On the capacity region for secure index coding. In *Proc. IEEE Inf. Theory Workshop (ITW)*. Guanzhou, China. https://arxiv.org/abs/1809.03615. (cited on pages 6, 8, 34, 101, 111, 117, 119, and 155)

LIU, Y. AND SADEGHI, P., 2019a. From alignment to acyclic chains: Lexicographic performance bounds for index coding. In *Proc. 57th Ann. Allerton Conf. Comm. Control Comput.*, 260–267. Minticello, IL. (cited on page 12)

LIU, Y. AND SADEGHI, P., 2019b. Generalized alignment chain: Improved converse results for index coding. In *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, 1242–1246. Paris, France. (cited on pages 12, 34, and 147)

LIU, Y.; SADEGHI, P.; ABOUTORAB, N.; AND SHARIFIFAR, A., 2020b. Secure index coding with security constraints on receivers. *arXiv preprint arXiv:2001.07296*, (2020). (cited on page 13)

LIU, Y.; SADEGHI, P.; ARBABJOLFAEI, F.; AND KIM, Y.-H., 2017. On the capacity for distributed index coding. In *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, 3055–3059. Aachen, Germany. (cited on pages 10, 11, 12, and 44)

LIU, Y.; SADEGHI, P.; ARBABJOLFAEI, F.; AND KIM, Y.-H., 2018b. Simplified composite coding for index coding. In *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, 456–460. Vail, CO. (cited on pages 10 and 11)

LIU, Y.; SADEGHI, P.; ARBABJOLFAEI, F.; AND KIM, Y.-H., 2020c. Capacity theorems for distributed index coding. *IEEE Trans. Inf. Theory*, (Mar. 2020). (cited on pages 10, 11, 12, and 113)

LIU, Y.; SADEGHI, P.; AND KIM, Y.-H., 2018c. Three-layer composite coding for index coding. In *Proc. IEEE Inf. Theory Workshop (ITW)*, 170–174. Guangzhou, China. (cited on pages 10 and 11)

LUBETZKY, E. AND STAV, U., 2009. Nonlinear index coding outperforming the linear optimum. *IEEE Trans. Inf. Theory*, 55, 8 (Aug. 2009), 3544–3551. (cited on pages 4 and 5)

MAKHDOUMI, A.; SALAMATIAN, S.; FAWAZ, N.; AND MÉDARD, M., 2014. From the information bottleneck to the privacy funnel. In *Proc. IEEE Inf. Theory Workshop (ITW)*, 501–505. (cited on pages 13, 117, 132, 141, and 148)

MALEKI, H.; CADAMBE, V. R.; AND JAFAR, S. A., 2014. Index coding: An interference alignment perspective. *IEEE Trans. Inf. Theory*, 60, 9 (2014), 5402–5432. (cited on pages 4, 5, 7, 11, 31, 32, 77, 115, 146, and 156)

MAZUMDAR, A., 2014. On a duality between recoverable distributed storage and index coding. In *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, 1977–1981. Honolulu, HI. (cited on page 8)

MOJAHEDIAN, M. M.; AREF, M. R.; AND GOHARI, A., 2017. Perfectly secure index coding. *IEEE Trans. Inf. Theory*, 63, 11 (2017), 7382–7395. (cited on pages 8 and 117)

NARAYANAN, V.; PRABHAKARAN, V. M.; RAVI, J.; MISHRA, V. K.; DEY, B. K.; AND KARAMCHANDANI, N., 2018. Private index coding. In *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, 596–600. Vail, CO. (cited on page 125)

NARAYANAN, V.; RAVI, J.; MISHRA, V. K.; DEY, B. K.; KARAMCHANDANI, N.; AND PRABHAKARAN, V. M., 2020. Private index coding. *arXiv preprint arXiv:2006.00257*, (2020). (cited on pages 8 and 117)

