UC Irvine

UC Irvine Electronic Theses and Dissertations

Title

Limited Feedback Design for Multiuser Networks

Permalink

https://escholarship.org/uc/item/40x3k1xg

Author

LIU, XIAOYI

Publication Date

2017

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA,
IRVINE

Limited Feedback Design for Multiuser Networks

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Electrical and Computer Engineering

by

Xiaoyi Liu

Dissertation Committee:
Professor Hamid Jafarkhani, Chair
Professor Ender Ayanoglu
Professor Syed Ali Jafar

2017

© 2017 Xiaoyi Liu

DEDICATION

To my boyfriend and my parents, for their unwavering support.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

This thesis would never have been possible without the support and guidance from my advisor, Professor Hamid Jafarkhani. Prof. Jafarkhani always gave me enough freedom to explore my own interest in the research, and provided valuable directions when I needed. His high standard towards research and other things has set a great example for me, and inspires me to be a better person in life.

I would like to thank other members in our research group over these years: Teddy Liangbin Li, Sina Poorkasmaei, Erdem Koyuncu, Mehdi Ganji, Jun Guo and Xun Zou. We have had so many insightful discussions and fun events, making my five years here fully filled with joy.

I want to thank my dearest friends Zach Dehua Zhao, Shu Xiong, Zhuojun Yang and Li Chen for accompanying me during these five years. We had such great memories together, and I am sincerely grateful for all of these. It is not easy to have some truly good friends that can understand you and tolerate you. We may not be able to see each other so often in the future, but I will remember all of them in my heart.

I would like to thank my parents for their selfless support. They have always believed in me, and given me the courage to explore the future.

Last but not least, I wouldn't have done any of this without the endless support from my boyfriend Dmitriy. He is a smart and adorable person with incredible wisdoms. He can always remind me of the things I neglect and gives me encouragement when I feel confused or frustrated. The best thing that ever happened in my life is to have him.

CURRICULUM VITAE

Xiaoyi Liu

EDUCATION

Doctor of Philosophy in Electrical Engineering and Computer Science **2017**
University of California, Irvine *Irvine, California*

Master of Science in Communication and Information System **2012**
Beijing University of Posts and Telecommunications *Beijing, China*

Bachelor of Science in Communication Engineering **2009**
Beijing University of Posts and Telecommunications *Beijing, China*

RESEARCH EXPERIENCE

Graduate Research Assistant **2012–2017**
University of California, Irvine *Irvine, California*

Research Intern **June 2015 - Sept. 2015**
Samsung Research America *Dallas, Texas*

Research Intern **June 2014 - Aug. 2014**
Bell Labs *Murray Hill, New Jersey*

TEACHING EXPERIENCE

Teaching Assistant **Jan. 2017 - Mar. 2017**
University of California, Irvine *Irvine, CA*

REFEREED JOURNAL PUBLICATIONS

Downlink Non-Orthogonal Multiple Access with Limited Feedback **2017**
Second-round review in IEEE Transactions on Wireless Communications

Outage-Optimized Multicast Beamforming with Distributed Limited Feedback **2017**
IEEE Transactions on Wireless Communications

Amplify-and-Forward Relay Networks with Variable-Length Limited Feedback **2016**
IEEE Transactions on Wireless Communications

Cooperative Quantization for Two-User Interference Channels **2015**
IEEE Transactions on Communications

Multicast Networks with Variable-Length Limited Feedback **2015**
IEEE Transactions on Wireless Communications

REFEREED CONFERENCE PUBLICATIONS

Two-User Downlink Non-Orthogonal Multiple Access with Limited Feedback **July 2017**
IEEE International Symposium on Information Theory

Outage-Optimized Distributed Quantizers for Multicast Beamforming **July 2016**
IEEE International Symposium on Information Theory

Variable-Length Limited Feedback for Amplify-and-Forward Relay Networks **Dec. 2015**
IEEE Global Communications Conference

Distributed Channel Quantization for Two-User Interference Networks **Dec. 2014**
IEEE Global Communications Conference

A Variable-Length Channel Quantizer for Multicast Networks with Two Users **Dec. 2014**
IEEE Global Communications Conference

ABSTRACT OF THE DISSERTATION

Limited Feedback Design for Multiuser Networks

By

Xiaoyi Liu

Doctor of Philosophy in Electrical and Computer Engineering

University of California, Irvine, 2017

Professor Hamid Jafarkhani, Chair

In this dissertation, the potential of limited feedback in multiuser/multinode networks is explored, and our goal is to design efficient and practical quantizers to mitigate the performance loss brought by limited feedback. For the multiple amplify-and-forward relay network, we propose variable-length quantizers (VLQs) with random infinite-cardinality codebooks in contrast to the fixed-length quantizers (FLQs) with finite-cardinality codebooks that cannot attain the full-channel-state-information (full-CSI) performance. We validate through both theoretical proofs and numerical simulations that the proposed VLQs can achieve the full-CSI outage probabilities with finite average feedback rates. We also apply the idea of VLQ to the multicast network, and show that the global VLQ can achieve the minimum full-CSI outage probability with a low average feedback rate. For the two-user interference network where interferences are treated as noise, we introduce the idea of cooperative quantization to allow multiple rounds of feedback communication in the form of conferencing between receivers. For both time-sharing and concurrent transmission strategies, the proposed cooperative quantizers are able to achieve the full-CSI network outage probability of sum-rate and the full-CSI network outage probability of minimum rate, respectively, with only finite average feedback rates. For non-orthogonal multiple access (NOMA) which is recognized as a key technique for 5G, we propose efficient quantizers using variable-length encoding, and prove that in the typical application with two receivers, the losses in the minimum rate

and outage probability decay at least exponentially with the minimum feedback rate. In addition, a sufficient condition for the quantizers to achieve the maximum diversity order is provided. For NOMA with K receivers where $K > 2$, the minimum rate maximization problem is solved within an accuracy of ϵ in time complexity of $O\left(K \log \frac{1}{\epsilon}\right)$.

Chapter 1

Introduction

The limited feedback design for multiuser networks is different in nature from that for the well-studied point-to-point system. As seen from Fig. 1.1, only one channel exists between the transmitter and receiver in the point-to-point system. From the view of the receiver, the local channel state information (CSI) is equivalent to the global one. Therefore, after the receiver obtains the perfect CSI through training sequences sent from the transmitter, it can then send the quantized feedback information to the transmitter. With the feedback information in hand, the transmitter can perform a lot of operations, i.e., power allocation, rate adaptation, to improve the overall transmission performance. Therefore, the limited feedback design in the point-to-point system can be treated as a scalar or vector quantization problem. There have been a plenty of literatures in the studies of limited feedback for the point-to-point systems, i.e., [1, 2, 3, 4].

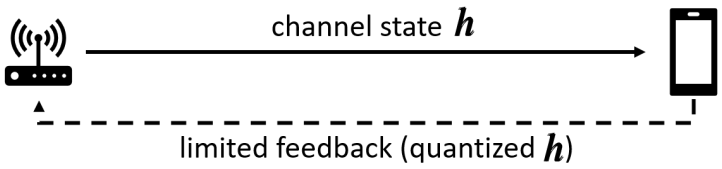


Figure 1.1: Limited feedback in point-to-point wireless systems.

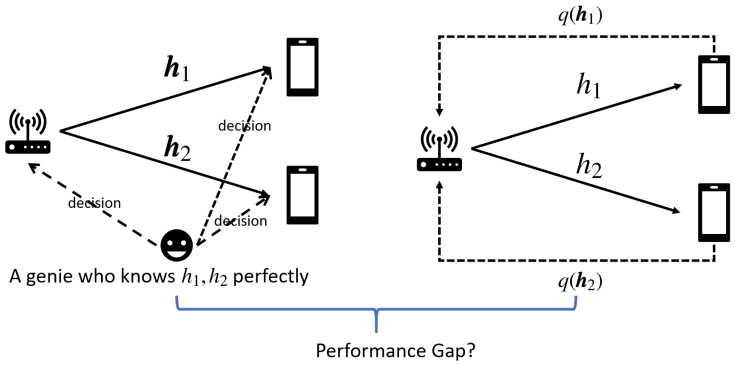


Figure 1.2: Limited feedback in multiuser networks.

Contrarily, as seen from the network with multiple receivers in Fig. 1.2, each receiver only has access to its own local CSI due to geographical separations. There is no such node like "a genie" that can acquire the perfect global CSI and make the decision as a center. Thus, each receiver could only quantize a part of the entire global CSI, and the transmitters need to make decision based on quantized versions of local CSI from receivers. In this scenario, how to most efficiently utilize the distributed quantized feedback to decrease the performance loss becomes a crucial problem for the mulituser networks [5, 6].

In the literatures, the most common approach of distributed quantization consists of two phases: In Phase 1, each receiver quantizes its local CSI independently based on pre-defined codebooks, and then sends a finite number of bits representing the quantized information to other receivers and transmitters. In Phase 2, after decoding the feedback information from all receivers, the transmitters apply the techniques such as beamforming or power control by treating the quantized CSI as the exact unquantized CSI. The quantization loss incurred in Phase 1 often causes more severe error propagation in Phase 2.

In this dissertation, we would like to consider the limited feedback design for multiuser networks from a different perspective. Note that the traditional fixed-length quantizers (VLQs) always bring in performance loss due to the finite size of the codebooks. Instead, we allow the codebooks to have infinite number of codewords, and apply the idea of variable

length encoding to encode their representations. We assign the commonly-used codeword with fewer number of bits, and assign the rare ones with more bits. In this way, the overall average feedback rate can still be finite (or even small). The relaxation of the finite codebook cardinality can greatly reduce or even eliminate the quantization loss of the local CSI, and soften the situation of error propagation in Phase 2. We apply this idea to design practical quantizers for different mulituser networks, and our main contributions are fourfold:

1) To get around the limitations of the traditional FLQs with a finite-cardinality codebook, we propose variable-length quantizers (VLQs) with random infinite-cardinality codebooks for the multiple amplify-and-forward (AF) relay networks subject to the sum or individual power constraints, and prove that VLQs can achieve the same minimum outage probability as the full-CSI case. For the first time, we provide a framework for analyzing the performance of random codebooks using variable-length limited feedback. The derivations based on random codebooks in this chapter can be applied to many other scenarios.

2) A VLQ is proposed for the multicast networks with two users, and the attained performance is the same as that of a system with full CSI. Our work is an important necessary first-step towards the goal of designing VLQs for multicast networks using only local CSI. The availability of a global quantizer that achieves the full-CSI performance opens the door for designing distributed quantizers. It can also be extended to the multicast networks with more than two users.

3) A novel strategy of cooperative quantization is proposed for two-user interference channels where interference signals are treated as noise, which allow multiple rounds of feedback communications in the form of conferencing between receivers. The full-CSI network outage probabilities of sum-rate and minimum-rate are achieved with only finite average feedback rates.

4) Efficient quantizers with variable-length encoding are proposed for the downlink non-orthogonal multiple access (NOMA) networks. We also prove that in the typical application with two receivers, the losses in the minimum rate and outage probability decay at least exponentially with the minimum feedback rate. Additionally, we provide a sufficient condition for the quantizer to achieve the maximum diversity order. For NOMA with K receivers where $K > 2$, the minimum rate maximization problem is solved within an accuracy of ϵ in time complexity of $O\left(K \log \frac{1}{\epsilon}\right)$.

Chapter 2

Amplify-and-Forward Relay Networks with Variable-Length Feedback

In this chapter, we study the channel quantization problem for amplify-and-forward (AF) relay networks and our target is to design a quantizer to minimize the outage probability. It is priorly known that any fixed-length quantizer with a finite-cardinality codebook cannot attain the same minimum outage probability as the case where all nodes in the AF relay networks have access to perfect channel state information (CSI). We propose variable-length quantizers with random infinite-cardinality codebooks for the sum and individual power constraints. We provide theoretical proofs and numerical simulations to validate that the proposed quantizers can achieve the full-CSI outage probabilities with finite average feedback rates.

2.1 Introduction

Cooperative diversity techniques have received significant attention since they can greatly enhance the spectral efficiency and extend the network coverage [7, 8]. In a wireless relay network, the destination node receives signals from the source node with the help of relay nodes in the form of "distributed antennas". Several cooperation strategies, such as amplify-and-forward (AF), decode-and-forward, and compress-and-forward have been proposed in the literature. Among these, AF is an attractive solution with very low complexity that requires no decoding at relay nodes.

In the case of point-to-point wireless communication, the performance of the system depends on the availability of channel state information (CSI) at the transmitter and the design of the corresponding finite-rate feedback [1, 2, 4]. Similarly, the performance of wireless relay networks depends on the availability of CSI at the relay nodes and the destination node [9, 10, 11]. The destination node can acquire the entire CSI through training sequences from the source node and relay nodes. Meanwhile, although each relay node can have the knowledge of its own receiving channel via training sequences from the source node, it does not have a direct access to the channel from itself to the destination node or the channels of other relays. Thus, the relay nodes rely on the feedback information from the destination node [12]. Perfect CSI at the relay nodes requires an "infinite" number of feedback bits from the destination node, which is unrealistic due to the limitations of the feedback links. Hence, in practice, it is desired to design efficient transmission schemes based on quantized CSI for wireless relay networks.

There has been a lot of work on quantized channel feedback in wireless relay networks. In a cooperative network with a single AF relay in [9], power control methods have been analyzed to minimize the outage probability with limited feedback available at the transmitter. When the cooperative network has multiple relays, it is shown in [11, 12] that using relay beam-

forming achieves the full-CSI performance. Relay beamforming based on quantized feedback from the receiver can be implemented in a distributed manner without complex coordination between relays. With the index fed back from the receiver, each relay can select the corresponding relay beamforming vector from the pre-defined codebook. Relay selection is possible to achieve the maximum diversity. However, it incurs an inevitable performance loss in terms of array gain compared to relay beamforming [13]. Therefore, we only consider the channel quantizers using relay beamforming in this chapter. In a cooperative network with multiple AF relays, the capacity loss and bit error probability with quantized feedback have been studied in [10], when each relay node is subject to an individual power constraint. Also, [11] has investigated the optimal beamforming vector for relay nodes in the full-CSI scenario and the outage probability in the limited feedback scenario when the sum power constraint is imposed on the relay nodes. Compared to the full-CSI scenario where all relay nodes know the perfect CSI, the schemes in [10] and [11] always suffer from performance loss.

All of these previous schemes have relied on fixed-length quantizers (FLQs), in which the receiver feeds back the same number of bits for every channel state. In general, the receiver can send a different number of feedback bits for different channel states, resulting in a variable-length quantizer (VLQ). Recently, a VLQ has been proposed to achieve the full-CSI outage probability with a finite feedback rate for the non-cooperative setting of a multiple-input single-output (MISO) system [3]. One can thus expect that a VLQ structure will similarly offer high performance gains in cooperative networks. On the other hand, the results of [3] for MISO systems are not directly applicable to the VLQ design problem in AF relay networks due to the following reasons: (i) In such AF relay networks, the relay nodes are geographically apart from each other, which, unlike the co-located transmit antennas in a MISO system, prevents direct access to the CSI of others. (ii) The amplification of both signal and noise from the first hop brings in a highly-nonlinear dependence on the relay beamforming vector and the channel values to the instantaneous signal-to-noise ratio (SNR).

However, in a MISO system, the SNR is simply given by the inner product of the beamforming and channel vectors. (iii) Both sum and individual power constraints are considered for the AF relay networks. As shown in [12], the individual power constraint causes severe non-convexity, which further hampers the limited feedback design for SNR optimization. Therefore, the distributed nature of the AF relay networks and the highly complicated SNR expressions result in great difficulties in the design and performance analysis of VLQs.

We overcome these difficulties by considering random quantizer codebooks instead of the structured codebooks presented in [3]. We also provide a framework for analyzing the performance of random codebooks using limited feedback in AF relay networks, and the derivations can be applied to many other scenarios with AF relays. We first prove that the outage probabilities of our proposed VLQs are the same as those of the full-CSI scenarios in the sum and individual power constraints, respectively. Then, for the average feedback rate of the proposed VLQ under the sum power constraint, we derive its upper bound to show it is finite. For the average feedback rate of the proposed VLQ under the individual power constraint, we are unable to theoretically prove it is finite due to the complicated SNR expression. Instead, we perform numerical simulations to verify it is finite and small.

Notations: Bold-face letters refer to vectors or matrices. For a vector or matrix $\boldsymbol{x}$, $\boldsymbol{x}^{\top}$ represents its transpose, $\boldsymbol{x}^{\dagger}$ represents its conjugate transpose, $||\boldsymbol{x}||$ is the l^2-norm, and $[\boldsymbol{x}]_i$ denotes its i-th element. The sets of complex, real, and natural numbers are denoted by C, R, and N, respectively. The probability and expectation are represented by $\Pr\{\cdot\}$ and $\mathtt{E}[\cdot]$, respectively. We use the notation $\mathbb{CN}(\boldsymbol{a},\boldsymbol{b})$ to stand for a circularly-symmetric complex Gaussian random vector with mean of $\boldsymbol{a}$ and variance of $\boldsymbol{b}$. Similarly, $\mathbb{N}(\boldsymbol{a},\boldsymbol{b})$ is for a real Gaussian random vector. For any $x \in \mathsf{R}$, $\lfloor x \rfloor$ is the largest integer that is less than or equal to x and $\lceil x \rceil$ is the smallest integer that is larger than or equal to x. For any $x \in \mathsf{C}$, x^* is the conjugate, $\mathrm{Real}(x)$ is the real part, $\mathrm{Imag}(x)$ is the imaginary part, $|x| = \sqrt{[\mathrm{Real}(x)]^2 + [\mathrm{Imag}(x)]^2}$ is the absolute value and $\arg(x) = \arctan\left(\frac{\mathrm{Imag}(x)}{\mathrm{Real}(x)}\right)$ is the argument. For a logical statement ST,

we let $\mathbf{1}\{\mathrm{ST}\} = 1$ when ST is true, and $\mathbf{1}\{\mathrm{ST}\} = 0$ otherwise. The column vector formed by stacking two column vectors $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ together is denoted as $[\boldsymbol{x}_1; \boldsymbol{x}_2]$. Finally, `rand`() returns a single uniformly distributed random number in the interval $(0, 1]$.

2.2 System Model and Problem Formulation

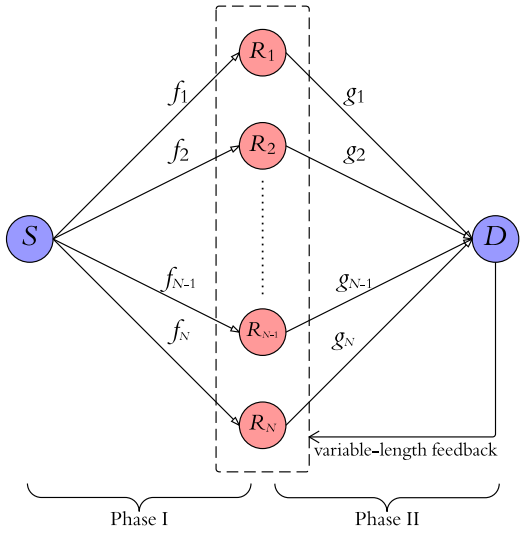


Figure 2.1: System block diagram of amplify-and-forward relay newtorks.

