University of California
Los Angeles

Robust Communication and Optimization over Dynamic Networks

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Electrical Engineering

by

Can Karakus

2018

ABSTRACT OF THE DISSERTATION

Robust Communication and Optimization over Dynamic Networks

by

Can Karakus
Doctor of Philosophy in Electrical Engineering
University of California, Los Angeles, 2018
Professor Suhas N. Diggavi, Chair

Many types of communication and computation networks arising in modern systems have fundamentally dynamic, time-varying, and ultimately unreliably available resources. Specifically, in wireless communication networks, such unreliability may manifest itself as variability in channel conditions, intermittent availability of undedicated resources (such as unlicensed spectrum), or collisions due to multiple-access. In distributed computing, and specifically in large-scale distributed optimization and machine learning, this phenomenon manifests itself in the form of communication bottlenecks, straggling or failed nodes, or running background processes which hamper or slow down the computational task. In this thesis, we develop information-theoretically-motivated approaches that make progress towards building robust and reliable communication and computation networks built upon unreliable resources.

In the first part of the thesis, we focus on three problems in wireless networks which involve opportunistically harnessing time-varying resources while providing theoretical performance guarantees. First, we show that in full-duplex uplink-downlink cellular networks, a simple, low-overhead user scheduling scheme that exploits the variations in channel conditions can be used to optimally mitigate inter-user interference in the many-user regime. Next, we consider the use of intermittently available links over unlicensed spectral bands to enhance communication over the licensed cellular band. We show that channel output feedback over such links, combined with quantize-map-forward relaying, provides generalized-degrees-of-freedom gain in interference networks. We characterize the information-theoretic capacity region of this model to within a constant gap. We finally consider the use of such intermittent links in device-to-device cooperation to aid cellular downlink. We develop an optimal dynamic resource allocation algorithm for such networks using stochastic approximation and graph theory techniques, and show that the resulting scheme results in up to 5-6x throughput gain for cell-edge users.

In the second part, we consider the problem of distributed optimization and machine learning over large-scale, yet unreliable clusters. Focusing on a master-worker architecture, where large-scale datasets are distributed across worker nodes which communicate with a central parameter server to optimize a global objective, we develop a framework for embedding redundancy in the dataset to combat node failures and delays. This framework consists of an efficient linear transformation (coding) of the dataset that results in an overcomplete representation, combined with a coding-oblivious application of a distributed optimization algorithm. We show that if the linear transformation is designed to satisfy certain spectral properties resembling the restricted isometry property, nodes that fail or delay their computation can be dynamically left out of the computational process, while still converging to a reasonable solution with fast convergence rates, obviating the need for explicit fault-tolerance mechanisms and significantly speeding up overall computation. We implement the techniques on Amazon EC2 clusters to demonstrate the applicability of the proposed technique to various machine learning problems, such as logistic regression, support vector machine, ridge regression, and collaborative filtering; as well as several popular optimization algorithms including gradient descent, L-BFGS, coordinate descent and proximal gradient methods.

The dissertation of Can Karakus is approved.

Wotao Yin

Gregory J. Pottie

Richard D. Wesel

Suhas N. Diggavi, Committee Chair

University of California, Los Angeles
2018

To my parents, Cilen and Ilhami…

Table of Contents

List of Figures

List of Tables

Acknowledgments

The past six and a half years, over which I have been working towards my Ph.D., have been an extremely transformative period for me, intellectually, personally, and emotionally. Throughout this period, I have learned not only a great deal of academic knowledge, but also learned about myself, my own strengths and weaknesses, and about how to face challenges and learn from mistakes. I have learned that research is more about asking the right questions and less about finding the right answers, and although ironically in the face of more knowledge I feel more ignorant than ever, I feel confident that during this challenging period, I have acquired the skills to adapt and persevere through future challenges, technical or otherwise.

In guiding me through this challenging yet rewarding process, I am extremely grateful to my advisor, Professor Suhas Diggavi. Since my first day in graduate school, I have continuously felt his kind support, which was critical in more instances throughout my studies than I can count. His open-minded approach to research, his vision to always foster and encourage original ideas, and his insightful advice is what enabled me to navigate through the unknown, and ultimately produce this thesis. I absolutely admire his genuine excitement and passion about research, which is forever imprinted upon me, and I am indebted to him for all that he taught me.

It would be impossible to complete this journey without the insightful input of my collaborators. I am especially grateful to Professor Wotao Yin, who in addition to serving in my doctoral committee, provided me with invaluable guidance in a research area I was not familiar with. His deep insight and expertise on numerical optimization and applied mathematics, as well as his kind personality is something I have always admired, and he has provided key support in the latter part of my PhD. I am also thankful to Professor I-Hsiang Wang, whose supervision in the earlier years of my PhD helped me take my first steps of doctoral research, and I learned a great deal discussing with him, both about information theory, and about research in general. I would also like to thank Dr. Yifan Sun, who helped

me venture into a completely new area of research, by supervising me as an intern, as well as providing crucial support in the experimental part of my work. I am also grateful to Professor Ashutosh Sabharwal for teaching me more about the practical aspects of wireless networks, and for his friendly and cheerful personality.

I would also like to thank Professors Greg Pottie and Rick Wesel for taking the time to serve in my doctoral committee, for their valuable comments, as well as for the inspiring classes they taught me.

I would also like to gratefully acknowledge National Science Foundation and Intel, whose funding supported my doctoral research[1].

I am thankful to the labmates that I worked side-by-side through these years: Nikhil, Shaunak, Jad, Mehrdad, Joyson, Yair, Wei, Ayan, Manikandan, and Cuneyd. I especially appreciate the lab spirit we created and good times we had in the earlier parts of my PhD with Nikhil, Shaunak, and Jad. A special thanks goes to Jad, who has been a very patient roommate to me for more than three years.

I have been very fortunate to meet some amazing friends during these years. I cannot thank Ayca Balkan enough for her enduring and genuine friendship, which I cherish deeply. Countless corners of UCLA and Los Angeles are filled with our memories, adventures, trips, and "bonding experiences" together, and her influence in my life during these years has been huge. I feel very lucky to have met such a great and loyal friend. I am also extremely thankful to Kasra Vakilinia, for being such a consistent source of joy and laughter in my life. I shared some great memories with him in multiple countries, and I look forward to being (almost) neighbors and exchanging insults (!) with him in the near future. An indispensible part of my life during these years was also my friendship with Omeed Paydar. We have been through so much, and I enjoyed every moment we spent together: being gym-mates, biking, partying, and everything else. I am very grateful to all the other great people who have been part of my life through this journey, who influenced me, amazed

[1] The work in this thesis was supported in part by NSF grants #1314937 and #1514531, and a gift from Intel Corp.

me, and inspired me: Iris Safaka, Bilgehan Avser, Marios Fanourakis, Martina Cardone, Evangelia Tsolaki, Anezina Solomonidou, Ryan Dollbaum; my biking friends, Dan Kim, Jason Yamanishi, Jimmy Ouyang, Qing Xu, Antonio Yang; as well as family members, Fatih Hekimoglu, Sultan Ercan, Nilgun Ercan, Birgul Yuksel, Suheyla Bilmez. Thank you all for being in my life.

I would like to specially thank Christina Vlachou, who has been a truly amazing girlfriend, and I am proud of her. We had so many unforgettable moments all around the world, and I am extremely lucky to have her in my life. Her continuous love, support, and *αγκαλίτσες* gave me the confidence to face the difficulties in life. I cannot wait to share more of the exciting moments with her.

Last but not least, I am immeasurably thankful to my parents, Cilen and Ilhami, for their unconditional love and understanding, and for always giving me comfort, regardless of situation, time and place. I know they have been very proud of me for coming this far, but I am also tremendously proud of them for being who they are, and for being such great parents. This thesis is dedicated to you, Mom and Dad.

Vita

2011-2018	Graduate Student Researcher, Electrical Engineering University of California, Los Angeles
Summer 2016	Research Intern Technicolor Research, Los Altos, CA
Summer 2015	Research Intern Qualcomm Research, San Diego, CA
March 2013	M.S., Electrical Engineering University of California, Los Angeles
June 2011	B.S., Electrical and Electronics Engineering Bilkent University, Turkey

Publications

C. Karakus, Y. Sun, S. Diggavi, W. Yin, "Redundancy techniques for straggler mitigation in distributed optimization and learning", *arXiv.org*, 2018.

C. Karakus, Y. Sun, S. Diggavi, W. Yin, "Straggler mitigation in distributed optimization through data encoding", *Advances in Neural Information Processing Systems (NIPS)*, 2017 *(Spotlight presentation)*

C. Karakus, Y. Sun, S. Diggavi, "Encoded distributed optimization", *IEEE International Symposium on Information Theory (ISIT)*, 2017

C. Karakus, S. Diggavi, “Enhancing multiuser MIMO through opportunistic D2D cooperation”, *IEEE Transactions on Wireless Communications*, 2017.

J. Sebastian, C. Karakus, and S. Diggavi, “Approximately achieving interference channel capacity with point-to-point codes”, *IEEE International Symposium on Information Theory (ISIT)*, 2016.

J. Sebastian, C. Karakus, S. Diggavi, and I.-H. Wang “Rate splitting is approximately optimal for fading gaussian interference channels”, *Allerton Conference on Communication, Control and Computing*, 2015.

C. Karakus, I.-H. Wang, and S. Diggavi, “Gaussian interference channel with intermittent feedback”, *IEEE Transactions on Information Theory*, 2015.

C. Karakus, S. Diggavi, “Opportunistic scheduling for full-duplex uplink-downlink networks”, *IEEE International Symposium on Information Theory (ISIT)*, 2015.

C. Karakus, I.-H. Wang, and S. Diggavi, “An achievable rate region for gaussian interference channel with intermittent feedback”, *Proc. Allerton Conference on Communication, Control and Computing*, 2013.

C. Karakus, I.-H. Wang, and S. Diggavi, “Interference channel with intermittent feedback”, *IEEE International Symposium on Information Theory (ISIT)*, 2013.

D. Ghita, C. Karakus, K. Argyraki, and P. Thiran, “Shifting network tomography toward a practical goal”, *ACM International Conference on Emerging Networking Experiments and Technologies (CoNEXT)*, 2011.

CHAPTER 1

Introduction

1.1 Dynamic Networks

Modern communication and computation networks have reached an unprecedented scale, both in terms of their size, and the amount of data they handle per second. Wireless communication networks support an ever-increasing volume of data traffic over billions of nodes, through a multitude of technologies and standards. Computation networks, consisting of large clusters of processors in data centers or networks of mobile devices, process massive amounts of data each second.

Such networks typically exhibit a high degree of dynamism, which can manifest itself in a variety of ways, including time-varying availability of network resources, failures of network components, and changes in network conditions and capacity. In designing any system built upon such time-varying and intermittently available network resources, one needs to explicitly account for such unreliability, since a system designed based on a static network assumption might fail in the face of dynamism.

The most straightforward solution to handle such dynamism is to constantly monitor network conditions, and change system operation whenever underlying network conditions change. For instance, in a wireless network, one can design a different communication strategy for different network conditions, and let the network choose the corresponding strategy based on the current network state. Unfortunately, in many cases, the sheer scale of the network, as well as time constraints might render accurate and timely tracking of global network state infeasible, let alone adapting strategies. Building reliable systems over such dynamic networks instead requires designing communication and computation strategies that are inherently robust to such changes in network conditions, and that are adaptive by design.

In this thesis, we explore such strategies under a variety of scenarios over certain types of communication and computation networks. First, we consider wireless communication networks, and focus on three specific scenarios where we develop and analyze such schemes; namely, interference networks with unreliable feedback, full-duplex uplink-downlink cellular networks, and device-to-device cooperation over downlink cellular networks. Next, we consider distributed optimization and learning over computing networks with master-worker architecture, and develop computing strategies that are robust against changing and unpredictable network failures and bottlenecks.

We next describe in more detail how dynamism manifests itself specifically in wireless networks and distributed computing.

1.1.1 Wireless networks

Variability is an integral part of most wireless networks. In the context of this thesis, it is useful to distinguish between two types of variability: the one in channel states, and the one in link availabilities.

In cellular networks, channel states typically vary in the order of milliseconds, due to effects such as user mobility, changing environment, and effects arising due to the characteristics of the wireless medium, such as multipath fading and shadowing. There is a vast body of research focusing on the design of wireless networks under such variability [TV05]. Although this is a well-researched topic, its effects and implications in emerging types of wireless networks are still fairly unexplored. In this thesis, we focus on one such network, namely full-duplex cellular networks, where the base stations are equipped with full-duplex radios, capable of transmitting and receiving over the same frequency band at the same time. We explore the effect of channel state variability in such networks, and design a communication strategy that specifically exploits the variations in channel state.

The second type of variability is the intermittent availability of certain communication links in the network. Such intermittence arises when undedicated resources are used for communication, such as unlicensed spectra. Since access to such resources are not coordinated, transmissions can face collisions from other networks, and get dropped. In addition, control and back-off mechanisms employed in higher layers of the protocol stack (*e.g.*, CSMA) can also make a certain link momentarily unusable from the physical-layer perspective. In this thesis, we consider schemes that make use of unlicensed bands to enhance communication over the licensed band, either through feedback or cooperation over such bands, while accounting for the intermittent availability of these resources. We show that, by taking a longer-term view of the network, it is possible to design schemes that harness these additional resources to reliably improve communication over the network.

1.1.2 Distributed computing

Large-scale computing networks can suffer from node failures and delays. In computing clusters in data centers, such failures and delays can arise from resources being shared with other applications, running background processes, communication bottlenecks on the network, or power limits [DB13]. In computations involving a large number of nodes, a small fraction of such slow nodes, called stragglers, can significantly slow down the overall computation.

Specifically in distributed optimization and learning, databases are stored across a large number of nodes, and these nodes alternate between processing their local data and communicating with a central server. In the case of node failures, the part of database stored in the failed nodes are effectively lost from the learning procedure, degrading the quality of the final solution. In the more common case of straggler nodes, learning is slowed down since the entire procedure waits for the slowest nodes. These effects are especially dominant in an emerging setting for distributed learning, called federated learning, where the local computations are done by mobile devices across the world, which communicate with a central server. In this case, communication delays and available computational resources can vary significantly across devices, and over time, and it is critical for the system to be robust to such variations.

In this thesis, we propose a framework of *encoded* distributed optimization, which adds redundancy in the data to counteract the effect of such failures and delays.

1.2 Thesis Outline

This thesis consists of the design and analysis of several communication and computation techniques that are provably robust against the types of variability outlined in this section.

In Chapter 2, we consider the use of intermittent channel output feedback for interference management in wireless networks. We develop a technique that harnesses such unreliable feedback, and show that it approximately achieves the information-theoretic capacity region for this model. The results of this chapter demonstrate that even unreliable feedback can be used to mitigate the harmful effect of interference in wireless networks.

In Chapter 3, we focus on full-duplex cellular networks, where base station serves uplink and downlink users over the same time-frequency resource, which results in interference from uplink users to downlink ones. By exploiting the variations in the network state, we design a simple, low-overhead joint uplink-downlink user scheduling algorithm, which is shown to optimally mitigate inter-user interference in the many-user regime.

We explore the use of device-to-device (D2D) cooperation over unlicensed bands in Chapter 4. We design a physical-layer D2D cooperation scheme that is shown to be information-theoretically approximately optimal, and using this cooperation mechanism as a building block, we develop a resource allocation policy that allocates the use of such D2D links in a network, while accounting for various constraints such as interference over the unlicensed band, and fairness considerations.

We turn to distributed computing in Chapter 5, and study node failures in distributed optimization under master-worker architecture. We introduce the encoded distributed optimization framework, which consists of encoding of the dataset to add redundancy, and a coding-oblivious application of any distributed optimization algorithm. We theoretically and numerically analyze the robustness achieved by this approach against node failures.

In Chapter 6, we develop the encoded distributed optimization framework further, accounting for stragglers and temporal variations in network conditions. We theoretically analyze the convergence of various popular optimization methods under this framework. We also discuss code design and other practical issues, and present our experimental results over computing clusters.

Finally, in Chapter 7, we present our conclusions and discuss open questions.

We point out that most of the material in this thesis has been published, or submitted for publication as of this date. The contents of Chapter 2 were published in [KWD13b, KWD13a, KWD15], those in Chapter 3 were published in [KD15], those in Chapter 4 can be found in [KD17], those in Chapter 5 were in [KSD17a], and the contents of Chapter 6 are partly published in [KSD17b], and partly submitted for publication in [KSD18].

CHAPTER 2

Opportunistic Feedback for Interference Management

2.1 Introduction

From a physical-layer standpoint, the fundamental bottleneck limiting performance in modern wireless networks is interference. The coding and transmission techniques developed in recent decades, such as MIMO [Tel99], LDPC codes [RSU01], turbo codes [BGT93], and polar codes [Ari09] boosted the point-to-point transmission rates near their theoretical limits, while rapid increase in wireless traffic demand resulted in dense deployment of networks. As a result, most wireless networks today operate in the interference-limited regime. This phenomenon has resulted in a myriad of research efforts in the last decade, aiming to mitigate the effects of interference in wireless networks.

One appealing idea along this direction is to use feedback from receiver to transmitter. Most modern wireless technologies and standards already employ the feedback mechanism for various purposes from multiple-access to incremental redundancy and hybrid ARQ techniques. The work in [ST11] showed that another type of feedback, namely channel output feedback, can be an effective mechanism to alleviate the effect of interference in wireless networks. The main idea underlying this approach is that channel output feedback informs that transmitter of interference signals in the past transmissions, which then allows it to reprocess and forward this interference signal in the future transmission slots, which is useful information for both its own intended destination -since it allows for the cancellation of past interference- and the intended destination of the interfering transmission.

Although being promising, the techniques described in this work fundamentally relies on the assumption that the feedback is always perfectly available at the transmitter. However, such perfect availability is typically not feasible in modern wireless networks, which have dynamically varying resources. In this work, we investigate the effect of feedback for interference management when the feedback is only intermittently available at the transmitter, due to such variations in resources. This model naturally leads to techniques that use feedback more opportunistically, which we show is still useful for interference management.

The chapter is organized as follows. We provide background and motivation for studying interference channel with intermittent feedback in Section 2.2. We formally state the problem and establish the notation in Section 2.3. We present our main results in Section 2.5 and give interpretations of them. We motivate our coding scheme and explain it through an example in Section 2.4. We give the analysis of the coding scheme in Section 2.6. The outer bound is developed in Section 2.7 and Section 2.8 concludes the chapter with a brief discussion of possible extensions of the work. Many of the detailed proofs are given in Appendix A.

2.2 Interference Channel with Intermittent Feedback

The simplest information-theoretic model for studying interference is the two-user Gaussian *interference channel* (IC). It has been shown that feedback can provide an unbounded gain in capacity for two-user Gaussian interference channels [ST11], in contrast to point-to-point memoryless channels, where feedback gives no capacity gain [Sha56], and multiple-access channels, where feedback can at most provide *power* gain [Oza84]. This has been demonstrated when the feedback is unlimited, perfect, and free of cost in [ST11]. Given the optimistic result obtained under this setting, a natural question arises: Can feedback be leveraged for interference management under imperfect feedback models?

There have been several pieces of work so far, attempting to answer this question. Vahid *et al.* [VSA12] considered a rate-limited feedback model, where the feedback links are modeled as fixed-capacity deterministic bit pipes. They developed a scheme based on decode-and-forward at transmitters and lattice coding to extract the helping information in the feedback links, and showed that it achieves the sum-capacity to within a constant gap. The work in [LTM12] studied a deterministic model motivated by passive feedback over AWGN channels, and [SAY09, SWT12] studied the two-way interference channel, where the feedback is provided through a backward interference channel that occupies the same resource as the forward channel. [LTM12, SAY09] and [SWT12] only dealt with the linear deterministic model [ADT11] of the Gaussian IC.

In this work, we investigate how to exploit *intermittent* feedback for managing interference. Such intermittent feedback could occur in several situations. For example, one could use a side-channel such as WiFi for feedback; in this case since the WiFi channel is best-effort, dropped packets might cause intermittent feedback. In other situations, control mechanisms in higher network layers could cause the feedback resource to be available intermittently. For the feedback links, Bernoulli processes $\{S_1[t]\}$ and $\{S_2[t]\}$ control the presence of feedback for user 1 and 2, respectively. The two processes can be dependent, but their joint distribution is i.i.d. over time. We assume that the receivers are *passive*: they simply feedback their received signals back to the transmitters without any processing. In other words, each transmitter receives from feedback an observation of the channel output of its own receiver through an erasure channel, with unit delay. We focus on the passive feedback model as the intermittence of feedback is motivated by the availability of feedback resources (either through use of best-effort WiFi for feedback or through feedback resource scheduling). Therefore, it might be that the time-variant statistics of the intermittent feedback are not *a priori* available at the receiver, precluding active coding. Moreover, the availability of the feedback resource may not be known ahead of transmission, therefore motivating the assumption of causal state information at the transmitter. If the receiver has *a priori* information about the feedback channel statistics, it can perform active coding, in which case, the intermittent feedback model reduces to the rate-limited model of [VSA12].

We study the effect of intermittent feedback for the two-user Gaussian IC inspired by ideas we develop for the linear deterministic IC model [ADT11]. Our main contribution is the approximate characterization of the capacity region of the interference channel with

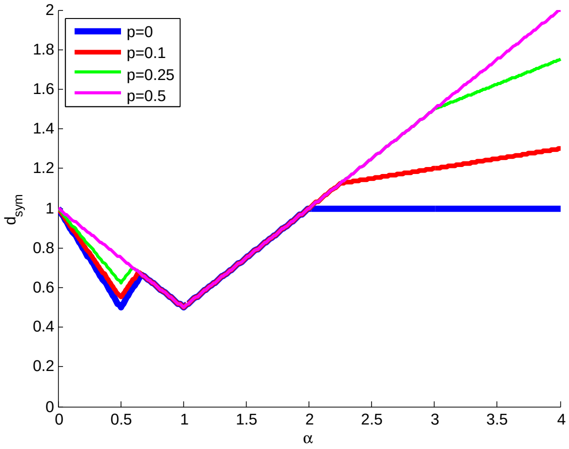


Figure 2.1: Generalized degrees of freedom with respect to interference strength $\alpha := \frac{\log \mathsf{INR}}{\log \mathsf{SNR}}$ for symmetric channel parameters.

intermittent feedback, under the Gaussian model. We also derive an exact characterization of the capacity region under the linear deterministic model, which agrees with the Gaussian result. The capacity characterizations under both models depend only on the forward channel parameters and the marginal distributions of S_1 and S_2; not on their joint distribution.

Our result shows that feedback can be harnessed to provide multiplicative gain in Gaussian interference channel capacity even when it is unreliable and intermittent. The result can be interpreted using the picture given in Figure 2.1, which is depicted (for convenience) in terms of symmetric generalized degrees of freedom for the special case of symmetric channel parameters. The given GDoF curves suggest that as the feedback probability increases, the achievable GDoF also increases for all interference regimes for which perfect feedback provides any GDoF gain. One can also observe from the figure that the capacity gain from intermittent feedback, which depends on the portion of time when the feedback is active, remains unbounded, similar to the perfect feedback case.

A consequence of this result is that when the feedback links are active with large enough probabilities, the sum-capacity of the perfect feedback channel can be achieved to within a constant gap. Similarly for the linear deterministic case, the perfect feedback capacity is exactly achieved even when there is only intermittent feedback, with large enough "on" probability. In particular, under the symmetric setting, this threshold is $1/2$ for each feedback link. This is also reflected in Figure 2.1, where the "V-curve" achievable with perfect feedback is already achievable when the feedback probability is only $1/2$.

Our achievable scheme has three main differences from the previous schemes developed in [ST11, VSA12] and [LTM12]. First, we use *quantize-map-and-forward* (QMF)[1] [ADT11] at the transmitters to send the information obtained through feedback, as opposed to (partial or complete) decode-and-forward, which has been used in [ST11, VSA12, LTM12]. This is because when there is intermittent feedback, the transmitters might not be able to decode the other user's (partial) message, but would still need to send useful information about the interference. A similar situation arises in a relay network, where QMF enables forwarding of evidence, without requiring decoding [ADT11]. Second, at the receivers, we perform forward decoding of blocks instead of backward decoding, which results in a better delay performance. Third, we do not use structured codes, *i.e.*, we only perform random coding.

We also develop novel outer bounds that are within a constant of the achievable rate region for the Gaussian IC and match the achievable region for the linear deterministic IC. These outer bounds are based on constructing an enhanced channel and appropriate side-information. These are illustrated in Section 2.7.

Lastly, we extend these results for packet transmission channels, modeled through parallel channels which are M-symbol extensions of the original model. This can be considered as a model for OFDM and packet drops over a best-effort channel.

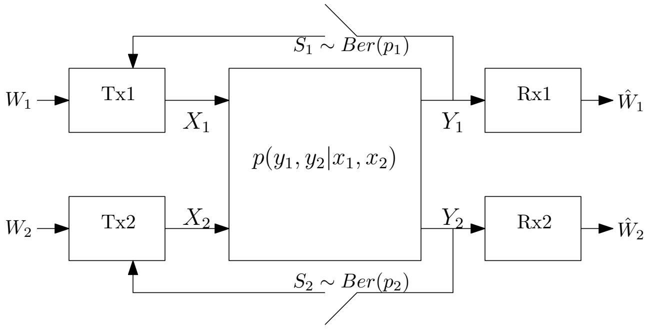


Figure 2.2: Two-user discrete memoryless interference channel with intermittent feedback

2.3 Model and Formulation

We consider the 2-user discrete memoryless interference channel (DM-IC) with intermittent feedback, illustrated in Figure 2.2. We assume Transmitter i (Txi) has a message W_i intended for Receiver i (Rxi), $i = 1, 2$. $W_1 \in \left[2^{NR_1}\right]$ and $W_2 \in \left[2^{NR_2}\right]$ are independent and uniformly distributed, where, for $n \in \mathbb{N}$, $[n] := \{k \in \mathbb{N} : k \leq n\}$. The signal transmitted by Txi at time t is denoted by $X_{i,t} \in \mathcal{X}_i$, while the channel output observed at Rxi is denoted by $Y_{i,t} \in \mathcal{Y}_i$, for $i = 1, 2$. For a block length N, the conditional probability distribution mapping the input codeword to the output sequence is given by

$$p(Y_1^N, Y_2^N|X_1^N, X_2^N) = \prod_{t=1}^{N} p\left(Y_{1,t}, Y_{2,t}|X_{1,t}, X_{2,t}\right)$$

The feedback state sequence pair $\underline{S} := \left(S_1^N, S_2^N\right)$ have the joint distribution

$$p\left(S_1^N, S_2^N\right) = \prod_{t=1}^{N} p\left(S_{1,t}, S_{2,t}\right).$$

and marginally, at time t, $S_{i,t} \sim \mathit{Bernoulli}(p_i)$, for $i = 1, 2$, for all t and N. Note that, for any fixed time slot t, the random variables $S_{1,t}$ and $S_{2,t}$ are not necessarily independent, that is, the joint distribution $p(S_{1,t}, S_{2,t})$ can be arbitrary. We assume that receivers have access to $\underline{S}$ strictly causally, that is, at time t, both receivers know the realization of $\underline{S}^{t-1}$.

[1] The QMF scheme of [ADT11] was generalized to DMCs in [LKE11] (and the scheme was called noisy network coding) and to lattices in [OD10, OD13]. In this work we develop the "short-messaging" version of QMF [KH11] instead of the "long-messaging" version first studied in [ADT11] and extended to DMCs in [LKE11]. For a longer discussion about this and other issues, refer to Section 2.8.

At the beginning of time t, Txi observes the channel output received by Rxi at time $t-1$ through an erasure channel, *i.e.*, it receives $\widetilde{Y}_{i,t-1} := S_{i,t-1}Y_{i,t-1}$, for $i = 1, 2$. Note that this is a *passive* feedback model, in that it does not allow the receiver to perform any processing on the channel output; it simply forwards the received signal Y_i at every time slot, which gets erased with probability $1 - p_i$.

For random variables A and B, we use the notation $A \overset{\text{f}}{=} B$ to denote that A is a deterministic function of B[2]. Then our channel model implies $X_{i,t} \overset{\text{f}}{=} \left(W_i, S_i^{t-1}, \widetilde{Y}_i^{t-1}\right)$.

A rate pair (R_1, R_2) is said to be achievable if there exists a pair of codebooks $(\mathcal{C}_1, \mathcal{C}_2)$ at Tx1 and Tx2, with rates R_1 and R_2, respectively, and pairs of encoding and decoding functions such that the average probability of error at any decoder goes to zero as the block length N goes to infinity. The capacity region with feedback probabilities p_1 and p_2, $\mathcal{C}(p_1, p_2)$, is defined as the closure of the set of all achievable rate pairs (R_1, R_2) when $S_1 \sim Bernoulli(p_1)$ and $S_2 \sim Bernoulli(p_2)$. Sum-capacity is defined by

$$C^{\text{sum}}(p_1, p_2) := \sup\left\{R_1 + R_2 : (R_1, R_2) \in \mathcal{C}(p_1, p_2)\right\}.$$

In this work, we consider two specific channel models (that is, two specific classes of

$$\left(\mathcal{X}_1, \mathcal{X}_2, \mathcal{Y}_1, \mathcal{Y}_2, p\left(y_1, y_2 | x_1, x_2\right)\right)$$

tuples), described in the following subsections.

2.3.1 Linear deterministic model

This channel model was introduced in [ADT11] and since then proved useful in providing insight into the nature of signal interactions many network information theory problems (see Figure 2.3).

We assume $X_{i,t} \in \mathbb{F}_2^q$, for $i = 1, 2$, where $\mathbb{F}_2$ is the binary field. The received signal at

[2] More formally, $A \overset{\text{f}}{=} B$ means that there exists a $\sigma(B)$-measurable function f such that $A = f(B)$ almost surely, where $\sigma(B)$ is the sigma-algebra generated by B.

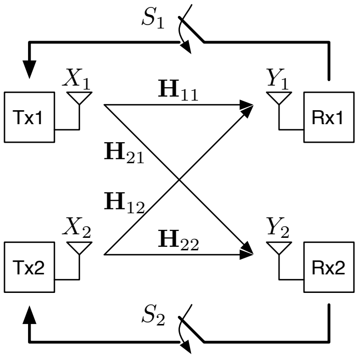


Figure 2.3: Two-user linear deterministic interference channel with intermittent feedback

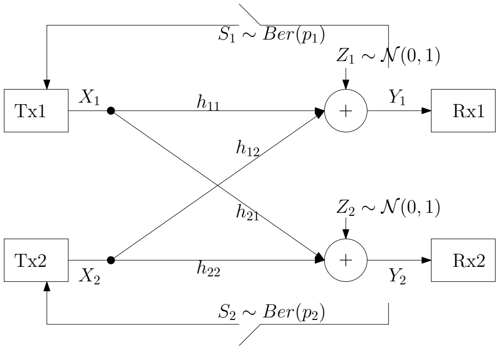


Figure 2.4: Two-user Gaussian interference channel with intermittent feedback

Rxi is given by

$$Y_{i,t} = \mathbf{H}_{ii} X_{i,t} + \mathbf{H}_{ij} X_{j,t}$$

for $(i,j) = (1,2),(2,1)$. The channel matrices are given by $\mathbf{H}_{ij} := \mathbf{S}^{q-n_{ij}}$ for $(i,j) \in \{1,2\}^2$, where $q = \max\{n_{11}, n_{12}, n_{21}, n_{22}\}$, and $\mathbf{S} \in \mathbb{F}_2^{q \times q}$ is the shift matrix $\begin{bmatrix} \mathbf{0}^T & 0 \\ \mathbf{I}_{q-1} & \mathbf{0} \end{bmatrix}$, where $\mathbf{0}$ is the zero vector in $\mathbb{F}_2^{q-1}$ and $\mathbf{I}_{q-1}$ is the identity matrix in $\mathbb{F}_2^{(q-1)\times(q-1)}$. We also define, for $(i,j) = (1,2),(2,1)$,

$$V_{i,t} = \mathbf{H}_{ji} X_{i,t}.$$

The capacity region for the linear deterministic model will be denoted by $\mathcal{C}_{LDC}(p_1,p_2)$, while its sum-capacity will be denoted by $C_{LDC}^{\text{sum}}(p_1,p_2)$.

2.3.2 Gaussian model

Under the canonical Gaussian model (see Figure 2.4), the channel outputs are related to the inputs through the equations

$$Y_{1,t} = h_{11} X_{1,t} + h_{12} X_{2,t} + Z_{1,t}$$

$$Y_{2,t} = h_{21} X_{1,t} + h_{22} X_{2,t} + Z_{2,t}$$

where $h_{ij} \in \mathbb{C}$, for $(i,j) \in \{1,2\}^2$, are channel gains, and $Z_{1,t}, Z_{2,t} \sim \mathcal{CN}(0,1)$ are circularly symmetric complex white Gaussian noise. We assume an average transmit power constraint of P_i at Txi, *i.e.*, for any length-N codeword X_i^N transmitted by Txi, $\frac{1}{N}\sum_{t=1}^{N} |X_{i,t}|^2 \leq P_i$, $i = 1,2$. We also define

$$\mathsf{SNR}_i := |h_{ii}|^2 P_i$$
$$\mathsf{INR}_i := |h_{ij}|^2 P_j$$

and

$$V_i := h_{ji} X_i + Z_j,$$
$$\widetilde{V}_i := S_j V_i,$$

for $(i,j) = (1,2),(2,1)$. Note that this definition of $V_{i,t}$ is consistent with its definition under linear deterministic model, in the sense that it is what remains out of the channel output when the intended signal is completely cancelled.

The capacity region for the Gaussian model will be denoted by $\mathcal{C}_G(p_1,p_2)$, while its sum-capacity will be denoted by $C_G^{\text{sum}}(p_1,p_2)$. We will also use the notation $C_{G,p}^{\text{sum}} := C_G^{\text{sum}}(1,1)$, denoting the sum-capacity under perfect feedback.

Gaussian parallel channel is described by the equations

$$\mathbf{Y}_{1,t} = h_{11}\mathbf{X}_{1,t} + h_{12}\mathbf{X}_{2,t} + \mathbf{Z}_{1,t} \tag{2.1}$$
$$\mathbf{Y}_{2,t} = h_{21}\mathbf{X}_{2,t} + h_{22}\mathbf{X}_{2,t} + \mathbf{Z}_{2,t} \tag{2.2}$$
$$\widetilde{\mathbf{Y}}_{1,t} = S_{1,t}\mathbf{Y}_{1,t} \tag{2.3}$$
$$\widetilde{\mathbf{Y}}_{2,t} = S_{2,t}\mathbf{Y}_{2,t} \tag{2.4}$$

where $\mathbf{X}_{i,t}, \mathbf{Y}_{i,t} \in \mathbb{C}^M$, $i = 1,2$, are the channel input and output, respectively, at user i; $\mathbf{Z}_{1,t}$ and $\mathbf{Z}_{2,t}$ are independent and distributed with $\mathcal{CN}(\mathbf{0},\mathbf{I})$; and $\widetilde{\mathbf{Y}}_{i,t}, i = 1,2$ is the output of the feedback channel of Txi, at time t. Note that the channel gains are scalars. It should also be noted that any given time, the same feedback state variable $S_{i,t}$ controls the presence of feedback for all sub-channels, *i.e.*, the feedback is present either for all M channels, or for none of them.

2.4 Insights from Linear Deterministic Model

In this section, we illustrate our coding scheme through an example over the linear deterministic channel. This example is intended to demonstrate how and why the proposed scheme works, and motivate the use of quantize-map-forward as a feedback strategy.

We consider the symmetric channel shown in Figures 2.5 and 2.6, with $n_{11} = n_{22} = 4$, $n_{12} = n_{21} = 2$, and $p_1 = p_2 = 0.5$, and focus on the achievable symmetric rate. In this example we will take a block length of $N = 2$ for illustration purposes. Although for this particular case, the probability of decoding error is large due to short block length, in general the same coding idea can be applied for a large block length, in which case arbitrarily small error probability can be achieved by taking advantage of the law of large numbers.

We focus on two blocks of transmission. At each block, the users split their messages into common and private parts. The common parts of the messages are decoded by both receivers, whereas the private part is only decoded by the intended receiver, as in Han-Kobayashi scheme for the interference channel without feedback [HK81]. In the first block, Tx1 sends linear combinations of its two common information symbols, a_1, a_2 on its two common (upper) levels, and linear combinations of its private information symbols, a_3, a_4, a_5, a_6, over its private (lower) two levels over a block of two time slots. Tx2 performs similar operations for its common symbols b_1, b_2, and its private symbols b_3, b_4, b_5, b_6.

Note that at this point, the receivers can decode the symbols sent at their upper two levels by solving the four equations in two unknowns.

After each time slot, the receivers feed back their channel outputs, but the transmitters wait until the end of the block to collect sufficient information from feedback. We consider a particular feedback channel realization $\left(S_1^N, S_2^N\right) = ((1,0),(0,1))$ for illustration purposes. After the first block, each transmitter gets from feedback two linear combinations of the interfering symbols of the previous block, by subtracting their own linear combinations from the channel outputs. In the second block, the transmitters perform further linear encoding of these two linear combinations. These additional linear combinations of the interference

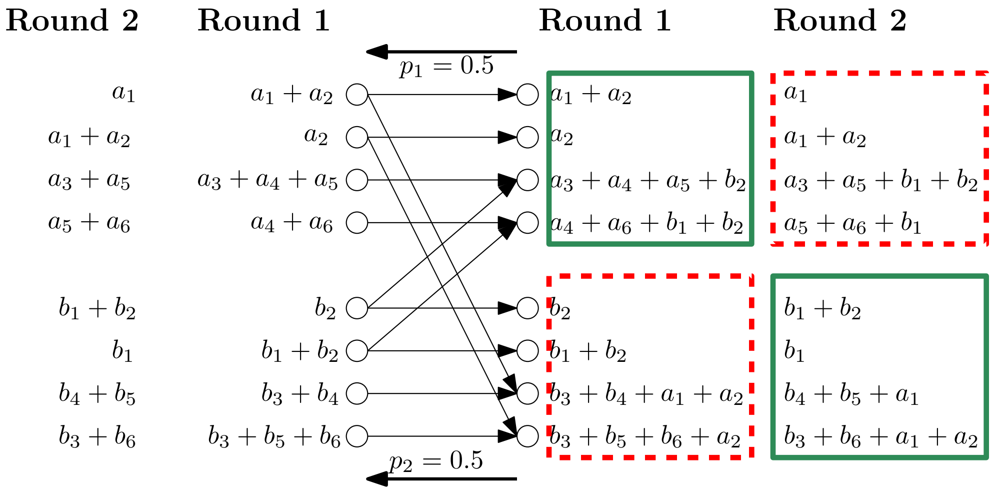


Figure 2.5: First block of transmissions for the example coding scheme over linear deterministic channel. Receptions enclosed in green/solid rectangles represent the channel outputs that the receivers are able to feed back; whereas those enclosed in red/dashed rectangles represent the channel outputs that gets erased through the feedback channel.

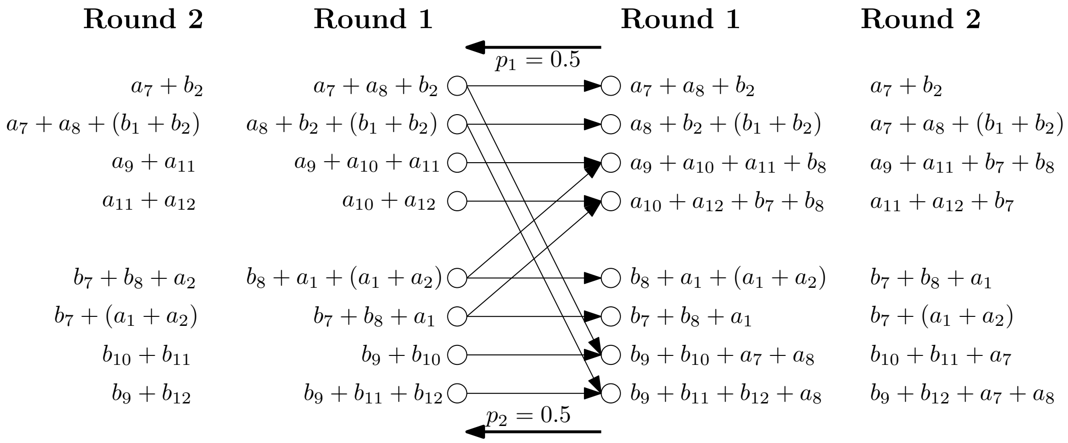


Figure 2.6: Second block of transmissions for the example coding scheme over linear deterministic channel. The helping information sent by the interfering transmitters (a_1, a_2 at Rx1, b_1, b_2 at Rx2) are omitted for brevity. Note that these are already known at the receivers from previous block, and hence can be cancelled.

symbols are superimposed on top of the linear combinations of the fresh common information symbols a_7, a_8 (and b_7, b_8 for Tx2) of the second block. On the private levels, linear combinations of new symbols $a_9, a_{10}, a_{11}, a_{12}$ at Tx1 and $b_9, b_{10}, b_{11}, b_{12}$ at Tx2 are sent, as in the first block.

After the second block of transmission, the receivers collect the four linear equations obtained in the lower two levels of the first block and the four linear equations obtained at the upper two uninterfered levels in the second block. It is easy to check that these eight equations are linearly independent, and hence the receivers can solve for the eight unknowns ($a_3, a_4, a_5, a_6, a_7, a_8, b_1, b_2$ for Tx1, and $b_3, b_4, b_5, b_6, b_7, b_8, a_1, a_2$ for Tx2).

Having decoded the private information (and interference) of the first block and the common information of the second block, the receivers next cancel the additional linear combinations of the previously decoded common information received at the lower two levels of the second block due to feedback. This means that Rx1 cancels the a_1 and a_2 symbols in the lower two levels, and Rx2 cancels the b_1 and b_2 symbols.

Since the transmitters can also cancel this information from the received feedback (because it is a function of their own symbols), the state of each terminal reduces to that in the end of the first block. Therefore, in each of the following blocks, the operation in the second block can be repeated, each time letting the receivers decode the private information of the previous block and the common information of the new block.

One caveat is that, the feedback channel realization will not be the same at each block. To address this point, we first note that the only decoding error event is when the channel realization is such that the resulting linear system in any of the receivers is not full rank. For the particular code in the example, it is easy to check that the probability of this event is zero for any feedback channel realization as long as $S_i^N \neq (0,0)$ for $i = 1, 2$. In general, for any $\epsilon > 0$, in order to achieve a symmetric rate $C_{\text{sym}} - \epsilon$, Tx$i$ needs to receive feedback for at least $N(p_i - \epsilon)$ time slots at each block. This condition is ensured by law of large numbers by letting $N \to \infty$, and arbitrarily small error probability can be achieved[3].

[3] Note that this does not prove the existence of a sequence of codes that allows arbitrarily small error probability for an arbitrary block length. The intention in this section is to give an illustration of the coding scheme; the precise achievability proof will be presented in Section 2.6.

To find the symmetric rate achieved by this scheme, we assume the scheme is run for B blocks. At the end, each receiver will have resolved $6B-4$ information bits in $2B$ time slots. Letting $B \to \infty$ gives a symmetric rate of 3 bits/time slot. Note that without feedback, a symmetric rate of at most 2 bits/time slot can be achieved. At the other extreme, it is also easy to verify from the results in [ST11] that symmetric capacity under perfect feedback is also 3 bits/time slot, which is in agreement with Figure 2.7 and Corollary 2.1.

This example also serves to demonstrate why we perform quantize-map-forward instead of decode-and-forward as a feedback strategy. In general, to achieve the symmetric capacity, Tx2 needs to send linear combinations of N information symbols on its common levels, while Tx1 receives $2Np_1$ of these linear combinations on the average. Hence, if $p_1 < 0.5$, Tx1 will not be able to decode the interference of the previous block. Instead, Tx1 performs a linear mapping of the received feedback information, which turns out to achieve the symmetric capacity.

Finally, we point out that decoding in this scheme is *sequential*, *i.e.*, the receiver decodes the blocks in the same order they are encoded[4]. This is in contrast to earlier feedback coding schemes proposed for interference channel, which perform backward coding. The obvious advantage of using sequential decoding is better delay performance, since the receiver does not need to wait for the end of the entire transmission to start to decode.

2.5 Capacity Results for Interference Channels with Intermittent Feedback

In this section, we present our results and discuss their consequences for both linear deterministic and Gaussian models.

[4] An alternate scheme based on backward decoding was presented in [KWD13b], for the case of linear deterministic channel.

$$R_1 \leq \min\left\{\max(n_{11}, n_{12}), n_{11} + p_2(n_{21} - n_{11})^+\right\} \tag{2.5}$$

$$R_2 \leq \min\left\{\max(n_{22}, n_{21}), n_{22} + p_1(n_{12} - n_{22})^+\right\} \tag{2.6}$$

$$R_1 + R_2 \leq \max(n_{11}, n_{12}) + (n_{22} - n_{12})^+ \tag{2.7}$$

$$R_1 + R_2 \leq \max(n_{22}, n_{21}) + (n_{11} - n_{21})^+ \tag{2.8}$$

$$\begin{aligned} R_1 + R_2 \leq{} & \max\left\{n_{12}, (n_{11} - n_{21})^+\right\} + \max\left\{n_{21}, (n_{22} - n_{12})^+\right\} \\ & + p_1 \min\left\{n_{12}, (n_{11} - n_{21})^+\right\} + p_2 \min\left\{n_{21}, (n_{22} - n_{12})^+\right\} \end{aligned} \tag{2.9}$$

$$\begin{aligned} 2R_1 + R_2 \leq{} & \max(n_{11}, n_{12}) + \max\left\{n_{21}, (n_{22} - n_{12})^+\right\} + (n_{11} - n_{21})^+ \\ & + p_2 \min\left\{n_{21}, (n_{22} - n_{12})^+\right\} \end{aligned} \tag{2.10}$$

$$\begin{aligned} R_1 + 2R_2 \leq{} & \max(n_{22}, n_{21}) + \max\left\{n_{12}, (n_{11} - n_{21})^+\right\} + (n_{22} - n_{12})^+ \\ & + p_1 \min\left\{n_{12}, (n_{11} - n_{21})^+\right\} \end{aligned} \tag{2.11}$$

2.5.1 Linear deterministic model

The following theorem captures our main result for the linear deterministic model.

Theorem 2.1. *The capacity region $\mathcal{C}_{LDC}(p_1, p_2)$ of the linear deterministic interference channel with intermittent feedback is given by the set of rate pairs (R_1, R_2) satisfying* (2.5)–(2.11)*.*

Proof. See Section 2.6 for achievability, and Section 2.7 for converse. □

The following corollary shows that it is possible to achieve perfect feedback sum-capacity even when feedback probabilities are less than one.

Corollary 2.1. *For $n_{12}, n_{21} > 0$, there exists $p^* < 1$ such that*

$$C_{LDC}^{sum}(p_1, p_2) = C_{LDC}^{sum}(1, 1)$$

for all $p_1, p_2 \geq p^*$.

Proof. See Appendix A.7. □

We illustrate Corollary 2.1 through an example. Let us assume $n_{12} = n_{21} = m$, $n_{11} = n_{22} = n$, and $p_1 = p_2 = p$. It is easy to see that if $p_1 = p_2 = 0.5$, the bounds on $R_1 + R_2$, $2R_1 + R_2$ and $R_1 + 2R_2$ that involve p_1 and p_2 become redundant, and the sum-capacity does not increase beyond this point, for all (m, n).

2.5.2 Gaussian model

We define, for any set $\mathcal{R}$ of rate pairs (R_1, R_2) and scalar $\delta \in \mathbb{R}$,

$$\begin{aligned}\mathcal{R} - \delta &:= \left\{(R_1, R_2) : (R_1 + \delta, R_2 + \delta) \in \mathcal{R}\right\},\\ \mathcal{R} + \delta &:= \left\{(R_1, R_2) : (R_1 - \delta, R_2 - \delta) \in \mathcal{R}\right\}.\end{aligned}$$

The following theorem captures our main result for the Gaussian model.

Theorem 2.2. *The capacity region $\mathcal{C}_G(p_1, p_2)$ of the Gaussian interference channel with intermittent feedback satisfies*

$$\bar{\mathcal{C}}(p_1, p_2) - \delta_1 \subseteq \mathcal{C}_G(p_1, p_2) \subseteq \bar{\mathcal{C}}(p_1, p_2) + \delta_2 \tag{2.12}$$

where $\bar{\mathcal{C}}\left(p_1, p_2\right)$ is the set of (R_1, R_2) satisfying (2.13)–(2.17) *for $(i, j) = (1, 2), (2, 1)$ and $\delta_1 < 2\log 3 + 3\left(p_1 + p_2\right)$ bits, and $\delta_2 < \log 3 + p_1 + p_2$ bits.*

Proof. Section 2.6 proves an inner bound region $\mathcal{R}_G^i(p_1, p_2)$, Section 2.7 proves an outer bound region $\mathcal{R}_G^o(p_1, p_2)$, and Appendix A.6 shows that $\bar{\mathcal{C}}(p_1, p_2) - \delta_1 \subseteq \mathcal{R}_G^i(p_1, p_2)$ and $\mathcal{R}_G^o(p_1, p_2) - \delta_2 \subseteq \bar{\mathcal{C}}(p_1, p_2)$. □

Remark 2.1. *Theorem 2.2 uniformly approximates the capacity region under Gaussian model to within a gap of $3\log 3 + 4\left(p_1 + p_2\right)$ bits, independent of channel parameters. To our knowledge, this is the first constant-gap capacity region characterization for interference channel with noisy feedback with arbitrary channel parameters.*

$$R_i < \log\left(1 + \mathsf{SNR}_i + \mathsf{INR}_i\right) \tag{2.13}$$

$$R_i < \log\left(1 + \mathsf{SNR}_i\right) + p_j \log\left(1 + \frac{\mathsf{INR}_j}{1 + \mathsf{SNR}_i}\right) \tag{2.14}$$

$$R_i + R_j < \log\left(1 + \frac{\mathsf{SNR}_i}{1 + \mathsf{INR}_j}\right) + \log\left(1 + \mathsf{SNR}_j + \mathsf{INR}_j\right) \tag{2.15}$$

$$\begin{aligned} R_i + R_j < {} & \log\left(1 + \frac{\mathsf{SNR}_i}{1 + \mathsf{INR}_j} + \mathsf{INR}_i\right) + \log\left(1 + \frac{\mathsf{SNR}_i}{1 + \mathsf{INR}_j} + \mathsf{INR}_i\right) \\ & + p_i \log\left(\frac{\left(1 + \mathsf{INR}_i\right)\left(1 + \frac{\mathsf{SNR}_i}{1+\mathsf{INR}_j}\right)}{1 + \frac{\mathsf{SNR}_i}{1+\mathsf{INR}_j} + \mathsf{INR}_i}\right) + p_j \log\left(\frac{\left(1 + \mathsf{INR}_j\right)\left(1 + \frac{\mathsf{SNR}_j}{1+\mathsf{INR}_i}\right)}{1 + \frac{\mathsf{SNR}_j}{1+\mathsf{INR}_i} + \mathsf{INR}_j}\right) \end{aligned} \tag{2.16}$$

$$\begin{aligned} 2R_i + R_j < {} & \log\left(1 + \frac{\mathsf{SNR}_i}{1 + \mathsf{INR}_j}\right) + \log\left(1 + \frac{\mathsf{SNR}_j}{1 + \mathsf{INR}_i} + \mathsf{INR}_j\right) \\ & + \log\left(1 + \mathsf{SNR}_i + \mathsf{INR}_i\right) + p_j \log\left(\frac{\left(1 + \mathsf{INR}_j\right)\left(1 + \frac{\mathsf{SNR}_j}{1+\mathsf{INR}_i}\right)}{1 + \frac{\mathsf{SNR}_j}{1+\mathsf{INR}_i} + \mathsf{INR}_j}\right) \end{aligned} \tag{2.17}$$

Remark 2.2. *As will be seen in the achievability proof, the proposed coding scheme achieves a smaller gap than what is given in Theorem 2.2; however, for simplicity in the achievability proof, we lower bound the achievable rate terms with computationally more tractable ones, which articifically contributes to the claimed gap. Moreover, one can optimize over the parameters of the proposed coding scheme, such as power allocation and quantization distortion, to further reduce the gap, but this issue will not be dealt with in this work.*

Theorem 2.2 allows us to characterize the symmetric generalized degrees of freedom under symmetric channel parameters, which is a metric often used to compare the capabilities of the interference channel under different settings.

Corollary 2.2 (Generalized Degrees of Freedom). *For symmetric channel parameters (*$\mathsf{SNR}_1 = \mathsf{SNR}_2 = \mathsf{SNR}$*,* $\mathsf{INR}_1 = \mathsf{INR}_2 = \mathsf{INR}$*,* $p_1 = p_2 = p$*), the symmetric generalized degrees of freedom of freedom, defined by*

$$d_{sym} := \lim_{\substack{\mathsf{SNR}\to\infty \\ \mathsf{INR}=\mathsf{SNR}^{\alpha}}} \frac{C_{sym}(\mathsf{SNR}, \mathsf{INR}, p)}{\log \mathsf{SNR}},$$

where $C_{sym}(\mathsf{SNR}, \mathsf{INR}, p) := \sup\{R : (R, R) \in \mathcal{C}_G(p, p)\}$, *is given by*

$$d_{sym} = \begin{cases} \min\{1 - \alpha/2, 1 - (1-p)\alpha\}, & \alpha \le 1/2 \\ \min\{1 - \alpha/2, p + (1-p)\alpha\}, & 1/2 \le \alpha \le 1 \\ \min\{\alpha/2, (1-p) + p\alpha\}, & \alpha \ge 1 \end{cases}$$

Figure 2.7 plots the available generalized degrees of freedom with respect to interference strength for various values of p. As can be observed, as p is increased, gradually better curves are obtained. It should be noted that once $p \ge 0.5$, the "V-curve" that is achieved by perfect feedback [ST11] is already achieved. Next, this observation will be made precise.

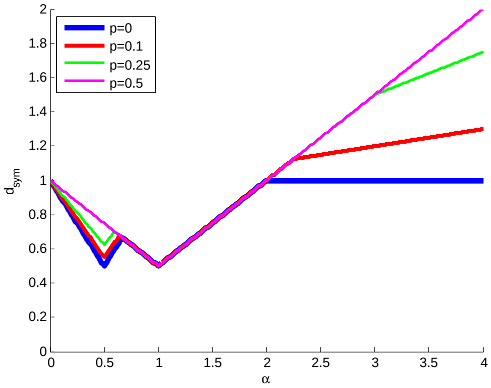


Figure 2.7: Generalized degrees of freedom with respect to interference strength $\alpha := \frac{\log \mathsf{INR}}{\log \mathsf{SNR}}$ for symmetric channel parameters.

The perfect feedback outer bound on the sum-capacity, $C_{G,p}^{\text{sum}}$, is given in Theorem 3 of [ST11] as follows.

$$C_{G,p}^{\text{sum}} \le \sup_{0\le\rho\le1} \min\{\zeta_1(\rho), \zeta_2(\rho)\}$$

$$\zeta_1(\rho) = \log\left(1 + \frac{(1-\rho^2)\mathsf{SNR}_1}{1 + (1-\rho^2)\mathsf{INR}_2}\right) + \log\left(1 + \mathsf{SNR}_2 + \mathsf{INR}_2 + 2\rho\sqrt{\mathsf{SNR}_2 \cdot \mathsf{INR}_2}\right)$$

$$\zeta_2(\rho) = \log\left(1 + \frac{(1-\rho^2)\mathsf{SNR}_2}{1 + (1-\rho^2)\mathsf{INR}_1}\right) + \log\left(1 + \mathsf{SNR}_1 + \mathsf{INR}_1 + 2\rho\sqrt{\mathsf{SNR}_1 \cdot \mathsf{INR}_1}\right)$$

The next corollary shows that when p_1 and p_2 are sufficiently large, the sum-capacity of the perfect feedback Gaussian channel can be achieved with intermittent feedback, to within a constant gap. Hence, this corollary is the Gaussian counterpart of the similar result given in Corollary 2.1, for the linear deterministic channel.

Corollary 2.3. *For* $\mathsf{INR}_1, \mathsf{INR}_2 > 0$*, there exists* $p^* < 1$ *such that*

$$C_{G,p}^{sum} - C_G^{sum}(p_1, p_2) \leq \delta_p$$

for all $p_1, p_2 \geq p^*$*, where* δ_p *is a constant independent of channel parameters.*

Proof. See Appendix A.7. □

In our intermittent feedback model, erasures are symbol-wise, that is, each symbol can get erased independently of others. However, in a best-effort channel, erasures might occur on *packet-level* instead. In order to study this scenario, we consider the parallel channel model described by the equations (2.1)–(2.4), which is simply the M-symbol extension of the Gaussian channel, where the channel parameters are the same for each subchannel. Each extended symbol over this channel models a packet. The result in Theorem 2.2 easily generalizes to parallel channel model, as shown by the following corollary.

Corollary 2.4 (Parallel channel)**.** *The capacity region* $\mathcal{C}_G^{(M)}(p_1, p_2)$ *of any parallel channel of size* M *with feedback probabilities* p_1 *and* p_2 *satisfies*

$$M\bar{\mathcal{C}}(p_1, p_2) - M\delta_1 \subseteq \mathcal{C}_G^{(M)}(p_1, p_2) \subseteq M\bar{\mathcal{C}}(p_1, p_2) + M\delta_2$$

where $\bar{\mathcal{C}}(p_1, p_2), \delta_1$ *and* δ_2 *are as defined in Theorem 2.2.*

Remark 2.3. *Although strictly speaking, the claim in Corollary 2.4 is more general than that in Theorem 2.2, the achievability and converse proofs for the scalar channel directly extend to the parallel channel without any non-trivial modification. Hence, for simplicity, we focus on the scalar case here, and omit a separate proof for the parallel channel.*

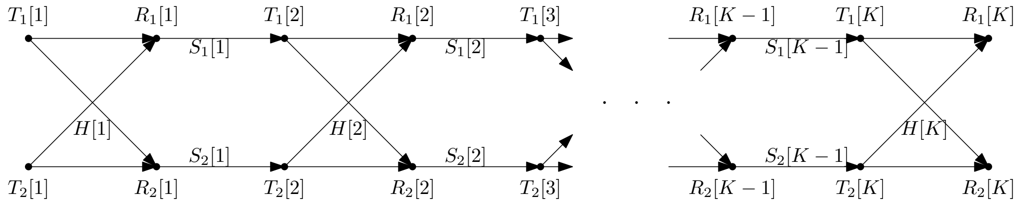


Figure 2.8: The interference network unfolded over a block of K time slots. The node $T_i[t]$ corresponds to the copy of Txi at time t, while $R_i[t]$ corresponds to the copy of Rxi at time t. The feedback channel for time t is an erasure channel controlled by $S_1[t]$ and $S_2[t]$, while the forward channel is a Gaussian interference channel with channel matrix $H[t]$.

2.5.3 Discussion of results

2.5.3.1 Feedback strategy

Our result shows that even unreliable feedback provides multiplicative gain in interference channels. The key insight in showing this result is using quantize-map-forward as a feedback strategy at the transmitters. This is in contrast to the schemes proposed for perfect feedback [ST11] and rate-limited feedback [VSA12], which use decode-and-forward to extract the feedback information. When the feedback channel is noisy, such schemes can result in rates arbitrarily far from optimality. In order to see this, consider unfolding the channel over time, as shown in Figure 2.8. This transformation effectively turns this channel into a relay network, where it is known that decode-and-forward based relaying schemes can give arbitrarily loose rates. This also motivates using quantize-map-forward as a feedback strategy, which has been shown to approximately achieve the relay network capacity [ADT11]. This observation also suggests that quantize-map-forward might be a promising feedback strategy for the additive white Gaussian noise (AWGN) feedback model of [LTM12] in order to uniformly achieve its capacity region to within a constant gap.

It is instructive to compare the achievable rate region for the case of $p_1 = p_2 = 1$ with the outer bound region of the perfect feedback model of [ST11]. Evaluating the region $\bar{\mathcal{C}}(p_1, p_2) - \delta_1$ with $p_1 = p_2 = 1$, we see that the perfect feedback bound (2.15) becomes redundant, and the achievable region comes within $(3 + 3\log 3)$ bits of the outer bound region of [ST11] (see Appendix A.6 for details). We note that this gap is larger than what is achieved by the decode-and-forward based scheme of [ST11]. This shows that uniform approximation of capacity region via quantize-map-forward comes at the expense of an additional (but constant) gap[5]. The source of this additional gap is the quantization step at the transmitters, which introduces a distortion in the feedback signal, and eventually incurs a constant rate penalty whose amount depends on the distortion level.

2.5.3.2 Perfect feedback sum capacity with intermittent feedback

Corollary 2.3 shows that for any set of channel parameters, there exists a threshold p^* on the feedback probability above which perfect feedback sum-capacity is achieved to within a constant gap. Although the exact closed-form expression of p^* is not clean, an examination of the symmetric case (see Figure 2.7) reveals that in some cases it can be as low as 0.5.

The intuition behind this result lies in the fact that it takes the transmitter forward-channel resources to send the information obtained through feedback. Note that the larger p is, the larger the amount of additional information about the past reception can be obtained through intermittent feedback at the transmitters. If the amount of such information is larger than a threshold, then sending it to the receivers will limit the rate for delivering fresh information. Hence, once this threshold is reached, having more feedback resource is no longer useful. However, this property is not observed for the entire capacity region, since if one of the users transmit at a low rate, then it will have sufficient slackness in rate to forward the entire feedback information.

[5] Although we stated that the quantize-map-forward scheme achieves a smaller gap than what is claimed in Theorem 2.2, the actual gap is still expected to be larger than that of the decode-and-forward based scheme for perfect feedback, due to quantization distortion.

2.6 Achievability

In this section, we describe the coding scheme in detail and derive an inner bound $\mathcal{R}_G^i\left(p_1, p_2\right)$ on the rate region.

2.6.1 Overview of the achievable strategy

The main idea of the coding scheme is the same as the one presented for the example in Section 2.4. However, it substantially generalizes the example scheme in order to account for possible channel noise, different interference regimes and an arbitrary target rate point in the achievable region.

The scheme consists of transmission over B blocks, each of length N. At the beginning of block b, upon reception of feedback, transmitters first remove their own contribution from the feedback signal and obtain a function of the interference and noise realization of block $b-1$. This signal is then quantized and mapped to a random codeword, which will be called the helping information. Finally, a new common codeword, which is to be decoded by both receivers, and a private codeword, to be decoded by only the intended receiver, are superimposed to the helping information, and transmitted.

The decoding operation depends on the desired rate point (see Figure 2.10). To achieve the rate points for which the common component of the message is large, the receiver simply performs a variation of Han-Kobayashi decoding [HK81], *i.e.*, it decodes the intended information jointly with the common part of the interference. Note that this does not make use of the helping information.

To achieve the remaining rate points, the helping information is used. For weak interference, at block b, we assume that the receiver has already decoded the intended common information of block $b-1$. After receiving the transmission of block b, the receivers jointly decode the intended private information and the interference of block $b-1$ jointly with the common information of block b, while using the helping information sent at block b as side information. For strong interference, the roles of intended common information and the interfering common information get switched.

Next, we present a detailed description of the coding scheme and proof of achievability.

2.6.2 Codebook generation

Fix $p(x_{ie})p(x_{ic})p(x_{ip})$ for $i = 1, 2$[6], and $p(u_i|\widetilde{v}_j)$ that achieves $\mathbb{E}\left[d(U_i, \widetilde{V}_j)\right] \leq D_i$ for $(i, j) = (1, 2), (2, 1)$, where $d : \mathcal{U} \times \mathcal{V} \to \mathbb{R}$ is the distortion measure, where $\mathcal{U}$ and $\mathcal{V}$ are the alphabets of U_i and $\widetilde{V}_j$, respectively. Generate 2^{Nr_i} quantization codewords U_i^N i.i.d. $\sim p(u_i) = \sum_{\widetilde{v}_j} p(u_i|\widetilde{v}_j)p(\widetilde{v}_j)$, for $(i, j) = (1, 2), (2, 1)$. For $i = 1, 2$, generate 2^{Nr_i} codewords X_{ie}^N i.i.d. $\sim p(x_{ie})$. Further generate, for $i = 1, 2$, $2^{NR_{ic}}$ codewords X_{ic}^N i.i.d. $\sim p(x_{ic})$ and $2^{NR_{ip}}$ codewords X_{ip}^N i.i.d. $\sim p(x_{ip})$. For $i = 1, 2$, define symbol-by-symbol mapping functions $x_i : \mathcal{X}_{if} \times \mathcal{X}_{ip} \to \mathcal{X}_i$ and $x_{if} : \mathcal{X}_{ie} \times \mathcal{X}_{ic} \to \mathcal{X}_{if}$, where $\mathcal{X}_{ie}$, $\mathcal{X}_{ic}$, $\mathcal{X}_{ip}$, and $\mathcal{X}_{if}$ are the alphabets for the symbols X_{ie}, X_{ic}, X_{ip}, and X_{if}, respectively.

2.6.3 Encoding

Encoding is performed over blocks (indexed by b) of length N. See Figure 2.9 for a system diagram. At the beginning of block b, Txi receives the punctured feedback signal $\widetilde{Y}_i^N(b-1) = S_i^N(b-1)Y_i^N(b-1)$ containing information about the channel output in block $b-1$, where the multiplication is element-wise. Upon reception of $\widetilde{Y}_i^N$, Txi first removes its own contribution from the feedback signal to obtain $\widetilde{V}_j^N(b-1) = S_i^N(b-1)V_j^N(b-1)$. For linear deterministic model, this is done by

$$\widetilde{V}_j^N(b-1) = \widetilde{Y}_i^N(b-1) - S_i^N(b-1)\mathbf{H}_{ii}X_i^N(b-1),$$

whereas for Gaussian model, it can be obtained by

$$\widetilde{V}_j^N(b-1) = \widetilde{Y}_i^N(b-1) - S_i^N(b-1)h_{ii}X_i^N(b-1)$$

for $(i, j) = (1, 2), (2, 1)$.

[6] Although the scheme loses beamforming gain by generating independent codebooks at the two users, this only results in a constant rate penalty.

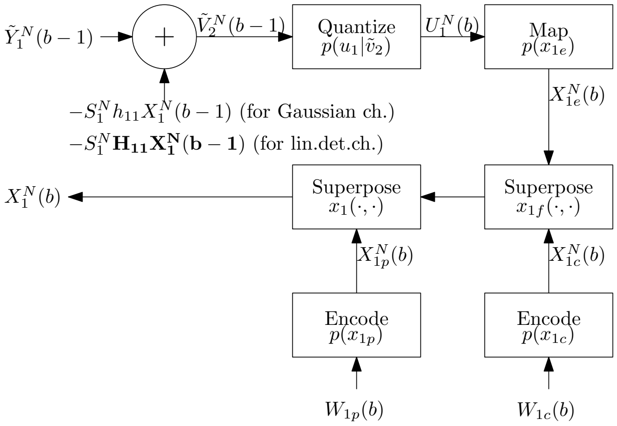


Figure 2.9: Encoder diagram at Tx1

The interference signal $\widetilde{V}_j^N(b-1)$ is then quantized by finding an index $Q_i(b)$ such that

$$\left(\widetilde{V}_j^N(b-1), U_i^N(Q_i(b))\right) \in \mathcal{A}_\epsilon^{(n)},$$

where $\mathcal{T}_\epsilon^{(N)}$ denotes the ϵ-typical set with respect to the distribution $p(\widetilde{v}_j)p(u_i|\widetilde{v}_j)$, and $p(\widetilde{v}_j)$ is induced by the channel and the input distributions. If such an index $Q_i(b)$ has been found, the codeword $X_{ie}^N(Q_i(b))$ that has the same index is chosen to be sent for block b. If there are multiple such indices, the smallest one is chosen. If no such index is found, the quantization index 1 is chosen.

Next, the message $W_i(b) \in \left[2^{NR_i}\right]$ to be sent at block b is split into common and private components $(W_{ic}(b), W_{ip}(b)) \in \left[2^{NR_{ic}}\right] \times \left[2^{NR_{ip}}\right]$. Depending on the desired message indices $(W_{ic}(b), W_{ip}(b))$, a common codeword $X_{ic}^N(W_{ic}(b))$, and a private codeword $X_{ip}^N(W_{ip}(b))$ is chosen from the respective codebooks.

Finally, the using the symbol-wise maps $x_{if}(\cdot,\cdot)$ and $x_i(\cdot,\cdot)$, we obtain the codewords

$$X_{if}^N(b) = x_{if}\left(X_{ie}^N(b), X_{ic}^N(b)\right)$$
$$X_i^N(b) = x_i\left(X_{if}^N(b), X_{ip}^N(b)\right)$$

where the functions are applied to vectors element-wise. $X_i^N(b)$ is sent at Txi over N channel uses.

2.6.4 Decoding

The message indices for common and private messages, and the quantization indices of Txi at block b will be denoted by $m_i(b)$, $n_i(b)$, and $q_i(b)$, respectively. When there are two quantization indices to be decoded from the same user, the second one will be denoted with $q'_i(b)$.

In order to describe the decoding process, we need to introduce some notation. Define the following sequence of sets:

$$\begin{aligned}&\mathcal{B}_i^{(N)}((q_j, m_j)(b-1))\\ &\quad := \Big\{q_i(b) : \big(\underline{S}^N(b-1), X_{jf}^N((q_j, m_j)(b-1)), (U_i^N, X_{ie}^N)(q_i(b))\big) \in \mathcal{T}_\epsilon^{(N)}\Big\}.\end{aligned}$$

for $(i, j) = (1, 2), (2, 1)$. Loosely, $\mathcal{B}_i^{(N)}$ is the set of quantization indices of Txi that are jointly typical with the interference of the previous round. If any of the indices (q_j, m_j) is known, we will suppress the dependence to that index, *e.g.*, if both are known, we simply denote

$$\mathcal{B}_i^{(N)}(b) := \Big\{q_i(b) : \big(\underline{S}^N(b-1), X_{jf}^N(b-1), (U_i^N, X_{ie}^N)(q_i(b))\big) \in \mathcal{T}_\epsilon^{(N)}\Big\}$$

where $X_{jf}^N(b-1)$ refers to the codeword corresponding to the known message indices.

We assume that the set $\mathcal{B}_i^{(N)}(b)$ has cardinality $2^{NK_i(b)}$. Specifically,

$$K_i(b) = \frac{\log\left|\Big\{q_i(b) : \Big(\widetilde{V}_j^N(b-1), U_i^N(q_i(b))\Big) \in \mathcal{A}_\epsilon^{(n)}\Big\}\right|}{N}$$

Note that due to random codebook generation, $K_i(b), i = 1, 2$, are random variables. The following lemma shows that $K_i(b)$ is almost surely bounded for sufficiently large N.

Lemma 2.1. *For any $\epsilon > 0$, there exists a block length N, and a quantization scheme such that $K_i(b) < \kappa_i + \delta(\epsilon)$, where*

$$\kappa_i := I(\widetilde{V}_j; U_i|S_i) - I(X_{jf}; U_i|S_i)$$

for $(i, j) = (1, 2), (2, 1)$, and $\delta(\epsilon)$ is such that $\delta(\epsilon) \to 0$ as $\epsilon \to 0$.

Proof. See Appendix A.1. □

Lemma 2.1 suggests that for each interference codeword, there is a constant number of plausible quantization codewords, for sufficiently large block length (to see that κ_i is a constant independent of channel parameters, refer to Appendix A.3). This means that the cost of jointly decoding the quantization indices together with the actual messages is a constant reduction in the achievable rate, which will be a useful observation in deriving the constant-gap result.

We also define $C_i = \kappa_i + 2\kappa_j$, for $(i, j) = (1, 2), (2, 1)$. The reason for this particular definition will become clear in the error analysis. Intuitively, C_i represents the rate cost associated with performing quantization to forward the feedback information, which introduces distortion. However, as we will show later in the proof, the upper bound given in Lemma 2.1 can be evaluated as a constant independent of channel parameters.

Given an input distribution, Rx1 is said to be in weak interference if $I(X_2; Y_1|X_1) \leq I(X_1; Y_1|X_2)$, and in strong interference otherwise. These regimes are defined similarly for Rx2.

Decoding operation depends on the interference regime and the desired operating point (R_1, R_2). In order to describe the relevant regimes of operating points, we define

$$I_{wi} := I(X_{if}; Y_i|X_{1e}, X_{2e}) - C_i, \tag{2.18}$$

$$I_{si} := I(X_{jf}; Y_i|X_{1e}, X_{2e}) - C_i, \tag{2.19}$$

for $(i, j) = (1, 2), (2, 1)$. In what follows, for clarity, we will focus only on Rx1. The operations performed at Rx2 are similar.

2.6.4.1 Weak interference $(I(X_2; Y_1|X_1) \leq I(X_1; Y_1|X_2))$

If, for the desired operating point, $R_{1c} > I_{w1}$, where I_{w1} is as defined in (2.18), the helping information is not used, and a slight modification of Han-Kobayashi scheme is employed. Otherwise, the helping information is used to decode the information of block $b - 1$. We describe the decoding for the two cases below.

$\mathbf{R_{1c} \geq I_{w1}}$: At block b, we assume that $X_{1e}^N(b)$ and $X_1^N(b - 1)$ are known. The decoder

Interference Regime	Operating Point	Decoding Operation
Weak Interference	$R_{1c} < I_{w1}$	Jointly decode $W_{1p}(b-1), W_{2c}(b-1), W_{1c}(b)$, and quantization index $Q_1(b)$
	$R_{1c} \geq I_{w1}$	Jointly decode $W_{1p}(b), W_{1c}(b)$, and $W_{2c}(b)$ (do not use helping information)
Strong Interference	$R_{2c} < I_{s1}$	Jointly decode $W_{1p}(b-1), W_{1c}(b-1), W_{2c}(b)$, and quantization index $Q_2(b)$
	$R_{2c} \geq I_{s1}$	Jointly decode $W_{1p}(b), W_{1c}(b)$, and $W_{2c}(b)$ (do not use helping information)

Figure 2.10: A high-level summary of the decoding policy at Rx1 (Details are omitted).

attempts to find unique indices $(m_1(b), n_1(b), m_2(b)) \in \left[2^{NR_{1c}}\right] \times \left[2^{NR_{1p}}\right] \times \left[2^{NR_{2c}}\right]$, and some $q_2(b) \in \left[2^{Nr_2}\right]$ such that

$$\left(\begin{array}{l} \underline{S}^N(b-1), X_{1f}^N(b-1), X_{1e}^N(b), X_{2e}^N(q_2(b)), X_{1f}^N(m_1(b)), \\ X_1^N(m_1(b), n_1(b)), X_{2f}^N(q_2(b), m_2(b)), Y_1^N(b) \end{array} \right) \in \mathcal{T}_\epsilon^{(N)} \tag{2.20}$$

where the known message indices are suppressed. If the receiver can find a unique collection of such indices, it declares them as the decoded message indices $\left(\widehat{W}_{1c}(b), \widehat{W}_{1p}(b), \widehat{W}_{2c}(b)\right)$; otherwise it declares an error.

After decoding, given the knowledge of $X_1^N(b)$, Rx1 reconstructs $X_{1e}^N(b+1)$ by imitating the steps taken by Tx1 at the beginning of block $b+1$, thereby maintaining the assumption that $X_{1e}^N(b)$ is known at the beginning of block b. Further, note that $X_{2e}^N(b)$ is not uniquely decoded, hence in block $b+1$, it will still be jointly (but still, non-uniquely) decoded with the variables of that block. We resort to non-unique decoding of this codeword since unique decoding imposes an additional rate constraint on the helping information, thereby limiting the amount of rate enhancement it can provide.

$\mathbf{R_{1c} < I_{w1}}$: At block b, it is assumed that $X_{1f}^N(b-1)$ and $X_1^N(b-2)$ are known at Rx1. To decode, Rx1 attempts to find unique indices $(m_1(b), n_1(b-1), m_2(b-1)) \in \left[2^{NR_{1c}}\right] \times$

$\left[2^{NR_{1p}}\right] \times \left[2^{NR_{2c}}\right]$ and some triple $(q_2(b-1), q_2(b), q_1(b)) \in \left[2^{Nr_2}\right] \times \left[2^{Nr_2}\right] \times \left[2^{Nr_1}\right]$ such that

$$\begin{pmatrix} \underline{S}^N(b-1), X_{1f}^N(b-2), X_{1f}^N(b-1), X_1^N(n_1(b-1)), \\ X_{2e}^N(q_2(b-1)), X_{2c}^N(m_2(b-1)), \left(U_1^N, X_{1e}^N\right)(q_1(b)), X_{2e}^N(q_2(b)), \\ X_{1c}^N(m_1(b)), Y_1^N(b-1), Y_1^N(b) \end{pmatrix} \in \mathcal{T}_\epsilon^{(N)} \tag{2.21}$$

If a unique collection of such indices exists, then these are declared as the decoded message indices $\left(\widehat{W}_{1c}(b), \widehat{W}_{1p}(b-1),\ \ \widehat{W}_{2c}(b-1)\right)$. Otherwise, an error is declared.

In (2.21), the dependence of $X_1^N(b-1)$ to the indices $q_1(b-1)$ and $m_1(b-1)$ is suppressed, since these indices correspond to messages that have already been decoded.

In words, the decoder jointly decodes the private information and the interference of block $b-1$ jointly with the helping information and common information from block b.

Note that non-unique decoding is performed for $X_{1e}^N(b)$, but we have assumed that $X_{1f}^N(b-1)$ $\left(\text{and thus, } X_{1e}^N(b-1)\right)$ is uniquely known at the beginning of block b. In order to maintain this assumption for the next block, $X_{1e}^N(b)$ is reconstructed at Rx1. To achieve this, given the knowledge of $X_1^N(b-1)$, and the quantization codebook, Rx1 imitates the operations performed by Tx1 at the beginning of block b.

2.6.4.2 Strong interference $(I(X_2;Y_1|X_1) > I(X_1;Y_1|X_2))$

As in the weak interference case, decoding depends on the operating point. For $R_{2c} < I_{s1}$, where I_{s1} is as defined in (2.19), helping information is used, otherwise, it is not used.

$\mathbf{R_{2c} \geq I_{s1}}$: The operations performed are identical to those for the case of $R_{1c} \geq I_{w1}$ under weak interference.

$\mathbf{R_{2c} < I_{s1}}$: We assume $X_1^N(b-2)$, $X_{1e}^N(b-1)$, and $X_{2c}^N(b-1)$ are known at Rx1 at block b.

To decode, Rx1 attempts to find unique indices $(m_1(b-1), n_1(b-1), m_2(b)) \in \left[2^{NR_{1c}}\right] \times$

$\left[2^{NR_{1p}}\right] \times \left[2^{NR_{2c}}\right]$ and some $(q_2(b-1), q_2(b), q_1(b)) \in \left[2^{Nr_2}\right] \times \left[2^{Nr_2}\right] \times \left[2^{Nr_1}\right]$ such that

$$\begin{pmatrix} \underline{S}^N(b-1), X_{1f}^N(b-2), X_{1e}^N(b-1), X_{2e}^N(q_2(b-1)), \\ X_{1c}^N(m_1(b-1)), X_{1p}^N(n_1(b-1)), X_{2c}^N(m_2(b)), \\ (U_2^N, X_{2e}^N)(q_2(b)), X_{1e}^N(q_1(b)), Y_1^N(b-1), Y_1^N(b) \end{pmatrix} \in \mathcal{T}_\epsilon^{(N)} \tag{2.22}$$

If a unique collection of such indices exists, they are declared as the decoded message indices $\left(\widehat{W}_{1c}(b-1), \widehat{W}_{1p}(b-1), \widehat{W}_{2c}(b)\right)$. Otherwise, an error is declared. Using the information of $X_1^N(b-1)$, Rx1 can now uniquely reconstruct $X_{1e}^N(b-1)$ by following the steps taken by Tx1 at the beginning of block b.

2.6.5 Error analysis

Without loss of generality, we only consider the error events occurring at Tx1 and Rx1. All arguments here will be applicable to the other Tx-Rx pair. We define the following decoding error events at Rx1, for block b and block length N:

$$\begin{aligned} D_{FB,w}(b,N) &= \Big\{\widehat{W}_{1c}(b) = W_{1c}(b), \widehat{W}_{1p}(b-1) = W_{1p}(b-1), \\ &\qquad \widehat{W}_{2c}(b-1) = W_{2c}(b-1)\Big\}^c \\ D_{FB,s}(b,N) &= \Big\{\widehat{W}_{1c}(b-1) = W_{1c}(b-1), \widehat{W}_{1p}(b-1) = W_{1p}(b-1), \\ &\qquad \widehat{W}_{2c}(b) = W_{2c}(b)\Big\}^c \\ D_{NFB}(b,N) &= \Big\{\widehat{W}_1(b) = W_1(b), \widehat{W}_{2c}(b) = W_{2c}(b)\Big\}^c \end{aligned}$$

The overall decoding error events at Rx1 is given by

$$D_{FB,w}(N) = \bigcup_{b=1}^{B} D_{FB,w}(b),\ D_{FB,s}(N) = \bigcup_{b=1}^{B} D_{FB,w}(b)$$

$$D_{NFB}(N) = \bigcup_{b=1}^{B} D_{NFB}(b)$$

We first prove that in order to find the rate achieved after transmission of B blocks, it is sufficient to focus on the error events at an arbitrary block b. Without loss of generality, consider the error event $D_{FB,w}(N)$. Assume that, after B blocks of transmission, the effective rate achieved by Txi is $\bar{R}_i$ (Note that at the end of block B, some of the information pertaining to block B is still undecoded), which can be lower bounded by $\bar{R}_i \geq \frac{B-2}{B} R_i$, by ignoring the partial information decoded in the first block and the last one. We can also upper bound the overall probability of error by

$$\begin{aligned}\mathbb{P}\left(D_{FB,w}\right) &\leq \sum_{b=2}^{B} \mathbb{P}\left(D_{FB,w}(b,N) \middle| \left\{D_{FB,w}^c(b',N)\right\}_{b'=2}^{b-1}\right) \\ &\leq B\mathbb{P}\left(D_{FB,w}(b,N) \middle| \left\{D_{FB,w}^c(b',N)\right\}_{b'=2}^{b-1}\right) \\ &=: B\mathbb{P}\left(\mathcal{D}_{FB,w}(b,N)\right)\end{aligned}$$

for an arbitrary block b, where the second line follows by the fact that the encoding and decoding processes are identical in each block, and we made a definition in the last line for brevity[7]. Setting $B = N = N'$, we see that for any N', an error probability less than $N'\mathbb{P}\left(\mathcal{D}_{FB,w}(b,N')\right)$ can be achieved with rate $\frac{N'-1}{N'}R_i$. Therefore, in order to show that rate R_i is achievable, it is sufficient to show that $N\mathbb{P}\left(\mathcal{D}_{FB,w}(b,N)\right) \to 0$ as $N \to \infty$. Using the same arguments, one can show the same result for $D_{FB,s}(N)$ and $D_{NFB}(N)$, and define $\mathcal{D}_{FB,s}(b,N)$ and $\mathcal{D}_{NFB}(b,N)$ similarly.

Now we analyze the weak and strong interference regimes separately.

2.6.5.1 Weak interference

The following lemmas characterize the rate constraints for reliable communication with Rx1 for feedback and non-feedback strategies, respectively, under weak interference.

Lemma 2.2. *For weak interference at Rx1,* $N\mathbb{P}\left(\mathcal{D}_{FB,w}(b,N)\right) \to 0$ *as* $N \to \infty$ *if*

$$R_{1c} < I(X_{1f};Y_1|\underline{S},X_{1e},X_{2e}) - C_1 \tag{2.23}$$

$$R_{1p} < I(X_1;Y_1|\underline{S},X_{1f},X_{2f}) - C_1 \tag{2.24}$$

$$R_{2c} < I(X_{2f};Y_1|\underline{S},X_{2e},X_1) - C_1 \tag{2.25}$$

[7] The event $\mathcal{D}_{FB,w}$ is defined in the filtered probability space formed by the conditioning.

$$R_{1p} + R_{2c} < \min\Big\{I(X_1, X_{2f}; Y_1, U_1|\underline{S}, X_{1f}, X_{2e}) - 2C_1,$$

$$I(X_1, X_{2f}; Y_1|\underline{S}, X_{1c}, X_{2e}) - C_1\Big\} \tag{2.26}$$

$$R_1 + R_{2c} < I(X_1, X_{2f}; Y_1|\underline{S}, X_{1e}, X_{2e}) - C_1 \tag{2.27}$$

Proof. See Appendix A.2. □

Lemma 2.3. *For weak interference at Rx1,* $N\mathbb{P}\left(\mathcal{D}_{NFB}(b, N)\right) \to 0$ *as* $N \to \infty$ *if*

$$R_{1c} > I(X_{1f}; Y_1|\underline{S}, X_{1e}, X_{2e}) - C_1 \tag{2.28}$$

$$R_{1p} < I(X_1; Y_1|\underline{S}, X_{1f}, X_{2f}) - \kappa_2 \tag{2.29}$$

$$R_{2c} < I(X_{2f}; Y_1|\underline{S}, X_{2e}, X_1) - \kappa_2 \tag{2.30}$$

$$R_1 < I(X_1; Y_1|\underline{S}, X_{2f}, X_{1e}) - \kappa_2 \tag{2.31}$$

$$R_1 + R_{2c} < I(X_1, X_{2f}; Y_1|\underline{S}, X_{1e}, X_{2e}) - C_1 - \kappa_2 \tag{2.32}$$

Proof. See Appendix A.2. □

2.6.5.2 Strong interference

The following lemmas give the rate constraints for the feedback and non-feedback modes under strong interference at Rxi.