NATARAJAN, L.; HONG, Y.; AND VITERBO, E., 2015a. Index codes for the gaussian broadcast channel using quadrature amplitude modulation. *IEEE Comm. Lett.*, 19, 8 (2015), 1291–1294. (cited on page 8)

NATARAJAN, L.; HONG, Y.; AND VITERBO, E., 2015b. Lattice index coding. *IEEE Trans. Inf. Theory*, 61, 12 (2015), 6505–6525. (cited on page 8)

NEELY, M. J.; TEHRANI, A. S.; AND ZHANG, Z., 2013. Dynamic index coding for wireless broadcast networks. *IEEE Trans. Inf. Theory*, 59, 11 (2013), 7525–7540. (cited on pages 4, 9, and 103)

ONG, L., 2014. Linear codes are optimal for index-coding instances with five or fewer receivers. In *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, 491–495. Honolulu, HI. (cited on page 16)

ONG, L.; HO, C. K.; AND LIM, F., 2016a. The single-uniprior index-coding problem: The single-sender case and the multi-sender extension. *IEEE Trans. Inf. Theory*, 62, 6 (Jun. 2016), 3165–3182. (cited on pages 5 and 30)

ONG, L.; KLIEWER, J.; AND VELLAMBI, B. N., 2018. Secure network-index code equivalence: Extension to non-zero error and leakage. In *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, 841–845. Vail, CO. (cited on pages 8 and 117)

ONG, L.; VELLAMBI, B. N.; AND KLIEWER, J., 2019a. Optimal-rate characterisation for pliable index coding using absent receivers. In *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, 522–526. Paris, France. (cited on page 8)

ONG, L.; VELLAMBI, B. N.; KLIEWER, J.; AND SADEGHI, P., 2019b. Improved lower bounds for pliable index coding using absent receivers. *arXiv preprint arXiv:1909.11850*, (2019). (cited on page 8)

ONG, L.; VELLAMBI, B. N.; KLIEWER, J.; AND YEOH, P. L., 2016b. An equivalence between secure network and index coding. In *Proc. IEEE Global Commun. Conf. Workshop on Netw. Coding and Appl.* Washington, DC. (cited on page 8)

ONG, L.; VELLAMBI, B. N.; YEOH, P. L.; KLIEWER, J.; AND YUAN, J., 2016c. Secure index coding: Existence and construction. In *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, 2834–2838. Barcelona, Spain. (cited on pages 8, 117, and 120)

PORTER, A. AND WOOTTERS, M., 2019. Embedded index coding. *arXiv preprint arXiv:1904.02179*, (2019). (cited on page 8)

RÉNYI, A., 1961. On measures of information and entropy. In *Proc. 4th Berkeley Symp. on Math Statist. Probability*, vol. 1, 547–561. (cited on page 134)

RIIS, S., 2007. Information flows, graphs and their guessing numbers. *Elec. J. Comb.*, 14, R44 (2007), 1–17. (cited on page 8)

SADEGHI, P.; ARBABJOLFAEI, F.; AND KIM, Y.-H., 2016. Distributed index coding. In *Proc. IEEE Inf. Theory Workshop (ITW)*. Cambridge, UK. (cited on pages 5, 7, 18, 30, 34, 65, and 112)

SASI, S. AND RAJAN, B. S., 2019. Code construction for pliable index coding. In *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, 527–531. Paris, France. (cited on page 8)

SHANMUGAM, K. AND DIMAKIS, A. G., 2014. Bounding multiple unicasts through index coding and locally repairable codes. In *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, 296–300. Honolulu, HI. (cited on page 8)

SHANMUGAM, K.; DIMAKIS, A. G.; AND LANGBERG, M., 2013. Local graph coloring and index coding. In *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, 1152–1156. Istanbul, Turkey. (cited on page 4)

SHANMUGAM, K.; DIMAKIS, A. G.; AND LANGBERG, M., 2014. Graph theory versus minimum rank for index coding. In *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, 291–295. Honolulu, HI. (cited on page 9)