In the AF relay network depicted in Fig. 2.1, a source node S transmits to a destination node D with the aid of N AF relay nodes $R_1, \ldots, R_N$, where $N \geq 2$. Each node is equipped with only a single antenna. Assume that there is no direct link between S and D. Denote the channels from S to R_n and R_n to D by $f_n \sim \mathbb{CN}\left(0, \sigma_{f_n}^2\right)$ and $g_n \sim \mathbb{CN}\left(0, \sigma_{g_n}^2\right)$, respectively. Without loss of generality, we assume $\sigma_{g_1}^2 \leq \sigma_{g_2}^2 \leq \ldots \leq \sigma_{g_N}^2$. The entire channel state is represented by $\boldsymbol{H} = [f_1, \ldots, f_N, g_1, \ldots, g_N]^\top \in \mathsf{C}^{2N \times 1}$. We assume a quasi-static channel model, in which the channels vary independently from one block to another, while remain constant within each block.

In Phase I, the received signal at the n-th relay node R_n is

$$y_{R_n} = \sqrt{P_S} f_n x + v_{R_n},$$

where x is the information bearing symbol sent by S with $\mathtt{E}\left[|x|^2\right] = 1$ for each channel state (the expectation is over all transmitted symbols), and P_S is the average transmit power at S.[1] The background noise v_{R_n} for $n = 1, \ldots, N$ is independent and modeled as $\mathbb{CN}(0, 1)$.

In Phase II, each relay node normalizes and retransmits its received signal y_{R_n}. The normalized signal to be re-transmitted with unit power at R_n is

$$x_{R_n} = \frac{y_{R_n}}{\sqrt{\mathtt{E}_{x, v_{R_n}}\left[|y_{R_n}|^2\right]}} = \frac{\sqrt{P_S} f_n x + v_{R_n}}{\sqrt{P_S |f_n|^2 + 1}}.$$

Thereafter, R_n sends $\sqrt{P_{R_n}} w_n^* x_{R_n}$, where P_{R_n} is the maximum transmit power at R_n and $P_{R_n} |w_n|^2$ is the actual-consumed transmit power. Without loss of generality, $P_S = P_{R_n} = P$ is assumed. Results for other values of P_S and P_{R_n} can be obtained similarly. The received signal at the destination node D is

$$\begin{aligned} y_D &= \sum_{n=1}^{N} g_n \sqrt{P} w_n^* x_{R_n} + v_D = \sum_{n=1}^{N} \frac{P w_n^* f_n g_n x}{\sqrt{P |f_n|^2 + 1}} + \sum_{n=1}^{N} \frac{\sqrt{P} w_n^* g_n v_{R_n}}{\sqrt{P |f_n|^2 + 1}} + v_D \\ &= \sqrt{P} \sum_{n=1}^{N} w_n^* \frac{f_n g_n}{\sqrt{|f_n|^2 + \frac{1}{P}}} x + \tilde{v}_D, \end{aligned} \tag{2.1}$$

where $\tilde{v}_D \triangleq \sum_{n=1}^{N} \frac{w_n^* g_n v_{R_n}}{\sqrt{|f_n|^2 + \frac{1}{P}}} + v_D$ and $v_D \sim \mathbb{CN}(0, 1)$ is the background noise at D. Given f_n and g_n, $\tilde{v}_D$ is distributed as $\tilde{v}_D \sim \mathbb{CN}\left(0, 1 + \sum_{n=1}^{N} |w_n|^2 \frac{|g_n|^2}{|f_n|^2 + \frac{1}{P}}\right)$. From (2.1), the signal-to-

[1] In the remainder of this chapter, we refer to P as the transmit power instead of the average power over all transmitted symbols for conciseness.

noise ratio (SNR) at D is given by

$$\Gamma(\boldsymbol{w}, \boldsymbol{H}) \triangleq P \frac{\left| \sum_{n=1}^{N} w_n^* \frac{f_n g_n}{\sqrt{|f_n|^2 + \frac{1}{P}}} \right|^2}{1 + \sum_{n=1}^{N} |w_n|^2 \frac{|g_n|^2}{|f_n|^2 + \frac{1}{P}}}, \tag{2.2}$$

where $\boldsymbol{w} = [w_1, \ldots, w_N]^\top$ is the relay beamforming vector.

2.2.1 Sum Power Constraint

Consider the sum power constraint for which the sum of the transmit power of all relay nodes is limited by P, i.e., w_n should satisfy $\sum_{n=1}^{N} |w_n|^2 = 1$, or $||\boldsymbol{w}||^2 = 1$ equivalently. The SNR expression of D in (2.2) can be reexpressed as

$$\Gamma(\boldsymbol{w}, \boldsymbol{H}) = P \frac{\boldsymbol{w}^\dagger \boldsymbol{h} \boldsymbol{h}^\dagger \boldsymbol{w}}{\boldsymbol{w}^\dagger (\boldsymbol{I} + \boldsymbol{D}) \boldsymbol{w}}, \tag{2.3}$$

where $\boldsymbol{h} = \left[\frac{f_1 g_1}{\sqrt{|f_1|^2 + \frac{1}{P}}}, \ldots, \frac{f_N g_N}{\sqrt{|f_N|^2 + \frac{1}{P}}} \right]^\top$, $\boldsymbol{I}$ is the $N \times N$ identity matrix and $\boldsymbol{D}$ is an $N \times N$ diagonal matrix with the n-th diagonal element being $\frac{|g_n|^2}{|f_n|^2 + \frac{1}{P}}$.

Consider outage probability as the performance measure throughout this chapter. For a target data rate τ, outage occurs if $\frac{1}{2} \log_2 (1 + \Gamma(\boldsymbol{w}, \boldsymbol{H})) < \tau$, or equivalently, $\Gamma(\boldsymbol{w}, \boldsymbol{H}) < 2^{2\tau} - 1 = \alpha$. In the rest of this chapter, we refer to α as the outage threshold.

In the full-CSI scenario where all nodes are aware of a perfect knowledge of $\boldsymbol{H}$, the optimal beamforming vector $\boldsymbol{w}_{\mathtt{SUM}}^\star$ that maximizes $\Gamma(\boldsymbol{w}, \boldsymbol{H})$ is $\boldsymbol{w}_{\mathtt{SUM}}^\star = \frac{(\boldsymbol{I}+\boldsymbol{D})^{-1}\boldsymbol{h}}{||(\boldsymbol{I}+\boldsymbol{D})^{-1}\boldsymbol{h}||}$ [11], and the maximum SNR is

$$\Gamma(\boldsymbol{w}_{\mathtt{SUM}}^\star, \boldsymbol{H}) = P \sum_{n=1}^{N} \underbrace{\frac{|f_n|^2 |g_n|^2}{|f_n|^2 + |g_n|^2 + \frac{1}{P}}}_{=\Gamma_n}. \tag{2.4}$$

The minimum outage probability is then given as

$$\mathtt{Out}\left(\mathtt{Full}_{\mathtt{SUM}}\right) \triangleq \Pr\left\{\Gamma\left(\boldsymbol{w}_{\mathtt{SUM}}^{\star}, \boldsymbol{H}\right) < \alpha\right\} = \Pr\left\{\sum_{n=1}^{N} \Gamma_n < \frac{\alpha}{P}\right\} = \mathtt{E}_{\boldsymbol{H}}\mathbf{1}\left\{\sum_{n=1}^{N} \Gamma_n < \frac{\alpha}{P}\right\}. \tag{2.5}$$

In the limited-feedback scenario, assume the n-th relay node R_n only knows $|f_n|$ and the destination node D knows $\boldsymbol{H}$ [10, 11].[2] Define $\mathcal{W}_{\mathtt{SUM}} \triangleq \left\{\boldsymbol{w} : \boldsymbol{w} \in \mathsf{C}^{N\times 1}, ||\boldsymbol{w}|| = 1\right\}$. With an arbitrary quantizer $Q_{\mathtt{SUM}} : \mathsf{C}^{2N\times 1} \to \mathcal{W}_{\mathtt{SUM}}$, D maps $\boldsymbol{H}$ to some beamforming vector $Q_{\mathtt{SUM}}\left(\boldsymbol{H}\right) \in \mathcal{W}_{\mathtt{SUM}}$, then, feeds the index of $Q_{\mathtt{SUM}}\left(\boldsymbol{H}\right)$ back to the relay nodes. The index of $Q_{\mathtt{SUM}}\left(\boldsymbol{H}\right)$ is decoded at each relay node and $Q_{\mathtt{SUM}}\left(\boldsymbol{H}\right)$ is recovered as the beamforming vector. The resulting SNR is $\Gamma\left(Q_{\mathtt{SUM}}\left(\boldsymbol{H}\right), \boldsymbol{H}\right)$, and the corresponding outage probability is

$$\mathtt{Out}\left(Q_{\mathtt{SUM}}\right) \triangleq \Pr\left\{\Gamma\left(Q_{\mathtt{SUM}}\left(\boldsymbol{H}\right), \boldsymbol{H}\right) < \alpha\right\}.$$

A closed-form expression for the outage probability is only known for one relay, given in [8, (5)] (by letting $\gamma = \alpha, \bar{\gamma}_1 = \sigma_{f_1}^2, \bar{\gamma}_2 = \sigma_{g_1}^2, m_1 = m_2 = 1$) as:

$$1 - \frac{2e^{-\left(\frac{1}{\sigma_{f_1}^2} + \frac{1}{\sigma_{g_1}^2}\right)\alpha}}{P\sqrt{\sigma_{f_1}^2\sigma_{g_1}^2}}\sqrt{\alpha\left(\alpha+1\right)}\boldsymbol{K}_1\left(\frac{2}{P}\sqrt{\frac{\alpha(\alpha+1)}{\sigma_{f_1}^2\sigma_{g_1}^2}}\right), \tag{2.6}$$

where $\boldsymbol{K}_v(z)$ is the modified bessel function of the second kind [14, (3.471.9)].

2.2.2 Individual Power Constraint

Alternatively, we assume a maximum transmit power constraint P is imposed on each relay node. With the relay beamforming vector $\boldsymbol{\mu} = [\mu_1, \ldots, \mu_N]^{\top}$ (we use $\boldsymbol{\mu}$ to distinguish it from the notation $\boldsymbol{w}$ used for the sum power constraint), the power consumed at the n-th relay

[2] One possible procedure of revealing the knowledge of $\boldsymbol{H}$ to the destination node D can be found in [10].

node R_n is $|\mu_n| P$, thus, $\boldsymbol{\mu}$ will be subject to $|\mu_n| \leq 1$ for $n = 1, \ldots, N$. The optimal solution $\boldsymbol{\mu}^{\star}_{\mathtt{IND}} = [\mu_1^{\star}, \ldots, \mu_N^{\star}]^{\top}$ that maximizes $\Gamma(\boldsymbol{\mu}, \boldsymbol{H})$ in (2.2) is given in [12, Theorem 1] as

$$\mu_n = \begin{cases} 1, & n = \tau_1, \ldots, \tau_{i_0}, \\ \lambda_{i_0} \phi_n, & n = \tau_{i_0+1}, \ldots, \tau_N, \end{cases} \tag{2.7}$$

where $\phi_n = \frac{|f_n|}{|g_n|} \sqrt{|f_n|^2 + \frac{1}{P}}$ for $n = 1, \ldots, N$ and $\phi_{N+1} = 0$; $(\tau_1, \ldots, \tau_N, \tau_{N+1})$ is an ordering of $(1, \ldots, N+1)$ satisfying $\phi_{\tau_1} \geq \phi_{\tau_2} \geq \cdots \geq \phi_{\tau_N} \geq \phi_{\tau_{N+1}}$ and $\tau_{N+1} = N+1$; $\lambda_i = \frac{1 + \sum_{m=1}^{i} \frac{|g_{\tau_m}|^2}{|f_{\tau_m}|^2 + \frac{1}{P}}}{\sum_{m=1}^{i} \frac{|f_{\tau_m} g_{\tau_m}|}{\sqrt{|f_{\tau_m}|^2 + \frac{1}{P}}}}$; i_0 is the smallest i such that $\lambda_i < \phi_{\tau_{i+1}}^{-1}$. Thus, the minimum outage probability is

$$\mathtt{Out}\left(\mathtt{Full}_{\mathtt{IND}}\right) \triangleq \Pr\left\{\Gamma\left(\boldsymbol{\mu}^{\star}_{\mathtt{IND}}, \boldsymbol{H}\right) < \alpha\right\}. \tag{2.8}$$

Define $\mathcal{U}_{\mathtt{IND}} \triangleq \left\{\boldsymbol{\mu} : \boldsymbol{\mu} \in \mathsf{C}^{N \times 1}, |\mu_n| \leq 1, n = 1, \ldots, N\right\}$. The relay beamforming vector selected by the quantizer $Q_{\mathtt{IND}} : \mathsf{C}^{2N \times 1} \rightarrow \mathcal{U}_{\mathtt{IND}}$ is $Q_{\mathtt{IND}}(\boldsymbol{H})$, then, the achieved SNR is $\Gamma\left(Q_{\mathtt{IND}}(\boldsymbol{H}), \boldsymbol{H}\right)$ and the outage probability is $\mathtt{Out}\left(Q_{\mathtt{IND}}\right) \triangleq \Pr\left\{\Gamma\left(Q_{\mathtt{IND}}(\boldsymbol{H}), \boldsymbol{H}\right) < \alpha\right\}$.

In the subsequent sections, we will propose two VLQs respectively for the sum and individual power constraints, and show that the full-CSI outage probabilities $\mathtt{Out}\left(\mathtt{Full}_{\mathtt{SUM}}\right)$ in (2.5) and $\mathtt{Out}\left(\mathtt{Full}_{\mathtt{IND}}\right)$ in (2.8) can be achieved with finite average feedback rates.

2.3 Variable-Length Limited Feedback for the Sum Power Constraint

In this section, we first describe the proposed VLQ for the relay networks subject to the sum power constraint. Afterwards, we show the proposed VLQ can achieve the full-CSI outage

probaiblity $\mathtt{Out}\left(\mathtt{Full}_{\mathtt{SUM}}\right)$ in (2.5) with a finite average feedback rate both theoretically and numerically.

2.3.1 Proposed VLQ

For any given $\boldsymbol{H}$, we propose a VLQ using the random codebook $\{\boldsymbol{w}_i\}_{\mathsf{N}}$, where $\boldsymbol{w}_i \in \mathcal{W}_{\mathtt{SUM}}$ is independent and identically distributed with a uniform distribution on $\mathcal{W}_{\mathtt{SUM}}$ for $i \in \mathsf{N}$ [15]. The random codebook provides a performance benchmark since if certain average performance is attained, one deterministic codebook can be found to surpass this average performance. Given $\{\boldsymbol{w}_i\}_{\mathsf{N}}$, the proposed VLQ is represented by

$$\mathrm{VLQ}_{\mathtt{SUM}} \triangleq \{\boldsymbol{w}_i, \mathcal{S}_i, \mathtt{b}_i\}, \tag{2.9}$$

where $\mathcal{S}_i$ denotes the channel partition region of $\boldsymbol{w}_i$ for $i \in \mathsf{N}$, $\boldsymbol{w}_i$ is the adopted relay beamforming vector when $\boldsymbol{H} \in \mathcal{S}_i$, and $\mathtt{b}_i$ is the binary feedback string representing the index of $\boldsymbol{w}_i$.

Different from the channel partition regions in FLQs which consist of channel states that achieve the best performance with the centroid codeword, the channel partition regions in $\mathrm{VLQ}_{\mathtt{SUM}}$ are set as

$$\mathcal{S}_i \triangleq \begin{cases} \{\boldsymbol{H} : \Gamma(\boldsymbol{w}_0, \boldsymbol{H}) \geq \alpha\} \cup \bigcap_{i \in \mathsf{N}} \{\boldsymbol{H} : \Gamma(\boldsymbol{w}_i, \boldsymbol{H}) < \alpha\}, & i = 0, \\ \{\boldsymbol{H} : \Gamma(\boldsymbol{w}_i, \boldsymbol{H}) \geq \alpha\} \cap \bigcap_{k=0}^{i-1} \{\boldsymbol{H} : \Gamma(\boldsymbol{w}_k, \boldsymbol{H}) < \alpha\}, & i \in \mathsf{N} - \{0\}. \end{cases} \tag{2.10}$$

For $i \in \mathsf{N}$, $\{\boldsymbol{H} : \Gamma(\boldsymbol{w}_i, \boldsymbol{H}) \geq \alpha\}$ is the set of channels that are in non-outage when $\boldsymbol{w}_i$ is the beamforming vector; $\{\boldsymbol{H} : \Gamma(\boldsymbol{w}_i, \boldsymbol{H}) < \alpha\}$ is its complementary set. For any $\boldsymbol{H}$ with $\Gamma(\boldsymbol{w}^{\star}_{\mathtt{SUM}}, \boldsymbol{H}) < \alpha$, all beamforming vectors lead to outage, then, $\mathrm{VLQ}_{\mathtt{SUM}}$ naively chooses

$\boldsymbol{w}_0$ as the beamforming vector; for any $\boldsymbol{H}$ with $\Gamma\left(\boldsymbol{w}_{\mathtt{SUM}}^{\star}, \boldsymbol{H}\right) \geq \alpha$, $\mathrm{VLQ}_{\mathtt{SUM}}$ examines each beamforming vector in $\{\boldsymbol{w}_i\}_{\mathsf{N}}$ sequentially until it finds some $\boldsymbol{w}_i$ satisfying $\Gamma\left(\boldsymbol{w}_i, \boldsymbol{H}\right) \geq \alpha$. In terms of outage probability, the contribution of such $\boldsymbol{w}_i$ is identical to that of the optimal beamforming vector $\boldsymbol{w}_{\mathtt{SUM}}^{\star}$.

Variable-length coding is applied to encode the indices of $\boldsymbol{w}_i$ for $i \in \mathsf{N}$. Concretely, we let $\mathtt{b}_0 \triangleq \{0\}$, $\mathtt{b}_1 \triangleq \{1\}$, $\mathtt{b}_2 \triangleq \{00\}$, $\mathtt{b}_3 \triangleq \{01\}$ and so on for all binary strings in the set $\{0, 1, 00, 01, 10, 11, \ldots\}$.[3] The length of $\mathtt{b}_i$ is $\lfloor \log_2(i+2) \rfloor$.