Lemma 2.4. *For strong interference at Rx1,* $N\mathbb{P}\left(\mathcal{D}_{FB,s}(b, N)\right) \to 0$ *as* $N \to \infty$ *if*

$$R_{2c} < I(X_{2f}; Y_1|\underline{S}, X_{1e}, X_{2e}) - C_1 \tag{2.33}$$

$$R_{1p} < I(X_1; Y_1|\underline{S}, X_{1f}, X_{2f}) - C_1 \tag{2.34}$$

$$R_1 < \min\{I(X_1; Y_1, U_2|\underline{S}, X_{1e}, X_{2f}), \tag{2.35}$$

$$I(X_1, X_{2e}; Y_1|\underline{S}, X_{1e}, X_{2c})\} - C_1 \tag{2.36}$$

$$R_1 + R_{2c} < I(X_1, X_{2f}; Y_1|\underline{S}, X_{1e}, X_{2e}) - C_1 \tag{2.37}$$

Proof. See Appendix A.2. □

Lemma 2.5. *For strong interference at Rx1,* $N\mathbb{P}\left(\mathcal{D}_{NFB}(b,N)\right) \to 0$ *as* $N \to \infty$ *if*

$$R_{2c} > I(X_{2f};Y_1|\underline{S},X_{1e},X_{2e}) - C_1 \tag{2.38}$$

$$R_{1p} < I(X_1;Y_1|\underline{S},X_{1f},X_{2f}) - \kappa_2 \tag{2.39}$$

$$R_{1p} + R_{2c} < I(X_1,X_{2f};Y_1|\underline{S},X_{1f},X_{2e}) - \kappa_2 \tag{2.40}$$

$$R_1 + R_{2c} < I(X_1,X_{2f};Y_1|\underline{S},X_{1e},X_{2e}) - C_1 - \kappa_2 \tag{2.41}$$

Proof. See Appendix A.2. □

2.6.6 Rate region evaluation

In this subsection, we first explicitly derive the set of achievable (R_1, R_2) pairs for linear deterministic and Gaussian models, from the results of the previous subsection.

We first find the conditions for decodability at Rx1 under weak interference. Recall that feedback mode is used at Rx1 only if (2.23) is satisfied; otherwise Han-Kobayashi decoding is performed. If we define $\underline{R} := (R_{1c}, R_{2c}, R_{1p})$, and

$$\begin{aligned}
\mathcal{R}^w_{FB} &:= \left\{\underline{R} : \text{(2.24)-(2.27) is satisfied}\right\},\\
\mathcal{R}^w_{NFB} &:= \left\{\underline{R} : \text{ (2.29)-(2.32) is satisfied}\right\},\\
\mathcal{R}^w_{d} &:= \left\{\underline{R} : \text{ (2.23) is satisfied}\right\},
\end{aligned}$$

then the set of rate points $\mathcal{R}^w$ that ensure decodability at Rx1 under weak interference contains

$$\begin{aligned}
\mathcal{R}^w &= (\mathcal{R}^w_{FB} \cap \mathcal{R}^w_d) \cup (\mathcal{R}^w_{NFB} \cap \mathcal{R}^{w,c}_d)\\
&\supseteq (\mathcal{R}^w_{NFB} \cap \mathcal{R}^w_{FB} \cap \mathcal{R}^w_d) \cup (\mathcal{R}^w_{NFB} \cap \mathcal{R}^w_{FB} \cap \mathcal{R}^{w,c}_d)\\
&= \mathcal{R}^w_{NFB} \cap \mathcal{R}^w_{FB}
\end{aligned}$$

where $\mathcal{R}^{w,c}_d$ is the complement of the set $\mathcal{R}^w_d$. Therefore, the rate constraints for decodability at Rx1 for the described strategy for weak interference are given by (2.24)-(2.27) and (2.29)-(2.32), for all joint distributions $\prod_{i=1}^{2} p(x_{ie})p(x_{ic})p(x_{ip})$, symbol-wise mappings $x_{if}(x_{ie}, x_{ic})$, $x_i(x_{if}, x_{ip})$, and $p(u_i|\tilde{v}_j)$, $(i,j) = (1,2),(2,1)$, consistent with the distortion constraints.

One can perform the same line of arguments as in the case of weak interference to show that the rate constraints for decodability at Rx1 for strong interference are given by (2.34)-(2.37) and (2.39)-(2.41), for all joint distributions $\prod_{i=1}^{2} p(x_{ie})p(x_{ic})p(x_{ip})$, symbol-wise mappings $x_{if}(x_{ie}, x_{ic})$, $x_i(x_{if}, x_{ip})$, and $p(u_i|\tilde{v}_j)$, $(i, j) = (1, 2), (2, 1)$, consistent with the distortion constraints.

Next, we consider linear deterministic and Gaussian models separately, and derive the achievable rate regions explicitly for both cases.

2.6.6.1 Rate region for linear deterministic model

To obtain the achievable rate region, we first evaluate the mutual information terms with specific input distributions. In particular, we choose the distributions and mappings

$$X_{ie} \sim Unif\left[\mathbb{F}_2^{n_{ji}}\right] \tag{2.42}$$

$$X_{ic} \sim Unif\left[\mathbb{F}_2^{n_{ji}}\right] \tag{2.43}$$

$$X_{ip} \sim Unif\left[\mathbb{F}_2^{(n_{ii}-n_{ji})^+}\right] \tag{2.44}$$

$$U_i = \widetilde{V}_j \tag{2.45}$$

$$\begin{aligned} x_{if} &: \mathbb{F}_2^{n_{ji}} \times \mathbb{F}_2^{n_{ji}} \rightarrow \mathbb{F}_2^{n_{ji}}, \\ x_i &: \mathbb{F}_2^{n_{ji}} \times \mathbb{F}_2^{(n_{ii}-n_{ji})^+} \rightarrow \mathbb{F}_2^{\max(n_{ii},n_{ji})}, \\ x_i &= \left[X_{if} \;\; X_{ip}\right]^T, \;\; x_{if}(a, b) = a + b \end{aligned} \tag{2.46}$$

for $(i, j) = (1, 2), (2, 1)$, where $Unif\,[\mathcal{A}]$ denotes uniform distribution over the set $\mathcal{A}$. Evaluating the mutual information terms of the previous subsection with this set of distributions, and applying Fourier-Motzkin elimination (see Appendix A.3 for details), we obtain the rate region given in (2.5)–(2.11).

2.6.6.2 Rate region for Gaussian model

Now we evaluate the rate constraints obtained in the previous section, and obtain the final achievable rate region. Assuming available power P_i at Txi, we assign the following input distributions, for $(i,j) = (1,2),(2,1)$:

$$X_{ie} \sim \mathcal{CN}(0, \frac{1}{2}P_i) \tag{2.47}$$

$$X_{ic} \sim \mathcal{CN}(0, \frac{1}{2}(1-P_{ip})P_i) \tag{2.48}$$

$$X_{ip} \sim \mathcal{CN}(0, \frac{1}{2}\min\left(\frac{1}{|h_{ji}|^2 P_i}, 1\right) P_i) \tag{2.49}$$

$$U_i|\widetilde{V}_j \sim \mathcal{CN}(\widetilde{V}_j, D_i) \tag{2.50}$$

$$x_{if} : \mathbb{C} \times \mathbb{C} \to \mathbb{C}, \;\; x_i : \mathbb{C} \times \mathbb{C} \to \mathbb{C},$$

$$x_{if}(a,b) = a+b, \;\; x_i(a,b) = a+b \tag{2.51}$$

where $D_i > 0$ are the distortion parameters. Using these input distributions, and applying Fourier-Motzkin elimination (See Appendix A.3 for details), we can show that the rate region (A.18)–(A.20), given in Appendix A.3, is achievable.

2.7 Converse

We now prove an outer bound region that exactly matches the region given in (2.5)–(2.11), and is within a constant gap of the region in (2.13)–(2.17).

The main idea between the novel bounds on R_1 and R_2 is based on a genie argument, where the receivers are provided with side-information about the messages. The bounds on $R_1 + R_2$, $2R_1 + R_2$ and $R_1 + 2R_2$ are proven through a channel enhancement technique, resembling the one used for the multiple-access channel in [KL13].

2.7.1 Bounds on R_1 and R_2

Since any outer bound for perfect feedback is also an outer bound for intermittent feedback, we have the perfect feedback bound

$$R_i \leq \max(n_{ii}, n_{ij}) \tag{2.52}$$

for linear deterministic model, and the bound

$$R_i \leq \sup_{0\leq\rho\leq 1} \log\left(1 + \mathsf{SNR}_i + \mathsf{INR}_i + 2\rho\sqrt{\mathsf{SNR}_i \cdot \mathsf{INR}_i}\right) \tag{2.53}$$

for Gaussian model, for $(i,j) = (1,2),(2,1)$, which are both proved in [ST11]. Next, we prove a novel bound for both models.

Without loss of generality, we focus on the bound on R_1. In order to prove the novel bound on R_1, the main idea is to provide $\left(W_2, \widetilde{V}_1^N\right)$ as side-information to Rx1. The intuition behind this particular choice is revealed when we consider the interference regime and operating point in which this bound is active. First, due to the structure of the capacity region, this bound is relevant only when the message (*i.e.*, the rate) of the interfering user is small enough. Hence, for that regime, W_2 does not carry too much information, and thus providing this to Rx1 still results in a tight outer bound. Second, note that this bound is only active in the strong interference regime, where feedback from Rx2 to Tx2 creates an alternative path for the transmission of W_1. Therefore, by forwarding this information, Tx2 indeed provides the information contained in $\widetilde{V}_1^N$ to Rx1.

Based on this idea, we prove the bound

$$R_i \leq n_{ii} + p_j\left(n_{ji} - n_{ii}\right)^+ \tag{2.54}$$

for linear deterministic model in Appendix A.4, and the bound

$$R_i \leq \log\left(1 + \mathsf{SNR}_i\right) + p_j \log\left(1 + \frac{\mathsf{INR}_j}{1 + \mathsf{SNR}_i}\right) \tag{2.55}$$

for the Gaussian model in Appendix A.5, for $(i,j) = (1,2),(2,1)$.

2.7.2 Bounds on $R_1 + R_2$, $2R_1 + R_2$ and $R_1 + 2R_2$

We have the perfect feedback outer bounds

$$R_i + R_j \leq \max\left(n_{ii}, n_{ij}\right) + \left(n_{jj} - n_{ji}\right)^+ \tag{2.56}$$

for linear deterministic model, and

$$R_i + R_j < \sup_{0\leq\rho\leq 1} \log\left(1 + \frac{(1-\rho^2)\mathsf{SNR}_i}{1 + (1-\rho^2)\mathsf{INR}_j}\right)$$

$$+\log\left(1+\mathsf{SNR}_j+\mathsf{INR}_j+2\rho\sqrt{\mathsf{SNR}_j\cdot\mathsf{INR}_j}\right) \tag{2.57}$$

for Gaussian model, for $(i,j)=(1,2),(2,1)$.

Next, we prove novel outer bounds on the capacity region. In order to prove these bounds, we first define a notion of enhanced channel. Considering our achievable scheme, feedback can be interpreted as a mechanism for the receivers to separate the interference and the intended signal, to the extent allowed by the erasure probability in the feedback channel. In the weak interference regime, this allows the receiver to cancel the interference. In the strong interference regime, through the alternate path created by the interfering user, it allows the reception of additional information about the intended message. Therefore we consider an enhanced channel where the receivers observe the interference and the intended signal individually whenever the feedback is available, and their sum otherwise. In addition to this enhancement, we provide Rxi with the side-information of V_i^N as well, as was done in [ETW08]. To make this more precise, we consider the two models separately.

2.7.2.1 Linear deterministic model

We define the enhanced linear deterministic channel with intermittent feedback by the following equations

$$\breve{Y}_i=\begin{cases} Y_i, & \text{if } S_i=0\\ \left(\mathbf{H}_{ii}X_i,V_j\right), & \text{if } S_i=1\end{cases}$$

for $(i,j)=(1,2),(2,1)$, where $\breve{Y}_i$ is the channel output of the enhanced channel at Rxi, Y_i is the channel output of the original channel, and X_i and V_j are as defined for the original channel. The output of the feedback channel is given by $\widetilde{Y}_i=S_iY_i$, *i.e.*, the same as the original channel. Note that any scheme that achieves arbitrarily small error probability in the original channel can also achieve arbitrarily small error probability for the enhanced channel, using the fact that $Y_i=\mathbf{H}_{ii}X_i+V_j$. This means that the capacity region of the original channel is a subset of that of the enhanced channel, and we can derive an outer bound for the enhanced channel instead.

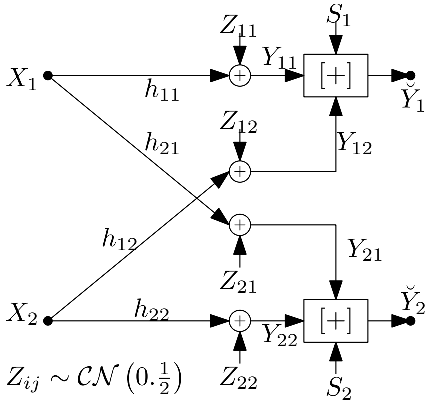


Figure 2.11: The enhanced channel for Gaussian model. The block [+] is a conditional adder, which outputs the sum of the other two inputs if $S_i = 0$, and outputs the two inputs separately otherwise.

It is easy to see that this enhancement is equivalent to providing the Rxi with $\widetilde{V}_2^N$, since for time slots where $S_i = 1$, Rxi can use this information to individually obtain the interference and the intended symbol.

Using the channel enhancement technique, we arrive at the following outer bounds on the capacity region of the linear deterministic interference channel with intermittent feedback, which are explicitly proved in Appendix A.4.

$$\begin{aligned}
R_1 + R_2 \leq\ & \max\left\{n_{12}, (n_{11} - n_{21})^+\right\} + \max\left\{n_{21}, (n_{22} - n_{12})^+\right\} \\
& + p_1 \min\left\{n_{12}, (n_{11} - n_{21})^+\right\} + p_2 \min\left\{n_{21}, (n_{22} - n_{12})^+\right\} \quad (2.58) \\
2R_1 + R_2 \leq\ & \max\left(n_{11}, n_{12}\right) + \max\left\{n_{21}, (n_{22} - n_{12})^+\right\} + (n_{11} - n_{21})^+ \\
& + p_2 \min\left\{n_{21}, (n_{22} - n_{12})^+\right\} \quad (2.59) \\
R_1 + 2R_2 \leq\ & \max\left(n_{22}, n_{21}\right) + \max\left\{n_{12}, (n_{11} - n_{21})^+\right\} + (n_{22} - n_{12})^+ \\
& + p_1 \min\left\{n_{12}, (n_{11} - n_{21})^+\right\} \quad (2.60)
\end{aligned}$$

2.7.2.2 Gaussian model

Next, we extend the enhanced channel idea to the Gaussian model. In this case, while splitting the interference and the intended signal, we also split the noise evenly between these two variables (see Figure 2.11). Specifically, we consider the channel defined by the equations

$$\breve{Y}_i = \begin{cases} \bar{Y}_i, & \text{if } S_i = 0 \\ \left(Y_{ii}, Y_{ij}\right), & \text{if } S_i = 1 \end{cases}$$

for $(i,j) = (1,2),(2,1)$, where $\breve{Y}_i$ is the output of the enhanced channel, and

$$\begin{aligned} Y_{ii} &= h_{ii}X_i + Z_{ii} \\ Y_{ij} &= h_{ij}X_j + Z_{ij} \\ \bar{Y}_i &= Y_{ii} + Y_{ij} = h_{ii}X_i + h_{ij}X_j + \bar{Z}_i \end{aligned}$$

with Z_{ij}, Z_{ii} are independent and distributed with $\mathcal{CN}(0, \frac{1}{2})$, and we define $\bar{Z}_i = Z_{ii} + Z_{ij}$. The output of the feedback channel at Txi is given by $S_i\bar{Y}_i = S_i \cdot (Y_{ii} + Y_{ij})$, *i.e.*, the same as the original channel. It is worth noting that unlike the linear deterministic case, this enhancement is not equivalent to providing Rxi with $\widetilde{V}_j^N$, since giving this side-information allows the receiver to completely cancel the noise for some time slots, resulting in an infinitely loose bound.

Let $\mathcal{C}_e(p_1, p_2)$ denote the capacity region of the enhanced channel.

The next lemma shows that the capacity region of the enhanced channel indeed dominates the original one.

Lemma 2.6. *For all* $0 \leq p_1, p_2 \leq 1$*,*

$$\mathcal{C}_G(p_1, p_2) \subseteq \mathcal{C}_e(p_1, p_2)$$

Proof. The proof has two steps. First, we consider an intermediate channel, with capacity region $\mathcal{C}_i(p_1, p_2)$, and the channel output at Rxi is given by

$$Y_i = h_{ii}X_i + h_{ij}X_j + \bar{Z}_i$$

for $(i,j) = (1,2),(2,1)$, where $\bar{Z}_i = Z_{ii} + Z_{ij}$ is the sum of two independent $\mathcal{CN}(0, \frac{1}{2})$ random variables as in the enhanced channel. Since $\bar{Z}_i$ and Z_i (the noise in the original channel) have the same probability distribution and are both i.i.d. processes across time and across users, the joint distribution of the channel $p(y_1, y_2|x_1, x_2)$ is identical for both channels, and hence they have the same feedback capacity region, *i.e.*, $\mathcal{C}_i(p_1, p_2) = \mathcal{C}_G(p_1, p_2)$.

Next, comparing the intermediate channel and the enhanced channel, we note that any rate pair (R_1, R_2) achievable in the intermediate channel is also achievable for the enhanced channel using the same pair of codes, using the fact that $\bar{Y}_i = Y_{ii} + Y_{ij}$. Therefore, $\mathcal{C}_i(p_1, p_2) \subseteq \mathcal{C}_e(p_1, p_2)$, which completes the proof. □

Remark 2.4. *We note that a similar channel enhancement technique has been applied before by Khisti and Lapidoth [KL13], for Gaussian multiple-access channel with intermittent feedback. In that work, the variances of the random variables Z_{ii} and Z_{ij} are not fixed, but are arbitrary, subject to the constraint that they sum to one. Although one can optimize over the noise variances in order to obtain the tightest bound, this only results in a small and constant improvement. Hence, for simplicity, we stick to the fixed variance of $\frac{1}{2}$ for the noise variables of the enhanced channel.*

Using Lemma 2.6, we can instead prove outer bounds for the enhanced channel. In Appendix A.5, we prove the following bounds.

$$\begin{aligned}
R_1 + R_2 \leq{}& \log\left(1 + \mathsf{INR}_1 + \frac{\mathsf{SNR}_1 + 2\sqrt{\mathsf{SNR}_1 \cdot \mathsf{INR}_1}}{1 + \mathsf{INR}_2}\right) \\
&+ \log\left(1 + \mathsf{INR}_2 + \frac{\mathsf{SNR}_2 + 2\sqrt{\mathsf{SNR}_2 \cdot \mathsf{INR}_2}}{1 + \mathsf{INR}_1}\right) \\
&+ p_1 \log\left(\frac{(1 + 2\mathsf{INR}_1)\left(1 + \frac{\mathsf{SNR}_1}{\mathsf{INR}_2 + \frac{1}{2}}\right)}{1 + \mathsf{INR}_1 + \frac{\mathsf{SNR}_1 + 2\sqrt{\mathsf{SNR}_1 \cdot \mathsf{INR}_1}}{1 + \mathsf{INR}_2}}\right) \\
&+ p_2 \log\left(\frac{(1 + 2\mathsf{INR}_2)\left(1 + \frac{\mathsf{SNR}_2}{\mathsf{INR}_1 + \frac{1}{2}}\right)}{1 + \mathsf{INR}_2 + \frac{\mathsf{SNR}_2 + 2\sqrt{\mathsf{SNR}_2 \cdot \mathsf{INR}_2}}{1 + \mathsf{INR}_1}}\right) \qquad (2.61) \\
2R_1 + R_2 \leq{}& \log\left(1 + \mathsf{SNR}_1 + \mathsf{INR}_1 + 2\sqrt{\mathsf{SNR}_1 \cdot \mathsf{INR}_1}\right) + \log\left(1 + \frac{\mathsf{SNR}_1}{\frac{1}{2} + \mathsf{INR}_2}\right) \\
&+ \log\left(1 + \mathsf{INR}_2 + \frac{\mathsf{SNR}_2 + 2\sqrt{\mathsf{SNR}_2 \cdot \mathsf{INR}_2}}{1 + \mathsf{INR}_1}\right) +
\end{aligned}$$

$$p_2 \log\left(\frac{(1+2\mathsf{INR}_2)\left(1+\frac{\mathsf{SNR}_2}{\mathsf{INR}_1+\frac{1}{2}}\right)}{1+\mathsf{INR}_2+\frac{\mathsf{SNR}_2+2\sqrt{\mathsf{SNR}_2\cdot\mathsf{INR}_2}}{1+\mathsf{INR}_1}}\right) \tag{2.62}$$

$$R_1+2R_2 \leq \log\left(1+\mathsf{SNR}_2+\mathsf{INR}_2+2\sqrt{\mathsf{SNR}_2\cdot\mathsf{INR}_2}\right)+\log\left(1+\frac{\mathsf{SNR}_2}{\frac{1}{2}+\mathsf{INR}_1}\right)$$
$$+\log\left(1+\mathsf{INR}_1+\frac{\mathsf{SNR}_1+2\sqrt{\mathsf{SNR}_1\cdot\mathsf{INR}_1}}{1+\mathsf{INR}_2}\right)+$$
$$p_1 \log\left(\frac{(1+2\mathsf{INR}_1)\left(1+\frac{\mathsf{SNR}_1}{\mathsf{INR}_2+\frac{1}{2}}\right)}{1+\mathsf{INR}_1+\frac{\mathsf{SNR}_1+2\sqrt{\mathsf{SNR}_1\cdot\mathsf{INR}_1}}{1+\mathsf{INR}_2}}\right) \tag{2.63}$$

2.8 Discussion and Extensions

We considered the interference channel with intermittent feedback, and derived an approximate characterization of the capacity region under Gaussian model, as well as an exact characterization for the linear deterministic case. The result shows that even intermittent feedback provides multiplicative gain in capacity in interference channels. The achievability result was based on quantize-map-forward relaying at the transmitters, and the outer bound result was based on a channel enhancement technique.

In this work, we considered short messaging, *i.e.*, a new message is sent at every block of transmission. An alternate approach one could try is long messaging, where the transmitters send codewords describing the same message at every block, and the receivers jointly decode all blocks to recover the message. The clear advantage of short messaging approach is better delay performance, since each message is decoded immediately after the transmission of the corresponding block, instead of waiting for the end of the entire transmission. However, combined with forward decoding, the rate region achievable by this strategy cannot approximate the entire capacity region by itself, as can be seen from the results of Section 2.6; we need to take the union with Han-Kobayashi rate region to approximate the entire capacity region. This is because while decoding block b, part of the message of block $b+1$ is jointly decoded by treating the interference of block $b+1$ as noise, which limits the rate in certain operating points. Hence, long-messaging approach would remove the need for taking union with Han-Kobayashi region and simplify the proof, since all blocks are jointly decoded. Such an approach has been taken in [Zai14] to derive an inner bound on the capacity region of interference channels with generalized feedback, which overlaps with the capacity region (2.5)–(2.11) for the special case of linear deterministic IC with intermittent feedback.

The extension to parallel channels is carried out for the special case of identical subchannels in this work. An important generalization can be the case where the channel gains of the subchannels are not necessarily the same. The main obstacle in generalizing our achievable scheme to this case is that it distinguishes the cases of weak and strong interference, although such a separation is not possible for vector channels. Again, long-messaging can be a strategy to circumvent this issue, since it removes the need for making such a distinction between weak and strong interference regimes [Zai14], albeit at the cost of a much larger delay.

Another important extension could be to the additive white Gaussian noise (AWGN) feedback model of [LTM12]. Since this model assumes passive feedback as well, our quantize-map-forward based scheme can be directly applied to to this channel model. The results of this chapter indicate that quantize-map-forward, as a feedback strategy, might be a promising candidate as an approximately-capacity-achieving scheme for AWGN feedback model. However, this investigation is not the focus of this paper, and is left as future work.

CHAPTER 3

Opportunistic Scheduling in Full-duplex Cellular Networks

3.1 Introduction

Full-duplex wireless communication is becoming closer to reality, in light of recent experimental results demonstrating its feasibility [DS10, CJS10]. Especially the development of massive MIMO can create opportunities for full-duplex communication, since all implementations of full-duplex use multiple antennas. The first application of full-duplex in a practical system is expected to be in base stations instead of mobile devices, due to relative flexibility in design. Since mobile devices remain half-duplex, the uplink-downlink nature of a cellular system is retained, even when the base station is full-duplex. By serving uplink and downlink simultaneously over the same band, a full-duplex cellular system might have the potential to double the spectral efficiency. However, in order to realize this gain, one is immediately faced with a challenge that is not present in half-duplex systems: uplink-to-downlink interference.

The problem of uplink-to-downlink interference management in full-duplex systems has been considered in [SDS13] and [BS13] with several interference management strategies proposed, based on interference alignment or message splitting. However, such sophisticated solutions, which require very tight coordination between nodes and a large amount of overhead to exchange channel information, are not well-suited for large-scale, dynamic networks, where a large number of high-mobility nodes can have rapidly changing channel states.

In this work we propose and analyze a solution that is much more suited to such dynamic networks. In fact, the solution is explicitly centered around exploiting this dynamic nature of the network, rather than attempting to manage it. This is based on adapting the opportunistic beamforming and scheduling [VTL02] ideas to this new scenario that arises in full-duplex networks with dynamically varying channel states. This approach enables us to design an opportunistic joint uplink-downlink scheduling algorithm that, in a homogeneous network with a large number of half-duplex users and a multi-antenna full-duplex base station, asymptotically achieves the sum of the capacities of the isolated uplink and downlink systems, thus doubling the spectral efficiency. The main idea underlying the result is to apply random transmit and receive beamforming at the base station [VTL02], and exploit the multiuser diversity in the system to schedule the uplink and downlink users that conflict the least with each other. Such a solution also has the advantage of requiring much less channel training overhead, as we will explore.

The chapter is organized as follows. Section 3.2 reviews existing work on the problem and contrasts our approach with it. Section 3.3 presents our mathematical model and notation. Section describes our proposed opportunistic joint uplink-downlink scheduling scheme, and presents our main results. Section discusses the case of clustered networks, how the presented approach might fail in such networks, and discusses user cooperation as a potential solution for this scenario. The proofs of these results are provided in Appendix B.

3.2 Related Work and Contributions

Many authors (including[SH05, YG06], among others) have studied the problem of MIMO downlink scheduling in the many-user regime, and it has been demonstrated that the same scaling law as the optimal dirty-paper coding sum rate can be achieved via beamforming with scheduling. It was also shown that the gap between the sum rate achievable with beamforming with scheduling and dirty-paper coding goes to zero [BK08, WLZ08]. There has also been works that explore how to exploit multiuser diversity in the presence of interference, under multi-cell downlink [LL06], and spectrum sharing cognitive radio [BCJ09] scenarios. However, the schemes developed in these works are either intended for an isolated downlink system, or fail to provide any theoretical performance guarantees on the overall system throughput when translated into a full-duplex system, where the goal is to *simultaneously* extract uplink and downlink multiuser diversity gains while dealing with the uplink-to-downlink interference. Based on the existing literature on opportunistic scheduling, it is not clear whether downlink sum rate optimality through scheduling is maintained in the presence of uplink interference, especially when the uplink sum rate optimality is also sought.

We have two main contributions in this work. First, we show that the asymptotic sum rate optimality in both uplink and downlink can be maintained individually, even in the presence of uplink-to-downlink interference. To achieve this, we develop a simple opportunistic scheduling algorithm based on random beamforming. The algorithm does not require the base station or the uplink users to have channel information about the interference links. Moreover, very little CSI is required at the base station due to random beamforming. We also show that the spatial multiplexing gain offered by the multiple antennas is retained in the full-duplex system when the number of antennas scale logarithmically with the number of users, as was shown for isolated downlink in [SH05].

This asymptotic decoupling result relies on there being sufficient channel diversity in the network. Although a homogeneous network with i.i.d. fading links provides sufficient diversity for this purpose, such diversity may not be present in a real network. For instance, there might be areas in a cell where users are densely clustered, and some other areas that are mostly deserted, resulting in a lack of sufficiently rich channel conditions. In a full-duplex system, in addition to diversity in channels to and from the base station, diversity in interference links is also required to realize the multiuser diversity gains. Our second contribution is to show that for a simple class of heterogeneous networks, it is not possible to achieve such gains, by deriving an upper bound on the achievable sum rate. In particular, the gap between the achievable sum rates of the full-duplex system and the decoupled system grows linearly with the number of antennas and logarithmically with downlink SNR. Although our heterogeneous network model is rather simple, it features the key property of the lack of channel

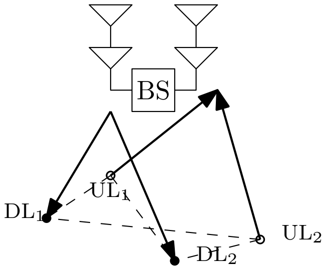


Figure 3.1: A cellular system with a full-duplex base station with $M = 2$ antennas and $n = 2$ uplink and downlink half-duplex users. Uplink users are represented with white dots, downlink users are represented with black dots, and the interference links are represented with dashed lines.

diversity. To address this limitation in heterogeneous networks, we demonstrate through an example that establishing device-to-device cooperation over orthogonal side-channels can be effective.

3.3 Model and Notation

We consider a cellular system with a single full-duplex base station, equipped with M antennas for uplink and M antennas for downlink communication (see Figure 1). We assume there are n uplink, and n downlink half-duplex users, each with a single antenna, requesting communication over the same band. We assume the base station is able to completely cancel self-interference, but the uplink transmission interferes with the received signal at the downlink users.

We first consider a homogeneous network, where all links in the network, including the interference links, are are generated i.i.d. from a $\mathcal{CN}(0,1)$ distribution; but once drawn, they remain fixed throughout the duration of transmission.

The uplink channel is described by the equation

$$\bar{y} = \bar{H}_n \bar{x} + \bar{z},$$

where $\bar{y} \in \mathbb{C}^{M\times 1}$ is the vector of channel outputs at the base station, $\bar{x} \in \mathbb{C}^{n\times 1}$ is the vector of channel inputs from n uplink users, subject to a per-user block power constraint $\frac{1}{T}\sum_{t=1}^{T}|\bar{x}_k[t]|^2 \leq \bar{P}$ for a block length of T, for $k = 1,\ldots,n$,

$$\bar{H}_n = \left[\begin{array}{ccc} \bar{h}_1 & \ldots & \bar{h}_n \end{array}\right] \in \mathbb{C}^{M\times n}$$

is the matrix of channel gains, with each element generated i.i.d. according to $\mathcal{CN}(0,1)$, and $\bar{z} \sim \mathcal{CN}(\mathbf{0},\mathbf{I}_M)$ is the vector of complex Gaussian noise. Throughout the chapter, we use the bar notation whenever a variable pertains to the uplink transmission, whereas we use plain letters for variables pertaining to downlink transmission, including the uplink-to-downlink interference link gain.

The downlink of the system is described by

$$y = H_n^* x + G_n \bar{x} + z,$$

where $y \in \mathbb{C}^{n\times 1}$ is the vector of channel output at the n downlink users, $x \in \mathbb{C}^{M\times 1}$ is the vector of channel inputs from M antennas, subject to a total block power constraint $\frac{1}{T}\sum_{t=1}^{T} x^*[t]x[t] \leq P$, $H_n \in \mathbb{C}^{n\times M}$ is the matrix of channel gains and $G_n \in \mathbb{C}^{n\times n}$ is the matrix of interference link gains, with each element of the matrices generated i.i.d. according to $\mathcal{CN}(0,1)$, and $z \sim \mathcal{CN}(\mathbf{0},\mathbf{I}_n)$ is the vector of complex Gaussian noise.

The set of all link gains in the network is denoted by $\mathcal{H}_n = \left(\bar{H}_n, H_n, G_n\right)$. Further, the rate of ith downlink (uplink) user is denoted by $R_i\left(\mathcal{H}_n\right)$ ($\bar{R}_i\left(\mathcal{H}_n\right)$), and the sum uplink and downlink rates are denoted by

$$\bar{\mathbf{R}}_n\left(\mathcal{H}_n\right) = \sum_{i=1}^{n} \bar{R}_i\left(\mathcal{H}_n\right), \quad \mathbf{R}_n\left(\mathcal{H}_n\right) = \sum_{i=1}^{n} R_i\left(\mathcal{H}_n\right).$$

All logarithms throughout the chapter are assumed to be in base e. We also define

$$[n] := \{k \in \mathbb{N} : 1 \leq k \leq n\}.$$

We impose the constraint that at most M uplink users can simultaneously transmit to the base station, *i.e.*, the vector $\bar{x}$ can only have M non-zero elements per time slot[1].

[1] This constraint is placed to prevent total uplink power in the system from growing unboundedly.

3.4 Opportunistic Scheduling for Homogeneous Networks

3.4.1 Opportunistic scheduling

We consider an opportunistic scheduling algorithm that performs random beamforming [VTL02] independently for uplink and downlink, and schedules the users whose channels best fit to the current beamforming patterns, and least interfere with each other. In particular, the base station first constructs a random unitary matrix $\bar{\Phi}$ and multiplies this with the received uplink channel output

$$\bar{\Phi}^*\bar{y} = \bar{\Phi}^*\bar{H}_n\bar{x} + \bar{\Phi}^*\bar{z}$$

Note that since $\bar{\Phi}$ is unitary, $\bar{\Phi}^*\bar{z}$ is still distributed as $\mathcal{CN}(\mathbf{0}, \mathbf{I}_M)$. We consider the scheduling of M uplink users for transmission at a given time. In particular, each element of the vector $\bar{\Phi}^*\bar{y}$ is assigned to a user, and the signal of that user is decoded from this component of the effective channel output, treating inter-stream interference as noise[2]. Note that this can be viewed as choosing an $M \times M$ submatrix of $\bar{\Phi}^*\bar{H}_n$. We use the following rule to choose the user $U_m \in \{1, \dots, n\}$ assigned to the mth stream:

$$\bar{U}_m = \arg\min_{k \in \bar{S}_m} \left|\bar{\phi}_m^* \bar{h}_k\right|^2$$

where

$$\bar{S}_m = \{1 \leq k \leq n : \left|\bar{\phi}_m^* \bar{h}_k\right|^2 \leq \epsilon_n, \forall r \neq m\}$$

for some ϵ_n such that $\epsilon_n \to 0$ as $n \to \infty$[3], where $\bar{\phi}_m$ is the mth column of $\bar{\Phi}$. Note that this scheduling algorithm first determines a set of candidate users for stream m, by eliminating all users whose interference to any other stream exceeds a certain threshold, and then picks

[2] Although successive cancellation decoding can also be used, this does not improve our main result, hence we treat interference as noise for simplicity.

[3] Note that ϵ_n must be scaled down slow enough to ensure that $\left|\bar{S}_m\right| > 0$ with high probability. The exact scaling of ϵ_n is left unspecified here, but in the proof of our main result, it will be seen that $\epsilon_n = O\left(\frac{1}{\log n}\right)$ is a good choice.

the user whose channel has the largest projection along the mth beamforming vector in the candidate set. We denote the set of uplink users scheduled in this way as $\bar{\mathcal{T}} = \{\bar{U}_m\}_{m=1}^{M}$.

Next, we consider the scheduling of downlink users, based on the uplink user selection. As in the uplink case, we begin by generating a random beamforming matrix Φ, and precode the transmitted signal with it, so that the vector of received signals at the n downlink users becomes

$$y = H_n^* \Phi x + G_n \bar{x} + z,$$

We use the following rule to choose the user $U_m \in \{1, \ldots, n\}$ assigned to the mth stream:

$$U_m = \arg \min_{k \in S_m} |\phi_m^* h_k|^2$$

where

$$\begin{aligned} S_m = \{1 \leq k \leq n : |\phi_m^* h_k|^2 \leq \epsilon_n, \forall r \neq m; \\ |g_{kj}|^2 \leq \epsilon_n, \forall j \in \bar{\mathcal{T}}\} \end{aligned}$$

for the same ϵ_n sequence as in the downlink, where ϕ_m is the mth column of Φ, *i.e.*, the candidate set of users for stream m are the users who receive bounded uplink interference as well as bounded inter-stream interference. We denote the set of uplink users scheduled in this way as $\mathcal{T} = \{U_m\}_{m=1}^{M}$.

Remark 3.1. *Originally, random beamforming was considered for downlink communication in order to artifically induce channel variations and realize the multiuser diversity effect [VTL02]. However, in a full-duplex system, one also needs to induce variations in the level of interference to each user to extract this gain. Since each user has a single antenna, this is not possible through random beamforming at the uplink user side. However, one can still perform receive beamforming for uplink at the base station, which results in scheduling a different subset of users at each time slot, which in turn causes variations in the aggregate interference strength observed at each downlink user, as desired.*

Remark 3.2. *Note that the base station or the uplink users do not require the channel knowledge of the interfering links for this scheme to work. If the downlink users are able to track the uplink interference strength they receive (which can potentially be arranged by overhearing the uplink pilots), they can send* SNR *feedback for their own channels only if the current interference level is below the threshold, and the base station can perform scheduling based only on this information.*

3.4.2 Asymptotic sum capacity for fixed number of antennas

Define the achieved uplink and downlink gaps from individual uplink and downlink capacities as

$$\begin{aligned}\bar{\eta}\left(\mathcal{H}_n\right) &:= \bar{\mathbf{C}}_n^{\text{MAC-M}}\left(\mathcal{H}_n\right) - \bar{\mathbf{R}}_n\left(\mathcal{H}_n\right) \\ \eta\left(\mathcal{H}_n\right) &:= \mathbf{C}_n^{\text{BC}}\left(\mathcal{H}_n\right) - \mathbf{R}_n\left(\mathcal{H}_n\right)\end{aligned}$$

respectively, where $\bar{\mathbf{C}}_n^{\text{MAC-M}}\left(\mathcal{H}_n\right)$ is the sum capacity of the multi-antenna MAC formed by considering the isolated uplink system, subject to the constraint that only M users can transmit simultaneously, and $\mathbf{C}_n^{\text{BC}}\left(\mathcal{H}_n\right)$ is the sum capacity of the multi-antenna broadcast channel formed by isolating downlink system, achieved by dirty-paper coding [WSS06].

Clearly, $\bar{\mathbf{C}}_n^{\text{MAC-M}}\left(\mathcal{H}_n\right) + \mathbf{C}_n^{\text{BC}}\left(\mathcal{H}_n\right)$ is an upper bound on the sum rate $\mathbf{R}_n\left(\mathcal{H}_n\right) + \bar{\mathbf{R}}_n\left(\mathcal{H}_n\right)$ achievable in the full-duplex system. Our main result is that in a homogeneous network, this upper bound is asymptotically achievable as the number of users n goes to infinity. This is more precisely stated in the following theorem.

Theorem 3.1. *For any $\delta > 0$,*

$$\lim_{n\to\infty} \mathbb{P}\left(\bar{\eta}\left(\mathcal{H}_n\right) + \eta\left(\mathcal{H}_n\right) > \delta\right) = 0$$

See Appendix B.1 for proof.

Theorem 3.1 implies that for a homogeneous network with sufficiently many users, the uplink-to-downlink interference can be mitigated through proper user scheduling to the extent that the uplink and downlink systems gets asymptotically decoupled. The main idea underlying this result is to exploit *multiuser diversity*, in terms of both the richness in the channel vectors to and from the base station, and richness in the strength of the interfering link.

Another important point in Theorem 3.1 is that not only does the sum rate has the same scaling law as the decoupled system (which scales as $M \log\log n$ for both uplink and downlink, as in the isolated uplink and downlink systems [SH05]), but the *additive* gap between the decoupled system sum capacity and the achievable full-duplex sum rate goes to zero. A similar behavior has been observed before for MIMO broadcast channels, where it has been shown that the achievable rate difference between zero-forcing beamforming and dirty-paper coding goes to zero as $n \to \infty$ [BK08]. Our result shows that through random beamforming, the same result can be obtained for simultaneous uplink and downlink, in the presence of uplink-to-downlink interference.

3.4.3 Scaling the number of antennas

An important assumption in Theorem 3.1 is that the number of antennas remain fixed as n grows. This is a crucial assumption, since as M grows, one would need to schedule a growing number of users simultaneously in order to realize the full multiplexing gain of the system, which would result in increasing uplink-to-downlink and inter-stream interference. Hence, an important question is whether a similar result would hold in the case where M is scaling. In [SH05], it is shown that for an isolated downlink system, the spatial multiplexing gain can be preserved if M is scaled like $O(\log n)$. Here, we show a similar result for the full-duplex system, which is given in the following theorem.

Theorem 3.2. *If* $\lim_{n\to\infty} \frac{M}{\log n} = \alpha$ *for some* $\alpha > 0$*,*

$$\lim_{n\to\infty} \frac{\bar{\mathbf{R}}_n(\mathcal{H}_n) + \mathbf{R}_n(\mathcal{H}_n)}{2M} = \beta$$

for some $\beta > 0$*, almost surely.*

See Appendix B.2 for proof.

Hence, even when the number of antennas grows to support the large number of users, the full sum degrees of freedom of the system can still be fully utilized despite the growing level of uplink interference, provided that the number of antennas does not scale faster than logarithmically in n.

3.5 D2D Cooperation for Clustered Full-Duplex Networks

The main idea underlying the result in Theorem 3.1 was to exploit the channel richness in the network to asymptotically decouple the uplink and downlink transmissions. We have seen that the homogeneous model described in Section 3.3 provides sufficient richness for this purpose. However, such homogeneity may not present in an actual network. Instead, users may be densely clustered in certain areas, and sparsely located in others. In such a scenario, it may not be possible to simultaneously approach uplink and downlink sum capacities, since the lack of channel diversity might force one to schedule an uplink-downlink user pair with significant interference in between.

In order to study this opposite regime, we consider a specific class of clustered networks that takes such non-homogeneity to the extreme, and prove that it is not possible to achieve the sum capacity of the decoupled system in such networks. Although the model of networks that we consider is rather specific, the main insight derived from this model might apply to more general heterogeneous networks.

3.5.1 Heterogeneous model

We consider a network with M clusters (see Figure 2), hosting a total of n uplink and n downlink users that are uniformly distributed among them. We consider a simplified model where each cluster is assigned a spatial direction h_i, with $h_i^* h_j = 0$ for $i \neq j$, and $\|h_i\| = h$ for all $1 \leq i \leq M$. We assume that all users (both uplink and downlink) within a cluster has the identical channel vector h_i. Further, we assume an all-or-none interference model, *i.e.*, if $\kappa(i)$ denotes the cluster index of user i, then the interference link gain magnitude from user

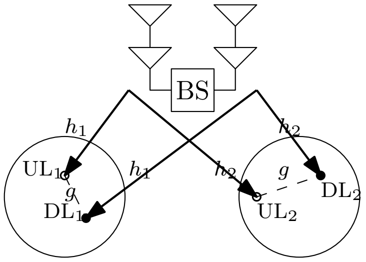


Figure 3.2: A heterogeneous cellular system with a full-duplex base station with $M = 2$ clusters and $n = 2$ uplink and downlink half-duplex users. The uplink and downlink users within the same cluster have the same channel, and users in different clusters do not interfere with each other.

j to user i is given by

$$|G_{ij}| = \begin{cases} g, & \text{if } \kappa(i) = \kappa(j) \\ 0, & \text{otherwise} \end{cases}$$

As in the homogeneous case, we impose the constraint that at most M uplink users can transmit simultaneously. Although this is a very simplified model, the unusual way in which the multi-antenna MAC and the BC interact with each other still makes this a non-trivial problem.

Henceforth, this model will be referred to as a (M, h, g)-clustered network. Next, we derive an upper bound on the sum capacity of the network.

3.5.2 Sum rate upper bound and the gap from the decoupled system capacity

Theorem 3.3. *If* $\left(\bar{\mathbf{R}}_n, \mathbf{R}_n\right)$ *is an achievable rate pair in a* $(M, h, g)-$*clustered network, then*

$$\begin{aligned} \bar{\mathbf{R}}_n + \mathbf{R}_n < {} & M \log\left(1 + \frac{h^2 \bar{P}}{1 + g^2 \bar{P}}\right) \\ & + M \log\left(1 + h^2 \frac{P}{M} + g^2 \bar{P} + 2gh\sqrt{\frac{\bar{P} P}{M}}\right) \end{aligned}$$

Proof. Let $y^{(m)}$ and $z^{(m)}$ denote the vector of channel outputs and the vector of noise at the users in cluster m. Since the downlink users do not cooperate, the capacity does not depend on the covariance matrix Σ_z of the noise at the downlink, as long as $\Sigma_z \geq 0$ and the diagonal consists of 1's [Sat78]. Hence, we assume that within the same cluster, all downlink users are subject to the same noise process, *i.e.*, $z_t^{(m)} \sim \mathcal{CN}(\mathbf{0}, \mathbf{1}\mathbf{1}^T)$, where $\mathbf{1}$ is the all ones vector. The noise processes at different clusters are independent[4]. Under these assumptions, using a genie-aided argument we show in Appendix B.3 that for a block length of N,

$$\begin{aligned} N\left(\bar{\mathbf{R}}_n + \mathbf{R}_n\right) < & \max_{\sum_{m=1}^{M} k_{m,t} \leq M} \quad \max_{\frac{1}{N}\sum_{(m,t)\in\mathcal{N}} P_{m,t} \leq P} \\ & \sum_{t=1}^{N}\sum_{m=1}^{M} \log\left(1 + h^2 P_{m,t} + k_{m,t} g^2 \bar{P} + 2gh\sqrt{k_{m,t} P_{m,t} \bar{P}}\right) \\ & + \log\left(1 + \frac{k_{m,t} h^2 \bar{P}}{1 + k_{m,t} g^2 \bar{P}}\right), \end{aligned} \tag{3.1}$$

where $k_{m,t}$ is the number of uplink users scheduled from cluster m at time t, $P_{m,t}$ is the power allocated to mth channel at the base station at time t, and $\mathcal{N} := [M] \times [N]$. It can be verified that the log terms in (3.1) are concave and monotonically increasing in $(P_{m,t}, k_{m,t})$, and hence the result follows by Jensen's inequality. □

It is easy to see that the sum of the isolated uplink and downlink capacities for a (M, h, g)-clustered network is given by

$$\bar{C}^{\text{MAC-M}} + C^{\text{BC}} = M \log\left(1 + h^2 \frac{P}{M}\right) + M \log\left(1 + h^2 \bar{P}\right) \tag{3.2}$$

Define the gaps form isolated systems, η and $\bar{\eta}$ as in the homogeneous case. Also set $\mathsf{SNR} := h^2 \frac{P}{M}$, $\mathsf{SNR}^{\alpha} := g^2 \bar{P}$, $\mathsf{SNR}^{\beta} := h^2 \bar{P}$. The following corollary of Theorem 3.3 characterizes the scaling behavior of $\eta + \bar{\eta}$.

Corollary 3.1. *For a (M, h, g)-clustered network with number of users $n \geq M$,*

$$\lim_{\mathsf{SNR}\to\infty} \frac{\eta + \bar{\eta}}{M \log \mathsf{SNR}} \geq 1$$

[4] For a general broadcast channel, it is known that assuming independent noise processes gives a loose bound while using Sato upper bound [VT03]; however, this does not matter in this case, since the links are orthogonal.

See Appendix B.4 for proof.

Note that this is the gap between an *upper bound* on the sum capacity and the decoupled system capacity. Hence, regardless of the scheme applied, the achieved sum rate can get arbitrarily far from the decoupled system sum capacity.

3.5.3 Potential for cooperation over side-channels

In order to remedy this inherent limitation in heterogeneous networks, we propose the use of device-to-device side-channels for user cooperation to resolve the full-duplex interference. In particular, we consider a system architecture where each uplink user is capable of allocating some $\lambda \in [0,1]$ fraction of its power to an orthogonal channel that is used for cooperation with the downlink users. The side-channels are modeled by

$$\widetilde{y}_i = g\widetilde{x}_j + \widetilde{z}_i$$

with the power constraint $\mathbb{E}\left|\widetilde{X}_j\right|^2 \leq \lambda\bar{P}$, for each uplink user j and downlink user i such that $\kappa(i) = \kappa(j)$, with $\widetilde{z}_i \sim \mathcal{CN}(0.1)$. Hence, the side-channels can be considered as orthogonal broadcast channels for each uplink user (we assume each broadcast channel operates over a different band, hence they do not interfere).

It is easy to see that cooperation over such orthogonal side-channels can help mitigate the device-to-device interference. Some schemes have been proposed in [BS13] regarding how to use such side-channels. Here, we focus on the following very simple scheme as an example to demonstrate that side-channels can indeed be effective in mitigating full-duplex interference in clustered networks.

Each uplink user j replicates its symbol over the main channel on the side-channel, with equal power allocation, *i.e.*, $\widetilde{x}_j = \bar{x}_j$, and $\lambda = \frac{1}{2}$. Each downlink user i subtracts the output received over the side-channel $\widetilde{y}_i$ from its output in the main channel y_i to obtain

$$y_i - \widetilde{y}_i = h_i x + z_i - \widetilde{z}_i$$

Note that as a result, the effective channels of each uplink and downlink gets isolated, but the signal-to-noise ratio gets halved for both uplink and downlink due to power allocation and noise superposition, respectively. Therefore, this scheme can achieve

$$\bar{\mathbf{R}}_n + \mathbf{R}_n < M \log\left(1 + h^2 \frac{P}{2M}\right) + M \log\left(1 + h^2 \frac{\bar{P}}{2}\right)$$

which is easily seen to be within $2M$ bits of the isolated system capacity with the side-channels (since the side-channel cannot increase capacity in the isolated case [BS13]), independent of SNR.

CHAPTER 4

Opportunistic D2D Cooperation in Cellular Networks

4.1 Introduction

One of the biggest challenges in wireless networks is to provide uniform connectivity experience throughout the service area. The problem is especially difficult at the cell-edge, where users with unfavorable channel conditions need to receive reliable and high-rate communications. One of the ambitious visions of 5G network design is to achieve 10x reduction in data rate variability in the cell [OHT16] (over existing 4G single-user MIMO OFDM architecture with proportional fair scheduling), without sacrificing the overall sum throughput in the system. In this chapter, we propose and study a solution that, realistic simulations indicate, can give up to approximately 6x improvement in data rate for cell-edge (bottom fifth-percentile) users while still improving the overall throughput under various system constraints.

The proposed solution is centered around opportunistically using the unlicensed band through device-to-device (D2D) cooperation to improve the performance of the licensed multiple-antenna downlink transmission. This solution can be enabled without the presence of any WiFi hotspots, or other data off-loading mechanisms. The main idea is an architecture where a multiple-antenna downlink channel is enhanced through out-of-band D2D relaying to provide multiple versions of the downlink channel outputs, forming virtual MIMO links, which is then opportunistically harnessed through scheduling algorithms designed for this architecture. Note that due to mobility of users, and the fact that the unlicensed band used for cooperation is undedicated, the opportunities for cooperation arise intermittently and unreliably, requiring the opportunistic use of the cooperative resources.

The architecture is predicated on two complementary developments. The first is that infrastructure is becoming more powerful, with the use of a growing number of multiple antennas through massive MIMO for 5G. The other development is on the user equipment (UE) side, with mobile devices becoming more powerful, both in terms of spectrum access and computational power. Most of the mobile devices currently in widespread use can access multiple bands over the ISM spectrum, including the 2.4GHz and 5GHz bands. Furthermore, dense clusters of users constitute a challenging scenario for increasing capacity through massive MIMO, which is precisely the scenario where D2D cooperation is the most useful, since the D2D links are much stronger.

The main technical question involving the architecture is that of how and when to enable the D2D links in a network with many users to boost the cell-edge gains. Our analysis, which uses the network utility maximization framework, leads to an optimal resource allocation algorithm for scheduling these links in a centralized manner, while accounting for system constraints such as limited network state knowlede at the base station; uncoordinated interference over the unlicensed band; fairness in throughput and fairness in the amount of relaying performed by users. Extensive simulations based on 3GPP channel models demonstrate that the proposed architecture combined with our resource allocation algorithm can yield up to approximately 6x throughput gain for the bottom fifth-percentile of users in the network and up to approximately 4x gain for median users over the state-of-the-art single-user MIMO (SU-MIMO) currently implemented in LTE systems, without degrading the throughput of the high-end users.

Since the architecture relies on opportunistically using the unlicensed ISM bands, an important question is how the D2D transmissions would affect other wireless technologies using the unlicensed bands, such as WiFi. As a co-existence mechanism, one can consider strategies similar to LTE-U [ZWC15]: a user can search for an available (unused) channel within the unlicensed band to use for D2D cooperation. If none exists, the user can either declare itself unavailable for D2D cooperation, or transmit only for a short duty cycle. We study the effect of a simplified co-existence mechanism that does the former through simulations, and find that the throughput loss in WiFi users is small compared with the gains in the cell-edge users, since the fraction of time D2D transmission is required from a given user is small.

The chapter is organized as follows. In Section 4.2, we review the literature and delineate our main contributions. In Section 4.3, we present our model and the proposed architecture. In Section 4.4, we present the physical-layer cooperation scheme, prove its approximate optimality, describe its extension to multiuser MIMO (MU-MIMO), and study the scaling behavior of the minimum effective SNR in the network. In Section 4.5, we formulate the downlink cooperative scheduling problem within the utility optimization framework and present our scheduling algorithm, along with the proposed cooperative utility metric, and in Section 4.6, we present our simulation results. Most of the lengthy proofs can be found in Appendix C.

4.2 Related Work and Contributions

The relevant literature can be broadly classified into three areas: (*i*) cooperative cellular communications; (*ii*) dynamic downlink scheduling; (*iii*) D2D in cellular communications; each of which we will summarize next.

In cooperative cellular communications, the idea is to allow users overhearing transmissions to perform relaying to increase spatial diversity and minimize outage probability. This line of work (for instance, [SEA03, NHH04, LTL06], and the references therein) typically focuses on uplink and in-band cooperation, where users that overhear other users' transmission over the licensed band relay their version to the base station. In contrast, we focus on downlink communication and out-of-band cooperation, where users perform relaying for each other's downlink traffic by *opportunistically* using the unlicensed band. As will be seen, the use of orthogonal bands for cooperation can significantly simplify coding schemes.

There is also a large literature in cellular downlink scheduling. Some of these works focus on scaling behavior of the achievable rate under various scheduling schemes [SH05, YG06], some focus on the low-complexity algorithms [DS05], while some others also account for fairness and various system constraints using the cross-layer optimization approach [LCS01, TG05, LSS06, GNT06, SCN10]. While our work uses the cross-layer optimization paradigm as well, none of the proposed resource allocation algorithms directly applicable to our cooperative scenario, since we consider an architecture where the broadcast nature of the wireless medium is explicitly used at the physical-layer, precluding an abstraction into isolated bit pipes in upper layers, which is a prevalent model in existing works on cross-layer optimization.

Embedding D2D communication in cellular network has also received considerable attention in the past (see [AWM14] for a comprehensive survey). A majority of these works (*e.g.*, [DRW09, LLG12, WTS13]) focus on direct proximal communication between devices, where one device directly transmits a message for another over the licensed band, skipping infrastructure nodes. This type of proximal D2D communication also has been part of the 4G LTE-Advanced standard [LKM15]. The main focus in this line of work is to do resource allocation and interference management across D2D and/or uplink/downlink message flows. In contrast, we focus on D2D *cooperation* to aid downlink communication, which is the use of physical D2D transmissions to assist downlink message flows intended for other devices. This can be considered as a new way the D2D capability can be used in the next-generation 5G networks, in addition to the existing proximal communication in 4G. Considering the fact that the volume of downlink traffic far exceeds the volume of proximal D2D communication traffic, the cooperation architecture has the potential to exploit the D2D capability to a much higher degree. This is also in line with one of the envisioned goals in 5G, which is to enable multihop communication in cellular networks [CZ14].

Conceptually, the most relevant work in the literature to our problem is the one in [AM13], where the authors propose an architecture where users form clusters through the use of unlicensed bands, and all communication with the base station is performed through the cluster head. In another line of work [WR13], the authors suggest using out-of-band D2D for traffic spreading, where a user performs sends request and receives downlink content on behalf of another user, in a base-station transparent manner. In both works, the authors numerically demonstrate various throughput, fairness and energy-efficiency benefits of D2D. In contrast to these works, our physical-layer scheme is not based on routing; it explicitly uses the direct link from the base station to the destination user in addition to the relay links. We also consider a much more general scheduling algorithm based on utility optimization and dynamic user pairing, while accounting for fairness and cooperation cost.

The main technical contributions of this work can be summarized as follows.

- We analyze a physical-layer scheme based on compress-and-forward relaying and MIMO Tx/Rx processing that approximately achieves (within 2 bits/s/Hz) the capacity of two-user downlink channel with D2D cooperation (Section 4.4.1), and describe how the scheme can be extended to MU-MIMO (Section 4.4.2). We characterize the gains in terms of cell-edge SNR-scaling due to D2D cooperation for a specific model of clustered networks (Section 4.4.3).

- We develop a resource allocation policy for selectively enabling such D2D links for cooperation, using the utility maximization framework (Section 4.5.1). Since the existing cross-layer design tools are not directly applicable in our scenario when D2D transmission conflicts are taken into account, we propose a novel scheduling policy for such D2D-enabled networks that takes into account such conflicts (Section 4.5.3). The policy consists of an extension of the single-user scheduling algorithm of [TG05] to the cooperative MU-MIMO scenario with incomplete network state knowledge, and a novel flow control component based on an explicit characterization of an inner bound on the stability region of the system. The proposed algorithm is shown to be optimal with respect to this inner bound on the stability region. We also introduce a novel class of utility functions for cooperative downlink communication, which incorporates the cost of cooperation and leads to desirable fairness properties (Section 4.5.5).

- We present an extensive simulation study using 3GPP specifications to study the performance of the proposed architecture (Section 4.6). The main results include (i) a throughput gain ranging from 4.3x up to 6.3x (depending on system constraints, channel estimation accuracy etc.) for the users in bottom fifth-percentile for MU-MIMO with D2D cooperation versus the state-of-the-art SU-MIMO, without degrading the throughput of the stronger users, (*ii*) a throughput gain ranging from 3.7x up to 4.9x for the bottom fifth-percentile users versus non-cooperative MU-MIMO without degrading throughput of stronger users, (*iii*) a reduction of more than 50% in the relaying load in the network through the use of novel utility functions, while still giving gains close to proportional fair case, (*iv*) a basic study of an architecture wherein D2D cooperation coexists (and interferes) with WiFi in the network via a simple co-existence mechanism where cooperation is disabled within WiFi range, where it is shown that despite the residual interference, the throughput loss in WiFi users is small (10% for median user) compared with the gains in the cell-edge users (130% for fifth-percentile user), since the fraction of time D2D transmission is required from a given user is small (in the simulation 80% of users performed relaying less than 10% of the time).

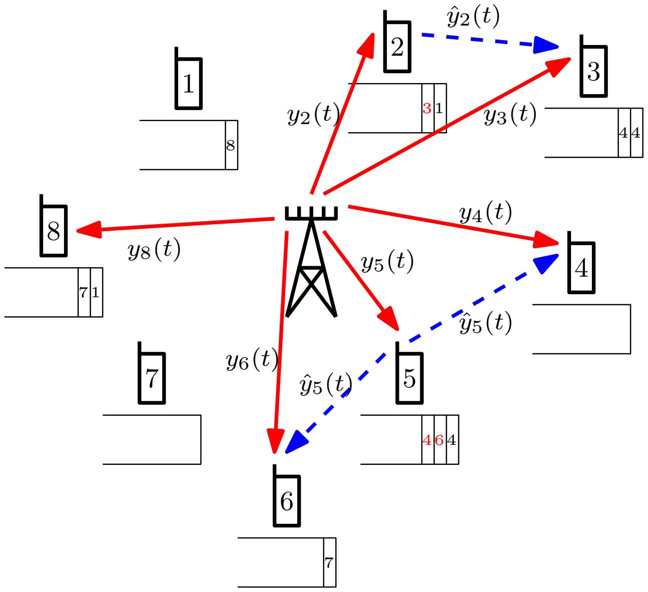
4.3 System Architecture and Model

4.3.1 Overview of the architecture

Consider a single cell in a multi-cell downlink cellular system[1] with a base station equipped with M antennas, and a set $\mathcal{N}$ of single-antenna users, where $|\mathcal{N}| = n$. An example operation is depicted in Figure 4.1. We assume slotted time, with m representing the physical-layer time index. A frame, indexed by t, is defined as T consecutive discrete time slots[2]. We will use the notation $m \sqsubset t$ to mean that the physical-layer slot m lies within the frame t, *i.e.*, $(t-1)T < m \leq tT$.

[1] Since the base stations are uncoordinated, for the purposes of designing a scheduling algorithm, it is sufficient to consider a single cell in isolation. We will consider the multi-cell system in Section 4.6 for evaluation purposes.

[2] We will use square brackets to denote physical-layer time indices, and round brackets for frame indices.



Destination	Relay
4	5
6	5
8	-
3	2

Figure 4.1: An example scheduling decision made by the base station, where the table reflects the selected active set. The red arrows denote the corresponding downlink transmissions, all taking place throughout frame t, and the dashed blue arrows represent scheduled side-channel transmissions, taking place at a later time, determined by multiple-access protocol $\mathcal{I}$. Once the active set is selected, the required side-channel transmissions are queued at the users (the transmissions scheduled in frame t are highlighted in red). In this example, user 5 is selected to share a function of its channel output to relay for users 4 and 6; user 2 is selected to relay for user 3; and user 8 is scheduled without any relays.

In the proposed architecture, the base station selects an active set $\mathcal{A}(t) \subseteq \mathcal{N}^2$ for each frame t, which consists of pairs (i, j) of users, where the first index i refers to the destination node scheduled for data, and the second index j refers to user assigned as a relay for user i. We define (i, i) to represent the case where user i is scheduled with no relay assigned. Note that a user can be designated as a relay for a stream and a destination for another stream simultaneously, as exemplified in Figure 4.1. It is also possible within this framework to assign multiple relays to the same destination by having $(i, j), (i, k) \in \mathcal{A}(t)$. We define $A_{ij}(t) = 1$ if $(i, j) \in \mathcal{A}(t)$, and $A_{ij}(t) = 0$ otherwise.

Once the selection $\mathcal{A}(t)$ is made, the base station transmits a sequence of vectors $\mathbf{x}[m] \in \mathbb{C}^M$, $m = 1, \ldots, T$, over M antennas and T time slots of the frame t. The channel output $y_i[m]$ at user i is given by

$$y_i[m] = \mathbf{h}_i^*(t)\mathbf{x}[m] + w[m], \tag{4.1}$$

for $m \sqsubset t$, where $\mathbf{h}_i(t) \in \mathbb{C}^M$ is the time-variant complex channel vector of user j at frame t (note that we are assuming that channel stays constant within a frame, but can arbitrarily vary over time slots), $\mathbf{x}[m]$ is the input vector to the channel at time m, and $w[m] \sim \mathcal{CN}(0, 1)$ is the circularly symmetric complex white Gaussian noise process. We assume an average power constraint $\frac{1}{T}\sum_{m=1}^{T} \mathbf{tr}\left(\mathbf{x}[m]\mathbf{x}^*[m]\right) \leq 1$, and define $H(t) := \left\{\mathbf{h}_i(t)\right\}_i$.

If user j is assigned as a relay for user i at frame t, a transmission from user j to i is queued at user j, to be transmitted at a later frame $\tau > t$. At frame τ, user j transmits the sequence $x_j[m] \in \mathbb{C}$, $m \sqsubset \tau$, which is a deterministic function of the receptions corresponding to earlier frame t, *i.e.*, $y_j[\widetilde{m}]$ for $\widetilde{m} \sqsubset t$. User i performs decoding by combining its own channel outputs $y_i[m]$, $m \sqsubset t$, with the receptions from j, $\bar{y}_j[\widetilde{m}]$, $\widetilde{m} \sqsubset \tau$, which is a function of $y_j[m]$, $m \sqsubset t$ (the specific D2D link model generating $\bar{y}_j[\widetilde{m}]$ will be discussed later). Note that user i can combine receptions corresponding to multiple frames to decode.

We will specify the details of the model and formulate the specific mathematical problem.

Table 4.1: Notation for variables corresponding to the D2D link (i, j)

Notation	**Explanation**
g_{ij}	D2D channel gain
Q_{ij}	State of the queue at relay j for destination i
ϕ_{ij}, Φ	Path-loss factor(s)
μ_{ij}	Binary service process (transmission permission indicator) for the queue Q_{ij}
ζ_{ij}, Z	Fading parameter(s)
A_{ij}	Binary arrival process (D2D link scheduling indicator) for the queue Q_{ij}
B_{ij}	D2D link availability indicator
J_{ij}	D2D interference indicator
C_{ij}	The capacity of the D2D link
β_{ij}	Arrival rate to the queue Q_{ij}

4.3.2 D2D link model and conflict graph

For any pair $(i, j) \in \mathcal{N}^2, i \neq j$, the time-variant channel gain is given by $g_{ij}(t) = \sqrt{\phi_{ij}}\zeta_{ij}(t)$, where $\phi_{ij} \in \mathbb{R}$ is the path loss component, and $\zeta_{ij}(t) \sim \mathcal{CN}(0, 1)$ is the fading component for the pair (i, j), i.i.d. across MAC layer slots. We assume reciprocal side-channels, *i.e.*, $g_{ij}(t) = g_{ji}(t)$, and define $Z(t) := \{\zeta_{ij}(t)\}_{i,j}$ and $\Phi := \{\phi_{ij}\}_{i,j}$.

We define $B_{ij}(t)$ as an i.i.d. *Bernoulli*(p_{ij}) process for each $(i, j) \in \mathcal{N}^2, i \neq j$, representing whether or not the link (i, j) is available at frame t. This models unavailability due to external transmissions (*e.g.*, WiFi access points, or another application on the same device attempting to use WiFi etc.) in the same unlicensed band. The realization of $B_{ij}(t)$ is known at the users strictly causally (at frame $t + 1$), and unknown at the base station. We define $B(t) := \{B_{ij}(t)\}_{i\neq j}$.

Define the connectivity graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ such that $\mathcal{V} = \mathcal{N}$, and $\mathcal{E}$ is such that $(i, j) \in \mathcal{E}$ if $i = j$ or $\phi_{ij} > \theta$ for some threshold $\theta > 0$ (*e.g.*, noise level). We further define the conflict graph $\mathcal{G}_c = (\mathcal{V}_c, \mathcal{E}_c)$ such that

$$\mathcal{V}_c := \left\{(i, j) \in [n]^2 : i \neq j\right\} \tag{4.2}$$

$$\mathcal{E}_c := \{((i, j), (k, \ell)) : (i, j) \neq (k, \ell) \text{ and } ((i, \ell) \in \mathcal{E} \text{ or } (j, k) \in \mathcal{E})\}.$$

The conflict graph represents the pairs of D2D transmissions $(i, j), (k, \ell)$ that are not allowed to simultaneously occur due to interference[3]. Given these definitions, the channel from user $j \in \mathcal{N}$ to user $i \in \mathcal{N} - \{j\}$ is modeled by

$$\bar{y}_i[m] = B_{ij}(t) J_{ij}(t) \left(g_{ij}(t) x_j[m] + \bar{w}_i[m] \right)$$

for $m \sqsubset t$, where

$$J_{ij}(t) = \begin{cases} 0, & \exists (k, \ell)\text{: } ((i,j),(k,\ell)) \in \mathcal{E}_c \text{ and } \|x_\ell[\tilde{m}]\|^2 > 0 \text{ for some } \tilde{m} \sqsubset t \\ 1, & \text{otherwise} \end{cases},$$

which captures interference between conflicting D2D transmissions, and $\bar{w}_i[m] \sim \mathcal{CN}(0,1)$ is the complex white Gaussian noise process. We assume an average power constraint $\frac{1}{T}\sum_{m=1}^{T} \|x_j[m]\|^2 \leq 1$, absorbing the input power into the channel gain. The capacity of the D2D link (i,j) at time t (assuming it is available) is given by $C_{ij}(t) := \log\left(1 + \|g_{ij}(t)\|^2\right)$. We assume the base station has knowledge of the average SNR, *i.e.*, the path-loss component ϕ_{ij} for each (i,j) pair, but has no knowledge of the fading realization $\zeta_{ij}(t)$.

4.3.3 D2D transmission queues

We assume that each user $j \in \mathcal{N}$ maintains $(n-1)$ queues, whose states are given by $Q_{ij}(t)$, $i \in \mathcal{N} - \{j\}$, each representing the number of slots of transmission[4] to be delivered to node i. We assume the queue states evolve according to

$$Q_{ij}(t+1) = \left(Q_{ij}(t) - B_{ij}(t) J_{ij}(t) \mu_{ij}(t) \right)^+ + A_{ij}(t), \tag{4.3}$$

where $\mu_{ij}(t)$ is a binary process that is induced by the multiple-access protocol $\mathcal{I}$ used by the nodes, indicating whether or not the flow (i,j) is granted permission for transmission at frame t. The protocol $\mathcal{I}$ is a mapping from the current queue states $\{Q_{ij}(t)\}_{i \neq j}$ and the D2D interference structure $\{J_{ij}(t)\}_{i \neq j}$ to the binary service processes $\{\mu_{ij}(t)\}_{i \neq j}$.

[3] The interference model that induces the conflict graph $\mathcal{G}_c$ as defined in (4.2) is similar to the two-hop interference model of [Ari84], but also takes into account the directionality of the transmission

[4] Note that $Q_{ij}(t)$ does not represent the number of *bits* to be transmitted, but the number of *slots of transmission.* This is because the reception of relay does not directly translate into information bits, but is rather a *refinement* of the reception of the destination node.

We define the average arrival rates as $\beta_{ij}(t) := \frac{1}{t}\sum_{\tau=1}^{t} A_{ij}(t)$, and $\beta_{ij} := \limsup_{t\to\infty} \beta_{ij}(t)$. For a given vector of arrival rates $\beta := \{\beta_{i,j}\}_{i\neq j}$, the system is said to be *stable* if the average queue sizes are bounded, *i.e.*, for all (i,j), $\limsup_{t\to\infty} \mathbb{E}\left[Q_{ij}(t)\right] < \infty$. The set of arrival-rate vectors β for which there exists service processes $\{\mu_{ij}(t)\}_{i\neq j}$ such that the system is stable is called the *stability region* of the queueing system, and will be denoted by Λ. Note that the arrival rates need to remain in the stability region in order to ensure that the D2D transmissions eventually occur with a finite delay. Within the scope of this work, we do not focus on the details of $\mathcal{I}$, and simply assume that the nodes implement a protocol $\mathcal{I}$ that achieves the stability region Λ, *i.e.*, if the arrival rates $\beta \in \Lambda$, protocol $\mathcal{I}$ can find a schedule for D2D transmissions such that each transmission is successfully delivered with finite delay[5].

4.3.4 Problem formulation

If the vector of arrival rates $\beta \in \Lambda$, we can assume that a noiseless logical link with capacity $\bar{R}_{ij}(t)$ is available at time t, where $\bar{R}_{ij}(t) = C_{ij}(\tau)$ for some finite $\tau \geq t$, where τ is the frame where the actual physical D2D transmission takes place, carrying traffic scheduled at frame t. Note that at frame t, the base station has no knowledge of $C_{ij}(\tau)$, but can still compute the average capacity $\mathbb{E}_{Z(\tau)}\left[C_{ij}(\tau)\,|\phi_{ij}\right]$ for a given link (i,j), for a transmission decision. We define $\bar{Z}(t) = Z(\tau)$. Let $\mathcal{C}(t)$ denote the instantaneous information-theoretic capacity region of the system consisting of the channels (4.1) and the set of logical links (i,j) with capacities $\bar{R}_{ij}(t)A_{ij}(t)$, with no knowledge of $\bar{Z}(t)$ at the base station[6]. A physical-layer strategy γ is a map $\left(H(t), \Phi, \bar{Z}(t)\right) \mapsto \{R_i(t)\}_i$ whose output vector (interpreted as the vector of information rates delivered to users, in bits/s/Hz) satisfies $\{R_i(t)\}_i \in \mathcal{C}(t)$ for all i and t.

[5] One can design such a protocol by having the nodes coordinate with the base station to circumvent the hidden terminal problem, and then use any of the existing stability-region-achieving distributed scheduling algorithms, *e.g.*, [TE92, JW10, MSZ06]

[6] Note that the D2D link is assumed to have zero capacity if $A_{ij}(t) = 0$, *i.e.*, if the base station did not schedule the link (i,j) at time t.

Note that even though the transmission decisions of the base station does not depend on the unknown components of the network state $Z(t)$, by allowing the rate vector $\{R_i(t)\}_i$ to be anywhere inside the instantaneous capacity region, we implicitly assume an idealized rate adaptation scenario, where once the transmission occurs, the capacity corresponding to the realization of $\bar{Z}(t)$ is achievable. In practice this can be implemented through incremental redundancy schemes such as hybrid ARQ.

Assume an infinite backlog of data to be transmitted to each user $i \in \mathcal{N}$. The long-term average rate of user i up to time t is defined as $r_i(t) = \frac{1}{t}\sum_{\tau=1}^{t} R_i(\tau)$, where $R_i(\tau)$ is the rate delivered to user i by the physical layer scheme $\gamma(t)$ chosen at time t. The long-term throughput of user i is $r_i = \liminf_{t\to\infty} r_i(t)$. Define $\mathbf{r}(t) = \{r_i(t)\}_i$, and $\mathbf{r} = \{r_i\}_i$.

Given the stability-region-achieving D2D MAC protocol $\mathcal{I}$, and a set of physical-layer strategies Γ, at every frame t, the base station chooses an active set $\mathcal{A}(t)$, and a strategy $\gamma(t) \in \Gamma$ consistent with $\mathcal{A}(t)$. A scheduling policy π is a collection of mappings

$$(\mathbf{r}(t-1), \beta(t-1), H(t), \Phi) \mapsto (\mathcal{A}(t), \gamma(t)),$$

indexed by t. If β^{π} represents the vector of arrival rates to the queues under policy π, and $\mathbf{r}^{\pi}$ the throughputs under policy π, then the policy π is called *stable* if $\beta^{\pi} \in \Lambda$. Our goal is to design a stable policy π that maximizes any given concave, twice-differentiable *network utility function* $U(\mathbf{r}^{\pi}, \beta^{\pi})$ of the throughputs and the fraction of time nodes spend relaying for others[7].

4.4 Physical-layer Cooperation

In this section, we describe a class of physical layer cooperation strategies that will be used as a building block for our proposed architecture, and derive its achievable rates. We will first focus on the two-user case, where we show the approximate information-theoretic optimality of the scheme. We consider the extension to MU-MIMO in Section 4.4.2.

[7] Note that since π is stable, the relaying fraction is the same quantity as the arrival rate β.

The main idea behind the cooperation strategy is that the D2D side-channel can be used by the destination node to access a quantized version of the channel output of the relay node, which combined with its own channel output, effectively forms a MIMO system. The base station can perform signaling based on singular value decomposition over this effective MIMO channel, to form two parallel AWGN channels accessible by the destination node. Next, we describe the strategy in detail, and derive the rate it achieves.

4.4.1 Cooperation strategy

We isolate a particular user pair (i, j), and without loss of generality assume $(i, j) = (1, 2)$. The effective network model is given by[8]

$$y_i = \mathbf{h}_i^* \mathbf{x} + z_i,\ i = 1, 2, \quad \bar{y}_1 = g_{12} x_2 + \bar{z}_1, \tag{4.4}$$

where $x_2[m]$ is a function of y_2^{m-1}, the past receptions of user 2, and user 1 has access to y_1 and $\bar{y}_1$.

By Wyner-Ziv Theorem [WZ76], if

$$\bar{R}_{12} \geq \min_{\substack{p(w|y_2) \\ \hat{y}_2(w,y_1)}:\mathbb{E}[\|\hat{y}_2 - y_2\|^2] \leq D} I(y_2; w|y_1)$$

for a given joint distribution of channel outputs $p(y_1, y_2)$, then given a block of outputs y_2^N, user 1 can recover a quantized version $\hat{y}_2^N$ of outputs such that[9] $\mathbb{E}\left[\|\hat{y}_2 - y_2\|^2\right] \leq D$.

Choosing $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}, \mathbf{Q})$, i.i.d. over time, we get $(y_1, y_2) \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma})$ i.i.d. over time, for some covariance matrix $\Sigma = \mathbf{HQH}^*$ induced by the channel, with $\mathbf{H} = [\mathbf{h}_1\ \mathbf{h}_2]^*$. We further choose $w = y_2 + q_2$, where $q_2 \sim \mathcal{CN}(0, D)$ is independent of all other variables, and we set the mapping $\hat{y}_2(w, y_1) = w$. We also choose $D = \frac{\sigma_{2|1}^2}{|g_{12}|^2}$, where $\sigma_{2|1}^2 = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$ is the

[8] We focus on a particular frame t to characterize the instantaneous capacity, *i.e.*, the achievable rate for a given set of network parameters.

[9] This is achieved by performing appropriate quantization and binning of the channel outputs at user 2 (see [WZ76] for details).

conditional variance of y_2 given y_1. With this set of choices, it can be shown that user 1 can access $\hat{y}_2 = y_2 + q_2$, where $q_2 \sim \mathcal{CN}(0, D)$.

Once user 1 recovers $\hat{y}_2$, it can construct the effective MIMO channel

$$\mathbf{y} = \begin{bmatrix} y_1 \\ \hat{y}_2 \end{bmatrix} = \mathbf{Hx} + \begin{bmatrix} z_1 \\ z_2 + q_2 \end{bmatrix}. \tag{4.5}$$

It follows that all rates $R < R_{\text{MIMO}}$ are achievable over the effective MIMO channel (4.5), where

$$R_{\text{MIMO}} = \max_{\text{tr}(\mathbf{Q}) \leq 1} \log \left| \mathbf{I}_2 + \mathbf{K}^{-1} \mathbf{HQH}^* \right|,$$

with $\mathbf{K} = \text{diag}\left(1, \;\; 1 + \frac{\sigma_{2|1}^2}{|g_{12}|^2}\right)$. Note that due to orthogonality of the links incoming to the destination, the encoding and decoding is significantly simplified compared to traditional Gaussian relay channel with superposition, since there is no need for complex schemes such as block Markov encoding and joint decoding, and point-to-point MIMO codes are sufficient from the point of view of the source.

Note that the MIMO channel (4.5) can be equivalently viewed as two parallel AWGN channels, using the singular value decomposition (SVD). It will also be useful to lower bound the rates individually achievable over these two parallel streams. Assuming $\mathbf{H} = \mathbf{USV}^*$ is an SVD, it can be shown that the rates

$$R_{\text{MIMO},d} = \log \left(1 + \frac{s_d^2 P_d}{1 + |u_{2d}|^2 \frac{\sigma_{2|1}^2}{|g_{12}|^2}} \right), \; d = 1, 2 \tag{4.6}$$

are achievable respectively[10], over the two streams, by transmit beamforming using the matrix $\mathbf{V}$ and receive beamforming using $\mathbf{U}^*$, where s_d is the dth singular value, $u_{2,d}$ is the $(2, d)$th element of $\mathbf{U}$, and the power allocation parameters satisfy $P_1 + P_2 \leq 1$.

The next theorem shows that the gap between the rate achievable with the cooperation scheme described in the previous subsection is universally within 2 bits/s/Hz of the capacity of the network.

[10] We perform the SVD on $\mathbf{H}$ directly, instead of performing on $\mathbf{K}^{-1/2}\mathbf{H}$, in order to obtain closed-form expressions for the subsequent analysis.

Theorem 4.1. *For any set of parameters* $\left(\mathbf{H}, \bar{R}_{12}, M\right)$*, the capacity* $\bar{C}$ *of the MIMO single relay channel with orthogonal links from relay to destination and from source to destination satisfies* $\bar{C} \geq R_{MIMO} \geq \bar{C} - 2$.

The proof is provided in the Appendix C.3.

Remark 4.1. *The relay channel with orthogonal links from relay to destination and from source to destination was studied by [LV05] and [ZME04]. In the former, the authors consider a relaying strategy based on decode-and-forward relaying, and focus on performance optimization problems such as optimal bandwidth allocation. The latter work focuses on linear relaying functions for such channels, and characterizes the achievable rates for scalar AWGN case. Here, we propose a relaying scheme based on compress-and-forward [Ct79] that achieves a rate that is within 2 bits/s/Hz of the information-theoretic capacity for the MIMO case.*

Remark 4.2. *Note that this strategy can also be implemented through quantize-map-forward relaying. Although the proposed architecture supports other relaying strategies (* e.g., *amplify-forward, decode-forward etc.), we stick with compress-forward (or quantize-map-forward implementation) due to the theoretical approximate optimality [ADT11] as well as practical feasibility, which was shown in [DSB13] through real testbed implementation.*

4.4.2 Cooperation with MU-MIMO

In this subsection, we demonstrate how the scheme described for two users in the previous subsection can be extended to MU-MIMO with pairs of cooperative users.

Given the set $\mathcal{N}$ of users, let us index all possible downlink streams that can be generated by the scheme by $(i, j, d) \in \mathcal{N}^2 \times \{1, 2\}$, where (i, j) is represents the cooperative pair, and d represents the stream index corresponding to this pair. We assume $d \neq 2$ if $i = j$, representing the case where user i is scheduled without a relay.

By a slight abuse of notation, we assume that a schedule set $\mathcal{S} \subseteq \mathcal{N}^2 \times \{1, 2\}$ is scheduled, consisting of such triples (i, j, d), where $(i, j, d) \in \mathcal{S}$ for some k if and only if $A_{ij} = 1$ (note that schedule set also contains the stream index unlike active set $\mathcal{A}$). Next, consider the "virtual users" $(i, j, d) \in \mathcal{S}$ with the channels

$$\tilde{y}_{ijd} := \mathbf{u}_{ijd}^* \begin{bmatrix} y_i \\ \hat{y}_j \end{bmatrix} = \mathbf{u}_{ijd}^* \mathbf{H}_{ij} \mathbf{x} + \mathbf{u}_{ijd}^* \begin{bmatrix} z_i \\ z_j + q_{ij} \end{bmatrix} := \tilde{h}_{ijd}^* \mathbf{x} + \tilde{z}_{ijd},$$

where $\mathbf{H}_{ij} = [\mathbf{h}_i \; \mathbf{h}_j]^*$, and assuming $\mathbf{H}_{ij} = \mathbf{U}_{ij}\mathbf{S}_{ij}\mathbf{V}_{ij}^*$ is an SVD of $\mathbf{H}_{ij}$, $\mathbf{u}_{ijd}$ is the kth column of $\mathbf{U}_{ij}$. By convention, we assume that $\mathbf{U}_{ii} = \begin{bmatrix} 1 & 0 \end{bmatrix}^*$. The variance of $\tilde{z}_{ijd}$ is given by $1 + |u_{ijd}(2)|^2 D_{ij}$, where $u_{ijd}(2)$ is the second element of $\mathbf{u}_{ijd}$, and $D_{ij} = \frac{\sigma_{j|i}^2}{|g_{ij}|^2}$ is the distortion introduced by quantization at node j. Note that, when $i = j$, we have $\tilde{\mathbf{h}}_{ijd} = \mathbf{h}_i$, and we set $\bar{R}_{ii} = \infty$ so that $D_{ii} = 0$.

Note that through the use of SVD over the virtual MIMO channel (4.5), we have reduced the system into a set of $|\mathcal{S}|$ single-antenna virtual users with channel vectors $\frac{1}{1+\left|u_{ijd}(2)\right|^2 D_{ij}} \tilde{\mathbf{h}}_{ijd}$. Given such a set of channel vectors, one can implement any MU-MIMO beamforming strategy (*e.g.*, zero-forcing, conjugate beamforming, SLR maximization etc.), by precoding the transmission with the corresponding beamforming matrix.

4.4.3 Scaling of SNR gain in clustered networks

In this subsection, we consider a specific clustered network model as an example, and characterize the achievable demodulation SNR gain due to D2D cooperation for the weakest user in the network, under this model. In this analysis, we use several simplifying assumptions on the channel and network model for analytical tractability, in order to get a feel for the scale of the possible gains that can be attained through cooperation. This simplification is limited to the scope of this particular subsection, and the results in the rest of the chapter do not depend on these assumptions.