SHANNON, C. E., 1948. A mathematical theory of communication. *The Bell system technical journal*, 27, 3 (1948), 379–423. (cited on page 23)

SIBSON, R., 1969. Information radius. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 14, 2 (1969), 149–160. (cited on pages 133 and 137)

SLONIM, N. AND TISHBY, N., 2000. Agglomerative information bottleneck. In *Advances in Neural Information Processing Systems (NIPS)*, 617–623. (cited on pages 13, 117, 132, 141, and 148)

SONG, L. AND FRAGOULI, C., 2017. A polynomial-time algorithm for pliable index coding. *IEEE Trans. Inf. Theory*, 64, 2 (2017), 979–999. (cited on page 8)

SONG, L.; FRAGOULI, C.; AND ZHAO, T., 2019. A pliable index coding approach to data shuffling. *IEEE Trans. Inf. Theory*, (2019). (cited on page 8)

SUN, H. AND JAFAR, S. A., 2015. Index coding capacity: How far can one go with only shannon inequalities? *IEEE Trans. Inf. Theory*, 61, 6 (2015), 3041–3055. (cited on pages 6, 34, and 147)

TAHMASBI, M.; SHAHRASBI, A.; AND GOHARI, A., 2015. Critical graphs in index coding. *IEEE J. Sel. Areas Commun.*, 33, 2 (2015), 225–235. (cited on pages 7 and 18)

TEHRANI, A. S.; DIMAKIS, A. G.; AND NEELY, M. J., 2012. Bipartite index coding. In *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, 2246–2250. Cambridge, MA. (cited on page 9)

THAKOR, S.; GRANT, A.; AND CHAN, T., 2016. Cut-set bounds on network information flow. *IEEE Trans. Inf. Theory*, (2016), 1850–1865. doi:10.1109/TIT.2016.2529643. (cited on pages 100, 190, 191, and 192)

THAPA, C.; ONG, L.; AND JOHNSON, S. J., 2016. Graph-theoretic approaches to two-sender index coding. In *Proc. IEEE Global Commun. Conf.*, 1–6. Washington, DC. (cited on pages 5 and 30)

THAPA, C.; ONG, L.; AND JOHNSON, S. J., 2017. Interlinked cycles for index coding: Generalizing cycles and cliques. *IEEE Trans. Inf. Theory*, 63, 6 (2017), 3692–3711. (cited on pages 4 and 5)

THAPA, C.; ONG, L.; JOHNSON, S. J.; AND LI, M., 2019. Structural characteristics of two-sender index coding. *Entropy*, 21, 6 (2019), 615. (cited on page 7)

THOMAS, A. AND RAJAN, B. S., 2017. Index coding with restricted information (icri) and interference alignment. In *Proc. IEEE Global Commun. Conf.*, 1–7. (cited on page 9)

UNAL, S. AND WAGNER, A. B., 2016. A rate–distortion approach to index coding. *IEEE Trans. Inf. Theory*, 62, 11 (2016), 6359–6378. (cited on page 9)

VERDÚ, S., 2015. α-mutual information. In *Proc. UCSD Inf. Theory and Applications Workshop (ITA)*, 1–6. (cited on pages 133 and 137)

WAN, K.; TUNINETTI, D.; AND PIANTANIDA, P., 2017. A novel index coding scheme and its application to coded caching. In *Proc. UCSD Inf. Theory and Applications Workshop (ITA)*, 1–6. (cited on page 9)

YEUNG, R. W., 2006. *Network coding theory*. Now Publishers Inc. (cited on page 7)

YEUNG, R. W., 2008. *Information theory and network coding*. Springer Science & Business Media. (cited on pages 6 and 7)

YU, H. AND NEELY, M. J., 2014. Duality codes and the integrality gap bound for index coding. *IEEE Trans. Inf. Theory*, 60, 11 (2014), 7256–7268. (cited on page 9)