Based on the random codebook $\{\boldsymbol{w}_i\}_{\mathsf{N}}$, the outage probability and average feedback rate of the proposed quantizer $\mathrm{VLQ}_{\mathtt{SUM}}$ are

$$\begin{aligned}
\mathtt{Out}\left(\mathrm{VLQ}_{\mathtt{SUM}}\right) &\triangleq \mathtt{E}_{\{\boldsymbol{w}_i\}_{\mathsf{N}}} \Pr\left\{\Gamma\left(\boldsymbol{w}_i, \boldsymbol{H}\right) < \alpha, \forall i \in \mathsf{N}\right\} \\
&= \mathtt{E}_{\boldsymbol{H}} \mathtt{E}_{\{\boldsymbol{w}_i\}_{\mathsf{N}}} \mathbf{1}\left\{\Gamma\left(\boldsymbol{w}_i, \boldsymbol{H}\right) < \alpha, \forall i \in \mathsf{N}\right\}, \qquad (2.11) \\
\mathtt{FR}\left(\mathrm{VLQ}_{\mathtt{SUM}}\right) &\triangleq \sum_{i=0}^{\infty} \lfloor \log_2(i+2) \rfloor \times \Pr\left\{\boldsymbol{H} \in \mathcal{S}_i\right\} \\
&= \sum_{i=0}^{\infty} \lfloor \log_2(i+2) \rfloor \times \mathtt{E}_{\boldsymbol{H}} \mathtt{E}_{\{\boldsymbol{w}_i\}_{\mathsf{N}}} \mathbf{1}\left\{\boldsymbol{H} \in \mathcal{S}_i\right\}. \qquad (2.12)
\end{aligned}$$

2.3.2 Outage Optimality

Theorem 2.1 states that the outage probability of our proposed quantizer $\mathrm{VLQ}_{\mathtt{SUM}}$ is the same as the full-CSI outage probability in (2.5). The proof of the theorem can be found in Appendix A.1.

[3]The proposed VLQ in (2.9) can be extended to the case of prefix-free codes. In other words, there is a prefix-free code for every quantizer designed in this chapter. Suppose $\{\boldsymbol{w}_i\}_{\mathsf{N}}$ is a fixed-structured infinite-cardinality codebook whose performance is no worse than that of random codebooks. Therefore, it can achieve the full-CSI outage probability for the relay newtork. Let the codeword length of $\boldsymbol{w}_i$ be $l_i \triangleq \lceil 2\log_2(i+1)+1 \rceil$ for $i \in \mathsf{N}$ [16, Example 1]. It is straightforward to show that $\sum_{i \in \mathsf{N}} 2^{-l_i} \leq 1$. According to the Kraft's inequality, this code is prefix-free. Moreover, since $l_i = \lceil 2\log_2(i+1)+1 \rceil \leq 2\log_2(i+1)+2$, the average feedback rate of this code is also finite following the same derivations in the proof of Theorem 2.2.

Theorem 2.1. *For any $P > 0$, we have*

$$\mathtt{Out}\left(\mathrm{VLQ}_{\mathtt{SUM}}\right) = \mathtt{Out}\left(\mathtt{Full}_{\mathtt{SUM}}\right). \tag{2.13}$$

In the following, we provide an intuitive explanation of the result in Theorem 2.1. For a given $\boldsymbol{H}$ with $\Gamma\left(\boldsymbol{w}^{\star}_{\mathtt{SUM}}, \boldsymbol{H}\right) > \alpha$, to achieve the same non-outage performance as the optimal beamforming vector $\boldsymbol{w}^{\star}_{\mathtt{SUM}}$, one should use a unit-normal vector $\boldsymbol{w} \in \mathcal{W}_{\mathtt{SUM}}$ that is "close" enough to $\boldsymbol{w}^{\star}_{\mathtt{SUM}}$ such that $\Gamma\left(\boldsymbol{w}, \boldsymbol{H}\right) \geq \alpha$. We show that there exists a non-zero probability region in the unit sphere where all the unit-normal vectors result in non-outage. However, to "closely" represent $\boldsymbol{w}^{\star}_{\mathtt{SUM}}$ for any such $\boldsymbol{H}$, we need infinitely many beamforming vectors in the codebook $\{\boldsymbol{w}_i\}_{\mathsf{N}}$ to capture at least one in that non-outage region. Obviously, a FLQ with a finite feedback rate will not succeed. Whereas our VLQ proposed in (2.9) includes infinitely many beamforming vectors to achieve the full-CSI outage probability while perserves a finite average feedback rate.

2.3.3 Average Feedback Rate

Theorem 2.2 provides an upper bound on the average feedback rate of $\mathrm{VLQ}_{\mathtt{SUM}}$, the proof of which is presented in Appendix A.2.

Theorem 2.2. *For any $P > 0$, we have*

$$\mathtt{FR}\left(\mathrm{VLQ}_{\mathtt{SUM}}\right) \leq C_0 + C_1 e^{-\frac{\alpha}{P \times \sigma^2_{g_N}}} \left[\frac{1}{P} + \frac{1}{P^N}\right] \left[1 + \left(\frac{\alpha}{P}\right)^N\right], \tag{2.14}$$

where $C_0, C_1 > 0$ are constants that are independent of α and P.

Since $e^{-\frac{\alpha}{P \times \sigma^2_{g_N}}} \left[\frac{1}{P} + \frac{1}{P^N}\right] \left[1 + \left(\frac{\alpha}{P}\right)^N\right]$ in (2.14) is bounded for any outage threshold $\alpha > 0$ and any transmit power $P > 0$, the average feedback rate of $\mathrm{VLQ}_{\mathtt{SUM}}$ is finite. As shown in

the numerical simulations, the average feedback rate can actually be very small.

2.3.4 Numerical Simulations

In this section, we provide numerical simulations of the outage probability and the average feedback rate of $\text{VLQ}_{\texttt{SUM}}$. We let $\alpha = 1$, and $\left(\sigma_{f_1}^2, \sigma_{f_2}^2\right) = (1, 0.8), \left(\sigma_{g_1}^2, \sigma_{g_2}^2\right) = (0.7, 0.9)$ for two relays; $\left(\sigma_{f_1}^2, \sigma_{f_2}^2, \sigma_{f_3}^2\right) = (1, 0.8, 0.6), \left(\sigma_{g_1}^2, \sigma_{g_2}^2, \sigma_{g_3}^2\right) = (0.5, 0.7, 0.9)$ for three relays; and $\left(\sigma_{f_1}^2, \sigma_{f_2}^2, \sigma_{f_3}^2, \sigma_{f_4}^2\right) = (1, 0.8, 0.6, 0.4), \left(\sigma_{g_1}^2, \sigma_{g_2}^2, \sigma_{g_3}^2, \sigma_{g_4}^2\right) = (0.3, 0.5, 0.7, 0.9)$ for four relays. Other values of α and channel variances will show similar simulation results. For each value of the transmit power P, a sufficiently large number of channel realizations are generated such that at least $1,000$ outage events can be observed. For each channel state realization with non-outage in the full-CSI case, a random relay beamforming vector $\boldsymbol{w} \in \mathcal{W}_{\texttt{SUM}}$ is generated repeatedly until a vector that results in non-outage is found. With such simulation settings, the average feedback rate is computed as the average number of feedback bits, and the simulated outage probability is the number of outage incidents divided by the number of all channel state realizations. No endless iteration has occurred when $\boldsymbol{w}$ is generated in any channel state realization.

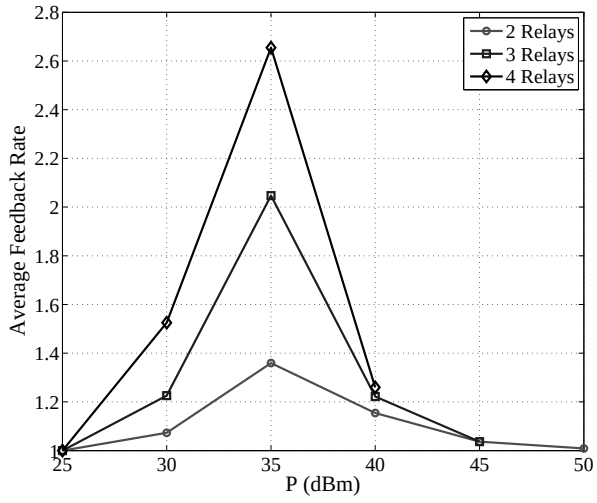


Figure 2.2: Simulated average feedback rates of $\text{VLQ}_{\texttt{SUM}}$ (dBm is 10log(P/1mW)).

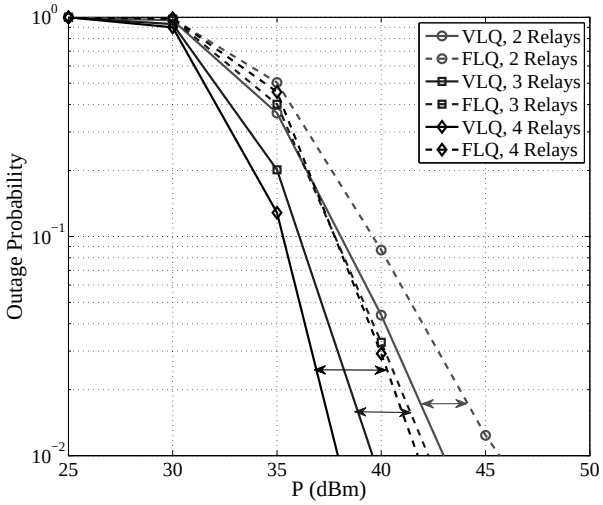


Figure 2.3: Simulated outage probabilities of $\text{VLQ}_{\texttt{SUM}}$ and $\text{FLQ}_{\texttt{SUM}}$ (dBm is 10log(P/1mW)).

In Fig. 2.2, when $N = 2, 3$ or 4, the simulated average feedback rate is no larger than 3 bits per channel state for any P. In Fig. 5.7, we compare the outage probabilities of $\text{VLQ}_{\texttt{SUM}}$ and the FLQ in [11] denoted by $\text{FLQ}_{\texttt{SUM}}$.[4] Given $\boldsymbol{H}$ and the random codebook $\{\boldsymbol{w}_i\}_{i=0,\ldots,2^B-1}$ where $\boldsymbol{w}_i \in \mathcal{W}_{\texttt{SUM}}$, $\text{FLQ}_{\texttt{SUM}}$ chooses the relay beamforming vector as $\text{FLQ}_{\texttt{SUM}}(\boldsymbol{H}) \triangleq \text{argmax}_{\boldsymbol{w}\in\{\boldsymbol{w}_i\}_{i=0,\ldots,2^B-1}} \Gamma(\boldsymbol{w}, \boldsymbol{H})$, thus, the feedback rate of $\text{FLQ}_{\texttt{SUM}}$ is B bits per channel state. We let $B = 2, 3, 3$ for $N = 2, 3, 4$, respectively. These values of B are close to (but still larger than) the average feedback rates of $\text{VLQ}_{\texttt{SUM}}$ with the same relay network configurations in Fig. 2.2. Therefore, $\text{VLQ}_{\texttt{SUM}}$ shows great improvement in outage probability compared to $\text{FLQ}_{\texttt{SUM}}$.

[4]Theorem 2.1 has shown $\text{VLQ}_{\texttt{SUM}}$ achieves the full-CSI outage probability in (2.5). Hence, the simulated outage probability of $\text{VLQ}_{\texttt{SUM}}$ in Fig. 5.7 is also the simulated full-CSI outage probability.

2.4 Variable-Length Limited Feedback for the Individual Power Constraint

In this section, we propose a VLQ design for the relay network subject to the individual power constraint and prove it can attain the optimal outage probability in (2.8). Due to the intractable theoretical analysis on the average feedback rate of the proposed VLQ, numerical simulations are presented to show it is finite.

2.4.1 Proposed VLQ

For any given $\boldsymbol{H}$, the relay beamforming vector $\boldsymbol{\mu}_i = [\mu_{i,1}, \ldots, \mu_{i,N}]^\top \in \mathcal{U}_{\mathtt{IND}}$ in the random codebook $\{\boldsymbol{\mu}_i\}_{\mathsf{N}}$ is constructed by

$$\mu_{i,n} = |\mu_{i,n}|\, e^{j \arg(\mu_{i,n})}, |\mu_{i,n}| = \mathtt{rand}(), \arg(\mu_{i,n}) = 2\pi \times \mathtt{rand}(). \tag{2.15}$$

The proposed VLQ for the individual power constraint is represented by

$$\mathrm{VLQ}_{\mathtt{IND}} \triangleq \{\boldsymbol{\mu}_i, \mathcal{P}_i, \mathtt{d}_i\}, \tag{2.16}$$

where $\boldsymbol{\mu}_i$ is the assigned relay beamforming vector when $\boldsymbol{H}$ falls in the channel partition region $\mathcal{P}_i$, and $\mathtt{d}_i$ is the binary representation for the index of $\boldsymbol{\mu}_i$. Similar to (2.10), the channel partition region $\mathcal{P}_i$ is given by

$$\mathcal{P}_i \triangleq \begin{cases} \{\boldsymbol{H} : \Gamma(\boldsymbol{\mu}_0, \boldsymbol{H}) \geq \alpha\} \cup \bigcap_{i \in \mathsf{N}} \{\boldsymbol{H} : \Gamma(\boldsymbol{\mu}_i, \boldsymbol{H}) < \alpha\}, & i = 0, \\ \{\boldsymbol{H} : \Gamma(\boldsymbol{\mu}_i, \boldsymbol{H}) \geq \alpha\} \cap \bigcap_{k=0}^{i-1} \{\boldsymbol{H} : \Gamma(\boldsymbol{\mu}_k, \boldsymbol{H}) < \alpha\}, & i \in \mathsf{N} - \{0\}. \end{cases} \tag{2.17}$$

The design for $\mathtt{d}_i$ can also be inherited from that for $\mathtt{b}_i$ in $\mathrm{VLQ}_{\mathtt{SUM}}$, thus, the length of $\mathtt{d}_i$ is $\lfloor \log_2(i+2) \rfloor$. The key difference between $\mathrm{VLQ}_{\mathtt{IND}}$ and $\mathrm{VLQ}_{\mathtt{SUM}}$ lies in the construction of the beamforming vectors in the random codebook.

With $\{\boldsymbol{\mu}_i\}_{\mathsf{N}}$, the outage probability and average feedback rate of $\mathrm{VLQ}_{\mathtt{IND}}$ are

$$\begin{aligned}
\mathtt{Out}\left(\mathrm{VLQ}_{\mathtt{IND}}\right) &\triangleq \mathtt{E}_{\{\boldsymbol{\mu}_i\}_{\mathsf{N}}} \Pr\left\{\Gamma\left(\boldsymbol{\mu}_i, \boldsymbol{H}\right) < \alpha, \forall i \in \mathsf{N}\right\} \\
&= \mathtt{E}_{\boldsymbol{H}} \mathtt{E}_{\{\boldsymbol{\mu}_i\}_{\mathsf{N}}} \left[\mathbf{1}\left\{\Gamma\left(\boldsymbol{\mu}_i, \boldsymbol{H}\right) < \alpha, \forall i \in \mathsf{N}\right\}\right], \qquad (2.18) \\
\mathtt{FR}\left(\mathrm{VLQ}_{\mathtt{IND}}\right) &\triangleq \sum_{i=0}^{\infty} \lfloor \log_2(i+2) \rfloor \times \Pr\left\{\boldsymbol{H} \in \mathcal{P}_i\right\} \\
&= \sum_{i=0}^{\infty} \lfloor \log_2(i+2) \rfloor \times \mathtt{E}_{\boldsymbol{H}} \mathtt{E}_{\{\boldsymbol{\mu}_i\}_{\mathsf{N}}} \left[\mathbf{1}\left\{\boldsymbol{H} \in \mathcal{P}_i\right\}\right]. \qquad (2.19)
\end{aligned}$$

2.4.2 Outage Optimality and Average Feedback Rate

The following theorem shows that in the relay network with the individual power constraint, our proposed VLQ achieves the full-CSI outage probability in (2.8). The proof of the theorem is provided in Appendix A.3.

Theorem 2.3. *For any $P > 0$, we have*

$$\mathtt{Out}\left(\mathrm{VLQ}_{\mathtt{IND}}\right) = \mathtt{Out}\left(\mathtt{Full}_{\mathtt{IND}}\right). \qquad (2.20)$$

Due to the highly complicated expression of $\boldsymbol{\mu}^{\star}_{\mathtt{IND}}$ in (2.7) which hinders from further tractable analysis, we are unable to provide a closed-form upper bound on the average feedback rate $\mathtt{FR}\left(\mathrm{VLQ}_{\mathtt{IND}}\right)$ to theoretically prove its finity. However, we can still perform numerical simulations to verify this, i.e., Fig. 2.4 shows the average feedback rate will be finite under different simulation parameters and network configurations.[5]

[5]We use the same parameters for channel variances and $\alpha = 1$ here as in Section 2.3.4.

In Fig. 2.5, we also compare the outage probabilities of $\text{VLQ}_{\texttt{IND}}$ and the FLQ in [10, Section V] denoted by $\text{FLQ}_{\texttt{IND}}$. The feedback rates of $\text{FLQ}_{\texttt{IND}}$ are chosen as $B = 2, 3, 4$ bits per channel state for $N = 2, 3, 4$, respectively. Although the average feedback rate of $\text{VLQ}_{\texttt{IND}}$ is smaller than that of $\text{FLQ}_{\texttt{IND}}$ with the same network configuration, $\text{VLQ}_{\texttt{IND}}$ has obtained much smaller outage probability compared to $\text{FLQ}_{\texttt{IND}}$.

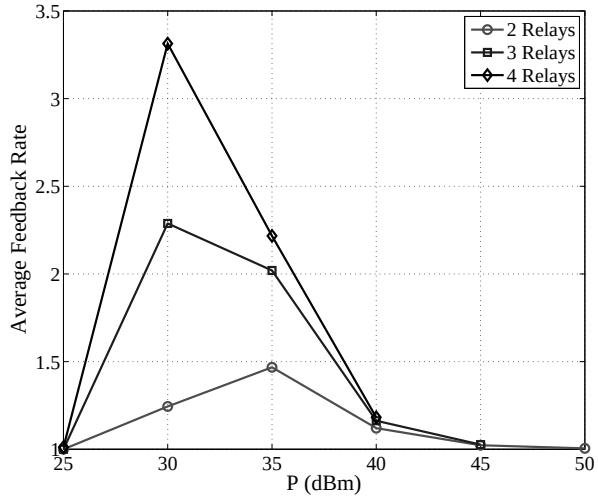


Figure 2.4: Simulated average feedback rates of $\text{VLQ}_{\texttt{IND}}$ (dBm is 10log(P/1mW)).

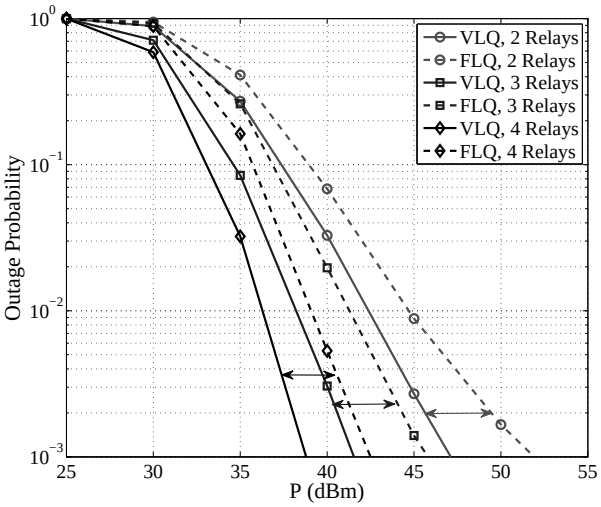


Figure 2.5: Simulated outage probabilities of $\text{VLQ}_{\texttt{IND}}$ and $\text{FLQ}_{\texttt{IND}}$ (dBm is 10log(P/1mW)).