Consider a network where users are clustered in a circular area of radius r, whose center is a distance d away from the base station, where $r \ll d$. The users are assumed to be uniformly distributed within the circular area. In general, a network might consist of several such clusters, but here we focus on one, assuming other clusters are geographically far relative to r.

We assume that the downlink channel vector of user i at time t is modeled by[11]

$$\mathbf{h}_i(t) = \sqrt{\rho} \sum_{k=1}^{P} \xi_{i,k}(t) \mathbf{e}\left(\theta_{i,k}(t)\right),$$

where ρ is the path loss factor (assumed constant across users in the same cluster since $r \ll d$), P is the number of signal paths, $\xi_{i,k}(t) \sim \mathcal{CN}(0,1)$ is the complex path gain for the kth path of user i at time t, $\theta_{i,k}$ is the angle of departure of the kth path of the ith user at time t, and $\mathbf{e}(\theta)$ is given by

$$\mathbf{e}(\theta) := \left[\begin{array}{ccccc} 1 & e^{j2\pi\Delta\cos(\theta)} & e^{j2\pi 2\Delta\cos(\theta)} & \dots & e^{j2\pi(M-1)\Delta\cos(\theta)} \end{array}\right]^*,$$

for an antenna separation Δ. The path gains $\xi_{i,k}(t)$ are i.i.d. across different i, k, and t.

Path loss between users is modeled by $\phi_{ij} = \phi_0 d_{ij}^c$ for some constant ϕ_0, where d_{ij} is the distance between i and j, and $c > 2$ is the path loss exponent.

For simplicity of analysis, in this example network we will assume that only one cooperative pair per time slot is scheduled. Our goal is to characterize the cooperation gains in SNR when one is allowed to choose the most suitable relay j for a given destination i.

Invoking (4.6), we define the cooperative SNR for the pair (i,j), SNR_{ij}^{coop} to be

$$\mathsf{SNR}_{ij}^{coop} := \frac{s_{ij1}^2}{1 + |u_{ij1}(2)|^2 \frac{\sigma_{j|i}^2}{|g_{ij}|^2}},$$

where s_{ij1} is the first singular value corresponding to the pair (i,j). Since we are interested in the achievable SNR gain, in defining this quantity, we have allocated all power to only one of the available streams, ignoring the multiplexing gain that could be achieved by scheduling two parallel streams to user i. The maximal non-cooperative SNR for user i is given by $\mathsf{SNR}_i^{non-coop} := \|\mathbf{h}_i\|^2$, achieved by beamforming along the direction of h_i. Minimum cooperative and non-cooperative SNRs in the network are respectively defined as

$$\mathsf{SNR}_{\min}^{coop} := \min_{i \in \mathcal{N}} \mathsf{SNR}_{ij^*(i)}^{coop}, \quad \mathsf{SNR}_{\min}^{non-coop} := \min_{i \in \mathcal{N}} \mathsf{SNR}_i^{non-coop},$$

[11] This is written for a uniform linear transmit array for simplicity, but our analysis using this model can be generalized for any array configuration.

where $j^*(i) = \arg\max_{j \in \mathcal{N}} \mathbb{E}\left[\mathsf{SNR}_{ij}^{coop} \middle| \phi_{ij}, \mathbf{h}_j\right]$, which arises due to relay selection, and the expectation is taken over the D2D side-channel fading $\zeta_{ij}(t)$.

The next theorem, whose proof is in Appendix C.1, summarizes our results on how the SNR of the weakest user in either case scales with the number of users n in the cluster.

Theorem 4.2.

$$\lim_{n\to\infty} \mathbb{P}\left(\mathsf{SNR}_{\min}^{coop} < \frac{1}{2}M\rho\left(\frac{1}{2}\log n - 2\log\log n\right) - 1\right) = O\left(e^{-\log^2 n + 2\log n}\right),$$

and

$$\lim_{n\to\infty} \mathbb{P}\left(\mathsf{SNR}_{\min}^{non-coop} > M\rho n^{-\frac{\gamma}{2P}}\psi(2P)\right) = O\left(e^{-n^{1-\gamma}}\right),$$

for any $0 < \gamma < 1$*, where* $\psi(\ell) = (\ell!)^{\frac{1}{\ell}}$*, and* P *is the number of signal paths.*

Theorem 4.2 highlights the importance of having multiple options in relay selection. In the non-cooperative case, the factor $n^{-\frac{\gamma}{2P}}$ appears due to the fact that as the number of users in the cluster grows, the minimum is taken over a larger set of users, and hence it is expected for the SNR of the weakest user to decay, in the absence of cooperation. On the other hand, in the presence of cooperation, the SNR of the weakest user actually *grows.* This is due to the *multiuser diversity gain*, which is present due to our ability to schedule the user with the most favorable channel conditions as a relay. In other words, as the number of users grows, so does the number of possible paths from the base station to each user, and thus the maximal SNR, even when the weakest user is considered.

4.5 Scheduling for Dynamic Networks Under D2D Cooperation

Although our analysis of the SNR gain with relay selection in the previous section is informative of the potential gains of cooperation, one should note that its scope is limited. For a more thorough understanding of how to perform relay selection, we formulate the problem within the network utility maximization framework, which has been extensively studied in the context of resource allocation and scheduling problems for wireless/wired networks [GNT06, LSS06].

Note that due to interference from other D2D links as well as from external sources, not all D2D users can transmit at a given time, which implicitly imposes a constraint on relay selection. In particular, one needs to ensure that the relays can find a slot for transmission to the destination user after a finite delay, *i.e.*, the relay queues remain stable. The existing cross-layer optimization algorithms, *e.g.*, [GNT06, LSS06] (*e.g.*, virtual queues, dynamic backpressure routing etc.) are not immediately applicable to this scenario. This is firstly because our physical-layer signaling is not based on routing, and makes explicit use of the broadcast nature of the wireless medium, by using both the direct link to the destination node, and the alternate link formed by relay. Consequently, the full network cannot be abstracted into a graph with isolated links, which is widely assumed in the literature. Second, since our utility metric is a function of the average amount of relaying done by users, different choices of relay for the same user results in different rewards, even when the rates offered in these choices are equal. Existing formulations do not capture this generalization, which necessitates a special treatment of the downlink resource allocation problem with D2D cooperation.

To achieve this, we take an approach consisting of

1. A generalization of the single-user scheduling algorithm of [TG05] based on the maximization of the derivative of the utility function to the cooperative scenario with relay selection, MU-MIMO, and incomplete network state knowledge,

2. A relay flow control scheme integrated into scheduling, which involves explicitly imposing a set of hard linear constraints on the relaying frequency of users,

3. A novel utility metric that is specific to the cooperative architecture, exhibiting desirable fairness properties.

In particular, the second point requires the use of a novel technique using exponential barrier functions to handle the stability constraint, and the generalizations of the first point requires several modifications to the proof of [TG05].

4.5.1 Utility maximization formulation

As discussed in Section 4.3, our goal is to design a stable policy π that maximizes a network utility function $U(\mathbf{r}, \beta) = \sum_{i=1}^{n} U_i(r_i, \beta_i)$, where $U_i : [0, \infty) \times [0, 1] \to \mathbb{R}$, for $i = 1, \ldots, n$, are twice continuously differentiable concave functions that are non-decreasing in the first argument, and non-increasing in the second argument. Note that unlike the existing works, the utility function is not only a function of the throughput (first argument), but also a function of the amount of relaying performed for others by the user (second argument). This definition naturally introduces a penalty each time a D2D link is scheduled, and thus the out-of-band resources are not "free". The utility function $U_i\left(r_i, \beta_i\right)$ then jointly captures the reward of having received an average throughput of r_i, and the cost of having relayed β_i fraction of time, for user i. We will consider a specific form of utility function in Section 4.5.5, and demonstrate its properties in terms of fairness and relaying cost.

Fixing the transmission strategy as the one described in Section 4.4.2, the problem of selecting the pair $(\mathcal{A}(t), \gamma(t))$ reduces to the selection of a schedule set $\mathcal{S}(t) \subseteq \mathcal{N}^2 \times \{1, 2\}$ for every frame t, which specifies the active set $\mathcal{A}(t)$ as well as the stream index corresponding to each pair $(i, j) \in \mathcal{A}(t)$. The schedule set chosen by policy π at frame t will be denoted by $\mathcal{S}_\pi(t)$.

Let the network state be represented by the pair $(K(t), Z(t))$, where $K(t) = (H(t), \Phi)$ represents the network parameters causally known at the base station, and $Z(t)$ is the fading parameter, which is unknown (all variables are as defined in Section 4.3, Table 4.1). We assume that $K(t)$ and $Z(t)$ take values over the arbitrarily large but finite sets $\mathcal{K}$ and $\mathcal{Z}$, respectively[12]. Define

$$\alpha^{\pi}_{skz}(t) = \frac{1}{t}\sum_{\tau=1}^{t} \mathbb{I}_{\mathcal{S}_\pi(\tau)=s}\mathbb{I}_{K(\tau)=k}\mathbb{I}_{Z(\tau)=z},$$

for $s \subseteq \mathcal{N}^2 \times \{1,2\}$, $k \in \mathcal{K}$, and $z \in \mathcal{Z}$, and $\mathbb{I}_E$ is the indicator variable for the event E; *i.e.*, $\alpha^{\pi}_{skz}(t)$ is the average fraction of time the network was in state (k,z), and the policy π chose the schedule set s up to time t. Under this definition, our joint scheduling/relay selection problem can be formulated as the following utility optimization problem.

$$\text{maximize} \sum_{i\in\mathcal{N}} U_i\left(r_i, \beta_i\right) \quad \text{s.t.} \quad (\mathbf{r}, \beta) \in \mathcal{R}, \quad \beta \in \Lambda, \tag{4.7}$$

where $\mathcal{R}$ is such that $(\mathbf{r}, \beta) \in \mathcal{R}$ if and only if there exists a scheduling policy π such that

$$\liminf_{t\to\infty} \sum_{s:i\in s_1}\sum_{k\in\mathcal{K}}\sum_{z\in\mathcal{Z}} R^{(i)}_{skz}\alpha^{\pi}_{skz}(t) = r_i, \quad \limsup_{t\to\infty} \sum_{s:i\in s_2}\sum_{k\in\mathcal{K}}\sum_{z\in\mathcal{Z}} \alpha^{\pi}_{skz}(t) = \beta_i,$$

almost surely for all $i \in \mathcal{N}$, where

$$s_1 := \{i : (i,j,d) \in s\},$$
$$s_2 := \{j : (i,j,d) \in s, i \neq j\}$$

, and $R^{(i)}_{skz}$ is the rate delivered to user i when $\mathcal{S}_\pi = s, \mathcal{K} = k, \mathcal{Z} = z$, which can be computed based on the results from Section 4.4. Note that in the optimization problem (4.7), the first constraint simply ensures feasibility of the pair $(\mathbf{r}, \beta)$, and the second one imposes the stability constraint for the relay queues, given the conflict graph $\mathcal{G}_c$ between the flows (i,j) available in the network.

4.5.2 Stability Region Structure

Let $\Lambda\left(\mathcal{G}_c\right)$ denote the stability region corresponding to the conflict graph $\mathcal{G}_c$. In general, an explicit characterization of $\Lambda\left(\mathcal{G}_c\right)$ is difficult to obtain. However, it turns out one can

[12] The finiteness assumption is made for technical convenience in proofs; however the proposed scheduling algorithm itself does not rely on this assumption. By assuming a large cardinality, one can model the general case with uncountable alphabets arbitrarily closely.

explicitly obtain a reasonably large inner bound by appropriately inserting edges in the conflict graph, and thus backing off from the optimal stability region. The following theorem characterizes this inner bound.

Theorem 4.3. *Given the conflict graph $\mathcal{G}_c = (\mathcal{V}_c, \mathcal{E}_c)$ and the non-zero link availability probabilities $\{p_{ij}\}$, there exists a polynomial-time algorithm that generates another graph $\bar{\mathcal{G}}_c = \left(\mathcal{V}_c, \bar{\mathcal{E}}_c\right)$ such that $\Lambda\left(\bar{\mathcal{G}}_c\right) \subseteq \Lambda\left(\mathcal{G}_c\right)$, and $\beta \in \Lambda\left(\bar{\mathcal{G}}_c\right)$ if and only if $\beta_Q := \sum_{(i,j)\in Q} \frac{\beta_{ij}}{p_{ij}} \leq 1$ for every maximal clique[13] Q of $\bar{\mathcal{G}}_c$. Further, the number of maximal cliques of $\bar{\mathcal{G}}_c$ is at most n^2, and these cliques can be listed in polynomial time.*

The proof of Theorem 4.3, given in Appendix C.6, relies on standard results from [TE92] specialized to our one-hop network consisting of user pairs, as well as certain graph-theoretic results on perfect graphs, *i.e.*, graphs whose chromatic numbers equal their clique number.

The relay flow control component of our scheduling algorithm uses the inner bound of Theorem 4.3 to ensure the stability of the relay queues. Defining $\bar{\Lambda} := \Lambda\left(\bar{\mathcal{G}}_c\right)$, we reformulate the optimization (4.7) as

$$\text{maximize} \sum_{i\in\mathcal{N}} U_i\left(r_i, \beta_i\right) \quad \text{s.t.}\, (\mathbf{r}, \beta) \in \mathcal{R}, \quad \beta \in \bar{\Lambda}. \tag{4.8}$$

The optimality of the proposed scheduling algorithm is with respect to (4.8).

4.5.3 Optimal scheduling

Let $\mathcal{Q}$ be the set of maximal cliques of $\mathcal{G}_c$. Consider the following policy, which we call π^*: Given $(\mathbf{r}(t-1), \beta(t-1), H(t), \Phi)$, choose the schedule set s^* such that $s^* = \underset{s\subseteq\bar{\mathcal{N}}(t)\times\{1,2\}}{\arg\max} f(s)$, where

$$f(s) = \sum_{(i,j,d)\in s} \mathbb{E}_{Z(t)}\left[R^{(i)}_{sK(t)Z(t)}\middle| K(t)\right] \frac{\partial U_i}{\partial r_i}\Big|_{\substack{r_i=r_i(t-1)\\ \beta_i=\beta_i(t-1)}} + \frac{\partial U_j}{\partial \beta_j}\Big|_{\substack{r_j=r_j(t-1)\\ \beta_j=\beta_j(t-1)}}, \tag{4.9}$$

[13] *A maximal clique is a clique that is not a subset of another clique.*

$\bar{\mathcal{N}}(t) := \{(i,j) \in \mathcal{N}^2 : \beta_Q(t) \leq 1 \text{ for all } Q \in \mathcal{Q} \text{ s.t. } (i,j) \in Q\}$, and $R^{(i)}_{sK(t)Z(t)} = R^{(i)}_{skz}$ with $K(t) = k$ and $Z(t) = z$. Note that $(i,i) \in \bar{\mathcal{N}}(t)$ is vacuously true for all i, corresponding to the scenario where user i is scheduled without relay.

There are a few key points to note in the definition of policy π^*. First, note that the maximization is performed over the available *streams* (i,j,d) in the network, as opposed to over the set of users themselves. Second, at any frame t, any stream (i,j,d) that involves a pair of users (i,j) that is part of a clique Q that currently violates its constraint $\beta_Q(t) \leq 1$ is ignored in the maximization, which is the relay flow control component of the algorithm to ensure stability of the relay queues. Third, the asymptotic optimality of π^* reveals that it is sufficient to average the rate $R^{(i)}_{sK(t)Z(t)}$ over the part of the network state $Z(t)$ that is unknown at the base station, which is consistent with the results in [SCN10].

Theorem 4.4. *Let the optimal value of the maximization in* (4.8) *be* OPT. *Define the empirical utility of* π^* *as* $U^*(t) = \sum_{i \in \mathcal{N}} U_i\left(r^*_i(t), \beta^*_i(t)\right)$, *where* $r^*_i(t)$ *and* $\beta^*_i(t)$ *correspond to variables* $r_i(t)$ *and* $\beta_i(t)$, *respectively, under policy* π^*. *Then the following events hold with probability 1 (*i.e., *almost surely) in the probability space generated by the random network parameters* $K(t)$ *and* $Z(t)$:

1. $\lim_{t\to\infty} \inf\left\{\|\beta^*(t) - \beta\|_1 : \beta \in \bar{\Lambda}\right\} = 0$,

2. $\lim_{t\to\infty} U^*(t) = \mathsf{OPT}$.

The proof outline is provided in Section 4.5.6, with details in Appendix C.2. Theorem 4.4 shows that policy π^* asymptotically achieves the optimum of (4.8).

4.5.4 Greedy implementation

Although converging to the optimal solution, policy π^* suffers from high computational complexity, since it involves an exhaustive search over all subsets of streams. To reduce the complexity, we consider a suboptimal greedy implementation of the policy, similar to [DS05] for non-cooperative MU-MIMO. The algorithm works by iteratively building the schedule set, at each step adding the stream (i^*, j^*, d^*) that contributes the largest amount to the objective $f(s)$, and committing to this choice in the following iterations, until there are no streams left that can result in a utility increment factor of $(1+\epsilon)$ to the existing schedule set (see Algorithm 1). The worst-case complexity of the algorithm is $O(NDn)$, where D is the maximum node degree in $\mathcal{G}$, and N is the maximum number of streams that can be scheduled at a time.

Algorithm 1 Greedy cooperative scheduling

```
$iter = 1$, $schedule\_set = \emptyset$, initialize $\epsilon > 0$.
while $iter \leq N$ do
    $(i^*, j^*, d^*) = \arg\max_{(i,j,d)\in\tilde{\mathcal{N}}(t)\times\{1,2\}} f\left(schedule\_set \cup (i,j,d)\right)$
    $f^*(iter) = f\left(schedule\_set \cup (i^*, j^*, d^*)\right)$
    if $f^*(iter) > (1+\epsilon) f^*(iter-1)$ then
        $schedule\_set = schedule\_set \cup (i^*, j^*, d^*)$
        $iter = iter + 1$
    else
        for all $Q \in \mathcal{Q}$ do
            $\beta_Q(t+1) = update\_clique\_states(\beta_Q(t), schedule\_set)$
        end for
        stop
    end if
end while
```

4.5.5 Choice of utility function

We focus on utility functions of the form[14]

$$U_i\left(r_i, \beta_i\right) = \log(r_i) + \kappa \log(1-\beta_i), \tag{4.10}$$

where $\kappa \geq 0$ is a parameter that controls the trade-off between fairness in throughput and fairness in relaying load. Using the concavity of the objective, it can be shown that (see Appendix C.5 for details) for any feasible pair $(\mathbf{r}, \beta)$, the optimum $\left(\widetilde{\mathbf{r}}, \widetilde{\beta}\right)$ with respect to

[14] Note that this choice means that the function is not defined for $\beta_i = 1$ and $r_i = 0$, but we ignore this since no user will operate at these points.

the objective (4.10) satisfies

$$\sum_i \frac{r_i - \widetilde{r}_i}{\widetilde{r}_i} \leq \kappa \sum_i \frac{(1-\widetilde{\beta}_i) - (1-\beta_i)}{1-\widetilde{\beta}_i}. \tag{4.11}$$

The condition (C.11) admits a meaningful interpretation. Note that the left-hand side represents the sum of the relative gains in throughput due to the perturbation, whereas the right hand-side represents the sum of the relative decrease in time spent idle (not relaying). The condition in (C.11) then suggests that any perturbation to the optimal values will result in a total percentage throughput gain that is less than the total percentage increase in relaying cost, with the parameter κ acting as a translation factor between throughput and relaying cost. This can be considered a generalization of well-studied proportional fairness, which implies that any perturbation to the optimal operating point results in a total percentage throughput loss. Our generalization allows for a positive total relative throughput change, albeit only at the expense of a larger total relative cost increase in relaying. For this utility function, we can evaluate the scheduling rule (4.9) as

$$s^* = \underset{s \subseteq \bar{\mathcal{N}}(t) \times \{1,2\}}{\arg\max} \frac{\mathbb{E}_{Z(t)}\left[R^{(i)}_{sK(t)Z(t)} \middle| K(t) \right]}{r_i(t)} - \frac{\kappa}{1-\beta_{ij}(t)}.$$

4.5.6 Proof outline of Theorem 4.4

We provide the outline for the proof of Theorem 4.4, leaving details to Appendix C.2.

We begin with the first claim. Due to Theorem 4.3, it is sufficient to show that for any maximal clique $Q \subseteq \mathcal{V}_c$, $\limsup \beta^*_Q(t) \leq 1$ almost surely. We state this in the following lemma, whose proof is relatively straightforward and provided in Appendix C.4.

Lemma 4.1. *For all maximal cliques Q of $\mathcal{G}_c$, $\limsup \beta^*_Q(t) \leq 1$ with probability 1 in the probability space generated by $K(t)$ and $Z(t)$.*

The proof of the second claim uses stochastic approximation techniques similar to the main proof in [TG05], but also features several key differences to account for D2D cooperation, multiuser MIMO, partial network knowledge, relay queue stability, and generalized utility functions. To prove the second claim, we first reformulate (4.8) in terms of the variables α_{skz}, as follows

$$\text{maximize } U(\mathbf{y}) := \sum_{i\in\mathcal{N}} U_i \left(\sum_{s:i\in s_1} \sum_{k\in\mathcal{K}} \sum_{z\in\mathcal{Z}} R^{(i)}_{skz}\alpha_{skz}, \sum_{s:i\in s_2} \sum_{k\in\mathcal{K}} \sum_{z\in\mathcal{Z}} \alpha_{skz} \right) \tag{4.12}$$

$$\text{s.t. } \alpha_{skz} \geq 0, \ \sum_{s} \alpha_{skz} \leq p_k q_z, \ \ \alpha_{skz} = q_z \sum_{z'} \alpha_{skz'}, \ \ \forall s, k, z \tag{4.13}$$

$$\sum_{(i,j)\in Q} \sum_{s: \substack{i\in s_1 \\ j\in s_2}} \sum_{k\in\mathcal{K}} \sum_{z\in\mathcal{Z}} \alpha_{skz} \leq 1, \ \ \forall Q \in \mathcal{Q}, \tag{4.14}$$

where $p_k = \mathbb{P}(K(t) = k)$, and $q_z = \mathbb{P}(Z(t) = z)$, where α_{skz} are deterministic; they represent the fraction of time spent in state (s, k, z) throughout the transmission. The last condition in (4.13) reflects the fact that the scheduling decision cannot depend on the realization of $Z(t)$, since this information is not available at the base station.

Lemma 4.2. *Let* OPT' *denote the optimal value of* (4.12). *Then* $\mathsf{OPT}' \geq \mathsf{OPT}$.

Lemma 4.2 is proved in Appendix C.4 using properties of compact sets.

Using Lemma 4.2, it is sufficient to show that $U^{\pi}(t)$ converges to the optimum value of (4.12). We state this in the following lemma, whose proof is provided in Appendix C.2.

Lemma 4.3. $\lim_{t\to\infty} U^*(t) = \mathsf{OPT}'$, *with prob. 1 in the probability space generated by* $(K(t), Z(t))$.

The proof of Lemma B.4 extends the stochastic approximation techniques from [TG05, BGT95] to our setup. In particular, we consider the relaxed version of the optimization problem by augmenting the objective with the stability constraint using a sequence of exponential barrier functions. We then determine the optimal policy for the relaxed problem, and take the limit in the slope of the barrier function to prove the result for the original problem.

Parameter	**Value**	**Parameter**	**Value**
Cellular bandwidth	40MHz	DL carrier freq.	2GHz
D2D bandwidth	40MHz	D2D carrier freq.	5GHz
# BS antennas	32 (linear array)	OFDM FFT size	2048
# UE antennas	1 cell.+1 ISM	Power allocation	equal
Antenna spacing	0.5λ	BS power	46dBm
BS antenna gain	0 dBi	UE power	23dBm
BS antenna pattern	Uniform	Penetration loss	0dB

Table 4.2: System parameters used in the simulations

	Large Cell	**Small/Hetero.**
Inter-site distance ($a\sqrt{3}$)	1732m	500m
No. cells (Ω)	5	19
No. active users/cell (n)	25	10
Cluster radius std. dev. (σ)	20m	10m
Mean # clusters ($\frac{3\sqrt{3}}{2}\lambda a$)	5	3
Utility trade-off param. (κ)	7	8

Table 4.3: Default cell-size-specific parameters

4.6 Numerical Results

4.6.1 Simulation setup

4.6.1.1 Geographic distribution

For the regular network model, we consider a hexagonal grid of Ω cells (see Figure 4.2), each of radius a, with a base station at the center, and n users at each cell. For each cell, we first generate a set of cluster centers according to a homogeneous Poisson point process with intensity λ. Next, we randomly assign each user to a cluster, where user locations for cluster i are chosen i.i.d. according to $\mathcal{CN}(\mathbf{c}_i, \sigma^2\mathbf{I}_2)$, where $\mathbf{c}_i$ is the i'th cluster center, with σ determining how localized the cluster is. In the heterogeneous network model (see Figure 4.3), we place the Ω base stations uniformly at random, generate cluster centers through a homogeneous Poisson process, and assign users to clusters uniformly at random. Next, each user associates with the nearest base station. In both cases, for each set of spatial parameters, we generate eight "drops", *i.e.*, instantiations of user distributions, and

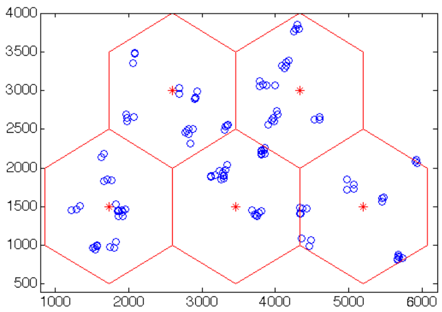


Figure 4.2: Sample geographic distribution of users for large cells..

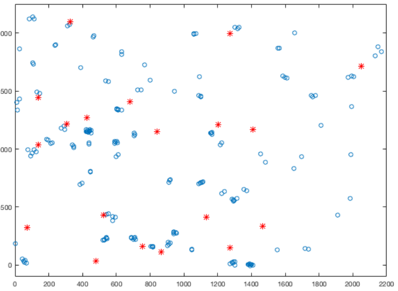


Figure 4.3: Example realization of user and base station realizations for the heterogeneous network model.

the CDFs are computed by aggregating the results across the drops.

4.6.1.2 Channel model

For each (BS, user) pair, we generate a time series of 100 channel vectors for each OFDM subcarrier using the 3GPP Spatial Channel Model (SCM) implementation [SDS05], assuming a user mobility of 3m/s. For each user pair, we use the models from 3GPP D2D Channel Model [TR314] to generate the path loss parameter ϕ_{ij} and the log-normal shadowing parameter χ_{ij}. The channel between the user pair (i, j) for each resource block (RB) is then computed as $\phi_{ij}\chi_{ij}\zeta_{ij}$, where $\zeta_{ij} \sim \mathcal{CN}(0, 1)$ is i.i.d. fading parameter for a given RB. The D2D fading parameters are assumed i.i.d. across RBs. For the main results, we use the line-of-sight (LOS) model, but we also explore the effect of non-line-of-sight links later in the section. For each drop, the channels are computed and stored *a priori*, and all the simulations are run for the same sequences of channel realizations.

4.6.1.3 System operation

Various system parameters are given in Table 4.2. We assume an infinite backlog of data to be transmitted for each user. At every time slot, the base station obtains an estimate of the

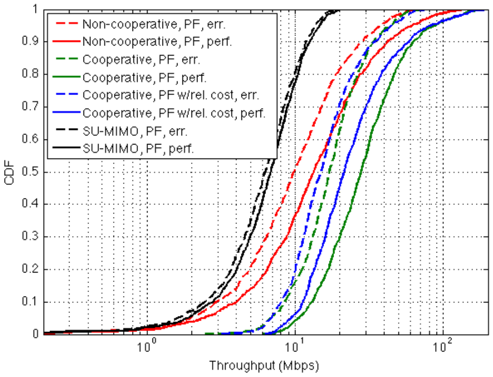


Figure 4.4: Throughput CDF for large cells.

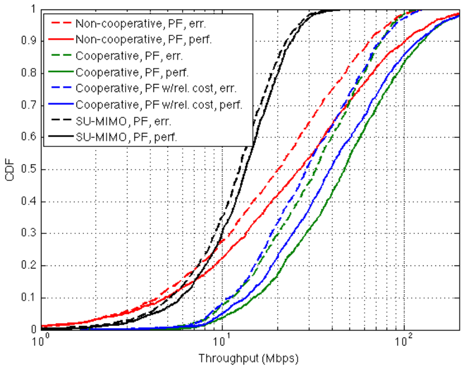


Figure 4.5: Throughput CDF for small cells.

current network state (estimation error modeled normally distributed with variance proportional to the total energy of the channel gains across the OFDM subcarriers, independently for each antenna), and makes a scheduling decision. The scheduling decision is made without knowledge of the inter-cell interference. In the cooperative case, scheduling is done according to Algorithm 4.5.4. In the non-cooperative case, we similarly use the greedy scheduling algorithm of [DS05]. Once the scheduling decision is made, the throughput is computed using the results of Section 4.4 based on the actual channel realizations with inter-cell interference, assuming regularized zero-forcing beamforming, and a 3dB SNR back-off to model practical coding performance. We also take into account various rate back-offs including OFDM cyclic prefix and guard intervals, channel training and uplink data bursts. After the transmission, user throughputs and relaying fractions are updated through exponentially-weighted moving average filters, with averaging window $T_w = 50$ frames.

4.6.2 Throughput distribution for regular cells

For the setup described, we simulate the system with and without cooperation, under the utility function introduced in Section 4.5, as well as conventional proportionally fair (PF) scheduler. We consider large and small cells, with parameters corresponding to either case provided in Table 4.3. For each case, we simulate the system with and without channel estimation errors, using $p_{ij} = 1$ for all (i, j) (we explore smaller values of p_{ij} later in the section).

The CDF of the long-term average throughput received by the users in the network is plotted in Figures 4.4 and 4.5 ("err." represents the case with channel estimation errors, and "perf." represents perfect channel estimation). These plots can be interpreted as a cumulative throughput histogram in the network, where the value on the vertical axis represents the fraction of users who experience a throughput that is less than or equal to the corresponding value on the horizontal axis.

One can observe from Figures 4.4 and 4.5 that, cooperation is most helpful for the weakest (cell-edge) users in the network, providing a throughput gain ranging from 3x up to 4.5x for the bottom fifth-percentile of users depending on cell size, channel estimation quality and utility function used, compared to non-cooperative MU-MIMO. The gain for the median user similarly ranges from 1.4x up to 2.1x depending on the scenario.

When the baseline is taken as non-cooperative SU-MIMO, the fifth percentile gain ranges from 3.5x to 5.7x, whereas the median gain ranges from 2.4x up to 4.1x.

4.6.3 Throughput distribution for heterogeneous networks

We consider the same setup under the heterogeneous network model (Figure 4.3), with the utility function of Section 4.5, and with the same cell-size specific parameters as those for small cells (see Table 4.3). Each user associates with the closest base station, and the resulting CDF is obtained by aggregating the results from independently generated drops, where the base station locations are different across drops. We observe that similar results can be obtained for randomly placed base stations of the heterogeneous model (see Figure 4.6). The fifth-percetile gain is 4.2x, while the median user gain is 1.8x, with respect to non-cooperative MU-MIMO.

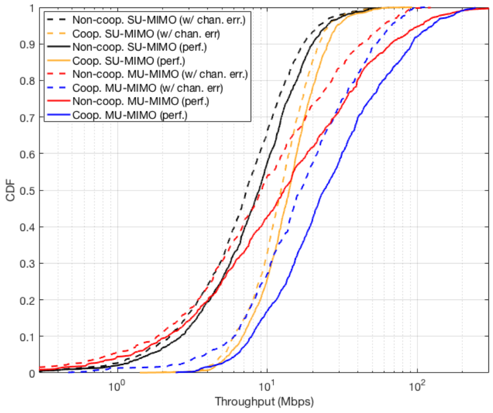


Figure 4.6: Throughput CDF for heterogeneous network.

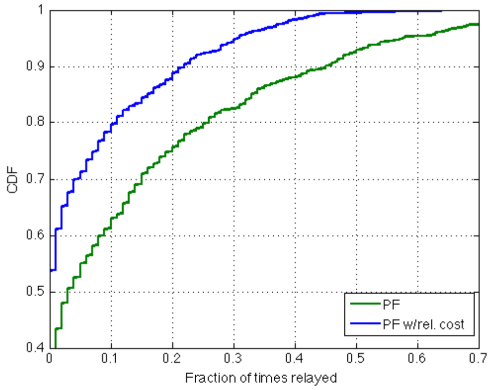


Figure 4.7: CDF for the fraction of time spent relaying for large cells.

4.6.4 Relaying cost

We consider the CDF of the fraction of time a user has performed relaying, for the same runs of simulation as in the previous subsection, in Figure 4.7. In this figure, the values on the vertical axis represent the fraction of users that perform relaying a fraction of time less than or equal to the corresponding value at the horizontal axis, *e.g.*, 90% of users perform relaying less than 22% of the time for PF with relaying cost, and less than 45% of the time for pure PF utility. We observe that our proposed utility function results in more than 50% drop in the total relaying load, with a relatively small penalty in throughput. In particular, the median throughput drop across users is 10%, and the maximum drop is 16%. Therefore, the novel utility function proposed in Section 4.5 enables a more efficient utilization of out-of-band resources, from a throughput-per-channel-access perspective.

4.6.5 D2D link intermittence

We re-run the simulation in Section 4.6.2 for smaller values of p_{ij}. The results are plotted in Figure 4.8, which suggests that the cell-edge gains are fairly robust to external interference of the D2D links, and the gains degrade gracefully with decreasing link availability, resulting in

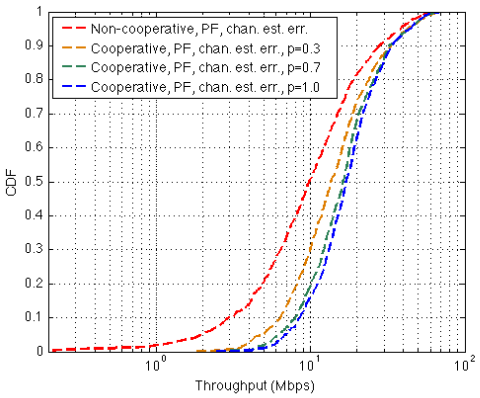


Figure 4.8: Throughput CDF for large cells, for intermittent side-channels.

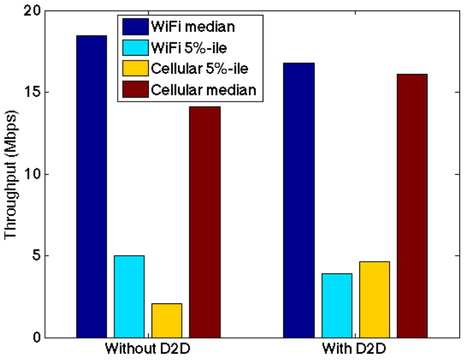


Figure 4.9: Throughput changes in WiFi and cellular users when D2D cooperation is enabled.

approximately 2.5x gain at the bottom fifth percentile even when the links are only available 30% of the time.

4.6.6 Co-existence with WiFi

Since the existing WiFi networks use the same band as D2D cooperation, an important question is whether co-existence of these technologies negates the possible gains due to interference. In this section, we study this scenario through simulations, and demonstrate that the combined overall benefit of WiFi access points (AP) and D2D dominates the loss due to interference, and thus WiFi and D2D cooperation can co-exist harmoniously.

To study this scenario, we consider a network model where an AP is placed at each cluster center $\mathbf{c}_i$. If a user is within the range of a AP, it only gets served by the AP, and is unavailable for D2D cooperation, since the unlicensed band is occupied by AP transmissions and we assume there is constant downlink traffic from the AP. Otherwise, the user is served by the base station and is potentially available for D2D cooperation. In practice, this co-existence mechanism can be implemented through a more aggressive policy, similar to LTE-U: having the user search for an available channel within the unlicensed band for a specified period of time, to use for D2D cooperation, and if none exists, having the user transmit for

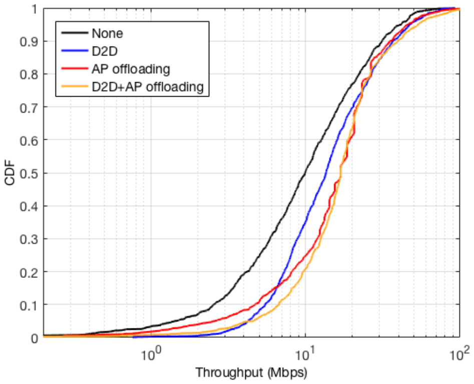


Figure 4.10: Throughput CDF for large cells with APs, with $\sigma = 100m$.

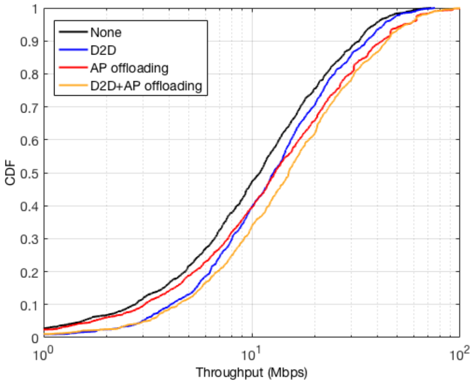


Figure 4.11: Throughput CDF for large cells with APs, with $\sigma = 200m$.

a short duty cycle. Note that the D2D transmissions from outside the AP range can still interfere with the receptions of AP users.

We consider a simplified model for the rates delivered by the AP. If there are ℓ users within the range of a given AP, then a user i at a distance d_i from the AP is offered a rate

$$R_i(t) = \eta J_i(t) \min \left(R\left(d_i\right), \frac{R_{\max}}{\ell} \right),$$

where $R\left(d\right)$ is a function that maps the user distance d from AP to the rate delivered to that user, $R_{\max}$ is the maximum rate the AP can deliver, $0 < \eta \leq 1$ is a back-off factor capturing various overheads in the system, and $J_k(t)$ is the binary variable that takes the value 0 if a neighbor of k in the connectivity graph is transmitting at time t, and 1 otherwise. We use the 802.11ac achievable rates reported in [Bro12] (3 streams, 80MHz, with rates normalized to 40MHz) for the $R\left(d_k\right)$ and $R_{\max}$ values, with $\eta = 0.5$. We reduce the device power to 17dBm for this setup. The throughput CDFs under this setup are given in Figures 4.10 and 4.11. If a user is served by WiFi, its throughput from WiFi is considered; otherwise, its throughput from the D2D-enhanced cellular network is considered.

The results suggest that when D2D cooperation and WiFi AP are simultaneously enabled, the performance is uniformly better than either of them individually enabled, despite the interference from D2D transmissions to AP users, and the relatively fewer D2D opportunities

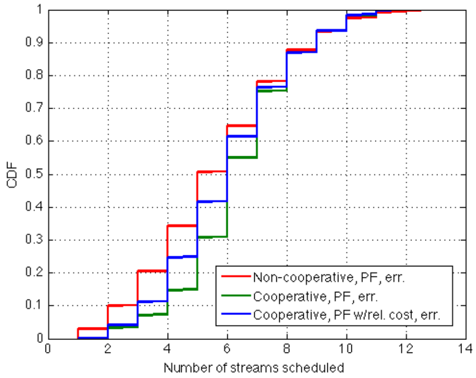


Figure 4.12: CDF for the number of streams scheduled for large cells.

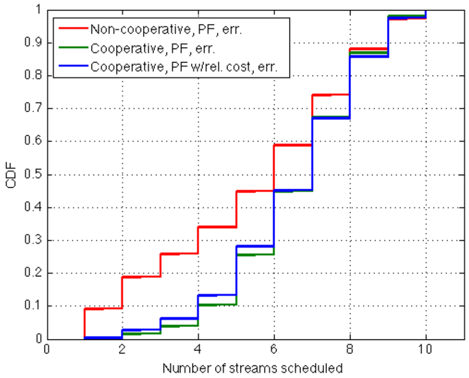


Figure 4.13: CDF for the number of streams scheduled for small cells.

due to users being served by AP. Note that this does not mean that the throughput of a given WiFi user is not reduced when D2D interference takes place (see Figure 4.9, where median WiFi user throughput drops by 10%, while the fifth-percentile cellular user throughput grows by 130%); it means that, if the user falls within the bottom x-percentile after the D2D interference, they are still better off than the bottom x-percentile when only WiFi is enabled. The main reason D2D does not hurt WiFi too much is that D2D cooperation is used for a relatively small fraction of time compared to WiFi for a given user (see Figure 4.7, which shows 80% of users relay less than 10% of the time), which limits the amount of interference. This may also suggest that the more aggressive LTE-U-type policies may also be feasible.

4.6.7 Number of streams scheduled

We compare the number of streams scheduled per time slot for cooperative and non-cooperative cases, in the CDF in Figures 4.12 and 4.13. This can also be understood as the number of steps it takes for Algorithm 1 to terminate.

One can observe that cooperation enables the base station to schedule 1-2 additional streams on average, compared to the non-cooperative case. The reason underlying this behavior is the richness in scheduling options, since data can be transmitted to a particular

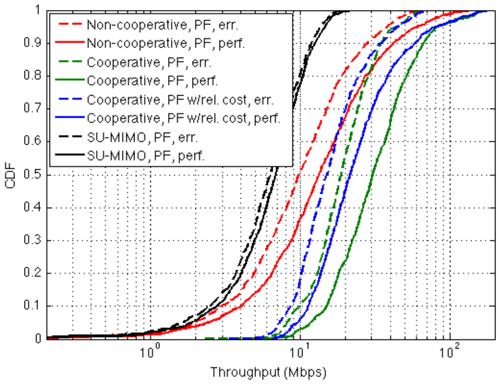


Figure 4.14: Throughput CDF for large cells (without stability constraint).

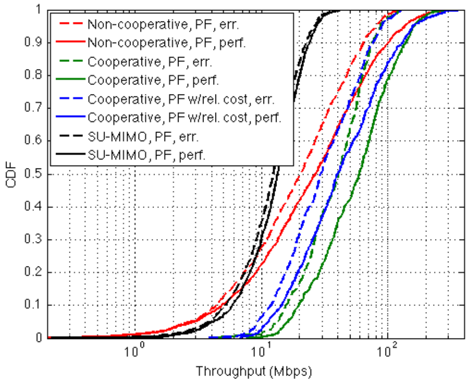


Figure 4.15: Throughput CDF for small cells (without stability constraint).

user through several relaying options, with a distinct beamforming vector corresponding to each option. Since it is easier to find a stream (beamforming vector) that is compatible (approximately orthogonal) with the already scheduled streams, on the average the algorithm is able to schedule a larger number of users per time slot.

4.6.8 Relaxing the stability constraint

In the scenario where the cellular bandwidth is sufficiently smaller than the D2D bandwidth, the interference constraint no longer active, since the devices can perform frequency-division multiplexing to orthogonalize their transmissions. This scenario can be modeled by removing the stability constraint, and performing the maximization in (4.9) over all $\mathcal{N}^2 \times \{1, 2\}$ streams available for scheduling. The resulting throughput CDFs are given in Figures 4.14 and 4.15. Comparing the result to those in Figures 4.4 and 4.5, we see that the stability constraint has a rather small effect on the cooperative cell-edge gains in throughput for large cells, and a relatively larger effect for small cells. This is because the users are located more densely in small cells, and thus the interference (and thus, the stability) constraint is more restrictive. We observe that under this setup, the fifth-percentile gains with respect to SU-MIMO baseline range from 3.5× up to 6.3×, depending on cell size, channel estimation quality and the utility function used. The median gain for large cells reaches almost 4.5×.

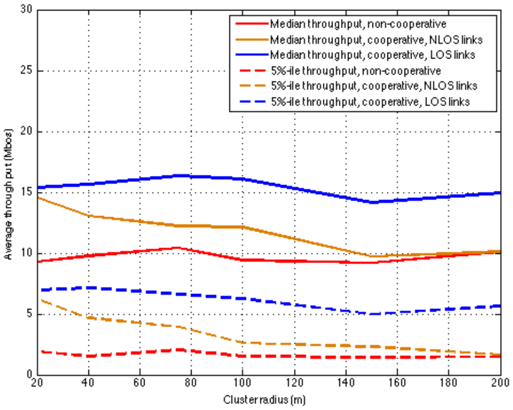


Figure 4.16: Median and 5-percentile throughput vs. cluster radius

The fifth-percentile gains with respect to non-cooperative MU-MIMO are similarly between 3.3× and 4.9×, and the median user gain ranges up to 2.3×.

4.6.9 Effect of clustering

For large cells, we vary the cluster radius σ to study its effect in the throughput CDF in the network. Figure 4.16 plots the throughputs corresponding to the median and the bottom fifth-percentile users in the network, for a range of cluster radii, cooperative and non-cooperative cases, and line-of-sight (LOS) and non-line-of-sight (NLOS) D2D links. We observe that at 23dBm device power, for LOS links, most of the median and fifth-percentile throughput gains are preserved up to a cluster radius of 200m[15]. The decay in throughput is much faster for NLOS D2D links, and the gain completely disappears at a cluster radius of 200m. The performance in a real scenario would be somewhere in between the LOS and NLOS curves, since in a real scenario only a fraction of the links would be LOS.

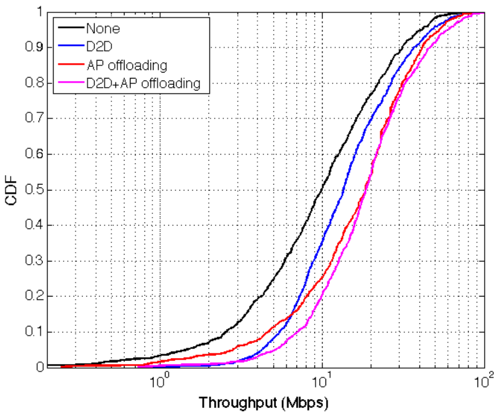


Figure 4.17: Throughput CDF for large cells with APs, with $\sigma = 100m$ (AP users served by AP and base station).

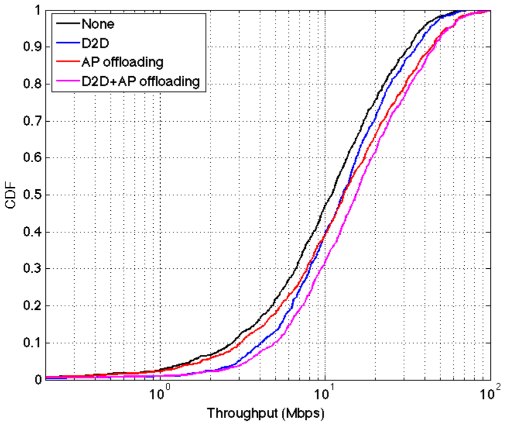


Figure 4.18: Throughput CDF for large cells with APs, with $\sigma = 200m$ (AP users served by AP and base station).

4.6.10 Co-existence with WiFi off-loading

One can also consider an off-loading scenario where the base station continues serving the WiFi users. In this case, the WiFi users are still not available for D2D cooperation, but they can receive from both the AP and directly from the base station whenever they are scheduled based on their past throughputs. We compute the rate delivered to a WiFi user as the sum of the rate that is delivered from the base station (whenever scheduled) and the rate that is delivered from the AP. Figures 4.17 and 4.18 plot the throughput CDFs under this scenario. The results follow a similar pattern to the case where WiFi users are served only by the AP, with a small additional gain in the curves with AP off-loading.

[15] Note that the cluster radius is the standard deviation of user locations from each cluster center. User pairs with pairwise distance much smaller than the cluster radius can still exist within the cluster.

CHAPTER 5

Encoded Distributed Optimization I: Node Failures

5.1 Introduction

Recent years have seen an enormous surge in interest for large-scale data analytics and machine learning. Typically, solving such large problems require storing data over a large number of distributed nodes and running optimization algorithms over these nodes. In such networks, an important concern is the sudden onset of unresponsive or failed nodes [DB13]. This can be caused by network failures, background processes, or (in the case of low-cost cloud computing) sudden deallocation of compute resources. In the case of short-term, or intermittent unavailability, such failures can significantly slow down the computation, since speed may be dictated by the slowest node. In longer-term unavailability, it might affect the accuracy of the final solution itself, since a fraction of data is effectively eliminated from the optimization process. In this chapter, we focus on this latter case.

A natural approach to combat node failure is to use redundancy in the form of additional nodes, for example, by simply replicating the data across multiple nodes. However, recently, distributed *coded* computing has received some attention from the information theory community [LMA16, LLP16, TLD16, DCG16]. In particular, [LLP16] used coding-theoretic ideas to provide robustness in two specific linear operations: distributed matrix multiplication and data shuffling. The work in [DCG16] also focused on linear operations, where the idea is to break up large dot products into shorter dot products, and perform redundant copies of the short dot products to provide resilience against failures. On the other hand, [TLD16] considers synchronous gradient descent, and proposes an architecture where each data sample is replicated s times across nodes, and designs a code such that the exact gradient can be recovered as long as fewer than s nodes fail.

In contrast to these works, which mainly focus on adding redundancy in the *implementation* of a distributed algorithm, we embed the redundancy in the *formulation* of the optimization problem. The idea is to linearly encode the data variables in the optimization, place the encoded data in the nodes, and let the nodes operate as if they are solving the original problem, ignoring failed nodes and stragglers. This is inspired by the randomized sketching techniques [Mah11] used for dimensionality reduction in optimization; however, the purpose, operating regime, and the tools used are different in our problem. The main observation underlying our approach is that one needs much less redundancy than in [TLD16] if one backs off from requiring exact recovery of the solution. For instance, for e node failures, the results in [TLD16] imply that one needs a redundancy factor of $e+1$ for exact recovery, whereas we show that the solution can be reasonably approximated with a redundancy factor of 2. Such relaxation is motivated by fields like machine learning, where approximate solutions that achieve good generalization error are sufficient. The main design objective then becomes how to design codes so that with increasing number of failed nodes, the solution accuracy degrades as slowly as possible. In particular, we observe (numerically and analytically) that equiangular tight frames (ETF) are attractive options as coding vectors, since (i) they contain inherent redundancy; (ii), the individual elements provide as much independent information as possible; and (iii), they allow reconstruction of the exact solution when no nodes fail. We also consider random codes, which asymptotically (data length) achieve good performance; however, as numerical evidence suggests, cannot achieve (iii) for finite lengths.

Our approach is not limited to a specific computational operation, but is applicable directly to large class of practically relevant optimization problems; specifically, any optimization that can be formulated as a least-squares minimization over a convex set, including linear regression, support vector machines, compressed sensing, projection *etc.* Further, since the nodes are oblivious to coding, the existing distributed computing infrastructure and software can be directly used without additional control/coordination messaging.

In this chapter we focus on a model where nodes become unavailable for the time frame of computation, where a failed node does not recover throughout the duration of the computation. This can also be thought of as a model where slow/straggling nodes are the same ones throughout the computation, and these nodes are ignored by the system. The case with asynchronous/intermittent failures and delays is a natural ongoing extension.

Our main contributions are as follows. First, we derive a general bound on relative objective error for encoding with tight frames, and specialize this to equiangular tight frames[1]. Second, using results from analytic number theory, we obtain a tighter bound for a specific construction with redundancy factor 2, which is constructed using Paley graphs [SH03]. To the best of our knowledge, this is the first analysis of the this particular tight frame construction in the context of robustness against erasures. We also present an error bound for random coding vectors. Bounds for other constructions with other redundancy factors are possible. Third, we prove a lower bound on the objective error for the special case of unconstrained least squares optimization. Fourth, we numerically demonstrate performance over three problems, two of which use real world datasets, namely, ridge regression, binary support vector machine classification, and low rank approximation. The results show that the Paley construction outperforms uncoded, replication, and random coding approaches.

The rest of the chapter is organized as follows: Section 5.2 presents our model and metrics of interest, Section 5.3 provides our results on encoding with tight frames and random codes, Section 5.4 gives lower bounds for general linear encoding, and Section 5.5 contains the numerical results on real datasets.

[1] Performance of frames under erasures have been studied in [GKK01, HP04, CK03], though not in the context of convex optimization. Further, these works either focus on exact reconstruction, or only one or two erasures, or otherwise do not provide a general error bound for arbitrary tight frames under arbitrary number of erasures.

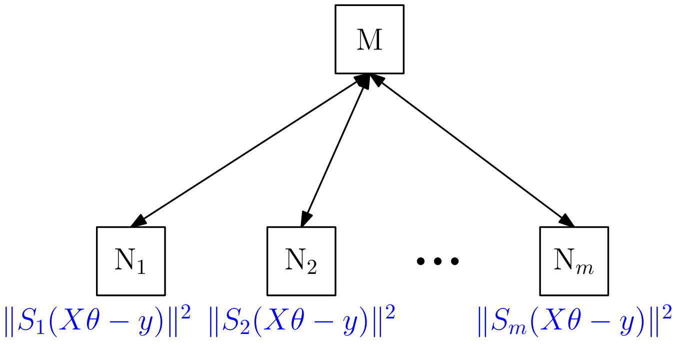


Figure 5.1: A distributed optimization network, where m nodes communicate directly with a centralized server. The local nodes compute terms specific to their data (such as gradients), and the central node aggregates such terms and computes simple steps, like small-dimension projections.

5.2 Model and Notation

Consider the minimization

$$\min_{\theta \in C} g(\theta) := \min_{\theta \in C} \|X\theta - y\|^2, \tag{5.1}$$

where $C \subseteq \mathbb{R}^d$ is an arbitrary convex set (that is globally known), $X \in \mathbb{R}^{n \times d}$ is the data matrix, and $y \in \mathbb{R}^n$ is the data vector. We will denote a solution of this optimization as θ^*.

Consider mapping this optimization problem into a distributed computing setup (see Figure 5.1), where the data variables X_i and y_i are collectively stored across m worker nodes, and a centralized server computes the solution without ever seeing the data itself. Such an architecture is present in most of the popular distributed computing and optimization frameworks [DG08, ZCD12]. Each worker node has sufficient memory to store $\ell(d+1)$ variables (*i.e.*, ℓ rows of data), where $m\ell \geq n$. We define the redundancy factor $\beta := \frac{m\ell}{n} \geq 1$, which captures the amount of additional storage space available. We consider a linear mapping of the data, where worker node $i \in [m]$ stores $Z_i = S_i \left[X \;\; y\right]$, where $S_i \in \mathbb{R}^{\ell \times n}$ is an encoding matrix. We define $S = \left[S_1^\top \;\; S_2^\top \;\; \ldots S_m^\top\right]^\top$. Note that, by setting $S = I_n$, or

$S = [I_n \;\; I_n \;\; \ldots]^\top$, this framework covers uncoded and repetition schemes as well[2].

We assume that after the data placement, a subset $A \subseteq [m]$ of the nodes are unavailable, and the data stored in the unavailable nodes is assumed to be lost throughout the duration of optimization, where $|A| = e$. We define, for a set $U \subseteq [m]$, $S_U = [S_i]_{i \in U}$, *i.e.*, S_U is the submatrix of S corresponding to the set of nodes U, and $A^c = [m] \backslash A$.

Given a mapping S of the data, the worker nodes directly communicate with the centralized server via (two-way) links with no communication constraints, but cannot communicate with each other. The worker nodes are also oblivious to the encoding (*i.e.*, they do not have access to $\{S_i\}$). These two assumptions imply that the nodes effectively attempt to solve the encoded problem $\min_{\theta \in C} \bar{g}(\theta)$, where

$$\bar{g}(\theta) := \|SX\theta - Sy\|^2 = \sum_{i=1}^{m} \|S_i X\theta - S_i y\|^2 \tag{5.2}$$

using any distributed optimization algorithm (*e.g.*, batch or stochastic gradient descent, L-BFGS, proximal gradient descent etc.). Since the objective function (5.2) is a sum of local terms, by having all worker nodes compute, for instance, local gradient terms, and summing them at the centralized server, the centralized solution of (5.2) can be achieved.

We also assume that the available nodes (A^c) are oblivious to the failed nodes (A), and they operate as if all nodes are available. This assumption, and the fact that the failed nodes (A) are unavailable throughout optimization imply that the effective problem whose solution is reached is

$$\min_{\theta \in C} \widetilde{g}(\theta) := \min_{\theta \in C} \|S_{A^c}(X\theta - y)\|^2. \tag{5.3}$$

We denote a solution to (5.3) as $\hat{\theta}(S; X, y; A)$. Given an encoding matrix S, data variables (X, y), and a failure pattern A, the *relative error* $\eta^*(S; X, y; A)$ is defined as the smallest

[2] From a technical standpoint, such linear encoding resembles the sketching technique [Mah11] used to approximate optimization problems by dimensionality reduction. However, sketching uses randomized, short and wide S matrices for dimensionality reduction; we use tall, deterministic S matrices to *increase* the problem dimensions and add redundancy.

$\eta \geq 1$ such that

$$\|X\hat{\theta} - y\|^2 \leq \eta \|X\theta^* - y\|^2.$$

For a given S, the *worst-case relative error* is given by

$$\gamma(S, e) := \sup_{X,y} \max_{A:|A|=e} \eta^*(S; X, y; A).$$

Our goal is to design a matrix S such that $\gamma(S, e)$ is minimized and grows slowly with e, *i.e.*, whose worst-case relative error degrades gracefully with increasing number of failed nodes.

5.3 Encoded Distributed Convex Programs

Intuitively, one would expect a good encoding matrix S to satisfy a number of properties. First, it must contain some form of redundancy in its set of encoding vectors (the rows $s_i^\top$ of S). Second, drawing from the intuition of the channel coding theorem, individual encoding vectors must provide as much independent information as possible. Third, the encoding matrix should not *add* error; that is, when there are no failures, the exact solution must be recoverable, assuming nodes are oblivious to coding.

5.3.1 Equiangular tight frames

Given such requirements, we turn to *equiangular tight frames* (ETF) as a natural choice of set of encoding vectors. Loosely speaking, ETFs constitute an overcomplete basis for $\mathbb{R}^n$, and whose individual elements are as decorrelated as possible. More formally, a (unit-norm) tight frame for $\mathbb{R}^n$ is a set $\{h_i\}_{i=1}^{n\beta} \subseteq \mathbb{R}^n$ of unit vectors (with $\beta \geq 1$), such that for any $u \in \mathbb{R}^n$,

$$\sum_{i=1}^{n\beta} |\langle h_i, u \rangle|^2 = \beta \|u\|^2. \tag{5.4}$$

The reader is referred to [Dau92, SH03] for more information on frames.

Define the maximal inner product of a tight frame H by

$$\epsilon(H) := \max_{\substack{h_i, h_j \in H \\ i \neq j}} |\langle h_i, h_j \rangle| .$$

A tight frame for which $|\langle h_i, h_j \rangle| = \epsilon(H)$ for every $i \neq j$ is called an *equiangular tight frame* (ETF).

Proposition 5.1 (Welch bound, [Wel74])**.** *Let $H = \{h_i\}_{i=1}^{n\beta}$ be a tight frame. Then $\epsilon(H) \geq \sqrt{\frac{\beta-1}{2n\beta-1}}$. Moreover, equality is satisfied if and only if H is an equiangular tight frame.*

Therefore, an ETF minimizes the correlation between its individual elements.

We define the tangent cone of the constraint set at the optimum by

$$\mathcal{K} := \operatorname{clconv} \left\{ u \in \mathbb{R}^d : u = t(\theta - \theta^*), t \geq 0, \theta \in C \right\},$$

where clconv denotes closure of the convex hull, and the linearly transformed cone is defined by $X\mathcal{K} := \{Xu : u \in \mathcal{K}\}$. We also define, for a set $\mathcal{U}$, and a symmetric matrix P,

$$\lambda_{\max}^{\mathcal{U}}(P) = \sup_{u \in \mathcal{U},\ \|u\|_2 = 1} \|Pu\|_2.$$

The case $\lambda_{\max}^{\mathbb{R}^n}(P) = \lambda_{\max}(P)$, the largest eigenvalue of P in absolute value (which is the spectral norm, since P is symmetric).

Our first result bounds the relative error under encoding with tight frames.

Theorem 5.1. *Let S be such that $\{s_i\}_{i=1}^{n\beta}$ is a tight frame over $\mathbb{R}^n$. Then for any encoded optimization problem in the form* (5.3)*,*

$$\eta^*(S; X, y; A) \leq \min_{0 \leq c \leq \beta} \left(1 + \frac{2\lambda_{\max}^{X\mathcal{K}} \left(S_A^\top S_A - cI \right)}{\beta - \lambda_{\max} \left(S_A^\top S_A \right)} \right)^2 .$$

Corollary 5.1. *Under the setup of Theorem 5.1,*

$$\gamma(S, e) \leq \left(\frac{\beta}{\beta - \max_{A : |A| = e} \left\| S_A^\top S_A \right\|_2} \right)^2 .$$

The proofs are given in Appendix D.1, which relies on techniques from [PW15], as well as convex optimality conditions and properties of tight frames. Note that the bound only depends on the spectral properties of the *lost* component of the encoding matrix S.

Theorem 5.1 and Corollary 5.1 show that when one encodes the data with tight frames, worst-case relative error can be uniformly bounded, and the error depends on the spectral properties of the relevant submatrices S_A of the encoding matrix. We note that (as expected), as the redundancy factor β grows, relative error goes to 1, and when $e = 0$, it is exactly 1, which implies perfect recovery when no failures occur. Note that this is not necessarily true for an arbitrary matrix S whose Gram matrix $S^\top S$ has non-zero eigenvalue spread, including random matrices. We also note that to minimize the error, one must design S such that any possible submatrix S_A has spectral norm close to 1.

Next we prove explicit bounds for *equiangular* tight frames, by bounding the spectral norm of the submatrices S_A. Although these bounds are non-trivial, numerical evidence suggests that tighter bounds may hold.

Theorem 5.2. *If the rows of S form an equiangular tight frame, then for $1 \leq e < \frac{\beta-1}{\alpha(m,n)}$,*

$$\gamma(S, e) \leq \left(\frac{\beta}{\beta - 1 - e\alpha(m,n)} \right)^2,$$

where $\alpha(m,n) = \frac{1}{m}\sqrt{\frac{n\beta(\beta-1)}{1-(n\beta)^{-1}}}$.

See Appendix D.2 for proof. For a specific construction obtained by using Paley conference matrices [SH03], we can in fact prove a tighter result that holds with high probability (under random failures). Let q be a prime number such that $q \equiv 1 \pmod 4$, and let $\mathbb{F}_q$ be the finite field of size q. Consider the graph G_q whose vertices are the elements of $\mathbb{F}_q$, and the elements $a \neq b$ are adjacent if and only if there exists $r \in \mathbb{F}_q$ such that $a - b \equiv r^2 \pmod q$ (in which case $a - b$ is called a *quadratic residue*, and G_q is known as Paley graph). It can be shown that [SH03] if A_{q+1} is the 0-1 adjacency matrix of the graph formed by combining G_q with an isolated node u, then the matrix $M_{q+1} := \frac{1}{\sqrt{q}}(J_{q+1} - I_{q+1} - 2A_{q+1}) + I_{q+1}$, where

J_{q+1} is the all-ones matrix, can be decomposed as

$$M_{q+1} = S_{q+1} S_{q+1}^{\top},$$

where the rows of S_{q+1} form an equiangular tight frame with $\epsilon(S_{q+1}) = \frac{1}{\sqrt{q}}$. Using number-theoretic results on multiplicative quadratic residue characters in finite fields (see Appendix D.3), we can obtain the following tighter bound for this construction.

Theorem 5.3. *Let $\breve{S}$ be an ETF constructed from Paley graph as above, where $q+1 = 2n$ (so that redundancy factor $\beta = 2$). Let $S = P\breve{S}$, where P is a random permutation matrix that is drawn uniformly random over all $(2n)!$ permutation matrices. Let A be uniformly random over all cardinality-e subsets of $[m]$. Then for $1 \leq e < \left(\frac{1}{c\tilde{\alpha}(m,n)}\right)^{4/3}$ and for any $c > 1$,*

$$\mathbb{P}\left(\eta(S; X, y; A) > \left(\frac{2}{1 - ce^{3/4}\tilde{\alpha}(m,n)}\right)^2\right) \leq \frac{1}{c^4},$$

where $\tilde{\alpha}(m,n) := \sqrt{\frac{2}{m - \frac{1}{\ell}}}\left(\frac{2n}{m}\right)^{1/4}$.

To the best of our knowledge, Theorem 5.3 is the first analysis of the erasure-robustness of Paley ETFs. This result shows that if we scale the number of nodes m faster than $n^{\frac{1}{3}}$, then the error is small with high probability, even under a large number of node failures. In fact, based on numerical evidence, we believe the following, even tighter, deterministic bound holds for this construction.

Conjecture 5.1. *If S is an ETF constructed from Paley graph as above, where $q+1 = 2n$, then for $1 \leq e < \frac{1}{\tilde{\alpha}^2(m)}$,*

$$\gamma(S, e) \leq \left(\frac{2}{1 - \sqrt{e}\tilde{\alpha}(m)}\right)^2,$$

where $\tilde{\alpha}(m) := \frac{c}{\sqrt{m}}$ for a universal constant c.

Note that there is no dependence on n in this bound.

5.3.2 Random coding

Another natural approach in designing S could be choosing its elements i.i.d. random, *e.g.*, with Gaussian entries. In particular, using results from [PW15], and the scaling behavior of singular values of i.i.d. Gaussian matrices [Sil85], it can be shown that the following holds (the details are in Appendix D.4).

Proposition 5.2. *For fixed $\beta = \frac{m\ell}{n}$, consider a family of encoding matrices $S_m \in \mathbb{R}^{m\ell \times \frac{m\ell}{\beta}}$, indexed by the number of worker nodes m. Choose all entries of S_m i.i.d. from $N\left(0, \frac{1}{n}\right)$. Denote the relative error for m machines as $\eta^*_m(S_m; X, y; A)$, for any A with $|A| = e < m^{\frac{\beta-1}{\beta}}$. Then, for any (X, y),*

$$\lim_{m\to\infty} \eta^*_m(S_m; X, y; A) \leq \left(\frac{\sqrt{\beta\left(1 - \frac{e}{m}\right)} + 1}{\sqrt{\beta\left(1 - \frac{e}{m}\right)} - 1} \right)^4 .$$

Note that random coding can achieve a bound independent of n as well, albeit asymptotically. In practice, however, we observe that the spectral norm of submatrices of Paley ETF grows slower than those of i.i.d. random matrices, and thus Paley ETF achieves a slightly tighter bound on relative error for finite data, as claimed in Conjecture 5.1, and further evidenced in the results of the next section.

5.4 Lower Bound for Unconstrained Optimization

Given the results of Section 5.3, one may wonder how they compare with the performance of other possible encoding techniques. In this section, we derive a lower bound on the relative error for unconstrained optimization ($C = \mathbb{R}^d$) for arbitrary linear encoding. The bound is not necessarily tight, but it still provides insight into how one should design the encoding matrix.

Theorem 5.4. *For any encoding matrix S, worst-case relative error for unconstrained op-*

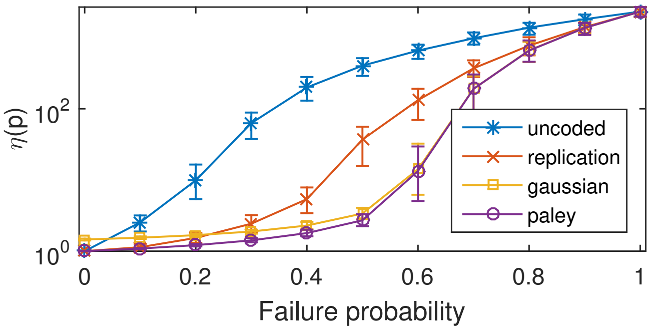


Figure 5.2: Performance for ridge regression, where X is 1000×750 and $\mu = 0.1$. There are 750 processors and S has 2000 rows.

timization is lower bounded by

$$\gamma(S, e) \geq \frac{1}{4}(1 + \max_{A:|A|=e} \kappa(S_{A^c}))^2$$

where $\kappa(Q)$ is the condition number of matrix Q.

The proof is provided in Appendix D.5, which is based on constructing an adversarial data pair (X, y) for any given encoding matrix S. Theorem 5.4 implies that in order to control the error, one needs to design the encoding matrix so that any relevant submatrix S_{A^c} is well-conditioned, which is similar to the restricted isometry condition in compressed sensing [CT05].

5.5 Numerical Results

We explore three machine learning problems, two of which use real world datasets. In each example, we compare four cases: uncoded ($S = I_n$), replication code, Gaussian ($S_{ij} \sim \mathcal{N}(0, 1)$), and Paley ETF. The redundancy factor $\beta = 2$ in each case except the uncoded one. In the simulations, we consider probabilistic availability of the nodes, where each node independently fails with probability p. In each case we plot relative error ($\eta(p)$, representing relative error at failure probability p) over 100 trials with different failure patterns, with error bars at a 95% interval.

5.5.1 Ridge regression

The encoded ridge regression problem solves

$$\underset{\theta}{\text{minimize}} \quad \|S(X\theta - y)\|_2^2 + \mu\|\theta\|_2^2, \tag{5.5}$$

where $\mu > 0$ is a regularization parameter. The rows of X and y represent data feature vectors and labels respectively, and the entries of the solution θ^* are the feature regressors.

Figure 5.2 shows the relative error performance with respect to failure probability, where $y = Xz + n$ and each element of X, y, and n is drawn independently from a Gaussian distribution. The data matrix X is 1000×750 and the generated encoding matrices have 1000 (uncoded), 2000 (replication, Gaussian, Paley[3]) rows. The problem is solved using gradient descent, where each worker node computes gradient terms corresponding to their own data and the central node only performs the aggregation and descent step.

5.5.2 Binary SVM classification

The MNIST dataset contains 28×28 binary images for handwritten digits 0-9 [LCB98]. We attempt to disambiguate 4's from 9's using binary support vector machines, by solving the reformulation suggested by [PW15, §3.4]:

$$\begin{array}{ll} \underset{\theta}{\text{minimize}} & \|W^T \text{diag}(d)\theta\|_2^2 + \mu\|\theta\|_2^2 = \|X\theta\|_2^2 \\ \text{subject to} & \sum_i \theta_i = 1, \;\; \theta_i \geq 0, \forall i. \end{array} \tag{5.6}$$

The rows W_i are ith vectorized binary images (demeaned), and $d_i \in \{1, -1\}$ indicates if the ith sample is a 4 or 9. The objective can be reformulated with $X = [\text{diag}(d)W, I]^T$, and the encoded problem has objective $\|SX\theta\|_2^2$.

We reduce the MNIST train and test dataset to only the digits 4 and 9, and additionally only use the first 1000 train samples ($W \in \mathbb{R}^{1000 \times 784}$). Fig. 5.3 shows the relative error

[3] Since Paley ETF has size $(q+1) \times (q+1)/2$ for prime q, we take the smallest prime s.t. $q \equiv 1 \pmod 4$ (in this case, 2017) larger than the required dimension, and take an arbitrary submatrix that matches the required dimensions. The error due to this subsampling is negligible.

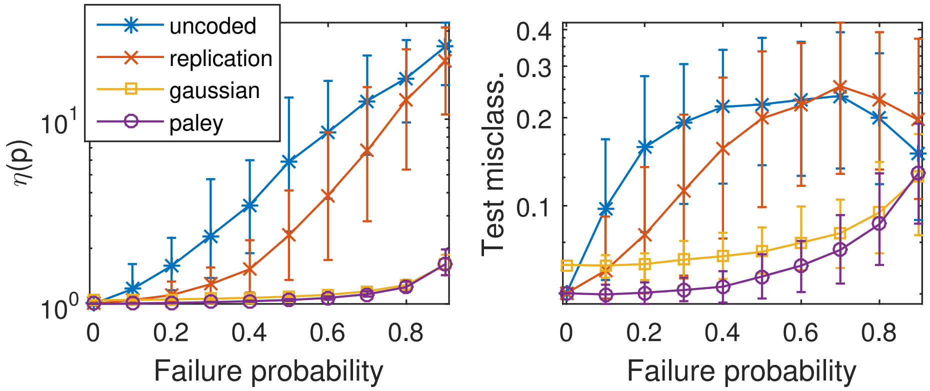


Figure 5.3: Performance for solving SVM on reduced MNIST set for 4 vs. 9 disambiguation. Here X is 1000×784 and $\mu = 0.1$. There are 500 processors.

performance, where (5.6) is solved using FISTA [BT09], where the worker nodes evaluate gradients and the centralized server aggregates terms and computes the projection on the simplex.

5.5.3 Low-rank approximation

The movielens ml-100k dataset [RK98] contains recommendations of users for movies. The task is, given ratings in a training set, predict the ratings in a separate test set. Given rating matrix R, where R_{ij} is the rating user i provided movie j (if exists in the training set), and find the nearest low rank approximate matrix completion of R. The following is an encoded version of a popular convex approximation of the rank-constrained matrix completion problem:

$$\begin{array}{ll} \underset{\Theta}{\text{minimize}} & \|SX\mathbf{vec}(\Theta - R)\|_F^2 \\ \text{subject to} & \|\Theta\|_* \leq \tau. \end{array} \tag{5.7}$$

Here, $\|Z\|_*$ is the nuclear norm (sum of the singular values of Z) and serves as a convex proxy for rank. The matrix X is such that $X\mathbf{vec}(R)$ selects only the provided ratings R_{ij}.

We subsample the movielens dataset to leave only users and movies that contribute the most ratings, resulting in 133 users and 56 movies, with 5,514 provided ratings evenly split between train and test sets. (Resulting X is 2757×7448.) (5.7) is solved using FISTA

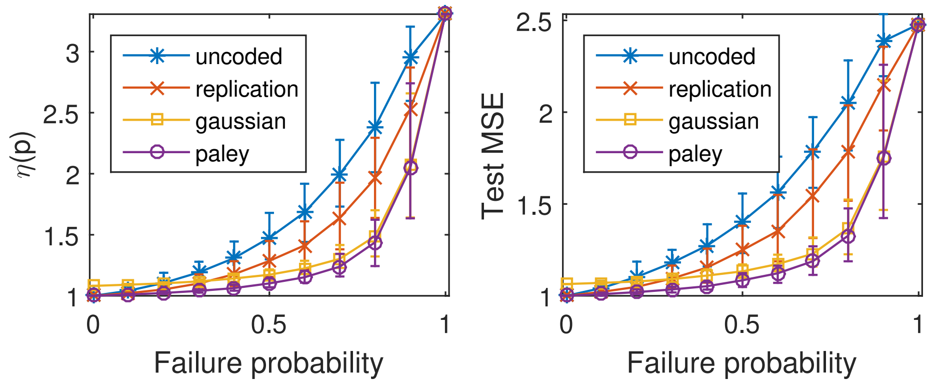


Figure 5.4: Performance for solving the matrix completion problem with the subsampled movielens dataset. Here, X is 2757×7448, and $\tau = 100$. There are 100 processors and S has twice the number of rows as X.

[BT09], with $\tau = 100$. Figure 5.4 shows relative error results and the mean squared error in test ratings, defined as $\frac{1}{|T|}\sum_{(i,j)\in T}((R_{\text{test}})_{ij} - X_{ij})^2$ where R_{test} is the test ratings matrix and T contains the (user, movie) pairs included in the test set.

In all three examples, it is clear that coding increases robustness in the presence of large numbers of node failures, both in the relative error of the objective and in test error metrics on real datasets. The tightness of the Paley frames is also observed; in all cases there is no degradation when no nodes fail, which is not true when using random encoding matrices.

CHAPTER 6

Encoded Distributed Optimization II: Node Delays

6.1 Introduction

Solving learning and optimization problems at present scale often requires parallel and distributed implementations to deal with otherwise infeasible computational and memory requirements. However, such distributed implementations often suffer from system-level issues such as slow communication and unbalanced computational nodes. The runtime of many distributed implementations are therefore throttled by that of a few slow nodes, called stragglers, or a few slow communication links, whose delays significantly encumber the overall learning task. In this chapter we further develop the encoded distributed optimization framework of the previous chapter, and generalize it to the case of node delays, with the aim of mitigating the effect of straggler nodes. In particular, we propose a distributed optimization framework based on proceeding with each iteration without waiting for the stragglers, and encoding the dataset across nodes to add redundancy in the system in order to mitigate the resulting potential performance degradation due to lost updates.

We consider the master-worker architecture, where the dataset is distributed across a set of worker nodes, which directly communicate to a master node to optimize a global objective. The encoding framework consists of an efficient linear transformation (coding) of the dataset that results in an overcomplete representation, which is then partitioned and distributed across the worker nodes. The distributed optimization algorithm is then performed directly on the encoded data, with all worker nodes oblivious to the encoding scheme, *i.e.*, no explicit decoding of the data is performed, and nodes simply solve the effective optimization problem after encoding. In order to mitigate the effect of stragglers, in each iteration, the master node only waits for the first k updates to arrive from the m worker nodes (where $k \leq m$ is a design parameter) before moving on; the remaining $m - k$ node results are effectively erasures, whose loss is compensated by the data encoding.

The framework is applicable to both the data parallelism and model parallelism paradigms of distributed learning, and can be applied to distributed implementations of several popular optimization algorithms, including gradient descent, limited-memory-BFGS, proximal gradient, and block coordinate descent. We show that if the linear transformation is designed to satisfy a spectral condition resembling the restricted isometry property, the iterates resulting from the encoded version of these algorithms deterministically converge to an exact solution for the case of model paralellism, and an approximate one under data parallelism, where the approximation quality only depends on the properties of encoding and the parameter k. These convergence guarantees are deterministic in the sense that they hold for any pattern of node delays, *i.e.*, even if an adversary chooses which nodes to delay at every iteration. In addition, the convergence behavior is independent of the tail behavior of the node delay distribution. Such a worst-case guarantee is not possible for the asynchronous versions of these algorithms, whose convergence rates deteriorate with increasing node delays. We point out that our approach is particularly suited to computing networks with a high degree of variability and unpredictability, where a large number of nodes can delay their computations for arbitrarily long periods of time.

Our contributions are as follows: (i) We propose the encoded distributed optimization framework, and prove deterministic convergence guarantees under this framework for gradient descent, L-BFGS, proximal gradient and block coordinate descent algorithms; (ii) we provide three classes of encoding matrices, and discuss their properties, and describe how to efficiently encode with such matrices on large-scale data; (iii) we implement the proposed technique on Amazon EC2 clusters and compare their performance to uncoded, replication, and asynchronous strategies for problems such as ridge regression, collaborative filtering, logistic regression, and LASSO. In these tasks we show that in the presence of stragglers, the technique can result in significant speed-ups (specific amounts depend on the underlying system, and examples are provided in Section 6.6) compared to the uncoded case when all workers are waited for in each iteration, to achieve the same test error.

The rest of the chapter is organized as follows. In Section 6.2, we review the relevant literature and contrast our work. In Section 6.3, we present our encoded distributed optimization framework that is generalized for node delays, and present the main optimization algorithms we consider. In Section 6.4, we prove our analytical convergence results. In Section 6.5, we discuss several families of code constructions, and discuss how to efficiently implement encoding. In Section 6.6, we present our numerical results from our distributed implementations on Amazon EC2 clusters.

6.2 Related work

The approaches to mitigating the effect of stragglers can be broadly classified into three categories: replication-based techniques, asynchronous optimization, and coding-based techniques.

Replication-based techniques consist of either re-launching a certain task if it is delayed, or pre-emptively assigning each task to multiple nodes and moving on with the copy that completes first. Such techniques have been proposed and analyzed in [GZD15, AGS13, SLR16, WJW15, YHG16], among others. Our framework does not preclude the use of such system-level strategies, which can still be built on top of our encoded framework to add another layer of robustness against stragglers. However, it is not possible to achieve the worst-case guarantees provided by encoding with such schemes, since it is still possible for both replicas to be delayed.

Perhaps the most popular approach in distributed learning to address the straggler problem is asynchronous optimization, where each worker node asynchronously pushes updates to and fetches iterates from a parameter server independently of other workers, hence the stragglers do not hold up the entire computation. This approach was studied in [RRW11, AD11, DCM12, LAP14] (among many others) for the case of data parallelism, and [LWR15, YLL16, PXY16, SHY17] for coordinate descent methods (model parallelism). Although this approach has been largely successful, all asynchronous convergence results depend on either a hard bound on the allowable delays on the updates, or a bound on the moments of the delay distribution, and the resulting convergence rates explicitly depend on such bounds. In contrast, our framework allows for completely unbounded delays. Further, as in the case of replication, one can still consider asynchronous strategies on top of the encoding, although we do not focus on such techniques within the scope of this work.

A more recent line of work that address the straggler problem is based on coding-theory-inspired techniques [TLD17, LLP16, DCG16, KSD17a, KSD17b, YGK17, HAS17, RPP17]. Some of these works focus exclusively on coding for distributed linear operations, which are considerably simpler to handle. The works in [TLD17, HAS17] propose coding techniques for distributed gradient descent that can be applied more generally. However, the approach proposed in these works require a redundancy factor of $r + 1$ in the code, to mitigate r stragglers. Our approach relaxes the exact gradient recovery requirement of these works, consequently reducing the amount of redundancy required by the code.

The proposed technique, especially under data parallelism, is also closely related to randomized linear algebra and sketching techniques in [Mah11, DMM11, PW15], used for dimensionality reduction of large convex optimization problems. The main difference between this literature and the proposed coding technique is that the former focuses on reducing the problem dimensions to lighten the computational load, whereas encoding *increases* the dimensionality of the problem to provide robustness. As a result of the increased dimensions, coding can provide a much closer approximation to the original solution compared to sketching techniques. In addition, unlike these works, our model allows for an arbitrary convex regularizer in addition to the encoded loss term.

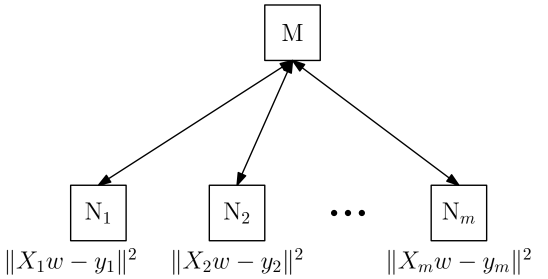


Figure 6.1: Uncoded distributed optimization with data parallelism, where X and y are partitioned as $X = [X_i]_{i\in[m]}$ and $y = [y_i]_{i\in[m]}$.

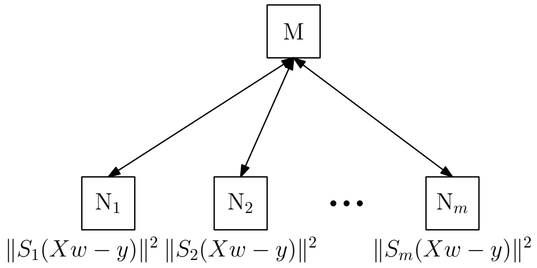


Figure 6.2: Encoded setup with data parallelism, where node i stores (S_iX, S_iy), instead of (X_i, y_i). The uncoded case corresponds to $S = I$.

6.3 Encoded Distributed Optimization for Straggler Mitigation

We will use the notation $[j] = \{i \in \mathbb{Z} : 1 \leq i \leq j\}$. All vector norms refer to 2-norm, and all matrix norms refer to spectral norm, unless otherwise noted. The superscript ^{c} will refer to complement of a subset, *i.e.*, for $A \subseteq [m]$, $A^c = [m]\backslash A$. For a sequence of matrices $\{M_i\}$ and a set A of indices, we will denote $[M_i]_{i\in A}$ to mean the matrix formed by stacking the matrices M_i vertically. The main notation used throughout the chapter is provided in Table 6.1.

We consider a distributed computing network where the dataset $\{(x_i, y_i)\}_{i=1}^n$ is stored across a set of m worker nodes, which directly communicate with a single master node. In practice the master node can be implemented using a fully-connected set of nodes, but this can still be abstracted as a single master node.

It is useful to distinguish between two paradigms of distributed learning and optimization; namely, data parallelism, where the dataset is partitioned across data samples, and model parallelism, where it is partitioned across features (see Figures 6.1 and 6.3). We will describe these two models in detail next.

Notation	**Explanation**
$[j]$	The set $\{i \in \mathbb{Z} : 1 \leq i \leq j\}$
m	Number of worker nodes
n, p	The dimensions of the data matrix $X \in \mathbb{R}^{n \times p}$, vector $y \in \mathbb{R}^{n \times 1}$
k_t	Number of updates the master node waits for in iteration t
η_t	Fraction of nodes waited for in iteration, *i.e.*, $\eta_t = \frac{k_t}{m}$
A_t	Subset of nodes $[m]$ which send the fastest k_t updates at iteration t
$f(w), \widetilde{f}(w)$	Original and encoded objectives, respectively, under data parallelism
$g(w) = \phi(Xw)$	Original objective under model parallelism
$\widetilde{g}(v) = \phi(XS^\top v)$	The encoded objective under model parallelism
$h(w)$	Regularization function (potentially non-smooth)
ν	Strong convexity parameter
L	Smoothness parameter for $h(w)$ (if smooth), and $g(w)$
λ	Regularization parameter
Ψ_t	Mapping from gradient updates to step $\{\nabla f_i(t)\}_{i \in A_t} \mapsto d_t$
d_t	Descent direction chosen by the algorithm
α_t, α	Step size
M, μ	Largest and smallest eigenvalues of $X^\top X$, respectively
β	Redundancy factor ($\beta \geq 1$)
S	Encoding matrix with dimensions $\beta n \times n$
S_i	ith row-block of S, corresponding to worker i
S_A	Submatrix of S formed by $\{S_i\}_{i \in A \subseteq [m]}$ stacked vertically

Table 6.1: Notation used in the chapter.