2.5 Conclusions

In this chapter, we have proposed VLQs for the AF relay networks respectively subject to the sum and individual power constraints, and showed the proposed VLQs can achieve the full-CSI outage probabilities with finite average feedback rates. In the future, we intend to work on the VLQ design for the multi-user relay networks with the sum or individual power constraint, and the goal is still to approach the full-CSI outage probability with a finite average feedback rate.

Chapter 3

Multicast Networks with Variable-Length Limited Feedback

We investigate the channel quantization problem for two-user multicast networks where the transmitter is equipped with multiple antennas and either receiver is equipped with only a single antenna. Our goal is to design a global quantizer to minimize the outage probability. It is known that any fixed-length quantizer with a finite-cardinality codebook cannot obtain the same minimum outage probability as the case where all nodes in the network know perfect channel state information (CSI). To achieve the minimum outage probability, we propose a variable-length global quantizer that knows perfect CSI and sends quantized CSI to the transmitter and receivers. With a random infinite-cardinality codebook, we prove that the proposed quantizer is able to achieve the minimum outage probability with a low average feedback rate. We also extend the proposed quantizer to the multicast networks with more than two users. Numerical simulations validate our theoretical analysis.

3.1 Introduction

It is known that using more than one antenna at the transmitters can greatly improve the performance of communication systems. However, the performance depends on the availability of channel state information (CSI) at the transmitters and receivers [1, 4, 10]. Receivers can obtain CSI through training sequences; however, the transmitters must rely on the feedback information from receivers to do so. Additionally, perfect CSI at the transmitters requires an "infinite" number of feedback bits, which is unrealistic due to the limitations of feedback links. Therefore, it is more practical to employ quantized CSI to design efficient transmission schemes for wireless networks.

There has been a lot of work on channel quantization in point-to-point multiple antenna systems. An overview of research on limited feedback can be found in [17]. In multiple-input single-output (MISO) systems, a fixed-length quantizer (FLQ) is proposed in [1] to maximize the capacity by applying the beamforming vector at the transmitter. In FLQs, the number of feedback bits per channel state is a fixed positive integer. Compared to the case that all the nodes know CSI perfectly, fixed-length quantization always suffers from some performance loss. On the other hand, [3] proposes a variable-length quantizer (VLQ) to achieve the full-CSI outage probability with a low average feedback rate. VLQs allow binary codewords of different lengths to represent different channel states. It has been shown in [3] that variable-length quantization does not suffer from performance loss in MISO systems.

In this chapter, we study the channel quantization problem in multicast networks with two receivers. We use transmit beamforming and consider the outage probability gap between the proposed quantizer and the full-CSI case. For a FLQ, the standard encoding rule is to choose the codeword "closest" to the channel state. For any finite-cardinality codebook, the outage probability of a FLQ is strictly worse than that of the full-CSI case [3]. To achieve the full-CSI outage probability with a finite average feedback rate, we propose a VLQ with

a codebook of infinite cardinality. We incorporate the idea of variable-length coding and expect that in such a VLQ, the codeword covering a larger partition of channel space can be represented by a fewer number of bits. In this way, the average feedback rate can be made finite.

Based on the above analysis, we propose a VLQ in multicast networks that has access to full CSI and sends quantized CSI to the transmitter and receivers via error-free and delay-free feedback links. We consider a random codebook with infinite cardinality that is tractable for analysis [18]. Also, if a random codebook can provide a certain level of performance, then one codebook that will surpass this performance can be found. We first prove that the outage probability for the VLQ is the same as that of the full-CSI case. Afterwards, through a derived upper bound on the average feedback rate, we will show that: (i) the average feedback rate is finite and small in the entire range of transmit power; (ii) the average feedback rate will converge to zero when the transmit power approaches infinity or zero. Moreover, we extend the proposed VLQ to the multicast networks with more than two users. In addition to theoretical analysis, numerical simulations are presented to verify the effectiveness of the proposed VLQ.

Our contributions in this chapter are threefold:

1. A novel VLQ is proposed for the multicast networks with two users. It can be extended to the multicast networks with more than two users. The performance of the proposed quantizer is the same as that of a system with full CSI.

2. For the first time, we provide a framework for analyzing the performance of random codebooks using variable-length limited feedback. The derivations based on random codebooks in this chapter can be applied to many other scenarios.

3. Our work is an important necessary first-step towards the goal of designing VLQs for multicast networks using only local CSI. The availability of a global quantizer

that achieves the full-CSI performance, as shown in this chapter, opens the door for designing distributed quantizers.

Notations: For a vector or matrix, $\intercal$ represents its transpose and $\dagger$ represents its conjugate transpose. $\mathbb{C}$ denotes the set of complex numbers and $\mathbb{C}^{m\times n}$ denotes the set of complex vectors or matrices. $\mathbb{CN}(a,b)$ represents a circulary-symmetric complex Gaussian random variable (r.v.) with mean a and covariance b. $\mathrm{E}[\cdot]$ denotes the expectation and $\mathrm{Prob}\{\cdot\}$ denotes the probability. $\mathbb{N}$ is the set consisting of all natural numbers. For any real number x, $\lfloor x \rfloor$ is the largest integer that is less than or equal to x. $\mathbf{1}_{\mathrm{ST}} = 1$ when the logical statement ST is true, and 0 otherwise. Finally, $f_X(\cdot)$ is the probability density function (PDF) for r.v. X.

3.2 System Model

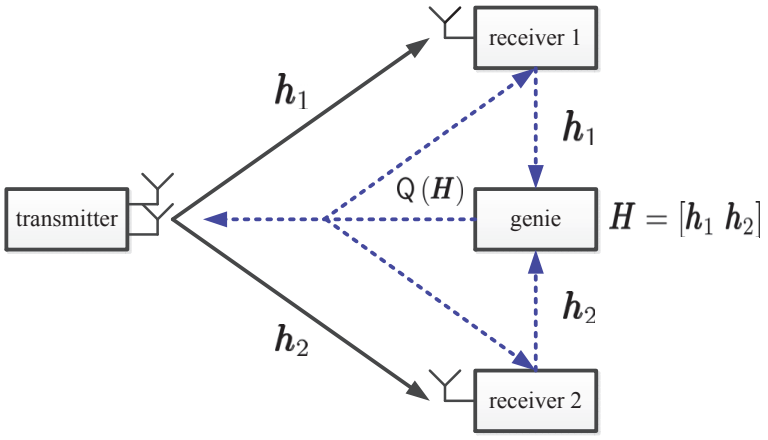


Figure 3.1: System block diagram (solid and dash lines represent signal transmission and feedback links, respectively. The "genie" stands for a global channel quantizer Q).

Consider the multicast network in Fig. 3.1, where a transmitter with t antennas ($t \geq 2$) is sending common information to two singe-antenna receivers. The channel vector from the transmitter to receiver m is denoted by $\boldsymbol{h}_m = [h_{m1} \cdots h_{mt}]^\intercal \in \mathbb{C}^{t\times 1}$, where $h_{mn} \simeq \mathbb{CN}(0,1)$ for $m = 1,2$, $n = 1,\ldots,t$. Let $\chi_m = ||\boldsymbol{h}_m||^2$ for $m = 1,2$. Then the entire channel state is

represented by $\boldsymbol{H} = [\boldsymbol{h}_1\, \boldsymbol{h}_2] \in \mathbb{C}^{t\times 2}$. We assume $\boldsymbol{h}_m$ is perfectly estimated at receiver m and consider a quasi-static block fading channel model in which the channel realizations vary independently from one block to another while remain constant within each block [4].

At the transmitter, $\boldsymbol{x} \in \mathcal{X} \triangleq \{\boldsymbol{x} : \boldsymbol{x} \in \mathbb{C}^{t\times 1}, ||\boldsymbol{x}||^2 = 1\}$ is employed as the beamforming vector and a scalar symbol $s \in \mathbb{C}$ is sent through t antennas. The received signal at receiver m is

$$y_m = \sqrt{P}\boldsymbol{x}^\dagger \boldsymbol{h}_m s + g_m,$$

where P denotes the transmit power and $g_m \simeq \mathbb{CN}(0,1)$ is the additive white Gaussian noise term. We assume $\mathrm{E}\left[|s|^2\right] = 1$. For the multicast network, the maximum achievable rate is $\log_2(1 + P\min_{m=1,2} |\boldsymbol{x}^\dagger \boldsymbol{h}_m|^2)$ [19].[1] Let $\gamma(\boldsymbol{x}, \boldsymbol{H}) = \min_{m=1,2} \left|\boldsymbol{x}^\dagger \boldsymbol{h}_m\right|^2$, then for the target data transmission rate ρ, an outage event will occur if $\log_2(1 + P\gamma(\boldsymbol{x}, \boldsymbol{H})) < \rho$, or equivalently, if $\gamma(\boldsymbol{x}, \boldsymbol{H}) < \frac{2^\rho - 1}{P}$. Without loss of generality, we assume $\rho = 1$ throughout this chapter. Thus, $\frac{2^\rho - 1}{P} = \frac{1}{P}$. Results for other values of ρ can be obtained similarly.

The full-CSI case where perfect CSI is known by all nodes in the multicast network is studied in [19], and the optimal beamforming vector is computed as $\mathtt{Full}(\boldsymbol{H}) = \mathrm{argmax}_{\boldsymbol{x}\in\mathcal{X}}\, \gamma(\boldsymbol{x}, \boldsymbol{H})$. [2] Then the full-CSI outage probability is

$$\mathtt{Out}(\mathtt{Full}) = \mathrm{Prob}\left\{\gamma\left(\mathtt{Full}\left(\boldsymbol{H}\right), \boldsymbol{H}\right) < \frac{1}{P}\right\} = \mathrm{E}_{\boldsymbol{H}} \mathbf{1}_{\gamma(\mathtt{Full}(\boldsymbol{H}), \boldsymbol{H}) < \frac{1}{P}}. \tag{3.1}$$

[1] In this chapter, we only consider the channel quantization problem for transmit beamforming. Although the precoding matrix can have higher rank than the beamforming vector, it can be inferred from [19, Theorem 1] and [19, Theorem 2] that optimal beamforming vector actually achieves the same maximum achievable rate as the optimal precoding matrix in multicast networks with two users. This also holds in the three-user case [20].

[2] For any $\boldsymbol{H}$, $\mathtt{Full}(\boldsymbol{H})$ exists because $\gamma(\boldsymbol{x}, \boldsymbol{H})$ is a continuous function on $\boldsymbol{x}$ and $\mathcal{X}$ is a bounded and closed set. There might exist more than one unit-normal vector that can achieve maximum value of $\gamma(\boldsymbol{x}, \boldsymbol{H})$ and $\mathtt{Full}(\boldsymbol{H})$ can be any one of them.

In contrast to the full-CSI case where the perfect CSI needs to be fed back to all nodes, we consider a global quantizer denoted by Q which only requires perfect CSI to be available at a "genie" in the multicast network. As depicted in Fig. 3.1, the "genie" first gathers $\boldsymbol{h}_1$ and $\boldsymbol{h}_2$ from receivers 1 and 2 via error-free and delay-free feedback links. Then it quantizes $\boldsymbol{H} = [\boldsymbol{h}_1\ \boldsymbol{h}_2]$ and sends limited feedback information $\mathsf{Q}(\boldsymbol{H})$ to both receivers and the transmitter. The "genie" does not have to be a specific node outside the network and it can be either receiver or the transmitter. For example, if receiver 1 plays the role of "genie", it only needs to collect $\boldsymbol{h}_2$ from receiver 2.

For an arbitrary global quantizer Q, the distortion with respect to the outage probability is defined as $\mathsf{Dist} = \mathtt{Out}(\mathsf{Q}) - \mathtt{Out}(\mathtt{Full})$. Since $\mathtt{Out}(\mathtt{Full})$ is invariant for fixed P, minimizing Dist is equivalent to designing a quantizer to minimize $\mathtt{Out}(\mathsf{Q})$. In the subsequent sections, we are going to propose a VLQ and show that even if perfect CSI is no longer available at all nodes, the full-CSI outage probability or zero distortion can still be achieved.

3.3 Channel Quantization and Encoding Rule

In the multicast network, we consider a global VLQ associated with a random codebook $\{\boldsymbol{x}_i\}_{i\in\mathbb{N}}$ where $\boldsymbol{x}_i \in \mathcal{X}$ is independent and identically distributed with a uniform distribution on $\mathcal{X}$ for $i \in \mathbb{N}$ [15]. The random codebook is generated each time the channel state changes and revealed to all nodes in the network. It provides a performance benchmark since if a random codebook can achieve certain performance, one deterministic codebook can be found to surpass this performance. For any realization of $\{\boldsymbol{x}_i\}_{i\in\mathbb{N}}$, the proposed VLQ is represented by

$$\mathsf{Q}_{\mathrm{VLQ}} = \{\boldsymbol{x}_i, \mathcal{R}_i, \mathtt{b}_i\}\,, \tag{3.2}$$

where $\mathcal{R}_i$ denotes the partition channel region of $\boldsymbol{x}_i$ for $i \in \mathbb{N}$. In other words, $\boldsymbol{x}_i$ is used as the transmit beamforming vector when $\boldsymbol{H} \in \mathcal{R}_i$. Also, $\mathtt{b}_i$ is the feedback binary string that represents the index $\boldsymbol{x}_i$. We shall later specify $\mathtt{b}_i$ explicitly for every $i \in \mathbb{N}$.

Let us now specify the partition regions $\mathcal{R}_i$. In this context, our main observation is that for a given H, it is not necessary to always choose the best codeword $\boldsymbol{x}^\star$ that maximizes $\gamma(\boldsymbol{x}, \boldsymbol{H})$ among $\boldsymbol{x} \in \{\boldsymbol{x}_i\}_{i\in\mathbb{N}}$. Any codeword $\boldsymbol{x}$ that enables $\gamma(\boldsymbol{x}, \boldsymbol{H}) \geq \frac{1}{P}$ can be applied. Hence, different from channel-partition regions of FLQs which consist of channel states that achieve the best performance with the "centroid" codeword, $\mathcal{R}_0$ in $\mathsf{Q}_{\mathrm{VLQ}}$ is set as

$$\mathcal{R}_0 = \left\{\boldsymbol{H} : \gamma(\boldsymbol{x}_0, \boldsymbol{H}) \geq \frac{1}{P}\right\} \cup \bigcap_{i\in\mathbb{N}} \left\{\boldsymbol{H} : \gamma(\boldsymbol{x}_i, \boldsymbol{H}) < \frac{1}{P}\right\}, \tag{3.3}$$

and $\mathcal{R}_i$ for $i \in \mathbb{N} - \{0\}$ is set as

$$\mathcal{R}_i = \left\{\boldsymbol{H} : \gamma(\boldsymbol{x}_i, \boldsymbol{H}) \geq \frac{1}{P}\right\} \cap \bigcap_{k=0}^{i-1} \left\{\boldsymbol{H} : \gamma(\boldsymbol{x}_k, \boldsymbol{H}) < \frac{1}{P}\right\}. \tag{3.4}$$

For any $\boldsymbol{x} \in \mathcal{X}$, $\left\{\boldsymbol{H} : \gamma(\boldsymbol{x}, \boldsymbol{H}) < \frac{1}{P}\right\}$ includes all channel states for which an outage incident will happen if $\boldsymbol{x}$ is employed as the beamforming vector, and $\left\{\boldsymbol{H} : \gamma(\boldsymbol{x}, \boldsymbol{H}) \geq \frac{1}{P}\right\}$ is the complement set. Thus $\mathcal{R}_0$ is the union set of channel states for which using any codeword in the codebook as the transmit beamforming vector cannot prevent outage and channel states for which using $\boldsymbol{x}_0$ will not result in outage.[3] For any $i \in \mathbb{N} - \{0\}$, $\mathcal{R}_i$ consists of channel states for which using $\boldsymbol{x}_i$ can prevent outage while using $\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{i-1}$ cannot. It can be easily inferred that $\{\mathcal{R}_i\}$ is a collection of disjoint sets and $\cup_{i\in\mathbb{N}}\mathcal{R}_i$ is equal to the entire channel space.

We apply variable-length coding to encode $\boldsymbol{x}_i$ for $i \in \mathbb{N}$. To be specific, we set $\mathtt{b}_0 = \epsilon$, which

[3]It will be shown in Appendix B.1 that $\mathcal{R}_0$ is equal to the expectation of $\left\{\boldsymbol{H} : \gamma(\boldsymbol{x}_0, \boldsymbol{H}) \geq \frac{1}{P}\right\} \bigcup \left\{\boldsymbol{H} : \gamma(\mathtt{Full}(\boldsymbol{H}), \boldsymbol{H}) < \frac{1}{P}\right\}$ with regard to the random codebook $\{\boldsymbol{x}_i\}_{i\in\mathbb{N}}$ with probability one. Therefore, $\mathsf{Q}_{\mathrm{VLQ}}$ can determine whether $\boldsymbol{H}$ belongs to this region or not based on the expression of the optimal beamforming vector given by [19, Theorem 2], rather than checking all codewords in $\{\boldsymbol{x}_i\}_{i\in\mathbb{N}}$.

is an empty codeword, [4] $\mathtt{b}_1 = \{0\}$, $\mathtt{b}_2 = \{1\}$, $\mathtt{b}_3 = \{00\}$, $\mathtt{b}_4 = \{01\}$ and sequentially so on for all codewords in the set $\{\epsilon, 0, 1, 00, 01, 10, 11, \ldots\}$. The length of $\mathtt{b}_i$ is $\lfloor \log_2(i+1) \rfloor$.

With perfect CSI and any realization of $\{\boldsymbol{x}_i\}_{i\in\mathbb{N}}$, $\mathtt{Q}_{\mathrm{VLQ}}$ first determines the partition channel region $\mathcal{R}_i$ in which the current channel state $\boldsymbol{H}$ falls according to (3.3) and (3.4). Then the corresponding codeword $\boldsymbol{x}_i$ is chosen and $\lfloor \log_2(i+1) \rfloor$ bits are fed back to notify the index of $\boldsymbol{x}_i$.[5] After decoding the feedback information, $\boldsymbol{x}_i$ is employed by the transmitter as the beamforming vector. Therefore, the average feedback rate of $\mathtt{Q}_{\mathrm{VLQ}}$ is

$$\mathtt{R}\left(\mathtt{Q}_{\mathrm{VLQ}}\right) = \sum_{i=1}^{\infty} \lfloor \log_2(i+1) \rfloor \mathrm{Prob}\left\{\boldsymbol{H} \in \mathcal{R}_i\right\} = \sum_{i=1}^{\infty} \lfloor \log_2(i+1) \rfloor \mathrm{E}_{\boldsymbol{H}} \mathrm{E}_{\{\boldsymbol{x}_i\}_{i\in\mathbb{N}}} \mathbf{1}_{\boldsymbol{H}\in\mathcal{R}_i}. \quad (3.5)$$

The outage probability is given by

$$\mathtt{Out}(\mathtt{Q}_{\mathrm{VLQ}}) = \mathrm{E}_{\{\boldsymbol{x}_i\}_{i\in\mathbb{N}}} \mathrm{Prob}\left\{\gamma\left(\boldsymbol{x}_i, \boldsymbol{H}\right) < \frac{1}{P}, \forall i \in \mathbb{N}\right\} = \mathrm{E}_{\boldsymbol{H}} \mathrm{E}_{\{\boldsymbol{x}_i\}_{i\in\mathbb{N}}} \mathbf{1}_{\gamma(\boldsymbol{x}_i,\boldsymbol{H})<\frac{1}{P}, \forall i\in\mathbb{N}}. \quad (3.6)$$

3.4 Outage Optimality

In this section, we show that the proposed VLQ in (3.2) will achieve the full-CSI outage probability.

Intuitively, to attain the full-CSI outage probability means for any $\boldsymbol{H}$ where strict non-outage achieved by the optimal beamforming vector $\mathtt{Full}\left(\boldsymbol{H}\right)$ (i.e., $\gamma\left(\mathtt{Full}\left(\boldsymbol{H}\right), \boldsymbol{H}\right) > \frac{1}{P}$), the proposed VLQ should return a unit-normal vector $\boldsymbol{x}$ that is "close" enough to $\mathtt{Full}\left(\boldsymbol{H}\right)$ so that $\boldsymbol{x}$ also succeeds in $\gamma\left(\boldsymbol{x}, \boldsymbol{H}\right) > \frac{1}{P}$.[6] For such $\boldsymbol{H}$, there exists a certain region in the unit

[4] An empty codeword is used here for illustration. Adding 1 bit to each codeword to avoid an empty codeword only increases the average feedback rate by 1 bit per channel realization, thereby not impacting the result of the average feedback rate being finite.

[5] We reemphasize that $\{\boldsymbol{x}_i\}_{i\in\mathbb{N}}$ refers to the infinite-cardinality codebook while $\boldsymbol{x}_i$ represents any beamforming vector selected from $\{\boldsymbol{x}_i\}_{i\in\mathbb{N}}$.

[6] We will show the channel state $\boldsymbol{H}$ satisfying $\gamma\left(\mathtt{Full}\left(\boldsymbol{H}\right), \boldsymbol{H}\right) = \frac{1}{P}$ has probability zero.

sphere of beamforming vectors with non-zero probability, where all the unit-normal vectors also result in strict non-outage. In order to "closely" represent $\texttt{Full}\left(\boldsymbol{H}\right)$ for any $\boldsymbol{H} \in \mathbb{C}^{t\times 2}$, we need infinitely many codewords in the codebook for the proposed VLQ, so that these infinite vectors ensure at least one efficient vector in that region will eventually be chosen to make $\boldsymbol{H}$ non-outage. This also tells why a FLQ with a finite feedback rate cannot achieve the full-CSI outage probability.

The following theorem says the outage probability of the proposed VLQ is equal to that of the full-CSI case, the proof of which is given in Appendix B.1.

Theorem 3.1. *For any $P > 0$, we have*

$$\texttt{Out}(\mathsf{Q}_{\mathrm{VLQ}}) = \texttt{Out}(\texttt{Full}). \tag{3.7}$$

3.5 Average Feedback Rate

In Section 3.4, we have shown the infinite codebook cardinality is the key to achieve the full-CSI outage probability. In this section, we will show that when variable-length design in Section 3.3 is applied to encode these infinite codewords, a finite average feedback rate is attainable.

Define

$$\begin{aligned}
\mathcal{H}_0 &= \left\{\boldsymbol{H} : \boldsymbol{H} \in \mathbb{C}^{t\times 2}, \chi_1 \geq \tfrac{1}{P}, \chi_2 \geq \tfrac{1}{P}\right\},\\
\mathcal{H}_1 &= \left\{\boldsymbol{H} : \boldsymbol{H} \in \mathcal{H}_0, \gamma\left(\texttt{Full}\left(\boldsymbol{H}\right), \boldsymbol{H}\right) < \tfrac{1}{P}\right\},\\
\mathcal{H}_2 &= \left\{\boldsymbol{H} : \boldsymbol{H} \in \mathcal{H}_0, \gamma\left(\texttt{Full}\left(\boldsymbol{H}\right), \boldsymbol{H}\right) = \tfrac{1}{P}\right\},\\
\mathcal{H}_3 &= \left\{\boldsymbol{H} : \boldsymbol{H} \in \mathcal{H}_0, \gamma\left(\texttt{Full}\left(\boldsymbol{H}\right), \boldsymbol{H}\right) > \tfrac{1}{P}\right\}.
\end{aligned}$$

As defined in Section 3.2, $\chi_m = ||\boldsymbol{h}_m||^2$ for $m = 1, 2$. Based on the encoding rules in (3.3), (3.4) and the random codebook $\{\boldsymbol{x}_i\}_{i\in\mathbb{N}}$, the feedback rate in (3.5) can be rewritten as

$$\mathtt{R}\left(\mathsf{Q}_{\mathrm{VLQ}}\right) \;=\; \sum_{l=1}^{3} \int_{\boldsymbol{H}\in\mathcal{H}_l} \Phi f_{\boldsymbol{H}}(\boldsymbol{H}) \mathrm{d}\boldsymbol{H}, \tag{3.8}$$

where

$$\Phi = \sum_{i=1}^{\infty} p^i (1-p) \lfloor \log_2(i+1) \rfloor, p = \mathrm{Prob}\left\{ \gamma(\boldsymbol{x}_i, \boldsymbol{H}) < \frac{1}{P} \right\}.$$

From the proof of Theorem 3.1 in Appendix B.1, it is directly obtained that $p = 1$ and $\Phi = 0$ for any $\boldsymbol{H} \in \mathcal{H}_1 \cup \mathcal{H}_2$. Hence, $\int_{\boldsymbol{H}\in\mathcal{H}_1} \Phi f_{\boldsymbol{H}}(\boldsymbol{H}) \mathrm{d}\boldsymbol{H} = \int_{\boldsymbol{H}\in\mathcal{H}_2} \Phi f_{\boldsymbol{H}}(\boldsymbol{H}) \mathrm{d}\boldsymbol{H} = 0$. Then $\mathtt{R}\left(\mathsf{Q}_{\mathrm{VLQ}}\right)$ in (3.8) is equivalent to

$$\mathtt{R}\left(\mathsf{Q}_{\mathrm{VLQ}}\right) = \int_{\boldsymbol{H}\in\mathcal{H}_3} \Phi f_{\boldsymbol{H}}(\boldsymbol{H}) \mathrm{d}\boldsymbol{H}. \tag{3.9}$$

The following lemma exhibits an upper bound on Φ, the proof of which is presented in Appendix B.2.

Lemma 3.1. *For any $0 \le p < 1$, we have*

$$\Phi \le p(1-p) + \left(\frac{6}{\log 2} + 2 \right) p^2 + \frac{2}{\log 2} p^2 \log \frac{1}{1-p}. \tag{3.10}$$

Substituting (3.10) into (3.9), it follows that

$$\mathtt{R}\left(\mathsf{Q}_{\mathrm{VLQ}}\right) \le I_1 + I_2 + I_3, \tag{3.11}$$

where

$$I_1 = C_1 \int_{\boldsymbol{H}\in\mathcal{H}_3} p(1-p) f_{\boldsymbol{H}}(\boldsymbol{H}) \mathrm{d}\boldsymbol{H},$$
$$I_2 = C_2 \int_{\boldsymbol{H}\in\mathcal{H}_3} p^2 f_{\boldsymbol{H}}(\boldsymbol{H}) \mathrm{d}\boldsymbol{H},$$
$$I_3 = C_3 \int_{\boldsymbol{H}\in\mathcal{H}_3} p^2 \left(\log \frac{1}{1-p}\right) f_{\boldsymbol{H}}(\boldsymbol{H}) \mathrm{d}\boldsymbol{H},$$

and $C_1 = 1$, $C_2 = \frac{6}{\log 2} + 2$, $C_3 = \frac{2}{\log 2}$. To further proceed, we also need useful bounds on p. For an upper bound on p, using [5, Lemma 2] and [21], we obtain

$$p \leq \sum_{m=1}^{2} \mathrm{Prob}\left\{ \left| \boldsymbol{x}_i^\dagger \boldsymbol{h}_m \right|^2 < \frac{1}{P} \right\} = \sum_{m=1}^{2} \left[1 - \left(1 - \frac{1}{P\chi_m} \right)^{t-1} \right],$$

where the last equality arises from $\mathrm{Prob}\left\{ \left| \boldsymbol{x}_i^\dagger \boldsymbol{h}_m \right|^2 < x \right\} = 1 - \left(1 - \frac{x}{\chi_m}\right)^{t-1}$ [**?**]. Since $(1-a)^b \geq 1 - ab$ for $0 < a < 1$ and $b \geq 1$, $\left(1 - \frac{1}{P\chi_m}\right)^{t-1} \geq 1 - \frac{t-1}{P\chi_m}$. Therefore, p is upper-bounded by

$$p \leq \frac{t-1}{P} \sum_{m=1}^{2} \frac{1}{\chi_m}. \tag{3.12}$$

Another upper bound on p obtained from Lemma B.4 in Appendices B.1 and B.4 is given as

$$p \leq 1 - (1 - \Pi)^{t-1}, \tag{3.13}$$

where

$$\Pi = 1 - \min_{m=1,2} \left[\frac{\left| \left[\mathtt{Full}\left(\boldsymbol{H}\right) \right]^\dagger \boldsymbol{h}_m \right|^2 - \frac{1}{P}}{\chi_m} \right]^2.$$

In addition, a lower bound on p (or equivalently, the upper bound on $1-p$) is given by

$$1-p \leq \text{Prob}\left\{\left|\boldsymbol{x}_i^{\dagger}\boldsymbol{h}_1\right|^2 \geq \frac{1}{P}\right\} = \left(1-\frac{1}{P\chi_1}\right)^{t-1}. \tag{3.14}$$

With bounds on p in (3.12), (3.13), (3.14) and based on (3.11), we deduce an upper bound on $\mathtt{R}\left(\mathsf{Q}_{\text{VLQ}}\right)$ and present it in the following theorem, the detailed proof of which is shown in Appendix B.3.

Theorem 3.2. *For any $P > 0$, we have*

$$\mathtt{R}\left(\mathsf{Q}_{\text{VLQ}}\right) \leq C_0 e^{-\frac{1}{P}}\left[\frac{1}{P}+\frac{1}{P^{2t}}+\frac{\log\left(1+P\right)}{P}\right], \tag{3.15}$$

where $C_0 > 0$ is a constant that is independent of P.

Remark 1: We mainly focus on showing how the number of average feedback bits for Q_{VLQ} changes with P. Therefore, it is beyond the scope of this chapter to find the tightest bound, i.e., the smallest value for C_0.

Remark 2: From (3.15), it can be seen that in the medium and high regions for P, the derived upper bound on average feedback rate is dominated by $e^{-\frac{1}{P}}\left[\frac{1}{P}+\frac{\log(1+P)}{P}\right]$; in the low region for P, it is dominated by $\frac{e^{-\frac{1}{P}}}{P^{2t}}$. Moreover, the upper bound will approach zero when $P \to \infty$ and $P \to 0$. The average feedback rate also behaves like this. This can be intuitively interpreted as follows: when $P \to \infty$, any vector in the codebook will not cause an outage event, while when $P \to 0$, any vector will result in outage. According to the encoding rule of Q_{VLQ}, only empty codewords will be fed back in both situations. Thus the average feedback rate approaches zero.

3.6 Numerical Simulations

In this section, we perform numerical simulations to verify the theoretical results for the outage probability and the average feedback rate.

Simulation Procedure:

1: Initialization: Given t, P. Set $\texttt{Out} = 0$, $\texttt{R} = 0$, $\texttt{Loop} = 0$;
2: **while** $\texttt{Out} < 1000$
3: $\quad \texttt{Index} = 0$;
4: $\quad \texttt{Loop} = \texttt{Loop} + 1$;
5: $\quad$ Generate a realization of $\boldsymbol{H}$;
6: $\quad$ **if** $\gamma\left(\texttt{Full}\left(\boldsymbol{H}\right), \boldsymbol{H}\right) < \frac{1}{P}$
7: $\quad\quad \texttt{Out} = \texttt{Out} + 1$;
8: $\quad$ **else**
9: $\quad\quad$ Randomly generate $\boldsymbol{x} \in \mathcal{X}$;
10: $\quad\quad$ **while** $\gamma\left(\boldsymbol{x}, \boldsymbol{H}\right) < \frac{1}{P}$
11: $\quad\quad\quad$ Randomly generate $\boldsymbol{y} \in \mathcal{X}$;
12: $\quad\quad\quad \boldsymbol{x} = \boldsymbol{y}$;
13: $\quad\quad\quad \texttt{Index} = \texttt{Index} + 1$;
14: $\quad\quad$ **end**
15: $\quad$ **end**
16: $\quad \texttt{R} = \texttt{R} + \lfloor \log_2(1 + \texttt{Index}) \rfloor$;
17: **end**
18: **return** $\texttt{R} = \frac{\texttt{R}}{\texttt{Loop}}$, $\texttt{Out} = \frac{\texttt{Out}}{\texttt{Loop}}$.

In the pseudo-code, a sufficiently large number of channel realizations will be generated in order to observe 1000 outage events for each t and P. Moreover, `Out` stands for the simulated outage probability, `R` refers to the simulated average feedback rate and `Loop` records the number of channel realizations. For each channel realization, whether the full-CSI case could prevent outage will be checked in line 6. If not, an outage event is declared in line 7; otherwise, in lines 9 to 14, a random unit-normal vector will be generated repeatedly until one that allows the current channel realization to prevent outage is found, and the index of the selected codeword is `Index`. Together with line 16, the simulated feedback rate is the average number of feedback bits calculated in line 18, where the simulated outage probability is computed as 1000 divided by the number of all channel realizations. In all the simulations,

no endless iteration has been detected, which is equivalent to say that as long as the channel state realization is able to avoid outage in the full-CSI case, a randomly-generated codeword that also prevents outage will eventually be found.

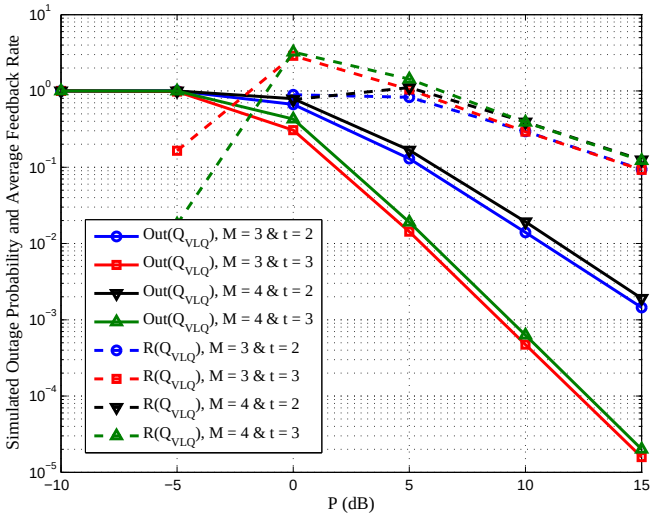


Figure 3.2: Simulated average feedback rates when $t = 2, 3, 4$ (t is the number of transmit antennas).

Fig. 3.2 shows the simulated average feedback rates for $t = 2, 3, 4$. The horizontal axis represents P in decibels. It can be observed that: (i) all the average feedback rates will decrease towards zero when P increases towards infinity or decreases to zero; (ii) the average feedback rate is finite and small for any P; (iii) the average feedback rates for $t = 2, 3, 4$ coincide in the high-P region. These observations correspond to the upper bound derived in Theorem 3.2.[7]

In Figs. 3.3(a) and 3.3(b), we compare the outage probabilities of $\mathsf{Q}_{\mathrm{VLQ}}$ and a traditional FLQ denoted by $\mathsf{Q}_{\mathrm{FLQ}}$. For any given $\boldsymbol{H}$, $\mathsf{Q}_{\mathrm{FLQ}}$ employs B bits to quantize $\boldsymbol{H}$ based on the random codebook $\left\{\boldsymbol{x}_i, i = 0, \ldots, 2^B - 1\right\}$ according to

$$\mathsf{Q}_{\mathrm{FLQ}}\left(\boldsymbol{H}\right) = \underset{\boldsymbol{x}\in\{\boldsymbol{x}_i, i=0,\ldots,2^B-1\}}{\operatorname{argmax}} \gamma\left(\boldsymbol{x}, \boldsymbol{H}\right).$$

[7]For (iii), the upper bound in Theorem 3.2 shows the average feedback rate is dominated by $e^{-\frac{1}{P}}\left[\frac{1}{P} + \frac{\log(1+P)}{P}\right]$ in the high-P region, which is independent of t.

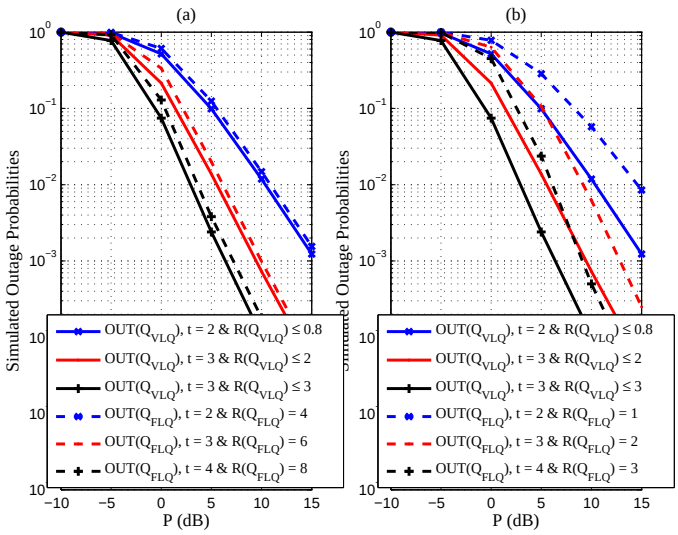


Figure 3.3: Simulated outage probabilities of $\mathsf{Q}_{\mathrm{VLQ}}$ and $\mathsf{Q}_{\mathrm{FLQ}}$ when $t = 2, 3, 4$ (t is the number of transmit antennas).

Then the outage probability is $\mathtt{Out}\left(\mathsf{Q}_{\mathrm{FLQ}}\right) = \mathrm{E}_{\{\boldsymbol{x}_i, i=0,\ldots,2^B-1\}}\mathrm{Prob}\left\{\gamma\left(\mathsf{Q}_{\mathrm{FLQ}}\left(\boldsymbol{H}\right), \boldsymbol{H}\right) < \frac{1}{P}\right\}$, and the average feedback rate is $\mathtt{R}\left(\mathsf{Q}_{\mathrm{FLQ}}\right) = B$. It is observed from Fig. 3.2 that $\mathtt{R}\left(\mathsf{Q}_{\mathrm{VLQ}}\right)$ is no larger than $0.8, 2$ or 3 bits per channel state when $t = 2, 3$ or 4, respectively. Thus in Fig. 3.3(a), we choose the number of feedback bits assigned to $\mathsf{Q}_{\mathrm{FLQ}}$ to be $B = 4, 6$ and 8 when $t = 2, 3, 4$, respectively. Curves in Fig. 3.3(a) demonstrate that $\mathsf{Q}_{\mathrm{VLQ}}$ outperforms $\mathsf{Q}_{\mathrm{FLQ}}$ even when the latter one has a much larger feedback rate. In Fig. 3.3(b), we let $B = 1, 2, 3$ for $t = 2, 3, 4$, which are close to (but still larger than) $\mathtt{R}\left(\mathsf{Q}_{\mathrm{VLQ}}\right)$. It can be seen that the outage probabilities of $\mathsf{Q}_{\mathrm{FLQ}}$ are much worse than those of $\mathsf{Q}_{\mathrm{VLQ}}$. Therefore, it is revealed from Figs. 3.3(a) and 3.3(b) that $\mathsf{Q}_{\mathrm{VLQ}}$ is superior to $\mathsf{Q}_{\mathrm{FLQ}}$.