6.3.1 Data parallelism

We focus on objectives of the form

$$f(w) = \frac{1}{2n}\|Xw - y\|^2 + \lambda h(w), \tag{6.1}$$

where X and y are the data matrix and data vector, respectively. We assume each row of X corresponds to a data sample, and the data samples and response variables can be horizontally partitioned as $X = \begin{bmatrix} X_1^\top \; X_2^\top \; \cdots \; X_m^\top \end{bmatrix}^\top$ and $y = \begin{bmatrix} y_1^\top \; y_2^\top \; \cdots \; y_m^\top \end{bmatrix}^\top$. In the uncoded setting, machine i stores the row-block X_i (Figure 6.1). We denote the largest and smallest eigenvalues of $X^\top X$ with $M > 0$, and $\mu \geq 0$, respectively. We assume $\lambda \geq 0$, and $h(w) \geq 0$ is a convex, extended real-valued function of w that does not depend on data. Since $h(w)$ can take the value $h(w) = \infty$, this model covers arbitrary convex constraints on the optimization.

The encoding consists of solving the proxy problem

$$\widetilde{f}(w) = \frac{1}{2n}\|S\left(Xw - y\right)\|^2 + \lambda h(w) = \frac{1}{2n}\sum_{i=1}^{m} \underbrace{\|S_i\left(Xw - y\right)\|^2}_{f_i(w)} + \lambda h(w), \tag{6.2}$$

instead, where $S \in \mathbb{R}^{\beta n \times n}$ is a designed encoding matrix with redundancy factor $\beta \geq 1$, partitioned as $S = \begin{bmatrix} S_1^\top \; S_2^\top \; \cdots \; S_m^\top \end{bmatrix}^\top$ across m machines. Based on this partition, worker node i stores $(S_i X, S_i y)$, and operates to solve the problem (6.2) in place of (6.1) (Figure 6.2). We will denote $\hat{w} \in \arg\min \widetilde{f}(w)$, and $w^* \in \arg\min f(w)$.

In general, the regularizer $h(w)$ can be non-smooth. We will say that $h(w)$ is L-smooth if $\nabla h(w)$ exists everywhere and satisfies

$$h(w') \leq h(w) + \langle \nabla h(w), w' - w \rangle + \frac{L}{2}\|w' - w\|^2$$

for some $L > 0$, for all w, w'. The objective f is ν-strongly convex if, for all x, y,

$$f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle + \frac{\nu}{2}\|x - y\|^2.$$

Once the encoding is done and appropriate data is stored in the nodes, the optimization process works in iterations. At iteration t, the master node broadcasts the current iterate

w_t to the worker nodes, and wait for k_t gradient updates $\nabla f_i(w)$ to arrive, corresponding to that iteration, and then chooses a step direction d_t and a step size α_t (based on algorithm Ψ_t that maps the set of gradients updates to a step) to update the parameters. We will denote $\eta_t = \frac{k_t}{m}$. We will also drop the time dependence of k and η whenever it is kept constant.

The set of fastest k_t nodes to send gradients for iteration t will be denoted as A_t. Once k_t updates have been collected, the remaining nodes, denoted A_t^c, are interrupted by the master node[1]. Algorithms 2 and 3 describe the generic mechanism of the proposed distributed optimization scheme at the master node and a generic worker node, respectively.

The intuition behind the encoding idea is that waiting for only $k_t < m$ workers prevents the stragglers from holding up the computation, while the redundancy provided by using a tall matrix S compensates for the information lost by proceeding without the updates from stragglers (the nodes in the subset A_t^c).

We next describe the three specific algorithms that we consider under data parallelism, to compute d_t.

Gradient descent. In this case, we assume that $h(w)$ is L-smooth. Then we simply set the descent direction

$$d_t = -\left(\frac{1}{2n\eta} \sum_{i \in A_t} \nabla f_i(w_t) + \lambda \nabla h(w_t) \right).$$

We keep $k_t = k$ constant, chosen based on the number of stragglers in the network, or based on the desired operating regime.

Limited-memory-BFGS. We assume that $h(w) = \|w\|^2$, and assume $\mu + \lambda > 0$. Although L-BFGS is traditionally a batch method, requiring updates from all nodes, its stochastic variants have also been proposed by [MR15, BNT16]. The key modification to ensure

[1] If the communication is already in progress at the time when k_t faster gradient updates arrive, the communication can be finished without interruption, and the late update can be dropped upon arrival. Otherwise, such interruption can be implemented by having the master node send an interrupt signal, and having one thread at each worker node keep listening for such a signal.

convergence in this case is that the Hessian estimate must be computed via gradient components that are common in two consecutive iterations, *i.e.*, from the nodes in $A_t \cap A_{t-1}$. We adapt this technique to our scenario. For $t > 0$, define $u_t := w_t - w_{t-1}$, and

$$r_t := \frac{m}{2n\,|A_t \cap A_{t-1}|} \sum_{i \in A_t \cap A_{t-1}} \left(\nabla f_i(w_t) - \nabla f_i(w_{t-1})\right).$$

Then once the gradient terms $\{\nabla f_i(w_t)\}_{i \in A_t}$ are collected, the descent direction is computed by $d_t = -B_t \widetilde{g}_t$, where $\widetilde{g}_t = \frac{1}{2\eta n} \sum_{i \in A_t} \nabla f_i(w_t)$, and B_t is the inverse Hessian estimate for iteration t, which is computed by

$$B_t^{(\ell+1)} = V_{j_{\ell,t}}^\top B_t^{(\ell)} V_{j_{\ell,t}} + \rho_{j_{\ell,t}} u_{j_{\ell,t}} u_{j_{\ell,t}}^\top, \quad \rho_j = \frac{1}{r_j^\top u_j}, \quad V_j = I - \rho_j r_j u_j^\top$$

with $j_{\ell,t} = t - \widetilde{\sigma} + \ell$, $B_t^{(0)} = \frac{r_t^\top r_t}{r_t^\top u_t} I$, and $B_t := B_t^{(\widetilde{\sigma})}$ with $\widetilde{\sigma} := \min\{t, \sigma\}$, where σ is the L-BFGS memory length. Once the descent direction d_t is computed, the step size is determined through exact line search[2]. To do this, each worker node computes $S_i X d_t$, and sends it to the master node. Once again, the master node only waits for the fastest k_t nodes, denoted by $D_t \subseteq [m]$ (where in general $D_t \neq A_t$), to compute the step size that minimizes the function along d_t, given by

$$\alpha_t = -\rho \frac{d_t^\top \widetilde{g}_t}{d_t^\top \widetilde{X}_D^\top \widetilde{X}_D d_t}, \tag{6.3}$$

where $\widetilde{X}_D = [S_i X]_{i \in D_t}$, and $0 < \rho < 1$ is a back-off factor of choice.

Proximal gradient. Here, we consider the general case of non-smooth $h(w) \geq 0, \lambda \geq 0$. The descent direction d_t is given by

$$d_t = \arg\min_w \widetilde{F}_t(w) - w_t,$$

where

$$\widetilde{F}_t(w) := \frac{1}{2\eta n} \sum_{i \in A_t} f_i(w_t) + \left\langle \frac{1}{2\eta n} \sum_{i \in A_t} \nabla f_i(w_t), w - w_t \right\rangle + \lambda h(w) + \frac{1}{2\alpha} \|w - w_t\|^2.$$

We keep the step size $\alpha_t = \alpha$ and $k_t = k$ constant.

[2] Note that exact line search is not more expensive than backtracking line search for a quadratic loss, since it only requires a single matrix-vector multiplication.

Algorithm 2 Generic encoded distributed optimization procedure under data parallelism, at the master node.

```
Given: $\Psi_t$, a sequence of functions that map gradients $\{\nabla f_i(w_t)\}_{i \in A_t}$ to a descent direction $d_t$
Initialize $w_0$, $\alpha_0$
for $t = 1, \dots, T$ do
    broadcast $w_t$ to all worker nodes
    wait to receive $k_t$ gradient updates $\{\nabla f_i(w_t)\}_{i \in A_t}$
    send interrupt signal the nodes in $A_t^c$
    compute the descent direction $d_t = \Psi_t \left( \{\nabla f_i \left( w_t \right)\}_{i \in A_t} \right)$
    determine step size $\alpha_t$
    take the step $w_{t+1} = w_t + \alpha_t d_t$
end for
```

Algorithm 3 Generic encoded distributed optimization procedure under data parallelism, at worker node i.

```
Given: $f_i(w) = \|S_i(Xw - y)\|^2$
for $t = 1, \dots, T$ do
    wait to receive $w_t$
    while not interrupted by master do
        compute $\nabla f_i(w_t)$
    end while
    if computation was interrupted then
        continue
    else
        send $\nabla f_i(w_t)$
    end if
end for
```

6.3.2 Model parallelism

Under the model parallelism paradigm, we focus on objectives of the form

$$\min g(w) := \min_w \phi\left(Xw\right) = \min_w \phi\left(\sum_{i=1}^m X_i w_i\right), \tag{6.4}$$

where the data matrix is partitioned as $X = [X_1 \ X_2 \ \cdots \ X_m]$, the parameter vector is partitioned as $w = \left[w_1^\top \ w_2^\top \ \cdots \ w_m^\top\right]^\top$, ϕ is convex, and $g(w)$ is L-smooth. Note that the data matrix X is partitioned horizontally, meaning that the dataset is split across features, instead of data samples (see Figure 6.3). Common machine learning models, such as any regression problem with generalized linear models, support vector machine, and many other convex problems fit within this model.

We encode the problem (6.4) by setting $w = S^\top v$, and solving the problem

$$\min_v \widetilde{g}(v) := \phi\left(XS^\top v\right) = \min_v \phi\left(\sum_{i=1}^m XS_i^\top v_i\right), \tag{6.5}$$

where $w \in \mathbb{R}^p$ and $S^\top = \left[S_1^\top \ S_2^\top \ \cdots \ S_m^\top\right] \in \mathbb{R}^{p\times\beta p}$ (see Figure 6.4). As a result, worker i stores the column-block $XS_i^\top$, as well as the iterate partition v_i. Note that we increase the dimensions of the parameter vector by multiplying the dataset X with a wide encoding matrix $S^\top$ from the right, and as a result we have redundant coordinates in the system. As in the case of data parallelism, such redundant coordinates provide robustness against erasures arising due to stragglers. Such increase in coordinates means that the problem is simply lifted onto a larger dimensional space, while preserving the original geometry of the problem. We will denote $u_{i,t} = XS_i^\top v_{i,t}$, where $v_{i,t}$ is the parameter iterates of worker i at iteration t. In order to compute updates to its parameters v_i, worker i needs the up-to-date value of $\widetilde{z}_i := \sum_{j\neq i} u_j$, which is provided by the master node at every iteration.

Let $\mathcal{S} = \arg\min_w g(w)$, and given w, let w^* be the projection of w onto $\mathcal{S}$. We will say that $g(w)$ satisfies ν-restricted-strong convexity ([LY13]) if

$$\langle \nabla g(w), w - w^* \rangle \geq \nu \|w - w^*\|^2$$

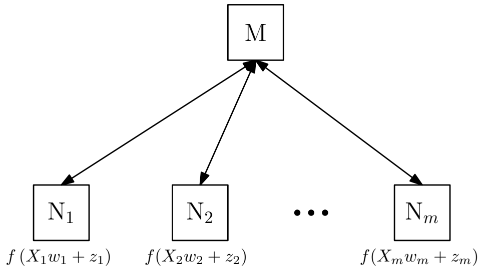


Figure 6.3: Uncoded distributed optimization with model parallelism, where ith node stores the ith partition of the model w_i. For $i = 1, \ldots, m$, $z_i = \sum_{j \neq i} X_j w_j$.

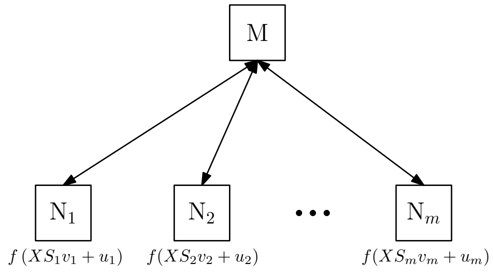


Figure 6.4: Encoded setup with model parallelism, where ith node stores the partition v_i of the model in the "lifted" space. For $i = 1, \ldots, m$, $\widetilde{z}_i = \sum_{j \neq i} u_j = \sum_{j \neq i} XS_j^\top v_j$.

for all w. Note that this is weaker than (implied by) strong convexity since w^* is restricted to be the projection of w, but unlike strong convexity, it is satisfied under the case where ϕ is strongly convex, but X has a non-trivial null space, *e.g.*, when it has more columns than rows.

For a given $w \in \mathbb{R}^p$, we define the level set of g at w as $D_g(w) := \{w' : g(w') \leq g(w)\}$. We will say that the level set at w_0 has diameter R if

$$\sup \left\{\|w - w'\| : w, w' \in D_g(w_0)\right\} \leq R.$$

As in the case of data parallelism, we assume that the master node waits for k updates at every iteration, and then moves onto the next iteration (see Algorithms 4 and 5). We similarly define A_t as the set of k fastest nodes in iteration t, and also define

$$I_{i,t} = \begin{cases} 1 & i \in A_t \\ 0 & i \notin A_t. \end{cases}$$

Under model parallelism, we consider block coordinate descent, described in Algorithm 4, where worker i stores the current values of the partition v_i, and performs updates on it, given the latest values of the rest of the parameters. The parameter estimate at time t is denoted

Algorithm 4 Encoded block coordinate descent at worker node i.

```
Given: $X_i$, $v_i$.
for $t = 1, \ldots, T$ do
    wait to receive $(I_{i,t-1}, \widetilde{z}_{i,t})$
    if $I_{i,t} == 1$ then
        take step $v_{i,t} = v_{i,t-1} + d_{i,t-1}$
    else
        set $v_{i,t} = v_{i,t-1}$
    end if
    while not interrupted by master do
        compute next step $d_{i,t} = \alpha S_i X^\top \nabla\phi\left(X S_i^\top v_{i,t} + \widetilde{z}_{i,t}\right)$
        compute $u_{i,t} = X S_i^\top v_{i,t}$
    end while
    if computation was interrupted then
        continue
    else
        send $u_{i,t}$ to master node
    end if
end for
```

Algorithm 5 Encoded block coordinate descent at the master node.

```
for $t = 1, \ldots, T$ do
    for $i = 1, \ldots, m$ do
        send $(I_{i,t-1}, \widetilde{z}_{i,t})$ to worker $i$
    end for
    wait to receive $k$ updated parameters $\{u_{i,t}\}_{i \in A_t}$
    send interrupt signal the nodes in $A_t^c$
    set $u_{i,t} = u_{i,t-1}$ for $i \in A_t^c$
    compute $\widetilde{z}_{i,t} = \sum_{j \neq i} u_{j,t}$ for all $i$
end for
```

by $v_{i,t}$, and we also define $\widetilde{z}_{i,t} = \sum_{j \neq i} u_{i,t} = \sum_{j \neq i} X S_j^\top v_j$. The iterates are updated by

$$v_{i,t} - v_{i,t-1} = \Delta_{i,t} := \begin{cases} -\alpha \nabla_i \widetilde{g}(v_{t-1}), & \text{if } i \in A_t \\ 0, & \text{otherwise,} \end{cases}$$

for a step size parameter $\alpha > 0$, where ∇_i refers to gradient only with respect to the variables v_i, *i.e.*, $\nabla \widetilde{g} = [\nabla_i \widetilde{g}]_{i \in [m]}$. Note that if $i \notin A_t$ then v_i does not get updated in worker i, which ensures the consistency of parameter values across machines. This is achieved by lines 4–8 in Algorithm 4. Worker i learns about this in the next iteration, when $I_{i,t-1}$ is sent by the master node.

6.4 Convergence Analysis

In this section, we prove convergence results for the algorithms described in Section 6.3. Note that since we modify the original optimization problem and solve it obliviously to this change, it is not obvious that the solution has any optimality guarantees with respect to the original problem. We show that, it is indeed possible to provide convergence guarantees in terms of the *original* objective under the encoded setup.

6.4.1 A spectral condition

In order to show convergence under the proposed framework, we require the encoding matrix S to satisfy a certain spectral criterion on S. Let S_A denote the submatrix of S associated with the subset of machines A, *i.e.*, $S_A = [S_i]_{i \in A}$. Then the criterion in essence requires that for any sufficiently large subset A, S_A behaves approximately like a matrix with orthogonal columns. We make this precise in the following statement.

Definition 6.1. *Let $\beta \geq 1$, and $\frac{1}{\beta} \leq \eta \leq 1$ be given. A matrix $S \in \mathbb{R}^{\beta n \times n}$ is said to satisfy the (m, η, ϵ)-block-restricted isometry property ((m, η, ϵ)-BRIP) if for any $A \subseteq [m]$ with $|A| = \eta m$,*

$$(1 - \epsilon) I_n \preceq \frac{1}{\eta} S_A^\top S_A \preceq (1 + \epsilon) I_n. \tag{6.6}$$

Note that this is similar to the restricted isometry property used in compressed sensing ([CT05]), except that we do not require (6.6) to hold for every submatrix of S of size $\mathbb{R}^{\eta n \times n}$. Instead, (6.6) needs to hold only for the submatrices of the form $S_A = [S_i]_{i \in A}$, which is a less restrictive condition. In general, it is known to be difficult to analytically prove that a structured, deterministic matrix satisfies the general RIP condition. Such difficulty extends to the BRIP condition as well. However, it is known that i.i.d. sub-Gaussian ensembles and randomized Fourier ensembles satisfy this property ([CT06]). In addition, numerical evidence suggests that there are several families of constructions for S whose submatrices have eigenvalues that mostly tend to concentrate around 1. We point out that although the strict BRIP condition is required for the theoretical analysis, in practice the algorithms perform well as long as the bulk of the eigenvalues of S_A lie within a small interval $(1-\epsilon, 1+\epsilon)$, even though the extreme eigenvalues may lie outside of it (in the non-adversarial setting). In Section 6.5, we explore several classes of matrices and discuss their relation to this condition.

6.4.2 Convergence of encoded gradient descent

We first consider the algorithms described under data parallelism architecture. The following theorem summarizes our results on the convergence of gradient descent for the encoded problem.

Theorem 6.1. *Let w_t be computed using encoded gradient descent with an encoding matrix that satisfies (m, η, ϵ)-BRIP, with step size $\alpha_t = \frac{2\zeta}{M(1+\epsilon)+L}$ for some $0 < \zeta \leq 1$, for all t. Let $\{A_t\}$ be an arbitrary sequence of subsets of $[m]$ with cardinality $|A_t| \geq \eta m$ for all t. Then, for f as given in* (6.1),

1.
$$\frac{1}{t}\sum_{\tau=1}^{t} f(w_\tau) - \kappa_1 f(w^*) \leq \frac{4\epsilon f(w_0) + \frac{1}{2\alpha}\|w_0 - w^*\|^2}{(1-7\epsilon)\, t}$$

2. *If f is in addition ν-strongly convex, then*
$$f(w_t) - \frac{\kappa_2^2(\kappa_2 - \gamma)}{1 - \kappa_2 \gamma} f\left(w^*\right) \leq \left(\kappa_2 \gamma\right)^t f(w_0), \quad t = 1, 2, \ldots,$$

where $\kappa_1 = \frac{1+3\epsilon}{1-7\epsilon}$, $\kappa_2 = \frac{1+\epsilon}{1-\epsilon}$, *and* $\gamma = \left(1 - \frac{4\nu\zeta(1-\zeta)}{M(1+\epsilon)+L}\right)$, *where* ϵ *is assumed to be small enough so that* $\kappa_2\gamma < 1$.

The proof is provided in Appendix E.1, which relies on the fact that the solution to the effective "instantaneous" problem corresponding to the subset A_t lies in a bounded set $\{w : f(w) \leq \kappa f(w^*)\}$ (where κ depends on the encoding matrix and strong convexity assumption on f), and therefore each gradient descent step attracts the iterate towards a point in this set, which must eventually converge to this set. Theorem 6.1 shows that encoded gradient descent can achieve the standard $O\left(\frac{1}{t}\right)$ convergence rate for the general case, and linear convergence rate for the strongly convex case, up to an approximate minimum. For the convex case, the convergence is shown on the running mean of past function values, whereas for the strongly convex case we can bound the function value at every step. Note that although the nodes actually minimize the encoded objective $\widetilde{f}(w)$, the convergence guarantees are given in terms of the original objective $f(w)$.

Theorem 6.1 provides deterministic, sample path convergence guarantees under any (adversarial) sequence of active sets $\{A_t\}$, which is in contrast to the stochastic methods, which show convergence typically in expectation. Further, the convergence rate is not affected by the tail behavior of the delay distribution, since the delayed updates of stragglers are not applied to the iterates.

Note that since we do not seek exact solutions under data parallelism, we can keep the redundancy factor β fixed regardless of the number of stragglers. Increasing number of stragglers in the network simply results in a looser approximation of the solution, allowing for a graceful degradation. This is in contrast to existing work [TLD17] seeking exact convergence under coding, which shows that the redundancy factor must grow linearly with the number of stragglers.

6.4.3 Convergence of encoded L-BFGS

We consider the variant of L-BFGS described in Section 6.3. For our convergence result for L-BFGS, we need another assumption on the matrix S, in addition to (6.6). Defining $\breve{S}_t = [S_i]_{i \in A_t \cap A_{t-1}}$ for $t > 0$, we assume that for some $\delta > 0$,

$$\delta I \preceq \breve{S}_t^\top \breve{S}_t \tag{6.7}$$

for all $t > 0$. Note that this requires that one should wait for sufficiently many nodes to send updates so that the overlap set $A_t \cap A_{t_1}$ has more than $\frac{1}{\beta}$ nodes, and thus the matrix $\breve{S}_t$ can be full rank. When the columns of X are linearly independent, this is satisfied if $\eta \geq \frac{1}{2} + \frac{1}{2\beta}$ in the worst-case, and in the case where node delays are i.i.d. across machines, it is satisfied in expectation if $\eta \geq \frac{1}{\sqrt{\beta}}$. One can also choose k_t adaptively so that $k_t = \min\left\{k : |A_t(k) \cap A_{t-1}| > \frac{1}{\beta}\right\}$. We note that although this condition is required for the theoretical analysis, the algorithm may perform well in practice even when this condition is not satisfied.

We first show that this algorithm results in stable inverse Hessian estimates under the proposed model, under arbitrary realizations of $\{A_t\}$ (of sufficiently large cardinality), which is done in the following lemma.

Lemma 6.1. *Let $\mu + \lambda > 0$. Then there exist constants $c_1, c_2 > 0$ such that for all t, the inverse Hessian estimate B_t satisfies $c_1 I \preceq B_t \preceq c_2 I$.*

The proof, provided in Appendix E.1, is based on the well-known trace-determinant method. Using Lemma 6.1, we can show the following convergence result.

Theorem 6.2. *Let $\mu + \lambda > 0$, and let w_t be computed using the L-BFGS method described in Section 6.3, with an encoding matrix that satisfies (m, η, ϵ)-BRIP. Let $\{A_t\}, \{D_t\}$ be arbitrary sequences of subsets of $[m]$ with cardinality $|A_t|, |D_t| \geq \eta m$ for all t. Then, for f as described in Section 6.3,*

$$f(w_t) - \frac{\kappa^2(\kappa - \gamma)}{1 - \kappa\gamma} f(w^*) \leq (\kappa\gamma)^t f(w_0),$$

where $\kappa = \frac{1+\epsilon}{1-\epsilon}$, *and* $\gamma = \left(1 - \frac{4(\mu+\lambda)c_1c_2}{(M+\lambda)(1+\epsilon)(c_1+c_2)^2}\right)$, *where* c_1 *and* c_2 *are the constants in Lemma 6.1.*

The proof is given in Appendix E.1. Similar to Theorem 6.1, the proof is based on the observation that the solution of the effective problem at time t lies in a bounded set around the true solution w^*. As in gradient descent, coding enables linear convergence deterministically, unlike the stochastic and multi-batch variants of L-BFGS, *e.g.*, [MR15, BNT16].

6.4.4 Convergence of encoded proximal gradient

Next we consider the encoded proximal gradient algorithm, described in Section 6.3, for objectives with potentially non-smooth regularizers $h(w)$. The following theorem characterizes our convergence results under this setup.

Theorem 6.3. *Let* w_t *be computed using encoded proximal gradient with an encoding matrix that satisfies* (m, η, ϵ)*-BRIP, with step size* $\alpha_t = \alpha < \frac{1}{M}$*, and where* $\epsilon < \frac{1}{7}$*. Let* $\{A_t\}$ *be an arbitrary sequence of subsets of* $[m]$ *with cardinality* $|A_t| \geq \eta m$ *for all* t*. Then, for* f *as described in Section 6.3,*

1. *For all* t*,*

$$\frac{1}{t}\sum_{\tau=1}^{t} f(w_\tau) - \kappa f(w^*) \leq \frac{4\epsilon f(w_0) + \frac{1}{2\alpha}\|w_0 - w^*\|^2}{(1-7\epsilon)\,t},$$

2. *For all* t*,*

$$f(w_{t+1}) \leq \kappa f(w_t),$$

where $\kappa = \frac{1+7\epsilon}{1-3\epsilon}$*.*

The proof is given in Appendix E.2. As in the previous algorithms, the convergence guarantees hold for arbitrary sequences of active sets $\{A_t\}$. Note that as in the gradient descent case, the convergence is shown on the mean of past function values. Since this does not prevent the iterates from having a sudden jump at a given iterate, we include the second part of the theorem to complement the main convergence result, which implies that the function value cannot increase by more than a small factor of its current value.

6.4.5 Convergence of encoded block coordinate descent

Finally, we consider the convergence of encoded block coordinate descent algorithm. The following theorem characterizes our main convergence result for this case.

Theorem 6.4. *Let $w_t = S^\top v_t$, where v_t is computed using encoded block coordinate descent as described in Section 6.3. Let S satisfy (m, η, ϵ)-BRIP, and the step size satisfy $\alpha < \frac{1}{L(1+\epsilon)}$. Let $\{A_t\}$ be an arbitrary sequence of subsets of $[m]$ with cardinality $|A_t| \geq \eta m$ for all t. Let the level set of g at the first iterate $D_g(w_0)$ have diameter R. Then, for $g(w) = \phi(Xw)$ as described in Section 6.3, the following hold.*

1. *If ϕ is convex, then*

$$g(w_t) - g(w^*) \leq \frac{1}{\frac{1}{\pi_0} + Ct},$$

where $\pi_0 = g(w_0) - g(w^*)$, and $C = \frac{(1-\epsilon)\alpha}{R}\left(1 - \frac{\alpha L'}{2}\right)$.

2. *If g is ν-restricted-strongly convex, then*

$$g(w_t) - g(w^*) \leq \left(1 - \frac{1}{\xi}\right)^t \left(g(w_0) - g(w^*)\right),$$

where $\xi = \frac{1}{\nu(1-\epsilon)\alpha}\left(1 - \frac{L(1+\epsilon)\alpha}{2}\right)^{-1}$.

The proof is given in Appendix E.3. Theorem 6.4 demonstrates that the standard $O\left(\frac{1}{t}\right)$ rate for the general convex, and linear rate for the strongly convex case can be obtained under the encoded setup. Note that unlike the data parallelism setup, we can achieve exact minimum under model parallelism, since the underlying geometry of the problem does not change under encoding; the same objective is simply mapped onto a higher-dimensional space, which has redundant coordinates. Similar to the previous cases, encoding allows for deterministic convergence guarantees under adversarial failure patterns. This comes at the expense of a small penalty in the convergence rate though; one can observe that a non-zero ϵ slightly weakens the constants in the convergence expressions. Still, note that this penalty in convergence rate only depends on the encoding matrix and not on the delay profile in the system. This is in contrast to the asynchronous coordinate descent methods; for instance, in [LWR15], the step size is required to shrink *exponentially* in the maximum allowable delay, and thus the guaranteed convergence rate can exponentially degrade with increasing worst-case delay in the system. The same is true for the linear convergence guarantee in [PXY16].

6.5 Code Design

6.5.1 Block RIP condition and code design

We first discuss two classes of encoding matrices with regard to the BRIP condition; namely equiangular tight frames, and random matrices.

Tight frames. A unit-norm *frame* for $\mathbb{R}^n$ is a set of vectors $F = \{a_i\}_{i=1}^{n\beta}$ with $\|a_i\| = 1$, where $\beta \geq 1$, such that there exist constants $\xi_2 \geq \xi_1 > 0$ such that, for any $u \in \mathbb{R}^n$,

$$\xi_1 \|u\|^2 \leq \sum_{i=1}^{n\beta} |\langle u, a_i \rangle|^2 \leq \xi_2 \|u\|^2.$$

The frame is *tight* if the above satisfied with $\xi_1 = \xi_2$. In this case, it can be shown that the constants are equal to the redundancy factor of the frame, *i.e.*, $\xi_1 = \xi_2 = \beta$. If we form $S \in \mathbb{R}^{(\beta n) \times n}$ by rows that form a *tight frame*, then we have $S^\top S = \beta I$, which ensures $\|Xw - y\|^2 = \frac{1}{\beta}\|SXw - Sy\|^2$. Then for any solution $\hat{w}$ to the encoded problem (with $k = m$),

$$\nabla \widetilde{f}(\hat{w}) = X^\top S^\top S(X\hat{w} - y) = \beta X^\top (X\hat{w} - y) = \beta \nabla f(\hat{w}).$$

Therefore, the solution to the encoded problem satisfies the optimality condition for the original problem as well:

$$-\nabla \widetilde{f}(\hat{w}) \in \partial h(\hat{w}), \quad \Leftrightarrow \quad -\nabla f(\hat{w}) \in \partial h(\hat{w}),$$

and if f is also strongly convex, then $\hat{w} = w^*$ is the unique solution. This means that for $k = m$, obliviously solving the encoded problem results in the same objective value as in the original problem.

Define the maximal inner product of a unit-norm tight frame $F = \{a_i\}_{i=1}^{n\beta}$, where $a_i \in \mathbb{R}^n, \forall i$, by

$$\omega(F) := \max_{\substack{a_i, a_j \in F \\ i \neq j}} |\langle a_i, a_j \rangle| \,.$$

A tight frame is called an *equiangular tight frame* (ETF) if $|\langle a_i, a_j \rangle| = \omega(F)$ for every $i \neq j$.

Proposition 6.1 ([Wel74]). *Let $F = \{a_i\}_{i=1}^{n\beta}$ be a tight frame. Then $\omega(F) \geq \sqrt{\frac{\beta-1}{n\beta-1}}$. Moreover, equality is satisfied if and only if F is an equiangular tight frame.*

Therefore, an ETF minimizes the correlation between its individual elements, making each submatrix $S_A^\top S_A$ as close to orthogonal as possible. This, combined with the property that tight frames preserve the optimality condition when all nodes are waited for ($k = m$), make ETFs good candidates for encoding, in light of the required property (6.6). We specifically evaluate the Paley ETF from [Pal33] and [GS67]; Hadamard ETF from [Szo13] (not to be confused with Hadamard matrix); and Steiner ETF from [FM12] in our experiments.

Although the derivation of tight eigenvalue bounds for subsampled ETFs is a long-standing problem, numerical evidence (see Figures 6.5, 6.6) suggests that they tend to have their eigenvalues more tightly concentrated around 1 than random matrices (also supported by the fact that they satisfy Welch bound, Proposition 6.1 with equality).

Note that our theoretical results focus on the extreme eigenvalues due to a worst-case analysis; in practice, most of the energy of the gradient lies on the eigen-space associated with the bulk of the eigenvalues, which the following proposition shows can be identically 1.

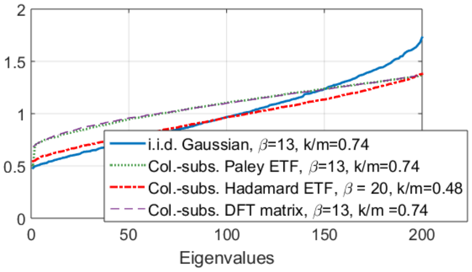


Figure 6.5: Sample spectrum of $S_A^\top S_A$ for various constructions with high redundancy, and small k (normalized).

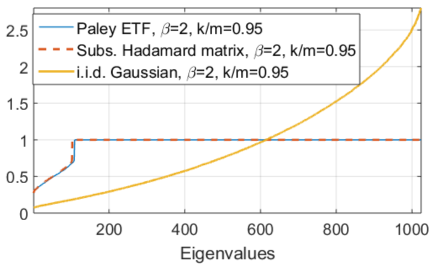


Figure 6.6: Sample spectrum of $S_A^\top S_A$ for various constructions with moderate redundancy, and large k (normalized).

Proposition 6.2. *If the rows of S are chosen to form an ETF with redundancy β, then for $\eta \geq 1 - \frac{1}{\beta}$, $\frac{1}{\beta} S_A^\top S_A$ has $n(1-\beta(1-\eta))$ eigenvalues equal to 1.*

This follows immediately from Cauchy interlacing theorem, using the fact that $S_A S_A^\top$ and $S_A^\top S_A$ have the same spectra except zeros. Therefore for sufficiently large η, ETFs have a mostly flat spectrum even for low redundancy, and thus in practice one would expect ETFs to perform well even for small amounts of redundancy. This is also confirmed by Figure 6.6, as well as our numerical results.

Random matrices. Another natural choice of encoding could be to use i.i.d. random matrices. Although encoding with such random matrices can be computationally expensive and may not have the desirable properties of encoding with tight frames, their eigenvalue behavior can be characterized analytically. In particular, using the existing results on the eigenvalue scaling of large i.i.d. Gaussian matrices from [Gem80, Sil85] and union bound, it can be shown that

$$\mathbb{P}\left(\max_{A:|A|=k} \lambda_{\max}\left(\frac{1}{\beta\eta n} S_A^\top S_A\right) > \left(1 + \sqrt{\frac{1}{\beta\eta}}\right)^2\right) \to 0 \tag{6.8}$$

$$\mathbb{P}\left(\min_{A:|A|=k} \lambda_{\min}\left(\frac{1}{\beta\eta n} S_A^\top S_A\right) < \left(1 - \sqrt{\frac{1}{\beta\eta}}\right)^2\right) \to 0, \tag{6.9}$$

as $n \to \infty$, if the elements of S_A are drawn i.i.d. from $N(0,1)$. Hence, for sufficiently large redundancy and problem dimension, i.i.d. random matrices are good candidates for encoding as well. However, for finite β, even if $k = m$, in general the optimum of the original problem is not recovered exactly, for such matrices.

6.5.2 Efficient encoding

In this section we discuss some of the possible practical approaches to encoding. Some of the practical issues involving encoding include the the computational complexity of encoding, as well as the loss of sparsity in the data due to the multiplication with S, and the resulting increase in time and space complexity. We address these issues in this section.

6.5.2.1 Efficient distributed encoding with sparse matrices

Let the dataset (X, y) lie in a database, accessible to each worker node, where each node is responsible for computing their own encoded partitions $S_i X$ and $S_i y$. We assume that S has a sparse structure. Given S, define $B_i(S) = \{j : S_{ij} \neq 0\}$ as the set of indices of the non-zero elements of the ith row of S. For a set $\mathcal{I}$ of rows, we define $B_{\mathcal{I}}(S) = \cup_{i \in \mathcal{I}} B_i(S)$.

Let us partition the set of rows of S, $[\beta n]$, into m machines, and denote the partition of machine k as $\mathcal{I}_k$, *i.e.*, $\bigsqcup_{k=1}^{m} \mathcal{I}_k = [\beta n]$, where $\sqcup$ denotes disjoint union. Then the set of non-zero columns of S_k is given by $B_{\mathcal{I}_k}(S)$. Note that in order to compute $S_k X$, machine k only requires the rows of X in the set $B_{\mathcal{I}_k}(S)$. In what follows, we will denote this submatrix of X by $\widetilde{X}_k$, *i.e.*, if $x_i^\top$ is the ith row of X, $\widetilde{X}_k := \left[x_i^\top\right]_{i \in B_{\mathcal{I}_k}(S)}$. Similarly $\widetilde{y}_k = [y_i]_{i \in B_{\mathcal{I}_k}(S)}$, where y_i is the ith element of y.

Consider the specific computation that needs to be done by worker k during the iterations, for each algorithm. Under the data parallelism setting, worker k computes the following gradient:

$$\nabla f_k(w) = X^\top S_k^\top S_k (Xw - y) \overset{\text{(a)}}{=} \widetilde{X}_k^\top S_k^\top S_k (\widetilde{X}_k w - \widetilde{y}_k) \tag{6.10}$$

where (a) follows since the rows of X that are not in $B_{\mathcal{I}_k}$ get multiplied by zero vector. Note that the last expression can be computed without any matrix-matrix multiplication.

This gives a natural storage and computation scheme for the workers. Instead of computing $S_k X$ offline and storing it, which can result in a loss of sparsity in the data, worker k can store $\widetilde{X}_k$ in uncoded form, and compute the gradient through (6.10) whenever needed, using only matrix-vector multiplications. Since S_k is sparse, the overhead associated with multiplications of the form $S_k v$ and $S_k^\top v$ is small.

Similarly, under model parallelism, the computation required by worker k is

$$\nabla_k \widetilde{g}(v) = S_k X^\top \nabla_k \phi \left(X S_k^\top v_k + \widetilde{z}_k \right) = S_k \widetilde{X}_k^\top \nabla_k \phi \left(\widetilde{X}_k S_k^\top v_k + \widetilde{z}_k \right), \tag{6.11}$$

and as in the data parallelism case, the worker can store $\widetilde{X}_k$ uncoded, and compute (6.11) online through matrix-vector multiplications.

Example: Steiner ETF. We illustrate the described technique through Steiner ETF, based on the construction proposed in [FM12], using $(2, 2, v)$-Steiner systems. Let v be a power of 2, let $H \in \mathbb{R}^{v \times v}$ be a real Hadamard matrix, and let h_i be the ith column of H, for $i = 1, \ldots, v$. Consider the matrix $V \in \{0,1\}^{v \times v(v-1)/2}$, where each column is the incidence vector of a distinct two-element subset of $\{1, \ldots, v\}$. For instance, for $v = 4$,

$$V = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}.$$

Note that each of the v rows have exactly $v - 1$ non-zero elements. We construct Steiner ETF S as a $v^2 \times \frac{v(v-1)}{2}$ matrix by replacing each 1 in a row with a distinct column of H, and normalizing by $\sqrt{v-1}$. For instance, for the above example, we have

$$S = \frac{1}{\sqrt{3}} \begin{bmatrix} h_2 & h_3 & h_4 & 0 & 0 & 0 \\ h_2 & 0 & 0 & h_3 & h_4 & 0 \\ 0 & h_2 & 0 & h_3 & 0 & h_4 \\ 0 & 0 & h_2 & 0 & h_3 & h_4 \end{bmatrix}.$$

We will call a set of rows of S that arises from the same row of V a block. In general, this procedure results in a matrix S with redundancy factor $\beta = \frac{2v}{v-1}$. In full generality, Steiner ETFs can be constructed for larger redundancy levels; we refer the reader to [FM12] for a full discussion of these constructions.

We partition the rows of the V matrix into m machines, so that each machine gets assigned $\frac{v}{m}$ rows of V, and thus the corresponding $\frac{v}{m}$ blocks of S.

This construction and partitioning scheme is particularly attractive for our purposes for two reasons. First, it is easy to see that for any node k, $|B_{\mathcal{I}_k}|$ is upper bounded by $\frac{v(v-1)}{m} = \frac{2n}{m}$, which means the memory overhead compared to the uncoded case is limited to a factor[3] of β. Second, each block of S_k consists of (almost) a Hadamard matrix, so the multiplication $S_k v$ can be efficiently implemented through Fast Walsh-Hadamard Transform.

Example: Haar matrix. Another possible choice of sparse matrix is column-subsampled Haar matrix, which is defined recursively by

$$H_{2n} = \frac{1}{\sqrt{2}} \begin{bmatrix} H_n \otimes [1\ 1] \\ I_n \otimes [1\ -1] \end{bmatrix}, \quad H_1 = 1,$$

where $\otimes$ denotes Kronecker product. Given a redundancy level β, one can obtain S by randomly sampling $\frac{n}{\beta}$ columns of H_n. It can be shown that in this case, we have $|B_{\mathcal{I}_k}| \leq \frac{\beta n \log(n)}{m}$, and hence encoding with Haar matrix incurs a memory cost by logarithmic factor.

6.5.2.2 Fast transforms

Another computationally efficient method for encoding is to use fast transforms: Fast Fourier Transform (FFT), if S is chosen as a subsampled DFT matrix, and the Fast Walsh-Hadamard Transform (FWHT), if S is chosen as a subsampled real Hadamard matrix. In particular, one can insert rows of zeroes at random locations into the data pair (X, y), and then take the

[3] In practice, we have observed that the convergence performance improves when the blocks are broken into multiple machines, so one can, for instance, assign half-blocks to each machine.

FFT or FWHT of each column of the augmented matrix. This is equivalent to a randomized Fourier or Hadamard ensemble, which is known to satisfy the RIP with high probability by [CT06]. However, such transforms do not have the memory advantages of the sparse matrices, and thus they are more useful for the setting where the dataset is dense, and the encoding is done offline.

6.5.3 Cost of encoding

Since encoding increases the problem dimensions, it clearly comes with the cost of increased space complexity. The memory and storage requirement of the optimization still increases by a factor of 2, if the encoding is done offline (for dense datasets), or if the techniques described in the previous subsection are applied (for sparse datasets)[4]. Note that the added redundancy can come by increasing the amount of effective data points per machine, by increasing the number of machines while keeping the load per machine constant, or a combination of the two. In the first case, the computational load per machine increases by a factor of β. Although this can make a difference if the system is bottlenecked by the computation time, distributed computing systems are typically communication-limited, and thus we do not expect this additional cost to dominate the speed-up from the mitigation of stragglers.

6.6 Numerical Results

We implement the proposed technique on four problems: ridge regression, matrix factorization, logistic regression, and LASSO.

[4] Note that the increase in space complexity is not higher for sparse matrices, since the sparsity loss can be avoided using the techniques described in Section 6.5.2

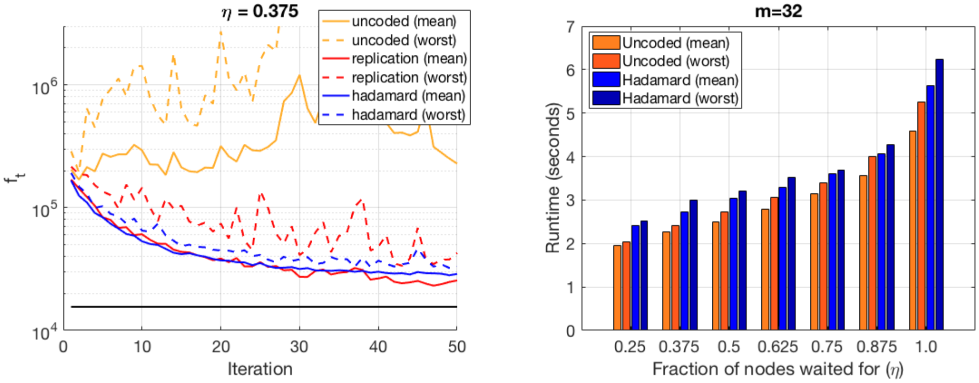


Figure 6.7: **Left:** Sample evolution of uncoded, replication, and Hadamard (FWHT)-coded cases, for $k = 12$, $m = 32$. **Right:** Runtimes of the schemes for different values of η, for the same number of iterations for each scheme. Note that this essentially captures the delay profile of the network, and does not reflect the relative convergence rates of different methods.

6.6.1 Ridge regression

We generate the elements of matrix X i.i.d. $\sim N(0,1)$, and the elements of y are generated from X and an i.i.d. $N(0,1)$ parameter vector w^*, through a linear model with Gaussian noise, for dimensions $(n,p) = (4096, 6000)$. We solve the problem $\min_w \frac{1}{2n} \|S(Xw - y)\|^2 + \frac{\lambda}{2}\|w\|^2$, for regularization parameter $\lambda = 0.05$. We evaluate column-subsampled Hadamard matrix with redundancy $\beta = 2$ (encoded using FWHT), replication and uncoded schemes. We implement distributed L-BFGS as described in Section 6.4 on an Amazon EC2 cluster using `mpi4py` Python package, over $m = 32$ `m1.small` instances as worker nodes, and a single `c3.8xlarge` instance as the central server.

Figure 6.7 shows the result of our experiments, which are aggregated from 20 trials. In addition to uncoded scheme, we consider data replication, where each uncoded partition is replicated $\beta = 2$ times across nodes, and the server discards the duplicate copies of a partition, if received in an iteration. It can be seen that for low η, uncoded L-BFGS may not converge when a fixed number of nodes are waited for, whereas the Hadamard-coded case stably converges. We also observe that the data replication scheme converges on average,

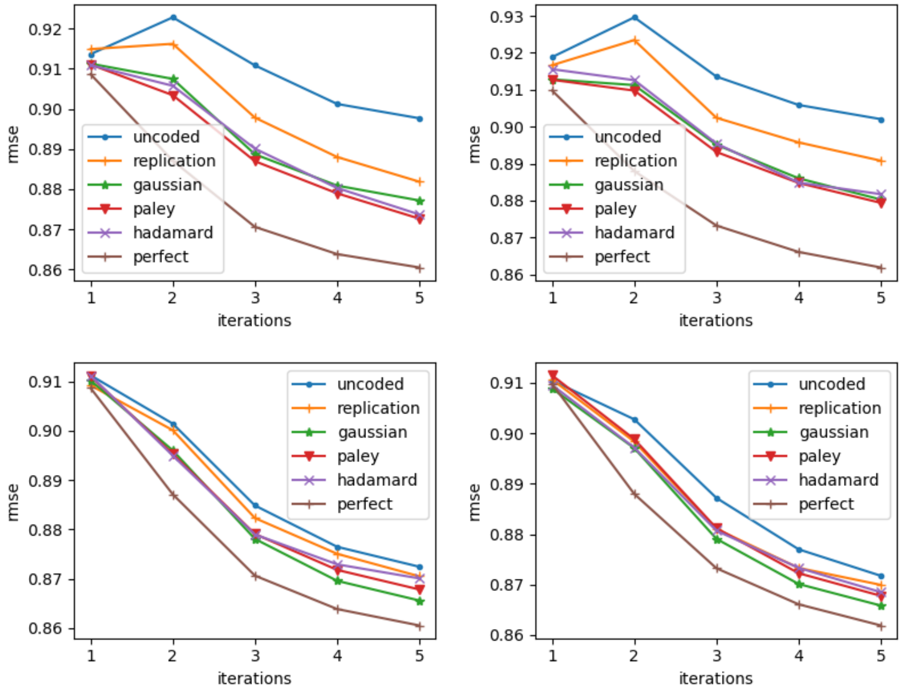


Figure 6.8: Test RMSE for $m = 8$ (left) and $m = 24$ (right) nodes, where the server waits for $k = m/8$ (top) and $k = m/2$ (bottom) responses. "Perfect" refers to the case where $k = m$.

but its performance may deteriorate if both copies of a partition are delayed. Figure 6.7 suggests that this performance can be achieved with an approximately 40% reduction in the runtime, compared to waiting for all the nodes.

6.6.2 Matrix factorization

We next apply matrix factorization on the MovieLens-1M dataset ([RK98]) for the movie recommendation task. We are given R, a sparse matrix of movie ratings 1–5, of dimension $\#users \times \#movies$, where R_{ij} is specified if user i has rated movie j. We withhold randomly 20% of these ratings to form an 80/20 train/test split. The goal is to recover user vectors $x_i \in \mathbb{R}^p$ and movie vectors $y_i \in \mathbb{R}^p$ (where p is the embedding dimension) such that $R_{ij} \approx x_i^T y_j + u_i + v_j + b$, where u_i, v_j, and b are user, movie, and global biases, respectively. The

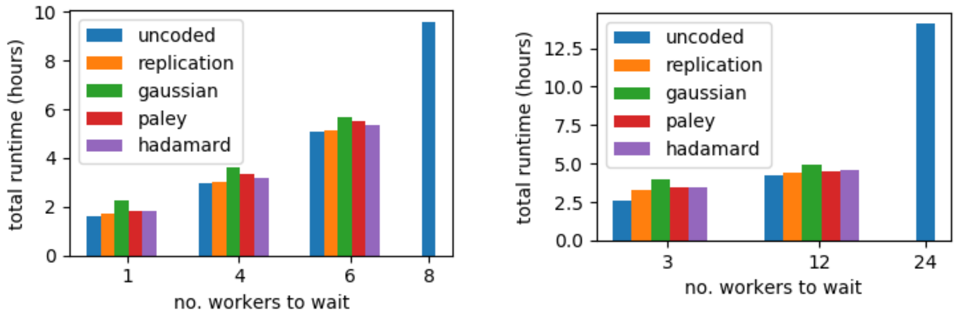


Figure 6.9: Total runtime with $m = 8$ and $m = 24$ nodes for different values of k, under fixed 100 iterations for each scheme.

optimization problem is given by

$$\min_{x_i,y_j,u_i,v_j} \sum_{i,j:\text{ observed}} (R_{ij} - u_i - v_j - x_i^T y_j - b)^2 + \lambda \left(\sum_i \|x_i\|_2^2 + \|u\|_2^2 + \sum_j \|y_j\|_2^2 + \|v\|_2^2 \right). \tag{6.12}$$

We choose $b = 3$, $p = 15$, and $\lambda = 10$, which achieves test RMSE 0.861, close to the current best test RMSE on this dataset using matrix factorization[5].

Problem (6.12) is often solved using alternating minimization, minimizing first over all (x_i, u_i), and then all (y_j, v_j), in repetition. Each such step further decomposes by row and column, made smaller by the sparsity of R. To solve for (x_i, u_i), we first extract $I_i = \{j \mid r_{ij} \text{ is observed}\}$, and minimize

$$\left(\left[y_{I_i}^T, \mathbf{1} \right] \begin{bmatrix} x_i \\ u_i \end{bmatrix} - (R_{i,I_i}^T - v_{I_i} - b\mathbf{1}) \right)^2 + \lambda \left(\sum_i \|x_i\|_2^2 + \|u\|_2^2 \right) \tag{6.13}$$

for each i, which gives a sequence of regularized least squares problems with variable $w = [x_i^T, u_i]^T$, which we solve distributedly using coded L-BFGS; and repeat for $w = [y_j^T, v_j]^T$, for all j.

[5] http://www.mymedialite.net/examples/datasets.html

The Movielens experiment is run on a single 32-core machine with Linux 4.4. In order to simulate network latency, an artificial delay of $\Delta \sim \exp(10 \text{ ms})$ is imposed each time the worker completes a task. Small problem instances ($n < 500$) are solved locally at the central server, using the built-in function `numpy.linalg.solve`. To reduce overhead, we create a bank of encoding matrices $\{S_n\}$ for Paley ETF and Hadamard ETF, for $n = 100, 200, \ldots, 3500$, and then given a problem instance, subsample the columns of the appropriate matrix S_n to match the dimensions. Overall, we observe that encoding overhead is amortized by the speed-up of the distributed optimization.

Figure 6.8 gives the final performance of our distributed L-BFGS for various encoding schemes, for each of the 5 epochs, which shows that coded schemes are most robust for small k. A full table of results is given in Appendix E.4.

6.6.3 Logistic regression

In our next experiment, we apply logistic regression for document classification for Reuters Corpus Volume 1 (`rcv1.binary`) dataset from [LYR04], where we consider the binary task of classifying the documents into corporate/industrial/economics vs. government/social/markets topics. The dataset has 697,641 documents, and 47,250 term frequency-inverse document frequency (tf-idf) features. We randomly select 32,500 features for the experiment, and reserve 100,000 documents for the test set. We use logistic regression with ℓ_2-regularization for the classification task, with the objective

$$\min_{w,b} \frac{1}{n} \sum_{i=1}^{n} \log\left(1 + \exp\left\{-z_i^\top w + b\right\}\right) + \lambda \|w\|^2,$$

where $z_i = y_i x_i$ is the data sample x_i multiplied by the label $y_i \in \{-1, 1\}$, and b is the bias variable. We solve this optimization using encoded distributed block coordinate descent as described in Section 6.3, and implement Steiner and Haar encoding as described in Section 6.5, with redundancy $\beta = 2$. In addition we implement the asynchronous coordinate descent, as well as replication, which represents the case where each partition Z_i is replicated across two nodes, and the faster copy is used in each iteration. We use $m = 128$ `t2.medium`

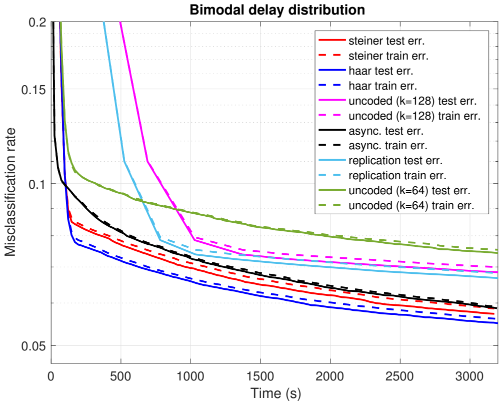


Figure 6.10: Test and train errors over time (in seconds) for each scheme, for the bimodal delay distribution. Steiner and Haar encoding is done with $k = 64$, $\beta = 2$.

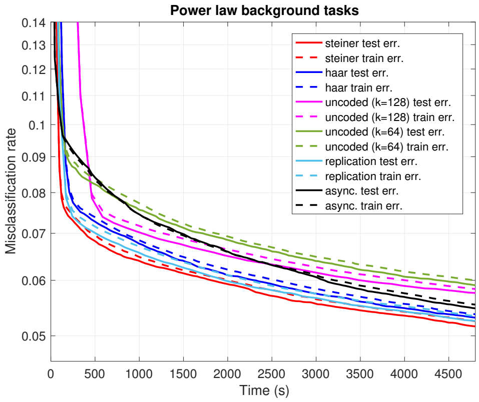


Figure 6.11: Test and train errors over time (in seconds) for each scheme. Number of background tasks follow a power law. Steiner and Haar encoding is done with $k = 80$, $\beta = 2$.

instances as worker nodes, and a single `c3.4xlarge` instance as the master node, which communicate using the `mpi4py` package. We consider two models for stragglers. In the first model, at each node, we add a random delay drawn from a Gaussian mixture distribution $q\mathcal{N}(\mu_1, \sigma_1^2) + (1-q)\mathcal{N}(\mu_2, \sigma_2^2)$, where $q = 0.5$, $\mu_1 = 0.5$s, $\mu_2 = 20$s, $\sigma_1 = 0.2$s, $\sigma_2 = 5$s. In the second model, we do not directly add any delay, but at each machine we launch a number of dummy background tasks (matrix multiplication) that are executed throughout the computation. The number of background tasks across the nodes is distributed according to a power law with exponent $\alpha = 1.5$. The number of background tasks launched is capped at 50.

Figures 6.10 and 6.11 shows the evolution of training and test errors as a function of wall clock time. We observe that for each straggler model, either Steiner or Haar encoded optimization dominates all schemes. Figures 6.12 and 6.13 show the statistics of how frequent each node participates in an update, for the case with background tasks, for encoded and asynchronous cases, respectively. We observe that the stark difference in the relative speeds of different machines result in vastly different update frequencies for the asynchronous case,

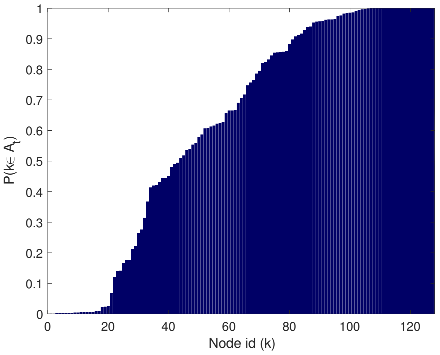


Figure 6.12: The fraction of iterations each worker node participates in (the empirical probability of the event $\{k \in A_t\}$), plotted for Steiner encoding with $k = 80$, $m = 128$. The number of background tasks are distributed by a power law with $\alpha = 1.5$ (capped at 50).

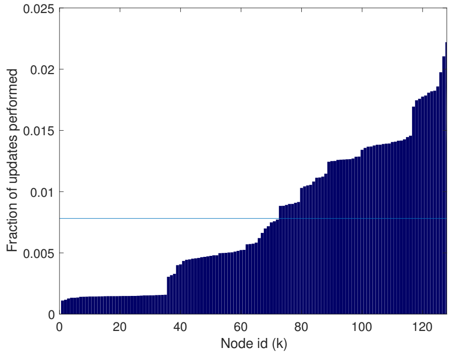


Figure 6.13: The fraction of updates performed by each node, for asynchronous block coordinate descent. The horizontal line represents the uniformly distributed case. The number of background tasks are distributed by a power law with $\alpha = 1.5$ (capped at 50).

which results in updates with large delays, and a corresponding performance loss.

6.6.4 LASSO

We solve the LASSO problem, with the objective

$$\min_{w} \frac{1}{2n}\|Xw - y\|^2 + \lambda\|w\|_1^2,$$

where $X \in \mathbb{R}^{130{,}000 \times 100{,}000}$ is a matrix with i.i.d. $N(0, 1)$ entries, and y is generated from X and a parameter vector w^* through a linear model with Gaussian noise:

$$y = Xw^* + \sigma z,$$

where $\sigma = 40$, $z \sim N(0, 1)$. The parameter vector w^* has 7695 non-zero entries out of 100,000, where the non-zero entries are generated i.i.d. from $N(0, 4)$. We choose $\lambda = 0.6$ and consider the sparsity recovery performance of the corresponding LASSO problem, solved using proximal gradient (iterative shrinkage/thresholding algorithm).

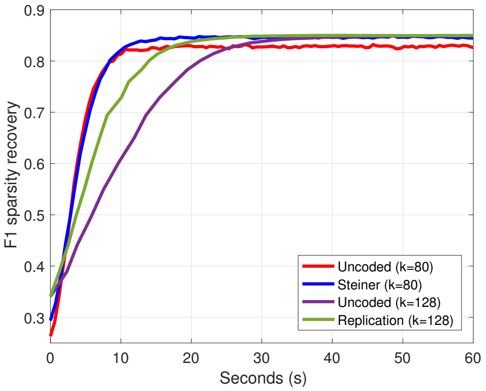


Figure 6.14: Evolution of F1 sparsity recovery performance for each scheme.

We implement the algorithm over 128 `t2.medium` worker nodes which collectively store the matrix X, and a `c3.4xlarge` master node. We measure the sparsity recovery performance of the solution using the F1 score, defined as the harmonic mean

$$F1 = \frac{2PR}{P+R},$$

where P and R are precision recall of the solution vector $\hat{w}$ respectively, defined as

$$P = \frac{|\{i : w_i^* \neq 0, \hat{w}_i \neq 0\}|}{|i : \hat{w}_i \neq 0|}, \quad R = \frac{|\{i : w_i^* \neq 0, \hat{w}_i \neq 0\}|}{|i : w_i^* \neq 0|}$$

.

Figure 6.14 shows the sample evolution of the F1 score of the model under uncoded, replication, and Steiner encoded scenarios, with artificial multi-modal communication delay distribution $q_1\mathcal{N}(\mu_1, \sigma_1^2) + q_2\mathcal{N}(\mu_2, \sigma_2^2) + q_3\mathcal{N}(\mu_3, \sigma_3^2)$, where $q_1 = 0.8$, $q_2 = 0.1$, $q_3 = 0.1$; $\mu_1 = 0.2$s, $\mu_2 = 0.6$s, $\mu_3 = 1$s; and $\sigma_1 = 0.1$s, $\sigma = 0.2$s, $\sigma_3 = 0.4$s, independently at each node. We observe that the uncoded case $k = 80$ results in a performance loss in sparsity recovery due to data dropped from delayed noes, and uncoded and replication with $k = 128$ converges slow due to stragglers, while Steiner coding with $k = 80$ is not delayed by stragglers, while maintaining almost the same sparsity recovery performance as the solution of the uncoded $k = 128$ case.

CHAPTER 7

Conclusions and Open Problems

We studied techniques to achieve robust communication and optimization over networks with unreliable and intermittently available resources. In the first part, we have focused on communication over wireless networks, and developed and analyzed schemes that make use of intermittently available links for feedback and cooperation. We have also considered the problem of uplink-downlink interference in full-duplex cellular networks, and proposed a simple scheduling scheme for networks with time-varying link strengths, to mitigate this interference. In the second part, we considered distributed optimization over networks with nodes that fail or delay their computation, and developed the encoded distributed optimization framework to counteract the harmful effect of such nodes on the optimization. Along all directions that are explored in this thesis, there are several open questions that are interesting to study in the future.

Specifically, an interesting research direction could be to gain a more comprehensive understanding of how unlicensed bands can be harnessed for communication over licensed bands, in upcoming 5G networks. This could be a particularly important question since 5G is envisioned to be a network of billions of devices, each of which have different requirements, computational capabilities, power limitations, hardware constraints, and spectral limitations, over the same infrastructure. Opportunistic use of such intermittent links to support networking over 5G could be one of the ways to get closer to this goal.

The encoded optimization framework we developed in this thesis assumes convex problems. Given the recent rise of deep learning, an important question is to understand how to extend these ideas to non-convex problems, and specifically to the training of deep neural networks. Especially under the emerging federated learning scenario, where the user devices which act as worker nodes can be highly unreliable, the problem of straggler mitigation becomes critical to the training performance and quality of the trained model. Therefore an exciting research question is whether we can extend the encoding idea to such a scenario.

Another worthwhile direction is to understand if the encoding idea can be used for privacy purposes, where a central server receives a version of the data that is jumbled through encoding. Ideally, such encoding would reveal little about the data itself, but its output would still be useful for the learning objective.

There are many exciting and interesting future research directions in understanding how to design robust communication and computation systems over unreliable networks. We hope that this thesis contributes to the development of the foundational ideas to build reliable future systems.

APPENDIX A

Proofs for Chapter 2

A.1 Proof of Lemma 2.1

Choose $\epsilon > 0$. We suppress the dependence of variables on block index b for simplicity. Consider, for $(i,j) = (1,2), (2,1)$,

$$\begin{aligned}
\mathbb{E}\left[2^{NK_i}\right] &= \mathbb{E}\left[\sum_{q_i=1}^{2^{Nr_i}} \mathbb{1}_{\left\{q_i:\left(X_{jf}^N, U_i^N(q_i)\right)\in\mathcal{A}_\epsilon^{(n)}\right\}}\right] \\
&= \sum_{q_i=1}^{2^{Nr_i}} \mathbb{P}\left(\left(X_{jf}^N, U_i^N(q_i)\right)\in\mathcal{A}_\epsilon^{(n)}\right) \\
&= 2^{Nr_i}\mathbb{P}\left(\left(X_{jf}^N, U_i^N(1)\right)\in\mathcal{A}_\epsilon^{(n)}\right)
\end{aligned}$$

Since $U_i^N(1)$ is generated independently from X_{jf}^N, by packing lemma [GK11], there exists $\delta(\epsilon)$ with $\delta(\epsilon)\to 0$ such that

$$\mathbb{E}\left[2^{NK_i}\right] \le 2^{Nr_i}2^{-N\left[I(X_{jf};U_i)-\delta(\epsilon)/3\right]}$$

for all N.

Next consider the variance of 2^{NK_i}.

$$\begin{aligned}
var(2^{NK_i}) &= var\left(\sum_{q_i=1}^{2^{Nr_i}} \mathbb{1}_{\left\{q_i:\left(X_{jf}^N, U_i^N(q_i)\right)\in\mathcal{A}_\epsilon^{(n)}\right\}}\right) \\
&\overset{\text{(a)}}{=} \sum_{q_i=1}^{2^{Nr_i}} var\left(\mathbb{1}_{\left\{q_i:\left(X_{jf}^N, U_i^N(q_i)\right)\in\mathcal{A}_\epsilon^{(n)}\right\}}\right) \\
&= \sum_{q_i=1}^{2^{Nr_i}} \mathbb{E}\left[\mathbb{1}_{\left\{q_i:\left(X_{jf}^N, U_i^N(q_i)\right)\in\mathcal{A}_\epsilon^{(n)}\right\}}\right]\cdot\left(1-\mathbb{E}\left[\mathbb{1}_{\left\{q_i:\left(X_{jf}^N, U_i^N(q_i)\right)\in\mathcal{A}_\epsilon^{(n)}\right\}}\right]\right)
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{q_i=1}^{2^{Nr_i}} \mathbb{P}\left(\left(X_{jf}^N, U_i^N(q_i)\right) \in \mathcal{A}_\epsilon^{(n)}\right) \cdot \left(1 - \mathbb{P}\left(\left(X_{jf}^N, U_i^N(q_i)\right) \in \mathcal{A}_\epsilon^{(n)}\right)\right) \\
&\overset{\text{(b)}}{=} 2^{Nr_i} p_N (1 - p_N) \leq 2^{Nr_i} p_N
\end{aligned}
$$

where (a) is due to independence of the indicator variables, and we have defined $p_N := \mathbb{P}\left(\left(X_{jf}^N, U_i^N(1)\right) \in \mathcal{A}_\epsilon^{(n)}\right)$ in (b). Hence, there exists N_1 such that for all $N > N_1$,

$$
var(2^{NK_i}) \leq 2^{Nr_i} 2^{-N\left[I(X_{jf};U_i) - \delta(\epsilon)\right]}
$$

for some $\delta(\epsilon)$ with $\delta(\epsilon) \to 0$ as $\epsilon \to \infty$.

Define $\eta := 2^{Nr_i} 2^{-N\left[I(X_{jf};U_i) - \delta(\epsilon)/3\right]}(2^{N\delta(\epsilon)/3} - 1)$, and the sequence of events

$$
\mathcal{E}_n := \left\{ \left| 2^{(n+N_1)K_i} - \mathbb{E}\left[2^{(n+N_1)K_i}\right] \right| > \eta \right\}
$$

indexed by $n \geq 1$.

Borel-Cantelli lemma [Dur10] states that if $\sum_{n=1}^\infty \mathbb{P}\left(\mathcal{E}_n\right) < \infty$, then

$$
\mathbb{P}\left(\mathcal{E}_n \text{ infinitely often}\right) = 0.
$$

Then consider

$$
\begin{aligned}
\sum_{n=1}^\infty \mathbb{P}\left(\mathcal{E}_n\right) &= \sum_{n=1}^\infty \mathbb{P}\left(\left| 2^{(n+N_1)K_i} - \mathbb{E}\left[2^{(n+N_1)K_i}\right] \right| > \eta \right) \\
&\overset{\text{(a)}}{\leq} \sum_{n=1}^\infty \frac{var\left(2^{(n+N_1)K_i}\right)}{\eta^2} \\
&\leq \sum_{n=1}^\infty \frac{1}{2^{n\left[r_i - I(X_{jf};U_i) + \delta(\epsilon)/3\right]}(2^{n\delta(\epsilon)/3} - 1)^2} \\
&\overset{\text{(b)}}{<} \infty
\end{aligned}
$$

where (a) follows by Chebyshev's inequality, and (b) is because exponentially decaying series converge, and $r_i > I(\widetilde{V}_j; U_i) \geq I(X_{jf}; U_i)$ where the first inequality is by covering lemma [GK11], and the second is by data processing inequality (recall that $X_{jf} - \widetilde{V}_j - U_i$ is a Markov chain). Therefore, with probability one, there exists a finite integer $N_2 \geq N_1$ such that for all $N \geq N_2$,

$$
2^{NK_i} < 2^{Nr_i} 2^{-N\left[I(X_{jf},U_i) - 2\delta(\epsilon)/3\right]}
$$

Choosing $r_i = I(\widetilde{V}_j; U_i) + \delta(\epsilon)/3$, taking the logarithm of both sides, and dividing by N, we get the desired result.

A.2 Proofs of Lemmas 2.2, 2.3, 2.4 and 2.5

A.2.1 Notation

We will often suppress the dependence on block index b and block length N for brevity.

For any given set of message indices (m_1, n_1, m_2), define the following events, with a little abuse of notation

$$T(m_1, n_1, m_2) := \{\exists\, (q_1, q_2, q_2') \text{ s.t. (2.21) holds for } (m_1, n_1, m_2, q_1, q_2, q_2')\},$$
$$T(m_1, n_1, m_2, q_1, q_2, q_2') := \{\text{(2.21) holds for the indices}(m_1, n_1, m_2, q_1, q_2, q_2')\}$$

We also define the following quantization error event at Txi

$$E_i = \left\{\widetilde{V}_j^N \in \mathcal{T}_{\epsilon'}^{(N)},\ \left(\widetilde{V}_j^N, U_i^N(q_i)\right) \notin \mathcal{A}_\epsilon^{(n)}\ \forall q_i\right\} \cup \left\{\widetilde{V}_j^N \notin \mathcal{T}_{\epsilon'}^{(N)}\right\}$$

for $(i,j) = (1,2), (2,1)$, and $E := E_1 \cup E_2$.

Without loss of generality, we assume that the correct message and quantization indices correspond to the index 1, *i.e.* $(W_{1c}, W_{1p}, W_{2c}, Q_1, Q_2) = (1,1,1,1,1)$ for all blocks. We introduce the notation

$$\bar{\mathcal{B}}_i(b) := \mathcal{B}_i(b) \setminus \{1\}$$

An arbitrary element of the set $\bar{\mathcal{B}}_i(b)$ will be denoted with $\bar{q}_i$, or $\bar{q}_i'$. In this analysis, we focus on an arbitrary block b, but we will also need to refer to variables from block $b-1$. The variables associated with block $b-1$ will be represented with a caron notation when in single letter form. For example, while $\check{X}_{2e}$ is the single letter form for $X_{2e}^N(b-1)$, X_{2e} is the single letter form for $X_{2e}^N(b)$. The feedback state pair $\underline{S} = (S_1, S_2)$ is assumed to be conditioned upon in all the mutual information terms (since the receivers have access to this information causally), but will be omitted for brevity.

A.2.2 Claims

In this subsection, we will prove two simple claims that will be useful in bounding the probability of decoding error.

Claim A.1. *Let A_k, $k = 1, 2, ...$ be a sequence of i.i.d. events. Let $\mathcal{S} \subset \mathbb{N}$ be a random subset of natural numbers (not necessarily independent from the events A_k) such that $|\mathcal{S}| \leq M$ a.s. for some real number M, and $\mathbb{P}(A_k|\mathcal{S}) = \mathbb{P}(A_m|\mathcal{S})$ a.s. for all (k, m) pairs. Then*

$$\mathbb{P}\left(\bigcup_{k\in\mathcal{S}} A_k\right) \leq M\mathbb{P}(A_j)$$

for an arbitrary j.

Proof.

$$\begin{aligned}\mathbb{P}\left(\bigcup_{k\in\mathcal{S}} A_k\right) &= \mathbb{E}\left[\mathbb{E}\left[\mathbb{1}_{\cup_{k\in\mathcal{S}} A_k}|\mathcal{S}\right]\right] \leq \mathbb{E}\left[\mathbb{E}\left[\sum_{k\in\mathcal{S}} \mathbb{1}_{A_k}\middle|\mathcal{S}\right]\right] \\ &\overset{(a)}{=} \mathbb{E}\left[\mathbb{E}\left[|\mathcal{S}|\,\mathbb{1}_{A_1}|\mathcal{S}\right]\right] = \mathbb{E}\left[|\mathcal{S}|\,\mathbb{E}\left[\mathbb{1}_{A_1}|\mathcal{S}\right]\right] \\ &\leq \mathbb{E}\left[M\mathbb{E}\left[\mathbb{1}_{A_1}|\mathcal{S}\right]\right] = M\mathbb{E}\left[\mathbb{E}\left[\mathbb{1}_{A_1}|\mathcal{S}\right]\right] = M\mathbb{P}(A_1)\end{aligned}$$

where (a) follows by the fact that $\mathbb{P}(A_k|\mathcal{S})$ is the same for all k. □

Claim A.2. *Let $\left(X^N, Y^N, Z^N\right)$ be distributed i.i.d. according to $p(x, y, z)$, and $\left(\widetilde{X}^N, \widetilde{Y}^N, \widetilde{Z}^N\right)$ be distributed i.i.d. according to $p(x)p(y)p(z)$. Then there exists $\delta(\epsilon)$ with $\lim_{\epsilon\to 0} \delta(\epsilon) = 0$ such that*

$$\mathbb{P}\left(\left(\widetilde{X}^N, \widetilde{Y}^N, \widetilde{Z}^N\right) \in \mathcal{A}_\epsilon^{(n)}\right) \leq 2^{-N[I(X;Y)+I(Z;X,Y)-\delta(\epsilon)]}$$

Proof.

$$\begin{aligned}\mathbb{P}\left(\left(\widetilde{X}^N, \widetilde{Y}^N, \widetilde{Z}^N\right) \in \mathcal{A}_\epsilon^{(n)}\right) &\leq 2^{-N\left[D\left(P_{X,Y,Z}||P_X P_Y P_Z\right)-\delta(\epsilon)\right]} \\ &= 2^{-N[I(X;Y)+I(Z;X,Y)-\delta(\epsilon)]}\end{aligned}$$

where $D(P||Q)$ is the relative entropy between probability distributions P and Q. □

A.2.3 Proof of Lemma 2.2

We will show that there exists a sequence of codes such that $\mathbb{P}\left(\mathcal{D}_{FB,w}\right) \to 0$ *exponentially*, if the given rate constraints are satisfied, which implies the claimed result. The probability of the decoding error event $\mathcal{D}_{FB,w}$ can be bounded by

$$\begin{aligned}
&\mathbb{P}\left(\mathcal{D}_{FB,w}\right) = \mathbb{P}\left(E\right)\mathbb{P}\left(\mathcal{D}_{FB,w}|E\right) + \mathbb{P}\left(E^c\right)\mathbb{P}\left(\mathcal{D}_{FB,w}|E^c\right) \\
&\leq \mathbb{P}\left(E\right) + \mathbb{P}\left(\mathcal{D}_{FB,w}|E^c\right) \\
&\leq \mathbb{P}\left(E_1\right) + \mathbb{P}\left(E_2\right) + \mathbb{P}\left(\mathcal{D}_{FB,w}|E^c\right) \qquad \text{(A.1)}
\end{aligned}$$

If we choose the rates of the quantization codebooks such that $r_i > I(\widetilde{V}_j;U_i)$, for $(i,j) = (1,2),(2,1)$, by covering lemma [GK11], $\mathbb{P}\left(E_1\right), \mathbb{P}\left(E_2\right) \to 0$. Therefore, it is sufficient to show that $\mathbb{P}\left(\mathcal{D}_{FB,w}|E^c\right)$ vanishes if the conditions in the lemma are satisfied.

The decoding error event $\mathcal{D}_{FB,w}$ can also be expressed as the following union of events.

$$\begin{aligned}
\mathcal{D}_{FB,w} = &\bigcup_{m_1\neq 1} T(m_1,1,1) \cup \bigcup_{n_1\neq 1} T(1,n_1,1) \cup \bigcup_{m_2\neq 1} T(1,1,m_2) \cup \bigcup_{\substack{m_1\neq 1\\ n_1\neq 1}} T(m_1,n_1,1) \\
&\cup \bigcup_{\substack{m_1\neq 1\\ m_2\neq 1}} T(m_1,1,m_2) \cup \bigcup_{\substack{n_1\neq 1\\ m_2\neq 1}} T(1,n_1,m_2) \cup \bigcup_{\substack{m_1\neq 1\\ n_1\neq 1\\ m_2\neq 1}} T(m_1,n_1,m_2) \cup T^c(1,1,1) \qquad \text{(A.2)}
\end{aligned}$$

Using the union bound on (A.2), probability of decoding error conditioned on quantization success can be bounded by

$$\begin{aligned}
&\mathbb{P}\left(\mathcal{D}_{FB,w}|E^c\right) = 2^{NR_{1c}}\mathbb{P}\left(T(m_1,1,1)|E^c\right) + 2^{NR_{1p}}\mathbb{P}\left(T(1,n_1,1)|E^c\right) \\
&\quad + 2^{NR_{2c}}\mathbb{P}\left(T(1,1,m_2)|E^c\right) + 2^{NR_1}\mathbb{P}\left(T(m_1,n_1,1)|E^c\right) \\
&\quad + 2^{N(R_{1c}+R_{2c})}\mathbb{P}\left(T(m_1,1,m_2)|E^c\right) + 2^{N(R_{1p}+R_{2c})}\mathbb{P}\left(T(1,n_1,m_2)|E^c\right) \\
&\quad + 2^{N(R_1+R_{2c})}\mathbb{P}\left(T(m_1,n_1,m_2)|E^c\right) + \mathbb{P}\left(T^c(1,1,1)|E^c\right) \qquad \text{(A.3)}
\end{aligned}$$

Note that conditioned on successful quantization, the relevant random variables are distributed i.i.d. over time according to the joint distribution

$$p(\check{x}_{1p},\check{x}_{2e},\check{x}_{2c},u_1,x_{1e},x_{1c},x_{2e},\check{y}_1,y_1) = p(\check{x}_{1p})p(\check{x}_{2e})p(\check{x}_{2c})p(x_{1e})p(x_{1c})p(x_{2e})$$

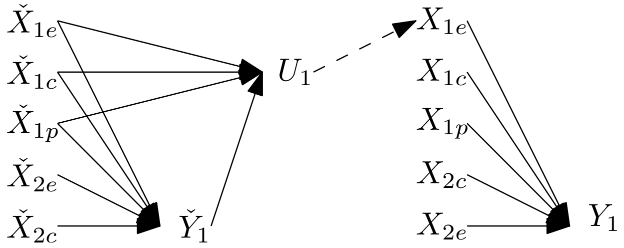


Figure A.1: Markov network showing the dependence of the relevant variables. The variables connected with the dashed arrow are independent in the single-letter form, although in multi-letter form they are not.

$$\cdot\, p(u_1|\check{y}_1, \check{x}_{1p})p(\check{y}_1|\check{x}_{2e}, \check{x}_{2c}, \check{x}_{1p})p(y_1|x_{1e}, x_{1c}, x_{2e}) \tag{A.4}$$

Next, we bound the error terms one by one. In what follows, joint typicality is sought with respect to the joint distribution in $(A.4)$. The first term is bounded by $\mathbb{P}\left(T^c(1,1,1)|E^c\right) < \epsilon$ by law of large numbers.

The second error term in (A.3) can be bounded as follows.

$$\begin{aligned}
&\mathbb{P}\left(T(m_1,1,1)|E^c\right) = \mathbb{P}\left(\bigcup_{q_1,q_2,q_2'} T(m_1,1,1,q_1,q_2,q_2')\middle| E^c\right)\\
&= \mathbb{P}\left(\bigcup_{\substack{q_1\neq 1, q_2\neq 1\\ q_2'\neq 1}} T(m_1,1,1,q_1,q_2,q_2') \cup \bigcup_{\substack{q_2\neq 1\\ q_2'\neq 1}} T(m_1,1,1,1,q_2,q_2')\right.\\
&\cup \bigcup_{\substack{q_1\neq 1\\ q_2'\neq 1}} T(m_1,1,1,q_1,1,q_2') \cup \bigcup_{\substack{q_1\neq 1\\ q_2\neq 1}} T(m_1,1,1,q_1,q_2,1) \cup \bigcup_{q_1\neq 1} T(m_1,1,1,q_1,1,1)\\
&\left.\cup \bigcup_{q_2\neq 1} T(m_1,1,1,1,q_2,1) \cup \bigcup_{q_2'\neq 1} T(m_1,1,1,1,1,q_2') \cup T(m_1,1,1,1,1,1)\ \middle|\ E^c\right)\\
&\overset{\text{(a)}}{=} \mathbb{P}\left(\bigcup_{\substack{q_1\in\bar{\mathcal{B}}_1(1,q_2)\\ q_2\in\bar{\mathcal{B}}_2(b-1), q_2'\in\bar{\mathcal{B}}_2(b)}} T(m_1,1,1,q_1,q_2,q_2') \cup \bigcup_{\substack{q_2\in\bar{\mathcal{B}}_2\\ q_2'\in\bar{\mathcal{B}}_2(b)}} T(m_1,1,1,1,q_2,q_2')\right.
\end{aligned}$$

$$
\begin{aligned}
&\cup \bigcup_{\substack{q_1\in\bar{\mathcal{B}}_1(b)\\ q_2'\in\bar{\mathcal{B}}_2(b)}} T(m_1,1,1,q_1,1,q_2') \cup \bigcup_{\substack{q_1\in\bar{\mathcal{B}}_1(1,q_2)\\ q_2\in\bar{\mathcal{B}}_2(b-1)}} T(m_1,1,1,q_1,q_2,1)\\
&\cup \bigcup_{q_1\in\bar{\mathcal{B}}_1(b)} T(m_1,1,1,q_1,1,1) \cup \bigcup_{q_2\in\bar{\mathcal{B}}_2(b-1)} T(m_1,1,1,1,q_2,1)\\
&\cup \bigcup_{q_2'\in\bar{\mathcal{B}}_2(b)} T(m_1,1,1,1,1,q_2') \cup T(m_1,1,1,1,1,1) \left| E^c \right)\\
\overset{(b)}{\le}\ & \mathbb{P}\left(\bigcup_{\substack{q_1\in\bar{\mathcal{B}}_1(1,q_2)\\ q_2\in\bar{\mathcal{B}}_2(b-1), q_2'\in\bar{\mathcal{B}}_2(b)}} T(m_1,1,1,q_1,q_2,q_2') \middle| E^c \right)\\
&+\mathbb{P}\left(\bigcup_{\substack{q_2\in\bar{\mathcal{B}}_2(b-1)\\ q_2'\in\bar{\mathcal{B}}_2(b)}} T(m_1,1,1,1,q_2,q_2') \middle| E^c \right)\\
&+\mathbb{P}\left(\bigcup_{\substack{q_1\in\bar{\mathcal{B}}_1(b)\\ q_2'\in\bar{\mathcal{B}}_2(b)}} T(m_1,1,1,q_1,1,q_2') \middle| E^c \right)\\
&+\mathbb{P}\left(\bigcup_{\substack{q_1\in\bar{\mathcal{B}}_1(1,q_2(b-1))\\ q_2\in\bar{\mathcal{B}}_2(b-1)}} T(m_1,1,1,q_1,q_2,1) \middle| E^c \right)\\
&+\mathbb{P}\left(\bigcup_{q_1\in\bar{\mathcal{B}}_1(b)} T(m_1,1,1,q_1,1,1) \middle| E^c \right) + \mathbb{P}\left(\bigcup_{q_2\in\bar{\mathcal{B}}_2(b-1)} T(m_1,1,1,1,q_2,1) \middle| E^c \right)\\
&+\mathbb{P}\left(\bigcup_{q_2'\in\bar{\mathcal{B}}_2(b)} T(m_1,1,1,1,1,q_2') \middle| E^c \right) + \mathbb{P}\left(T(m_1,1,1,1,1,1)|E^c\right)\\
\overset{(c)}{=}\ & 2^{N(\kappa_1+2\kappa_2)}\mathbb{P}\left(T(m_1,1,1,q_1,q_2,q_2')|E^c\right) + 2^{2N\kappa_2}\mathbb{P}\left(T(m_1,1,1,1,q_2,q_2')|E^c\right)\\
&+2^{N(\kappa_1+\kappa_2)}\mathbb{P}\left(T(m_1,1,1,q_1,1,q_2')|E^c\right) + 2^{N(\kappa_1+\kappa_2)}\mathbb{P}\left(T(m_1,1,1,q_1,q_2,1)|E^c\right)\\
&+2^{N\kappa_1}\mathbb{P}\left(T(m_1,1,1,q_1,1,1)|E^c\right) + 2^{N\kappa_2}\mathbb{P}\left(T(m_1,1,1,1,q_2,1)|E^c\right)\\
&+2^{N\kappa_2}\mathbb{P}\left(T(m_1,1,1,1,1,q_2')|E^c\right) + \mathbb{P}\left(T(m_1,1,1,1,1,1)|E^c\right)
\end{aligned}
$$

$$
\begin{aligned}
&\overset{\text{(d)}}{\leq} 2^{N(\kappa_1+2\kappa_2)}2^{-N\left[I(U_1;\check{X}_{2e})+I(\check{X}_{2e},U_1,X_{1f},X_{2e};Y_1,\check{Y}_1,\check{X}_{2c}|\check{X}_1)-\delta(\epsilon)\right]}\\
&\quad + 2^{2N\kappa_2}2^{-N\left[I(X_{1f},X_{2e},\check{X}_{2e};\check{Y}_1,Y_1|U_1,X_{1e},\check{X}_{2c},\check{X}_1)-\delta(\epsilon)\right]}\\
&\quad + 2^{N(\kappa_1+\kappa_2)}2^{-N\left[I(U_1;\check{X}_{2e})+I(\check{X}_{2e},U_1,X_{1f};Y_1,\check{Y}_1,\check{X}_{2c}|\check{X}_1,X_{2e})-\delta(\epsilon)\right]}\\
&\quad + 2^{N(\kappa_1+\kappa_2)}2^{-N\left[I(X_{1f},X_{2e},U_1;Y_1,\check{Y}_1,\check{X}_{2f}|\check{X}_1)-\delta(\epsilon)\right]}\\
&\quad + 2^{N\kappa_1}2^{-N\left[I(X_{1f},U_1;Y_1,\check{Y}_1,\check{X}_{2f}|\check{X}_1,X_{2e})-\delta(\epsilon)\right]}\\
&\quad + 2^{N\kappa_2}2^{-N\left[I(X_{1f},\check{X}_{2e};U_1,Y_1,\check{Y}_1|X_{1e},X_{2e},\check{X}_1,\check{X}_{2c})-\delta(\epsilon)\right]}\\
&\quad + 2^{N\kappa_2}2^{-N\left[I(X_{1f},X_{2e};Y_1|X_{1e},U_1,\check{Y}_1,\check{X}_1,\check{X}_{2f})-\delta(\epsilon)\right]}\\
&\quad + 2^{-N\left[I(X_{1f};Y_1|X_{1e},U_1,\check{Y}_1,\check{X}_1,\check{X}_{2f},X_{2e})-\delta(\epsilon)\right]}\\
&\overset{\text{(e)}}{=} 2^{N(\kappa_1+2\kappa_2)}2^{-N\left[I(U_1;\check{X}_{2e})+I(X_{1f},X_{2e};Y_1)+I(\check{X}_{2e};\check{Y}_1|\check{X}_1,\check{X}_{2c})-\delta(\epsilon)\right]}\\
&\quad + 2^{-N\left[I(X_{1f},X_{2e},\check{X}_{2e};\check{Y}_1,Y_1|U_1,X_{1e},\check{X}_{2c},\check{X}_1)-2\kappa_2-\delta(\epsilon)\right]}\\
&\quad + 2^{N(\kappa_1+\kappa_2)}2^{-N\left[I(U_1;\check{X}_{2e})+I(X_{1f};Y_1|X_{2e})+I(\check{X}_{2e};\check{Y}_1|\check{X}_1,\check{X}_{2c})-\delta(\epsilon)\right]}\\
&\quad + 2^{-N\left[I(X_{1f},X_{2e};Y_1)+I(U_1;\check{Y}_1|\check{X}_1,\check{X}_{2f})-\kappa_1-\kappa_2-\delta(\epsilon)\right]}\\
&\quad + 2^{-N\left[I(X_{1f};Y_1|X_{2e})+I(U_1;\check{Y}_1|\check{X}_1,\check{X}_{2f})-\kappa_1-\delta(\epsilon)\right]}\\
&\quad + 2^{-N\left[I(X_{1f},\check{X}_{2e};U_1,Y_1,\check{Y}_1|X_{1e},X_{2e},\check{X}_1,\check{X}_{2c})-\kappa_2-\delta(\epsilon)\right]}\\
&\quad + 2^{-N\left[I(X_{1f},X_{2e};Y_1|X_{1e})-\kappa_2-\delta(\epsilon)\right]}\\
&\quad + 2^{-N\left[I(X_{1f};Y_1|X_{1e},X_{2e})-\delta(\epsilon)\right]}\\
&\overset{\text{(f)}}{\leq} 8\cdot 2^{-N\left[I(X_{1f};Y_1|X_{1e},X_{2e})-C_1-\delta(\epsilon)\right]}
\end{aligned}
$$

where

- (a) is since $T(m_1,1,1,q_1,q_2,q_2')$ is empty set for $q_1 \notin \mathcal{B}_1(b)$, $q_2 \notin \mathcal{B}_2(b-1)$, or $q_2' \notin \mathcal{B}_2((q_2,m_2)(b))$, since for random variables $(X,Y,Z) \sim p(x,y,z)$, $(X,Y,Z) \in \mathcal{A}_\epsilon^{(n)}$ implies $(X,Y) \in \mathcal{A}_\epsilon^{(n)}$,
- (b) follows by union bound,
- (c) follows by Claim A.1, where the upper bound on the size of the $\bar{\mathcal{B}}_i$ sets for sufficiently large N is given by Lemma 2.1,

- (d) follows by packing lemma and Claim A.2,
- (e) is by manipulating the mutual information terms using the dependence structure of the involved variables (see Figure A.1),
- (f) is by upper bounding each of the eight terms with the same bound, using chain rule and non-negativity of mutual information.