3.7 Generalization to Multicast Networks with More than Two Users

The quantizer proposed for multicast networks with two users can be applied to the multicast networks with more than two users after slight modifications. We still name it $\mathsf{Q}_{\mathrm{VLQ}}$ for

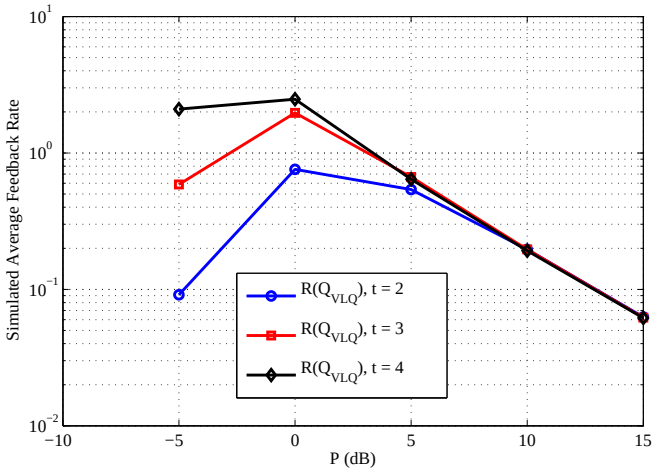


Figure 3.4: Simulated outage probabilities and average feedback rates when $M = 3, 4$ and $t = 2, 3$ (M is the number of receivers and t is the number of transmit antennas).

simplicity. Denote the number of receivers by M. When $M \geq 3$, $\boldsymbol{h}_m = [h_{m1} \cdots h_{mt}]^{\mathsf{T}}$ stands for the channel vector from the transmitter to receiver m with $h_{mn} \simeq \mathbb{CN}(0, 1)$ for $m = 1, \ldots, M$ and $n = 1, \ldots, t$. Then $\boldsymbol{H} = [\boldsymbol{h}_1 \cdots \boldsymbol{h}_M] \in \mathbb{C}^{t \times M}$ represents the entire channel state. Let $\gamma(\boldsymbol{x}, \boldsymbol{H}) = \min_{m=1,\ldots,M} \left|\boldsymbol{x}^{\dagger} \boldsymbol{h}_m\right|^2$ for any $\boldsymbol{x} \in \mathcal{X}$. With such modifications, we can apply the proposed quantizer $\mathsf{Q}_{\mathrm{VLQ}}$ in (3.2) with encoding rules in (3.3), (3.4) to the multicast networks with more than two users.

For the two-user case, we have rigorously proved $\mathsf{Q}_{\mathrm{VLQ}}$ could achieve the optimal outage probability with a finite average feedback rate, and the proofs rely on the closed-form expression of $\mathtt{Full}(\boldsymbol{H})$ given in [19]. But when the multicast network has more than two users, there is no optimal solution for $\mathtt{Full}(\boldsymbol{H})$ in the literature. Thus, we cannot apply the same method to prove $\mathsf{Q}_{\mathrm{VLQ}}$ could achieve the optimal outage probability with a finite average feedback rate in the general case with an arbitrary number of users.

Nevertheless, our proposed quantizer $\mathsf{Q}_{\mathrm{VLQ}}$ is still effective in the multicast networks with more than two users. The "closest" solution we have found for $\mathtt{Full}(\boldsymbol{H})$ is in [22], which uses approximation but generates optimal solutions in many scenarios. In Fig. 3.4, we simulate the outage probabilities and average feedback rates according to the simulation procedure in

Section VI when the numbers of users are $M = 3, 4$ and the numbers of transmit antennas are $t = 2, 3$, respectively. We use the solution in [22] as the base for the simulation procedure, thus its outage probability is treated as the full-CSI performance. We believe that if the exactly optimal solution for $\texttt{Full}\,(\boldsymbol{H})$ is found, our proposed quantizer will also yield the optimal outage probability. Fig. 3.4 shows that $\mathsf{Q}_{\mathrm{VLQ}}$ could attain the full-CSI outage probability using finite average feedback rates when there are more than two users.

3.8 Conclusions and Future Work

In this chapter, we have proved that in the two-user multicast network, the proposed VLQ can achieve the full-CSI outage probability with a low average feedback rate. We have also extended the proposed VLQ to the multicast networks with more than two users. In the future, we intend to work on a distributed quantizer for the multicast network by localizing the proposed VLQ. In this scenario, each receiver only feedbacks its local channel information and no node can acquire the full CSI. We aim to approach or even achieve the full-CSI outage probability at the cost of a finite average feedback rate.

Chapter 4

Cooperative Quantization for Two-User Interference Channels

In this chapter, we introduce cooperative quantizers for two-user interference channels where interference signals are treated as noise. Compared with the conventional quantizers where each receiver quantizes its own channel independently, the proposed cooperative quantizers allow multiple rounds of feedback communication in the form of conferencing between receivers. For both time-sharing and concurrent transmission strategies, we propose different cooperative quantizers to achieve the full-channel-state-information (full-CSI) network outage probability of sum-rate and the full-CSI network outage probability of minimum rate, respectively. Our proposed quantizers only require finite average feedback rates, while the conventional quantizers require infinite rate to achieve the full-CSI performance. For the minimum rate, we also design cooperative quantizers for a joint time-sharing and concurrent transmission strategy that can approach the previously-established optimal network outage probability with a negligible gap. Numerical simulations confirm our cooperative quantizers based on conferencing outperform the conventional quantizers.

4.1 Introduction

Channel quantization in a network with multiple receivers is fundamentally different from that in a point-to-point system. In a point-to-point system, the receiver can acquire the entire channel state information (CSI) and send the corresponding quantized feedback information to the transmitter [1, 2, 3, 4]. On the other hand, in a network with multiple receivers, each receiver only has access to its own local CSI due to different geographical locations of the different receivers. Each receiver can thus, quantize only a part of the entire global CSI, which results in a distributed quantization problem [5, 6].

In the existing work on distributed quantization for networks with multiple receivers [6, 23**?**], each receiver first quantizes its local CSI independently, then, sends a finite number of bits representing quantized information through feedback links to other terminals. After decoding the feedback information from all receivers, each terminal reconstructs the quantized version of the global CSI. Finally, transmission techniques such as beamforming or power control are adopted by treating the global quantized CSI as the exact unquantized CSI. For example, throughput maximization for interference networks based on separate quantized feedback information from receivers is analyzed in [23]. In [6], MMSE-based beamformers are designed for the K-user MIMO interference channels with independent quantized information from each receiver. In [24], after receivers feed back the quantized CSI of both the desired and interfering channels, beamforming vectors are designed in order to maximize the sum-rate. In [25], the technique of inter-cell interference nulling is applied to cellular networks based on quantized feedback channels to reduce outage and improve data rates. The design of distributed quantizers for beamforming in relay-interference networks has been studied in [5]. The performance of these quantizers heavily depends on the number of feedback bits assigned to each receiver for quantization, and always suffers loss when compared with the full-CSI performance.

Another related work is the splitting algorithm for relay selection proposed in [26], in which there is only one receiver in the system model with a one-dimensional source to be quantized. The splitting algorithm performs the quantization through rounds of bit exchanges. However, the splitting algorithm cannot be directly applied to solving the distributed quantization problem in a network with multiple receivers, which usually involves higher-dimensional sources to be quantized.

In this chapter, we propose a novel distributed quantization strategy with multiple rounds of feedback communication in the form of conferencing between receivers. We assume (i) each transmitter has the quantized instantaneous CSI provided by feedback from the receivers through the feedback links between itself and receivers; (ii) receivers are co-located. Thus, each receiver could have the quantized feedback information from others through inter-receiver feedback links. This network model with co-located receivers and inter-receiver feedback links is a valid model for many different scenarios. One example is a cellular network with two neighboring cells and one mobile user at each cell. If the two mobile users are at the cell edges and are geographically close enough, the feedback links between these mobile users (receivers) could be utilized. Another example is that of two Wi-Fi access points coexisting in the same room with different target wireless devices. Since the wireless devices are clearly co-located, one can similarly make use of the feedback links between them for a better performance.[1]

To illustrate the idea of conferencing, we consider the distributed quantization problem for two-user interference channels with the transmission strategies of time-sharing and concurrent transmission. In time-sharing, only one transmitter can be active at any time within

[1]We would however also like to emphasize that our quantizer designs are applicable to networking models where inter-receiver feedback links are not available. In fact, after each receiver broadcasts its quantized local CSI, we can schedule one transmitter to broadcast those feedback bits back to the receivers. In this manner, each receiver can have access to the feedback information from the other receiver. This only incurs at most a doubling of the feedback bits relative to the scenario where inter-receiver feedback links are available. In other words, for every rate-R distributed quantizer that allows inter-receiver conferencing, we can synthesize a rate-$2R$ distributed quantizer that achieves the same performance as the rate-R quantizer and does not need any inter-receiver communication.

the transmission block, thus, no interference exists. On the contrary, in concurrent transmission, two transmitters are allowed to send signals concurrently during the entire transmission block. We consider the treatment of interference as noise, since (i) it allows low complexity transceivers (which is most desired in cellular communications, ad hoc and sensor networks [27, 28]), and (ii) it incurs very little performance loss compared to the information-theoretic outer bounds in [29] that require cooperation among the transmitters and/or multi-user decoding at the receivers. We propose cooperative quantizers that achieve the full-CSI network outage probabilities of sum-rate or the full-CSI network outage probabilities of minimum rate in both time-sharing and concurrent transmission strategies with only finite number of feedback bits. To approach the full-CSI network outage probability of minimum rate, we further propose a joint strategy that combines the quantizers proposed for time-sharing and concurrent transmission. Through numerical simulations, we verify the superiority of our quantizers by comparing them with the conventional ones.

Notations: Bold-face letters refer to vectors or matrices. We use $\top$ to denote the matrix transpose. The sets of complex, real, and natural numbers are represented by $\mathtt{C}$, $\mathtt{R}$, and $\mathtt{N}$, respectively. $\Pr\{\cdot\}$ and $\mathrm{E}[\cdot]$ represent the probability and expectation, respectively. The sets of complex $n \times 1$ vectors and complex $m \times n$ matrices are denoted by $\mathtt{C}^{n\times 1}$ and $\mathtt{C}^{m\times n}$, respectively. We use the notation $\mathtt{CN}(a)$ to represent a circularly-symmetric complex Gaussian random variable (r.v.) with 0 mean and variance $a/2$ per complex dimension. For a r.v. X, $f_X(\cdot)$ is its probability density function (pdf). For sets $\mathcal{A}$ and $\mathcal{B}$, $\mathcal{A}-\mathcal{B} = \{x : x \in \mathcal{A}, x \notin \mathcal{B}\}$, and $\overline{\mathcal{A}}$ is the complement of $\mathcal{A}$. For any $x \in \mathtt{R}$, $\lfloor x \rfloor$ is the largest integer that is less than or equal to x, and $\lceil x \rceil$ is the smallest integer that is larger than or equal to x. For any $x \in \mathtt{C}$, $\mathrm{Real}(x)$ is the real part of x. For any logical statement ST, we let $\mathbf{1}(\mathsf{ST}) = 1$ when ST is true, and $\mathbf{1}(\mathsf{ST}) = 0$ otherwise. Finally, for $b_1, \ldots, b_N \in \{0, 1\}$ where $N \in \mathtt{N} - \{0\}$, the real number $[0.b_1 \cdots b_N]_2$ is the base-2 representation of $\sum_{n=1}^{N} b_n 2^{-n}$.

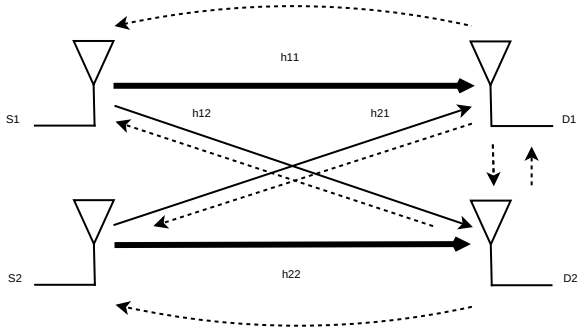


Figure 4.1: The two-user interference channel with limited feedback. The thick solid lines, the thin solid lines, and the dashed lines represent the desired signals, interference signals, and the feedback links, respectively.

4.2 Preliminaries

4.2.1 System model

Consider the interference channel in Fig. 4.1, where transmitters S_1 and S_2 send independent signals to receivers D_1 and D_2 concurrently. All terminals are equipped with only a single antenna. The channel gain from S_k to D_l is denoted by h_{kl}, for $k, l = 1, 2$. We assume that $h_{11}, h_{22} \simeq \mathtt{CN}(1)$ and $h_{12}, h_{21} \simeq \mathtt{CN}(\eta)$ for some $0 \leq \eta \leq 1$. The role of the parameter η is to model different interference signal strengths.[2] The local CSI at Receiver k is $\boldsymbol{h}_k = \left[|h_{1k}|^2, |h_{2k}|^2\right]^\top \in \mathtt{R}^{2\times 1}$, and $\boldsymbol{H} = [\boldsymbol{h}_1, \boldsymbol{h}_2] \in \mathtt{R}^{2\times 2}$ represents the global CSI. We assume that the additive noises at the receivers are distributed as $\mathtt{CN}(1)$ and they are independent of $\boldsymbol{H}$.

We assume a quasi-static block fading channel in which the channel states vary independently from one block to another and remain constant within each block. Either receiver is assumed to perfectly estimate its local CSI and send the associated quantized local CSI to the other receiver and the transmitters in a broadcast manner via error-free and delay-free feedback links [4, 30]. More particularly, in Fig. 4.1, the feedback bits from Receiver 1 will be received

[2]For a simpler exposition, we assume throughout the chapter that both h_{12} and h_{21} have the same variance η. Our results can easily be generalized to the scenario where the channels h_{12} and h_{21} have different variances.

by the two transmitters via the feedback links between Receiver 1 and transmitters, and by the other receiver via the inter-receiver feedback link, as will the feedback bits from Receiver 2.

4.2.2 Transmission strategies

In this chapter, we consider two transmission strategies, namely time-sharing and concurrent transmission. Time-sharing means either transmitter only occupies a proportion of the block to transmit while remains silent in the rest, thus, no interference exists. concurrent transmission refers to the scenario where both transmitters send signals within the entire block, thereby causing interference to each other. Unless otherwise specified, we assume interference signals are dealt with as additive noises throughout this chapter [27, 28].

Let $s_k \simeq \mathtt{CN}(P_k)$ denote the information bearing symbol sent by Transmitter k with the per-transmitter power constraint $\mathrm{E}\left[|s_k|^2\right] = P_k \leq P$ for each channel state, where the expectation is over all transmitted symbols. In other words, the average energy of the symbols of either transmitter is always constrained by P. In time-sharing, let $t_k \in [0,1]$ be the percentage of time within the entire block in which only S_k is active, with the constraint $t_1 + t_2 = 1$. It is optimal for transmitters to use full power under the condition of no interference, i.e., $P_k = P$ for $k = 1, 2$. Therefore, for a given $\boldsymbol{H}$, the data rate at Receiver k is

$$R_{\mathsf{TS},k}(t_k) \triangleq t_k \log_2\left(1 + P\left|h_{kk}\right|^2\right). \tag{4.1}$$

In concurrent transmission, let $P_k = p_k P$, where $p_k \in [0,1]$. For $k, l = 1, 2$ and $k \neq l$, the data rate at Receiver k is

$$R_{\mathsf{CT},k}(p_1, p_2) \triangleq \log_2\left(1 + \frac{p_k P\left|h_{kk}\right|^2}{p_l P\left|h_{lk}\right|^2 + 1}\right) = \log_2\left(1 + \frac{p_k\left|h_{kk}\right|^2}{p_l\left|h_{lk}\right|^2 + \frac{1}{P}}\right). \tag{4.2}$$

We will also consider a joint time-sharing and concurrent transmission strategy, the details of which will be discussed later on.

4.2.3 Network Outage Probability

We consider the network outage probability, i.e., the fraction of channel states at which the network rate measure falls below a target data rate ρ, as our performance measure. Such a performance metric is well-suited for applications where a given constant data rate needs to be sustained for every channel state [25, 31]. Two kinds of rate measures will be considered, namely sum-rate and minimum rate. Specifically, in time-sharing, the network outage probability of sum-rate and the network outage probability of minimum rate are defined as $\Pr\left\{\sum_{k=1}^{2} R_{\mathsf{TS},k}(t_k) < \rho\right\}$ and $\Pr\left\{\min\left\{R_{\mathsf{TS},1}(t_1), R_{\mathsf{TS},2}(t_2)\right\} < \rho\right\}$, respectively; in concurrent transmission, the network outage probability of sum-rate and the network outage probability of minimum rate are defined as $\Pr\left\{\sum_{k=1}^{2} R_{\mathsf{CT},k}(p_1, p_2) < \rho\right\}$ and $\Pr\left\{\min\left\{R_{\mathsf{CT},1}(p_1, p_2), R_{\mathsf{CT},2}(p_1, p_2)\right\} < \rho\right\}$, respectively. For illustrative simplicity and without loss of generality, we assume the sum-rate outage threshold to be $\rho = 2$ bits per channel state and the minimum-rate outage threshold (for either receiver) to be $\rho = 1$. Results for other values of ρ can be obtained likewise.

Our goal is to design efficient quantizers that can achieve the full-CSI network outage probabilities in terms of sum-rate or minimum rate for both time-sharing and concurrent transmission strategies.

4.3 Cooperative Quantization for Network Outage Probability of Sum-Rate

We first study quantizers for concurrent transmission. In this scenario, the sum-rate is given by

$$SR_{\mathsf{CT}}\left(p_1,p_2\right) \triangleq \sum_{k=1}^{2} R_{\mathsf{CT},k}(p_1,p_2).$$

Therefore, we defined the corresponding network outage probability as

$$\mathsf{OUT}_{SR,\mathsf{CT}} \triangleq \Pr\left\{SR_{\mathsf{CT}}\left(p_1,p_2\right)<2\right\}.$$

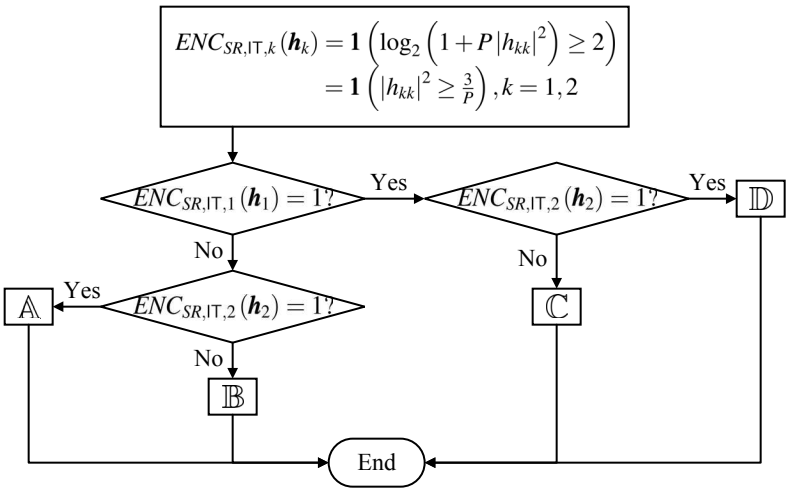
It is proved in [32] that when interference signals are treated as noise, the maximum sum-rate is

$$\max\left\{SR_{\mathsf{CT}}\left(1,0\right),SR_{\mathsf{CT}}\left(0,1\right),SR_{\mathsf{CT}}\left(1,1\right)\right\}.$$

Therefore, the full-CSI network outage probability is

$$\mathsf{OUT}^{\text{opt}}_{SR,\mathsf{CT}} = \Pr\left\{\max\left\{SR_{\mathsf{CT}}\left(1,0\right),SR_{\mathsf{CT}}\left(0,1\right),SR_{\mathsf{CT}}\left(1,1\right)\right\}<2\right\}.$$

In the following, we design a cooperative quantizer, namely $CQ_{SR,\mathsf{CT}}$, that can achieve $\mathsf{OUT}^{\text{opt}}_{SR,\mathsf{CT}}$ exactly with only 1 feedback bit per receiver. The quantizer $CQ_{SR,\mathsf{CT}}$ consists of two local encoders and a unique decoder. For $k=1,2$, the k-th encoder $ENC_{SR,\mathsf{CT},k}:\mathtt{R}^{2\times 1}\rightarrow\{0,1\}$ is located at Receiver k, which will map $\boldsymbol{h}_k$ to 0 or 1 according to $ENC_{SR,\mathsf{CT},k}\left(\boldsymbol{h}_k\right)=0$ or 1. Then, Receiver k will feed back the binary codeword "1" if $ENC_{SR,\mathsf{CT},k}\left(\boldsymbol{h}_k\right)=1$, and otherwise, it will feed back "0". Afterwards, the decoder $DEC_{SR,\mathsf{CT}}$ that is shared by all



$\mathbb{A}$	set $CQ_{SR,\mathsf{CT}}(\boldsymbol{H}) = (0,1)$
$\mathbb{B}$	set $CQ_{SR,\mathsf{CT}}(\boldsymbol{H}) = (1,1)$
$\mathbb{C}$	set $CQ_{SR,\mathsf{CT}}(\boldsymbol{H}) = (1,0)$
$\mathbb{D}$	set $CQ_{SR,\mathsf{CT}}(\boldsymbol{H}) = (1,0)$ or $(0,1)$

Figure 4.2: The flow chart of $CQ_{SR,\mathsf{CT}}$.

terminals will decode the two feedback bits and recover the values of $ENC_{SR,\mathsf{CT},k}(\boldsymbol{h}_k)$ for $k = 1,2$. The flow chart of $CQ_{SR,\mathsf{CT}}$ is shown in Fig. 4.2.

Theorem 4.1. *The quantizer $CQ_{SR,\mathsf{CT}}$ achieves the full-CSI network outage probability using only one bit of feedback per receiver for every channel state.*

Proof. Note that the quantizer $CQ_{SR,\mathsf{CT}}$ operates using only one bit of feedback per receiver per channel state, for a total of two bits per channel state. Moreover, with $CQ_{SR,\mathsf{CT}}$, an outage event will occur only when $SR_{\mathsf{CT}}(p_1,p_2) < 2$ for any $(p_1,p_2) \in \{(1,0),(0,1),(1,1)\}$, or equivalently, when both receivers feed "0" back and the determined power pair $CQ_{SR,\mathsf{CT}}(\boldsymbol{H}) = (1,1)$ still leads to outage. Therefore, the outage probability with the quantizer $CQ_{SR,\mathsf{CT}}$ is the same as the minimum possible outage probability $\mathsf{OUT}^{\text{opt}}_{SR,\mathsf{CT}}$. □

The design of $CQ_{SR,\mathsf{CT}}$ utilizes the fact that checking whether $(p_1,p_2) = (1,0)$ or $(0,1)$ causes outage only requires the knowledge of local CSI at either receiver. We let Receiver k send

one bit to indicate whether $R_{\mathsf{CT},k}(p_k = 1, p_l = 0) = \log_2\left(1 + P\left|h_{kk}\right|^2\right) \geq 2$, thus, two bits from the two receivers provide adequate information for choosing the right pair (p_1, p_2) to attain the best performance.

We now proceed to time-sharing. In this case, the network outage probability for sum-rate is

$$\mathsf{OUT}_{SR,\mathsf{TS}} \triangleq \Pr\left\{ SR_{\mathsf{TS}}\left(t_1, t_2\right) \triangleq \sum_{k=1}^{2} R_{\mathsf{TS},k}(t_k) < 2 \right\}.$$

Under the constraint $t_1 + t_2 = 1$, the maximum sum-rate is

$$\max\left\{SR_{\mathsf{TS}}\left(1,0\right), SR_{\mathsf{TS}}\left(0,1\right)\right\}.$$

Therefore, the full-CSI network outage probability is

$$\mathsf{OUT}^{\mathrm{opt}}_{SR,\mathsf{TS}} = \Pr\left\{SR_{\mathsf{TS}}\left(1,0\right) < 2, SR_{\mathsf{TS}}\left(0,1\right) < 2\right\}.$$

Noticing that $SR_{\mathsf{TS}}\left(1,0\right) = SR_{\mathsf{CT}}\left(1,0\right)$ and $SR_{\mathsf{TS}}\left(0,1\right) = SR_{\mathsf{CT}}\left(0,1\right)$ from (4.1) and (4.2), then, applying the same idea as in the construction of $CQ_{SR,\mathsf{CT}}$, we can design a cooperative quantizer for time-sharing that also achieves $\mathsf{OUT}^{\mathrm{opt}}_{SR,\mathsf{TS}}$ with only one bit of feedback per receiver (the details are omitted). On the other hand, $SR_{\mathsf{TS}}(1,0) = SR_{\mathsf{CT}}(1,0)$ and $SR_{\mathsf{TS}}(0,1) = SR_{\mathsf{CT}}(0,1)$ implies $\mathsf{OUT}^{\mathrm{opt}}_{SR,\mathsf{CT}} \leq \mathsf{OUT}^{\mathrm{opt}}_{SR,\mathsf{TS}}$. Hence, we only need to consider concurrent transmission when we wish to minimize the network outage probability of the sum-rate.

4.4 Cooperative Quantization for Network Outage Probability of Minimum Rate

In this section, we design cooperative quantizers that minimize the outage probability of minimum rate. We first study the minimum possible network outage probabilities with time-sharing and concurrent transmission, respectively.

Given a possibly channel-dependent time-sharing pair (t_1, t_2), the network outage probability of minimum rate can be expressed as

$$\mathsf{OUT}_{MR,\mathsf{TS}} \triangleq \Pr\left\{ MR_{\mathsf{TS}}(t_1,t_2) \triangleq \min\left\{R_{\mathsf{TS},1}(t_1), R_{\mathsf{TS},2}(t_2)\right\} < 1 \right\}.$$

Likewise, for concurrent transmission, given (p_1, p_2), the network outage probability can be expressed as

$$\mathsf{OUT}_{MR,\mathsf{CT}} \triangleq \Pr\left\{ MR_{\mathsf{CT}}(p_1,p_2) \triangleq \min\{R_{\mathsf{CT},1}(p_1,p_2), R_{\mathsf{CT},2}(p_1,p_2)\} < 1 \right\}.$$

In the following two propositions, whose proofs are provided in Appendix C.1, we respectively determine the optimal time-sharing and concurrent transmission pairs that minimize the network outage probability of minimum rate.

Proposition 4.1. *Let* $(t_1^\star, t_2^\star) = \operatorname{argmax}_{(t_1,t_2)} MR_{\mathsf{TS}}(t_1,t_2)$. *We have*

$$(t_1^\star, t_2^\star) = \left(\frac{\log_2\left(1+P\left|h_{22}\right|^2\right)}{\log_2\left(1+P\left|h_{11}\right|^2\right)+\log_2\left(1+P\left|h_{22}\right|^2\right)}, \frac{\log_2\left(1+P\left|h_{11}\right|^2\right)}{\log_2\left(1+P\left|h_{11}\right|^2\right)+\log_2\left(1+P\left|h_{22}\right|^2\right)} \right). \tag{4.3}$$

Proposition 4.2. *Let* $(p_1^\star, p_2^\star) = \operatorname{argmax}_{(p_1,p_2)} MR_{\mathsf{CT}}(p_1,p_2)$. *If* $\frac{|h_{11}|^2}{|h_{21}|^2+\frac{1}{P}} > \frac{|h_{22}|^2}{|h_{12}|^2+\frac{1}{P}}$, *we*

have

$$(p_1^\star, p_2^\star) = \left(\frac{\sqrt{\frac{4P^2|h_{12}|^2|h_{21}|^2|h_{22}|^2+4P|h_{22}|^2|h_{12}|^2}{|h_{11}|^2}+1}-1}{2P\,|h_{12}|^2}, 1 \right), \tag{4.4}$$

and otherwise, if $\frac{|h_{11}|^2}{|h_{21}|^2+\frac{1}{P}} \leq \frac{|h_{22}|^2}{|h_{12}|^2+\frac{1}{P}}$*, we have*

$$(p_1^\star, p_2^\star) = \left(1, \frac{\sqrt{\frac{4P^2|h_{11}|^2|h_{12}|^2|h_{21}|^2+4P|h_{11}|^2|h_{21}|^2}{|h_{22}|^2}+1}-1}{2P\,|h_{21}|^2} \right). \tag{4.5}$$

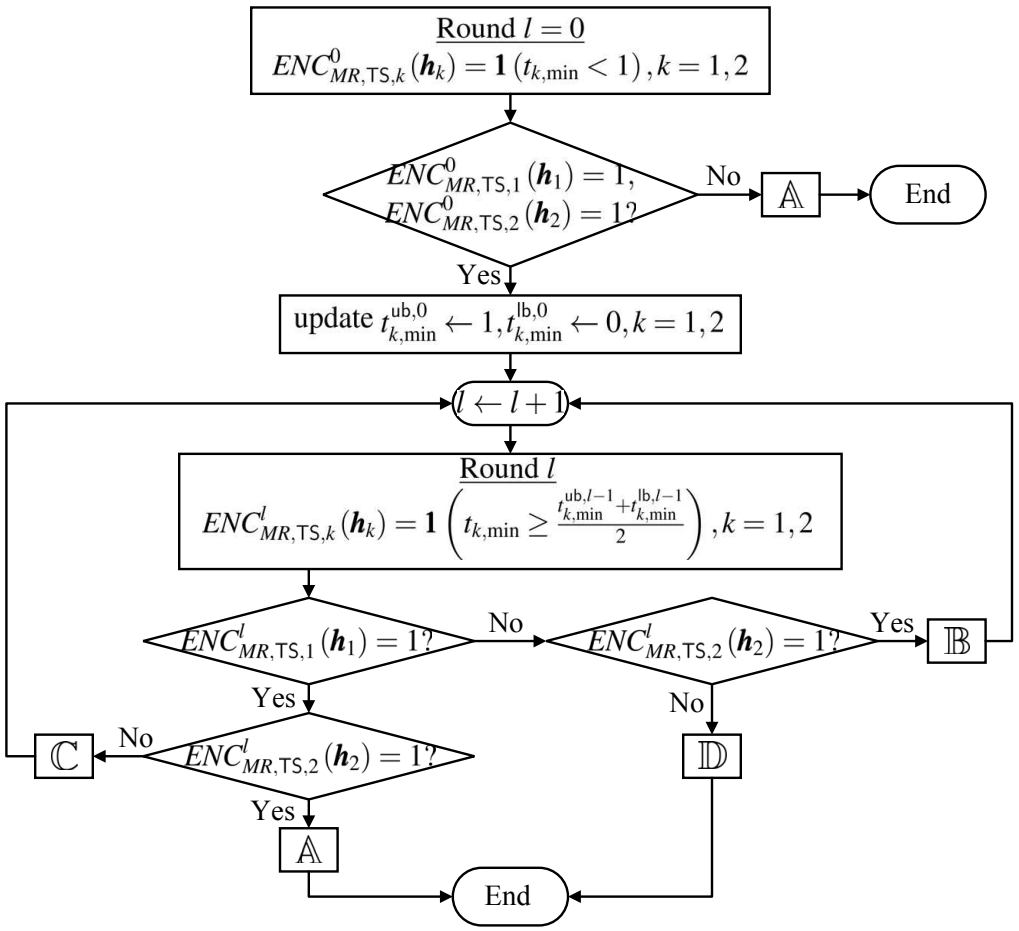
In particular, the full-CSI network outage probabilities of minimum rate for time-sharing and concurrent transmission are given by

$$\mathsf{OUT}^{\text{opt}}_{MR,\mathsf{TS}} = \Pr\left\{MR_{\mathsf{TS}}(t_1^\star, t_2^\star) < 1\right\}, \mathsf{OUT}^{\text{opt}}_{MR,\mathsf{CT}} = \Pr\left\{MR_{\mathsf{CT}}(p_1^\star, p_2^\star) < 1\right\}.$$

We now propose two cooperative quantizers, namely $CQ_{MR,\mathsf{TS}}$ and $CQ_{MR,\mathsf{CT}}$, that can respectively achieve the optimal outage probabilities $\mathsf{OUT}^{\text{opt}}_{MR,\mathsf{TS}}$ and $\mathsf{OUT}^{\text{opt}}_{MR,\mathsf{CT}}$ with finite average feedback rates per receiver.

4.4.1 Time-Sharing

For a given $\boldsymbol{H}$, the minimum time percentage for Receiver k to prevent outage is given by $t_{k,\min} = \frac{1}{\log_2\left(1+P|h_{kk}|^2\right)}$, which can be calculated locally by Receiver k. Denote by $CQ_{MR,\mathsf{TS}}(\boldsymbol{H})$ the time-sharing pair (t_1, t_2) chosen by $CQ_{MR,\mathsf{TS}}$.



$\mathbb{A}$	set $CQ_{MR,\mathsf{TS}}(\boldsymbol{H})=\left(\frac{1}{2},\frac{1}{2}\right)$
$\mathbb{B}$	set $CQ_{MR,\mathsf{TS}}(\boldsymbol{H})=\left(\frac{t^{\mathrm{ub},l-1}_{1,\min}+t^{\mathrm{lb},l-1}_{1,\min}}{2},\frac{t^{\mathrm{ub},l-1}_{2,\min}+t^{\mathrm{lb},l-1}_{2,\min}}{2}\right)$, update $t^{\mathrm{ub},l}_{1,\min}\leftarrow\frac{t^{\mathrm{ub},l-1}_{1,\min}+t^{\mathrm{lb},l-1}_{1,\min}}{2}, t^{\mathrm{lb},l}_{1,\min}\leftarrow t^{\mathrm{lb},l-1}_{1,\min}$, $t^{\mathrm{ub},l}_{2,\min}\leftarrow t^{\mathrm{ub},l-1}_{2,\min}, t^{\mathrm{lb},l}_{2,\min}\leftarrow\frac{t^{\mathrm{ub},l-1}_{2,\min}+t^{\mathrm{lb},l-1}_{2,\min}}{2}$
$\mathbb{C}$	set $CQ_{MR,\mathsf{TS}}(\boldsymbol{H})=\left(\frac{t^{\mathrm{ub},l-1}_{1,\min}+t^{\mathrm{lb},l-1}_{1,\min}}{2},\frac{t^{\mathrm{ub},l-1}_{2,\min}+t^{\mathrm{lb},l-1}_{2,\min}}{2}\right)$, update $t^{\mathrm{ub},l}_{1,\min}\leftarrow t^{\mathrm{ub},l-1}_{1,\min}, t^{\mathrm{lb},l}_{1,\min}\leftarrow\frac{t^{\mathrm{ub},l-1}_{1,\min}+t^{\mathrm{lb},l-1}_{1,\min}}{2}$, $t^{\mathrm{ub},l}_{2,\min}\leftarrow\frac{t^{\mathrm{ub},l-1}_{2,\min}+t^{\mathrm{lb},l-1}_{2,\min}}{2}, t^{\mathrm{lb},l}_{2,\min}\leftarrow t^{\mathrm{lb},l-1}_{2,\min}$
$\mathbb{D}$	set $CQ_{MR,\mathsf{TS}}(\boldsymbol{H})=\left(\frac{t^{\mathrm{ub},l-1}_{1,\min}+t^{\mathrm{lb},l-1}_{1,\min}}{2},\frac{t^{\mathrm{ub},l-1}_{2,\min}+t^{\mathrm{lb},l-1}_{2,\min}}{2}\right)$

Figure 4.3: The flow chart of $CQ_{MR,\mathsf{TS}}$.

The proposed quantizer $CQ_{MR,\mathsf{TS}}$ consists of two local encoders with the k-th encoder $ENC_{MR,\mathsf{TS},k}:\mathtt{R}^{2\times 1}\rightarrow\{0,1\}$ located at Receiver k and a unique decoder $DEC_{MR,\mathsf{TS}}$ function-

ing at all terminals. We add the superscript "l" to indicate their operations at Round l for $l \in \mathtt{N}$. At round l, Receiver k will feed back "1" if $ENC^{l}_{MR,\mathsf{TS},k}\left(\boldsymbol{h}_k\right) = 1$, and otherwise, it will feed back "0". The decoder $DEC^{l}_{MR,\mathsf{TS}}$ decodes the two feedback bits and recovers the values of $ENC^{l}_{MR,\mathsf{TS},k}\left(\boldsymbol{h}_k\right)$ for $k = 1, 2$. Two variables, namely $t^{\mathrm{ub}}_{k,\min}, t^{\mathrm{lb}}_{k,\min}$, will be stored and dynamically updated at all terminals; $t^{\mathrm{ub},l}_{k,\min}, t^{\mathrm{lb},l}_{k,\min}$ will represent the values of $t^{\mathrm{ub}}_{k,\min}, t^{\mathrm{lb}}_{k,\min}$ after Round l. The flow chart of $CQ_{MR,\mathsf{TS}}$ is presented in Fig. 4.3.

Remark 1: To provide some intuitions on the structure of $CQ_{MR,\mathsf{TS}}$, note that the outage condition $MR_{\mathsf{TS}}\left(t_1^\star, t_2^\star\right) < 1$ is the same as $t_{1,\min} + t_{2,\min} > 1$ since $t_{1,\min} + t_{2,\min} = \frac{1}{MR_{\mathsf{TS}}\left(t_1^\star, t_2^\star\right)}$. Equivalently, $CQ_{MR,\mathsf{TS}}$ needs to determine whether $t_{1,\min} + t_{2,\min} > 1$. To accomplish this, Receiver k quantizes $t_{k,\min}$ in a successively refinable way with increasing l. The variables $t^{\mathrm{ub}}_{k,\min}, t^{\mathrm{lb}}_{k,\min}$ serve as the upper and lower bounds on $t_{k,\min}$, which are updated based on the two bits fed back by the two receivers in each round. This inter-receiver conferencing process will end if the two feedback bits in some round satisfy the ending criterion shown in Fig. 4.3. When the ending criterion has been satisfied, the updated $t^{\mathrm{ub}}_{k,\min}, t^{\mathrm{lb}}_{k,\min}$ will be adequate to make a judgement of whether $t_{1,\min} + t_{2,\min} > 1$ holds.