Next, we bound the term $\mathbb{P}\left(T(1,n_1,m_2)|E^c\right)$. We apply steps (a)-(d), which are also applicable here, to obtain the following.

$$\begin{aligned}
\mathbb{P}\left(T(1,n_1,m_2)|E^c\right) &\leq 2^{N(\kappa_1+2\kappa_2)}2^{-N\left[I(U_1;\check{X}_{2f})+I(\check{X}_1,\check{X}_{2f},X_{2e},U_1,X_{1e};Y_1,\check{Y}_1,X_{1c}|\check{X}_{1f})-\delta(\epsilon)\right]} \\
&\quad + 2^{2N\kappa_2}2^{-N\left[I(\check{X}_1,\check{X}_{2f},X_{2e};\check{Y}_1,Y_1,U_1|X_{1f},\check{X}_{1f})-\delta(\epsilon)\right]} \\
&\quad + 2^{N(\kappa_1+\kappa_2)}2^{-N\left[I(U_1;\check{X}_{2f})+I(\check{X}_{2f},\check{X}_1,U_1,X_{1e};Y_1,\check{Y}_1|\check{X}_{1f},X_{1c},X_{2e})-\delta(\epsilon)\right]} \\
&\quad + 2^{N(\kappa_1+\kappa_2)}2^{-N\left[I(\check{X}_{2f},\check{X}_1,U_1,X_{1e},X_{2e};Y_1,\check{Y}_1|\check{X}_{1f},X_{1c},\check{X}_{2e})-\delta(\epsilon)\right]} \\
&\quad + 2^{N\kappa_1}2^{-N\left[I(\check{X}_1,\check{X}_{2f},U_1,X_{1e};Y_1,\check{Y}_1,\check{X}_{2e}|\check{X}_{1f},X_{1c},X_{2e})-\delta(\epsilon)\right]} \\
&\quad + 2^{N\kappa_2}2^{-N\left[I(\check{X}_1,\check{X}_{2f};U_1,Y_1,\check{Y}_1|X_{1f},X_{2e},\check{X}_{1f})-\delta(\epsilon)\right]} \\
&\quad + 2^{N\kappa_2}2^{-N\left[I(\check{X}_1,\check{X}_{2f},X_{2e};Y_1,\check{Y}_1,U_1|\check{X}_{1f},X_{1f},\check{X}_{2e})-\delta(\epsilon)\right]} \\
&\quad + 2^{-N\left[I(\check{X}_1,\check{X}_{2f};\check{Y}_1,U_1|Y_1,\check{X}_{1f},\check{X}_{2e},X_{1e},X_{2e})-\delta(\epsilon)\right]} \\
&\overset{\text{(e)}}{\leq} 4\cdot 2^{-N\left[I(\check{X}_{2f},\check{X}_1,U_1,X_{1e};Y_1,\check{Y}_1|\check{X}_{1f},\check{X}_{2e},X_{1c},X_{2e})-\delta(\epsilon)\right]} \\
&\quad + 4\cdot 2^{-N\left[I(\check{X}_{2f},\check{X}_1;Y_1,\check{Y}_1,U_1|\check{X}_{1f},\check{X}_{2e},X_{1f},X_{2e})-\delta(\epsilon)\right]} \\
&\overset{\text{(f)}}{\leq} 4\cdot 2^{-N\left[I(X_{2f},X_1;Y_1|X_{1c},X_{2e})-\delta(\epsilon)\right]} \\
&\quad + 4\cdot 2^{-N\left[I(\check{X}_{2f},\check{X}_1;\check{Y}_1,U_1|\check{X}_{1f},\check{X}_{2e})-\delta(\epsilon)\right]}
\end{aligned}$$

where (e) is by upper bounding the first, third, fourth and fifth terms with the first term in (j), and the rest of the terms with the second; (f) is by rearranging the mutual information terms using chain rule and the fact that the distribution of variables is the same for each block. The conditioning on $\check{X}_{1f}$ is because the messages corresponding to this variable has already been decoded in the previous block.

In order to bound the term $\mathbb{P}\left(T(1,n_1,1)|E^c\right)$ in (A.3), we note that the the joint distribution (A.4) has a similar structure with respect to X_{1c} and $\check{X}_{1p}$, with the following mapping between random variables

$$\begin{aligned}\check{X}_{1p} &\leftrightarrow X_{1c},\\ \left(\check{Y}_1, U_1\right) &\leftrightarrow Y_1,\\ \left(\check{X}_{2e}, \check{X}_{2c}\right) &\leftrightarrow (X_{1e}, X_{2e})\end{aligned}$$

Therefore, one can perform the steps (a)-(f) for the third error term as well, by switching the variables as above, to obtain the following bound

$$\begin{aligned}\mathbb{P}\left(T(1,n_1,1)|E^c\right) &\leq 4\cdot 2^{-N\left[I(\check{X}_1;\check{Y}_1,U_1|\check{X}_{1f},\check{X}_{2f})-C_1-\delta(\epsilon)\right]}\\ &\quad + 4\cdot 2^{-N\left[I(\check{X}_1;\check{Y}_1,U_1|\check{X}_{1c},\check{X}_{2f})-C_1-\delta(\epsilon)\right]}\\ &\leq 8\cdot 2^{-N\left[I(\check{X}_1;\check{Y}_1|\check{X}_{1f},\check{X}_{2f})-C_1-\delta(\epsilon)\right]}\\ &= 8\cdot 2^{-N\left[I(X_1;Y_1|X_{1f},X_{2f})-C_1-\delta(\epsilon)\right]}\end{aligned}$$

We have dropped the U_1 variable from the mutual information term for the sake of simplicity in evaluating the rate region, since its contribution is small. In the final step, we used the fact that the distribution of variables is the same for each block. We can obtain the following bounds for each error term in a similar way, by exploiting the structure of the joint distribution as done above and noting that the steps (a)-(f) are applicable with an appropriate mapping between the variables.

$$\begin{aligned}\mathbb{P}\left(T(1,1,m_2)|E^c\right) &\leq 8\cdot 2^{-N\left[I(\check{X}_{2f};\check{Y}_1,U_1|\check{X}_1,\check{X}_{2e})-C_1-\delta(\epsilon)\right]}\\ &\leq 8\cdot 2^{-N\left[I(\check{X}_{2f};\check{Y}_1|\check{X}_1,\check{X}_{2e})-C_1-\delta(\epsilon)\right]}\\ &= 8\cdot 2^{-N\left[I(X_{2f};Y_1|X_1,X_{2e})-C_1-\delta(\epsilon)\right]}\\ \mathbb{P}\left(T(m_1,n_1,1)|E^c\right) &\leq 8\cdot 2^{-N\left[I(\check{X}_1,X_{1f};\check{Y}_1,Y_1,U_1|\check{X}_{1f},\check{X}_{2f},X_{1e},X_{2e})-C_1-\delta(\epsilon)\right]}\\ &\leq 8\cdot 2^{-N\left[I(\check{X}_1,X_{1f};\check{Y}_1,Y_1|\check{X}_{1f},\check{X}_{2f},X_{1e},X_{2e})-C_1-\delta(\epsilon)\right]}\\ &= 8\cdot 2^{-N\left[I(\check{X}_1;\check{Y}_1|\check{X}_{1f},\check{X}_{2f})+I(X_{1f};Y_1|X_{1e},X_{2e})-C_1-\delta(\epsilon)\right]}\\ &= 8\cdot 2^{-N\left[I(X_1;Y_1|X_{1f},X_{2f})+I(X_{1f};Y_1|X_{1e},X_{2e})-C_1-\delta(\epsilon)\right]}\end{aligned}$$

$$
\begin{aligned}
\mathbb{P}\left(T(m_1,1,m_2)|E^c\right) &\leq 8\cdot 2^{-N\left[I(\check{X}_{2f},X_{1f};\check{Y}_1,Y_1,U_1|\check{X}_1,\check{X}_{2e},X_{1e},X_{2e})-C_1-\delta(\epsilon)\right]}\\
&\leq 8\cdot 2^{-N\left[I(\check{X}_{2f},X_{1f};\check{Y}_1,Y_1|\check{X}_1,\check{X}_{2e},X_{1e},X_{2e})-C_1-\delta(\epsilon)\right]}\\
&= 8\cdot 2^{-N\left[I(\check{X}_{2f};\check{Y}_1|\check{X}_1,\check{X}_{2e})+I(X_{1f};Y_1|X_{1e},X_{2e})-C_1-\delta(\epsilon)\right]}\\
&= 8\cdot 2^{-N\left[I(X_{2f};Y_1|X_1,X_{2e})+I(X_{1f};Y_1|X_{1e},X_{2e})-C_1-\delta(\epsilon)\right]}\\
\mathbb{P}\left(T(m_1,n_1,m_2)|E^c\right) &\leq 8\cdot 2^{-N\left[I(\check{X}_1,\check{X}_{2f},X_{1f};\check{Y}_1,Y_1,U_1|\check{X}_{1f},\check{X}_{2e},X_{1e},X_{2e})-C_1-\delta(\epsilon)\right]}\\
&\leq 8\cdot 2^{-N\left[I(\check{X}_1,\check{X}_{2f},X_{1f};\check{Y}_1,Y_1|\check{X}_{1f},\check{X}_{2e},X_{1e},X_{2e})-C_1-\delta(\epsilon)\right]}\\
&= 8\cdot 2^{-N\left[I(\check{X}_1,\check{X}_{2f};\check{Y}_1|\check{X}_{1f},\check{X}_{2e})+I(X_{1f};Y_1|X_{1e},X_{2e})-C_1-\delta(\epsilon)\right]}\\
&= 8\cdot 2^{-N\left[I(X_1,X_{2f};Y_1|X_{1f},X_{2e})+I(X_{1f};Y_1|X_{1e},X_{2e})-C_1-\delta(\epsilon)\right]}\\
&= 8\cdot 2^{-N\left[I(X_1,X_{2f};Y_1|X_{1e},X_{2e})-C_1-\delta(\epsilon)\right]}
\end{aligned}
$$

Using these bounds in (A.3), it is easy to see that if the following are satisfied, then $\mathbb{P}\left(\mathcal{D}_{FB,w}|E^c\right) \to 0$ as $N \to \infty$ (Note that the bounds on $R_{1p} + R_{1c}$ and $R_{1c} + R_{2c}$ are redundant, as they can be expressed as a sum of other bounds),

$$
\begin{aligned}
R_{1p} &< I(X_1;Y_1|X_{1f},X_{2f}) - C_1 && \text{(A.5)}\\
R_{1c} &< I(X_{1f};Y_1|X_{1e},X_{2e}) - C_1 && \text{(A.6)}\\
R_{1p} + R_{2c} &< \min\{I(X_{2f},X_1;Y_1|X_{1c},X_{2e}),\\
&\qquad I(X_{2f},X_1;Y_1,U_1|X_{1f},X_{2e})\} - C_1 && \text{(A.7)}\\
R_1 + R_{2c} &< I(X_1,X_{2f};Y_1|X_{1e},X_{2e}) - C_1 && \text{(A.8)}
\end{aligned}
$$

The rate constraint on $R_{1p} + R_{2c}$ provided in the lemma is slightly stricter, which allows us to show the redundancy of some of the bounds obtained later.

A.2.4 Proof of Lemma 2.4

We will show that there exists a sequence of codes such that $\mathbb{P}\left(\mathcal{D}_{FB,s}\right) \to 0$ *exponentially*, if the given rate constraints are satisfied, which implies the claimed result.

Similar to the case of weak interference, choosing the quantization rates such that $r_i >$

$I(\widetilde{V}_j; U_i)$, for $(i,j) = (1,2), (2,1)$, probability of decoding error can be bounded by

$$\begin{aligned}\mathbb{P}\left(\mathcal{D}_{FB,s}|E^c\right) = {} & \mathbb{P}\left(T^c(1,1,1)|E^c\right) + 2^{NR_{1p}}\mathbb{P}\left(T(1,n_1,1)|E^c\right) \\ & + 2^{NR_{2c}}\mathbb{P}\left(T(1,1,m_2)|E^c\right) + 2^{NR_1}\mathbb{P}\left(T(m_1,n_1,1)|E^c\right) \\ & + 2^{N(R_{1p}+R_{2c})}\mathbb{P}\left(T(1,n_1,m_2)|E^c\right) \\ & + 2^{N(R_1+R_{2c})}\mathbb{P}\left(T(m_1,n_1,m_2)|E^c\right) \end{aligned} \tag{A.9}$$

Note that conditioned on successful quantization, the relevant random variables are distributed i.i.d. over time according to the joint distribution

$$\begin{aligned} p(\check{x}_{1c}, \check{x}_{1p}, \check{x}_{2e}, x_{2c}, u_2, x_{2e}, x_{1e}, \check{y}_1, y_1) = {} & p(\check{x}_{1c})p(\check{x}_{1p})p(\check{x}_{2e})p(x_{1e})p(x_{1c})p(x_{2e}) \\ & \cdot p(u_2|\check{x}_{1c}, \check{x}_{1p})p(\check{y}_1|\check{x}_{1c}, \check{x}_{1c}, \check{x}_{2e})p(y_1|x_{1e}, x_{2e}, x_{2c}). \end{aligned} \tag{A.10}$$

Next, we bound the error terms one by one. In what follows, joint typicality is sought with respect to the joint distribution in (A.10). The first term is bounded by $\mathbb{P}\left(T^c(1,1,1)|E^c\right) < \epsilon$ by law of large numbers.

Now we take the third term, which is bounded as follows.

$$\begin{aligned} \mathbb{P}\left(T(1,n_1,1)|E^c\right) &= \mathbb{P}\left(\bigcup_{q_1,q_2,q_2'} T(1,n_1,1,q_1,q_2,q_2')|E^c\right) \\ &= \mathbb{P}\Bigg(\bigcup_{\substack{q_1\neq 1, q_2\neq 1 \\ q_2'\neq 1}} T(1,n_1,1,q_1,q_2,q_2') \cup \bigcup_{\substack{q_2\neq 1 \\ q_2'\neq 1}} T(1,n_1,1,1,q_2,q_2') \\ &\quad \cup \bigcup_{\substack{q_1\neq 1 \\ q_2'\neq 1}} T(1,n_1,1,q_1,1,q_2') \cup \bigcup_{\substack{q_1\neq 1 \\ q_2\neq 1}} T(1,n_1,1,q_1,q_2,1) \cup \bigcup_{q_1\neq 1} T(1,n_1,1,q_1,1,1) \\ &\quad \cup \bigcup_{q_2\neq 1} T(1,n_1,1,1,q_2,1) \cup \bigcup_{q_2'\neq 1} T(1,n_1,1,1,1,q_2') \cup T(1,n_1,1,1,1,1) \,\Bigg|\, E^c \Bigg) \\ &\overset{\text{(a)}}{=} \mathbb{P}\Bigg(\bigcup_{\substack{q_1\in\bar{\mathcal{B}}_1(q_2), q_2\in\bar{\mathcal{B}}_2(b-1) \\ q_2'\in\bar{\mathcal{B}}_2(b)}} T(1,n_1,1,q_1,q_2,q_2') \cup \bigcup_{\substack{q_2\in\bar{\mathcal{B}}_2(b-1) \\ q_2'\in\bar{\mathcal{B}}_2(b)}} T(1,n_1,1,1,q_2,q_2') \end{aligned}$$

$$
\begin{aligned}
&\cup \bigcup_{\substack{q_1\in\bar{\mathcal{B}}_1(q_2)\\ q_2'\in\bar{\mathcal{B}}_2(b)}} T(1,n_1,1,q_1,1,q_2') \cup \bigcup_{\substack{q_1\in\bar{\mathcal{B}}_1(q_2)\\ q_2\in\bar{\mathcal{B}}_2(b-1)}} T(1,n_1,1,q_1,q_2,1)\\
&\cup \bigcup_{q_1\in\bar{\mathcal{B}}_1(q_2)} T(1,n_1,1,q_1,1,1) \cup \bigcup_{q_2\in\bar{\mathcal{B}}_2(b-1)} T(1,n_1,1,1,q_2,1)\\
&\cup \bigcup_{q_2'\in\bar{\mathcal{B}}_2((b)} T(1,n_1,1,1,1,q_2') \cup T(1,n_1,1,1,1,1) \Bigg| E^c \Bigg)\\
&\overset{\text{(b)}}{\leq} \mathbb{P}\Bigg(\bigcup_{\substack{q_1\in\bar{\mathcal{B}}_1(q_2),q_2\in\bar{\mathcal{B}}_2(b-1)\\ q_2'\in\bar{\mathcal{B}}_2(b)}} T(1,n_1,1,q_1,q_2,q_2') \Bigg| E^c \Bigg)\\
&+ \mathbb{P}\Bigg(\bigcup_{\substack{q_2\in\bar{\mathcal{B}}_2(b-1)\\ q_2'\in\bar{\mathcal{B}}_2(b)}} T(1,n_1,1,1,q_2,q_2') \Bigg| E^c \Bigg)\\
&+ \mathbb{P}\Bigg(\bigcup_{\substack{q_1\in\bar{\mathcal{B}}_1(q_2)\\ q_2'\in\bar{\mathcal{B}}_2(b)}} T(1,n_1,1,q_1,1,q_2') \Bigg| E^c \Bigg) + \mathbb{P}\Bigg(\bigcup_{\substack{q_1\in\bar{\mathcal{B}}_1(q_2)\\ q_2\in\bar{\mathcal{B}}_2(b-1)}} T(1,n_1,1,q_1,q_2,1) \Bigg| E^c \Bigg)\\
&+ \mathbb{P}\Bigg(\bigcup_{q_1\in\bar{\mathcal{B}}_1(q_2)} T(1,n_1,1,q_1,1,1) \Bigg| E^c \Bigg) + \mathbb{P}\Bigg(\bigcup_{q_2\in\bar{\mathcal{B}}_2(b-1)} T(1,n_1,1,1,q_2,1) \Bigg| E^c \Bigg)\\
&+ \mathbb{P}\Bigg(\bigcup_{q_2'\in\bar{\mathcal{B}}_2(b)} T(1,n_1,1,1,1,q_2') \Bigg| E^c \Bigg) + \mathbb{P}\left(T(1,n_1,1,1,1,1)|E^c\right)\\
&\overset{\text{(c)}}{\leq} 2^{N(\kappa_1+2\kappa_2)}\mathbb{P}\left(T(1,n_1,1,q_1,q_2,q_2')|E^c\right) + 2^{2N\kappa_2}\mathbb{P}\left(T(1,n_1,1,1,q_2,q_2')|E^c\right)\\
&\quad + 2^{N(\kappa_1+\kappa_2)}\mathbb{P}\left(T(1,n_1,1,q_1,1,q_2')|E^c\right) + 2^{N(\kappa_1+\kappa_2)}\mathbb{P}\left(T(1,n_1,1,q_1,q_2,1)|E^c\right)\\
&\quad + 2^{N\kappa_1}\mathbb{P}\left(T(1,n_1,1,\bar{q}_1,1,1)|E^c\right) + 2^{N\kappa_2}\mathbb{P}\left(T(1,n_1,1,1,\bar{q}_2,1)|E^c\right)\\
&\quad + 2^{N\kappa_2}\mathbb{P}\left(T(1,n_1,1,1,1,\bar{q}_2')|E^c\right) + \mathbb{P}\left(T(1,n_1,1,1,1,1)|E^c\right)\\
&\overset{\text{(d)}}{\leq} 2^{N(\kappa_1+2\kappa_2)}2^{-NI(U_2;\check{X}_{1p}|\check{X}_{1e})}2^{-N\left[I(\check{X}_{1p},\check{X}_{2e},U_2,X_{1e},X_{2e};Y_1,\check{Y}_1,X_{2c},\check{X}_{1c}|\check{X}_{1e},\check{X}_{2c})-\delta(\epsilon)\right]}\\
&\quad + 2^{2N\kappa_2}2^{-NI(U_2;\check{X}_{1p})}2^{-N\left[I(\check{X}_{1p},\check{X}_{2e},U_2,X_{2e};Y_1,\check{Y}_1,X_{1e},X_{2c},\check{X}_{1c}|\check{X}_{1e},\check{X}_{2c})-\delta(\epsilon)\right]}\\
&\quad + 2^{N(\kappa_1+\kappa_2)}2^{-NI(U_2;\check{X}_{1p})}2^{-N\left[I(\check{X}_{1p},U_2,X_{1e},X_{2e};Y_1,\check{Y}_1,\check{X}_{2e},X_{2c},\check{X}_{1c}|\check{X}_{1e},\check{X}_{2c})-\delta(\epsilon)\right]}
\end{aligned}
$$

$$
\begin{aligned}
&+ 2^{N(\kappa_1+\kappa_2)} 2^{-N\left[I(\check{X}_{1p},\check{X}_{2e},X_{1e};Y_1,\check{Y}_1,U_2,X_{2f},\check{X}_{1c}|\check{X}_{1e},\check{X}_{2c})-\delta(\epsilon)\right]} \\
&+ 2^{N\kappa_1} 2^{-N\left[I(\check{X}_{1p},X_{1e};Y_1,\check{Y}_1,\check{X}_{2f},U_2,X_{2f},\check{X}_{1c}|\check{X}_{1e},\check{X}_{2c})-\delta(\epsilon)\right]} \\
&+ 2^{N\kappa_2} 2^{-NI(U_2;\check{X}_{1p}|\check{X}_{1e})} 2^{-N\left[I(\check{X}_1,U_2,X_{2e};Y_1,\check{Y}_1,\check{X}_{2e},X_{2c},X_{1e},\check{X}_{1c}|\check{X}_{1e},\check{X}_{2c})-\delta(\epsilon)\right]} \\
&+ 2^{N\kappa_2} 2^{-N\left[I(\check{X}_{1p},\check{X}_{2e};Y_1,\check{Y}_1,X_{2c},U_2,X_{1e},X_{2e},\check{X}_{1c}|\check{X}_{1e},\check{X}_{2c})-\delta(\epsilon)\right]} \\
&+ 2^{-N\left[I(\check{X}_{1p};Y_1,\check{Y}_1,\check{X}_{2c},\check{X}_{2e},U_2,X_{1e},X_{2e},\check{X}_{1c}|\check{X}_{1e},\check{X}_{2c})-\delta(\epsilon)\right]} \\
\overset{\text{(e)}}{\leq}\ & 2^{N(\kappa_1+2\kappa_2)} 2^{-N\left[I(\check{X}_{1p},\check{X}_{2e},X_{1e},X_{2e};\check{Y}_1,Y_1,U_2|\check{X}_{1f},X_{2c},\check{X}_{2c})-\delta(\epsilon)\right]} \\
&+ 2^{2N\kappa_2} 2^{-N\left[I(\check{X}_{1p},\check{X}_{2e},U_2,X_{2e};Y_1,\check{Y}_1|\check{X}_{1f},X_{1e},X_{2c},\check{X}_{2c})-\delta(\epsilon)\right]} \\
&+ 2^{N(\kappa_1+\kappa_2)} 2^{-N\left[I(\check{X}_{1p},X_{1e},X_{2e};\check{Y}_1,Y_1,U_2|\check{X}_{1e},\check{X}_{1c},\check{X}_{2f},X_{2c})-\delta(\epsilon)\right]} \\
&+ 2^{N(\kappa_1+\kappa_2)} 2^{-N\left[I(\check{X}_{1p},\check{X}_{2e},X_{1e};Y_1,\check{Y}_1,U_2|\check{X}_{1f},\check{X}_{2c},X_{2f})-\delta(\epsilon)\right]} \\
&+ 2^{N\kappa_1} 2^{-N\left[I(\check{X}_{1p},X_{1e};Y_1,\check{Y}_1,U_2|\check{X}_{1f},\check{X}_{2f},X_{2f})-\delta(\epsilon)\right]} \\
&+ 2^{N\kappa_2} 2^{-N\left[I(\check{X}_{1p},X_{2e};\check{Y}_1,Y_1,U_2|\check{X}_{1f},\check{X}_{2f},X_{1e},X_{2c})-\delta(\epsilon)\right]} \\
&+ 2^{N\kappa_2} 2^{-N\left[I(\check{X}_{1p},\check{X}_{2e};Y_1,\check{Y}_1,U_2|\check{X}_{1f},\check{X}_{2c},X_{1e},X_{2f})-\delta(\epsilon)\right]} \\
&+ 2^{-N\left[I(\check{X}_{1p};Y_1,\check{Y}_1,U_2|\check{X}_{1f},\check{X}_{2f},X_{1e},X_{2e})-\delta(\epsilon)\right]} \\
\overset{\text{(f)}}{\leq}\ & 4 \cdot 2^{-N\left[I(\check{X}_1;\check{Y}_1,U_2|\check{X}_{1f},\check{X}_{2f})-C_1-\delta(\epsilon)\right]} \\
&+ 4 \cdot 2^{-N\left[I(\check{X}_1,X_{2e};\check{Y}_1,Y_1|\check{X}_{1e},\check{X}_{2f},X_{1e},X_{2c})-C_1-\delta(\epsilon)\right]} \\
\overset{\text{(g)}}{=}\ & 4 \cdot 2^{-N\left[I(X_1;Y_1,U_2|X_{1f},X_{2f})-C_1-\delta(\epsilon)\right]} \\
&+ 4 \cdot 2^{-N\left[I(X_1,X_{2e};Y_1|X_{1f},X_{2c})-C_1-\delta(\epsilon)\right]} \\
\overset{\text{(h)}}{\leq}\ & 8 \cdot 2^{-N\left[I(X_1;Y_1|X_{1f},X_{2f})-C_1-\delta(\epsilon)\right]}
\end{aligned}
$$

where

- (a) is since $T(1,n_1,1,q_1,q_2,q_2')$ is empty set for $q_1 \notin \mathcal{B}_1(b)$, $q_2 \notin \mathcal{B}_2(b-1)$, or $q_2' \notin \mathcal{B}_2((q_2,m_2)(b))$, since for random variables $(X,Y,Z) \sim p(x,y,z)$, $(X,Y,Z) \in \mathcal{A}_\epsilon^{(n)}$ implies $(X,Y) \in \mathcal{A}_\epsilon^{(n)}$,

- (b) follows by union bound,

- (c) follows by Claim A.1, where the upper bound on the number of terms is given by

Lemma 2.1,

- (d) is by packing lemma, Claim A.2, and the fact that $X_{1e}^N(b-1)$ is already known at the decoder,
- (e) is by rearranging mutual information terms using chain rule and independence (see Figure A.1),
- (f) follows by upper bounding four of the terms with the first expression, the remaining terms with the second expression, and using the definition of C_1,
- (g) is because the distribution of variables is the same for all blocks,
- (h) is by upper bounding the two terms with the same expression.

Once again, we use the structure of the joint distribution (A.10) to show that a similar bounding can be performed for other error terms as follows.

$$\begin{aligned}
\mathbb{P}\left(T(1,1,m_2)|E^c\right) &\leq 8 \cdot 2^{-N\left[I(X_{2f};Y_1|X_{1e},X_{2e})-C_1-\delta(\epsilon)\right]} \\
\mathbb{P}\left(T(m_1,n_1,1)|E^c\right) &\leq 4 \cdot 2^{-N\left[I(X_1;Y_1,U_2|X_{1e},X_{2f})-C_1-\delta(\epsilon)\right]} \\
&\quad + 4 \cdot 2^{-N[I(X_1,X_{2e};Y_1|X_{1e},X_{2c})-C_1-\delta(\epsilon)]} \\
\mathbb{P}\left(T(1,n_1,m_2)|E^c\right) &\leq 8 \cdot 2^{-N\left[I(\check{X}_1,X_{2f};\check{Y}_1,Y_1,U_2|\check{X}_{1f},\check{X}_{2e},X_{1e},X_{2e})-C_1-\delta(\epsilon)\right]} \\
&= 8 \cdot 2^{-N\left[I(X_1,X_{2f};Y_1,U_2|X_{1f},X_{2e})-C_1-\delta(\epsilon)\right]} \\
&\leq 8 \cdot 2^{-N\left[I(X_1,X_{2f};Y_1|X_{1f},X_{2e})-C_1-\delta(\epsilon)\right]} \\
\mathbb{P}\left(T(m_1,n_1,m_2)|E^c\right) &\leq 8 \cdot 2^{-N\left[I(\check{X}_1,X_{2f};\check{Y}_1,Y_1,U_2|\check{X}_{1e},\check{X}_{2e},X_{1e},X_{2e})-C_1-\delta(\epsilon)\right]} \\
&= 8 \cdot 2^{-N\left[I(X_1,X_{2f};Y_1,U_2|X_{1e},X_{2e})-C_1-\delta(\epsilon)\right]} \\
&\leq 8 \cdot 2^{-N\left[I(X_1,X_{2f};Y_1|X_{1e},X_{2e})-C_1-\delta(\epsilon)\right]}
\end{aligned}$$

Using these bounds in (A.9), we see that if the conditions in the lemma are satisfied, $\mathbb{P}\left(\mathcal{D}_{FB,s}|E^c\right) \to 0$ as $N \to \infty$.

A.2.5 Proof of Lemmas 2.3 and 2.5

Extending the notation defined in the first subsection, we define

$$T(m_1, n_1, m_2, q_2) := \{(2.20) \text{ holds for the indices } (m_1, n_1, m_2, q_2)\}$$

We will show that there exists a code such that $\mathbb{P}\left(\mathcal{D}_{NFB}\right) \to 0$ *exponentially*, if the given rate constraints are satisfied, which implies the claimed result. The decoding error event $\mathcal{D}_{NFB}$ can be expressed as follows.

$$\begin{aligned}\mathcal{D}_{NFB} = &\left(\bigcap_{q_2} T^c(1,1,1,q_2)\right) \cup \bigcup_{n_1 \neq 1} T(1,n_1,1,1) \cup \bigcup_{m_2 \neq 1} T(1,1,m_2,1) \\ &\cup \bigcup_{\substack{m_1 \neq 1 \\ n_1 \neq 1}} T(m_1,n_1,1,1) \cup \bigcup_{\substack{n_1 \neq 1 \\ m_2 \neq 1}} T(1,n_1,m_2,1) \cup \bigcup_{\substack{m_2 \neq 1 \\ q_2 \neq 1}} T(1,1,m_2,q_2) \\ &\cup \bigcup_{\substack{m_1 \neq 1 \\ n_1 \neq 1 \\ m_2 \neq 1}} T(m_1,n_1,m_2,1) \cup \bigcup_{\substack{n_1 \neq 1 \\ m_2 \neq 1 \\ q_2 \neq 1}} T(1,n_1,m_2,q_2) \cup \bigcup_{\substack{m_1 \neq 1, n_1 \neq 1 \\ m_2 \neq 1, q_2 \neq 1}} T(m_1,n_1,m_2,q_2)\end{aligned}$$

Similar to the previous proofs, choosing $r_i > I(\widetilde{V}_j; U_i)$ ensures quantization success with high probability. Then since

$$\mathbb{P}\left(\mathcal{D}_{NFB}\right) \leq \mathbb{P}\left(E_1\right) + \mathbb{P}\left(E_2\right) + \mathbb{P}\left(\mathcal{D}_{NFB} | E^c\right),$$

it is sufficient to show that $\mathbb{P}\left(\mathcal{D}_{NFB}|E^c\right) \to 0$. Using union bound, packing lemma, Lemma 2.1, and Claim A.1, we can upper bound the probability of decoding error conditioned on quantization success by

$$\begin{aligned}\mathbb{P}\left(\mathcal{D}_{NFB}|E^c\right) \leq\ &\epsilon_N + 2^{NR_{1p}} 2^{-N\left[I(X_1;Y_1|X_{1f},X_{2f})-\delta(\epsilon)\right]} + 2^{NR_{2c}} 2^{-N\left[I(X_{2f};Y_1|X_1,X_{2e})-\delta(\epsilon)\right]} \\ &+ 2^{NR_1} 2^{-N\left[I(X_1;Y_1|X_{1e},X_{2f})-\delta(\epsilon)\right]} + 2^{N(R_{1p}+R_{2c})} 2^{-N\left[I(X_1,X_{2f};Y_1|X_{1f},X_{2e})-\delta(\epsilon)\right]} \\ &+ 2^{N(R_{2c}+C_2')} 2^{-N\left[I(X_{2f};Y_1|X_1)-\delta(\epsilon)\right]} + 2^{N(R_1+R_{2c})} 2^{-N\left[I(X_1,X_{2f};Y_1|X_{1e},X_{2e})-\delta(\epsilon)\right]} \\ &+ 2^{N(R_{1p}+R_{2c}+C_2')} 2^{-N\left[I(X_1,X_{2f};Y_1|X_{1f})-\delta(\epsilon)\right]} + 2^{N(R_1+R_{2c}+C_2')} 2^{-N\left[I(X_1,X_{2f};Y_1|X_{1e})-\delta(\epsilon)\right]}.\end{aligned}$$

where $\epsilon_N \to 0$ as $N \to \infty$. Note that the conditions in both lemmas are sufficient to ensure $\mathbb{P}\left(\mathcal{D}_{NFB}|E^c\right) \to 0$ as $N \to \infty$.

A.3 Evaluation of Rate Regions

In this section, we consider the set of rate conditions derived in Section 2.6 for decodability (*i.e.*, (2.24)–(2.27), (2.29)–(2.32) for weak interference; (2.34)–(2.37), (2.39)–(2.41) for strong interference), and obtain an explicit rate region for both linear deterministic and Gaussian models.

A.3.1 Rate region for linear deterministic model

Under the input distribution given by (2.42)–(2.46), the set of rate constraints for decodability at Rx1 are evaluated as follows:

$$\begin{aligned}
R_{1p} &\leq H\left(Y_1|V_1,V_2\right) = \left(n_{11}-n_{21}\right)^+ \\
R_{2c} &\leq H\left(Y_1|X_1\right) = n_{12} \\
R_{1p}+R_{2c} &\leq \min\left\{H\left(Y_1,\widetilde{V}_2|V_1\right), H\left(Y_1\right)\right\} \\
&= \min\left\{p_1\left(n_{11}-n_{21}\right)^+ + (1-p_1)\max\left\{n_{12},\left(n_{11}-n_{21}\right)^+\right\} + p_1 n_{12},\right. \\
&\quad \left.\max\left(n_{11},n_{12}\right)\right\} \\
R_1+R_{2c} &\leq H\left(Y_1\right) = \max\left(n_{11},n_{12}\right)
\end{aligned}$$

for weak interference ($n_{12} \leq n_{11}$), and

$$\begin{aligned}
R_{1p} &\leq H\left(Y_1|V_1,V_2\right) = \left(n_{11}-n_{21}\right)^+ \\
R_{2c} &\leq H\left(Y_1|X_1\right) = n_{12} \\
R_1 &\leq \min\left\{H\left(Y_1,\widetilde{V}_1|V_2\right), H\left(Y_1\right)\right\} \\
&= \min\left\{p_2\left(n_{11}-n_{21}\right)^+ + (1-p_2)n_{11} + p_2 n_{21}, \max\left(n_{11},n_{12}\right)\right\} \\
R_1+R_{2c} &\leq H\left(Y_1\right) = \max\left(n_{11},n_{12}\right)
\end{aligned}$$

for strong interference ($n_{12} > n_{11}$). Note that the set of conditions given above can be summarized into the following five inequalities, valid for any interference regime.

$$R_{1p} \leq \left(n_{11}-n_{21}\right)^+$$

$$R_{ip} < \mathsf{A}_i := \log\left(3 + \frac{\mathsf{SNR}_i}{1+\mathsf{INR}_j}\right) - \log 3 - C_i \tag{A.11}$$

$$R_{jc} < \mathsf{B}_i := \log\left(2 + \mathsf{INR}_i\right) - \log 3 - C_i \tag{A.12}$$

$$R_i < \mathsf{C}_i := \log\left(3 + \mathsf{SNR}_i + \mathsf{INR}_i\right) - \log 3 - C_i \tag{A.13}$$

$$\begin{aligned} R_i < \mathsf{D}_i &:= \log\left(3 + \mathsf{SNR}_i\right) \\ &\quad + \mathbb{1}_{\{\mathsf{SNR}_i \leq \mathsf{INR}_i\}} p_j \left[\log\left(1 + \frac{\mathsf{INR}_j}{3+\mathsf{SNR}_i}\right) - \log\frac{5}{3}\right] - \log 3 - C_i \end{aligned} \tag{A.14}$$

$$R_{ip} + R_{jc} < \mathsf{E}_i := \log\left(2 + \mathsf{SNR}_i + \mathsf{INR}_i + \frac{\mathsf{SNR}_i}{1+\mathsf{INR}_j}\right) - \log 3 - C_i \tag{A.15}$$

$$\begin{aligned} R_{ip} + R_{jc} < \mathsf{F}_i &:= \log\left(2 + \mathsf{INR}_i + \frac{\mathsf{SNR}_i}{1+\mathsf{INR}_j}\right) \\ &\quad + \mathbb{1}_{\{\mathsf{SNR}_i \geq \mathsf{INR}_i\}} p_i \left[\log\left(\frac{\left(2+\mathsf{INR}_i\right)\left(3 + \frac{\mathsf{SNR}_i}{1+\mathsf{INR}_j}\right)}{2 + \frac{\mathsf{SNR}_i}{1+\mathsf{INR}_j} + \mathsf{INR}_i}\right) - \log 6\right] \\ &\quad - \log 3 - C_i - \mathbb{1}_{\{\mathsf{SNR}_i \geq \mathsf{INR}_i\}} C_i \end{aligned} \tag{A.16}$$

$$R_i + R_{jc} < \mathsf{G}_i := \log\left(2 + \mathsf{SNR}_i + \mathsf{INR}_i\right) - \log 3 - C_i - \kappa_j \tag{A.17}$$

$$C_i := p_i + 2p_j,\ \kappa_j = p_j$$

$$\begin{aligned} R_1 &\leq n_{11} + p_2\left(n_{21} - n_{11}\right)^+ \\ R_{2c} &\leq n_{12} \\ R_{1p} + R_{2c} &\leq \max\left\{n_{12}, \left(n_{11} - n_{21}\right)^+\right\} + p_1 \min\left\{n_{12}, \left(n_{11} - n_{21}\right)^+\right\} \\ R_1 + R_{2c} &\leq \max\left(n_{11}, n_{12}\right) \end{aligned}$$

Combining these inequalities with their Rx2 counterparts, and applying Fourier-Motzkin elimination, we arrive at the set of inequalities given in (2.5)–(2.11).

A.3.2 Rate region for Gaussian model

We consider the input distributions (2.47)–(2.51), and the set of input-output relationships given by

$$Y_i = h_{ii}X_i + h_{ij}X_j + Z_i$$
$$U_i = S_i\left(h_{ij}X_j + Z_i\right) + Q_i$$

for $(i,j) = (1,2),(2,1)$, where $Q_i \sim \mathcal{CN}(0, D_i)$. Choosing $D_1 = D_2 = \frac{3}{2}$, and using standard techniques, it is straightforward to evaluate the rate inequalities derived in Section 2.6, and show that the set of rate triples (R_{1p}, R_{1c}, R_{2c}) defined by (A.11)–(A.17), for $(i,j) = (1,2)$ are contained in the set defined by (2.24)–(2.27), (2.29)–(2.32) for weak interference, and (2.34)–(2.37), (2.39)–(2.41) for strong interference. In (A.11)–(A.17), we used indicator functions to unify the constraints for weak and strong interference.

In order to find the set of achievable (R_1, R_2) points, we first note that $\mathsf{E}_i \geq \mathsf{G}_i$ and $\mathsf{C}_i \geq \mathsf{G}_i$, and hence the bounds E_i and C_i are redundant. Considering the remaining bounds for $(i,j) = (1,2),(2,1)$, noting that $\mathsf{F}_i \leq \mathsf{A}_i + \mathsf{B}_i$, and applying Fourier-Motzkin elimination, we find that the set of (R_1, R_2) points that satisfy the following are achievable.

$$R_i < \min\left\{\mathsf{A}_i + \mathsf{B}_j, \mathsf{D}_i\right\} \tag{A.18}$$

$$R_i + R_j < \min\left\{\mathsf{A}_i + \mathsf{G}_j, \mathsf{F}_i + \mathsf{F}_j\right\} \tag{A.19}$$

$$2R_i + R_j < \mathsf{A}_i + \mathsf{F}_j + \mathsf{G}_i \tag{A.20}$$

for $(i,j) = (1,2),(2,1)$.

A.4 Proofs of Outer Bounds (2.54), (2.58), (2.59), and (2.60)

In this section, we prove the outer bounds for the linear deterministic channel, based on the ideas presented in Section 2.7. We first prove four claims that will be useful in the main proof.

A.4.1 Proof of the bound (2.54)

By symmetry, we only focus on the bound on R_1. By Fano's inequality,

$$\begin{aligned}
N\left(R_1-\epsilon_N\right) &\leq I\left(W_1;Y_1^N\underline{S}^N\right) = I\left(W_1;Y_1^N,\widetilde{V}_1^N,W_2,\underline{S}^N\right)\\
&= I\left(W_1;Y_1^N,\widetilde{V}_1^N|W_2,\underline{S}^N\right) \overset{\text{(a)}}{=} H\left(Y_1^N,\widetilde{V}_1^N|W_2,\underline{S}^N\right)\\
&= H\left(Y_1^N|\widetilde{V}_1^N,W_2,\underline{S}^N\right)+H\left(\widetilde{V}_1^N|W_2,\underline{S}^N\right)\\
&\overset{\text{(b)}}{=} H\left(Y_1^N|\widetilde{V}_1^N,W_2,X_2^N,\underline{S}^N\right)+H\left(\widetilde{V}_1^N|W_2,\underline{S}^N\right)\\
&\leq H\left(Y_1^N|\widetilde{V}_1^N,X_2^N,\underline{S}^N\right)+H\left(\widetilde{V}_1^N|\underline{S}^N\right)\\
&\overset{\text{(c)}}{\leq} n_{11}+p_2\left(n_{21}-n_{11}\right)^+
\end{aligned}$$

where (a) follows by the fact that channel is deterministic and hence all variables are completely determined given $\left(W_1,W_2,\underline{S}^N\right)$; (b) follows by Claim A.3, and (c) follows by Claim A.7.

A.4.2 Proof of the bound (2.58)

By Fano's inequality,

$$\begin{aligned}
N\left(R_1+R_2-\epsilon_N\right) &\leq I(W_1;Y_1^N,\underline{S}^N)+I(W_2;Y_2^N,\underline{S}^N)\\
&= I(W_1;Y_1^N|\underline{S}^N)+I(W_2;Y_2^N|\underline{S}^N)\\
&\leq I(W_1;Y_1^N,V_1^N,\widetilde{V}_2^N|\underline{S}^N)+I(W_2;Y_2^N,V_2^N,\widetilde{V}_1^N|\underline{S}^N)\\
&= H\left(Y_1^N,V_1^N,\widetilde{V}_2^N|\underline{S}^N\right)+H\left(Y_2^N,V_2^N,\widetilde{V}_1^N|\underline{S}^N\right)\\
&\quad -H\left(Y_1^N,V_1^N,\widetilde{V}_2^N|W_1,\underline{S}^N\right)-H\left(Y_2^N,V_2^N,\widetilde{V}_1^N|W_2,\underline{S}^N\right)\\
&\overset{\text{(a)}}{=} H\left(Y_1^N|V_1^N,\widetilde{V}_2^N,\underline{S}^N\right)+H\left(Y_2^N|V_2^N,\widetilde{V}_1^N,\underline{S}^N\right)\\
&\quad +H\left(V_1^N,\widetilde{V}_2^N|\underline{S}^N\right)+H\left(V_2^N,\widetilde{V}_1^N|\underline{S}^N\right)\\
&\quad -H\left(V_2^N,\widetilde{V}_1^N|W_1,\underline{S}^N\right)-H\left(V_1^N,\widetilde{V}_2^N|W_2,\underline{S}^N\right)\\
&= H\left(Y_1^N|V_1^N,\widetilde{V}_2^N,\underline{S}^N\right)+H\left(Y_2^N|V_2^N,\widetilde{V}_1^N,\underline{S}^N\right)
\end{aligned}$$

$$\begin{aligned}
&+ I\left(W_2; V_1^N, \widetilde{V}_2^N | \underline{S}^N\right) + I\left(W_1; V_2^N, \widetilde{V}_1^N | \underline{S}^N\right) \\
&\overset{\text{(b)}}{\leq} N \max\left\{n_{12}, (n_{11} - n_{21})^+\right\} + N \max\left\{n_{21}, (n_{22} - n_{12})^+\right\} \\
&\quad + N p_1 \min\left\{n_{12}, (n_{11} - n_{21})^+\right\} + N p_2 \min\left\{n_{21}, (n_{22} - n_{12})^+\right\}
\end{aligned}$$

where (a) follows by Claim A.4, and (b) follows by Claims A.6 and A.7.

A.4.3 Proof of the bounds (2.59) and (2.60)

By symmetry, it is sufficient to prove (2.59). To prove this bound, we consider two copies of Rx1, where one of the copies are enhanced as decribed in Section 2.7, while the other one is provided with the output of the original channel. The only copy of Rx2 receives the enhanced channel output as well. We would like to prove a sum rate bound for this three-receiver channel. By Fano's inequality,

$$\begin{aligned}
&N\left(2R_1 + R_2 - \epsilon_N\right) \\
&\leq I\left(W_1; Y_1^N, \underline{S}^N\right) + I\left(W_2; Y_2^N, \underline{S}^N\right) + I\left(W_1; Y_1^N, \underline{S}^N\right) \\
&= I\left(W_1; Y_1^N | \underline{S}^N\right) + I\left(W_2; Y_2^N | \underline{S}^N\right) + I\left(W_1; Y_1^N | \underline{S}^N\right) \\
&\leq I\left(W_1; Y_1^N | \underline{S}^N\right) + I\left(W_2; Y_2^N, V_2^N, \widetilde{V}_1^N | \underline{S}^N\right) + I\left(W_1; Y_1^N, V_1^N | \underline{S}^N, W_2\right) \\
&\overset{\text{(a)}}{=} H\left(Y_1^N | \underline{S}^N\right) - H\left(V_2^N, \widetilde{V}_1^N | \underline{S}^N, W_1\right) + H\left(Y_2^N, V_2^N, \widetilde{V}_1^N | \underline{S}^N\right) \\
&\quad - H\left(Y_2^N, V_2^N, \widetilde{V}_1^N | \underline{S}^N, W_2\right) \\
&\quad + H\left(Y_1^N, V_1^N | \underline{S}^N, W_2\right) \\
&\overset{\text{(b)}}{=} H\left(Y_1^N | \underline{S}^N\right) - H\left(V_2^N, \widetilde{V}_1^N | \underline{S}^N, W_1\right) + H\left(V_2^N, \widetilde{V}_1^N | \underline{S}^N\right) \\
&\quad + H\left(Y_2^N | V_2^N, \widetilde{V}_1^N\right) - H\left(V_1^N | \underline{S}^N, W_2\right) + H\left(Y_1^N, V_1^N | \underline{S}^N, W_2\right) \\
&= H\left(Y_1^N | \underline{S}^N\right) + I\left(W_1; V_2^N, \widetilde{V}_1^N | \underline{S}^N\right) \\
&\quad H\left(Y_2^N | V_2^N, \widetilde{V}_1^N\right) + H\left(Y_1^N | \underline{S}^N, W_2, V_1^N\right) \\
&\overset{\text{(c)}}{\leq} \max\left(n_{11}, n_{12}\right) + \max\left\{n_{21}, (n_{22} - n_{12})^+\right\} \\
&\quad + (n_{11} - n_{21})^+ + p_2 \min\left\{n_{21}, (n_{22} - n_{12})^+\right\}
\end{aligned}$$

where (a) follows by Claim A.4 , (b) follows by Claim A.5, (c) follows by Claims A.6, A.7 and A.8.

A.4.4 Claims

Claim A.3. *For* $(i,j) = (1,2),(2,1)$,

$$X_{i,t} \overset{\mathrm{f}}{=} \left(W_i, \widetilde{V}_j^{t-1}, \underline{S}^{t-1}\right) \overset{\mathrm{f}}{=} \left(W_i, V_j^{t-1}, \underline{S}^{t-1}\right)$$

Proof. We focus on the case $(i,j) = (1,2)$ without loss of generality. Note that

$$X_{1,1} \overset{\mathrm{f}}{=} W_1$$

and by the definition of the channel,

$$X_{1,t} \overset{\mathrm{f}}{=} \left(W_1, \widetilde{Y}_1^{t-1}, \underline{S}^{t}\right) \overset{(\mathrm{a})}{\overset{\mathrm{f}}{=}} \left(W_1, \widetilde{V}_2^{t-1}, X_1^{t-1}\underline{S}^{t}\right),$$

hence the result follows by induction on t. (a) follows because

$$\widetilde{Y}_1^{t-1} = S_1^{t-1}\mathbf{H}_{11}X_1^{t-1} + \widetilde{V}_2^{t-1}.$$

□

Claim A.4. *For* $(i,j) = (1,2),(2,1)$,

$$H\left(Y_i^N | W_i, \underline{S}^N\right) = H\left(V_j^N, \widetilde{V}_i^N | W_i, \underline{S}^N\right).$$

Proof. Let us focus on the case $(i,j) = (1,2)$.

$$\begin{aligned} H\left(Y_1^N | W_1, \underline{S}^N\right) &= \sum_{t=1}^{N} H\left(Y_{1,t} | W_1, \underline{S}^N, Y_1^{t-1}\right) \\ &\overset{(\mathrm{a})}{=} \sum_{t=1}^{N} H\left(Y_{1,t} | W_1, \underline{S}^N, Y_1^{t-1}, X_1^t\right) \\ &= \sum_{t=1}^{N} H\left(V_{2,t} | W_1, \underline{S}^N, V_2^{t-1}, X_1^t\right) \end{aligned}$$

$$\overset{\text{(b)}}{=}\sum_{t=1}^{N} H\left(V_{2,t}|W_1,\underline{S}^N,V_2^{t-1}\right)$$
$$= H\left(V_2^N,\widetilde{V}_1^N|W_1,\underline{S}^N\right),$$

where (a) is by definition, (b) is due to Claim A.3. The other holds similarly. □

Claim A.5. *For $(i,j)=(1,2),(2,1)$,*

$$H\left(Y_i^N,V_i^N,\widetilde{V}_j^N|W_i,\underline{S}^N\right)=H\left(V_j^N,\widetilde{V}_i^N|W_i,\underline{S}^N\right)=H\left(V_j^N|W_i,\underline{S}^N\right)$$

Proof. Let us focus on the case $(i,j)=(1,2)$.

$$\begin{aligned}
&H\left(Y_1^N,V_1^N,\widetilde{V}_2^N|W_1,\underline{S}^N\right)\\
&=\sum_{t=1}^{N} H\left(Y_{1,t},V_{1,t},\widetilde{V}_{2,t}|W_1,\underline{S}^N,Y_1^{t-1},V_1^{t-1},\widetilde{V}_2^{t-1}\right)\\
&\overset{\text{(a)}}{=}\sum_{t=1}^{N} H\left(Y_{1,t},V_{1,t},\widetilde{V}_{2,t}|W_1,\underline{S}^N,Y_1^{t-1},V_1^{t-1},\widetilde{V}_2^{t-1},X_1^t\right)\\
&\overset{\text{(b)}}{=}\sum_{t=1}^{N} H\left(V_{2,t},\widetilde{V}_{1,t}|W_1,\underline{S}^N,V_2^{t-1},\widetilde{V}_1^{t-1},X_1^t\right)\\
&\overset{\text{(c)}}{=}\sum_{t=1}^{N} H\left(V_{2,t}|W_1,\underline{S}^N,V_2^{t-1},X_1^t\right)
\end{aligned}$$

where (a) follows by the fact that $X_{1,t}\overset{\text{f}}{=}\left(W_1,\widetilde{Y}_1^N,\underline{S}^N\right)$, (b) follows by subtracting $X_{1,t}$ from $Y_{1,t}$ and because $V_{1,t}\overset{\text{f}}{=}X_{1,t}$ and $\widetilde{V}_{2,t}\overset{\text{f}}{=}(\underline{S}_t,V_{2,t})$. Similarly, (c) follows since $\widetilde{V}_{1,t}\overset{\text{f}}{=}(X_{1,t},\underline{S}_t)$.

Now, the two equalities in the claim can be easily obtained from (b) and (c) respectively, by removing X_1^t from the conditioning by virtue of Claim A.3, and using chain rule. □

Claim A.6. *For $(i,j)=(1,2),(2,1)$,*

$$I\left(W_i;V_j^N,\widetilde{V}_i^N|\underline{S}^N\right)\leq Np_jn_{ji}.$$

Proof. Let us focus on the case $(i,j)=(1,2)$.

$$I\left(W_1;V_2^N,\widetilde{V}_1^N|\underline{S}^N\right)\overset{\text{(a)}}{\leq} I\left(W_1;W_2,\widetilde{V}_1^N|\underline{S}^N\right)=I\left(W_1;\widetilde{V}_1^N|\underline{S}^N,W_2\right)$$

$$
\begin{aligned}
&= H\left(\widetilde{V}_1^N|\underline{S}^N, W_2\right) \le H\left(\widetilde{V}_1^N|S_2^N\right) \\
&= \mathbb{E}_{S_2^N}\left[H\left((s_2V_1)^N\right)|S_2^N = s_2^N\right] \\
&\le \mathbb{E}_{S_2^N}\left[\sum_{t=1}^N H\left(s_{2,t}V_{1,t}\right)\middle|S_2^N = s_2^N\right] \\
&\le \mathbb{E}_{S_2^N}\left[N_1\left(s_2^N\right)n_{21}\middle|S_2^N = s_2^N\right] = Np_2n_{21}.
\end{aligned}
$$

Here $N_1(\cdot)$ denotes the number of 1's in the sequence. (a) follows because

$$
V_2^N \overset{\mathrm{f}}{=} \left(W_2, \widetilde{V}_1^N, \underline{S}^N\right).
$$

□

Claim A.7. *For* $(i,j) = (1,2), (2,1)$,

$$
\begin{aligned}
N^{-1}H\left(Y_i^N|V_i^N, \widetilde{V}_j^N, \underline{S}^N\right) &\le p_i(n_{ii} - n_{ji})^+ + (1-p_i)\max\left\{n_{ij}, (n_{ii}-n_{ji})^+\right\}, \\
N^{-1}H\left(Y_i^N|V_j^N, \widetilde{V}_i^N, \underline{S}^N\right) &\le p_j(n_{ii} - n_{ji})^+ + (1-p_j)n_{ii},
\end{aligned}
$$

Proof. Let us focus on the case $(i,j) = (1,2)$.

$$
\begin{aligned}
&H\left(Y_1^N|V_1^N, \widetilde{V}_2^N, \underline{S}^N\right) \le H\left(Y_1^N|V_1^N, \widetilde{V}_2^N, S_1^N\right) \\
&= \mathbb{E}_{S_1^N}\left[H\left(Y_1^N|V_1^N, (s_1V_2)^N\right)|S_1^N = s_1^N\right] \\
&\le \mathbb{E}_{S_1^N}\left[\sum_{t=1}^N H\left(Y_{1,t}|V_{1,t}, s_{1,t}V_{2,t}\right)\middle|S_1^N = s_1^N\right] \\
&\le \mathbb{E}_{S_1^N}\left[\begin{array}{c} N_1\left(s_1^N\right)(n_{11}-n_{21})^+ \\ +N_0\left(s_1^N\right)\max\left\{n_{12}, (n_{11}-n_{21})^+\right\}\end{array}\middle|S_1^N = s_1^N\right] \\
&= Np_1(n_{11}-n_{21})^+ + N(1-p_1)\max\left\{n_{12}, (n_{11}-n_{21})^+\right\}
\end{aligned}
$$

Here $N_1(\cdot)$ and $N_0(\cdot)$ denote the number of 1's and 0's respectively in the sequence.

For the second inequality,

$$
H\left(Y_1^N|V_2^N, \widetilde{V}_1^N, \underline{S}^N\right) \le H\left(Y_1^N|V_2^N, \widetilde{V}_1^N, S_2^N\right)
$$

$$
\begin{aligned}
&= \mathbb{E}_{S_2^N}\left[H\left(Y_1^N|V_2^N,(s_2V_1)^N\right)|S_2^N=s_2^N\right] \\
&\leq \mathbb{E}_{S_2^N}\left[\sum_{t=1}^{N} H\left(Y_{1,t}|V_{2,t},s_{2,t}V_{1,t}\right)\middle|S_2^N=s_2^N\right] \\
&\leq \mathbb{E}_{S_2^N}\left[N_1\left(s_2^N\right)(n_{11}-n_{21})^+ + N_0\left(s_2^N\right)n_{11}\middle|S_2^N=s_2^N\right] \\
&= Np_2(n_{11}-n_{21})^+ + N(1-p_2)n_{11}.
\end{aligned}
$$

The case $(i,j)=(2,1)$ follows similarly. □

Claim A.8. *For* $(i,j)=(1,2),(2,1)$,

$$H\left(Y_i^N|\underline{S}^N,W_j,V_i^N\right) \leq N\left(n_{ii}-n_{ji}\right)^+$$

Proof. We focus on $(i,j)=(1,2)$ without loss of generality.

$$
\begin{aligned}
H\left(Y_i^N|\underline{S}^N,W_j,V_i^N\right) &\overset{\text{(a)}}{=} H\left(Y_i^N|\underline{S}^N,W_j,V_i^N,V_j^N\right) \\
&\leq H\left(Y_i^N|\underline{S}^N,V_i^N,V_j^N\right) \leq N\left(n_{11}-n_{21}\right)^+
\end{aligned}
$$

where (a) follows because $V_j^N \overset{\text{f}}{=} X_j^N \overset{\text{f}}{=} \left(W_j,V_i^N,\underline{S}^N\right)$ by Claim A.3. □

A.5 Proofs of Outer Bounds (2.55), (2.61), (2.62), and (2.63)

In this section, we prove an outer bound region for the enhanced channel defined in Section 2.7.

A.5.1 Notation

We define

$$\breve{V}_i = \begin{cases} \bar{V}_i, & \text{if } S_j=0 \\ Y_{ji}, & \text{if } S_j=1 \end{cases}$$

for $(i,j)=(1,2),(2,1)$, where $\bar{V}_i = Y_{ji}+Z_{jj} = h_{ji}X_i+\bar{Z}_j$, and

$$M_i = \left|\{t : S_{i,t}=1\}\right|,$$

$$L_i = N - M_i.$$

For any random vector E^N, we define

$$\begin{aligned}
E^{(t)} &= \{E_{t'}\}_{t':S_{i,t'}=1,t'\leq t} \\
E^{[t]} &= \{E_{t'}\}_{t':S_{i,t'}=0,t'\leq t} \\
E_{i,(t)} &= \begin{cases} \emptyset, & \text{if } S_i = 0 \\ E_{i,t}, & \text{if } S_i = 1 \end{cases} \\
E_{i,[t]} &= \begin{cases} E_{i,t}, & \text{if } S_i = 0 \\ \emptyset, & \text{if } S_i = 1 \end{cases}
\end{aligned}$$

for $i = 1$ or 2. Note that in vector form, this notation omits any reference to user index i for the sake of brevity. That is, although it is not clear whether $E^{(t)}$ is defined with respect to S_1 or S_2, in the proof this will be clear from the context. For instance, $Y_1^{(t)}$ and $V_2^{(t)}$ are defined with respect to S_1, since these variables refer to signals that pass through the feedback channel controlled by S_1. The partial average power for Txi, $P_i^{(jk)}$, is a random variable defined as

$$\begin{aligned}
P_i^{(j0)} &= \frac{1}{L_i} \sum_{t:S_{j,t=0}} P_{i,t} \\
P_i^{(j1)} &= \frac{1}{M_i} \sum_{t:S_{j,t=1}} P_{i,t}
\end{aligned}$$

for $j = 1, 2$, where $P_{i,t}$ ie the power used by Txi at time slot t.

Finally, we define $h_S(\cdot) := h\left(\cdot|\underline{S}^N = S^N\right)$ for convenience, where $h(\cdot)$ denotes differential entropy, and S^N is a particular realization of $\underline{S}^N$. Similarly, we define $I_S(\cdot;\cdot) := I_S\left(\cdot;\cdot|\underline{S}^N = S^N\right)$.

A.5.2 Proof of the Bound (2.55)

We focus on the case $(i, j) = (1, 2)$. By Fano's inequality,

$$N(R_1 - \epsilon_N) \leq I(W_1; Y_1^N, \underline{S}^N) \leq I(W_1; Y_1^N, \widetilde{V}_1^N, W_2, \underline{S}^N)$$

$$
\begin{aligned}
&\le I(W_1; Y_1^N, \widetilde{V}_1^N, \underline{S}^N | W_2) \\
&\le \sum_{t=1}^{N} I(W_1; Y_{1,t}, \widetilde{V}_{1,t}, \underline{S}_t | W_2, Y_1^{t-1}, \widetilde{V}_1^{t-1}, \underline{S}^{t-1}) \\
&= \sum_{t=1}^{N} I(W_1; Y_{1,t}, \widetilde{V}_{1,t} | W_2, Y_1^{t-1}, \widetilde{V}_1^{t-1}, \underline{S}^{t}) \\
&\overset{(a)}{=} \sum_{t=1}^{N} I(W_1; Y_{1,t}, \widetilde{V}_{1,t} | W_2, Y_1^{t-1}, \widetilde{V}_1^{t-1}, \underline{S}^{t}, X_{2,t}) \\
&= \sum_{t=1}^{N} I(W_1; Y_{1,t} | W_2, Y_1^{t-1}, \widetilde{V}_1^{t}, \underline{S}^{t}, X_{2,t}) + I(W_1; \widetilde{V}_{1,t} | W_2, Y_1^{t-1}, \widetilde{V}_1^{t-1}, \underline{S}^{t}, X_{2,t}) \\
&= \sum_{t=1}^{N} h(Y_{1,t} | W_2, Y_1^{t-1}, \widetilde{V}_1^{t}, \underline{S}^{t}, X_{2,t}) - h(Y_{1,t} | W_2, Y_1^{t-1}, \widetilde{V}_1^{t}, \underline{S}^{t}, X_{2,t}, W_1) \\
&\quad + I(W_1; \widetilde{V}_{1,t} | W_2, Y_1^{t-1}, \widetilde{V}_1^{t-1}, \underline{S}^{t}, X_{2,t}) \\
&\overset{(b)}{=} \sum_{t=1}^{N} h(Y_{1,t} | W_2, Y_1^{t-1}, \widetilde{V}_1^{t}, \underline{S}^{t}, X_{2,t}) - h(Y_{1,t} | W_2, Y_1^{t-1}, \widetilde{V}_1^{t}, \underline{S}^{t}, X_{2,t}, W_1, X_{1,t}) \\
&\quad + \sum_{t=1}^{N} I(W_1; \widetilde{V}_{1,t} | W_2, Y_1^{t-1}, \widetilde{V}_1^{t-1}, \underline{S}^{t}, X_{2,t}) \\
&= \sum_{t=1}^{N} h(Y_{1,t} | W_2, Y_1^{t-1}, \widetilde{V}_1^{t}, \underline{S}^{t}, X_{2,t}) - h(Z_{1,t} | W_2, Y_1^{t-1}, \widetilde{V}_1^{t}, \underline{S}^{t}, X_{2,t}, W_1, X_{1,t}) \\
&\quad + I(W_1; \widetilde{V}_{1,t} | W_2, Y_1^{t-1}, \widetilde{V}_1^{t-1}, \underline{S}^{t}, X_{2,t}) \\
&\overset{(c)}{=} \sum_{t=1}^{N} h(Y_{1,t} | W_2, Y_1^{t-1}, \widetilde{V}_1^{t}, \underline{S}^{t}, X_{2,t}) - h(Z_{1,t}) + I(W_1; \widetilde{V}_{1,t} | W_2, Y_1^{t-1}, \widetilde{V}_1^{t-1}, \underline{S}^{t}, X_{2,t}) \\
&\le \sum_{t=1}^{N} h(Y_{1,t} | \widetilde{V}_{1,t}, S_{2,t}, X_{2,t}) - h(Z_{1,t}) + I(W_1; \widetilde{V}_{1,t} | W_2, Y_1^{t-1}, \widetilde{V}_1^{t-1}, \underline{S}^{t}, X_{2,t}) \\
&\le \sum_{t=1}^{N} h(Y_{1,t} | \widetilde{V}_{1,t}, S_{2,t}, X_{2,t}) - h(Z_{1,t}) + I(W_1, X_{1,t}; \widetilde{V}_{1,t} | W_2, Y_1^{t-1}, \widetilde{V}_1^{t-1}, \underline{S}^{t}, X_{2,t}) \\
&= \sum_{t=1}^{N} h(Y_{1,t} | \widetilde{V}_{1,t}, S_{2,t}, X_{2,t}) - h(Z_{1,t}) + I(X_{1,t}; \widetilde{V}_{1,t} | W_2, Y_1^{t-1}, \widetilde{V}_1^{t-1}, \underline{S}^{t}, X_{2,t}) \\
&\quad + I(W_1; \widetilde{V}_{1,t} | W_2, Y_1^{t-1}, \widetilde{V}_1^{t-1}, \underline{S}^{t}, X_{2,t}, X_{1,t}) \\
&\overset{(d)}{=} \sum_{t=1}^{N} h(Y_{1,t} | \widetilde{V}_{1,t}, S_{2,t}, X_{2,t}) - h(Z_{1,t}) + I(X_{1,t}; \widetilde{V}_{1,t} | W_2, Y_1^{t-1}, \widetilde{V}_1^{t-1}, \underline{S}^{t}, X_{2,t})
\end{aligned}
$$

$$
\begin{aligned}
= \quad & \sum_{t=1}^{N} h(Y_{1,t}|\widetilde{V}_{1,t}, S_{2,t}, X_{2,t}) - h(Z_{1,t}) + h(\widetilde{V}_{1,t}|W_2, Y_1^{t-1}, \widetilde{V}_1^{t-1}, \underline{S}^t, X_{2,t}) \\
& - h(\widetilde{V}_{1,t}|W_2, Y_1^{t-1}, \widetilde{V}_1^{t-1}, \underline{S}^t, X_{2,t}, X_{1,t}) \\
& \leq \sum_{t=1}^{N} h(Y_{1,t}|\widetilde{V}_{1,t}, S_{2,t}, X_{2,t}) - h(Z_{1,t}) + h(\widetilde{V}_{1,t}|S_{2,t}) \\
& - h(\widetilde{V}_{1,t}|W_2, Y_1^{t-1}, \widetilde{V}_1^{t-1}, \underline{S}^t, X_{2,t}, X_{1,t}) \\
& \overset{\text{(e)}}{=} \sum_{t=1}^{N} h(Y_{1,t}|\widetilde{V}_{1,t}, S_{2,t}, X_{2,t}) - h(Z_{1,t}) + h(\widetilde{V}_{1,t}|\underline{S}_t) \\
& - h(\widetilde{V}_{1,t}|S_{2,t}, X_{1,t}) \\
& = \sum_{t=1}^{N} h(Y_{1,t}|\widetilde{V}_{1,t}, S_{2,t}, X_{2,t}) - h(Z_{1,t}) + I(X_{1,t}; \widetilde{V}_{1,t}|S_{2,t}) \\
& \overset{\text{(f)}}{=} p_2 \log\left(1 + \frac{\mathsf{SNR}_1}{1 + \mathsf{INR}_2}\right) + (1 - p_2)\log\left(1 + \mathsf{SNR}_1\right) + p_2 \log\left(1 + \mathsf{INR}_2\right) \\
& = \log\left(1 + \mathsf{SNR}_1\right) + p_2 \log\left(1 + \frac{\mathsf{INR}_2}{1 + \mathsf{SNR}_1}\right)
\end{aligned}
$$

where

- (a) is due to Lemma A.1
- (b) is because $X_{1,t} \overset{\text{f}}{=} \left(\underline{S}^{t-1}, W_1, \widetilde{Y}_1^{t-1}\right) \overset{\text{f}}{=} \left(\underline{S}^{t-1}, W_1, Y_1^{t-1}\right)$,
- (c) is because $Z_{1,t}$ is independent from all past signals and messages,
- (d) is because $W_1 - X_{1,t} - \widetilde{V}_{1,t}$ is a Markov chain, hence conditioned on $X_{1,t}$, $\widetilde{V}_{1,t}$ is independent from W_1 and all the other past signals,
- (e) is because given $(S_{2,t}, X_{1,t})$, $\widetilde{V}_{1,t}$ is independent from all the other variables in the conditioning,
- (f) is due to Lemma A.2

A.5.3 Proof of Bound (2.61)

In this section, we exclusively focus on the enhanced channel defined in Section 2.7. By Fano's inequality.

$$\begin{aligned}
N(R_1 + R_2 - \epsilon_N) &\leq I(W_1; \breve{Y}_1^N, \underline{S}^N) + I(W_2; \breve{Y}_2^N, \underline{S}^N) \\
&= I(W_1; \breve{Y}_1^N | \underline{S}^N) + I(W_2; \breve{Y}_2^N | \underline{S}^N) \\
&\leq I(W_1; \breve{Y}_1^N, \breve{V}_1^N | \underline{S}^N) + I(W_2; \breve{Y}_2^N, \breve{V}_2^N | \underline{S}^N) \\
&= h(\breve{Y}_1^N | \underline{S}^N, \breve{V}_1^N) + h(\breve{Y}_2^N | \underline{S}^N, \breve{V}_2^N) && \text{(A.21)} \\
&\quad + h(\breve{V}_1^N | \underline{S}^N) + h(\breve{V}_2^N | \underline{S}^N) && \text{(A.22)} \\
&\quad - h(\breve{Y}_1^N, \breve{V}_1^N | \underline{S}^N, W_1) - h(\breve{Y}_2^N, \breve{V}_2^N | \underline{S}^N, W_2) && \text{(A.23)}
\end{aligned}$$

Let us take one term from (A.21).

$$\begin{aligned}
h(\breve{Y}_1^N | \underline{S}^N, \breve{V}_1^N) &= \mathbb{E}_{S^N}\left[h_S\left(\breve{Y}_1^N | \breve{V}_1^N\right)\right] \\
&\overset{\text{(a)}}{=} \mathbb{E}_{S^N}\left[h_S\left(\bar{Y}_1^{L_1}, Y_{11}^{M_1}, Y_{12}^{M_1} | \breve{V}_1^N\right)\right] \\
&\leq \mathbb{E}_{S^N}\left[h_S\left(\bar{Y}_1^{L_1} | \breve{V}_1^N\right)\right] + \mathbb{E}_{S^N}\left[h_S\left(Y_{11}^{M_1} | \breve{V}_1^N\right)\right] \\
&\quad + \mathbb{E}_{S^N}\left[h_S\left(Y_{12}^{M_1} | \breve{V}_1^N\right)\right]
\end{aligned}$$

where (a) follows by (with a slight abuse of notation) decomposing $\breve{Y}_1^N$ into $\left(Y_{11}^{M_1}, Y_{12}^{M_1}\right)$ for time slots where $S_{1,t} = 1$, and into $\bar{Y}_1^{L_1}$ for time slots for which $S_{1,t} = 0$.

The other term in (A.21) can be bounded similarly. Let us take one term from (A.23).

$$\begin{aligned}
-h(\breve{Y}_1^N, \breve{V}_1^N | \underline{S}^N, W_1) &= -\mathbb{E}_{S^N}\left[h_S\left(\breve{Y}_1^N, \breve{V}_1^N | W_1\right)\right] \\
&\overset{\text{(b)}}{=} -\mathbb{E}_{S^N}\left[h_S\left(\breve{V}_2^N, Z_{11}^{M_1}, Z_2^{L_2}, Z_{21}^{M_2} | W_1\right)\right] \\
&\overset{\text{(c)}}{=} -\sum_{t=1}^{N} \mathbb{E}_{S^N}\Big[h\Big(\breve{V}_{2,t}, Z_{11,(t)}, Z_{2,[t]}, Z_{21,(t)} \\
&\qquad\qquad \Big| W_1, \breve{V}_2^{t-1}, Z_{11}^{(t-1)}, Z_2^{[t-1]}, Z_{21}^{(t-1)}\Big)\Big] \\
&\overset{\text{(d)}}{=} -\mathbb{E}\left[\sum_{t=1}^{N} h_S\left(\breve{V}_{2,t} | W_1, \breve{V}_2^{t-1}, Z_{11}^{(t-1)}, Z_2^{[t-1]} Z_{21}^{(t-1)}\right)\right.
\end{aligned}$$

$$
\begin{aligned}
&+ \sum_{t:S_{1,t}=1} h_S\left(Z_{11,t}|W_1, \breve{V}_2^{t-1}, Z_{11}^{(t-1)}, Z_2^{[t-1]} Z_{21}^{(t-1)}\right) \\
&+ \sum_{t:S_{2,t}=0} h_S\left(Z_{2,t}|W_1, \breve{V}_2^{t-1}, Z_{11}^{(t-1)}, Z_2^{[t-1]} Z_{21}^{(t-1)}\right) \\
&\left. + \sum_{t:S_{2,t}=1} h_S\left(Z_{21,t}|W_1, \breve{V}_2^{t-1}, Z_{11}^{(t-1)}, Z_2^{[t-1]} Z_{21}^{(t-1)}\right)\right] \\
\overset{\text{(e)}}{=} &-\mathbb{E}_{S^N}\left[\sum_{t=1}^{N} h_S\left(\breve{V}_{2,t}|W_1, \breve{V}_2^{t-1}, Z_{11}^{(t-1)}, Z_2^{[t-1]} Z_{21}^{(t-1)}\right)\right] \\
&- \mathbb{E}_{S^N}\left[L_2 h(Z_2) + M_2 h(Z_{21}) + M_1 h(Z_{11})\right] \\
= &-\mathbb{E}_{S^N}\left[\sum_{t=1}^{N} h_S\left(\breve{V}_{2,t}|W_1, \breve{V}_2^{t-1}, Z_{11}^{(t-1)}, Z_2^{[t-1]} Z_{21}^{(t-1)}\right)\right] \\
&- N(1-p_2)h(Z_2) - Np_2 h(Z_{21}) - Np_1 h(Z_{11})
\end{aligned}
$$

where (b) follows by Lemma A.6, (c) follows by chain rule, (d) follows by the fact that for a given time slot t, the involved noise terms are independent from each other and from $\breve{V}_{2,t}$; and (e) is because the signals up to time $t-1$ are independent from the noise at time t, and because noise processes are i.i.d.

The other term in (A.23) can be bounded similarly.

Putting everything together, we have

$$
\begin{aligned}
N(R_1 + R_2 - \epsilon_N) \leq\ & \mathbb{E}_{S^N}\left[h_S\left(\bar{Y}_1^{L_1}|\breve{V}_1^N\right)\right] + \mathbb{E}_{S^N}\left[h_S\left(Y_{11}^{M_1}|\breve{V}_1^N\right)\right] \\
&+ \mathbb{E}_{S^N}\left[h_S\left(\bar{Y}_2^{L_2}|\breve{V}_2^N\right)\right] + \mathbb{E}_{S^N}\left[h_S\left(Y_{22}^{M_2}|\breve{V}_2^N\right)\right] \\
&+ \mathbb{E}_{S^N}\left[h_S\left(\breve{V}_1^N, Y_{12}^{M_1}\right)\right] & \text{(A.24)} \\
&+ \mathbb{E}_{S^N}\left[h_S\left(\breve{V}_2^N, Y_{21}^{M_2}\right)\right] & \text{(A.25)} \\
&- \mathbb{E}_{S^N}\left[\sum_{t=1}^{N} h_S\left(\breve{V}_{2,t}|W_1, \breve{V}_2^{t-1}, Z_{11}^{(t-1)}, Z_2^{[t-1]} Z_{21}^{(t-1)}\right)\right] & \text{(A.26)} \\
&- \mathbb{E}_{S^N}\left[\sum_{t=1}^{N} h_S\left(\breve{V}_{1,t}|W_2, \breve{V}_1^{t-1}, Z_{22}^{(t-1)}, Z_1^{[t-1]} Z_{12}^{(t-1)}\right)\right] & \text{(A.27)} \\
&- N(1-p_2)h(Z_2) - Np_2 h(Z_{21}) - Np_1 h(Z_{11})
\end{aligned}
$$

$$- N(1 - p_1)h(Z_1) - Np_1h(Z_{12}) - Np_2h(Z_{22})$$

Let us combine (A.24) and (A.27).

$$\begin{aligned}
&(\text{A.24}) + (\text{A.27}) \\
&= \mathbb{E}_{S^N}\left[h_S\left(\breve{V}_1^N, Y_{12}^{M_1}\right) - \sum_{t=1}^{N} h_S\left(\breve{V}_{1,t}|W_2, \breve{V}_1^{t-1}, Z_{22}^{(t-1)}, Z_1^{[t-1]}Z_{12}^{(t-1)}\right)\right] \\
&= \mathbb{E}_{S^N}\Bigg[\sum_{t=1}^{N} h_S\left(\breve{V}_{1,t}|\breve{V}_1^{t-1}, Y_{12}^{(t-1)}\right) - h_S\left(\breve{V}_{1,t}|W_2, \breve{V}_1^{t-1}, Z_{22}^{(t-1)}, Z_1^{[t-1]}Z_{12}^{(t-1)}\right) \\
&\quad + \sum_{t:S_{1,t}=1} h_S\left(Y_{12,t}|\breve{V}_1^{t}, Y_{12}^{(t-1)}\right)\Bigg] \\
&\leq \mathbb{E}_{S^N}\left[\sum_{t:S_{1,t}=1} h_S\left(Y_{12,t}\right) + \sum_{t=1}^{N} I_S(\breve{V}_{1,t}; W_2, Z_{22}^{(t-1)}, Z_1^{[t-1]}Z_{12}^{(t-1)}|\breve{V}_1^{t-1}, Y_{12}^{(t-1)}\right]
\end{aligned}$$

Similarly, we combine (A.25) with (A.26) to obtain the same expression with user indices swapped. Plugging these back, we get

$$\begin{aligned}
N(R_1 + R_2 - \epsilon_N) \leq\ & \mathbb{E}_{S^N}\left[h_S\left(\bar{Y}_1^{L_1}|\breve{V}_1^N\right)\right] + \mathbb{E}_{S^N}\left[h_S\left(\bar{Y}_{11}^{M_1}|\breve{V}_1^N\right)\right] \\
& + \mathbb{E}_{S^N}\Bigg[h_S\left(\bar{Y}_2^{L_2}|\breve{V}_2^N\right) + h_S\left(\bar{Y}_{22}^{M_2}|\breve{V}_2^N\right) \\
& + \sum_{t=1}^{N} I_S(\breve{V}_{1,t}; W_2, Z_{22}^{(t-1)}, Z_1^{[t-1]}Z_{12}^{(t-1)}|\breve{V}_1^{t-1}, Y_{12}^{(t-1)}) \\
& + \sum_{t=1}^{N} I_S(\breve{V}_{2,t}; W_1, Z_{11}^{(t-1)}, Z_2^{[t-1]}Z_{21}^{(t-1)}|\breve{V}_2^{t-1}, Y_{21}^{(t-1)})\Bigg] \\
& + \mathbb{E}_{S^N}\left[\sum_{t:S_{1,t}=1} h_S\left(Y_{12,t}\right)\right] + \mathbb{E}_{S^N}\left[\sum_{t:S_{2,t}=1} h_S\left(Y_{21,t}\right)\right] \\
& - N(1 - p_2)h(Z_2) - Np_2h(Z_{21}) - Np_1h(Z_{11}) \\
& - N(1 - p_1)h(Z_1) - Np_1h(Z_{12}) - Np_2h(Z_{22})
\end{aligned}$$

We use Lemmas A.3, A.4, A.7 and A.8 to bound each of these terms, and use the fact that noise distribution is Gaussian to obtain the desired bound.

A.5.4 Proof of the Bounds (2.62) and (2.63)

We exclusively focus on the enhanced channel, defined in Section 2.7. By symmetry, it is sufficient to prove (2.62). In addition to the enhanced interference channel in the case of sum rate bound, we consider an additional copy of Receiver 1, who always receives $\bar{Y}_1$ (*i.e.*, as in the original channel). The feedback signal of Tx1 is still given by $S_1 \cdot \bar{Y}_1$, *i.e.*, the same as the original channel. We would like to prove a sum rate upper bound on this new channel.

By Fano's inequality,

$$\begin{aligned} N(2R_1 + R_2 - \epsilon_N) &\leq I(W_1; \bar{Y}_1^N, \underline{S}^N) + I(W_2; \breve{Y}_2^N, \underline{S}^N) + I(W_1; \breve{Y}_1^N, \underline{S}^N) \\ &= I(W_1; \bar{Y}_1^N | \underline{S}^N) + I(W_2; \breve{Y}_2^N | \underline{S}^N) + I(W_1; \breve{Y}_1^N | \underline{S}^N) \\ &= I(W_1; \bar{Y}_1^N | \underline{S}^N) + I(W_2; \breve{Y}_2^N, \bar{V}_2^N | \underline{S}^N) \\ &\quad + I(W_1; \breve{Y}_1^N, \breve{V}_1^N | W_2, \underline{S}^N) \end{aligned} \tag{A.28}$$

The first mutual information term in (A.28) can be bounded as follows

$$\begin{aligned} I(W_1; \bar{Y}_1^N | \underline{S}^N) &= h(\bar{Y}_1^N | \underline{S}^N) - h(\bar{Y}_1^N | W_1, \underline{S}^N) \\ &= h(\bar{Y}_1^N | \underline{S}^N) - \sum_{t=1}^{N} h(\bar{Y}_{1,t} | W_1, \bar{Y}_1^{t-1}, \underline{S}^N) \\ &\overset{(a)}{=} h(\bar{Y}_1^N | \underline{S}^N) - \sum_{t=1}^{N} h(\bar{V}_{2,t} | W_1, \bar{V}_2^{t-1}, \underline{S}^N) \\ &\leq \sum_{t=1}^{N} h(\bar{Y}_{1,t}) - h(\bar{V}_{2,t} | W_1, \bar{V}_2^{t-1}, \underline{S}^N) \end{aligned}$$

where (a) follows by the fact that $X_{1,t} \overset{\text{f}}{=} \left(W_1, \bar{Y}_1^{t-1}, \underline{S}^{t-1}\right)$ and by subtracting $X_{1,t}$ from $\bar{Y}_{1,t}$.

Let us consider the second mutual information term from (A.28).

$$\begin{aligned} & I(W_2; \breve{Y}_2^N, \bar{V}_2^N | \underline{S}^N) = \mathbb{E}_{S^N}\left[h_S\left(\breve{Y}_2^N, \bar{V}_2^N\right) - h_S\left(\breve{Y}_2^N, \bar{V}_2^N | W_2\right)\right] \\ & \overset{(a)}{=} \mathbb{E}_{S^N}\left[h_S\left(Y_{22}^{M_2}, Y_{21}^{M_2}, \bar{Y}_2^{L_2}, \bar{V}_2^N\right) - h_S\left(\breve{V}_1^N, Z_1^N, Z_{22}^{M_2} | W_2\right)\right] \\ & = \mathbb{E}_{S^N}\left[h_S\left(Y_{22}^{M_2}, \bar{Y}_2^{L_2} | Y_{21}^{M_2}, \bar{V}_2^N\right)\right] \\ & \quad + \mathbb{E}_{S^N}\left[h_S\left(Y_{21}^{M_2}, \bar{V}_2^N\right) - h_S\left(\breve{V}_1^N, Z_1^N, Z_{22}^{M_2} | W_2\right)\right] \end{aligned}$$

$$
\begin{aligned}
&\overset{\text{(b)}}{\leq} \mathbb{E}_{S^N}\left[h_S\left(Y_{22}^{M_2}|\bar{V}_2^N\right) + h_S\left(\bar{Y}_2^{L_2}|\bar{V}_2^N\right)\right] \\
&\quad + \mathbb{E}_{S^N}\left[\sum_{t=1}^{N} h_S\left(\bar{V}_{2,t}|\bar{V}_2^{t-1}, Y_{21}^{(t-1)}\right) + \sum_{t:S_{2,t}=1}^{N} h_S\left(Y_{21,t}|\bar{V}_2^{t-1}, Y_{21}^{(t-1)}\right)\right. \\
&\quad \left. - \sum_{t=1}^{N} h_S\left(\breve{V}_{1,t}, Z_{1,t}, Z_{22,(t)}|W_2, \breve{V}_1^{t-1}, Z_1^{t-1}, Z_{22}^{(t-1)}\right)\right] \\
&\overset{\text{(c)}}{=} \mathbb{E}_{S^N}\left[h_S\left(Y_{22}^{M_2}|\bar{V}_2^N\right) + h_S\left(\bar{Y}_2^{L_2}|\bar{V}_2^N\right)\right] \\
&\quad + \mathbb{E}_{S^N}\left[\sum_{t=1}^{N} h_S\left(\bar{V}_{2,t}|\bar{V}_2^{t-1}, Y_{21}^{(t-1)}\right) + \sum_{t:S_{2,t}=1}^{N} h_S\left(Y_{21,t}|\bar{V}_2^{t-1}, Y_{21}^{(t-1)}\right)\right] \\
&\quad - \mathbb{E}_{S^N}\left[\sum_{t=1}^{N} h_S\left(\breve{V}_{1,t}|W_2, \breve{V}_1^{t-1}, Z_1^{t-1}, Z_{22}^{(t-1)}\right)\right. \\
&+ \sum_{t=1}^{N} h_S\left(Z_{1,t}|W_2, \breve{V}_1^{t-1}, Z_1^{t-1}, Z_{22}^{(t-1)}\right) \\
&\quad \left. + \sum_{t:S_{2,t}=1} h_S\left(Z_{22,t}|W_2, \breve{V}_1^{t-1}, Z_1^{t-1}, Z_{22}^{(t-1)}\right)\right] \\
&\overset{\text{(d)}}{\leq} \mathbb{E}_{S^N}\left[h_S\left(Y_{22}^{M_2}|\bar{V}_2^N\right) + h_S\left(\bar{Y}_2^{L_2}|\bar{V}_2^N\right)\right] \\
&\quad + \mathbb{E}_{S^N}\left[\sum_{t=1}^{N} h_S\left(\bar{V}_{2,t}|\bar{V}_2^{t-1}, Y_{21}^{(t-1)}\right) + \sum_{t:S_{2,t}=1} h_S\left(Y_{21,t}\right)\right] \\
&\quad - \mathbb{E}_{S^N}\left[\sum_{t=1}^{N} h_S\left(\breve{V}_{1,t}|W_2, \breve{V}_1^{t-1}, Z_1^{t-1}, Z_{22}^{(t-1)}\right)\right] \\
&\quad - Nh(Z_1) - Np_2 h(Z_{22})
\end{aligned}
$$

where (a) is by decomposing $\breve{Y}_2^N$ into $\left(Y_{22}^{M_2}, Y_{21}^{M_2}\right)$ for time slots where $S_{2,t} = 1$, and to $\bar{Y}_2^{L_2}$ for time slots where $S_{2,t} = 0$, and by Lemma A.6. (b) is because conditioning reduces entropy and by chain rule. (c) is because for a given time slot t, the noise terms involved are independent from each other and from $\breve{V}_{1,t}$. (d) is because conditioning reduces entropy and the noise processes are i.i.d., and because the noise terms are independent from the signals up to time $t-1$.

Next, we consider the third mutual information term from (A.28).

$$\begin{aligned}
&I(W_1;\breve{Y}_1^N,\breve{V}_1^N|W_2,\underline{S}^N)=\mathbb{E}_{S^N}\left[I_S(W_1;\breve{Y}_1^N,\breve{V}_1^N|W_2)\right]\\
&\overset{\text{(a)}}{\leq}\mathbb{E}_{S^N}\left[I(W_1;\breve{Y}_1^N,\breve{V}_1^N|W_2,Z_{22}^{M_2},Z_{11}^{M_1})\right]\\
&=\mathbb{E}_{S^N}\left[h_S\left(\breve{Y}_1^N,\breve{V}_1^N|W_2,Z_{22}^{M_2},Z_{11}^{M_1}\right)\right]\\
&-\mathbb{E}_{S^N}\left[h_S\left(\breve{Y}_1^N,\breve{V}_1^N|W_2,Z_{22}^{M_2},Z_{11}^{M_1},W_1\right)\right]\\
&=\sum_{t=1}^{N}\mathbb{E}_{S^N}\left[h_S\left(\breve{Y}_{1,t},\breve{V}_{1,t}|\breve{Y}_1^{t-1},\breve{V}_1^{t-1},W_2,Z_{22}^{M_2},Z_{11}^{M_1}\right)\right]\\
&\quad-\mathbb{E}_{S^N}\left[h_S\left(\breve{Y}_{1,t},\breve{V}_{1,t}|\breve{Y}_1^{t-1},\breve{V}_1^{t-1},W_2,Z_{22}^{M_2},Z_{11}^{M_1},W_1\right)\right]\\
&=\sum_{t=1}^{N}\mathbb{E}_{S^N}\left[h_S\left(\breve{Y}_{1,t}|\breve{Y}_1^{t-1},\breve{V}_1^{t},W_2,Z_{22}^{M_2},Z_{11}^{M_1}\right)\right]\\
&\quad+\mathbb{E}_{S^N}\left[h_S\left(\breve{V}_{1,t}|\breve{Y}_1^{t-1},\breve{V}_1^{t-1},W_2,Z_{22}^{M_2},Z_{11}^{M_1}\right)\right]\\
&-\mathbb{E}_{S^N}\left[h_S\left(\breve{Y}_{1,t},\breve{V}_{1,t}|\breve{Y}_1^{t-1},\breve{V}_1^{t-1},W_2,Z_{22}^{M_2},Z_{11}^{M_1},W_1\right)\right]\\
&\overset{\text{(b)}}{=}\sum_{t=1}^{N}\mathbb{E}_{S^N}\left[h_S\left(\breve{Y}_{1,t}|\breve{Y}_1^{t-1},\breve{V}_1^{t},W_2,Z_{22}^{M_2},Z_{11}^{M_1},X_{2,t}\right)\right]\\
&+\mathbb{E}_{S^N}\left[h_S\left(\breve{V}_{1,t}|\breve{Y}_1^{t-1},\breve{V}_1^{t-1},W_2,Z_{22}^{M_2},Z_{11}^{M_1}\right)\right]\\
&-\mathbb{E}_{S^N}\left[h_S\left(\breve{Y}_{1,t},\breve{V}_{1,t}|\breve{Y}_1^{t-1},\breve{V}_1^{t-1},W_2,Z_{22}^{M_2},Z_{11}^{M_1},W_1,X_{1,t},X_{2,t}\right)\right]\\
&\overset{\text{(c)}}{\leq}\sum_{t=1}^{N}\mathbb{E}_{S^N}\left[h_S\left(\breve{Y}_{1,t}|\breve{V}_{1,t},X_{2,t}\right)\right]+\mathbb{E}_{S^N}\left[h_S\left(\breve{V}_{1,t}|Y_{12}^{(t-1)},\breve{V}_1^{t-1},W_2,Z_{22}^{M_2},Z_{11}^{M_1}\right)\right]\\
&-\mathbb{E}_{S^N}\left[\sum_{t:S_{2,t}=0}h_S\left(Z_{2,t}\right)+\sum_{t:S_{2,t}=1}h_S\left(Z_{21,t}\right)\right]\\
&-\mathbb{E}_{S^N}\left[\sum_{t:S_{1,t}=0}h_S\left(Z_{1,t}\right)+\sum_{t:S_{1,t}=1}h_S\left(Z_{11,t},Z_{12,t}\right)\right]\\
&\overset{\text{(d)}}{\leq}\sum_{t=1}^{N}\mathbb{E}_{S^N}\left[h_S\left(\breve{Y}_{1,t}|\breve{V}_{1,t},X_{2,t}\right)\right]\\
&+\mathbb{E}_{S^N}\left[h_S\left(\breve{V}_{1,t}|Y_{12}^{(t-1)},\breve{V}_1^{t-1},W_2,Z_{22}^{(t-1)},Z_{11}^{(t-1)}\right)\right]\\
&-N(1-p_1)h(Z_1)-Np_1h(Z_{11})-Np_1h(Z_{12})-Np_2h(Z_{21})-N(1-p_2)h(Z_2)
\end{aligned}$$

where (a) follows by the fact that $\left(Z_{11}^{M_1}, Z_{22}^{M_2}\right)$ is independent from W_1 given W_2, (b) is because $X_{1,t} \overset{\text{f}}{=} \left(W_1, \breve{Y}_1^{t-1}, \underline{S}^{t-1}\right)$ and $X_{2,t} \overset{\text{f}}{=} \left(W_2, \breve{V}_1^{t-1}, Z_{22}^{(t-1)}, \underline{S}^{t-1}\right)$. In (c), the first two terms are upper bounded using the fact that

$$\breve{Y}_1^{t-1} = \left(Y_{11}^{(t-1)}, Y_{12}^{(t-1)}, \bar{Y}_1^{[t-1]}\right),$$

and that conditioning reduces entropy. The noise terms are obtained by subtracting $X_{1,t}$ and $X_{2,t}$ from $\breve{Y}_{1,t}$ and $\breve{V}_{1,t}$, and using the fact that noise variables at time t are independent from the variables up to time $t-1$. (d) is because conditioning reduces entropy and because noise processes are i.i.d.

Putting everything back together, we have

$$\begin{aligned}
N(2R_1 + R_2 - \epsilon_N) \leq \sum_{t=1}^{N} h(\bar{Y}_{1,t}) &+ \mathbb{E}_{S^N}\left[h_S\left(Y_{22}^{M_2}|\bar{V}_2^N\right) + h_S\left(\bar{Y}_2^{L_2}|\bar{V}_2^N\right)\right] \\
&+ \mathbb{E}_{S^N}\left[\sum_{t=1}^{N} I_S(\bar{V}_{2,t}; W_1|\bar{V}_2^{t-1}, Y_{21}^{(t-1)})\right. \\
&+ \sum_{t=1}^{N} I_S(\breve{V}_{1,t}; Z_1^{t-1}|Y_{12}^{(t-1)}, \breve{V}_1^{t-1}, W_2, Z_{22}^{(t-1)}, Z_{11}^{(t-1)}) \\
&\left.+ \sum_{t=1}^{N} h_S\left(\breve{Y}_{1,t}|\breve{V}_{1,t}, X_{2,t}\right) + \sum_{t:S_{2,t}=1} h_S\left(Y_{21,t}\right)\right] \\
&- N(1-p_1)h(Z_1) - Np_1h(Z_{11}) - Np_1h(Z_{12}) \\
&- Np_2h(Z_{21}) - N(1-p_2)h(Z_2) - Nh(Z_1) \\
&- Np_2h(Z_{22})
\end{aligned}$$

Using Lemma A.3, A.4, A.5, A.7, A.8 and A.9 to bound each of these terms, we get the desired bound.

A.5.5 Lemmas

In this subsection, we prove the lemmas that have been used in the proofs of the previous subsections.

Lemma A.1. $X_{2,t} \overset{\mathrm{f}}{=} \left(W_2, \widetilde{V}_1^{t-1}, \underline{S}^t\right)$

Proof. Note that

$$X_{2,1} \overset{\mathrm{f}}{=} W_2$$

and by the definition of the channel,

$$X_{2,t} \overset{\mathrm{f}}{=} \left(W_2, \widetilde{Y}_2^{t-1}, \underline{S}^t\right) \overset{\overset{\text{(a)}}{\mathrm{f}}}{=} \left(W_2, \widetilde{V}_1^{t-1}, X_2^{t-1}, \underline{S}^t\right),$$

hence the result follows by induction on t. (a) follows because

$$\widetilde{Y}_2^{t-1} = S_2^{t-1} h_{22} X_2^{t-1} + \widetilde{V}_1^{t-1}.$$

□

Lemma A.2.

$$\sum_{t=1}^{N} h(Y_{1,t}|\widetilde{V}_{1,t}, S_{2,t}, X_{2,t}) \leq N p_2 \log\left(1 + \frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right) + N(1-p_2)\log\left(1+\mathsf{SNR}_1\right)$$

Proof.

$$\begin{aligned}
&\sum_{t=1}^{N} h(Y_{1,t}|\widetilde{V}_{1,t}, S_{2,t}, X_{2,t}) = \sum_{t=1}^{N} p_2 h(Y_{1,t}|V_{1,t}, X_{2,t}) + (1-p_2) h(Y_{1,t}|X_{2,t}) \\
&\overset{\text{(a)}}{=} \sum_{t=1}^{N} p_2 h(Y_{1,Q}|V_{1,Q}, X_{2,Q}, Q=t) + (1-p_2) h(Y_{1,Q}|X_{2,Q}, Q=t) \\
&= N p_2 h(Y_{1,Q}|V_{1,Q}, X_{2,Q}, Q) + N(1-p_2) h(Y_{1,Q}|X_{2,Q}, Q) \\
&\overset{\text{(b)}}{=} N p_2 h(Y_1|V_1, X_2, Q) + N(1-p_2) h(Y_1|X_2, Q) \\
&\overset{\text{(c)}}{\leq} N p_2 \log\left(1 + \frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right) + N(1-p_2)\log\left(1+\mathsf{SNR}_1\right)
\end{aligned}$$

where (a) follows by introducing a time-sharing variable Q uniformly distributed between 1 and N, (b) is by defining $Y_i := Y_{i,Q}$, $X_i := X_{i,Q}$ and $V_i := V_{i,Q}$. (c) follows by the fact that choosing jointly Gaussian input distribution with correlation coefficient $\rho = 0$ for $p(x_1, x_2)$ maximizes the given conditional differential entropy. □

Lemma A.3.

$$\mathbb{E}_{S^N}\left[\sum_{t:S_{i,t}=1} h_S\left(Y_{ij,t}\right)\right] = Np_i \log\left(2\pi e\left(\frac{1}{2}+\mathsf{INR}_i\right)\right)$$

for $(i,j)=(1,2),(2,1)$.

Proof.

$$\begin{aligned}
&\mathbb{E}_{S^N}\left[\sum_{t:S_{i,t}=1} h_S\left(Y_{ij,t}\right)\right] = \mathbb{E}_{S^N}\left[M_i\frac{1}{M_i}\sum_{t:S_{i,t}=1} h_S\left(Y_{ij,t}\right)\right]\\
&\le \mathbb{E}_{S^N}\left[M_i\frac{1}{M_i}\sum_{t:S_{i,t}=1}\log 2\pi e\left(\frac{1}{2}+|h_{ij}|^2P_{j,t}\right)\right]\\
&= \mathbb{E}\left[M_i\mathbb{E}\left[\frac{1}{M_i}\sum_{t:S_{i,t}=1}\log 2\pi e\left(\frac{1}{2}+|h_{ij}|^2P_{j,t}\right)\middle| M_i\right]\right]\\
&\overset{\text{(a)}}{\le} \mathbb{E}_{M_i}\left[M_i\mathbb{E}\left[\log 2\pi e\left(\frac{1}{2}+|h_{ij}|^2P_j^{(i1)}\right)\middle| M_i\right]\right]\\
&\overset{\text{(b)}}{\le} \mathbb{E}_{M_i}\left[M_i\log 2\pi e\left(\frac{1}{2}+|h_{ij}|^2\mathbb{E}\left[P_j^{(i1)}|M_i\right]\right)\right]\\
&\overset{\text{(c)}}{\le} \mathbb{E}_{M_i}\left[M_i\log 2\pi e\left(\frac{1}{2}+|h_{ij}|^2P_j\right)\right]\\
&= \mathbb{E}\left[M_i\right]\log\left(2\pi e\left(\frac{1}{2}+\mathsf{INR}_i\right)\right)\\
&= Np_i\log\left(2\pi e\left(\frac{1}{2}+\mathsf{INR}_i\right)\right)
\end{aligned}$$

where (a) and (b) follow by Jensen's inequality (since $\log(\cdot)$ is concave), and (c) follows since $P_j^{(i1)}$ averaged over the realizations of S_i is the average power, which is less than the power constraint P_j. □

Lemma A.4.

$$\begin{aligned}
&\mathbb{E}_{S^N}\left[I_S(V_{i,t};W_j,Z_{jj}^{(t-1)},Z_i^{[t-1]},Z_{ij}^{(t-1)}|V_i^{t-1},Y_{ij}^{(t-1)})\right]\\
&\quad=\mathbb{E}_{S^N}\left[I_S(V_{i,t};Z_i^{t-1}|Y_{ij}^{(t-1)},V_i^{t-1},W_j,Z_{jj}^{(t-1)},Z_{ii}^{(t-1)})\right]=0
\end{aligned}$$

for $(i,j)=(1,2),(2,1)$.

Proof. Since all variables involved are related to $V_{i,t}$ through $X_{i,t}$; by data processing inequality,

$$\begin{aligned}
&\mathbb{E}_{S^N}\left[I_S(V_{i,t};W_j,Z_{jj}^{(t-1)},Z_i^{[t-1]},Z_{ij}^{(t-1)}|V_i^{t-1},Y_{ij}^{(t-1)})\right]\\
&\quad\leq \mathbb{E}_{S^N}\left[I_S(X_{i,t};W_j,Z_{jj}^{(t-1)},Z_i^{[t-1]},Z_{ij}^{(t-1)}|V_i^{t-1},Y_{ij}^{(t-1)})\right]\\
&\quad\leq \mathbb{E}_{S^N}\left[I_S(W_i,Z_{ii}^{(t-1)};W_j,Z_{jj}^{(t-1)},Z_i^{[t-1]},Z_{ij}^{(t-1)}|V_i^{t-1},Y_{ij}^{(t-1)})\right]
\end{aligned}$$

where the latter inequality follows by the fact that $X_{i,t} \overset{\text{f}}{=} \left(W_i, Y_{ij}^{(t-1)}, Z_{ii}^{(t-1)}, \underline{S}^{t-1}\right)$. Similarly, the second mutual information can be bounded by

$$\begin{aligned}
&\mathbb{E}_{S^N}\left[I_S(V_{i,t};Z_i^{t-1}|Y_{ij}^{(t-1)},V_i^{t-1},W_j,Z_{jj}^{(t-1)},Z_{ii}^{(t-1)})\right]\\
&\quad\leq \mathbb{E}_{S^N}\left[I_S(X_{i,t};Z_i^{[t-1]},Z_{ij}^{(t-1)}|Y_{ij}^{(t-1)},V_i^{t-1},W_j,Z_{jj}^{(t-1)},Z_{ii}^{(t-1)})\right]\\
&\quad\leq \mathbb{E}_{S^N}\left[I_S(W_i,Z_{ii}^{(t-1)};Z_{jj}^{(t-1)},Z_i^{[t-1]},Z_{ij}^{(t-1)},W_j|V_i^{t-1},Y_{ij}^{(t-1)})\right]
\end{aligned}$$

where the first step is because $Z_i^{t-1} \overset{\text{f}}{=} \left(Z_i^{[t-1]}, Z_{ij}^{(t-1)}, Z_{ii}^{(t-1)}, \underline{S}^{t-1}\right)$, and second step is because for random variables A, B, C; $I(A,B;C) \geq I(A;C|B)$; and $X_{i,t} \overset{\text{f}}{=} \left(W_i, Y_{ij}^{(t-1)}, Z_{ii}^{(t-1)}, \underline{S}^{t-1}\right)$. Note that we have the same upper bound for both mutual information terms. We will next show that this upper bound is zero.

To show conditional independence, we will use the property that X and Y are independent given Z if and only if the probability distribution $p(X,Y,Z)$ can be factorized as

$$p(X,Y,Z) = f(X,Z)g(Y,Z)$$

for some functions f and g. Consider the joint distribution of all the variables involved in the above mutual information (we define $p_S(\cdot) := p(\cdot|\underline{S}^N = S^N)$).

$$\begin{aligned}
&p_S(W_i,Z_{ii}^{(t-1)},Z_{jj}^{(t-1)},Z_i^{[t-1]},Z_{ij}^{(t-1)},W_j,V_i^{t-1},Y_{ij}^{(t-1)})\\
&= p(W_i)p(W_j)\prod_{\tau=1}^{t-1} p_S(Z_{ii,(\tau)},Z_{jj,(\tau)},Z_{i,[\tau]},Z_{ij,(\tau)},V_{i,\tau},Y_{ij,(\tau)}\\
&\qquad\qquad |Z_{ii}^{(\tau-1)},Z_{jj}^{(\tau-1)},Z_i^{[\tau-1]},Z_{ij}^{(\tau-1)},V_i^{\tau-1},Y_{ij}^{(\tau-1)},W_i,W_j)\\
&\overset{\text{(a)}}{=} p(W_i)p(W_j)\prod_{\tau=1}^{t-1} p_S(Z_{ii,(\tau)})p_S(Z_{jj,(\tau)})p_S(Z_{i,[\tau]})p_S(Z_{ij,(\tau)})
\end{aligned}$$

$$
\begin{aligned}
&\cdot p_S(V_{i,\tau}|Z_{ii}^{(\tau-1)}, Y_{ij}^{(\tau-1)}, W_i) p_S(Y_{ij,(\tau)}|Z_{jj}^{(\tau-1)}, Z_{ij}^{(\tau)}, V_i^{\tau-1}, W_j)\\
&= f(W_i, Z_{ii}^{(t-1)}, V_i^{t-1}, Y_{ij}^{(t-1)}) \cdot g(Z_{jj}^{(t-1)}, Z_i^{(t-1)}, Z_{ij}^{(t-1)}, W_j, V_i^{t-1}, Y_{ij}^{(t-1)})
\end{aligned}
$$

where (a) follows since

$$
Y_{ij,(\tau)} \overset{\mathrm{f}}{=} \left(X_{j,\tau}, Z_{ij}^{(\tau)}, \underline{S}^{\tau}\right) \overset{\mathrm{f}}{=} \left(Z_{jj}^{(\tau-1)}, Z_{ij}^{(\tau)}, V_i^{\tau-1}, \underline{S}^{\tau}, W_j\right)
$$

and

$$
V_{i,\tau} \overset{\mathrm{f}}{=} \left(X_{i,\tau}, Z_{j,(\tau)}, Z_{ji,(\tau)}\right) \overset{\mathrm{f}}{=} \left(Z_{ii}^{(\tau-1)}, Y_{ij}^{(\tau-1)}, \underline{S}^{\tau-1}, W_i, Z_{j,(\tau)}, Z_{ji,(\tau)}\right)
$$

and $\left(Z_{j,(\tau)}, Z_{ji,(\tau)}\right)$ is independent of everything else. In the last line, we define

$$
\begin{aligned}
f(W_i, Z_{ii}^{(t-1)}, V_i^{t-1}, Y_{ij}^{(t-1)}) &= p(W_i) \prod_{\tau=1}^{t-1} p_S(Z_{ii,(\tau)}) p_S(V_{i,\tau}|Z_{ii}^{(\tau-1)}, Y_{ij}^{(\tau-1)}, W_i)\\
g(Z_{jj}^{(t-1)}, Z_i^{(t-1)}, Z_{ij}^{(t-1)}, W_j, V_i^{t-1}, Y_{ij}^{(t-1)}) &= p(W_j) \prod_{\tau=1}^{t-1} p_S(Z_{jj,(\tau)}) p_S(Z_{i,(\tau)})\\
&\quad \cdot_S (Z_{ij,(\tau)}) p_S(Y_{ij,\tau}|Z_{jj}^{(\tau-1)}, Z_{ij}^{(\tau)}, V_i^{\tau-1}, W_j)
\end{aligned}
$$

from which the result follows. □

Lemma A.5.

$$
\mathbb{E}_{S^N}\left[I_S(\bar{V}_{j,t}; W_i|\bar{V}_j^{t-1}, Y_{ji}^{(t-1)})\right] = 0
$$

Proof.

$$
\begin{aligned}
\mathbb{E}_{S^N}\left[I_S(\bar{V}_{j,t}; W_i|\bar{V}_j^{t-1}, Y_{ji}^{(t-1)})\right] &\leq \mathbb{E}_{S^N}\left[I_S(X_{j,t}; W_i|\bar{V}_j^{t-1}, Y_{ji}^{(t-1)})\right]\\
&\leq \mathbb{E}_{S^N}\left[I_S(W_j, Z_{jj}^{(t-1)}; W_i|\bar{V}_j^{t-1}, Y_{ji}^{(t-1)})\right]
\end{aligned}
$$

where the first step follows by data processing inequality, and the second one follows by $X_{j,t} \overset{\mathrm{f}}{=} \left(W_j, Y_{ji}^{(t-1)}, Z_{jj}^{(t-1)}, \underline{S}^{t-1}\right)$. The proof technique is similar to that of Lemma A.4. The probability distribution of the involved variables is

$$
p(W_i, W_j, Z_{jj}^{(t-1)}, \bar{V}_j^{t-1}, Y_{ji}^{(t-1)}) = p(W_i) p(W_j) \prod_{\tau=1}^{t-1} p_S(Z_{jj,(\tau)})
$$

$$
\begin{aligned}
&p_S(\bar{V}_{j,\tau}, Y_{ji,(\tau)} | \bar{V}_j^{\tau-1}, Y_{ji}^{(\tau-1)}, Z_{jj}^{(\tau-1)}, W_i, W_j) \\
&\overset{\text{(a)}}{=} p(W_i)p(W_j) \prod_{\tau=1}^{t-1} p_S(Z_{jj,(\tau)}) \\
&\quad p_S(\bar{V}_{j,\tau} | Y_{ji}^{(\tau-1)}, Z_{jj}^{(\tau-1)}, W_j) p_S(Y_{ji,(\tau)} | W_i, \bar{V}_j^{t-1})
\end{aligned}
$$

where (a) follows by the fact that

$$
\begin{aligned}
\bar{V}_{j,\tau} &\overset{\text{f}}{=} \left(W_j, Y_{ji}^{(\tau-1)}, Z_{jj}^{(t-1)}, \underline{S}^{t-1}, Z_{i,\tau} \right), \\
Y_{ji,(\tau)} &\overset{\text{f}}{=} \left(W_i, \bar{V}_j^{t-1}, \underline{S}^{t-1}, Z_{ji,(\tau)} \right)
\end{aligned}
$$

and that $Z_{i,\tau}$ and $Z_{ji,(\tau)}$ are independent of everything else. Then the result follows by defining

$$
\begin{aligned}
f(W_i, \bar{V}_j^{t-1}, Y_{ji}^{(t-1)}) &= p(W_i) \prod_{\tau=1}^{t-1} p_S(Y_{ji,(\tau)} | W_i, \bar{V}_j^{t-1}) \\
g(W_j, Z_{jj}^{(t-1)}, \bar{V}_j^{t-1}, Y_{ji}^{(t-1)}) &= p(W_j) \prod_{\tau=1}^{t-1} p_S(Z_{jj,(\tau)}) p_S(\bar{V}_{j,\tau} | Y_{ji}^{(\tau-1)}, Z_{jj}^{(\tau-1)}, W_i)
\end{aligned}
$$

and noting that the above probability distribution factorizes as $f \cdot g$. □

Lemma A.6.

$$
\begin{aligned}
\mathbb{E}_{S^N} \left[h_S \left(\breve{Y}_i^N, \breve{V}_i^N | W_i \right) \right] &= \mathbb{E}_{S^N} \left[h_S \left(\breve{V}_j^N, Z_{ii}^{M_i}, Z_j^{L_j}, Z_{ji}^{M_j} | W_i \right) \right] \\
\mathbb{E}_{S^N} \left[h_S \left(\breve{Y}_i^N, \bar{V}_i^N | W_i \right) \right] &= \mathbb{E}_{S^N} \left[h_S \left(\breve{V}_j^N, Z_{ii}^{M_i}, Z_j^N | W_i \right) \right]
\end{aligned}
$$

for $(i, j) = (1, 2), (2, 1)$.

Proof.

$$
\begin{aligned}
&\mathbb{E}_{S^N} \left[h_S \left(\breve{Y}_i^N, \breve{V}_i^N | W_i \right) \right] = \mathbb{E}_{S^N} \left[\sum_{t=1}^{N} h_S \left(\breve{Y}_{i,t}, \breve{V}_{i,t} | W_i, \breve{Y}_i^{t-1}, \breve{V}_i^{t-1} \right) \right] \\
&\overset{\text{(a)}}{=} \mathbb{E}_{S^N} \left[\sum_{t=1}^{N} h_S \left(\breve{Y}_{i,t}, \breve{V}_{i,t} | W_i, \breve{Y}_i^{t-1}, \breve{V}_i^{t-1}, X_i^t \right) \right] \\
&= \mathbb{E}_{S^N} \left[\sum_{t=1}^{N} h_S \left(\bar{Y}_{i,[t]}, Y_{ii,(t)}, Y_{ij,(t)}, \bar{V}_{i,[t]}, Y_{ji,(t)} \right. \right.
\end{aligned}
$$

$$\Big| W_i, \bar{Y}_i^{[t-1]}, Y_{ii}^{(t-1)}, Y_{ij}^{(t-1)}, \bar{V}_i^{[t-1]}, Y_{ji}^{(t-1)}, X_i^t \Big) \Bigg]$$

$$= \mathbb{E}_{S^N} \left[\sum_{t=1}^{N} h_S \left(\bar{V}_{j,[t]}, Z_{ii,(t)}, Y_{ij,(t)}, Z_{j,[t]}, Z_{ji,(t)} \right.\right.$$

$$\left.\left. \Big| W_i, \bar{V}_j^{[t-1]}, Z_{ii}^{(t-1)}, Y_{ij}^{(t-1)}, Z_j^{[t-1]}, Z_{ji}^{(t-1)} \Big) \right]$$

$$= \mathbb{E}_{S^N} \left[h_S \left(\bar{V}_j^{L_i}, Z_{ii}^{M_i}, Y_{ij}^{M_i}, Z_j^{L_j}, Z_{ji}^{M_j} | W_i \right) \right]$$

$$\overset{\text{(b)}}{=} \mathbb{E}_{S^N} \left[h_S \left(\breve{V}_j^{N}, Z_{ii}^{M_i}, Z_j^{L_j}, Z_{ji}^{M_j} | W_i \right) \right]$$

where (a) is because $X_i^t \overset{\text{f}}{=} \left(W_i, \breve{Y}_i^{t-1}, \mathcal{S}^{t-1} \right)$, and (b) is because $\breve{V}_j^N = \left(\bar{V}_j^{L_i}, Y_{ij}^{M_i} \right)$. The second equality can be proved using similar steps. □

Lemma A.7.

$$\mathbb{E}_{S^N} \left[h_S \left(\bar{Y}_i^{L_i} | \breve{V}_i^N \right) \right] \leq a$$

$$\mathbb{E}_{S^N} \left[h_S \left(\bar{Y}_i^{L_i} | \bar{V}_i^N \right) \right] \leq a$$

for $(i, j) = (1, 2), (2, 1)$, *where*

$$a = N(1 - p_i) \log 2\pi e \left(1 + \mathsf{INR}_i + \frac{\mathsf{SNR}_i + 2\sqrt{\mathsf{SNR}_i \cdot \mathsf{INR}_i}}{1 + \mathsf{INR}_j} \right)$$

Proof.

$$\mathbb{E}_{S^N} \left[h_S \left(\bar{Y}_i^{L_i} | \breve{V}_i^N \right) \right] \leq \mathbb{E}_{S^N} \left[\sum_{t : S_{i,t} = 0} h_S \left(\bar{Y}_{i,t} | \breve{V}_{i,t} \right) \right]$$

$$\overset{\text{(a)}}{=} \mathbb{E}_{S^N} \left[L_i \frac{1}{L_i} \sum_{t : S_{i,t} = 0} h_S \left(\bar{Y}_{i,Q} | \breve{V}_{i,Q}, Q = t \right) \right]$$

$$\overset{\text{(b)}}{=} \mathbb{E}_{S^N} \left[L_i h_S \left(\bar{Y}_i | \breve{V}_i, Q \right) \right]$$

$$= \mathbb{E}_{S^N} \left[L_i \left(p_j h_S \left(\bar{Y}_i | Y_{ji}, Q \right) + (1 - p_j) h_S \left(\bar{Y}_i | \bar{V}_i, Q \right) \right) \right]$$

$$\overset{\text{(c)}}{\leq} \mathbb{E} \left[L_i p_j \log 2\pi e \left(1 + \mathsf{INR}_i + \frac{\mathsf{SNR}_i + 2\sqrt{\mathsf{SNR}_i \cdot \mathsf{INR}_i}}{1 + 2\mathsf{INR}_j} \right) \right.$$

$$\left. + L_i (1 - p_j) \log 2\pi e \left(1 + \mathsf{INR}_i + \frac{\mathsf{SNR}_i + 2\sqrt{\mathsf{SNR}_i \cdot \mathsf{INR}_i}}{1 + \mathsf{INR}_j} \right) \right]$$

$$\leq (1-p_i)\log 2\pi e\left(1+\mathsf{INR}_i+\frac{\mathsf{SNR}_i+2\sqrt{\mathsf{SNR}_i\cdot\mathsf{INR}_i}}{1+\mathsf{INR}_j}\right)$$

where (a) is by introducing a time-sharing random variable Q with uniform distribution over the set $\{t : S_{i,t}=0\}$, and (b) follows by setting $\bar{Y}_i=\bar{Y}_{i,Q}$ and $\breve{V}_i=\breve{V}_{i,Q}$. (c) follows by the fact that choosing jointly Gaussian input distribution with correlation coefficient $\rho=0$ for $p(x_1,x_2)$ maximizes the given conditional differential entropy.

By following similar steps, we can show that

$$\begin{aligned}\mathbb{E}_{S^N}\left[h_S\left(\bar{Y}_i^{L_i}|\bar{V}_i^N\right)\right] &\leq \mathbb{E}_{S^N}\left[L_i h_S\left(\bar{Y}_i|\bar{V}_i,Q\right)\right]\\ &\leq (1-p_i)\log 2\pi e\left(1+\mathsf{INR}_i+\frac{\mathsf{SNR}_i+2\sqrt{\mathsf{SNR}_i\cdot\mathsf{INR}_i}}{1+\mathsf{INR}_j}\right)\end{aligned}$$

□

Lemma A.8.

$$\mathbb{E}_{S^N}\left[h_S\left(Y_{ii}^{M_i}|\breve{V}_i^N\right)\right]\leq b$$
$$\mathbb{E}_{S^N}\left[h_S\left(Y_{ii}^{M_i}|\bar{V}_i^N\right)\right]\leq b$$

where

$$b=Np_i\log 2\pi e\left(\frac{1}{2}+\frac{\mathsf{SNR}_i}{2\mathsf{INR}_j+1}\right)$$

for $(i,j)=(1,2),(2,1)$.

Proof.

$$\begin{aligned}\mathbb{E}_{S^N}\left[h_S\left(Y_{ii}^{M_i}|\breve{V}_i^N\right)\right] &\leq \mathbb{E}_{S^N}\left[\sum_{t:S_{i,t}=1} h_S\left(Y_{ii,t}|\breve{V}_{i,t}\right)\right]\\ &\overset{\text{(a)}}{=} \mathbb{E}_{S^N}\left[M_i\frac{1}{M_i}\sum_{t:S_{i,t}=1} h_S\left(Y_{ii,Q}|V_{i,Q},Q=t\right)\right]\\ &\overset{\text{(b)}}{=} \mathbb{E}_{S^N}\left[M_i h_S\left(Y_{ii}|\breve{V}_i,Q\right)\right]\\ &= \mathbb{E}_{S^N}\left[M_i\left((1-p_j)h_S\left(Y_{ii}|\bar{V}_i,Q\right)+p_j h_S\left(Y_{ii}|Y_{ji},Q\right)\right)\right]\end{aligned}$$

$$
\begin{aligned}
&\overset{\text{(c)}}{\le} \mathbb{E}_{S^N}\left[M_i\left((1-p_j)\log 2\pi e\left(\frac{1}{2}+\frac{2\mathsf{SNR}_i+\frac{1}{2}}{2\mathsf{INR}_j+1}\right)+p_j\log\left(\pi e+\frac{2\pi e\cdot\mathsf{SNR}_i}{2\mathsf{INR}_j+1}\right)\right)\right]\\
&\le p_i\log 2\pi e\left(\frac{1}{2}+\frac{\mathsf{SNR}_i}{2\mathsf{INR}_j+1}\right)
\end{aligned}
$$

where (a) is by introducing a time-sharing random variable Q with uniform distribution over the set $\{t : S_{i,t}=1\}$, and (b) follows by setting $Y_{ii}=Y_{ii,Q}$ and $\breve{V}_i=\breve{V}_{i,Q}$. (c) follows by the fact that choosing jointly Gaussian input distribution with correlation coefficient $\rho=0$ for $p(x_1,x_2)$ maximizes the given conditional differential entropy. Similarly,

$$
\mathbb{E}_{S^N}\left[h_S\left(Y_{ii}^{M_i}|\bar{V}_i^N\right)\right]\le\mathbb{E}_{S^N}\left[M_i h_S\left(Y_{ii}|\bar{V}_i,Q\right)\right]\le p_i\log 2\pi e\left(\frac{1}{2}+\frac{\mathsf{SNR}_i}{2\mathsf{INR}_j+1}\right)
$$

□

Lemma A.9.

$$
\begin{aligned}
\mathbb{E}_{S^N}\left[\sum_{t=1}^{N}h_S\left(\breve{Y}_{i,t}|\breve{V}_{i,t},X_{j,t}\right)\right]&\le p_i\log 2\pi e\left(\frac{1}{2}+\frac{\mathsf{SNR}_i}{2\mathsf{INR}_j+1}\right)\\
&\quad+p_i\log 2\pi e\frac{1}{2}+(1-p_i)\log 2\pi e\left(1+\frac{\mathsf{SNR}_i}{1+\mathsf{INR}_j}\right)
\end{aligned}
$$

for $(i,j)=(1,2),(2,1)$.

Proof.

$$
\begin{aligned}
\mathbb{E}_{S^N}\left[\sum_{t=1}^{N}h_S\left(\breve{Y}_{i,t}|\breve{V}_{i,t},X_{j,t}\right)\right]&\overset{\text{(a)}}{=}\mathbb{E}_{S^N}\left[N\frac{1}{N}\sum_{t=1}^{N}h_S\left(\breve{Y}_{i,Q}|\breve{V}_{i,Q},X_{j,Q},Q=t\right)\right]\\
&=\mathbb{E}_{S^N}\left[Nh_S\left(\breve{Y}_{i,Q}|\breve{V}_{i,Q},X_{j,Q},Q\right)\right]\\
&\overset{\text{(b)}}{=}Nh(\breve{Y}_i|\breve{V}_i,X_j,Q)\\
&=p_ih(Y_{ii},Y_{ij}|\breve{V}_i,X_j,Q)+(1-p_i)h(\bar{Y}_i|\breve{V}_i,X_j,Q)\\
&\le p_ih(Y_{ii}|V_i)+p_ih(Y_{ij}|X_j)\\
&\quad+(1-p_i)\left[(1-p_j)h(\bar{Y}_i|\bar{V}_i,X_j)+p_jh(\bar{Y}_i|Y_{ji},X_j)\right]\\
&\overset{\text{(c)}}{\le}p_i\log 2\pi e\left(\frac{1}{2}+\frac{\mathsf{SNR}_i}{2\mathsf{INR}_j+1}\right)+p_i\log 2\pi e\frac{1}{2}\\
&\quad+(1-p_i)\log 2\pi e\left(1+\frac{\mathsf{SNR}_i}{1+\mathsf{INR}_j}\right)
\end{aligned}
$$

where (a) is by introducing a uniformly distributed time-sharing random variable Q, and (b) is by defining $\breve{Y}_i = \breve{Y}_{i,Q}$, $\breve{V}_i = \breve{V}_{i,Q}$ and $X_j = X_{j,Q}$. (c) follows by the fact that choosing jointly Gaussian input distribution with correlation coefficient $\rho = 0$ for $p(x_1, x_2)$ maximizes the given conditional differential entropy. □

A.6 Gap Analysis

In this section, we give upper bounds for the gap terms δ_1 and δ_2 from Theorem 2.2. We also compare our achievable region with the outer bound of [ST11] for the case $p_1 = p_2 = 1$.

A.6.1 Bounding δ_1

We will show that, each of the bounds (A.18), (A.19), and (A.20) are within a constant gap of the region given in (2.13)–(2.17). Without loss of generality, we focus on the case $(i,j) = (1,2)$, and start with the first bound in (A.18).

$$\begin{aligned}
\mathsf{A}_1 + \mathsf{B}_2 &= \log\left(3 + \frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right) + \log\left(2 + \mathsf{INR}_2\right) - 2\log 3 - C_1 - C_2 \\
&\geq \log\left(1 + \frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right) + \log\left(1 + \mathsf{INR}_2\right) - 2\log 3 - C_1 - C_2 \\
&= \log\left(1 + \mathsf{SNR}_1 + \mathsf{INR}_2\right) - 2\log 3 - C_1 - C_2 \\
&= \log\left(1 + \mathsf{SNR}_1\right) + \log\left(1 + \frac{\mathsf{INR}_2}{1+\mathsf{SNR}_1}\right) - 2\log 3 - C_1 - C_2 \\
&\geq (2.14)_{(1,2)} - 2\log 3 - C_1 - C_2
\end{aligned}$$

where $(2.13)_{(1,2)}$ refers to bound in (2.14), evaluated with $(i,j) = (1,2)$. Next, we consider the second bound in (A.18). If $\mathsf{SNR}_1 \geq \mathsf{INR}_1$,

$$\begin{aligned}
\mathsf{D}_1 &= \log\left(3 + \mathsf{SNR}_1\right) - \log 3 - C_1 \\
&\geq \log\left(1 + \mathsf{SNR}_1\right) - \log 3 - C_1 \\
&\geq \log\left(1 + \mathsf{SNR}_1 + \mathsf{INR}_1\right) - \log 3 - C_1 - 1 \\
&= (2.13)_{(1,2),L} - \log 3 - C_1 - 1
\end{aligned}$$

where $(2.13)_{(1,2)}$ refers to bound in (2.13), evaluated with $(i,j)=(1,2)$. If $\mathsf{SNR}_1 < \mathsf{INR}_1$,

$$\begin{aligned}\mathsf{D}_1 &= \log\left(3+\mathsf{SNR}_1\right) + p_2 \log\left(1+\frac{\mathsf{INR}_2}{3+\mathsf{SNR}_1}\right) - p_2\log\frac{5}{3} - \log 3 - C_1\\ &\geq \log\left(1+\mathsf{SNR}_1\right) + p_2 \log\left(1+\frac{\mathsf{INR}_2}{1+\mathsf{SNR}_1}\right) - p_2\log\frac{5}{3} - p_2\log 3 - \log 3 - C_1\\ &= (2.13)_{(1,2),R} - p_2\log 5 - \log 3 - C_1\end{aligned}$$

Next, we consider the first bound in (A.19).

$$\begin{aligned}\mathsf{A}_1 + \mathsf{G}_2 &= \log\left(3+\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right) + \log\left(2+\mathsf{SNR}_2+\mathsf{INR}_2\right) - 2\log 3 - C_1 - C_2 - \kappa_1\\ &\geq \log\left(1+\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right) + \log\left(1+\mathsf{SNR}_2+\mathsf{INR}_2\right) - 2\log 3 - C_1 - C_2 - \kappa_1\\ &= (2.15)_{(1,2)} - 2\log 3 - C_1 - C_2 - \kappa_1\end{aligned}$$

where $(2.15)_{(1,2)}$ refers to the bound (2.15) evaluated with $(i,j)=(1,2)$. For the bound $\mathsf{F}_1+\mathsf{F}_2$, we first consider the case when $\mathsf{INR}_1 > \mathsf{SNR}_1$.

$$\begin{aligned}\mathsf{F}_1 + \mathsf{F}_2 &\geq \log\left(2+\mathsf{INR}_1+\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right) + \log\left(2+\mathsf{INR}_2+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}\right)\\ &\quad - 2\log 3 - C_1 - C_2\\ &\geq \log\left(1+\mathsf{INR}_1\right) + \log\left(1+\mathsf{INR}_2+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}\right) - 2\log 3 - C_1 - C_2\\ &\geq \log\left(1+\mathsf{SNR}_1+\mathsf{INR}_1\right) + \log\left(1+\mathsf{INR}_2+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}\right) - 2\log 3 - C_1 - C_2 - 1\\ &= (2.15)_{(2,1)} - 2\log 3 - C_1 - C_2 - 1\end{aligned}$$

By symmetry, we can show that when $\mathsf{INR}_2 > \mathsf{SNR}_2$,

$$\mathsf{F}_1 + \mathsf{F}_2 \geq (2.15)_{(1,2)} - 2\log 3 - C_1 - C_2 - 1$$

Next we consider the only remaining case of $\mathsf{INR}_1 \leq \mathsf{SNR}_1$, $\mathsf{INR}_2 \leq \mathsf{SNR}_2$.

$$\begin{aligned}\mathsf{F}_1 + \mathsf{F}_2 &= \log\left(2+\mathsf{INR}_1+\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right) + \log\left(2+\mathsf{INR}_2+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}\right)\\ &\quad + p_1\log\left(\frac{\left(2+\mathsf{INR}_1\right)\left(3+\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right)}{2+\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}+\mathsf{INR}_1}\right)\end{aligned}$$

$$
\begin{aligned}
&+ p_2 \log\left(\frac{(2+\mathsf{INR}_2)\left(3+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}\right)}{2+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}+\mathsf{INR}_2}\right) - 2\log 3 - 2C_1 - 2C_2 - (p_1+p_2)\log 6 \\
&\overset{(a)}{\geq} \log\left(1+\mathsf{INR}_1+\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right) + \log\left(1+\mathsf{INR}_2+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}\right) \\
&+ p_1 \log\left(\frac{(1+\mathsf{INR}_1)\left(1+\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right)}{1+\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}+\mathsf{INR}_1}\right) \\
&+ p_2 \log\left(\frac{(1+\mathsf{INR}_2)\left(1+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}\right)}{1+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}+\mathsf{INR}_2}\right) - 2\log 3 - 2C_1 - 2C_2 - (p_1+p_2)\log 6 \\
&= (2.16) - 2\log 3 - 2C_1 - 2C_2 - (p_1+p_2)\log 6
\end{aligned}
$$

where in (a), we used the fact that the function $\log\left(\frac{x+a}{x+a+b}\right)$ is monotonically increasing in x, for $x, a, b > 0$. Finally, we consider the bound (A.20). Again, we distinguish the cases $\mathsf{INR}_2 > \mathsf{SNR}_2$ and $\mathsf{INR}_2 \leq \mathsf{SNR}_2$. For the former case,

$$
\begin{aligned}
\mathsf{A}_1 + \mathsf{F}_2 + \mathsf{G}_1 &= \log\left(3+\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right) + \log\left(1+\mathsf{INR}_2+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}\right) \\
&\quad + \log(2+\mathsf{SNR}_1+\mathsf{INR}_1) - 3\log 3 - 2C_1 - C_2 - \kappa_2 \\
&\geq \log\left(3+\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right) + \log(1+\mathsf{INR}_2) + \log(2+\mathsf{SNR}_1+\mathsf{INR}_1) \\
&\quad - 3\log 3 - 2C_1 - C_2 - \kappa_2 \\
&\geq \log\left(3+\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right) + \log(1+\mathsf{INR}_2+\mathsf{SNR}_2) + \log(2+\mathsf{SNR}_1+\mathsf{INR}_1) \\
&\quad - 3\log 3 - 2C_1 - C_2 - \kappa_2 - 1 \\
&\geq \log\left(1+\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right) + \log(1+\mathsf{INR}_2+\mathsf{SNR}_2) + \log(1+\mathsf{SNR}_1+\mathsf{INR}_1) \\
&\quad - 3\log 3 - 2C_1 - C_2 - \kappa_2 - 1 \\
&= (2.15)_{(1,2)} + (2.13)_{(1,2)} - 3\log 3 - 2C_1 - C_2 - \kappa_2 - 1
\end{aligned}
$$

For the case $\mathsf{INR}_2 \leq \mathsf{SNR}_2$,

$$
\begin{aligned}
\mathsf{A}_1 + \mathsf{F}_2 + \mathsf{G}_1 &= \log\left(3+\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right) + \log\left(1+\mathsf{INR}_2+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}\right) \\
&\quad + \log(2+\mathsf{SNR}_1+\mathsf{INR}_1)
\end{aligned}
$$

$$
\begin{aligned}
&+ p_2 \log\left(\frac{(2+\mathsf{INR}_2)\left(3+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}\right)}{2+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}+\mathsf{INR}_2}\right) - 3\log 3 - 2C_1 - C_2 - p_2\log 6 - \kappa_2 \\
&\overset{(a)}{\geq} \log\left(1+\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right) + \log\left(1+\mathsf{INR}_2+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}\right) + \log\left(1+\mathsf{SNR}_1+\mathsf{INR}_1\right) \\
&+ p_2 \log\left(\frac{(1+\mathsf{INR}_2)\left(1+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}\right)}{1+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}+\mathsf{INR}_2}\right) - 3\log 3 - 2C_1 - C_2 - p_2\log 6 - \kappa_2 \\
&= (2.17) - 3\log 3 - 2C_1 - 2C_2 - p_2\log 6 - \kappa_2
\end{aligned}
$$

where in (a), as before, we used the fact that the function $\log\left(\frac{x+a}{x+a+b}\right)$ is monotonically increasing in x, for $x, a, b > 0$. By symmetry, similar gaps apply to the case $(i, j) = (2, 1)$. Now, we can upper bound δ_1 by noting that it cannot be larger than the maximum of the gaps found above (after proper normalization, *e.g.*, the gap found for the bound on $R_1 + R_2$ is divided by 2, and the one on $2R_1 + R_2$ is divided by 3). Hence, using the fact that $C_i = 2p_j + p_i$ and $\kappa_i = p_i$, we find

$$\delta_1 < 2\log 3 + 3\left(p_1 + p_2\right) \text{ bits.}$$

A.6.2 Bounding δ_2

In order to bound δ_2, we compare the bounds obtained in Section 2.7 with the bounds (2.13)–(2.17) one by one. Without loss of generality, we focus on $(i, j) = (1, 2)$, and begin with the bound in (2.13).

$$
\begin{aligned}
(2.13)_{(1,2)} &= \log\left(1+\mathsf{SNR}_1+\mathsf{INR}_1\right) \\
&\geq \log\left(1+\mathsf{SNR}_1+\mathsf{INR}_1+2\sqrt{\mathsf{SNR}_1\cdot\mathsf{INR}_1}\right) - \log 3 \\
&\geq (2.53) - \log 3
\end{aligned}
$$

Next, we consider (2.14), and note that (2.55) = (2.14). We now consider the bound (2.15).

$$
\begin{aligned}
(2.15) &= \log\left(1+\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right) + \log\left(1+\mathsf{SNR}_2+\mathsf{INR}_2\right) \\
&\geq \log\left(1+\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right) + \log\left(1+\mathsf{SNR}_2+\mathsf{INR}_2+2\sqrt{\mathsf{SNR}_2\cdot\mathsf{INR}_2}\right) - \log 3
\end{aligned}
$$

$$\geq (2.57) - \log 3$$

Let us take (2.16).

$$\begin{aligned}
(2.16) &= \log\left(1+\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}+\mathsf{INR}_1\right)+\log\left(1+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}+\mathsf{INR}_2\right)\\
&\quad+p_1\log\left(\frac{(1+\mathsf{INR}_1)\left(1+\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right)}{1+\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}+\mathsf{INR}_1}\right)+p_2\log\left(\frac{(1+\mathsf{INR}_2)\left(1+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}\right)}{1+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}+\mathsf{INR}_2}\right)\\
&\geq \log\left(1+\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}+\mathsf{INR}_1+2\sqrt{\mathsf{SNR}_1\cdot\mathsf{INR}_1}\right)\\
&\quad+\log\left(1+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}+\mathsf{INR}_2+2\sqrt{\mathsf{SNR}_2\cdot\mathsf{INR}_2}\right)\\
&\quad+p_1\log\left(\frac{(1+\mathsf{INR}_1)\left(1+\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right)}{1+\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}+\mathsf{INR}_1}\right)+p_2\log\left(\frac{(1+\mathsf{INR}_2)\left(1+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}\right)}{1+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}+\mathsf{INR}_2}\right)\\
&\quad-2\log 3\\
&\geq \log\left(1+\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}+\mathsf{INR}_1+2\sqrt{\mathsf{SNR}_1\cdot\mathsf{INR}_1}\right)\\
&\quad+\log\left(1+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}+\mathsf{INR}_2+2\sqrt{\mathsf{SNR}_2\cdot\mathsf{INR}_2}\right)\\
&\quad+p_1\log\left(\frac{(1+\mathsf{INR}_1)\left(1+\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right)}{1+\frac{\mathsf{SNR}_1+2\sqrt{\mathsf{SNR}_1\cdot\mathsf{INR}_1}}{1+\mathsf{INR}_2}+\mathsf{INR}_1}\right)\\
&\quad+p_2\log\left(\frac{(1+\mathsf{INR}_2)\left(1+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}\right)}{1+\frac{\mathsf{SNR}_2+2\sqrt{\mathsf{SNR}_2\cdot\mathsf{INR}_2}}{1+\mathsf{INR}_1}+\mathsf{INR}_2}\right)-2\log 3\\
&\geq (2.61)-2\log 3-2p_1-2p_2
\end{aligned}$$

Finally, we consider (2.17)

$$\begin{aligned}
(2.17) &= \log\left(1+\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right)+\log\left(1+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}+\mathsf{INR}_2\right)\\
&\quad+\log\left(1+\mathsf{SNR}_1+\mathsf{INR}_1\right)+p_2\log\left(\frac{(1+\mathsf{INR}_2)\left(1+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}\right)}{1+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}+\mathsf{INR}_2}\right)\\
&\geq \log\left(1+\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right)+\log\left(1+\frac{\mathsf{SNR}_2+2\sqrt{\mathsf{SNR}_2\cdot\mathsf{INR}_2}}{1+\mathsf{INR}_1}+\mathsf{INR}_2\right)
\end{aligned}$$

$$
\begin{aligned}
&+\log\left(1+\mathsf{SNR}_1+\mathsf{INR}_1+2\sqrt{\mathsf{SNR}_1\cdot\mathsf{INR}_1}\right)\\
&+p_2\log\left(\frac{(1+\mathsf{INR}_2)\left(1+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}\right)}{1+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}+\mathsf{INR}_2}\right)-2\log 3\\
\geq{}&\log\left(1+\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right)+\log\left(1+\frac{\mathsf{SNR}_2+2\sqrt{\mathsf{SNR}_2\cdot\mathsf{INR}_2}}{1+\mathsf{INR}_1}+\mathsf{INR}_2\right)\\
&+\log\left(1+\mathsf{SNR}_1+\mathsf{INR}_1+2\sqrt{\mathsf{SNR}_1\cdot\mathsf{INR}_1}\right)\\
&+p_2\log\left(\frac{(1+\mathsf{INR}_2)\left(1+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}\right)}{1+\frac{\mathsf{SNR}_2+2\sqrt{\mathsf{SNR}_2\cdot\mathsf{INR}_2}}{1+\mathsf{INR}_1}+\mathsf{INR}_2}\right)-2\log 3\\
\geq{}&(2.62)-2\log 3-1-2p_2
\end{aligned}
$$

In order to bound δ_2, we note that it cannot be larger than the maximum of the gaps found above, after normalization as done with bounding δ_1. Hence, we find

$$\delta_2 < \log 3 + p_1 + p_2 \text{ bits.}$$

A.6.3 Comparison with Suh-Tse outer bound

In this subsection, we compare our inner bound for the case $p_1 = p_2 = 1$ with the perfect feedback outer bound of [ST11]. Looking at the region (2.13)–(2.17), we see that if we set $p_1 = p_2 = 1$, then the bounds (2.16) and (2.17) become redundant, and the region reduces to the outer bound region of [ST11], with the following differences:

- The outer bounds in [ST11] are parameterized by the parameter ρ, which captures the correlation between the symbols of two users. In the region (2.13)–(2.17), supremum values over all possible values of ρ is given.
- The bounds in (2.13)–(2.17) terms $2\rho\sqrt{\mathsf{SNR}_i\cdot\mathsf{INR}_i}$ that arise from beamforming gain, which appear in the outer bounds of [ST11].

It is easy to see that the first item does not result in a rate penalty, while the second one gives a penalty of $\log 3$ since

$$\begin{aligned}(2.13) &= \log\left(1+\mathsf{SNR}_1+\mathsf{INR}_1\right)\\ &\geq \log\left(1+\mathsf{SNR}_1+\mathsf{INR}_1+2\sqrt{\mathsf{SNR}_1\cdot\mathsf{INR}_1}\right)-\log 3\end{aligned}$$

Hence, the region $\bar{\mathcal{C}}(1,1)$ is within at most $\log 3$ bits of the outer bound region of [ST11]. From the results of this section, we also know that our scheme achieves the region $\bar{\mathcal{C}}(1,1)-\delta_1$. Evaluating δ_1 for $p_1=p_2=1$, we see that the proposed scheme achieves within $3+3\log 3\approx 7.75$ bits of the Suh-Tse outer bound region.

A.7 Proofs of Corollaries 2.1 and 2.3

In this section, we prove that when the feedback probabilities are sufficiently high, perfect feedback sum-capacity can be achieved (approximately for Gaussian case, exactly for linear deterministic case). The precise statements for the two models are given in Corollaries 2.1 and 2.3.

A.7.1 Proof of Corollary 2.1

We will show that when feedback is perfect, the bounds on the sum rate that involve the feedback probabilities become strictly redundant. That is, setting $p_1=p_2=p$, we will prove that the bounds (2.9), (2.5) + (2.6), $\frac{(2.5)+(2.11)}{2}$, $\frac{(2.6)+(2.10)}{2}$, and $\frac{(2.10)+(2.11)}{3}$ are all strictly larger than the perfect feedback bounds (2.7)–(2.8) when $p=1$ and $n_{12},n_{21}>0$. Then the result follows by noting that all such bounds are continuous and monotonically increasing functions of p, and hence there must exist a $p^*<1$ such that whenever $p=p^*$, perfect feedback sum-rate bounds (2.7)–(2.8) are exactly matched by these bounds.

We first prove a claim that will be used in the main proof.

Claim A.9. *For $n_{21},n_{12}>0$,*

$$(2.7)<n_{12}+n_{21}+\left(n_{11}-n_{21}\right)^+ +\left(n_{22}-n_{12}\right)^+$$

$$(2.8) < n_{12} + n_{21} + (n_{22} - n_{12})^+ + (n_{11} - n_{21})^+$$

Proof. By symmetry, we only prove the first statement.

$$\begin{aligned}
&n_{12} + n_{21} + (n_{11} - n_{21})^+ + (n_{22} - n_{12})^+ \\
&\quad = n_{12} + \max(n_{11}, n_{21}) + (n_{22} - n_{12})^+ \\
&\quad = \max(n_{12} + n_{11}, n_{12} + n_{21}) + (n_{22} - n_{12})^+ \\
&\quad > \max(n_{11}, n_{12}) + (n_{22} - n_{12})^+ \geq \min\{(2.7), (2.8)\}
\end{aligned}$$

where the strict inequality follows by the fact that $n_{21}, n_{12} > 0$. □

Next, we consider the bound (2.9).

$$\begin{aligned}
(2.9) &= \max\left\{n_{12}, (n_{11} - n_{21})^+\right\} + \max\left\{n_{21}, (n_{22} - n_{12})^+\right\} \\
&\quad + \min\left\{n_{12}, (n_{11} - n_{21})^+\right\} + \min\left\{n_{21}, (n_{22} - n_{12})^+\right\} \\
&= n_{12} + n_{21} + (n_{11} - n_{21})^+ + (n_{22} - n_{12})^+ \\
&> \min\{(2.7), (2.8)\}
\end{aligned}$$

where the last line follows by Claim A.9. Hence, the bound (2.9) becomes strictly redundant. Next, consider

$$\begin{aligned}
&(2.10) + (2.11) = \max(n_{11}, n_{12}) + \max(n_{22}, n_{21}) + (n_{11} - n_{21})^+ + (n_{22} - n_{12})^+ \\
&\quad + \max\left\{n_{12}, (n_{11} - n_{21})^+\right\} + \max\left\{n_{21}, (n_{22} - n_{12})^+\right\} \\
&\quad + \min\left\{n_{12}, (n_{11} - n_{21})^+\right\} + \min\left\{n_{21}, (n_{22} - n_{12})^+\right\} \\
&= \max(n_{11}, n_{12}) + \max(n_{22}, n_{21}) \\
&\quad + 2(n_{11} - n_{21})^+ + 2(n_{22} - n_{12})^+ n_{12} + n_{21} \\
&\geq 2\min\left\{\max(n_{11}, n_{12}) + (n_{22} - n_{12})^+, \max(n_{22}, n_{21}) + (n_{11} - n_{21})^+\right\} \\
&\quad + (n_{11} - n_{21})^+ + (n_{22} - n_{12})^+ n_{12} + n_{21}
\end{aligned}$$

$$> 3 \cdot \min\{(2.7), (2.8)\}$$

where the last line follows by Claim A.9.

Next, we consider the sum of individual rate bounds. Since these bounds consist of the minimum of two terms, we consider each case separately. In what follows, $(2.5)_R$ denotes the term on the right-hand side of the minimization in (2.5), while $(2.5)_L$ denotes the term on the left-hand side ($(2.6)_R$ and $(2.6)_L$ are also defined similarly). By symmetry, it is sufficient to prove that $(2.5)_R + (2.6)_R$ and $(2.5)_L + (2.6)_R$ are strictly redundant for $p_1 = p_2 = 1$. We show this as follows.

$$\begin{aligned}
&(2.5)_R + (2.6)_R = n_{11} + n_{22} + (n_{21} - n_{11})^+ + (n_{12} - n_{22})^+ \\
&= \max(n_{11}, n_{21}) + \max(n_{22}, n_{12}) \\
&= n_{12} + n_{21} + (n_{11} - n_{21})^+ + (n_{22} - n_{12})^+ > \min\{(2.7), (2.8)\}
\end{aligned}$$

by Claim A.9, and

$$\begin{aligned}
&(2.5)_L + (2.6)_R = \max(n_{11}, n_{12}) + n_{22} + (n_{12} - n_{22})^+ \\
&= \max(n_{11}, n_{12}) + \max(n_{22}, n_{12}) \\
&> \max(n_{11}, n_{12}) + (n_{22} - n_{12})^+ \, (2.7)
\end{aligned}$$

since $n_{12} > 0$.

Finally, we consider the bounds (2.10) + (2.6), and (2.11) + (2.5). By symmetry, it is sufficient to show the redundancy of $(2.10) + (2.6)_R$ and $(2.10) + (2.6)_L$. The former is shown by

$$\begin{aligned}
&(2.10) + (2.6)_R = \max(n_{11}, n_{12}) + (n_{11} - n_{21})^+ + \max\left\{n_{21}, (n_{22} - n_{12})^+\right\} \\
&\quad + \min\left\{n_{21}, (n_{22} - n_{12})^+\right\} + n_{22} + (n_{12} - n_{22})^+ \\
&= \max(n_{11}, n_{12}) + (n_{11} - n_{21})^+ + n_{21} + (n_{22} - n_{12})^+ + \max(n_{22}, n_{12}) \\
&= \max(n_{11}, n_{12}) + + (n_{22} - n_{12})^+ + n_{21} + n_{12} + (n_{11} - n_{21})^+ (n_{22} - n_{12})^+
\end{aligned}$$

$$
\begin{aligned}
&\geq (2.7) + n_{21} + n_{12} + (n_{11} - n_{21})^+ (n_{22} - n_{12})^+ \\
&> 2 \cdot (2.7)
\end{aligned}
$$

where the last line follows by Claim A.9, and

$$
\begin{aligned}
(2.10) + (2.6)_L &= \max (n_{11}, n_{12}) + (n_{11} - n_{21})^+ + \max \left\{n_{21}, (n_{22} - n_{12})^+\right\} \\
&\quad + \min \left\{n_{21}, (n_{22} - n_{12})^+\right\} + \max \{n_{22}, n_{21}\} \\
&= \max (n_{11}, n_{12}) + (n_{11} - n_{21})^+ + n_{21} + (n_{22} - n_{12})^+ + \max (n_{22}, n_{21}) \\
&> \max (n_{11}, n_{12}) + (n_{11} - n_{21})^+ + \max (n_{22}, n_{21}) + (n_{22} - n_{12})^+ \\
&\geq 2 \min \left\{\max (n_{11}, n_{12}) + (n_{22} - n_{12})^+, \max (n_{22}, n_{21}) + (n_{11} - n_{21})^+\right\} \\
&= 2 \cdot \min \{(2.7), (2.8)\}
\end{aligned}
$$

where the strict inequality follows by the fact that $n_{21} > 0$.

A.7.2 Proof of Corollary 2.3

Similar to the proof of Corollary 2.1, we will show that when feedback is perfect, the bounds on the sum rate that involve the feedback probabilities become redundant for the set $\bar{\mathcal{C}}(p_1, p_2)$. Since the capacity region is within a constant gap of the region $\bar{\mathcal{C}}(p_1, p_2)$, for all channel parameters, the result will follow.

Specifically, setting $p_1 = p_2 = p$, we will show that when $p = 1, \mathsf{INR}_1, \mathsf{INR}_2 > 0$, the bounds (2.16), $\min\left\{(2.13)_{(1,2)}, (2.14)_{(1,2)}\right\} + \min\left\{(2.13)_{(2,1)}, (2.14)_{(2,1)}\right\}$, $\frac{(2.13)_{(1,2)}+(2.17)_{(2,1)}}{2}$, $\frac{(2.13)_{(2,1)}+(2.17)_{(1,2)}}{2}$, and $\frac{(2.17)_{(1,2)}+(2.17)_{(2,1)}}{3}$ are all strictly larger than the perfect feedback bound (2.15), where the subscript (a, b) denotes the evaluation of the relevant bound with $(i, j) = (a, b)$.

We first prove a claim that will be useful in the proof of the corollary.

Claim A.10. *For* $\mathsf{INR}_1, \mathsf{INR}_2 > 0$,

$$
\min_{(i,j)=(1,2),(2,1)} (2.15) < \log (1 + \mathsf{INR}_1) + \log (1 + \mathsf{INR}_2)
$$

$$+\log\left(1+\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right)+\log\left(1+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}\right)$$

Proof.

$$\begin{aligned}
&\log(1+\mathsf{INR}_1)+\log(1+\mathsf{INR}_2)+\log\left(1+\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right)+\log\left(1+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}\right)\\
&=\log\left(1+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}\right)+\log(1+\mathsf{INR}_1+\mathsf{INR}_2+\mathsf{SNR}_1+\mathsf{INR}_1\mathsf{INR}_2+\mathsf{INR}_1\mathsf{SNR}_1)\\
&>\log(1+\mathsf{SNR}_1+\mathsf{INR}_1)+\log\left(1+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}\right)\\
&\geq\min_{(i,j)=(1,2),(2,1)}(2.15)
\end{aligned}$$

□

Next, we show that under the condition $\mathsf{INR}_1,\mathsf{INR}_2>0$ and $p=1$, all of the mentioned bounds are strictly redundant. We start with (2.16):

$$\begin{aligned}
(2.16)&=\log(1+\mathsf{INR}_1)+\log(1+\mathsf{INR}_2)\\
&\quad+\log\left(1+\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right)+\log\left(1+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}\right)\\
&>\min_{(i,j)=(1,2),(2,1)}(2.15)
\end{aligned}$$

by Claim A.10. Next,

$$\begin{aligned}
&(2.17)_{(1,2)}+(2.17)_{(2,1)}=\log(1+\mathsf{SNR}_1+\mathsf{INR}_1)+\log(1+\mathsf{SNR}_2+\mathsf{INR}_2)\\
&\quad+2\log\left(1+\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right)+2\log\left(1+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}\right)\\
&\quad+\log(1+\mathsf{INR}_1)+\log(1+\mathsf{INR}_2)\\
&=(2.15)_{(1,2)}+(2.15)_{(2,1)}+\log(1+\mathsf{INR}_i)+\log(1+\mathsf{INR}_j)\\
&\quad+\log\left(1+\frac{\mathsf{SNR}_i}{1+\mathsf{INR}_j}\right)+\log\left(1+\frac{\mathsf{SNR}_j}{1+\mathsf{INR}_i}\right)\\
&>2\cdot(2.15)_{(1,2)}+(2.15)_{(2,1)}\\
&\geq 3\cdot\min_{(i,j)=(1,2),(2,1)}(2.15)
\end{aligned}$$

by Claim A.10. Next, we consider the bounds $(2.14)_{(1,2)} + (2.14)_{(2,1)}$, $(2.13)_{(1,2)} + (2.14)_{(2,1)}$, and $(2.14)_{(1,2)} + (2.13)_{(2,1)}$. By symmetry, it is sufficient to show the redundancy of the former two.

$$\begin{aligned}(2.14)_{(1,2)} + (2.14)_{(2,1)} &= \log\left(1 + \mathsf{INR}_1\right) + \log\left(1 + \mathsf{INR}_2\right)\\ &\quad + \log\left(1 + \frac{\mathsf{SNR}_1}{1 + \mathsf{INR}_2}\right) + \log\left(1 + \frac{\mathsf{SNR}_2}{1 + \mathsf{INR}_1}\right)\\ &> \min_{(i,j)=(1,2),(2,1)} (2.15)\end{aligned}$$

by Claim A.10, and

$$\begin{aligned}&(2.13)_{(1,2)} + (2.14)_{(2,1)} = \log\left(1 + \mathsf{SNR}_1 + \mathsf{INR}_1\right) + \log\left(1 + \mathsf{SNR}_2 + \mathsf{INR}_1\right)\\ &> \log\left(1 + \mathsf{SNR}_1 + \mathsf{INR}_1\right) + \log\left(1 + \frac{\mathsf{SNR}_2}{1 + \mathsf{INR}_1}\right)\\ &\geq \min_{(i,j)=(1,2),(2,1)} (2.15)\end{aligned}$$

since $\mathsf{INR}_1, \mathsf{INR}_2 > 0$. Finally, we show the redundancy of $\frac{(2.13)_{(1,2)}+(2.17)_{(2,1)}}{2}$ and $\frac{(2.14)_{(1,2)}+(2.17)_{(2,1)}}{2}$. Then

$$\begin{aligned}&(2.13)_{(1,2)} + (2.17)_{(2,1)} = \log\left(1 + \mathsf{SNR}_1 + \mathsf{INR}_1\right) + \log\left(1 + \mathsf{SNR}_2 + \mathsf{INR}_2\right)\\ &\quad + \log\left(1 + \frac{\mathsf{SNR}_2}{1 + \mathsf{INR}_1}\right) + \log\left(1 + \mathsf{INR}_1\right) + \log\left(1 + \frac{\mathsf{SNR}_1}{1 + \mathsf{INR}_2}\right)\\ &> \log\left(1 + \mathsf{SNR}_1 + \mathsf{INR}_1\right) + \log\left(1 + \mathsf{SNR}_2 + \mathsf{INR}_2\right)\\ &\quad + \log\left(1 + \frac{\mathsf{SNR}_2}{1 + \mathsf{INR}_1}\right) + \log\left(1 + \frac{\mathsf{SNR}_1}{1 + \mathsf{INR}_2}\right)\\ &= (2.15)_{(1,2)} + (2.15)_{(2,1)}\\ &\geq \min_{(i,j)=(1,2),(2,1)} (2.15)\end{aligned}$$

and

$$\begin{aligned}&(2.14)_{(1,2)} + (2.17)_{(2,1)} = \log\left(1 + \mathsf{INR}_2\right) + \log\left(1 + \mathsf{SNR}_2 + \mathsf{INR}_2\right)\\ &\quad + \log\left(1 + \frac{\mathsf{SNR}_2}{1 + \mathsf{INR}_1}\right) + \log\left(1 + \mathsf{INR}_1\right) + 2\log\left(1 + \frac{\mathsf{SNR}_1}{1 + \mathsf{INR}_2}\right)\\ &> (2.15)_{(2,1)} + \log\left(1 + \mathsf{INR}_2\right) + \log\left(1 + \mathsf{INR}_1\right)\end{aligned}$$

$$
+ \log\left(1 + \frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}\right) + \log\left(1 + \frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right)
$$

$$
> 2 \cdot \min_{(i,j)=(1,2),(2,1)} (2.15)
$$

APPENDIX B

Proofs for Chapter 3

B.1 Proof of Theorem 3.1

Assume the M streams are decoded in the order $(1, \ldots, M)$ for both uplink and downlink, and denote the rate achieved on the mth uplink (downlink) stream by $\bar{\mathbf{R}}_n^{(m)}$ $\left(\bar{\mathbf{R}}_n^{(m)}\right)$, with $\sum_{m=1}^{M} \bar{\mathbf{R}}_n^{(m)} = \bar{\mathbf{R}}_n$ and $\sum_{m=1}^{M} \mathbf{R}_n^{(m)} = \mathbf{R}_n$. Note that these rates are all random variables due to their dependence on $\mathcal{H}_n$, Φ and $\bar{\Phi}$, but in this proof we will suppress this dependence for brevity.

Define $p_n := \mathbb{P}\left(k \in S_m\right)$ and $\bar{p}_n := \mathbb{P}\left(k \in \bar{S}_m\right)$ for an arbitrary user $1 \leq k \leq n$ and arbitrary $1 \leq m \leq M$. Note that $p_n, \bar{p}_n \to 0$ as $n \to \infty$.

Define $\delta'_n = \frac{p_n}{c}$ for a large constant $c > 0$, define $q_n := p_n - \delta'_n$ and $\bar{q}_n := \bar{p}_n - \bar{\delta}'_n$, and the events

$$\bar{\mathcal{F}}_m := \left\{\left|\bar{S}_m\right| \geq n\bar{q}_n\right\} \qquad \bar{\mathcal{G}}_m := \left\{\max_{k \in \bar{S}_m} \left|\bar{\phi}_m^* \bar{h}_k\right|^2 > \epsilon_n\right\}$$

$$\mathcal{F}_m := \left\{\left|S_m\right| \geq nq_n\right\} \qquad \mathcal{G}_m := \left\{\max_{k \in S_m} \left|\phi_m^* h_k\right|^2 > \epsilon_n\right\}$$

Let us choose $\epsilon_n = O\left(\frac{1}{\log n}\right)$. Then

$$\begin{aligned}
\mathbb{P}\left(\bar{\eta} + \eta > \delta\right) &\overset{\text{(a)}}{\leq} \mathbb{P}\left(\bar{\eta} > \frac{\delta}{2}\right) + \mathbb{P}\left(\eta > \frac{\delta}{2}\right) \\
&\overset{\text{(b)}}{\leq} \sum_{m=1}^{M} \mathbb{P}\left(\frac{1}{M}\bar{\mathbf{R}}_n^{\text{MAC-M}} - \bar{\mathbf{R}}_n^{(m)} > \frac{\delta}{2M}\right) \\
&\quad + \sum_{m=1}^{M} \mathbb{P}\left(\frac{1}{M}\bar{\mathbf{R}}_n^{\text{BC}} - \mathbf{R}_n^{(m)} > \frac{\delta}{2M}\right)
\end{aligned}$$

$$\overset{(c)}{=} M\mathbb{P}\left(\frac{1}{M}\bar{\mathbf{R}}_n^{\text{MAC-M}} - \bar{\mathbf{R}}_n^{(m)} > \frac{\delta}{2M}\right)$$

$$+ M\mathbb{P}\left(\frac{1}{M}\bar{\mathbf{R}}_n^{\text{BC}} - \mathbf{R}_n^{(m)} > \frac{\delta}{2M}\right)$$

$$\overset{(d)}{\leq} M\mathbb{P}\left(\frac{1}{M}\bar{\mathbf{R}}_n^{\text{MAC-M}} - \bar{\mathbf{R}}_n^{(m)} > \frac{\delta}{2M}\middle| \bar{\mathcal{F}}_m, \bar{\mathcal{G}}_m\right) \tag{B.1}$$

$$+ M\mathbb{P}\left(\bar{\mathcal{G}}_m^c \middle| \bar{\mathcal{F}}_m\right) + M\mathbb{P}\left(\bar{\mathcal{F}}_m^c\right) \tag{B.2}$$

$$+ M\mathbb{P}\left(\frac{1}{M}\mathbf{R}_n^{\text{BC}} - \mathbf{R}_n^{(m)} > \frac{\delta}{2M}\middle| \mathcal{F}_m, \mathcal{G}_m\right) \tag{B.3}$$

$$+ M\mathbb{P}\left(\mathcal{G}_m^c \middle| \mathcal{F}_m\right) + M\mathbb{P}\left(\mathcal{F}_m^c\right) \tag{B.4}$$

where (a) and (b) follow by the fact that $\sum_{k=1}^{K} a_k > x \Rightarrow \bigvee_{k=1}^{K} (a_k > x/K)$ and by union bound; (c) follows because uniformly random selection of $\bar{\Phi}$ and Φ from the space of unitary matrices induces exchangeable distributions $p\left(\bar{\phi}_1, \ldots, \bar{\phi}_M\right)$ and $p\left(\phi_1, \ldots, \phi_M\right)$ on their respective columns; and (d) follows by the law of total probability and by upper bounding probabilities by one. Of the remaining terms, we will focus only on (B.3) and (B.4) here, to avoid repetition. The uplink counterparts of these terms, given in (B.1) and (B.2), are bounded in exactly the same way in what follows, except where noted.

First consider (B.3). Note that the conditioning on $\mathcal{G}_m$ implies that $k^* \notin S_r$ for $r \neq m$, where k^* is the strongest user in S_m, *i.e.*, $k^* = \arg\max_{k \in S_m} \left|\phi_m^* h_k\right|^2$. This ensures that the user that is scheduled for stream m is not already scheduled for another stream, and hence, using independent Gaussian codebooks and allocating equal power for each downlink stream,

$$\mathbf{R}_n^{(m)} \geq \log\left(1 + \frac{P}{M}\frac{\max_{k \in S_m} \left|\phi_m^* h_k\right|^2}{1 + (2M-1)\epsilon_n}\right), \tag{B.5}$$

almost surely. Therefore,

$$\mathbb{P}\left(\frac{1}{M}\mathbf{R}_n^{DPC} - \mathbf{R}_n^{(m)} > \frac{\delta}{2M}\middle| \mathcal{F}_m, \mathcal{G}_m\right)$$

$$\overset{(a)}{\leq} \mathbb{P}\left(\log\left(\frac{1 + \frac{P}{M}\max_{1 \leq k \leq n} \left\|h_k\right\|^2}{1 + \frac{P}{M}\frac{\max_{k \in S_m} \left|\phi_m^* h_k\right|^2}{1 + (2M-1)\epsilon_n}}\right) > \frac{\delta}{2M}\middle| \mathcal{F}_m, \mathcal{G}_m\right)$$

$$\overset{(b)}{\leq} \mathbb{P}\left(\frac{\max_{1 \leq k \leq n} \left\|h_k\right\|^2}{\max_{k \in S_m} \left|\phi_m^* h_k\right|^2} > \frac{1 + \frac{\delta}{2M}}{1 + (2M-1)\epsilon_n}\middle| \mathcal{F}_m, \mathcal{G}_m\right)$$

$$\overset{\text{(c)}}{\leq} \mathbb{P}\left(\frac{\max_{1\leq k\leq n} \|h_k\|^2}{\max_{k\in S_m} \|h_k\|^2} + \frac{\max_{k\in S_m} \|h_k\|^2}{\max_{k\in S_m} |\phi_m^* h_k|^2} \right.$$

$$\left. > 2\sqrt{\frac{1+\frac{\delta}{2M}}{1+(2M-1)\epsilon_n}} \right| \mathcal{F}_m, \mathcal{G}_m \Bigg)$$

$$\overset{\text{(d)}}{\leq} \mathbb{P}\left(\frac{\max_{1\leq k\leq n} \|h_k\|^2}{\max_{k\in S_m} \|h_k\|^2} > 1+\gamma \middle| \mathcal{F}_m, \mathcal{G}_m \right) \tag{B.6}$$

$$+ \mathbb{P}\left(\frac{\max_{k\in S_m} \|h_k\|^2}{\max_{k\in S_m} |\phi_m^* h_k|^2} > 1+\gamma \middle| \mathcal{F}_m, \mathcal{G}_m \right) \tag{B.7}$$

where (a) follows by using Lemma 3 in [SH07] for downlink and Lemma B.1 for uplink (replace $\frac{P}{M}$ with $\bar{P}$ for uplink); (b) follows by the inequality $e^x \geq 1+x$ and by the fact that $\frac{x}{y} \geq \frac{1+x}{1+y}$ for $x \geq y$; (c) follows by the fact $ab \geq x \Rightarrow a+b \geq 2\sqrt{x}$ (by AM-GM inequality); (d) follows by the fact that $\sum_{k=1}^K a_k > x \Rightarrow \bigvee_{k=1}^K (a_k > x/K)$, by the union bound, and by defining $\gamma > 0$ such that

$$(1+(2M-1)\epsilon_n)(1+\gamma)^2 < 1 + \frac{\delta}{2M}$$

for sufficiently large n.

Next, we bound the terms (B.6) and (B.7) separately. Consider (B.6) first.

$$\text{(B.6)} \leq \mathbb{P}\left(\frac{\max_{1\leq k\leq n} \|h_k\|^2}{\max_{k\in S_m} \|h_k\|^2} > 1+\gamma \middle| \mathcal{F}_m, \mathcal{G}_m \right)$$

$$\overset{\text{(a)}}{=} \mathbb{P}\left(\frac{\max_{k\in S_m^c} \|h_k\|^2}{\max_{k\in S_m} \|h_k\|^2} > 1+\gamma \middle| \mathcal{F}_m, \mathcal{G}_m \right)$$

$$\leq \mathbb{P}\left(\frac{\max_{k\in S_m^c} \|h_k\|^2}{\max_{k\in S_m} |\phi_m^* h_k|^2} > 1+\gamma \middle| \mathcal{F}_m, \mathcal{G}_m \right)$$

$$\overset{\text{(b)}}{\leq} \frac{1}{1-\epsilon_n'} \mathbb{P}\left(\frac{\max_{k\in S_m^c} \|h_k\|^2}{\max_{k\in S_m} |\phi_m^* h_k|^2} > 1+\gamma \middle| |S_m| > nq_n \right)$$

$$= \frac{1}{1-\epsilon_n'} \sum_{s=\lceil nq_n \rceil}^{n} \sum_{\mathcal{A}_s \subseteq [n]:|\mathcal{A}_s|=s} \mathbb{P}\left(S_m = \mathcal{A}_s \middle| |S_m| \geq nq_n \right)$$

$$\cdot \mathbb{P}\left(\frac{\max_{k\in \mathcal{A}_s^c} \|h_k\|^2}{\max_{k\in \mathcal{A}_s} |\phi_m^* h_k|^2} > 1+\gamma \middle| |\mathcal{A}_s| \geq nq_n, S_m = \mathcal{A}_s \right)$$

$$\overset{\text{(c)}}{=} \frac{1}{1-\epsilon_n'} \sum_{s=\lceil nq_n \rceil}^{n} \sum_{\mathcal{A}_s \subseteq [n]:|\mathcal{A}_s|=s} \mathbb{P}\left(S_m = \mathcal{A}_s \middle| |S_m| \geq nq_n \right)$$

$$\cdot \mathbb{P}\left(\frac{\max_{k\in \mathcal{A}_s^c} \|h_k\|^2}{\max_{k\in \mathcal{A}_s} |\phi_m^* h_k|^2} > 1+\gamma \middle| |\mathcal{A}_s| \geq nq_n, S_m = \mathcal{A}_s, \right.$$

$$
\begin{aligned}
&\left\{\exists r \neq m : \left|\phi_r^* h_{k^*}\right|^2 > \epsilon_n \ \vee \ \exists j \in \bar{\mathcal{T}} : \left|h_{k^* j}\right|^2 > \epsilon_n\right\}\Bigg) \\
&\overset{\text{(d)}}{=} \frac{1}{1-\epsilon_n'} \sum_{s=\lceil n q_n \rceil}^{n} \mathbb{P}\left(|S_m| = s \middle| \, |S_m| > n q_n\right) \\
&\cdot \mathbb{P}\left(\frac{\max_{k \in \mathcal{A}_s^c} \|h_k\|^2}{\max_{k \in \mathcal{A}_s} \left|\phi_m^* h_k\right|^2} > 1+\gamma \middle| \, |\mathcal{A}_s| = s, \right. \\
&\left\{\exists r \neq m : \left|\phi_r^* h_{k^*}\right|^2 > \epsilon_n \ \vee \ \exists j \in \bar{\mathcal{T}} : \left|h_{k^* j}\right|^2 > \epsilon_n\right\}\Bigg) \\
&\overset{\text{(e)}}{\leq} \frac{1}{1-\epsilon_n'} \mathbb{P}\left(\frac{\max_{k \in \tilde{\mathcal{A}}_s^c} \|h_k\|^2}{\max_{k \in \tilde{\mathcal{A}}_s} \left|\phi_m^* h_k\right|^2} > 1+\gamma \middle| \right. \\
&\left\{\exists r \neq m : \left|\phi_r^* h_{k^*}\right|^2 > \epsilon_n \ \vee \ \exists j \in \bar{\mathcal{T}} : \left|h_{k^* j}\right|^2 > \epsilon_n\right\}\Bigg) \\
&\overset{\text{(f)}}{\leq} \frac{1}{\left(1-\epsilon_n'\right)^2} \mathbb{P}\left(\frac{\max_{k \in \tilde{\mathcal{A}}^c} \|h_k\|^2}{\max_{k \in \tilde{\mathcal{A}}} \left|\phi_m^* h_k\right|^2} > 1+\gamma \right) \\
&\overset{\text{(g)}}{\leq} \frac{1}{\left(1-\epsilon_n'\right)^2} \left[\mathbb{P}\left(\max_{k \in \tilde{\mathcal{A}}} \left|\phi_m^* h_k\right|^2 < 2 \log\left(\frac{n q_n}{\log\left(n q_n\right)} \right) \right) \right. \\
&+ \mathbb{P}\left(\max_{k \in \tilde{\mathcal{A}}^c} \|h_k\|^2 > 2 \log\left(n q_n\right) + (2M+2) \log\log\left(n q_n\right) \right) \\
&\left. + \mathbb{P}\left(\frac{2 \log\left(n q_n\right) + (2M+2) \log\log\left(n q_n\right)}{2 \log\left(n q_n\right) - 2 \log\log\left(n q_n\right)} > 1+\gamma \right) \right] \\
&\overset{\text{(h)}}{\leq} \frac{1}{\left(1-\epsilon_n'\right)^2} \left[\frac{1}{n q_n} + O\left(\frac{1}{\log\left(n q_n\right)} \right) + 0 \right] \\
&\overset{\text{(i)}}{=} O\left(\frac{1}{\log n} \right)
\end{aligned}
$$

where

- (a) follows by the fact that the ratio can be larger than $(1+\gamma)$ only if the maximum in the numerator occurs for a $k \in S_m^c$ (otherwise the ratio is 1);
- (b) is by the fact that for events A, B; $\mathbb{P}\left(A|B\right) \leq \frac{\mathbb{P}(A)}{\mathbb{P}(B)}$ and by Lemma B.7, where $\epsilon_n' \to 0$;
- (c) is because $S_m = \mathcal{A}_s$ implies the newly conditioned event, which is that for any user outside the set $\mathcal{A}_s$, there must exist an r such that $\left|\phi_r^* h_{k^*}\right|^2 > \epsilon_n$ or an uplink user whose interference strength is larger than ϵ_n, by the construction of the set S_m, where

we have defined $k^* := \arg\max_{k \in \mathcal{A}_s^c} \|h_k\|^2$ (for the uplink case the second part of the event is removed);

- (d) follows by the fact that the probability on the right-hand side does not depend on $\mathcal{A}_s$, as long as $|\mathcal{A}_s|$ is fixed, owing to the fact that the user channel vectors $\bar{h}_k$ are i.i.d.;
- (e) follows because the given probability is a monotonically decreasing function of s, and $\tilde{\mathcal{A}}$ is any arbitrary subset of users such that $\left|\tilde{\mathcal{A}}\right| = \lceil nq_n \rceil$;
- (f) is by Lemma B.7;
- (g) is by the fact that for events A, B, C; $\mathbb{P}(A) \leq \mathbb{P}(B^c) + \mathbb{P}(C^c) + \mathbb{P}(A|B, C)$ by union bound and law of total probability;
- (h) is because of Lemmas B.5 and B.6, and by the fact that the last probability is that of the elements of a deterministic sequence converging to 1 being larger than $1 + \gamma$ for sufficiently large n;
- (i) is because we chose $\epsilon_n = O\left(\frac{1}{\log n}\right)$, and thus $\bar{q}_n = O\left(\frac{1}{\log^{M-1} n}\right)$ by Lemma B.2 for the uplink and $q_n = O\left(\frac{1}{\log^{2M-1} n}\right)$ by Lemma B.3 for the downlink.

Next, we move on to analyze the term (B.7).

$$
\begin{aligned}
\text{(B.7)} &\leq \mathbb{P}\left(\frac{\max_{k \in S_m} \|h_k\|^2}{\max_{k \in S_m} |\phi_m^* h_k|^2} > 1 + \gamma \middle| \mathcal{F}_m, \mathcal{G}_m \right) \\
&\overset{\text{(a)}}{=} \mathbb{P}\left(\frac{\max_{k \in S_m} \sum_{r=1}^{M} |\phi_r^* h_k|^2}{\max_{k \in S_m} |\phi_m^* h_k|^2} > 1 + \gamma \middle| \mathcal{F}_m, \mathcal{G}_m \right) \\
&\overset{\text{(b)}}{\leq} \mathbb{P}\left(\frac{\max_{k \in S_m} |\phi_m^* h_k|^2 + M\epsilon_n}{\max_{k \in S_m} |\phi_m^* h_k|^2} > 1 + \gamma \middle| \mathcal{F}_m, \mathcal{G}_m \right) \\
&= \mathbb{P}\left(\max_{k \in S_m} |\phi_m^* h_k|^2 < \frac{M\epsilon_n}{\gamma} \middle| \mathcal{F}_m, \mathcal{G}_m \right) \\
&\overset{\text{(c)}}{\leq} \frac{1}{1 - \epsilon_n'} \mathbb{P}\left(\max_{k \in S_m} |\phi_m^* h_k|^2 < \frac{M\epsilon_n}{\gamma} \middle| |S_m| \geq nq_n \right) \\
&\overset{\text{(d)}}{\leq} \frac{1}{1 - \epsilon_n'} \left(1 - \exp\left\{ -\frac{M\epsilon_n}{2\gamma} \right\} \right)^{nq_n}
\end{aligned}
$$

$$\overset{\text{(e)}}{=} O\left(\frac{1}{(\log n)^{n/\log n}}\right)$$

where

- (a) follows by the fact that Φ is unitary and thus $\|\Phi h_k\| = \|h_k\|$;
- (b) is by construction of the set S_m;
- (c) is by Lemma B.7;
- (d) is by Lemma B.4;
- (e) is because we chose $\epsilon_n = O\left(\frac{1}{\log n}\right)$, thus $q_n = O\left(\frac{1}{\log^{2M-1} n}\right)$ by Lemma B.3, and $\bar{q}_n = O\left(\frac{1}{\log^{M-1} n}\right)$ by Lemma B.2.

Therefore (B.3) goes to zero as $n \to \infty$. Next, we consider the terms in (B.4). Note that the first term goes to zero since

$$\begin{aligned}\mathbb{P}\left(\mathcal{G}^c|\mathcal{F}\right) &= \left(1 - \exp\left\{-\frac{\epsilon_n}{2}\right\}\right)^{nq_n} \\ &= O\left(\frac{1}{(\log n)^{n/\log n}}\right)\end{aligned}$$

by Lemma B.4 and by the choice of ϵ_n. The second term in (B.4) goes to zero by weak law of large numbers for triangular arrays [Dur10], applied to the binomial random variable $|S_m|$ with mean np_n.

Since all terms in (B.1), (B.2), (B.3), and (B.4) go to zero, the result follows.

B.2 Proof of Theorem 3.2

Let us choose $\epsilon_n = \epsilon > 0$, *i.e.*, a constant. Then, as in the proof of Theorem 3.1,

$$\begin{aligned}\mathbb{P}\left(\mathbf{R}_n + \bar{\mathbf{R}}_n < 2M\beta\right) \leq \sum_{m=1}^{M} &\mathbb{P}\left(\mathbf{R}_n^{(m)} < \beta\right) \\ &+ \mathbb{P}\left(\bar{\mathbf{R}}_n^{(m)} < \beta\right)\end{aligned}$$

by the fact that $\sum_{k=1}^{K} a_k < x \Rightarrow \bigvee_{k=1}^{K} (a_k < x/K)$ and by union bound. We only consider the first term, associated with downlink. The uplink term is bounded the same way, except where noted. By law of total probability, and by upper bounding probabilities by one,

$$\begin{aligned}\mathbb{P}\left(\mathbf{R}_n^{(m)} < \beta\right) &\leq \mathbb{P}\left(\mathcal{F}_m^c\right) + \mathbb{P}\left(\mathcal{G}_m^c | \mathcal{F}_m\right) \\ &\quad + \mathbb{P}\left(\mathbf{R}_n^{(m)} < \beta \middle| \mathcal{F}_m, \mathcal{G}_m\right) \end{aligned} \tag{B.8}$$

Since ϵ_n is a constant, $\mathbb{P}\left(\mathcal{F}_m^c\right)$ goes to zero exponentially by Hoeffding's inequality. $\mathbb{P}\left(\mathcal{G}_m^c | \mathcal{F}_m\right)$ is upper bounded by

$$\begin{aligned}\mathbb{P}\left(\mathcal{G}_m^c | \mathcal{F}_m\right) &\leq \left(1 - \exp\left\{-\frac{\epsilon}{2}\right\}\right)^{nq_n} \\ &= a^{\frac{n^{1+2\log a}}{a}},\end{aligned}$$

by Lemmas B.4 and B.3, where $a = (1 - \exp\{-\epsilon/2\})$. Note that the last term goes to zero exponentially if $1 + 2\log a > 0$, which is satisfied for sufficiently large $\epsilon > 0$. We consider the first term. Conditioned on $\mathcal{G}_m$, a different user is scheduled for each stream, hence

$$\begin{aligned}&\mathbb{P}\left(\mathbf{R}_n^{(m)} < \beta \middle| \mathcal{F}_m, \mathcal{G}_m\right) \\ &\leq \mathbb{P}\left(\log\left(\frac{\max_{k \in S_m} \left|\phi_m^* h_k\right|^2}{1 + (2M-1)\epsilon}\right) < \beta \middle| \mathcal{F}_m, \mathcal{G}_m\right) \\ &\overset{\text{(a)}}{=} \frac{1}{1-\epsilon_n'} \mathbb{P}\left(\log\left(\frac{\max_{k \in S_m} \left|\phi_m^* h_k\right|^2}{1 + (2M-1)\epsilon}\right) < \beta \middle| \mathcal{F}_m\right) \\ &\leq \frac{1}{1-\epsilon_n'} \mathbb{P}\left(\max_{k \in S_m} \left|\phi_m^* h_k\right|^2 < \beta e \left(1 + 2\epsilon \log n\right) \middle| \mathcal{F}_m\right) \\ &\overset{\text{(b)}}{\leq} \frac{1}{1-\epsilon_n'} \left(1 - \exp\left\{-\frac{\beta e \left(1 + 2\epsilon \log n\right)}{2}\right\}\right)^{nq_n} \\ &\overset{\text{(c)}}{=} \Theta\left(e^{-n^{\gamma}}\right)\end{aligned}$$

where (a) follows by Lemma B.7, (b) follows by Lemma B.4, and (c) follows, for some $0 < \gamma < 1$, by Lemma B.3 with the choice $\epsilon_n = \epsilon$, and by letting $M = \alpha \log n$ for sufficiently small $\alpha > 0$. Since all terms in (B.8) go to zero exponentially as $n \to \infty$,

$$\sum_n \mathbb{P}\left(\mathbf{R}_n + \bar{\mathbf{R}}_n < 2M\right) < \infty$$

and thus by Borel-Cantelli Lemma [Dur10], the result follows.

B.3 Proof of (3.1)

Let us denote message of the kth uplink user as $\bar{W}_k$, the message intended for the kth downlink user by W_k, and for any set S, define $W_S = \{W_k : k \in S\}$. We also define $v_t^{(m)} = y_t^{(m)} - \mathbf{1} h_m^* x_t$, where $\mathbf{1}$ is the vector of ones, *i.e.*, $v_t^{(m)}$ is the vector of interference signals at the downlink users of cluster m at time t. Set $v_t = \left[v_t^{(1)}, \ldots, v_t^{(M)}\right]^*$.

We consider a block length of N, and as explained in Section 3.5, assume $z_t^{(m)} \sim \mathcal{CN}(\mathbf{0}, \mathbf{1}\mathbf{1}^T)$, where $\mathbf{1}$ is the all ones vector, for $m \in [M]$. We also assume that the downlink users within each cluster cooperate, since this cannot reduce capacity. Then, by Fano's inequality,

$$
\begin{aligned}
& N\left(\mathbf{R}_n + \bar{\mathbf{R}}_n\right) \leq I\left(W_{[n]}; y^N\right) + I\left(\bar{W}_{[n]}; \bar{y}^N\right) \\
& \leq I\left(W_{[n]}; y^N\right) + I\left(\bar{W}_{[n]}; \bar{y}^N, y^N, W_{[n]}\right) \\
& \overset{\text{(a)}}{=} I\left(W_{[n]}; y^N\right) + I\left(\bar{W}_{[n]}; \bar{y}^N, y^N \left| W_{[n]}\right.\right) \\
& = h\left(y^N\right) - h\left(y^N \left| W_{[n]}\right.\right) + h\left(\bar{y}^N, y^N \left| W_{[n]}\right.\right) \\
& \qquad - h\left(\bar{y}^N, y^N \left| W_{[n]}, \bar{W}_{[n]}\right.\right) \\
& = h\left(y^N\right) + h\left(\bar{y}^N \left| W_{[n]}, y^N\right.\right) - h\left(\bar{y}^N, y^N \left| W_{[n]}, \bar{W}_{[n]}\right.\right) \\
& = \sum_{t=1}^{N} h\left(y_t | y^{t-1}\right) + h\left(\bar{y}_t \left| W_{[n]}, y^N, \bar{y}^{t-1}\right.\right) \\
& \qquad - h\left(\bar{y}_t, y_t \left| W_{[n]}, \bar{W}_{[n]}, \bar{y}^{t-1}, y^{t-1}\right.\right) \\
& \overset{\text{(b)}}{=} \sum_{t=1}^{N} h\left(y_t | y^{t-1}\right) + h\left(\bar{y}_t \left| W_{[n]}, y^N, \bar{y}^{t-1}, x_t\right.\right) \\
& \qquad - h\left(\bar{y}_t, y_t \left| W_{[n]}, \bar{W}_{[n]}, \bar{y}^{t-1}, y^{t-1}, \bar{x}_t, x_t\right.\right) \\
& \overset{\text{(c)}}{\leq} \sum_{t=1}^{N} h\left(y_t\right) + h\left(\bar{y}_t \left| y_t, x_t\right.\right) - h\left(\bar{z}_t, z_t\right) \\
& \overset{\text{(d)}}{=} \sum_{t=1}^{N} h\left(y_t\right) + h\left(\bar{y}_t \left| v_t, x_t\right.\right) - h\left(\bar{z}_t\right) - h\left(z_t\right) \\
& \leq \sum_{t=1}^{N} h\left(y_t\right) + h\left(\bar{y}_t \left| v_t\right.\right) - h\left(\bar{z}_t\right) - h\left(z_t\right)
\end{aligned}
$$

$$\overset{\text{(e)}}{\leq} \sum_{t=1}^{N} \left(\sum_{m=1}^{M} h\left(y_t^{(m)} \right) \right) + h\left(\bar{y}_t \left| v_t \right. \right) - h\left(\bar{z}_t \right)$$
$$- \left(\sum_{m=1}^{M} h\left(z_t^{(m)} \right) \right)$$

where (a) follows by independence of messages; (b) follows by the fact that x_t is a deterministic function of $\left(W_{[n]}, \bar{y}^{t-1}\right)$ and $\bar{x}_t$ is a deterministic function of $\bar{W}_{[n]}$; (c) follows because conditioning reduces entropy and by subtracting x_t and $\bar{x}_t$ from y_t and $\bar{y}_t$; (d) is because $v_t = y_t - \mathbf{1}H^*x$ and by independence of uplink and downlink noise; (e) is by the fact that conditioning reduces entropy, and that noise processes at different clusters are independent. Since $\{h_m\}$ are orthogonal, $\left\{ \bar{y}_t^{(m)} \right\}$ can be uniquely expressed as $\bar{y}_t = \sum_{m=1}^{M} \frac{h_m^*}{\|h_m\|} \bar{y}_t^{(m)}$, *i.e.*, this transformation is a bijection. Let us define the matrix $\widetilde{H} := \left[\frac{h_1}{\|h_1\|} \; \ldots \; \frac{h_M}{\|h_M\|} \right]$. Then

$$h\left(\bar{y}_t\right) = h\left(\widetilde{H}^* \bar{y}_t \right) = h\left(\bar{y}_t^{(1)}, \ldots, \bar{y}_t^{(M)} \right) + \log\left| \widetilde{H} \right|$$
$$= h\left(\bar{y}_t^{(1)}, \ldots, \bar{y}_t^{(M)} \right)$$

since $\widetilde{H}$ is unitary. Similarly, $\bar{z}_t = \sum_{m=1}^{M} h_m \bar{z}_t^{(m)}$, and $\left\{ \bar{z}_t^{(m)} \right\}$ are still distributed i.i.d. $\mathcal{CN}(0,1)$. Hence, also using the fact that conditioning reduces entropy,

$$N\left(\mathbf{R}_n + \bar{\mathbf{R}}_n \right) \leq \sum_{t=1}^{N} \sum_{m=1}^{M} h\left(y_t^{(m)} \right) + h\left(\bar{y}_t^{(m)} \left| v_t^{(m)} \right. \right)$$
$$- h\left(\bar{z}_t^{(m)} \right) - h\left(z_t^{(m)} \right)$$

Let $k_{m,t}$ denote the number of uplink users scheduled in cluster m at time t, with $\sum_{m=1}^{M} k_{m,t} \leq M$, for all t. Note that given any power allocation, there is a covariance constraint on $\left[\begin{array}{cc} \bar{y}_t^{(m)} & v_t^{(m)} \end{array} \right]^*$ given by

$$K = I + k_{m,t} \bar{P} \left[\begin{array}{c} h_m \\ g \end{array} \right] \left[\begin{array}{cc} h_m^* & g* \end{array} \right].$$

Hence, $h(\bar{y}_t | v_t)$ is maximized when $(\bar{y}_t, v_t) \sim \mathcal{CN}\left(0, K\right)$, with

$$h(\bar{y}_t | v_t) = \log 2\pi e \left| K_{\bar{y}|v} \right|,$$

where $K_{\bar{y}|v}$ is the conditional covariance matrix of $\bar{y}_t^{(m)}$ given $v_t^{(m)}$. Therefore, evaluating the differential entropy terms with Gaussian input distributions[1], and using the fact that $z_t^{(m)} \sim \mathcal{CN}(\mathbf{0}, \mathbf{1}\mathbf{1}^T)$, we find (3.1).

B.4 Proof of Corollary 3.1

Using Theorem 3.3 and (3.2), $\eta + \bar{\eta}$ can be lower bounded by

$$\begin{aligned}\eta + \bar{\eta} > M \log\left(\frac{1+\mathsf{SNR}^\beta}{1+\frac{\mathsf{SNR}^\beta}{1+\mathsf{SNR}^\alpha}}\right)\\ - M\log\left(1+\frac{\mathsf{SNR}^\alpha}{1+\frac{1}{M}\mathsf{SNR}}\right) - M\log 3\end{aligned}$$

If we use the notation $f(\mathsf{SNR}) \doteq g(\mathsf{SNR})$ to mean that $\lim_{\mathsf{SNR}\to\infty} \frac{f(\mathsf{SNR})}{g(\mathsf{SNR})} = 1$, then it is easy to see that

$$\begin{aligned}&\log\left(\frac{1+\mathsf{SNR}^\beta}{1+\frac{\mathsf{SNR}^\beta}{1+\mathsf{SNR}^\alpha}}\right) - \log\left(1+\frac{\mathsf{SNR}^\alpha}{1+\frac{1}{M}\mathsf{SNR}}\right) - \log 3\\ &\doteq \log \mathsf{SNR}^\alpha - \log \mathsf{SNR}^{\alpha-1}\\ &= \log \mathsf{SNR}\end{aligned}$$

Hence, the result follows.

B.5 Auxiliary Lemmas

Lemma B.1.

$$\bar{\mathbf{R}}_n^{MAC\text{-}M}(\mathcal{H}_n) \leq M \log\left(1 + P \max_{1\leq k\leq n} \|\bar{h}_k\|^2\right)$$

Proof. The capacity of a MIMO MAC with a per-user power constraint $\bar{P}$, and an active

[1] We evaluate $h(y_t^{(m)})$ assuming a *joint* Gaussian distribution on x_t and $\bar{x}_t$ with arbitrary correlation, since x_t is a function of both $W_{[n]}$ and $\bar{y}_{t-1}$.

user constraint M is given by

$$\begin{aligned}\bar{\mathbf{R}}_n^{\text{MAC-M}}(\mathcal{H}_n) &= \max_{\mathcal{A}\subseteq[n]:|\mathcal{A}|=M} \log\left|I_M + \bar{P}\bar{H}_{\mathcal{A}}\bar{H}_{\mathcal{A}}^*\right| \\ &= \max_{\mathcal{A}\subseteq[n]:|\mathcal{A}|=M} \log\left|I_M + \bar{P}\sum_{k\in\mathcal{A}}\bar{h}_k\bar{h}_k^*\right|\end{aligned}$$

Using the inequality $|A| \leq \left(\frac{\operatorname{tr}(A)}{M}\right)^M$ (which is a direct consequence of AM-GM inequality applied to the eigenvalues of A),

$$\begin{aligned}&\bar{\mathbf{R}}_n^{\text{MAC-M}}(\mathcal{H}_n) \\ &= \max_{\mathcal{A}\subseteq[n]:|\mathcal{A}|=M} M\log\left(\frac{\operatorname{tr}\left(I_M + \bar{P}\sum_{k\in\mathcal{A}}\bar{h}_k\bar{h}_k^*\right)}{M}\right) \\ &= \max_{\mathcal{A}\subseteq[n]:|\mathcal{A}|=M} M\log\left(1 + \frac{\bar{P}\sum_{k\in\mathcal{A}}\operatorname{tr}\left(\bar{h}_k\bar{h}_k^*\right)}{M}\right) \\ &= \max_{\mathcal{A}\subseteq[n]:|\mathcal{A}|=M} M\log\left(1 + \bar{P}\frac{\sum_{k\in\mathcal{A}}\|\bar{h}_k\|^2}{M}\right) \\ &= M\log\left(1 + \bar{P}\max_{\mathcal{A}\subseteq[n]:|\mathcal{A}|=M}\frac{\sum_{k\in\mathcal{A}}\|\bar{h}_k\|^2}{M}\right) \\ &\leq M\log\left(1 + \bar{P}\max_{1\leq k\leq n}\|\bar{h}_k\|^2\right)\end{aligned}$$

□

Lemma B.2. *For an arbitrary uplink user* $1 \leq k \leq n$*, and arbitrary* $1 \leq m \leq M$*,*

$$\mathbb{P}\left(k \in \bar{S}_m\right) = \left(1 - \exp\{-\epsilon_n/2\}\right)^{M-1}$$

Proof.

$$\begin{aligned}\mathbb{P}\left(k \in \bar{S}_m\right) &= \mathbb{P}\left(\left|\bar{\phi}_r^*\bar{h}_k\right|^2 \leq \epsilon_n,\ \forall r \neq m\right) \\ &\overset{(a)}{=} \left[\mathbb{P}\left(\left|\bar{\phi}_1^*\bar{h}_k\right|^2 \leq \epsilon_n\right)\right]^{M-1} \\ &\overset{(b)}{=} \left(1 - \exp\{-\epsilon_n/2\}\right)^{M-1}\end{aligned}$$

where (a) follows by the fact that the components of $\bar{\Phi}\bar{h}_k$ are i.i.d. distributed because $\bar{\Phi}$ is unitary; and (b) follows by the fact that $\left|\bar{\phi}_1^*\bar{h}_k\right|^2$ is $\chi^2(2)$ distributed. □

Lemma B.3. *For an arbitrary downlink user* $1 \leq k \leq n$*, and arbitrary* $1 \leq m \leq M$*,*

$$\mathbb{P}\left(k \in S_m\right) = \left(1 - \exp\left\{-\epsilon_n/2\right\}\right)^{2M-1}$$

Proof.

$$\begin{aligned}
&\mathbb{P}\left(k \in S_m\right) \\
&= \mathbb{P}\left(\left|\phi_r^* h_k\right|^2 \leq \epsilon_n,\ \forall r \neq m;\ \ \left|h_{kj}\right|^2 \leq \epsilon_n,\ \forall j \in \bar{\mathcal{T}}\right) \\
&= \sum_{\mathcal{A} \subseteq [n]:|\mathcal{A}|=M} \mathbb{P}\left(\bar{\mathcal{T}} = \mathcal{A}\right) \\
&\cdot \mathbb{P}\left(\left|\phi_r^* h_k\right|^2 \leq \epsilon_n\ \forall r \neq m;\ \ \left|h_{kj}\right|^2 \leq \epsilon_n,\ \forall j \in \mathcal{A} \middle| \bar{\mathcal{T}} = \mathcal{A}\right) \\
&\overset{\text{(a)}}{=} \sum_{\mathcal{A} \subseteq [n]:|\mathcal{A}|=M} \mathbb{P}\left(\bar{\mathcal{T}} = \mathcal{A}\right) \\
&\cdot \mathbb{P}\left(\left|\phi_r^* h_k\right|^2 \leq \epsilon_n\ \forall r \neq m;\ \ \left|h_{kj}\right|^2 \leq \epsilon_n,\ \forall j \in \mathcal{A}\right) \\
&= \mathbb{P}\left(\left|\phi_r^* h_k\right|^2 \leq \epsilon_n\ \forall r \neq m;\ \ \left|h_{kj}\right|^2 \leq \epsilon_n,\ \forall j \in \mathcal{A}\right) \\
&\overset{\text{(b)}}{=} \left[\mathbb{P}\left(\left|\bar{\phi}_r^* \bar{h}_k\right|^2 \leq \epsilon_n\right)\right]^{M-1} \left[\mathbb{P}\left(\left|h_{k1}\right|^2 \leq \epsilon\right)\right]^{M} \\
&\overset{\text{(c)}}{=} \left(1 - \exp\left\{-\epsilon_n/2\right\}\right)^{2M-1}
\end{aligned}$$

where (a) follows by the fact that $\bar{\mathcal{T}}$ is a function of $\left\{\bar{\phi}_m^* \bar{h}_k\right\}_{m,k}$, and all links are independent, and thus the event $\left\{\bar{\mathcal{T}} = \mathcal{A}\right\}$ is independent, (defining $\tilde{\mathcal{A}}$ to be an arbitrary subset of uplink users s.t. $\left|\tilde{\mathcal{A}}\right| = M$); (b) follows because the components of Φh_k are i.i.d. distributed and all links are independent; and (c) follows because both $\left|\phi_1^* h_k\right|^2$ and $\left|h_{k1}\right|^2$ are $\chi^2(2)$ distributed. □

Lemma B.4.

$$\mathbb{P}\left(\max_{k \in S_m} \left|\phi_m^* h_k\right|^2 < x \middle| \left|S_m\right| \geq n q_n\right) \leq \left(1 - e^{-\frac{x}{2}}\right)^{n q_n}$$

Proof.

$$\mathbb{P}\left(\max_{k \in S_m} \left|\phi_m^* h_k\right|^2 < x \middle| \left|S_m\right| \geq n q_n\right)$$

$$
\begin{aligned}
&= \sum_{s=\lceil n(\bar{p}-\delta)\rceil}^{n} \sum_{\substack{\mathcal{A}_s\subseteq[n]:\\ |\mathcal{A}_s|=s}} \mathbb{P}\left(S_m = \mathcal{A}_s|\, |S_m| \geq nq_n\right) \\
&\qquad \cdot \mathbb{P}\left(\max_{k\in\mathcal{A}_s} |\phi_m^* h_k|^2 < x \middle| |\mathcal{A}_s| \geq nq_n, S_m = \mathcal{A}_s\right) \\
&\overset{\text{(a)}}{=} \sum_{s=\lceil nq_n\rceil}^{n} \mathbb{P}\left(|S_m| = s\, ||S_m| \geq nq_n\right) \\
&\qquad \cdot \mathbb{P}\left(\max_{k\in\mathcal{A}_s} |\phi_m^* h_k|^2 < x \middle| |\mathcal{A}_s| = s\right) \\
&\overset{\text{(b)}}{\leq} \mathbb{P}\left(\max_{k\in\mathcal{A}_s} |\phi_m^* h_k|^2 < x \middle| |\mathcal{A}_s| = nq_n\right) \\
&\overset{\text{(c)}}{=} \left(1 - e^{-\frac{x}{2}}\right)^{nq_n}
\end{aligned}
$$

where (a) follows by the fact that the probability on the right-hand side does not depend on $\mathcal{A}_s$ as long as $|\mathcal{A}_s|$ is fixed, owing to the fact that the user channel vectors h_k are i.i.d., and since $\Phi h_k \sim \mathcal{CN}(0, I)$; (b) follows because the given probability is a monotonically decreasing function of s; and (c) is because $\left\{|\phi_m^* h_k|^2\right\}$ are i.i.d. $\chi^2(2)$ distributed; □

Lemma B.5. *Let $X_1, \ldots, X_N$ be i.i.d. $\chi^2(2)$ distributed random variables. Then*

$$\mathbb{P}\left(\max_{1\leq i\leq N} X_i < 2\log N - 2\log\log N\right) \leq \frac{1}{N}$$

Proof.

$$
\begin{aligned}
&\mathbb{P}\left(\max_{1\leq i\leq N} X_i < 2\log N - \log\log N\right) \\
&= \left[\mathbb{P}\left(X_1 < 2\log N - \log\log N\right)\right]^N \\
&= \left(1 - \exp\left\{-\log N + \log\log N\right\}\right)^N = \left(1 - \frac{\log N}{N}\right)^N \\
&= \exp\left\{N \log\left(1 - \frac{\log N}{N}\right)\right\} \\
&= \exp\left\{N\left(-\frac{\log N}{N} - O\left(\frac{\log^2 N}{N^2}\right)\right)\right\} \leq \frac{1}{N}
\end{aligned}
$$

□

Lemma B.6. *Let $X_1, \ldots, X_N$ be i.i.d. $\chi^2(2M)$ distributed random variables. Then for N sufficiently large,*

$$\mathbb{P}\left(\max_{1\leq i\leq N} X_i > 2\log N + (2M+2)\log\log N\right)$$
$$= O\left(\frac{1}{\log N}\right)$$

Proof. Chernoff bound for a $\chi^2(2M)$ random variable Z is given by

$$\mathbb{P}\left(Z > x\right) \leq \left(\frac{x}{2M}e^{1-\frac{x}{2M}}\right)^M,$$

for $x > 2M$. Then, assuming N is large enough,

$$\begin{aligned}
&\mathbb{P}\left(\max_{1\leq i\leq N} X_i > 2\log N + (2M+2)\log\log N\right)\\
&= 1 - \mathbb{P}\left(\max_{1\leq i\leq N} X_i \leq 2\log N + (2M+2)\log\log N\right)\\
&= 1 - \left[\mathbb{P}\left(X_1 \leq 2\log N + (2M+2)\log\log N\right)\right]^N\\
&= 1 - \left[1 - \mathbb{P}\left(X_1 > 2\log N + (2M+2)\log\log N\right)\right]^N\\
&\leq 1 - \left(1 - \left(\frac{2\log N + (2M+2)\log\log N}{2M}\right.\right.\\
&\qquad \left.\left.\exp\left\{1 - \frac{2\log N + (2M+2)\log\log N}{2M}\right\}\right)^M\right)^N\\
&= 1 - \left(1 - \frac{\left(2\log N + (2M+2)\log\log N\right)^M e^M}{\left(2M\right)^M N\log^{M+1} N}\right)^N\\
&\doteq 1 - \exp\left\{-\frac{\left(2\log N + (2M+2)\log\log N\right)^M}{\left(2M/e\right)^M \log^{M+1} N}\right\}\\
&\overset{\text{(a)}}{\leq} \left(\frac{e}{2M}\right)^M \frac{\left(2\log N + (2M+2)\log\log N\right)^M}{\log^{M+1} N}\\
&= \left(\frac{e}{2M}\right)^M \frac{O\left(\log^M N\right)}{\log^{M+1} N} = O\left(\frac{1}{\log N}\right)
\end{aligned}$$

where (a) is by the inequality $1 - x \leq e^{-x}$. □

Lemma B.7. *If $N_n \to \infty$ and $\epsilon_n \to 0$ as $n \to \infty$, then for i.i.d. $\chi^2(2)$ distributed $X_i, \ldots, X_{N_n}$,*

$$\lim_{n\to\infty} \mathbb{P}\left(\max_{1\leq k\leq N_n} X_k > \epsilon_n\right) = 1.$$

The proof for Lemma B.7 is trivial and omitted here.

APPENDIX C

Proofs for Chapter 4

C.1 Proof of Theorem 4.2

Proposition C.1. *Let $X_i, i = 1, \ldots, n$, be i.i.d. $\chi^2(2P)$ random variables. Then*

$$\mathbb{P}\left(\min_{1\le i\le n} X_i > n^{-\frac{\gamma}{2P}}\psi(2P)\right) = O\left(e^{-n^{1-\gamma}}\right), \quad \textit{for } 0 < \gamma < 1.$$

Proof. Using the Taylor series for the upper incomplete Gamma function, as $x \to 0$,

$$\mathbb{P}\left(X_i > x\right) = 1 - \frac{x^{2P}}{(2P)!} + O\left(x^{2P+1}\right).$$

Therefore,

$$\begin{aligned}\mathbb{P}\left(\min_{1\le i\le n} X_i > n^{-\frac{\gamma}{2P}}\psi(2P)\right) &= \left(\mathbb{P}\left(X_i > n^{-\frac{\gamma}{2P}}\psi(2P)\right)\right)^n \\ &= \left(1 - n^{-\gamma}\right)^n = O\left(e^{-n^{1-\gamma}}\right).\end{aligned}$$

□

We will first derive a lower bound on SNR_{ij}^{coop}, defined by

$$\mathsf{SNR}_{ij}^{coop} = \frac{s_{ij1}^2}{1 + \left|u_{ij1}(2)\right|^2 \frac{\sigma_{j|i}^2}{\|g_{ij}\|^2}}.$$

Using the fact that $\left|u_{ij1}(2)\right|^2 \le 1$ and $\sigma_{j|i}^2 \le \sigma_j^2$, where σ_j^2 is the variance of y_2,

$$\mathsf{SNR}_{ij}^{coop} \ge \frac{s_{ij1}^2}{1 + \frac{\sigma_j^2}{\|g_{ij}\|^2}} = \frac{s_{ij1}^2}{1 + \frac{1+\|\mathbf{h}_j\|^2}{\|g_{ij}\|^2}}. \tag{C.1}$$

Next, since s_{ij1}^2 is the larger eigenvalue of the matrix $\mathbf{H}_{ij}\mathbf{H}_{ij}^*$, using the closed form expressions for the eigenvalues of 2×2 matrices,

$$\begin{aligned} s_{ij1}^2 &= \frac{1}{2}\left(\|\mathbf{h}_i\|^2 + \|\mathbf{h}_j\|^2 + \sqrt{\|\mathbf{h}_i\|^4 + \|\mathbf{h}_j\|^4 + 2\|\mathbf{h}_i\|^2\|\mathbf{h}_j\|^2\cos(2\Theta)}\right) \\ &\geq \frac{1}{2}\left(\|\mathbf{h}_i\|^2 + \|\mathbf{h}_j\|^2 + \left|\|\mathbf{h}_i\|^2 - \|\mathbf{h}_j\|^2\right|\right) = \max\left(\|\mathbf{h}_i\|^2, \|\mathbf{h}_j\|^2\right), \end{aligned}$$

where $\Theta = \cos^{-1}\frac{\mathbf{h}_i^*\mathbf{h}_j}{\|\mathbf{h}_i\|\|\mathbf{h}_j\|}$ is the angle between $\mathbf{h}_i$ and $\mathbf{h}_j$, and the lower bound is obtained by setting $\cos(2\Theta) = -1$. Using this lower bound in (C.1), we get

$$\begin{aligned} \mathsf{SNR}_{ij}^{coop} &\geq \frac{\max\left(\|\mathbf{h}_i\|^2, \|\mathbf{h}_j\|^2\right)}{1 + \frac{1+\|\mathbf{h}_j\|^2}{\|g_{ij}\|^2}} \geq \frac{\|\mathbf{h}_j\|^2}{1 + \frac{1+\|\mathbf{h}_j\|^2}{\|g_{ij}\|^2}} \geq \frac{\left(\|\mathbf{h}_j\|^2 + 1\right)\|g_{ij}\|^2}{1 + \|\mathbf{h}_j\|^2 + \|g_{ij}\|^2} - 1 \\ &\geq \frac{1}{2}\min\left(\|\mathbf{h}_j\|^2, \|g_{ij}\|^2\right) - 1. \end{aligned}$$

Therefore, to prove the first claim in Theorem 4.2, it is sufficient to prove that

$$\begin{aligned} &\mathbb{P}\left(\min_{i\in\mathcal{N}}\min\left(\|\mathbf{h}_{j^*(i)}\|^2, \|g_{ij^*(i)}\|^2\right) > M\rho\left(\frac{1}{2}\log n - 2\log\log n\right)\right) \\ &\qquad = O\left(e^{-\log^2 n + 2\log n}\right). \end{aligned}$$

Define $\mathcal{P}_n = \{j : \|\mathbf{h}_j\|^2 \geq M\rho\left(\frac{1}{2}\log n - 2\log\log n\right)\}$, and $\mathcal{R}_n(i) = \{j \in \mathcal{P}_n : \phi_{ij} \geq n^{\frac{c}{4}}\}$.

Proposition C.2. $\mathbb{P}\left(\mathcal{R}_n(i) = \varnothing \textit{ for some } i\right) = O\left(e^{-\log^2 n + 2\log n}\right)$.

Therefore, if $\mathcal{R}_n(i) \neq \varnothing$ for all i,

$$\begin{aligned} 1 + \mathsf{SNR}_{\min}^{coop} &\geq \frac{1}{2}\min_{i\in\mathcal{N}}\min\left(\|\mathbf{h}_{j^*(i)}\|^2, \|g_{ij^*(i)}\|^2\right) \\ &\geq \frac{1}{2}\min_{i\in\mathcal{N}}\min\left(M\rho\left(\frac{1}{2}\log n - 2\log\log n\right), n^{\frac{c}{4}}\|\zeta_{ij^\dagger(i)}\|^2\right) \\ &= \frac{1}{2}\min\left(M\rho\left(\frac{1}{2}\log n - 2\log\log n\right), n^{\frac{c}{4}}\min_{i\in\mathcal{N}}\|\zeta_{ij^\dagger(i)}\|^2\right), \end{aligned}$$

where $j^\dagger(i) = \arg\max_{j\in\mathcal{R}_n(i)}\mathbb{E}\left[\mathsf{SNR}_{ij}^{coop}\middle|\,\phi_{ij}, \mathbf{h}_j\right]$, and thus

$$\begin{aligned} &\mathbb{P}\left(\mathsf{SNR}_{\min}^{coop} < \frac{1}{2}M\rho\left(\frac{1}{2}\log n - 2\log\log n\right) - 1\middle|\,\mathcal{R}_n(i) \neq \varnothing\ \forall i\right) \\ &\qquad = O\left(e^{-n^{1-\gamma}}\right), \end{aligned} \tag{C.2}$$

for all $0 < \gamma < 1$, by Proposition C.1, by the fact that $\|\zeta_{ij}\|^2$ is a $\chi^2(2)$ random variable, and that $j^\dagger(i)$ is independent of $\|\zeta_{ij}\|^2$. Then (C.2), together with Proposition C.2 implies the first claim of the theorem.

It remains to prove Proposition C.2. To achieve this, we will first lower bound the tail probability $\mathbb{P}\left(\|\mathbf{h}_j\|^2 > a\right)$. Define $\hat{\mathbf{e}}_{j,k} := \frac{\mathbf{e}_{j,k}}{\|\mathbf{e}_{j,k}\|} = \frac{\mathbf{e}_{j,k}}{\sqrt{M}}$, $\mathbf{E}_j := [\hat{\mathbf{e}}_{j,1} \ \ldots \ \hat{\mathbf{e}}_{j,P}]$, and $\xi_j := [\xi_{j,k}]_k$. Letting $\mathbf{E}_j = \mathbf{Q}_j \Lambda_j \mathbf{Q}_j^*$ be an eigendecomposition of $\mathbf{E}_j$,

$$\begin{aligned}\|\mathbf{h}_j\|^2 &= \rho \left\| \sum_{k=1}^{P} \xi_{j,k} \mathbf{e}(\theta_{j,k}) \right\|^2 = M\rho \left\| \sum_{k=1}^{P} \xi_{j,k} \hat{\mathbf{e}}(\theta_{j,k}) \right\|^2 \\ &= M\rho \left(\mathbf{E}_j \xi_j\right)^* \left(\mathbf{E}_j \xi_j\right) = M\rho \xi_j^* \left(\mathbf{E}_j^* \mathbf{E}_j\right) \xi_j = M\rho \sum_{k=1}^{P} \lambda_k \left(\mathbf{E}_j^* \mathbf{E}_j\right) \left|\left(\mathbf{Q}_j \xi_j\right)_k\right|^2,\end{aligned}$$

where $\lambda_k\left(\mathbf{E}_j^*\mathbf{E}_j\right)$ is the kth eigenvalue of $\mathbf{E}_j^*\mathbf{E}_j$, and $\left(\mathbf{Q}_j\xi_j\right)_k$ is the kth element of $\mathbf{Q}_j\xi_j$. Since $\sum_{k=1}^{P} \lambda_k\left(\mathbf{E}_j^*\mathbf{E}_j\right) = \operatorname{tr}\left(\mathbf{E}_j^*\mathbf{E}_j\right) = P$, there must exist a k, say k^*, such that $\lambda_{k^*}\left(\mathbf{E}_j^*\mathbf{E}_j\right) \geq 1$. Hence,

$$\|\mathbf{h}_j\|^2 = M\rho \sum_{k=1}^{P} \lambda_k \left(\mathbf{E}_j^* \mathbf{E}_j\right) \left|\left(\mathbf{Q}_j \xi_j\right)_k\right|^2 \geq M\rho \left|\left(\mathbf{Q}_j \xi_j\right)_{k^*}\right|^2.$$

Since $\mathbf{E}_j$ is independent from ξ_j, and since the distributions of i.i.d. Gaussian vectors are invariant under orthogonal transformations, $\left|\left(\mathbf{Q}_j\xi_j\right)_{k^*}\right|^2$ has the same distribution as $\|\xi_{j,k}\|^2$ for an arbitrary k, *i.e.*, $\chi^2(2)$ distribution, or equivalently, exponential distribution with mean 1. Therefore, the tail probability of $\|\mathbf{h}_j\|^2$ can be lower bounded by $\mathbb{P}\left(\|\mathbf{h}_j\|^2 > M\rho a\right) \geq e^{-a}$. Hence,

$$\begin{aligned}&\mathbb{P}\left(|\mathcal{P}_n| \leq (1-\delta)\sqrt{n}\right) \\ &\quad = \mathbb{P}\left(\sum_{j=1}^{n} \mathbb{I}\left(\|\mathbf{h}_j\|^2 \geq M\rho \left(\frac{1}{2}\log n - 2\log\log n \right) \right) \leq (1-\delta)\sqrt{n} \right)\end{aligned}$$

Using the tail lower bound on $\|\mathbf{h}_j\|^2$, we see that each indicator variable is i.i.d. with mean at least $\frac{\log^2 n}{\sqrt{n}}$. Therefore, using Chernoff bound,

$$\mathbb{P}\left(|\mathcal{P}_n| \leq (1-\delta)\sqrt{n}\log^2 n\right) \leq O\left(e^{-\delta^2\sqrt{n}\log^2 n}\right)$$

Next, we consider the probability $\mathbb{P}\left(\mathcal{R}_n(1) = \varnothing \,\middle|\, |\mathcal{P}_n| \geq (1-\delta)\sqrt{n}\log^2 n\right)$. Since the users are uniformly distributed in a circle of radius R, $\mathbb{P}\left(r_{ij} \leq r\right) = \frac{r^2}{R^2}$ for sufficiently small $r > 0$, and consequently $\mathbb{P}\left(\phi_{ij} \geq x\right) = \frac{1}{R^2} x^{-\frac{2}{c}}$. Since h_j is independent from ϕ_{1j},

$$\begin{aligned}
&\mathbb{P}\left(\mathcal{R}_n(1) = \varnothing \,\middle|\, |\mathcal{P}_n| \geq (1-\delta)\sqrt{n}\log^2 n\right) = \left(1 - \mathbb{P}\left(\phi_{1j} \geq n^{\frac{c}{4}}\right)\right)^{(1-\delta)\sqrt{n}\log^2 n} \\
&= \left(1 - n^{-\frac{1}{2}}\right)^{(1-\delta)\sqrt{n}\log^2 n} \qquad = O\left(e^{-(1-\delta)\log^2 n}\right).
\end{aligned}$$

Then, choosing $\delta = \frac{1}{\log n}$, and by using independence of channels across i's,

$$\begin{aligned}
\mathbb{P}\left(\mathcal{R}_n(1) \neq \varnothing \;\; \forall i\right) &= \left(1 - O\left(e^{-(1-\delta)\log^2 n}\right) - O\left(e^{-\delta^2\sqrt{n}\log^2 n}\right)\right)^n \\
&= 1 - O\left(e^{-\log^2 n + 2\log n}\right)
\end{aligned}$$

which concludes our proof of the first claim.

To prove the second claim, we note that

$$\|h_i\|^2 = \rho \left\| \sum_{k=1}^{P} \xi_{i,k} \mathbf{e}(\theta_{i,k}) \right\|^2 \leq \rho \sum_{k=1}^{P} |\xi_{i,k}|^2 \, \|\mathbf{e}(\theta_{i,k})\|^2 = M\rho X_i,$$

where $X_i \sim \chi^2(2P)$. The second claim then follows by Proposition C.1.

C.2 Proof of Lemma B.4

Define $\alpha^Q = \sum_{(i,j)\in Q} \sum_{s:\substack{i\in s_1\\ j\in s_2}} \sum_{k\in\mathcal{K}} \sum_{z\in\mathcal{Z}} \alpha_{skz}$, and consider the following sequence of optimization problems, indexed by n (with a slight abuse of notation):

$$\text{maximize} \quad U_n\left(\alpha\right) := \sum_{i\in\mathcal{N}} U_i\left(\alpha\right) - \sum_{Q\in\mathcal{Q}} \exp\left\{n\left(\alpha^Q - 1\right)\right\} \tag{C.3}$$

$$\text{s.t.} \;\; \alpha_{skz} \geq 0, \;\; \sum_{s} \alpha_{skz} \leq p_k q_z, \;\; \alpha_{skz} = q_z \sum_{z'} \alpha_{skz'}, \;\; \forall s, k, z. \tag{C.4}$$

We will denote the optimal value of the optimization (C.3) with OPT_n. Further consider the corresponding sequence of scheduling policies π_n, that choose $s^* = \arg\max_{s\subseteq\mathcal{N}^2\times\{1,2\}} \widetilde{f}_n(s)$, where

$$\widetilde{f}_n(s) = \sum_{(i,j,m)\in s} \left(\mathbb{E}\left[R^{(i)}_{sK(t)Z(t)}\right] \frac{\partial U_i}{\partial r_i} + \frac{\partial U_j}{\partial \beta_j}\right)\Bigg|_{\substack{r_j = r_j(t-1)\\ \beta_j = \beta_j(t-1)}} - n \sum_{Q: s_{12}\cap Q \neq \varnothing} e^{n\left(\alpha^Q(t) - 1\right)}, \tag{C.5}$$

The empirical utility of the policy π_n up to time t is denoted by $U_n(t)$.

Proposition C.3. $\lim_{n\to\infty} \mathsf{OPT}_n = \mathsf{OPT}'$.

Proof. We first show that for any $\epsilon > 0$, $\mathsf{OPT}_n \geq \mathsf{OPT}' - \epsilon$ for large enough n. Consider the optimization (4.12), with the condition (4.14) replaced by

$$\alpha^Q \leq 1 + \Delta,\ \forall Q \in \mathcal{Q}, \tag{C.6}$$

and denote the optimal value of the resulting maximization as OPT^{Δ}. By continuity of the objective function, for any $\epsilon > 0$, there exists $\delta > 0$ such that $\left|\mathsf{OPT}^{-\delta} - \mathsf{OPT}\right| < \frac{\epsilon}{2}$. For such δ, choose n large enough so that $e^{-n\delta} < \frac{\epsilon}{2|\mathcal{Q}|}$. Similarly, denote the maximal value of (C.3) subject to (C.6) as OPT_n^{Δ}. Then

$$\mathsf{OPT}_n \geq \mathsf{OPT}_n^{-\delta} \geq \mathsf{OPT}^{-\delta} - \frac{\epsilon}{2} \geq \mathsf{OPT}' - \epsilon.$$

Next, we show that for large enough n, $\mathsf{OPT}_n \leq \mathsf{OPT}' + \epsilon$. Choose $\delta > 0$ small enough so that $\left|\mathsf{OPT}^{\delta} - \mathsf{OPT}'\right| < \epsilon$. Hence

$$\mathsf{OPT}' + \epsilon \geq \mathsf{OPT}^{\delta} \geq \mathsf{OPT}_n^{\delta}.$$

Therefore it is sufficient to show that $\mathsf{OPT}_n^{\delta} = \mathsf{OPT}_n$ for large enough n. If we choose n large enough so that

$$\left.\frac{\partial U_n(\alpha)}{\partial \alpha^Q}\right|_{\alpha^Q > 1+\delta} = \sum_{i\in\mathcal{N}} \frac{\partial U_i(\alpha)}{\partial \alpha^Q} - n e^{n\left(\alpha^Q - 1\right)}\Big|_{\alpha^Q > 1+\delta} < 0,$$

then concavity implies $\mathsf{OPT}_n^{\delta} = \mathsf{OPT}_n$, since the derivative would have to be monotonically decreasing with increasing α^Q. Such a choice of n is possible since $\left.\frac{\partial U_i(\alpha)}{\partial \alpha^Q}\right|_{\alpha^Q = 1+\delta} < \infty$, similarly by concavity and twice continuous differentiability, which concludes the proof. □

Proposition C.4. $\lim_{n\to\infty} U_n(t) = U(t)$.

Proof. It is sufficient to show that for a given t, for a sufficiently large n, all the control actions taken by policies π_n and π^* up to time t are identical. Note that since the sets $\mathcal{K}$, $\mathcal{Z}$ and $\mathcal{N}$ are finite, for a finite t, there are finitely many values $\alpha_{skz}(t)$, and therefore $f_s(t)$ can take. Therefore we can choose n large enough so that

1. For any $\tau \leq t$, if $\alpha^Q(\tau) > 1$ for some Q, then

   $$f_s(\tau) - n \exp\left\{n\left(\alpha^Q - 1\right)\right\} < f_{s^*}(\tau)$$

   for all subsets s such that $s_{12} \cap Q \neq \varnothing$,

2. For each pair of subsets $s, t \subseteq \bar{\mathcal{N}}(t) \times \{1, 2\}$ such that $f_s(\tau) > f_t(\tau)$ and $\alpha^Q(\tau) < 1$ for all Q s.t. $s_{12} \cap Q \neq \varnothing$ and $t_{12} \cap Q \neq \varnothing$,

   $$\begin{aligned} f_s(\tau) - n &\sum_{Q: s_{12} \cap Q \neq \varnothing} \exp\{n\left(\alpha^Q(\tau) - 1\right)\} \\ &> f_t(\tau) - n \sum_{Q: t_{12} \cap Q \neq \varnothing} \exp\{n\left(\alpha^Q(\tau) - 1\right)\}. \end{aligned}$$

Here, the first condition ensures that a subset that violates any of the clique constraints is never scheduled, and the second condition ensures that for the subsets whose scheduling does not violate any of the clique constraints, the order with respect to f is preserved, and hence the subset that maximizes f remains the same. This is possible since for $x > 0$, e^{nx} can be made arbitrarily large, whereas for $x < 0$, it can be made arbitrarily small by scaling n. For such n, all scheduling decisions of π^* and π_n up to time t are identical, and thus $U_n(t) = U(t)$ for n sufficiently large. □

Proposition C.5. $\lim_{t\to\infty} U_n(t) = \mathsf{OPT}_n$.

Proof. The proof uses Lyapunov optimization techniques from [BGT95, TG05]. We will make use of the following theorem from [BGT95] to show the result.

Theorem C.1. *Consider a stochastic sequence in $\mathbb{R}^p$ satisfying the recursion*

$$\alpha(t) = \alpha(t-1) + \frac{1}{t}\mathbf{g}(t),$$

and let $\{\mathcal{F}_t\}_{t\geq 0}$ be a non-decreasing family of filtrations of the underlying σ-algebra, such that $\mathbf{g}(t)$ is $\mathcal{F}_t$-measurable.

Assume the following are satisfied.

1. *There exists a compact set $\mathcal{A} \subseteq \mathbb{R}^p$ such that*

$$\lim_{t\to\infty} \inf\{\|\alpha(t) - \alpha\|_1 : \alpha \in \mathcal{A}\} = 0,$$

2. *There exists $K > 0$ such that for all t, $\|\mathbf{g}(t)\|_1 \le K$,*

3. *There exists a twice continuously differentable function $V : \mathbb{R}^p \to \mathbb{R}$ such that*

$$\mathbb{E}\left[\mathbf{g}^\top(t+1)|\mathcal{F}_t\right] \nabla V\left(\alpha(t)\right) < -V\left(\alpha(t)\right),$$

   where $^\top$ represents vector transpose.

Then the function V in condition 3 satisfies $\lim_{t\to\infty} V\left(\alpha(t)\right)^+ = 0$.

Consider the sequence of vectors $\alpha(t) = \{\alpha_{skz}(t)\}_{s,k,z}$, whose entries satisfy the recursion

$$\alpha_{skz}(t) = \alpha_{skz}(t-1) + \frac{1}{t}\left(\mathbb{I}_{\mathcal{S}(t)=s}\mathbb{I}_{K(t)=k}\mathbb{I}_{Z(t)=z} - \alpha_{skz}(t-1)\right).$$

Note that the vector $\alpha(t)$ converges to the compact set defined by (4.13)–(4.14), by the first claim of Theorem 4.4, and the entries of the corresponding update sequence $\mathbf{g}(t)$ in this case is bounded by 1. Following the strategy of [TG05], we choose

$$\begin{aligned}
V(y(t)) = &\sum_{i\in\mathcal{N}} U_i\left(\sum_{s:i\in s_1}\sum_{k\in\mathcal{K}}\sum_{z\in\mathcal{Z}} R^{(i)}_{skz}\alpha^*_{skz}, \sum_{s:i\in s_2}\sum_{k\in\mathcal{K}}\sum_{z\in\mathcal{Z}} \alpha^*_{skz}\right) \\
&- \sum_{i\in\mathcal{N}} U_i\left(\sum_{s:i\in s_1}\sum_{k\in\mathcal{K}}\sum_{z\in\mathcal{Z}} R^{(i)}_{skz}\alpha_{skz}(t), \sum_{s:i\in s_2}\sum_{k\in\mathcal{K}}\sum_{z\in\mathcal{Z}} \alpha_{skz}(t)\right) \\
&- \sum_{Q\in\mathcal{Q}} \exp\left\{n\left(\alpha^{*Q} - 1\right)\right\} + \sum_{Q\in\mathcal{Q}} \exp\left\{n\left(\alpha^{Q}(t) - 1\right)\right\},
\end{aligned}$$

where α^* is the solution to (C.3)[1]. Then, if we verify the third condition for this choice of V, then the proof is concluded using Theorem C.1.

We first evaluate the terms in the left-hand side of the third condition.

$$\mathbb{E}\left[g^\top_{skz}(t+1)|\mathcal{F}_t\right] = \mathbb{E}\left[\mathbb{I}_{\mathcal{S}(t+1)=s}\mathbb{I}_{K(t+1)=k}\mathbb{I}_{Z(t+1)=z}|\mathcal{F}_t\right] - \alpha_{skz}(t) =$$

[1] Since (C.3) is the maximization of a continuous function over a compact set, the extreme values are attained within the feasible set.

$$\sum_{b\in\mathcal{K},c\in\mathcal{Z}} \mathbb{E}\left[\mathbb{I}_{\mathcal{S}(t+1)=s}\mathbb{I}_{K(t+1)=k}\mathbb{I}_{Z(t+1)=z}|K(t+1)=b,Z(t+1)=c,\mathcal{F}_t\right]p_bq_c-\alpha_{skz}(t)$$
$$=\mathbb{E}\left[\mathbb{I}_{\mathcal{S}(t+1)=s}|K(t+1)=k,Z(t+1)=z,\mathcal{F}_t\right]p_kq_z-\alpha_{skz}(t)$$
$$=\begin{cases} p_kq_z-\alpha_{skz}(t), & \text{if } s=s^*,\\ -\alpha_{skz}(t), & \text{otherwise}\end{cases}$$

where $s^* = \arg\max_{\widetilde{s}\in\bar{\mathcal{N}}(t+1)\times\{1,2\}} \widetilde{f}_n(\widetilde{s})$. Since a single entry of $\nabla V\left(\alpha(t)\right)$ is given by

$$\begin{aligned}D_{skz} &:= \frac{\partial V\left(\alpha(t)\right)}{\partial\alpha_{skz}(t)}\\ &= -\sum_{i\in s_1} R^{(i)}_{skz}\frac{\partial U_i}{\partial r_i}\Big|_{r_i=r_i(t)} - \sum_{i\in s_2}\frac{\partial U_i}{\partial\beta_i}\Big|_{\beta_i=\beta_i(t)} + n\sum_{(i,j)\in s_{12}}\sum_{Q:(i,j)\in Q} e^{n\left(\alpha^Q(t)-1\right)},\end{aligned}$$

and the inner product on the left-hand side of the third condition can be expressed as

$$\begin{aligned}
&\mathbb{E}\left[g^\top(t+1)|\mathcal{F}_t\right]\nabla V\left(\alpha(t)\right) = -\sum_{k\in\mathcal{K}}\sum_{z\in\mathcal{Z}} D_{s^*kz}p_kq_z + \sum_{k\in\mathcal{K}}\sum_{z\in\mathcal{Z}}\sum_{s} D_{skz}\alpha_{skz}(t)\\
&= -\sum_{k\in\mathcal{K}}\mathbb{E}_L\left[D_{s^*kL}\right]p_k + \sum_{k\in\mathcal{K}}\sum_{z\in\mathcal{Z}}\sum_{s} D_{skz}\alpha_{skz}(t)\\
&\leq -\sum_{k\in\mathcal{K}}\sum_{z\in\mathcal{Z}}\sum_{s}\mathbb{E}_L\left[D_{s^*kL}\right]\alpha^*_{skz} + \sum_{k\in\mathcal{K}}\sum_{z\in\mathcal{Z}}\sum_{s} D_{skz}\alpha_{skz}(t)\\
&\leq -\sum_{k\in\mathcal{K}}\sum_{z\in\mathcal{Z}}\sum_{s}\mathbb{E}_L\left[D_{skL}\right]\alpha^*_{skz} + \sum_{k\in\mathcal{K}}\sum_{z\in\mathcal{Z}}\sum_{s} D_{skz}\alpha_{skz}(t)\\
&\leq -\sum_{k\in\mathcal{K}}\sum_{z\in\mathcal{Z}}\sum_{s}\sum_{z'\in\mathcal{Z}} q_{z'}D_{skz'}\alpha^*_{skz} + \sum_{k\in\mathcal{K}}\sum_{z\in\mathcal{Z}}\sum_{s} D_{skz}\alpha_{skz}(t)\\
&\overset{\text{(a)}}{=} -\sum_{k\in\mathcal{K}}\sum_{s}\sum_{z'\in\mathcal{Z}} D_{skz'}\alpha^*_{skz'} + \sum_{k\in\mathcal{K}}\sum_{z\in\mathcal{Z}}\sum_{s} D_{skz}\alpha_{skz}(t)\\
&= -\sum_{k\in\mathcal{K}}\sum_{s}\sum_{z\in\mathcal{Z}}\frac{\partial V\left(\alpha(t)\right)}{\partial\alpha_{skz}(t)}\left(\alpha^*_{skz}-\alpha_{skz}(t)\right) \overset{\text{(b)}}{\leq} -V\left(\alpha(t)\right)
\end{aligned}$$

where (a) follows by the third constraint in (C.4), and (b) follows by convexity. □

Finally, we can prove that $U(t)\to\mathsf{OPT}'$. Note that this is equivalent to the statement

$$\lim_{n\to\infty}\lim_{t\to\infty}U_n(t) = \lim_{t\to\infty}\lim_{n\to\infty}U_n(t).$$

Given $\epsilon > 0$, using Propositions C.3, C.4, and C.5, we can find sufficiently large n and t such that

$$|U(t) - \mathsf{OPT}| \leq |U(t) - U_n(t)| + |U_n(t) - \mathsf{OPT}_n| + |\mathsf{OPT}_n - \mathsf{OPT}| < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon,$$

which concludes the proof.

C.3 Proof of Theorem 4.1

Proof. The upper bound follows by the fact that R_{MIMO} is achievable. To prove the lower bound, we first note that for any input convariance matrix $\mathbf{Q}$,

$$\sigma^2_{2|1} = \frac{|\Sigma|}{\Sigma_{11}} = \frac{|\mathbf{I} + \mathbf{HQH}^*|}{1 + \|\mathbf{h}_1\|^2}, \tag{C.7}$$

and that $\mathbf{K}^{-1} = \operatorname{diag}(1, \eta)$, where $\eta = \frac{1}{1+\frac{\sigma^2_{2|1}}{|g_{12}|^2}}$. Next, we lower bound R_{MIMO} as follows.

$$\begin{aligned} R_{\text{MIMO}} &= \log\left|\mathbf{I}_2 + \mathbf{K}^{-1}\mathbf{HQH}^*\right| \overset{\text{(a)}}{\geq} \log\left|\mathbf{K}^{-1} + \mathbf{K}^{-1}\mathbf{HQH}^*\right| \\ &\geq \log|\mathbf{I}_2 + \mathbf{HQH}^*| + \log\eta, \end{aligned} \tag{C.8}$$

To see why (a) holds, define $\mathbf{P} := \mathbf{K}^{-1} - \mathbf{I}_2$, and denote by $\lambda_k(\mathbf{A})$ the k'th largest eigenvalue for a matrix $\mathbf{A}$. Then by Weyl's inequality, since $\eta \leq 1$,

$$\begin{aligned} \lambda_k\left(\mathbf{P} + \mathbf{I}_2 + \mathbf{K}^{-1}\mathbf{HQH}^*\right) &\leq \lambda_k(\mathbf{I}_2 + \mathbf{K}^{-1}\mathbf{HQH}^*) + \lambda_1(\mathbf{P}) \\ &= \lambda_k(\mathbf{I}_2 + \mathbf{K}^{-1}\mathbf{HQH}^*), \end{aligned}$$

which implies latter determinant in (C.8) is smaller. Next, note that η can be lower bounded by

$$\eta \leq \begin{cases} \frac{|g_{12}|^2}{2\sigma^2_{2|1}} & \text{if } \sigma^2_{2|1} \geq |g_{12}|^2 \\ \frac{1}{2} & \text{otherwise} \end{cases} \tag{C.9}$$

Then, combining (C.7), (C.8), and (C.9), we can show that R_{MIMO} is lower bounded by

$$R_{\text{MIMO}} \geq \min\left\{\max_{\operatorname{tr}(\mathbf{Q})\leq 1} \log|\mathbf{I}_2 + \mathbf{HQH}^*|, \log\left(1 + \|\mathbf{h}_1\|^2\right) + \log^+\left(|g_{12}|^2\right)\right\} - 1,$$

where $\log^+(x) := \max(0, \log(x))$. We conclude the proof by noting that for any $x \geq 0$, $\log^+(x) \geq \log(1+x) - 1$, and by the fact that the capacity $\bar{C}$ is upper bounded by the cut-set bound [Ct79], given by

$$\bar{C} \leq \min \left\{ \max_{\operatorname{tr}(\mathbf{Q}) \leq 1} \log |\mathbf{I}_2 + \mathbf{H}\mathbf{Q}\mathbf{H}^*| , \log\left(1 + \|\mathbf{h}_1\|^2\right) + \log\left(1 + |g_{12}|^2\right) \right\}.$$

□

C.4 Proofs of Lemmas 4.1 and 4.2

C.4.1 Proof of Lemma 4.2

For any $n \in \mathbb{N}$, let π_n be a feasible policy such that $\liminf_{t\to\infty} U^{\pi_n}(t) \geq \mathsf{OPT} - \frac{1}{2n}$. Then by definition, there must exist T_n such that for $t > T_n$, $U^{\pi_n}(t) \geq \mathsf{OPT} - \frac{1}{n}$. Consider the sequence $\alpha^{\pi_n}(T_n)$, where $U^{\pi_n}(t) = U(\alpha^{\pi_n}(t))$. Let the set of vectors α defined by (4.13) and (4.14) be $\mathcal{Y}$. Then strong law of large numbers, and the independence of $(\mathcal{S}(t), K(t))$ from $Z(t)$ implies $\lim_{n\to\infty} \inf \{\|\alpha - \alpha^{\pi_n}(T_n)\| : \alpha \in \mathcal{Y}\} = 0$. Therefore, there exists a sequence $\{\alpha_n\} \in \mathcal{Y}$ such that $\lim_{n\to\infty} \|\alpha_n - \alpha^{\pi_n}(T_n)\| = 0$. Since $\mathcal{Y}$ is closed and bounded, it is compact, and therefore α_n must have a subsequence, say α_{n_k}, that converges to a point $\alpha^* \in \mathcal{Y}$, which implies

$$\lim_{k\to\infty} \alpha^{\pi_{n_k}}(T_{n_k}) = \alpha^* \in \mathcal{Y}.$$

Since the function U is continuous, we have

$$\mathsf{OPT} = \lim_{k\to\infty} U\left(\alpha^{\pi_{n_k}}(T_{n_k})\right) = U\left(\lim_{k\to\infty} \alpha^{\pi_{n_k}}(T_{n_k})\right) = U\left(\alpha^*\right).$$

Since α^* is in the feasible set $\mathcal{Y}$, it must be that $\mathsf{OPT}' \geq U(\alpha^*) = \mathsf{OPT}$.

C.4.2 Proof of Lemma 4.1

Assume that there exists $\epsilon > 0$, $Q \in \mathcal{Q}$, such that for any N, there exists $t > N$ that satisfies $\beta^*_Q(t) > 1 + \epsilon$. Note that

$$\beta^*_Q(t) \leq \frac{t-1}{t}\beta^*_Q(t-1) + \frac{|Q|}{tp}\mathbb{I}_{\beta^*_Q(t-1)<1}, \tag{C.10}$$

with $p = \min_{(i,j)\in Q} p_{ij}$, where the upper bound is obtained by observing that the maximal increase in $\beta^*_Q(t)$ is achieved when all flows $(i,j) \in Q$ are scheduled at slot t. Choosing $N = \frac{|Q|}{\epsilon p}$, there must exist $t > N$ s.t. $\beta^*_Q(t) > 1 + \epsilon$. Letting $t^* \geq N$ to be the smallest of such indices, it must be that $\beta^*_Q(t^* - 1) \leq 1$, since otherwise the increment $\beta^*_Q(t) - \beta^*_Q(t-1)$ cannot be positive, by construction. But by (C.10) and by the choice of N,

$$\beta^*_Q(t^*) \leq \frac{t^*-1}{t^*}\beta^*_Q(t^*-1) + \epsilon\mathbb{I}_{\beta^*_Q(t^*-1)<1} \leq 1 + \epsilon,$$

which is a contradiction.

C.5 Utility Function with Relaying Cost

For an arbitrary κ, let $\left(\widetilde{\mathbf{r}}, \widetilde{\beta}\right)$ solve the optimization (4.8) with $U_i\left(r_i, \beta_i\right) = \log(r_i) + \kappa\log(1-\beta_i)$, where

$$r_i = \sum_{s:i\in s_1}\sum_{k\in\mathcal{K}}\sum_{z\in\mathcal{Z}} R^{(i)}_{skz}\alpha_{skz}, \quad \beta_i = \sum_{s:i\in s_2}\sum_{k\in\mathcal{K}}\sum_{z\in\mathcal{Z}}\alpha_{skz}.$$

Note that here α_{skz} has no time dependence and refers to a deterministic quantity, *i.e.*, the fraction of time for which $\mathcal{S}(t) = s, K(t) = k, Z(t) = z$, throughout the (infinite) duration of transmission. Then, for any feasible perturbation $\delta\alpha$ that pushes the operating point from $\left(\widetilde{\mathbf{r}}, \widetilde{\beta}\right)$ to $(\mathbf{r}, \beta)$, it must be that $\sum_{s,k,z}\delta\alpha_{skz}\sum_i \frac{\partial U_i}{\partial\alpha_{skz}} \leq 0$ by concavity, which, using the facts

$$r_i - \widetilde{r}_i = \delta r_i = \sum_{s:i\in s_1}\sum_{(k,z)\in\mathcal{K}\times\mathcal{Z}} R^{(i)}_{skz}\delta\alpha_{skz}, \qquad \beta_i - \widetilde{\beta}_i = \delta\beta_i = \sum_{s:i\in s_2}\sum_{(k,z)\in\mathcal{K}\times\mathcal{Z}}\delta\alpha_{skz}$$

can be re-arranged into

$$\sum_i \frac{r_i - \widetilde{r}_i}{\widetilde{r}_i} \leq \kappa\sum_i \frac{(1-\widetilde{\beta}_i) - (1-\beta_i)}{1-\widetilde{\beta}_i}. \tag{C.11}$$

C.6 Proof of Theorem 4.3

Before we present the proof, we need several definitions.

Definition C.1. *The chromatic number $\chi(\mathcal{G})$ is the minimum number of colors needed to color graph $\mathcal{G}$.*

Definition C.2. *The clique number $\omega(\mathcal{G})$ is the maximum clique size in $\mathcal{G}$.*

Definition C.3. *A* perfect graph *is a graph whose chromatic number equals its clique number,* i.e., *$\chi(\mathcal{G}) = \omega(\mathcal{G})$.*

Definition C.4. *A graph is* chordal *if, for every cycle of length larger than three, there is an edge that is not part of the cycle, connecting two of the vertices of the cycle.*

Given these definitions, we are ready for the proof. The results in [TE92] can be used to show that the stability region of the constrained queueing network formed by the n users is given by

$$\Lambda = \left\{\beta : \mathbf{D}^{-1}\beta \in conv\,(\Pi)\right\}, \tag{C.12}$$

where $\mathbf{D}$ is a diagonal matrix with p_{ij} values on the diagonal ($p_{ij} > 0$ without loss of generality), $conv(\cdot)$ represents the convex hull of a set of vectors, and Π is the set of incidence vectors of the independent sets of $\mathcal{G}_c$, *i.e.*, a vector $\mathbf{s}$ whose elements are indexed by (i, j) is contained in Π if $\left\{(i, j) : s_{(i,j)} = 1\right\}$ is an independent set of ${\mathcal{G}_c}^2$.

The set Λ as defined in (C.12) is known as the stable set polytope of the graph $\mathcal{G}_c$. The exact characterization of Λ is not known in general [Reb08]. However, stable set polytopes of perfect graphs can be completely described in terms of their maximal cliques, as characterized in the following theorem.

[2] The boundary of the stability region is included in the set Λ for technical convenience. Note that this does not change the supremum value in the optimization (4.7) since the objective function is continuous.

Theorem C.2. *[Chv75] Let $\mathcal{Q}$ be the set of maximal cliques of a perfect graph $\mathcal{G}$. Then the stable set polytope of $\mathcal{G}$ is the set of vectors $x \in [0,1]^{|\mathcal{V}|}$ satisfying $\sum_{v \in Q} x_v \leq 1$ for all $Q \in \mathcal{Q}$.*

Therefore, to complete the proof, it is sufficient to show that there exists a polynomial-time procedure that adds edges in $\mathcal{G}_c$ such that the resulting graph $\bar{\mathcal{G}}_c$ is perfect[3].

It is known that chordal graphs are perfect [Ber61], and any graph can be made into a chordal one in polynomial time by inserting edges[4]. Further, the number of maximal cliques in a chordal graph is upper bounded by the number of nodes (equal to $n(n-1)$ for $\mathcal{G}_c$) [Gav74], and the maximal cliques of a chordal graph can be listed in polynomial time [RS07], which concludes the proof.

[3] The fact that $\Lambda\left(\bar{\mathcal{G}}_c\right) \subseteq \Lambda\left(\mathcal{G}_c\right)$ follows directly from the fact that $\mathcal{E}_c \subseteq \bar{\mathcal{E}}_c$

[4] For instance, one can iterate over the vertices, in each iteration connecting all the previously unvisited neighbors of the current vertex to each other. It is easy to show that such a procedure outputs a chordal graph.

APPENDIX D

Proofs for Chapter 5

D.1 Proofs of Theorem 5.1 and Corollary 5.1

The proof is based on a variation of the proof of the main result in [PW15]; however, unlike the proof therein, we make use of the properties of tight frames.

Fix a failure pattern A. We first note that since the rows of S form a tight frame, $S^\top S = \beta I_n$. Recalling that $s_i^\top$ is the ith row of S,

$$\begin{aligned} S_A^\top S_A &= \sum_{i \in A} s_i s_i^\top = \sum_{i=1}^{n\beta} s_i s_i^\top - \sum_{i \notin A} s_i s_i^\top \\ &= S^\top S - S_{A^c}^\top S_{A^c} = \beta I_n - S_{A^c}^\top S_{A^c}. \end{aligned} \tag{D.1}$$

Denoting the minimum and maximum eigenvalues of a matrix by $\lambda_{\min}(\cdot)$ and $\lambda_{\max}(\cdot)$ respectively, and using (D.1), any unit vector u satisfies

$$\|S_{A^c} u\|^2 \geq \lambda_{\min}\left(S_{A^c}^\top S_{A^c}\right) = \beta - \lambda_{\max}\left(S_A^\top S_A\right). \tag{D.2}$$

Defining $e = \hat{\theta} - \theta^*$, we have

$$\|X\hat{\theta} - y\| \leq \left(1 + \frac{\|Xe\|}{\|X\theta^* - y\|}\right) \|X\theta^* - y\|,$$

by triangle inequality. Therefore

$$\eta(S; X, y; \alpha) \leq \left(1 + \frac{\|Xe\|}{\|X\theta^* - y\|}\right)^2. \tag{D.3}$$

For any $0 \leq c \leq \beta$, consider

$$\|Xe\|^2 \overset{\text{(a)}}{\leq} \frac{\|S_{A^c} X e\|^2}{\beta - \lambda_{\max}\left(S_A^\top S_A\right)}$$

$$\begin{aligned}
&\overset{\text{(b)}}{\leq} -2\frac{e^\top X^\top S_{A^c}^\top S_{A^c}(X\theta^* - y)}{\beta - \lambda_{\max}\left(S_A^\top S_A\right)} \\
&= -2\frac{e^\top X^\top \left(S_{A^c}^\top S_{A^c} - (\beta - c)\, I\right)(X\theta^* - y)}{\beta - \lambda_{\max}\left(S_A^\top S_A\right)} \\
&\qquad\qquad - \frac{2\,(\beta - c)}{\beta - \lambda_{\max}} e^\top X^\top (X\theta^* - y) \\
&\overset{\text{(c)}}{\leq} -2\frac{e^\top X^\top \left(S_{A^c}^\top S_{A^c} - (\beta - c)\, I\right)(X\theta^* - y)}{\beta - \lambda_{\max}\left(S_A^\top S_A\right)} \\
&\overset{\text{(d)}}{=} 2\frac{e^\top X^\top \left(S_A^\top S_A - cI\right)(X\theta^* - y)}{\beta - \lambda_{\max}\left(S_A^\top S_A\right)} \\
&\overset{\text{(e)}}{\leq} 2\frac{\left\|e^\top X^\top \left(S_A^\top S_A - cI\right)\right\|}{\beta - \lambda_{\max}\left(S_A^\top S_A\right)} \|X\theta^* - y\| \\
&\overset{\text{(f)}}{\leq} \frac{2\lambda_{\max}^{X\mathcal{K}}\left(S_A^\top S_A - cI\right)}{\beta - \lambda_{\max}\left(S_A^\top S_A\right)} \|Xe\|\|X\theta^* - y\|,
\end{aligned}$$

where (a) follows by (D.2); (b) follows by re-arranging $\|S_{A^c}(X\hat{\theta} - y)\|^2 \leq \|S_{A^c}(X\theta^* - y)\|^2$, which is true because of the optimality of $\hat{\theta}$ for the encoded problem; (c) follows by the convex optimality condition

$$\langle X^\top (X\theta^* - y), e\rangle = \langle \nabla g(\theta^*), \hat{\theta} - \theta^*\rangle \geq 0;$$

(d) follows by (D.1); (e) follows by Cauchy-Schwarz inequality; and (f) follows by the definition of $\lambda_{\max}^{X\mathcal{K}}$, and the fact that $\hat{\theta}$ is feasible, so $e \in \mathcal{K}$. This bound, together with (D.3), implies Theorem 5.1 by minimizing over all possible choices of c.

To prove Corollary 5.1, first note that the bound is maximized when $X\mathcal{K}$ contains the eigenvector of $\left(S_A^\top S_A - cI\right)$ corresponding to the largest eigenvalue. Choose X to map an arbitrary $e \in \mathcal{K}$ to this eigenvector, which implies $\lambda_{\max}^{X\mathcal{K}}(S_A^\top S_A - cI) = \lambda_{\max}(S_A^\top S_A - cI)$ (recall that $\lambda_{\max}$ refers to the maximum *absolute value* of the eigenvalues, hence equivalent to operator norm for any symmetric matrix). Further choose $c = \frac{1}{2}\lambda_{\max}(S_A^\top S_A)$ to get

$$\begin{aligned}
\gamma(S, e) &\leq \min_{0 \leq c \leq \beta} \max_{|A| = e} \left(1 + \frac{2\lambda_{\max}\left(S_A^\top S_A - cI\right)}{\beta - \lambda_{\max}\left(S_A^\top S_A\right)}\right)^2 \\
&= \max_{|A| = e} \left(1 + \frac{2\lambda_{\max}\left(S_A^\top S_A - \frac{1}{2}\lambda_{\max}(S_A^\top S_A) I\right)}{\beta - \lambda_{\max}\left(S_A^\top S_A\right)}\right)^2
\end{aligned}$$

$$\overset{\text{(g)}}{=} \max_{|A|=e} \left(1 + \frac{\lambda_{\max}\left(S_A^\top S_A\right)}{\beta - \lambda_{\max}\left(S_A^\top S_A\right)} \right)^2$$

$$= \max_{|A|=e} \left(\frac{\beta}{\beta - \lambda_{\max}\left(S_A^\top S_A\right)} \right)^2 ,$$

where (g) follows by the fact that all eigenvalues of $S_A^\top S_A$ are between 0 and $\lambda_{\max}(S_A^\top S_A)$ and thus the absolute values of all eigenvalues of $S_A^\top S_A - \frac{1}{2}\lambda_{\max}(S_A^\top S_A)I$ are upper bounded by $\frac{1}{2}\lambda_{\max}(S_A^\top S_A)$.

D.2 Proof of Theorem 5.2

First we would like to bound $\left\|S_A S_A^\top - I_{e\ell}\right\|_2$. Note that the (i,j)th element of $S_A S_A^\top$ is given by $\langle s_i, s_j \rangle$ for $i \neq j$, where $s_i^\top$ is the ith row of S_A, and the diagonal of $S_A S_A^\top - I_{e\ell}$ consists of zeros. Since S is equiangular, Proposition 5.1 implies that $|\langle s_i, s_j\rangle| = \sqrt{\frac{\beta-1}{n\beta-1}}$. Then by Gershgorin circle theorem, all eigenvalues $\{\lambda_k\}$ of $S_A S_A^\top - I_{e\ell}$ satisfy

$$|\lambda_k| \leq \sum_{j=1}^{e\ell} |\langle s_i, s_j \rangle| = e\ell \sqrt{\frac{\beta-1}{n\beta-1}},$$

which, using the fact $\ell = \frac{n\beta}{m}$ implies,

$$\left\|S_A S_A^\top - I_{e\ell}\right\|_2 \leq \frac{e}{m}\sqrt{\frac{n\beta(\beta-1)}{1-\frac{1}{n\beta}}}.$$

Using triangle inequality,

$$\left\|S_A^\top S_A\right\|_2 = \left\|S_A S_A^\top\right\|_2 \leq 1 + \frac{e}{m}\sqrt{\frac{n\beta(\beta-1)}{1-\frac{1}{n\beta}}}.$$

Plugging in this bound in Corollary 5.1 gives the desired result.

D.3 Proof of Theorem 5.3

Recall that by construction, $2n-1 = q$ is a prime such that $q \equiv 1 \pmod 4$. For any row index $1 \leq i \leq q$, define $\kappa(i)$ as the index of the node row i of S corresponds to, *i.e.*, $\kappa(i) := \lceil \frac{i}{\ell} \rceil$.

Further define $\pi : [2n] \to [2n]$ to be a random permutation of the integers $\{1, \ldots, 2n\}$, which is uniform over each of the $(2n)!$ realizations.

Let J_i be the 0-1 indicator variable denoting whether node i is unavailable, (*i.e.*, $J_i = 1$ if and only if $i \in A$), and $J := \{J_i\}_{i=1}^m$. Given e, we assume J takes uniformly at random one of the $\binom{m}{e}$ vector values consisting of e 1's, and $m - e$ 0's. Note that J_i and J_j are not independent for $i \neq j$.

Given a finite field $\mathbb{F}_q$, $a \in \mathbb{F}_q$ is called a *quadratic residue* if there exists $r \in \mathbb{F}_q$ such that $a \equiv r^2 \pmod{q}$. Construct the matrix $L \in \{-1, 0, 1\}^{2n \times 2n}$ such that

$$L_{ij} = \begin{cases} \chi(i-j), & 1 \leq i, j \leq q \\ \mathbb{I}_{i \neq j}, & \text{if } i = q+1 \text{ or } j = q+1 \end{cases}$$

where χ is the quadratic residue character in $\mathbb{F}_q$, defined by

$$\chi(x) = \begin{cases} 0, & \text{if } x = 0, \\ 1, & \text{if } x \neq 0 \text{ is a quadratic residue in } \mathbb{F}_q, \\ -1, & \text{otherwise.} \end{cases}$$

In the above definition, we have assumed that the $(q+1)$th index corresponds to the isolated node appended to the Paley graph.

Characters are important objects of study in analytic number theory (see, *e.g.*, [Apo13] for more information on characters). In particular, quadratic residue character χ is a *multiplicative* character, satisfying the following properties, which can be easily verified:

1. $\chi(1) = 1$,
2. For $a, b \in \mathbb{F}_q$, $\chi(a)\chi(b) = \chi(ab)$,
3. For $a \in \mathbb{F}_q$, $\chi(a) = \chi(a^{-1})$.

Proposition D.1 ([Apo13]). *Let q be an odd prime. The quadratic residue character χ over $\mathbb{F}_q$ satisfies*

$$\sum_{a \in \mathbb{F}_q} \chi(a) = 0.$$

Define

$$\bar{L} := \left[L_{ij} J_{\kappa(\pi(i))} J_{\kappa(\pi(j))}\right]_{i,j}.$$

Note that the matrix

$$\sqrt{q}\left(S_A S_A^\top - I\right) \tag{D.4}$$

is identical to the realization of $\bar{L}$ corresponding to J such that $J_i = 1 \Leftrightarrow i \in A$, up to padding with zeroes. Therefore, they have the same spectrum and the problem reduces to characterizing the expected spectral norm of $\bar{L}$.

We will prove the following lemma.

Lemma D.1. *Let $a, b \in \mathbb{F}_q$. Then*

$$\sum_{x \in \mathbb{F}_q} \chi(a-x)\chi(b-x) = \left(-1 + q\mathbb{I}_{a=b}\right).$$

Proof. The case $a = b$ easily follows by the fact that

$$\sum_{x \in \mathbb{F}_q} \chi(a-x)\chi(a-x) = \sum_{x \in \mathbb{F}_q} \mathbb{I}_{a \neq x} = q - 1.$$

If $a \neq b$, using properties of $\chi(\cdot)$,

$$\begin{aligned}\sum_{x \in \mathbb{F}_q} \chi(a-x)\chi(b-x) &= \sum_{x \neq a,b} \chi(a-x)\chi(b-x) \\ = \sum_{x \neq a,b} \chi(a-x)\chi\left((b-x)^{-1}\right) &= \sum_{x \neq a,b} \chi\left(\frac{a-x}{b-x}\right) \\ = \sum_{x \neq a,b} \chi\left(1 + \frac{a-b}{b-x}\right) &\overset{(a)}{=} \sum_{y \neq 0,1} \chi(y) \overset{(b)}{=} -\chi(1) = -1,\end{aligned}$$

which completes the proof. (a) follows because in $\mathbb{F}_q$ every non-zero element has a unique multiplicative inverse, hence the argument of the character will take every value except 0 (since $x \neq a$) and 1 (since $a \neq b$); (b) follows by Proposition D.1. □

Now, consider

$$\mathbb{E}\left[\mathbf{tr}\left(\bar{L}^4\right)\right]$$

$$= \mathbb{E}\left[\sum_{i_1,\ldots,i_4} L_{i_1i_2}L_{i_2i_3}L_{i_3i_4}L_{i_4i_1}J_{\kappa(\pi(i_1))}\ldots J_{\kappa(\pi(i_4))}\right]$$

$$= \sum_{i_1,\ldots,i_4} L_{i_1i_2}L_{i_2i_3}L_{i_3i_4}L_{i_4i_1}\mathbb{E}\left[J_{\kappa(\pi(i_1))}\ldots J_{\kappa(\pi(i_4))}\right].$$

Note that, since π is uniformly random, we have

$$\mathbb{E}\left[J_{\kappa(\pi(i_1))}\ldots J_{\kappa(\pi(i_4))}\right] = \frac{\binom{e\ell}{s}}{\binom{m\ell}{s}} \leq \left(\frac{e}{m}\right)^s,$$

where s is the number of unique elements in the tuple (i_1, i_2, i_3, i_4). Therefore

$$\mathbb{E}\left[\mathbf{tr}\left(\bar{L}^4\right)\right] \leq \sum_{s=1}^{4}\left(\frac{e}{m}\right)^s \sum_{\substack{i_1,\ldots,i_4:\\ \{i_1,i_2,i_3,i_4\}=s}} L_{i_1i_2}L_{i_2i_3}L_{i_3i_4}L_{i_4i_1}$$

$$=: \sum_{s=1}^{4}\left(\frac{e}{m}\right)^s \phi(s),$$

where we have defined the inner sum as $\phi(s)$. First, note that $\phi(1) = 0$ by the fact that this would require all i_j to be equal, and $L_{i_ji_j} = 0$ by definition. Next, consider

$$\phi(2) = \sum_{\substack{i_1,\ldots,i_4:\\ \{i_1,i_2,i_3,i_4\}=2}} L_{i_1i_2}L_{i_2i_3}L_{i_3i_4}L_{i_4i_1}$$

$$= \sum_{a\neq b} L_{ab}L_{ba}L_{ab}L_{ba} = \sum_{a\neq b} L_{ab}^4 = q(q+1)$$

by the fact that L is symmetric and all the off-diagonal elements are ±1. Then

$$\phi(3) = \sum_{\substack{i_1,\ldots,i_4:\\ \{i_1,i_2,i_3,i_4\}=3}} L_{i_1i_2}L_{i_2i_3}L_{i_3i_4}L_{i_4i_1}$$

$$= \sum_{\substack{a\neq b\\ a\neq c\\ b\neq c}} L_{ab}L_{bc}L_{cb}L_{ba} + \sum_{\substack{a\neq b\\ a\neq c\\ b\neq c}} L_{ab}L_{ba}L_{ac}L_{ca}$$

$$= \sum_{\substack{a\neq b\\ a\neq c\\ b\neq c}} L_{ab}^2L_{bc}^2 + \sum_{\substack{a\neq b\\ a\neq c\\ b\neq c}} L_{ab}^2L_{ac}^2 = 2(q+1)q(q-1),$$

similarly by the symmetry and unit modulus of the elements of L. Now,

$$\sum_{s=1}^{4}\phi(s) = \sum_{i_1,\ldots,i_4} L_{i_1i_2}L_{i_2i_3}L_{i_3i_4}L_{i_4i_1}$$

$$
\begin{aligned}
&= \sum_{i_1,i_3} \left(\sum_{i_2} L_{i_1i_2} L_{i_2i_3} \right) \left(\sum_{i_4} L_{i_3i_4} L_{i_4i_1} \right) \\
&= q^2 \sum_{i_1,i_3} \mathbb{I}_{i_1=i_3} = q^2(q+1),
\end{aligned}
$$

where the third equality follows by Lemma D.1 and the definition of L_{ij}, which implies

$$
\sum_{j=1}^{q+1} L_{ij} L_{jk} = q\mathbb{I}_{i=k}.
$$

The above results then imply

$$
\begin{aligned}
\phi(4) &= q^2(q+1) - 2(q+1)q(q-1) - q(q+1) \\
&= -q(q+1)(q-1).
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\mathbb{E}\left[\mathbf{tr}\left(\bar{L}^4\right)\right] &= \sum_{s=1}^{4} \left(\frac{e}{m}\right)^s \phi(s) \\
&= \left(\frac{e}{m}\right)^2 q(q+1) \left(1 + \frac{e}{m}(q-1)\left(2 - \frac{e}{m}\right)\right) \\
&\leq 4 \left(\frac{e}{m}\right)^3 (q+1)^3.
\end{aligned}
$$

Then, defining λ_i to be the ith largest eigenvalue of $\bar{L}$, for any $a > 0$,

$$
\begin{aligned}
\mathbb{P}\left(\max_i |\lambda_i| > a\right) &= \mathbb{P}\left(\max_i |\lambda_i|^4 > a^4\right) \\
&\leq \mathbb{P}\left(\sum_i |\lambda_i|^4 > a^4\right) = \mathbb{P}\left(\mathbf{tr}\left(\bar{L}^4\right) > a^4\right) \\
&\overset{(a)}{\leq} \frac{\mathbb{E}\left[\mathbf{tr}\left(\bar{L}^4\right)\right]}{a^4} \leq \frac{4\left(\frac{e}{m}\right)^3 (q+1)^3}{a^4},
\end{aligned}
$$

where (a) is by Markov inequality. Therefore, setting

$$
a = c\sqrt{2}\left(\frac{e}{m}\right)^{3/4} (q+1)^{3/4} = c\sqrt{2}\,(e\ell)^{3/4}
$$

for some constant $c > 0$, using that for A uniformly random among all $\binom{m}{e}$ possible e failures,

$$
\lambda_{\max}\left(S_A^\top S_A\right) = \lambda_{\max}\left(S_A S_A^\top\right) = 1 + \frac{1}{\sqrt{q}} \lambda_{\max}\left(\bar{L}\right),
$$

we get

$$\mathbb{P}\left(\lambda_{\max}\left(S_A^\top S_A\right) > 1 + c\sqrt{\frac{2}{m-\frac{1}{\ell}}e^{3/4}\ell^{1/4}}\right) \leq \frac{1}{c^4},$$

which directly implies the result using Corollary 5.1.

D.4 Proof of Proposition 5.2

We will use the following result (slightly loosened and rephrased in our notation).

Lemma D.2 ([PW15], Lemma 1). *For any $c > 0$,*

$$\eta\left(S_m; X, y; A\right) \leq \left(1 + 2\frac{\|S_A^\top S_A - cI\|_2}{\lambda_{\min}\left(S_A^\top S_A\right)}\right)^2.$$

Using the results from [Sil85, Gem80], we know that

$$\begin{aligned}\lambda_{\max}\left(\left(S_m\right)_A^\top\left(S_m\right)_A\right) &\to \left(\sqrt{\beta\left(1-\frac{e}{m}\right)}+1\right)^2\\ \lambda_{\min}\left(\left(S_m\right)_A^\top\left(S_m\right)_A\right) &\to \left(\sqrt{\beta\left(1-\frac{e}{m}\right)}-1\right)^2,\end{aligned}$$

almost surely as $m \to \infty$. Plugging these in Lemma D.2, and using $c = 1 + \beta\left(1 - \frac{e}{m}\right)$, we get the desired result.

D.5 Proof of Theorem 5.4

Given an S, we will construct a data pair (X, y) so that the quantity

$$\frac{\|X\hat{\theta} - y\|^2}{\|X\theta^* - y\|^2}$$

is maximized, where we choose (X, y) so that $\|X\theta^* - y\|^2 > 0$ by design, so the above is well-defined.

To this end, let us first fix θ^*, and assume $y = X\theta^* + r$, where $r^\top X = 0$, by the optimality condition. We can equivalently construct the pair (X, r). Then the relative error can be written as

$$\frac{\|X\hat{\theta}-y\|^2}{\|X\theta^*-y\|^2}=\frac{\|X\left(\hat{\theta}-\theta^*\right)+r\|^2}{\|r\|^2}\overset{\text{(a)}}{=}1+\frac{\|X\left(\hat{\theta}-\theta^*\right)\|^2}{\|r\|^2}$$

$$\overset{\text{(b)}}{=}1+\frac{\|X\left(X^\top S_{A^c}^\top S_{A^c}X\right)^{-1}X^\top S_{A^c}^\top S_{A^c}y-X\theta^*\|^2}{\|r\|^2}=$$

$$1+\frac{\|X\left(X^\top S_{A^c}^\top S_{A^c}X\right)^{-1}X^\top S_{A^c}^\top S_{A^c}(X\theta^*+r)-X\theta^*\|^2}{\|r\|^2}$$

$$=1+\frac{\|X\left(X^\top S_{A^c}^\top S_{A^c}X\right)^{-1}X^\top S_{A^c}^\top S_{A^c}r\|^2}{\|r\|^2}$$

where (a) follows by the fact that $r^\top X=0$, and (b) follows by plugging in the analytic expression for $\hat{\theta}=(S_{A^c}X)^\dagger(S_{A^c}y)$. Let $S_{A^c}^\top S_{A^c}=Q^\top\Lambda Q$ be the eigendecomposition of $S_{A^c}^\top S_{A^c}$, and define $Z=QX$ and $t=Qr$, where we reduced the problem to constructing (Z,t). Then

$$\frac{\|X\hat{\theta}-y\|^2}{\|X\theta^*-y\|^2}\overset{\text{(a)}}{=}1+\frac{\|QX\left(X^\top S_{A^c}^\top S_{A^c}X\right)^{-1}X^\top S_{A^c}^\top S_{A^c}r\|^2}{\|Qr\|^2}$$

$$=1+\frac{\|Z\left(Z^\top\Lambda Z\right)^{-1}Z^\top\Lambda t\|^2}{\|t\|^2}$$

where (a) follows by the fact that ℓ_2 norm is invariant under orthogonal transformations. Note that since we require $r^\top X=0$, we have $t^\top Z=0$. Therefore we set $t=(I-ZZ^\dagger)v$, where there is no constraint on v. Plugging in this value for t and simplifying, and also using the non-expansiveness of the projection, which implies $\|v\|^2\geq\|t\|^2$, we have

$$\sup_{X,y}\frac{\|X\hat{\theta}-y\|^2}{\|X\theta^*-y\|^2}$$

$$\geq\sup_{Z,v}\left(1+\frac{\|Z\left(Z^\top\Lambda Z\right)^{-1}Z^\top\Lambda v\|^2-\|U^\top v\|^2}{\|v\|^2}\right)$$

$$=\sup_{Z,v}\frac{\|Z\left(Z^\top\Lambda Z\right)^{-1}Z^\top\Lambda v\|^2}{\|v\|^2},$$

where U is a $n\times d$ matrix with orthonormal columns, whose columns span the column space of Z. In the last equality, we have used the fact that U is orthogonal.

Now, note that we can assume, without loss of generality, $S_{A^c}^\top S_{A^c}$ is positive definite, since otherwise we can construct (X,y) with unbounded error, by choosing columns of X in the eigenspace of $S_{A^c}^\top S_{A^c}$ associated with zero eigenvalues. Therefore, we can assume Λ is invertible. Define $B = \Lambda^{1/2} Z$, and $P = B\left(B^\top B\right)^{-1} B^\top$ to be the projection matrix on the range space of B. We pick an X such that

$$P = \begin{bmatrix} \frac{1}{2} & 0^\top & \frac{1}{2} \\ 0 & \widetilde{P} & 0 \\ \frac{1}{2} & 0^\top & \frac{1}{2} \end{bmatrix}$$

where 0 is the 0-vector and $\widetilde{P}$ is some other idempotent matrix of the appropriate size. Then P is an appropriate projection matrix for the choice of B as

$$B = \begin{bmatrix} 0^\top & \frac{1}{2} \\ \widetilde{B} & 0 \\ 0^\top & \frac{1}{2} \end{bmatrix}, \quad \tilde{P} = \widetilde{B}\left(\widetilde{B}^\top \widetilde{B}^\top\right)^{-1} \widetilde{B}^\top.$$

We additionally pick $v = \alpha[1, 0, \ldots, 0]^\top$ for any scalar α. Then, denoting with λ_i the ith largest eigenvalue in Λ,

$$\begin{aligned} &\sup_{X,y} \frac{\|X\hat{\theta} - y\|^2}{\|X\theta^* - y\|^2} \geq \sup_{B,v} \frac{\|\Lambda^{-1/2} B\left(B^\top B\right)^{-1} B^\top \Lambda^{1/2} v\|^2}{\|v\|^2} \\ &= \left(\frac{\lambda_1^{1/2} + \lambda_n^{1/2}}{2\lambda_n^{1/2}}\right)^2 = \frac{1}{4}(1 + \kappa(S_{A^c}))^2 \end{aligned}$$

where $\kappa(S_{A^c}) = \frac{\sqrt{\lambda_1}}{\sqrt{\lambda_n}}$ is the condition number of S_{A^c}.

APPENDIX E

Proofs and Tables for Chapter 6

E.1 Proofs of Theorems 6.1 of 6.2

In the proofs, we will ignore the normalization constants on the objective functions for brevity. We will assume the normalization $\frac{1}{\sqrt{\eta}}$ is absorbed into the encoding matrix S_A. Let $\widetilde{f}_t^A := \|S_{A_t}(Xw_t - y)\|^2 + \lambda h(w)$, and $\widetilde{f}^A(w) := \|S_{A_t}(Xw - y)\|^2 + \lambda h(w)$, where we set $A \equiv A_t$. Let $\widetilde{w}_t^*$ denote the solution to the effective "instantaneous" problem at iteration t, *i.e.*, $\widetilde{w}_t^* = \arg\min_w \widetilde{f}^A(w)$.

Throughout this appendix, we will also denote

$$w^* = \arg\min_w \|Xw - y\|^2 + \lambda h(w)$$
$$\hat{w} = \arg\min_w \|S_A(Xw - y)\|^2 + \lambda h(w)$$

unless otherwise noted, where A is a fixed subset of $[m]$.

E.1.1 Lemmas

Lemma E.1. *If S satisfies* (6.6) *for any $A \subseteq [m]$ with $|A| \geq k$, for any convex set C,*

$$\|X\hat{w} - y\|^2 \leq \kappa^2 \|Xw^* - y\|^2,$$

where $\kappa = \frac{1+\epsilon}{1-\epsilon}$, $\hat{w} = \arg\min_{w\in C} \|S_A(Xw - y)\|^2$, and $w^* = \arg\min_{w\in C} \|Xw - y\|^2$.

Proof. Define $e = \hat{w} - w^*$ and note that

$$\|X\hat{w} - y\| = \|Xw^* - y + Xe\| \leq \|Xw^* - y\| + \|Xe\|$$

by triangle inequality, which implies

$$\|X\hat{w}-y\|^2 \leq \left(1+\frac{\|Xe\|}{\|Xw^*-y\|}\right)^2 \|Xw^*-y\|^2 = \left(1+\frac{\|Xe\|}{\|Xw^*-y\|}\right)^2 \|Xw^*-y\|^2. \quad \text{(E.1)}$$

Now, for any $c > 0$, consider

$$\begin{aligned}
\|Xe\|^2 &\leq \frac{\|S_A Xe\|^2}{1-\epsilon} \overset{\text{(a)}}{\leq} -2\frac{e^\top X^\top S_A^\top S_A (Xw^*-y)}{1-\epsilon} \\
&= -2\frac{e^\top X^\top \left(S_A^\top S_A - cI\right)(Xw^*-y)}{1-\epsilon} - \frac{2c}{1-\epsilon} e^\top X^\top (Xw^*-y) \\
&\overset{\text{(b)}}{\leq} -2\frac{e^\top X^\top \left(S_A^\top S_A - cI\right)(Xw^*-y)}{1-\epsilon} \\
&\overset{\text{(c)}}{\leq} 2\frac{\left\|e^\top X^\top \left(cI - S_A^\top S_A\right)\right\|}{1-\epsilon} \|Xw^*-y\| \\
&\overset{\text{(d)}}{\leq} 2\frac{\left\|cI - S_A^\top S_A\right\|}{1-\epsilon} \|Xw^*-y\| \|Xe\|,
\end{aligned}$$

where (a) follows by expanding and re-arranging $\left\|S_A\left(X\hat{w}-y\right)\right\|^2 \leq \left\|S_A\left(Xw^*-y\right)\right\|^2$, which is true since $\hat{w}$ is the minimizer of this function; (b) follows by the fact that since $\hat{w} \in C$, e represents a feasible direction of the constrained optimization, and thus the convex optimality condition implies $\langle \nabla f(w^*), \hat{w}-w^* \rangle = e^\top X^\top (Xw^*-y) \geq 0$; (c) follows by Cauchy-Schwarz inequality; and (d) follows by the definition of matrix norm.

Since this is true for any $c > 0$, we make the minimizing choice $c = \frac{\lambda_{\max}+\lambda_{\min}}{2}$ (where $\lambda_{\max}$ and $\lambda_{\min}$ represent the largest and smallest eigenvalues of $S_A^\top S_A$, respectively), which gives

$$\frac{\|Xe\|}{\|X\hat{w}-y\|} \leq \frac{\lambda_{\max}-\lambda_{\min}}{\lambda_{\min}} \leq \frac{2\epsilon}{1-\epsilon}.$$

Plugging this back in (E.1), we get the desired result. □

Lemma E.2. *If S satisfies* (6.6) *for any $A \subseteq [m]$ with $|A| \geq k$,*

$$f(\hat{w}) \leq \kappa^2 f(w^*),$$

where $\kappa = \frac{1+\epsilon}{1-\epsilon}$, $\hat{w} = \arg\min_w \|S_A(Xw-y)\|^2 + \lambda h(w)$, and $w^* = \arg\min_w \|Xw-y\|^2 + \lambda h(w)$.

Proof. Consider a fixed $A_t = A$, and a corresponding

$$\hat{w} = \widetilde{w}_t^* \in \arg\min_w \|S_A(Xw - y)\|^2 + \lambda h(w)$$

Define

$$\begin{aligned} \hat{w}(r) &= \arg\min_{w:\lambda h(w) \leq r} \|S_A(Xw - y)\|^2 \\ w^*(r) &= \arg\min_{w:\lambda h(w) \leq r} \|Xw - y\|^2. \end{aligned}$$

Finally, define

$$r^* = \arg\min_r \|Xw^*(r) - y\|^2 + r.$$

Now, consider

$$\begin{aligned} f(\hat{w}) = \|X\hat{w} - y\|^2 + \lambda h(w) &= \min_r \left(\|X\hat{w}(r) - y\|^2 + r\right) \\ &\leq \|X\hat{w}(r^*) - y\|^2 + r^* \overset{\text{(a)}}{\leq} \kappa^2 \|Xw^*(r^*) - y\|^2 + r^* \\ &\leq \kappa^2 \left(\|Xw^*(r^*) - y\|^2 + r^*\right) = \kappa^2 f(w^*), \end{aligned}$$

which shows the desired result, where (a) follows by Lemma E.1, and by the fact that the set $\{w : \lambda h(w) \leq r\}$ is a convex set. □

Lemma E.3. *If*

$$\widetilde{f}_{t+1}^A - \widetilde{f}^A(\widetilde{w}_t^*) \leq \gamma \left(\widetilde{f}_t^A - \widetilde{f}^A(\widetilde{w}_t^*)\right)$$

for all $t > 0$, and for some $0 < \gamma < 1$, where $\widetilde{w}_t^* \in \arg\min_w \widetilde{f}_t^A$, then

$$f(w_t) \leq (\kappa\gamma)^t f(w_0) + \frac{\kappa^2(\kappa - \gamma)}{1 - \kappa\gamma} f(w^*),$$

where $\kappa = \frac{1+\epsilon}{1-\epsilon}$.

Proof. Since for any w,

$$(1 - \epsilon)\|Xw - y\|^2 \leq (Xw - y)^\top S_A^\top \widetilde{S}_A (Xw - y),$$

we have

$$(1-\epsilon)\, f(w) \leq \widetilde{f}^A(w).$$

Similarly $\widetilde{f}^A(w) \leq (1+\epsilon)\, f(w)$, and therefore, using the assumption of the theorem

$$(1-\epsilon)\, f(w_{t+1}) - (1+\epsilon)\, f\left(\widetilde{w}_t^*\right) \leq \gamma\left((1+\epsilon)\, f(w_t) - (1-\epsilon)\, f\left(\widetilde{w}_t^*\right)\right),$$

which can be re-arranged into the linear recursive inequality

$$f(w_{t+1}) \leq \kappa\gamma f_t + (\kappa-\gamma) f\left(\widetilde{w}_t^*\right) \overset{\text{(a)}}{\leq} \kappa\gamma f(w_t) + \kappa^2(\kappa-\gamma) f\left(w^*\right),$$

where $\kappa = \frac{1+\epsilon}{1-\epsilon}$ and (a) follows by Lemma E.2. By considering such inequalities for $0 \leq \tau \leq t$, multiplying each by $(\kappa\gamma)^{t-\tau}$ and summing, we get

$$\begin{aligned} f(w_t) &\leq (\kappa\gamma)^t f(w_0) + \kappa^2(\kappa-\gamma) f\left(w^*\right) \sum_{\tau=0}^{t-1} (\kappa\gamma)^\tau \\ &\leq (\kappa\gamma)^t f(w_0) + \frac{\kappa^2\left(\kappa-\gamma\right)}{1-\kappa\gamma} f\left(w^*\right). \end{aligned}$$

□

Lemma E.4. *Under the assumptions of Theorem 6.2, $\widetilde{f}^A(w)$ is $(1-\epsilon)\,(\mu+\lambda)$-strongly convex.*

Proof. It is sufficient to show that the minimum eigenvalue of $\widetilde{X}_A^\top \widetilde{X}_A$ is bounded away from zero. This can easily be shown by the fact that

$$u^\top \widetilde{X}_A^\top \widetilde{X}_A u = u^\top X^\top S_A^\top S_A X u \geq (1-\epsilon)\, \|Xu\|^2 \geq (1-\epsilon)\, \mu \|u\|^2,$$

for any unit vector u. □

Lemma E.5. *Let $M \in \mathbb{R}^{p\times p}$ be a symmetric positive definite matrix, with the condition number (ratio of maximum eigenvalue to the minimum eigenvalue) given by κ. Then, for any unit vector u,*

$$\frac{u^\top M u}{\|Mu\|} \geq \frac{2\sqrt{\kappa}}{\kappa+1}.$$

Proof. We point out that this is a special case of Kantorovich inequality, but provide a dedicated proof here for completeness.

Let M have the eigen-decomposition $M = Q^\top D Q$, where Q has orthonormal columns, and D is a diagonal matrix with positive, decreasing entries $d_1 \geq d_2 \geq \cdots \geq d_n$, with $\frac{d_1}{d_n} = \kappa$. Let $y = (Qu)^{\circ 2}$, where ${}^{\circ 2}$ denotes entry-wise square. Then the quantity we are interested in can be represented as

$$\frac{\sum_{i=1}^n d_i y_i}{\sqrt{\sum_{i=1}^n d_i^2 y_i}},$$

which we would like to minimize subject to a simplex constraint $\mathbf{1}^\top y = 1$. Using Lagrange multipliers, it can be seen that the minimum is attained where $y_1 = \frac{1}{1+\kappa}$, $y_n = \frac{\kappa}{1+\kappa}$, and $y_i = 0$ for $i \neq 1, n$. Plugging this back the objective, we get the desired result

$$\frac{u^\top M u}{\|Mu\|} \geq \frac{2\sqrt{\kappa}}{\kappa + 1}.$$

□

Proof of Lemma 1. Define $\breve{S}_t := S_{A_t \cap A_{t-1}}$. First note that

$$\begin{aligned} r_t^\top u_t &= \left(X^\top \breve{S}_t^\top \breve{S}_t \left[(X w_t - y) - (X w_{t-1} - y) \right] \right)^\top (w_t - w_{t-1}) \\ &= (w_t - w_{t-1})^\top X^\top \breve{S}_t^\top \breve{S}_t X (w - w_{t-1}) \\ &\geq \delta \mu \|u_t\|^2, \end{aligned} \tag{E.2}$$

by (5) Also consider

$$\frac{\|r_t\|^2}{r_t^\top u_t} = \frac{(w_t - w_{t-1})^\top \left(X^\top \breve{S}_t^\top \breve{S}_t X \right)^2 (w_t - w_{t-1})}{(w_t - w_{t-1})^\top X^\top \breve{S}_t^\top \breve{S}_t X (w_t - w_{t-1})},$$

which implies

$$\epsilon \mu \leq \frac{\|r_t\|^2}{r_t^\top u_t} \leq (1 + \epsilon) M,$$

again by (4). Now, setting $j_\ell = t - \widetilde{\sigma} + \ell$, consider the trace

$$\mathbf{tr}\left(B_t^{(\ell+1)} \right) = \mathbf{tr}\left(B_t^{(\ell)} \right) - \mathbf{tr}\left(\frac{B_t^{(\ell)} u_{j_\ell} u_{j_\ell}^\top B_t^{(\ell)}}{u_{j_\ell}^\top B_t^{(\ell)} u_{j_\ell}} \right) + \mathbf{tr}\left(\frac{r_{j_\ell} r_{j_\ell}^\top}{r_{j_\ell}^\top u_{j_\ell}} \right)$$

$$
\begin{aligned}
&\leq \mathbf{tr}\left(B_t^{(\ell)}\right) + \mathbf{tr}\left(\frac{r_{j_\ell} r_{j_\ell}^\top}{r_{j_\ell}^\top u_{j_\ell}}\right) \\
&= \mathbf{tr}\left(B_t^{(\ell)}\right) + \frac{\|r_{j_\ell}\|^2}{r_{j_\ell}^\top u_{j_\ell}} \\
&\leq \mathbf{tr}\left(B_t^{(\ell)}\right) + (1+\epsilon)M,
\end{aligned}
$$

which implies $\mathbf{tr}\left(B_t\right) \leq (1+\epsilon)M\left(\widetilde{\sigma}+d\right)$. It can also be shown (similar to [BNT16]) that

$$
\begin{aligned}
\det\left(B_t^{(\ell+1)}\right) &= \det\left(B_t^{(\ell)}\right) \cdot \frac{r_{j_\ell}^\top u_{j_\ell}}{u_{j_\ell}^\top B_t^{(\ell)} u_{j_\ell}} \\
&= \det\left(B_t^{(\ell)}\right) \cdot \frac{r_{j_\ell}^\top u_{j_\ell}}{\|u_{j_\ell}\|^2} \cdot \frac{\|u_{j_\ell}\|^2}{u_{j_\ell}^\top B_t^{(\ell)} u_{j_\ell}} \\
&\geq \det\left(B_t^{(\ell)}\right) \frac{\delta\mu}{(1+\epsilon)M\left(\widetilde{\sigma}+d\right)},
\end{aligned}
$$

which implies $\det\left(B_t\right) \geq \det\left(B_t^{(0)}\right)\left(\frac{\delta\mu}{(1+\epsilon)M(\widetilde{\sigma}+d)}\right)^{\widetilde{\sigma}}$. Since $B_t \geq 0$, its trace is bounded above, and its determinant is bounded away from zero, there must exist $0 < c_1 \leq c_2$ such that

$$
c_1 I \preceq B_t \preceq c_2 I.
$$

□

E.1.2 Proof of Theorem 6.1

The proof of the first part of the theorem is a special case of the proof of Theorem 6.3 (with $\lambda = 0$, and the smooth regularizer incorporated into $p(w)$) and thus we omit this proof and refer the reader to Appendix E.2. We prove the second part here.

Note that because of the condition in (6.6), we have

$$
\begin{aligned}
(1-\epsilon) \preceq S_A^\top S_A \preceq (1+\epsilon)I, \\
(1-\epsilon) \preceq S_D^\top S_D \preceq (1+\epsilon)I.
\end{aligned}
$$

Using smoothness of the objective, and the choices $d_t = -\nabla \widetilde{f}^A(w_t)(w_t)$ and $\alpha_t = \alpha$, we have

$$
\widetilde{f}^A\left(w_{t+1}\right) - \widetilde{f}^A(w_t) \leq \alpha \nabla \widetilde{f}^A(w_t)(w_t)^\top d_t + \frac{1}{2}\alpha^2 d_t^\top X^\top S_A^\top S_A X d_t + \frac{L}{2}\alpha^2 \|d_t\|^2
$$

$$\leq -\alpha\left(1-\frac{(1+\epsilon)M+L}{2}\alpha\right)\left\|\nabla\widetilde{f}^A(w_t)\right\|^2 = -\frac{2\zeta\left(1-\zeta\right)}{(1+\epsilon)M+L}\|\nabla\widetilde{f}^A(w_t)\|^2$$
$$\overset{\text{(a)}}{\leq} -\frac{4\nu\zeta\left(1-\zeta\right)}{M\left(1+\epsilon\right)+L}\left(\widetilde{f}^A\left(w_t\right)-\widetilde{f}^A\left(\widetilde{w}_t^*\right)\right),$$

where (a) follows by strong convexity. Re-arranging this inequality, and using the definition of γ, we get

$$\widetilde{f}_{t+1}^A-\widetilde{f}^A\left(\widetilde{w}_t^*\right)\leq\gamma\left(\widetilde{f}_t^A-\widetilde{f}^A\left(\widetilde{w}_t^*\right)\right),$$

which, using Lemma E.3, implies the result.

E.1.3 Proof of Theorem 6.2

Since $h(w)$ is constrained to be quadratic, we can absorb this term into the error term to get

$$\min_w\left\|\begin{bmatrix} S & 0 \\ 0 & I\end{bmatrix}\left(\begin{bmatrix} X \\ \sqrt{\lambda}I\end{bmatrix}w-\begin{bmatrix} y \\ 0\end{bmatrix}\right)\right\|.$$

Note that as long as S satisfies (6.6), the effective encoding matrix $\operatorname{diag}\left([S,I]\right)$ also satisfies the same. Therefore, without loss of generality we can ignore $h(w)$, and assume

$$(\mu+\lambda)I\preceq X^\top X\preceq(M+\lambda)I.$$

We also define $\lambda_{\min}=1-\epsilon$ and $\lambda_{\max}=1+\epsilon$ for convenience. Using convexity and the closed-form expression for the step size, we have

$$\widetilde{f}^A\left(w_{t+1}\right)-\widetilde{f}^A(w_t)\leq\alpha_t\nabla\widetilde{f}^A(w_t)^\top d_t+\frac{1}{2}\alpha_t^2 d_t^\top X^\top S_A^\top S_A X d_t$$
$$=-\frac{\rho\left(\nabla\widetilde{f}^A(w_t)^\top d_t\right)^2}{d_t^\top X^\top S_D^\top S_D X d_t}+\frac{1}{2}\frac{\rho^2\left(\nabla\widetilde{f}^A(w_t)^\top d_t\right)^2}{d_t^\top X^\top S_D^\top S_D X d_t}\cdot\frac{d_t^\top X^\top S_A^\top S_A X d_t}{d_t^\top X^\top S_D^\top S_D X d_t}$$
$$=\left(\frac{d_t^\top X^\top\left(\rho^2 S_A^\top S_A-2\rho S_D^\top S_D\right)X d_t}{2\left(d_t^\top X^\top S_D^\top S_D X d_t\right)^2}\right)\left(d_t^\top\nabla\widetilde{f}^A(w_t)\right)^2$$
$$\overset{\text{(a)}}{=}-\rho\left(\frac{z^\top\left(S_D^\top S_D-\frac{\rho}{2}S_A^\top S_A\right)z}{\left(z^\top S_D^\top S_D z\right)^2}\right)\frac{\left(d_t^\top\nabla\widetilde{f}^A(w_t)\right)^2}{\|X d_t\|^2}$$

$$
\begin{aligned}
&\overset{(b)}{\leq} -\rho\left(\frac{\lambda_{\min}-\frac{\rho}{2}\lambda_{\max}}{\lambda_{\min}^2}\right)\frac{\left(d_t^\top\nabla\widetilde{f}^A(w_t)\right)^2}{\|Xd_t\|^2} \overset{(c)}{\leq} -\frac{\rho}{M+\lambda}\left(\frac{\lambda_{\min}-\frac{\rho}{2}\lambda_{\max}}{\lambda_{\min}^2}\right)\frac{\left(d_t^\top\nabla\widetilde{f}^A(w_t)\right)^2}{\|d_t\|^2}\\
&\overset{(d)}{=} -\frac{\rho}{M+\lambda}\left(\frac{\lambda_{\min}-\frac{\rho}{2}\lambda_{\max}}{\lambda_{\min}^2}\right)\frac{\left(\nabla\widetilde{f}^A(w_t)^\top B_t\nabla\widetilde{f}^A(w_t)\right)^2}{\|B_t\nabla\widetilde{f}^A(w_t)\|^2}\\
&\overset{(e)}{\leq} -\frac{4\rho}{M+\lambda}\left(\frac{\lambda_{\min}-\frac{\rho}{2}\lambda_{\max}}{\lambda_{\min}^2}\right)\frac{c_1c_2}{(c_1+c_2)^2}\|\nabla\widetilde{f}^A(w_t)\|^2\\
&\overset{(f)}{\leq} -\frac{8(\mu+\lambda)\rho}{M+\lambda}\left(\frac{\lambda_{\min}-\frac{\rho}{2}\lambda_{\max}}{\lambda_{\min}^2}\right)\frac{c_1c_2}{(c_1+c_2)^2}\left(\widetilde{f}\left(w_t\right)-\widetilde{f}\left(\widetilde{w}_t^*\right)\right)\\
&\overset{(g)}{=} -\frac{4(\mu+\lambda)c_1c_2}{(M+\lambda)(1+\epsilon)\left(c_1+c_2\right)^2}\left(\widetilde{f}\left(w_t\right)-\widetilde{f}\left(\widetilde{w}_t^*\right)\right) \overset{(h)}{=} -\left(1-\gamma\right)\left(\widetilde{f}^A\left(w_t\right)-\widetilde{f}^A\left(\widetilde{w}_t^*\right)\right).
\end{aligned}
$$

where (a) follows by defining $z = \frac{Xd_t}{\|Xd_t\|}$; (b) follows by (6.6); (c) follows by the assumption that $X^\top X \preceq (M+\lambda)I$; (d) follows by the definition of d_t; (e) follows by Lemmas E.5 and 6.1; (f) follows by strong convexity of $\widetilde{f}$ (by Lemma E.4), which implies $\|\nabla\widetilde{f}^A(w_t)\|^2 \geq 2(\mu+\lambda)\left(\widetilde{f}\left(\theta_t\right)-\widetilde{f}\left(\widetilde{w}_t^*\right)\right)$; (g) follows by choosing $\rho = \frac{\lambda_{\min}}{\lambda_{\max}}$; and (h) follows using the definition of γ.

Re-arranging the inequality, we obtain

$$\widetilde{f}_{t+1}^A - \widetilde{f}^A\left(\widetilde{w}_t^*\right) \leq \gamma\left(\widetilde{f}_t^A - \widetilde{f}^A\left(\widetilde{w}_t^*\right)\right),$$

and hence applying first Lemma E.3, we get the desired result.

E.2 Proof of Theorem 6.3

Throughout this appendix, we will define $p(w) = \frac{1}{2}\|Xw-y\|^2$ and $\widetilde{p}_t(w) = \frac{1}{2}\|S_{A_t}\left(Xw-y\right)\|^2$ for convenience, where the normalization by $\sqrt{\eta}$ is absorbed into S_A. We will omit the normalization by n for brevity. Let us also define

$$w^* = \arg\min_w p(w) + \lambda h(w)$$

to be the true solution of the optimization problem.

By M-smoothness of $p(w)$,

$$p(w_{t+1}) \leq p(w_t) + \langle\nabla p(w_t), w_{t+1}-w_t\rangle + \frac{M}{2}\|w_{t+1}-w_t\|^2$$

$$
\begin{aligned}
&\leq p(w^*) - \langle \nabla p(w_t), w^* - w_t \rangle + \langle \nabla p(w_t), w_{t+1} - w_t \rangle + \frac{M}{2} \|w_{t+1} - w_t\|^2 \\
&\leq p(w^*) - \langle \nabla p(w_t), w^* - w_t \rangle + \langle \nabla p(w_t), w_{t+1} - w_t \rangle + \frac{1}{2\alpha} \|w_{t+1} - w_t\|^2 \qquad \text{(E.3)}
\end{aligned}
$$

where the second line follows by convexity of p, and the third line follows since $\alpha < \frac{1}{M}$. Since $w_{t+1} = \arg\min_w \widetilde{F}_t(w)$, by optimality conditions

$$
0 \in \partial h(w_{t+1}) + \nabla \widetilde{p}_t(w_t) + \frac{1}{\alpha} \left(w_{t+1} - w_t \right). \qquad \text{(E.4)}
$$

Since h is convex, any subgradient $g \in \partial h$ at $w = w_{t+1}$ satisfies

$$
h(w^*) \geq h(w_{t+1}) + \langle g, w^* - w_{t+1} \rangle,
$$

and therefore (E.4) implies

$$
h(w^*) \geq h(w_{t+1}) - \langle \nabla \widetilde{p}_t(w_t), w^* - w_{t+1} \rangle - \frac{1}{\alpha} \langle w_{t+1} - w_t, w^* - w_{t+1} \rangle. \qquad \text{(E.5)}
$$

Combining (E.3) and (E.5),we have

$$
\begin{aligned}
f(w_{t+1}) &\leq f(w^*) + \langle \nabla p(w_t) - \nabla \widetilde{p}_t(w_t), w_{t+1} - w^* \rangle \\
&\quad - \frac{1}{\alpha} \langle w_t - w_{t+1}, w^* - w_{t+1} \rangle + \frac{1}{2\alpha} \|w_t - w_{t+1}\|^2 \\
&= f(w^*) + \langle \nabla p(w_t) - \nabla \widetilde{p}_t(w_t), w_{t+1} - w^* \rangle \\
&\quad + \frac{1}{2\alpha} \left(\|w_t\|^2 - 2 w_t^\top w^* + \|w^*\|^2 + 2 w_{t+1}^\top w^* - \|w^*\|^2 - \|w_{t+1}\|^2 \right) \\
&= f(w^*) + \langle \nabla p(w_t) - \nabla \widetilde{p}_t(w_t), w_{t+1} - w^* \rangle \\
&\quad + \frac{1}{2\alpha} \left(\|w_t - w^*\|^2 - \|w_{t+1} - w^*\|^2 \right) \qquad \text{(E.6)}
\end{aligned}
$$

Define $\Delta = I - S_A^\top S_A$, and consider the second term on the right-hand side of (E.6).

$$
\begin{aligned}
&\langle \nabla p(w_t) - \nabla \widetilde{p}_t(w_t), w_{t+1} - w^* \rangle = \left\langle X^\top \Delta (X w_t - y), w_{t+1} - w^* \right\rangle \\
&\quad = \langle \Delta (X w_t - y), X w_{t+1} - y \rangle - \langle \Delta (X w_t - y), X w^* - y \rangle \\
&\quad = \frac{1}{2} \Big[\left(X \left(w_t + w_{t+1} \right) - 2y \right)^\top \Delta \left(X \left(w_t + w_{t+1} \right) - 2y \right) \\
&\qquad - \left(X w_{t+1} - y \right)^\top \Delta \left(X w_{t+1} - y \right) + \left(X w^* - y \right)^\top \Delta \left(X w^* - y \right)
\end{aligned}
$$

$$
\begin{aligned}
&- \left(X\left(w_t + w^*\right) - 2y\right)^\top \Delta \left(X\left(w_t + w^*\right) - 2y\right)\Big] \\
&= 2\left(X\left(\frac{w_t + w_{t+1}}{2}\right) - y\right)^\top \Delta \left(X\left(\frac{w_t + w_{t+1}}{2}\right) - y\right) \\
&\quad - 2\left(X\left(\frac{w_t + w^*}{2}\right) - y\right)^\top \Delta \left(X\left(\frac{w_t + w^*}{2}\right) - y\right) \\
&\quad - \frac{1}{2}\left(Xw_{t+1} - y\right)^\top \Delta \left(Xw_{t+1} - y\right) + \frac{1}{2}\left(Xw^* - y\right)^\top \Delta \left(Xw^* - y\right) \\
&\leq 4\epsilon p\left(\frac{w_t + w_{t+1}}{2}\right) + 4\epsilon p\left(\frac{w_t + w^*}{2}\right) + \epsilon p(w_{t+1}) + \epsilon p(w^*) \\
&\overset{(a)}{\leq} \epsilon\left[4p(w_t) + 3p(w_{t+1}) + 3p(w^*)\right] \\
&\leq \epsilon\left[4f(w_t) + 3f(w_{t+1}) + 3f(w^*)\right],
\end{aligned}
$$

where (a) if by convexity of $p(w)$ and Jensen's inequality, and the last line follows by non-negativity of h. Plugging this back in (E.6),

$$
(1 - 3\epsilon) f(w_{t+1}) - 4\epsilon f(w_t) \leq (1 + 3\epsilon) f(w^*) + \frac{1}{2\alpha}\left(\|w_t - w^*\|^2 - \|w_{t+1} - w^*\|^2\right).
$$

Adding this for $t = 1, \ldots, (T-1)$,

$$
\begin{aligned}
(1 - 7\epsilon)\sum_{t=1}^{T} f(w_t) &\leq (T - 1)(1 + 3\epsilon) f(w^*) + 4\epsilon f(w_0) + \frac{1}{2\alpha}\left(\|w_0 - w^*\|^2 - \|w_T - w^*\|^2\right) \\
&\leq T(1 + 3\epsilon) f(w^*) + 4\epsilon f(w_0) + \frac{1}{2\alpha}\|w_0 - w^*\|^2.
\end{aligned}
$$

Defining $\bar{f}_t = \frac{1}{T}\sum_{t=1}^{T} f(w_t)$, and $\kappa = \frac{1+3\epsilon}{1-7\epsilon}$, we get

$$
\bar{f}_T - \kappa f(w^*) \leq \frac{4\epsilon f(w_0) + \frac{1}{2\alpha}\|w_0 - w^*\|^2}{(1 - 7\epsilon)T},
$$

which proves the first part of the theorem. To establish the second part of the theorem, note that the convexity of h implies

$$
h(w_t) \geq h(w_{t+1}) + \langle g, w_t - w_{t+1}\rangle,
$$

where $g \in \partial h(w_{t+1})$. By the optimality condition (E.4), this implies

$$
h(w_t) \geq h(w_{t+1}) - \langle \nabla\widetilde{p}_t(w_t), w_t - w_{t+1}\rangle + \frac{1}{\alpha}\|w_{t+1} - w_t\|^2.
$$

Combining this with the smoothness condition of $p(w)$,

$$p(w_{t+1}) \leq p(w_t) + \langle \nabla p(w_t), w_{t+1} - w_t \rangle + \frac{M}{2} \|w_{t+1} - w_t\|^2$$

and using the fact that $\alpha < \frac{1}{M}$, we have

$$f(w_{t+1}) \leq f(w_t) + \langle \nabla p(w_t) - \nabla \widetilde{p}_t(w_t), w_{t+1} - w_t \rangle - \frac{1}{2\alpha} \|w_t - w_{t+1}\|^2.$$

As in the previous analysis, we can show that

$$\langle \nabla p(w_t) - \nabla \widetilde{p}_t(w_t), w_{t+1} - w_t \rangle \leq \epsilon \left[7f(w_t) + 3f(w_{t+1}) \right],$$

and therefore

$$\begin{aligned} f(w_{t+1}) &\leq \frac{1+7\epsilon}{1-3\epsilon} f(w_t) - \frac{1}{2\alpha(1-3\epsilon)} \|w_t - w_{t+1}\|^2 \\ &\leq \frac{1+7\epsilon}{1-3\epsilon} f(w_t). \end{aligned}$$

E.3 Proof of Theorem 6.4

For an iterate v_t, let $w_t := Sv_t$. Define the solution set $\mathcal{S} = \arg\min_w g(w)$, and $w_t^* = \mathcal{P}_{\mathcal{S}}(w_t)$, where $\mathcal{P}_{\mathcal{S}}(\cdot)$ is the projection operator onto the set $\mathcal{S}$. Let v_t^* be such that $w_t^* = S^\top v_t^*$, which always exists since S has full column rank.

We also define $L' := L(1+\epsilon)$, and $g^* = \min_w g(w) = g(w_t^*)$ for any t.

E.3.1 Lemmas

Lemma E.6. *$\widetilde{g}(v)$ is L'-smooth.*

Proof. For any u, v,

$$\begin{aligned} \widetilde{g}(u) = g(S^\top u) &\leq g(S^\top v) + \langle \nabla g(S^\top v), S^\top (u - v) \rangle + \frac{L}{2} \|S^\top (u - v)\|^2, \\ &\overset{(a)}{\leq} g(S^\top v) + \langle S \nabla g(S^\top v), u - v \rangle + \frac{L(1+\epsilon)}{2} \|u - v\|^2, \end{aligned}$$

$$\overset{(b)}{=} \widetilde{g}(v) + \langle \nabla \widetilde{g}(v), u - v \rangle + \frac{L(1+\epsilon)}{2}\|u - v\|^2,$$

where (a) follows from smoothness of g, and from (m, η, ϵ)-BRIP property, and (b) is by the chain rule of derivatives and the definition of $\widetilde{g}(v)$. Therefore $\widetilde{g}$ is $L(1+\epsilon)$-smooth. □

Lemma E.7. *For any t,*

$$\widetilde{g}^* := \min_v \widetilde{g}(v) = \min_w g(w) =: g^*.$$

Proof. It is clear that

$$\min_v \tilde{g}(v) = \min_v g(S^\top v) \geq \min_w g(w).$$

To show the other direction, set $v^* = S(S^\top S)^{-1} w^*$, where $S^\top S$ is invertible since S has full column rank. Then $g(w^*) = \tilde{g}(v^*) \geq \min_v \tilde{g}(v)$. □

Lemma E.8. *If g is ν-restricted-strongly convex, then*

$$g(w) - g^* \geq \nu \|w - w^*\|^2,$$

where $w^* = \mathcal{P}_{\mathcal{S}}(w)$.

Proof. We follow the proof technique in [ZY13]. We have

$$\begin{aligned} g(w) &= g^* + \int_0^1 \langle \nabla g(w^* + \tau(w - w^*)), w - w^* \rangle d\tau \\ &= g^* + \int_0^1 \frac{1}{\tau} \langle \nabla g(w^* + \tau(w - w^*)), \tau(w - w^*) \rangle d\tau \\ &\geq g^* + \int_0^1 \frac{1}{\tau} \nu \tau^2 \|w - w^*\|^2 d\tau \\ &= g^* + \nu \|w - w^*\|^2, \end{aligned}$$

which is the desired result, where in the third line we used ν-restricted strong convexity, and the fact that

$$\mathcal{P}_{\mathcal{S}}(w^* + \tau(w - w^*)) = w^*,$$

for all $\tau \in [0, 1]$, since $w^* = \mathcal{P}_{\mathcal{S}}(w)$ is thr orthogonal projection. □

E.3.2 Proof of Theorem 6.4

Recall that the step for block i at time t, $\Delta_{i,t}$, is defined by

$$\Delta_{i,t} := \begin{cases} -\alpha \nabla_i \widetilde{g}(v_{t-1}), & \text{if } i \in A_t \\ 0, & \text{otherwise.} \end{cases}$$

By smoothness and definition of Δ_t,

$$\begin{aligned} \widetilde{g}(v_{t+1}) - \widetilde{g}(v_t) &\leq \langle \nabla \widetilde{g}(v_t), \Delta_t \rangle + \frac{L'}{2} \|\Delta_t\|^2 \\ &= \sum_{i \in A_t} \left(\langle \nabla_i \widetilde{g}(v_t), \Delta_{i,t} \rangle + \frac{L'}{2} \|\Delta_{i,t}\|^2 \right) \\ &= \sum_{i \in A_t} \left(-\frac{1}{\alpha} \langle \Delta_{i,t}, \Delta_{i,t} \rangle + \frac{L'}{2} \|\Delta_{i,t}\|^2 \right) \\ &= -\left(\frac{1}{\alpha} - \frac{L'}{2} \right) \|\Delta_t\|^2. \end{aligned} \tag{E.7}$$

Now, for any t,

$$\begin{aligned} \widetilde{g}(v_t) - \widetilde{g}^* &\leq \langle \nabla \widetilde{g}(v_t), v_t^* - v_t \rangle = \left\langle S \nabla g(S^\top v_t), v_t^* - v_t \right\rangle \\ &\overset{(a)}{\leq} \left\| \nabla g(S^\top v_t) \right\| \cdot \left\| S^\top (v_t^* - v_t) \right\| = \left\| \nabla g(S^\top v_t) \right\| \cdot \|w_t^* - w_t\|, \end{aligned} \tag{E.8}$$

where (a) is due to Cauchy-Schwartz inequality. Using

$$\Delta_t = -\alpha P_t \begin{bmatrix} S_{A_t} \nabla g(S^\top v_t) \\ 0 \end{bmatrix},$$

where P_t is a block permutation matrix mapping $\{1, \ldots, k\}$ to the node indices in A_t, we have

$$\|\Delta_t\|^2 = \alpha^2 \nabla g(S^\top v_t)^\top S_{A_t}^\top P_t^\top P_t S_{A_t} \nabla g(S^\top v_t) \geq (1 - \epsilon) \alpha^2 \left\| \nabla g(S^\top v_t) \right\|^2. \tag{E.9}$$

Because of (E.7), we have

$$\widetilde{g}(v_{t+1}) - \widetilde{g}(v_t) = g(w_{t+1}) - g(w_t) \leq 0,$$

and hence w_t is contained in the level set defined by the initial iterate for all t, *i.e.*,

$$w_t \in \{w : g(w) \leq g(w_0)\}.$$

By the diameter assumption on this set, we have $\|w_t - w_t^*\| \leq R$ for all t. Using this and (E.9) in (E.8), we get

$$\widetilde{g}(v_t) - \widetilde{g}^* \leq \frac{R}{\alpha}\sqrt{\frac{1}{1-\epsilon}}\|\Delta_t\|.$$

Combining this with (E.7),

$$\widetilde{g}(v_{t+1}) - \widetilde{g}(v_t) \leq -\frac{(1-\epsilon)\alpha}{R}\left(1 - \frac{\alpha L'}{2}\right)\left(\widetilde{g}(v_t) - \widetilde{g}^*\right)^2.$$

Defining $\pi_t := \widetilde{g}(v_t) - \widetilde{g}^*$, and $C := \frac{(1-\epsilon)\alpha}{R}\left(1 - \frac{\alpha L'}{2}\right)$, this implies

$$\pi_{t+1} \leq \pi_t - C\pi_t^2.$$

Dividing both sides by $\pi_t \pi_{t+1}$, and noting that $\pi_{t+1} \leq \pi_t$ due to (E.7),

$$\frac{1}{\pi_t} \leq \frac{1}{\pi_{t+1}} - C\frac{\pi_t}{\pi_{t+1}} \leq \frac{1}{\pi_{t+1}} - C$$

Therefore

$$\frac{1}{\pi_t} \geq \frac{1}{\pi_0} + Ct,$$

which implies

$$\pi_t \leq \frac{1}{\frac{1}{\pi_0} + Ct}.$$

Since $g(w_t) = g(S^\top v_t) = \widetilde{g}(v_t)$ by definition, and $g^* = \widetilde{g}^*$ by Lemma E.7, $\pi_t = g(w_t) - g^*$, and therefore we have established the first part of the theorem.

To prove the second part, we make the additional assumption that g satisfies ν-restricted-strong convexity, which, through Lemma E.8, implies $g(w) - g^* \geq \nu\|w - w^*\|^2$, for $w^* = \mathcal{P}_{\mathcal{S}}(w)$. Plugging in $w = w_t$ then gives the bound

$$\|w_t - w_t^*\|^2 \leq \frac{1}{\nu}\pi_t.$$

Using this bound as well as (E.9) in (E.8), we have

$$\pi_t^2 \leq \frac{\|\Delta_t\|^2}{\nu(1-\epsilon)\alpha^2}\pi_t.$$

Using (E.7), this gives

$$\pi_t \leq \frac{1}{\nu(1-\epsilon)\alpha^2}\left(\frac{1}{\alpha}-\frac{L'}{2}\right)^{-1}\left(\pi_t-\pi_{t+1}\right),$$

which, defining $\xi = \frac{1}{\nu(1-\epsilon)\alpha}\left(1-\frac{L'\alpha}{2}\right)^{-1}$, results in

$$\pi_t \leq \left(1-\frac{1}{\xi}\right)^t \pi_0,$$

which shows the desired result.

E.4 Full Results of the Matrix Factorization Experiment

Tables E.1 and E.2 give the test and train RMSE for the Movielens 1-M recommendation task, with a random 80/20 train/test split.

	uncoded	replication	gaussian	paley	hadamard
	$m = 8$, $k = 1$				
train RMSE	0.804	0.783	0.781	**0.775**	0.779
test RMSE	0.898	0.889	0.877	**0.873**	0.874
runtime	1.60	1.76	2.24	1.82	1.82
	$m = 8$, $k = 4$				
train RMSE	0.770	0.766	0.765	**0.763**	0.765
test RMSE	0.872	0.872	**0.866**	0.868	0.870
runtime	2.96	3.13	3.64	3.34	3.18
	$m = 8$, $k = 6$				
train RMSE	0.762	0.760	0.762	**0.758**	0.760
test RMSE	0.866	0.871	0.864	**0.860**	0.864
runtime	5.11	4.59	5.70	5.50	5.33

Table E.1: Full results for Movielens 1-M, distributed over $m = 8$ nodes total. Runtime is in hours. An uncoded scheme running full batch L-BFGS has a train/test RMSE of 0.756 / 0.861, and a runtime of 9.58 hours.

	uncoded	replication	gaussian	paley	hadamard
	$m = 24$, $k = 3$				
train RMSE	0.805	0.791	0.783	**0.780**	0.782
test RMSE	0.902	0.893	0.880	**0.879**	0.882
runtime	2.60	3.22	3.98	3.49	3.49
	$m = 24$, $k = 12$				
train RMSE	0.770	**0.764**	0.767	**0.764**	0.765
test RMSE	0.872	0.870	**0.866**	0.868	0.868
runtime	4.24	4.38	4.92	4.50	4.61

Table E.2: Full results for Movielens 1-M, distributed over $m = 24$ nodes total. Runtime is in hours. An uncoded scheme running full batch L-BFGS has a train/test RMSE of 0.757 / 0.862, and a runtime of 14.11 hours.

References

[AD11] Alekh Agarwal and John C Duchi. "Distributed delayed stochastic optimization." In *Advances in Neural Information Processing Systems*, pp. 873–881, 2011.

[ADT11] Amir Salman Avestimehr, Suhas N. Diggavi, and David N. C. Tse. "Wireless Network Information Flow: A Deterministic Approach." *IEEE Transactions on Information Theory*, **57**(4):1872–1905, April 2011.

[AGS13] Ganesh Ananthanarayanan, Ali Ghodsi, Scott Shenker, and Ion Stoica. "Effective Straggler Mitigation: Attack of the Clones." In *NSDI*, volume 13, pp. 185–198, 2013.

[AM13] Arash Asadi and Vincenzo Mancuso. "On the compound impact of opportunistic scheduling and D2D communications in cellular networks." In *Proceedings of the 16th ACM international conference on Modeling, analysis & simulation of wireless and mobile systems*, pp. 279–288. ACM, 2013.

[Apo13] Tom M Apostol. *Introduction to analytic number theory.* Springer Science & Business Media, 2013.

[Ari84] Erdal Arikan. "Some complexity results about packet radio networks (Corresp.)." *IEEE Transactions on Information Theory*, **30**(4):681–685, 1984.

[Ari09] Erdal Arikan. "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels." *IEEE Transactions on Information Theory*, **55**(7):3051–3073, 2009.

[AWM14] Arash Asadi, Qing Wang, and Vincenzo Mancuso. "A survey on device-to-device communication in cellular networks." *IEEE Communications Surveys & Tutorials*, **16**(4):1801–1819, 2014.

[BCJ09] Tae Won Ban, Wan Choi, Bang Chul Jung, and Dan Keun Sung. "Multi-user diversity in a spectrum sharing system." *IEEE Transactions on Wireless Communications*, **8**(1):102–106, 2009.

[Ber61] Claude Berge. "Färbung von Graphen, deren sämtliche bzw. deren ungerade Kreise starr sind." *Wiss. Z. Martin-Luther-Univ. Halle-Wittenberg Math.-Natur. Reihe*, **10**(114):88, 1961.

[BGT93] Claude Berrou, Alain Glavieux, and Punya Thitimajshima. "Near Shannon limit error-correcting coding and decoding: Turbo-codes. 1." In *IEEE International Conference on Communications (ICC)*, volume 2, pp. 1064–1070. IEEE, 1993.

[BGT95] Partha P Bhattacharya, Leonidas Georgiadis, and Pantelis Tsoucas. "Problems of adaptive optimization in multiclass M/GI/1 queues with bernoulli feedback." *Mathematics of Operations Research*, **20**(2):355–380, 1995.

[BK08] Alireza Bayesteh and Amir K. Khandani. "On the user selection for MIMO broadcast channels." *IEEE Transactions on Information Theory*, **54**(3):1086–1107, 2008.

[BNT16] Albert S Berahas, Jorge Nocedal, and Martin Takác. "A Multi-Batch L-BFGS Method for Machine Learning." In *Advances in Neural Information Processing Systems*, pp. 1055–1063, 2016.

[Bro12] "Worlds First 5G WiFi 802.11ac SoC." Technical report, Broadcom Corporation, 2012.

[BS13] Jingwen Bai and Ashutosh Sabharwal. "Distributed full-duplex via wireless side-channels: Bounds and protocols." *IEEE Transactions on Wireless Communications*, **12**(8):4162–4173, 2013.

[BT09] Amir Beck and Marc Teboulle. "A fast iterative shrinkage-thresholding algorithm for linear inverse problems." *SIAM Journal on Imaging Sciences*, **2**(1):183–202, 2009.

[Chv75] Vašek Chvátal. "On certain polytopes associated with graphs." *Journal of Combinatorial Theory, Series B*, **18**(2):138–154, 1975.

[CJS10] Jung Il Choi, Mayank Jain, Kannan Srinivasan, Phil Levis, and Sachin Katti. "Achieving Single Channel, Full Duplex Wireless Communication." In *Proceedings International Conference on Mobile Computing and Networking (Mobicom)*, MobiCom '10, pp. 1–12, New York, NY, USA, 2010. ACM.

[CK03] Peter G Casazza and Jelena Kovačević. "Equal-norm tight frames with erasures." *Advances in Computational Mathematics*, **18**(2-4):387–430, 2003.

[Ct79] Thomas M. Cover and Abbas A. El Gamal. "Capacity Theorems for the Relay Channel." *IEEE Transactions on Information Theory*, **25**(5):572–584, September 1979.

[CT05] Emmanuel J Candes and Terence Tao. "Decoding by linear programming." *IEEE Transactions on Information Theory*, **51**(12):4203–4215, 2005.

[CT06] Emmanuel J Candes and Terence Tao. "Near-optimal signal recovery from random projections: Universal encoding strategies?" *IEEE Transactions on Information Theory*, **52**(12):5406–5425, 2006.

[CZ14] Shanzhi Chen and Jian Zhao. "The requirements, challenges, and technologies for 5G of terrestrial mobile telecommunication." *IEEE Communications Magazine*, **52**(5):36–43, 2014.

[Dau92] Ingrid Daubechies. *Ten lectures on wavelets*. SIAM, 1992.

[DB13] Jeffrey Dean and Luiz André Barroso. "The tail at scale." *Communications of the ACM*, **56**(2):74–80, 2013.

[DCG16] Sanghamitra Dutta, Viveck Cadambe, and Pulkit Grover. "Short-Dot: Computing Large Linear Transforms Distributedly Using Coded Short Dot Products." In *Advances In Neural Information Processing Systems*, pp. 2092–2100, 2016.

[DCM12] Jeffrey Dean, Greg Corrado, Rajat Monga, Kai Chen, Matthieu Devin, Mark Mao, Andrew Senior, Paul Tucker, Ke Yang, Quoc V Le, et al. "Large scale distributed deep networks." In *Advances in Neural Information pProcessing Systems*, pp. 1223–1231, 2012.

[DG08] Jeffrey Dean and Sanjay Ghemawat. "MapReduce: simplified data processing on large clusters." *Communications of the ACM*, **51**(1):107–113, 2008.

[DMM11] Petros Drineas, Michael W Mahoney, S Muthukrishnan, and Tamás Sarlós. "Faster least squares approximation." *Numerische mathematik*, **117**(2):219–249, 2011.

[DRW09] Klaus Doppler, Mika Rinne, Carl Wijting, Cássio B Ribeiro, and Klaus Hugl. "Device-to-device communication as an underlay to LTE-advanced networks." *IEEE Communications Magazine*, **47**(12), 2009.

[DS05] Goran Dimić and Nicholas D Sidiropoulos. "On downlink beamforming with greedy user selection: performance analysis and a simple new algorithm." *IEEE Transactions on Signal Processing*, **53**(10):3857–3868, 2005.

[DS10] Melissa Duarte and Ashutosh Sabharwal. "Full-duplex wireless communications using off-the-shelf radios: Feasibility and first results." In *Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, pp. 1558–1562. IEEE, 2010.

[DSB13] Melissa Duarte, Ayan Sengupta, Siddhartha Brahma, Christina Fragouli, and Suhas Diggavi. "Quantize-map-forward (QMF) relaying: an experimental study." In *Proceedings of the fourteenth ACM international symposium on Mobile ad hoc networking and computing*, pp. 227–236. ACM, 2013.

[Dur10] Rick Durrett. *Probability: theory and examples*, volume 3. Cambridge university press, 2010.

[ETW08] Raul Etkin, David N. C. Tse, and Hua Wang. "Gaussian Interference Channel Capacity to within One Bit." *IEEE Transactions on Information Theory*, **54**(12):5534–5562, December 2008.

[FM12] Matthew Fickus and Dustin G Mixon. "Numerically erasure-robust frames." *Linear Algebra and its Applications*, **437**(6):1394–1407, 2012.

[Gav74] Fnic Gavril. "The intersection graphs of subtrees in trees are exactly the chordal graphs." *Journal of Combinatorial Theory, Series B*, **16**(1):47–56, 1974.

[Gem80] Stuart Geman. "A limit theorem for the norm of random matrices." *The Annals of Probability*, pp. 252–261, 1980.

[GK11] Abbas El Gamal and Young-Han Kim. *Network Information Theory*. Cambridge University Press, 2011.

[GKK01] Vivek K Goyal, Jelena Kovačević, and Jonathan A Kelner. "Quantized frame expansions with erasures." *Applied and Computational Harmonic Analysis*, **10**(3):203–233, 2001.

[GNT06] Leonidas Georgiadis, Michael J. Neely, and Leandros Tassiulas. *Resource allocation and cross-layer control in wireless networks*. Now Publishers Inc., 2006.

[GS67] J.M. Goethals and J Jacob Seidel. "Orthogonal matrices with zero diagonal." *Canad. J. Math*, 1967.

[GZD15] Kristen Gardner, Samuel Zbarsky, Sherwin Doroudi, Mor Harchol-Balter, and Esa Hyytia. "Reducing latency via redundant requests: Exact analysis." *ACM SIGMETRICS Performance Evaluation Review*, **43**(1):347–360, 2015.

[HAS17] Wael Halbawi, Navid Azizan-Ruhi, Fariborz Salehi, and Babak Hassibi. "Improving distributed gradient descent using reed-solomon codes." *arXiv preprint arXiv:1706.05436*, 2017.

[HK81] Te Sun Han and Kingo Kobayashi. "A New Achievable Rate Region for the Interference Channel." *IEEE Transactions on Information Theory*, **27**(1):49–60, January 1981.

[HP04] Roderick B Holmes and Vern I Paulsen. "Optimal frames for erasures." *Linear Algebra and its Applications*, **377**:31–51, 2004.

[JW10] Libin Jiang and Jean Walrand. "A distributed CSMA algorithm for throughput and utility maximization in wireless networks." *IEEE/ACM Transactions on Networking (ToN)*, **18**(3):960–972, 2010.

[KD15] Can Karakus and Suhas Diggavi. "Opportunistic scheduling for full-duplex uplink-downlink networks." In *IEEE International Symposium on Information Theory (ISIT)*, pp. 1019–1023. IEEE, 2015.

[KD17] Can Karakus and Suhas Diggavi. "Enhancing multiuser MIMO through opportunistic D2D cooperation." *IEEE Transactions on Wireless Communications*, **16**(9):5616–5629, 2017.

[KH11] Gerhard Kramer and Jie Hou. "On message lengths for noisy network coding." In *IEEE Information Theory Workshop (ITW)*, pp. 430–431. IEEE, 2011.

[KL13] Ashish Khisti and Amos Lapidoth. "Multiple access channels with intermittent feedback and side information." In *IEEE International Symposium on Information Theory Proceedings (ISIT)*, pp. 2631–2635. IEEE, 2013.

[KSD17a] Can Karakus, Yifan Sun, and Suhas Diggavi. "Encoded distributed optimization." In *2017 IEEE International Symposium on Information Theory (ISIT)*, pp. 2890–2894. IEEE, 2017.

[KSD17b] Can Karakus, Yifan Sun, Suhas Diggavi, and Wotao Yin. "Straggler mitigation in distributed optimization through data encoding." In *Advances in Neural Information Processing Systems*, pp. 5440–5448, 2017.

[KSD18] Can Karakus, Yifan Sun, Suhas Diggavi, and Wotao Yin. "Redundancy techniques for straggler mitigation in distributed optimization and learning." *Submitted for publication. Preprint available, htttp://arxiv.org.*, 2018.

[KWD13a] Can Karakus, I-Hsiang Wang, and Suhas Diggavi. "An achievable rate region for Gaussian interference channel with intermittent feedback." In *Communication, Control, and Computing (Allerton), 2013 51st Annual Allerton Conference on*, pp. 203–210, Oct 2013.

[KWD13b] Can Karakus, I-Hsiang Wang, and Suhas Diggavi. "Interference channel with intermittent feedback." In *IEEE International Symposium on Information Theory Proceedings (ISIT)*, pp. 26–30. IEEE, 2013.

[KWD15] Can Karakus, I-Hsiang Wang, and Suhas Diggavi. "Gaussian interference channel with intermittent feedback." *IEEE Transactions on Information Theory*, **61**(9):4663–4699, 2015.

[LAP14] Mu Li, David G Andersen, Jun Woo Park, Alexander J Smola, Amr Ahmed, Vanja Josifovski, James Long, Eugene J Shekita, and Bor-Yiing Su. "Scaling Distributed Machine Learning with the Parameter Server." In *OSDI*, volume 14, pp. 583–598, 2014.

[LCB98] Yann LeCun, Corinna Cortes, and Christopher JC Burges. "The MNIST database of handwritten digits.", 1998.

[LCS01] Xiaojun Liu, Edwin K. P. Chong, and Ness B. Shroff. "Opportunistic transmission scheduling with resource-sharing constraints in wireless networks." *IEEE Journal on Selected Areas in Communications*, **19**(10):2053–2064, 2001.

[LKE11] Sung Lim, Young-Han Kim, Abbas El Gamal, and Sae-Young Chung. "Noisy network coding." *IEEE Transactions on Information Theory*, **57**(5):3132–3152, 2011.

[LKM15] Jiajia Liu, Nei Kato, Jianfeng Ma, and Naoto Kadowaki. "Device-to-device communication in LTE-advanced networks: A survey." *IEEE Communications Surveys & Tutorials*, **17**(4):1923–1940, 2015.

[LL06] Guoqing Li and Hui Liu. "Downlink radio resource allocation for multi-cell OFDMA system." *IEEE Transactions on Wireless Communications*, **5**(12):3451–3459, 2006.

[LLG12] James CF Li, Ming Lei, and Feifei Gao. "Device-to-device (D2D) communication in MU-MIMO cellular networks." In *IEEE Global Communications Conference (GLOBECOM)*, pp. 3583–3587. IEEE, 2012.

[LLP16] Kangwook Lee, Maximilian Lam, Ramtin Pedarsani, Dimitris Papailiopoulos, and Kannan Ramchandran. "Speeding up distributed machine learning using codes." In *IEEE International Symposium on Information Theory (ISIT)*, pp. 1143–1147. IEEE, 2016.

[LMA16] Songze Li, Mohammad Ali Maddah-Ali, and A Salman Avestimehr. "Fundamental tradeoff between computation and communication in distributed computing." In *IEEE International Symposium on Information Theory (ISIT)*, pp. 1814–1818. IEEE, 2016.

[LSS06] Xiaojun Lin, Ness B Shroff, and Rayadurgam Srikant. "A tutorial on cross-layer optimization in wireless networks." *IEEE Journal on Selected Areas in Communications*, **24**(8):1452–1463, 2006.

[LTL06] Pei Liu, Zhifeng Tao, Zinan Lin, Elza Erkip, and Shivendra Panwar. "Cooperative wireless communications: a cross-layer approach." *IEEE Wireless Communications*, **13**(4):84–92, 2006.

[LTM12] Sy-Quoc Le, Ravi Tandon, Mehul Motani, and H. Vincent Poor. "The Capacity Region of the Symmetric Linear Deterministic Interference Channel with Partial Feedback." *Proceedings of Allerton Conference on Communication, Control, and Computing*, October 2012.

[LV05] Yingbin Liang and Venugopal V Veeravalli. "Gaussian orthogonal relay channels: Optimal resource allocation and capacity." *IEEE Transactions on Information Theory*, **51**(9):3284–3289, 2005.

[LWR15] Ji Liu, Stephen J Wright, Christopher Ré, Victor Bittorf, and Srikrishna Sridhar. "An asynchronous parallel stochastic coordinate descent algorithm." *The Journal of Machine Learning Research*, **16**(1):285–322, 2015.

[LY13] Ming-Jun Lai and Wotao Yin. "Augmented ℓ_1 and nuclear-norm models with a globally linearly convergent algorithm." *SIAM Journal on Imaging Sciences*, **6**(2):1059–1091, 2013.

[LYR04] David D Lewis, Yiming Yang, Tony G Rose, and Fan Li. "Rcv1: A new benchmark collection for text categorization research." *Journal of machine learning research*, **5**(Apr):361–397, 2004.

[Mah11] Michael W Mahoney et al. "Randomized algorithms for matrices and data." *Foundations and Trends® in Machine Learning*, **3**(2):123–224, 2011.

[MR15] Aryan Mokhtari and Alejandro Ribeiro. "Global convergence of online limited memory BFGS." *Journal of Machine Learning Research*, **16**:3151–3181, 2015.

[MSZ06] Eytan Modiano, Devavrat Shah, and Gil Zussman. "Maximizing throughput in wireless networks via gossiping." In *ACM SIGMETRICS Performance Evaluation Review*, volume 34, pp. 27–38. ACM, 2006.

[NHH04] Aria Nosratinia, Todd E Hunter, and Ahmadreza Hedayat. "Cooperative communication in wireless networks." *IEEE communications Magazine*, **42**(10):74–80, 2004.

[OD10] Ayfer Ozgur and Suhas Diggavi. "Approximately achieving Gaussian relay network capacity with lattice codes." *arXiv preprint arXiv:1005.1284*, 2010.

[OD13] A Ozgur and S.N. Diggavi. "Approximately Achieving Gaussian Relay Network Capacity With Lattice-Based QMF Codes." *IEEE Transactions on Information Theory*, **59**(12):8275–8294, Dec 2013.

[OHT16] David Ott, Nageen Himayat, and Shilpa Talwar. *5G: Transforming theUser Wireless Experience*, pp. 34–51. John Wiley & Sons, Ltd, 2016.

[Oza84] Lawrence H Ozarow. "The capacity of the white Gaussian multiple access channel with feedback." *IEEE Transactions on Information Theory*, **30**(4):623–629, 1984.

[Pal33] Raymond EAC Paley. "On orthogonal matrices." *Studies in Applied Mathematics*, **12**(1-4):311–320, 1933.

[PW15] Mert Pilanci and Martin J Wainwright. "Randomized sketches of convex programs with sharp guarantees." *IEEE Transactions on Information Theory*, **61**(9):5096–5115, 2015.

[PXY16] Zhimin Peng, Yangyang Xu, Ming Yan, and Wotao Yin. "ARock: an algorithmic framework for asynchronous parallel coordinate updates." *SIAM Journal on Scientific Computing*, **38**(5):A2851–A2879, 2016.

[Reb08] Steffen Rebennack. "Stable set problem: Branch & cut algorithms stable set problem: Branch & cut algorithms." In *Encyclopedia of Optimization*, pp. 3676–3688. Springer, 2008.

[RK98] J Riedl and J Konstan. "Movielens dataset.", 1998.

[RPP17] Amirhossein Reisizadeh, Saurav Prakash, Ramtin Pedarsani, and Salman Avestimehr. "Coded computation over heterogeneous clusters." In *Information Theory (ISIT), 2017 IEEE International Symposium on*, pp. 2408–2412. IEEE, 2017.

[RRW11] Benjamin Recht, Christopher Re, Stephen Wright, and Feng Niu. "Hogwild: A lock-free approach to parallelizing stochastic gradient descent." In *Advances in Neural Information Processing Systems*, pp. 693–701, 2011.

[RS07] Bill Rosgen and Lorna Stewart. "Complexity results on graphs with few cliques." *Discrete Mathematics and Theoretical Computer Science*, **9**(1), 2007.

[RSU01] Thomas J Richardson, Mohammad Amin Shokrollahi, and Rüdiger L Urbanke.

"Design of capacity-approaching irregular low-density parity-check codes." *IEEE transactions on information theory*, **47**(2):619–637, 2001.

[Sat78] Hiroshi Sato. "An outer bound to the capacity region of broadcast channels." *IEEE Transactions on Information Theory*, **24**:374–377, 1978.

[SAY09] Achaleshwar Sahai, Vaneet Aggarwal, Melda Yuksel, and Ashutosh Sabharwal. "On channel output feedback in deterministic interference channels." In *IEEE Information Theory Workshop (ITW)*, pp. 298–302. IEEE, 2009.

[SCN10] Hooman Shirani-Mehr, Giuseppe Caire, and Michael J Neely. "MIMO downlink scheduling with non-perfect channel state knowledge." *IEEE Transactions on Communications*, **58**(7):2055–2066, 2010.

[SDS05] J. Salo, G. Del Galdo, J. Salmi, P. Kysti, M. Milojevic, D. Laselva, and C. Schneider. "MATLAB implementation of the 3GPP Spatial Channel Model (3GPP TR 25.996).", Jan. 2005.

[SDS13] Achaleshwar Sahai, Suhas Diggavi, and Ashutosh Sabharwal. "On degrees-of-freedom of full-duplex uplink/downlink channel." In *IEEE Information Theory Workshop (ITW)*, pp. 1–5. IEEE, 2013.

[SEA03] Andrew Sendonaris, Elza Erkip, and Behnaam Aazhang. "User cooperation diversity. Part I. System description." *IEEE transactions on communications*, **51**(11):1927–1938, 2003.

[SH03] Thomas Strohmer and Robert W Heath. "Grassmannian frames with applications to coding and communication." *Applied and Computational Harmonic Analysis*, **14**(3):257–275, 2003.

[SH05] Masoud Sharif and Babak Hassibi. "On the capacity of MIMO broadcast channels with partial side information." *IEEE Transactions on Information Theory*, **51**(2):506–522, 2005.

[SH07] Masoud Sharif and Babak Hassibi. "A comparison of time-sharing, DPC, and beamforming for MIMO broadcast channels with many users." *IEEE Transactions on Communications*, **55**(1):11–15, 2007.

[Sha56] Claude E Shannon. "The zero error capacity of a noisy channel." *IRE Transactions on Information Theory*, **2**(3):8–19, 1956.

[SHY17] Tao Sun, Robert Hannah, and Wotao Yin. "Asynchronous Coordinate Descent under More Realistic Assumptions." In *Advances in Neural Information Processing Systems*, pp. 6183–6191, 2017.

[Sil85] Jack W Silverstein. "The smallest eigenvalue of a large dimensional Wishart matrix." *The Annals of Probability*, pp. 1364–1368, 1985.

[SLR16] Nihar B Shah, Kangwook Lee, and Kannan Ramchandran. "When do redundant requests reduce latency?" *IEEE Transactions on Communications*, **64**(2):715–722, 2016.

[ST11] Changho Suh and David N. C. Tse. "Feedback Capacity of the Gaussian Interference Channel to Within 2 Bits." *IEEE Transactions on Information Theory*, **57**(5):2667–2685, May 2011.

[SWT12] Changho Suh, I-Hsiang Wang, and David N. C. Tse. "Two-way Interference Channels." *IEEE International Symposium on Information Theory (ISIT)*, pp. 2811–2815, July 2012.

[Szo13] Ferenc Szöllősi. "Complex Hadamard matrices and equiangular tight frames." *Linear Algebra and its Applications*, **438**(4):1962–1967, 2013.

[TE92] Leandros Tassiulas and Anthony Ephremides. "Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks." *IEEE Transactions on Automatic Control*, **37**(12):1936–1948, 1992.

[Tel99] Emre Telatar. "Capacity of Multi-antenna Gaussian Channels." *Transactions on Emerging Telecommunications Technologies*, **10**(6):585–595, 1999.

[TG05] Vagelis Tsibonis and Leonidas Georgiadis. "Optimal downlink scheduling policies for slotted wireless time-varying channels." *IEEE Transactions on Wireless Communications*, **4**(4):1808–1817, 2005.

[TLD16] Rashish Tandon, Qi Lei, Alexandros G Dimakis, and Nikos Karampatziakis. "Gradient Coding." *NIPS ML Systems Workshop (MLSyS)*, 2016.

[TLD17] Rashish Tandon, Qi Lei, Alexandros G Dimakis, and Nikos Karampatziakis. "Gradient coding: Avoiding stragglers in distributed learning." In *International Conference on Machine Learning*, pp. 3368–3376, 2017.

[TR314] "Study on LTE device to device proximity services; Radio aspects." Technical Report TR 36.843, 3GPP, Mar 2014.

[TV05] David Tse and Pramod Viswanath. *Fundamentals of wireless communication.* Cambridge university press, 2005.

[VSA12] Alireza Vahid, Changho Suh, and Amir Salman Avestimehr. "Interference Channels With Rate-Limited Feedback." *IEEE Transactions on Information Theory*, **58**(5):2788–2812, May 2012.

[VT03] Pramod Viswanath and David N. C. Tse. "Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality." *IEEE Transactions on Information Theory*, **49**(8):1912–1921, 2003.

[VTL02] Pramod Viswanath, David N. C. Tse, and Rajiv Laroia. "Opportunistic beamforming using dumb antennas." *IEEE Transactions on Information Theory*, **48**(6):1277–1294, 2002.

[Wel74] Lloyd Welch. "Lower bounds on the maximum cross correlation of signals (Corresp.)." *IEEE Transactions on Information theory*, **20**(3):397–399, 1974.

[WJW15] Da Wang, Gauri Joshi, and Gregory Wornell. "Using straggler replication to reduce latency in large-scale parallel computing." *ACM SIGMETRICS Performance Evaluation Review*, **43**(3):7–11, 2015.

[WLZ08] Jianqi Wang, David J. Love, and Michael D. Zoltowski. "User selection with zero-forcing beamforming achieves the asymptotically optimal sum rate." *IEEE Transactions on Signal Processing*, **56**(8):3713–3726, 2008.

[WR13] Qing Wang and Balaji Rengarajan. "Recouping opportunistic gain in dense base station layouts through energy-aware user cooperation." In *IEEE 14th International Symposium and Workshops on a World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, pp. 1–9. IEEE, 2013.

[WSS06] Hanan Weingarten, Yossef Steinberg, and Shlomo Shamai. "The Capacity Region of the Multiple-Input-Multiple-Output Broadcast Channel." *IEEE Transactions on Information Theory*, **52**(9):3936–3964, September 2006.

[WTS13] Xinzhou Wu, Saurabha Tavildar, Sanjay Shakkottai, Tom Richardson, Junyi Li, Rajiv Laroia, and Aleksandar Jovicic. "FlashLinQ: A synchronous distributed scheduler for peer-to-peer ad hoc networks." *IEEE/ACM Transactions on Networking (ToN)*, **21**(4):1215–1228, 2013.

[WZ76] Aaron D Wyner and Jacob Ziv. "The rate-distortion function for source coding with side information at the decoder." *IEEE Transactions on Information Theory*, **22**(1):1–10, 1976.

[YG06] Taesang Yoo and Andrea Goldsmith. "On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming." *IEEE Journal on Selected Areas in Communications*, **24**(3):528–541, 2006.

[YGK17] Yaoqing Yang, Pulkit Grover, and Soummya Kar. "Coded Distributed Computing for Inverse Problems." In *Advances in Neural Information Processing Systems*, pp. 709–719, 2017.

[YHG16] N J. Yadwadkar, B. Hariharan, J. Gonzalez, and R H. Katz. "Multi-Task Learning for Straggler Avoiding Predictive Job Scheduling." *Journal of Machine Learning Research*, **17**(4):1–37, 2016.

[YLL16] Yang You, Xiangru Lian, Ji Liu, Hsiang-Fu Yu, Inderjit S Dhillon, James Demmel, and Cho-Jui Hsieh. "Asynchronous parallel greedy coordinate descent." In *Advances in Neural Information Processing Systems*, pp. 4682–4690, 2016.

[Zai14] Abdellatif Zaidi. "Achievable Regions for Interference Channels with Generalized and Intermittent Feedback." In *IEEE International Symposium on Information Theory Proceedings (ISIT)*, pp. 1026–1030. IEEE, 2014.

[ZCD12] Matei Zaharia, Mosharaf Chowdhury, Tathagata Das, Ankur Dave, Justin Ma, Murphy McCauley, Michael J Franklin, Scott Shenker, and Ion Stoica. "Resilient distributed datasets: A fault-tolerant abstraction for in-memory cluster computing." In *Proceedings of the 9th USENIX conference on Networked Systems Design and Implementation*, pp. 2–2. USENIX Association, 2012.

[ZME04] Sina Zahedi, Mehdi Mohseni, and Abbas El Gamal. "On the capacity of AWGN relay channels with linear relaying functions." In *IEEE International Symposium on Information Theory (ISIT)*, pp. 399–399, 2004.

[ZWC15] Ran Zhang, Miao Wang, Lin X Cai, Zhongming Zheng, Xuemin Shen, and Liang-Liang Xie. "LTE-unlicensed: the future of spectrum aggregation for cellular networks." *IEEE Wireless Communications*, **22**(3):150–159, 2015.

[ZY13] Hui Zhang and Wotao Yin. "Gradient methods for convex minimization: better rates under weaker conditions." *arXiv preprint arXiv:1303.4645*, 2013.