Let $\mathsf{OUT}\left(CQ_{MR,\mathsf{TS}}\right)$ and $\mathsf{FR}\left(CQ_{MR,\mathsf{TS}}\right)$ be the network outage probability and average feedback rate of $CQ_{MR,\mathsf{TS}}$, respectively. The following theorem, whose proof can be found in Appendix C.2, shows that whenever the optimal time shairing pair $(t_1^\star, t_2^\star)$ in Proposition 4.1 can avoid outage, the quantized time-sharing pair chosen by $CQ_{MR,\mathsf{TS}}$ will also avoid outage with probability one and with a finite average feedback rate.

Theorem 4.2.

(a) Let $CQ_{MR,\mathsf{TS}}\left(\boldsymbol{H}\right) = \left(\hat{t}_1, \hat{t}_2\right)$. *Then, we have* $0 < \hat{t}_1, \hat{t}_2 < 1$ *and* $\hat{t}_1 + \hat{t}_2 = 1$.

(b) For any $P > 0$, *we have*

$$\mathsf{OUT}\left(CQ_{MR,\mathsf{TS}}\right) = \mathsf{OUT}^{\mathrm{opt}}_{MR,\mathsf{TS}}, \tag{4.6}$$

$$\mathsf{FR}\left(CQ_{MR,\mathsf{TS}}\right) \leq 2 + 2e^{-\frac{1}{P}}\left(1 + \frac{C_0}{P}\right), \tag{4.7}$$

where C_0 is a bounded positive constant that is independent of P.[3]

Theorem 4.2.(a) states the feasibility of $CQ_{MR,\mathsf{TS}}(\boldsymbol{H})$ as a time-sharing pair; Theorem 4.2.(b) shows the zero-distortion can actually be achieved in network outage probability by a finite average feedback rate. Therefore, there is no need for infinite feedback bits per channel state as the existing methods suggest. This surprising result comes from conferencing between receivers in our design.

Moreover, notice that when $P \to \infty$, $\mathsf{FR}\left(CQ_{MR,\mathsf{TS}}\right) \to 4$. This is because the probability that $t_{k,\min} < \frac{1}{2}$ converges to 1 for large P. Hence, after two rounds (which require a total of 4 bits of feedback), with probability 1, the pair $\left(\frac{1}{2}, \frac{1}{2}\right)$ will be chosen as the time-sharing pair. Similarly, when $P \to 0$, the probability that $t_{k,\min} > 1$ will converge to 1. Therefore, after Round 0 (with 2 bits of feedback), conferencing will end because the outage is inevitable almost surely.

Furthermore, if a maximum number of conferencing rounds is imposed on $CQ_{MR,\mathsf{TS}}$ (i.e., conferencing ends either when the ending criterion in Fig. 4.3 is satisfied or the maximum number of conferencing rounds $L \geq 2$ is reached), we have the following upper bounds on $\mathsf{OUT}\left(CQ_{MR,\mathsf{TS}}\right)$ and $\mathsf{FR}\left(CQ_{MR,\mathsf{TS}}\right)$. We omit their proofs here since they are similar to that of Theorem 4.2.

Corollary 4.1. *When a maximum number of conferencing rounds $L \geq 2$ is imposed on*

[3]Since we focus on showing the average feedback rate is finite for any P, it is beyond the scope of our paper to derive the tightest bound, i.e., the smallest value for C_0. Also, it is worth mentioning that the feedback rate here can be interpreted as the time cost of conferencing. Feeding one bit back by either receiver consumes one time unit, thus, the time cost of conferencing is equivalent to the total number of feedback bits from both receivers.

$CQ_{MR,\mathsf{TS}}$, for any $P > 0$, we have

$$\mathsf{OUT}\left(CQ_{MR,\mathsf{TS}}\right) \leq \mathsf{OUT}^{\text{opt}}_{MR,\mathsf{TS}} + C_{0,1}\frac{e^{-\frac{1}{P}}}{2^{L-2}P}, \tag{4.8}$$

$$\mathsf{FR}\left(CQ_{MR,\mathsf{TS}}\right) \leq 2 + 2e^{-\frac{1}{P}}\left(1 + C_{0,2}\frac{1}{P} + C_{0,3}\frac{L}{2^{L-2}P}\right), \tag{4.9}$$

where $C_{0,1}, C_{0,2}$ and $C_{0,3}$ are bounded positive constants independent of P.

When $P \to \infty$ in (4.8), $\mathsf{OUT}\left(CQ_{MR,\mathsf{TS}}\right) \to \mathsf{OUT}^{\text{opt}}_{MR,\mathsf{TS}}$. This arises from the fact that when $P \to \infty$, the data rate at either receiver will be infinitely large for any chosen time-sharing pair, thus, outage is eliminated almost surely.

4.4.2 Concurrent transmission

We have shown in the proof of Proposition 4.2 in Appendix C.1 that for any power pair (p_1, p_2) with $0 < p_1, p_2 < 1$, the minimum rate can be further increased by scaling p_1 and p_2 simultaneously till either p_1 or p_2 reaches 1. Therefore, an optimal choice of (p_1, p_2) should at least satisfy $p_1 = 1$ or $p_2 = 1$, as shown in Proposition 4.2.

For $k, l = 1, 2$ and $k \neq l$, if we let $p_k = 1$, the maximum allowed power at Transmitter l that will not cause outage to Receiver k, and the minimum required power at Transmitter l so as to avoid outage at Receiver l can be calculated at Receivers k and l, respectively, as

$$\begin{aligned} R_{\mathsf{CT},k}(p_k = 1, p_l) &= \log_2\left(1 + \frac{|h_{kk}|^2}{p_l|h_{lk}|^2 + \frac{1}{P}}\right) \geq 1 \Longrightarrow p_l \leq p_{l,\max} = \frac{|h_{kk}|^2 - \frac{1}{P}}{|h_{lk}|^2}, \\ R_{\mathsf{CT},l}(p_k = 1, p_l) &= \log_2\left(1 + \frac{p_l|h_{ll}|^2}{|h_{kl}|^2 + \frac{1}{P}}\right) \geq 1 \Longrightarrow p_l \geq p_{l,\min} = \frac{|h_{kl}|^2 + \frac{1}{P}}{|h_{ll}|^2}. \end{aligned} \tag{4.10}$$

Thus, when Transmitter k consumes full power, the first task of $CQ_{MR,\mathsf{CT}}$ is to determine

whether $p_{l,\min} \leq p_l \leq p_{l,\max}$ holds for some $p_l \in [0,1]$. Equivalently, $CQ_{MR,\mathsf{CT}}$ should make sure whether $p_{l,\min} \leq 1$ and $p_{l,\min} \leq p_{l,\max}$ are simultaneously satisfied. If so, the second task of $CQ_{MR,\mathsf{CT}}$ is to find a value between $p_{l,\max}$ and $p_{l,\min}$ for p_l. Whereas if $CQ_{MR,\mathsf{CT}}$ declares outage under $p_k = 1$, the other scenario where $p_l = 1$ should be attempted subsequently.

The quantization procedure of $CQ_{MR,\mathsf{CT}}$ is thus composed of two stages. Stage 0 will address the scenario where $p_1 = 1$ to decide whether $p_{2,\min} \leq p_2 \leq p_{2,\max}$ holds for some $p_2 \in [0,1]$, and find such a p_2 if it exists. Similarly, conditioned on $p_2 = 1$, Stage 1 will determine whether $p_{1,\min} \leq p_1 \leq p_{1,\max}$ holds for some $p_1 \in [0,1]$ and find such a p_1 if it exists. In addition, Stage 1 will be initiated only if Stage 0 cannot find an appropriate value of p_2 that avoids outage. If both stages declare outage, an outage event will be unavoidable for $CQ_{MR,\mathsf{CT}}$.

More specifically, $CQ_{MR,\mathsf{CT}}$ comprises of two local encoders and a universal decoder. The k-th encoder $ENC_{MR,\mathsf{CT},k} : \mathbb{R}^{2\times 1} \to \{0,1\}$ is at Receiver k, and the universal decoder $DEC_{MR,\mathsf{CT}}$ is at all terminals. We add the superscript "m,l" to represent their operations at Round l in Stage m for $m = 0,1$ and $l \in \mathbb{N}$. For $k = 1,2$, variables q_k^{ub} and q_k^{lb} will be stored and updated at all terminals, with $q_k^{\mathsf{ub},m,l}$ and $q_k^{\mathsf{lb},m,l}$ being their instantaneous values after Round l in Stage m. The flow chart of $CQ_{MR,\mathsf{CT}}$ is shown in Fig. 4.4. Since Stage 1 is similar to Stage 0, the procedure of Stage 1 is omitted for conciseness. Also, it should be noted that in Stage 1 where $p_2 = 1$, if $\left(ENC_{MR,\mathsf{CT},1}^{1,0}(\boldsymbol{h}_1), ENC_{MR,\mathsf{CT},2}^{1,0}(\boldsymbol{h}_2)\right) \neq (1,1)$, or $ENC_{MR,\mathsf{CT},1}^{1,l}(\boldsymbol{h}_1) = ENC_{MR,\mathsf{CT},2}^{1,l}(\boldsymbol{h}_2) = 1$ for $l \geq 2$, $CQ_{MR,\mathsf{CT}}$ ends and the power pair is $CQ_{MR,\mathsf{CT}}(\boldsymbol{H}) = (1,1)$.

Remark 2: The core of $CQ_{MR,\mathsf{CT}}$ lies in the procedure of Round l in Stage 0 or 1 for $l \geq 2$. In Stage 0, when it comes to Round 2, we have $0 < p_{2,\max} < 1$ and $0 < p_{2,\min} \leq 1$. Instead of determing whether $p_{2,\min} \leq p_{2,\max}$ directly, $CQ_{MR,\mathsf{CT}}$ determines whether $(1-p_{2,\max})+p_{2,\min} \leq 1$ holds, which can be done by treating $1-p_{2,\max}$ and $p_{2,\min}$ in $CQ_{MR,\mathsf{CT}}$ as $t_{1,\min}$ and $t_{2,\min}$ in $CQ_{MR,\mathsf{TS}}$, respectively. The conferencing-based procedure of $CQ_{MR,\mathsf{TS}}$

is thus applied to $CQ_{MR,\mathsf{CT}}$, where the variables $q_1^{\mathsf{lb},m,l}, q_1^{\mathsf{ub},m,l}$ and $q_2^{\mathsf{lb},m,l}, q_2^{\mathsf{ub},m,l}$ represent the lower and upper bounds on $1 - p_{2,\max}$ and $p_{2,\min}$ updated during the conferencing process, respectively.

Let $\mathsf{OUT}\left(CQ_{MR,\mathsf{CT}}\right)$ denote the network outage probability with $CQ_{MR,\mathsf{CT}}$, and let $\mathsf{FR}\left(CQ_{MR,\mathsf{CT}}\right)$ be the average feedback rate of $CQ_{MR,\mathsf{CT}}$. The following theorem shows that $CQ_{MR,\mathsf{CT}}$ achieves the optimal outage probability with full CSI and provides an upper bound on $\mathsf{FR}\left(CQ_{MR,\mathsf{CT}}\right)$. The proof of this theorem is provided in Appendix C.3.

Theorem 4.3.

(a) Let $CQ_{MR,\mathsf{CT}}(\boldsymbol{H}) = (\hat{p}_1, \hat{p}_2)$. *Then, we have* $0 \le \hat{p}_1, \hat{p}_2 \le 1$.

(b) For any $P > 0$, *we have*

$$\mathsf{OUT}\left(CQ_{MR,\mathsf{CT}}\right) = \mathsf{OUT}_{MR,\mathsf{CT}}^{\mathrm{opt}}, \tag{4.11}$$

$$\mathsf{FR}\left(CQ_{MR,\mathsf{CT}}\right) \le 4 + 4e^{-\frac{1}{P}}\left(C_1\eta + 1\right), \tag{4.12}$$

where C_1 *is a bounded positive constant that is independent of* P.

We also consider the situation where the numbers of conferencing rounds in both stages of $CQ_{MR,\mathsf{CT}}$ are limited by $L \ge 3$. With this constraint, in Stage 0, if the ending criterion in Fig. 4.4 is satisfied no later than Round $L-1$, conferencing will end; if conferencing in Stage 0 proceeds to Round $L-1$ while the ending criterion has not been satisfied, Stage 1 will be initiated, then, conferencing ends either when the ending criterion for Stage 1 has been met or conferencing reaches Round $L-1$. The upper bounds on $\mathsf{OUT}\left(CQ_{MR,\mathsf{CT}}\right)$ and $\mathsf{FR}\left(CQ_{MR,\mathsf{CT}}\right)$ for this case are given in Corollary 4.2. The proof of Corollary 4.2 is skipped since it is similar to that of Theorem 4.3.

Corollary 4.2. *When a maximum number of conferencing rounds* $L \ge 3$ *is imposed on*

both stages in $CQ_{MR,\mathsf{CT}}$, for any $P > 0$, we have

$$\mathsf{OUT}\left(CQ_{MR,\mathsf{CT}}\right) \leq \mathsf{OUT}^{\text{opt}}_{MR,\mathsf{CT}} + \frac{\eta e^{-\frac{1}{P}}}{2^{L-3}},$$
$$\mathsf{FR}\left(CQ_{MR,\mathsf{CT}}\right) \leq 4 + 4e^{-\frac{1}{P}}\left(1 + C_{1,0}\eta + C_{1,1}\eta\frac{L}{2^{L-4}}\right),$$

where $C_{1,0}$ and $C_{1,1}$ are bounded positive constants independent of P.

4.4.3 A Joint Time-sharing and concurrent transmission Strategy

We recall from Section 4.3 that for the network outage probability of sum-rate, concurrent transmission is always superior to time-sharing for any given channel state. On the other hand, for the network outage probability of minimum rate, depending on P, either time-sharing or concurrent transmission can be superior for different channel states. In this section, we thus propose a simple joint time-sharing and concurrent transmission strategy that can outperform either time-sharing or concurrent transmission alone.

Suppose for now that the transmitters and the receivers have full CSI. Then, for a given channel state, as a first step, we can apply either time-sharing or concurrent transmission. Only if this first step gives rise to an outage event, we apply the other strategy. It is straightforward to show that such a joint strategy will result in a network outage probability of

$$\mathsf{OUT}^{\text{joint}}_{MR} = \Pr\left\{MR_{\mathsf{TS}}\left(t_1^\star, t_2^\star\right) < 1, MR_{\mathsf{CT}}\left(p_1^\star, p_2^\star\right) < 1\right\}. \tag{4.13}$$

At this stage, it is instructive to compare the outage performance $\mathsf{OUT}^{\text{joint}}_{MR}$ of our joint strategy with the achievable outage probability with *any* scheme that treats interference as

noise. In this context, under the assumption of treating interference signals as noise, the rate region frontiers of two-user interference channels are provided in [27]. According to the results of [27], for a given $\boldsymbol{H}$, the maximum minimum rate of the interference channel can be expressed as[4]

$$MR_{\text{opt}} = \max\left\{MR_{\mathsf{TS}}\left(t_1^\star, t_2^\star\right), MR_{\mathsf{CT}}\left(p_1^\star, p_2^\star\right), MR_1, MR_2\right\}, \tag{4.14}$$

where the expressions of MR_1 and MR_2 are given at the top of the next page. Hence, the full-CSI network outage probability for maximum minimum rate is

$$\mathsf{OUT}_{MR,\text{opt}} = \Pr\left\{MR_{\mathsf{TS}}\left(t_1^\star, t_2^\star\right) < 1, MR_{\mathsf{CT}}\left(p_1^\star, p_2^\star\right) < 1, MR_1 < 1, MR_2 < 1\right\}. \tag{4.15}$$

Comparing (4.13) with (4.15), it is clear that we potentially have $\mathsf{OUT}_{MR,\text{opt}} < \mathsf{OUT}_{MR}^{\text{joint}}$. However, through numerical simulations, we find that the difference between $\mathsf{OUT}_{MR,\text{opt}}$ and $\mathsf{OUT}_{MR}^{\text{joint}}$ is very small as the differentiating outage event (i.e. the event of $MR_1 < 1$ and $MR_2 < 1$) is a very rare event. Define the percentage of optimality for our joint strategy as

$$\Pr\left\{\boldsymbol{H} \in \left\{\boldsymbol{H} : \mathbf{1}\left(MR_{\mathsf{TS}}\left(t_1^\star, t_2^\star\right) < 1, MR_{\mathsf{CT}}\left(p_1^\star, p_2^\star\right) < 1\right) = \mathbf{1}\left(MR_{\text{opt}} < 1\right)\right\}\right\}.$$

In Fig. 4.5, the simulated percentage of optimality shows for every P and η, in more than 99% of channel realizations, our joint strategy has the same outcome in terms of outage vs non-outage as the best possible strategy that induces the maximum minimum rate MR_{opt} in (4.14). In other words, only a negligible gap exists between their network outage probabilities.

[4]The maximum minimum rate is obtained by finding the rate pair (R_1, R_2) on the rate region frontier that satisfies $R_1 = R_2$.

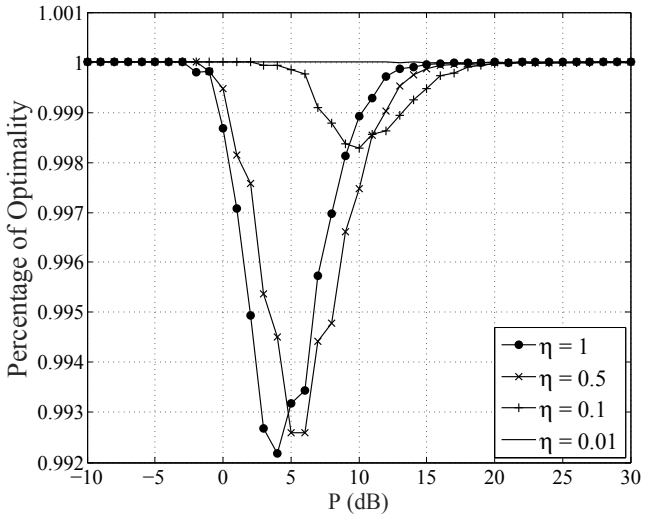


Figure 4.5: Simulated percentage of optimality.

The highly complicated expressions for MR_{opt} makes the design of a quantizer that can achieve $\mathsf{OUT}_{MR,\text{opt}}$ very challenging. However, $\mathsf{OUT}_{MR}^{\text{joint}}$ can be easily achieved in a distributed and quantized manner denoted by CQ_{MR}^{joint}. We can first apply the proposed quantizer $CQ_{MR,\mathsf{TS}}$ for time-sharing. In the second stage, $CQ_{MR,\mathsf{CT}}$ is applied only if $CQ_{MR,\mathsf{TS}}$ cannot avoid outage.[5] In the previous sections, we have shown that $CQ_{MR,\mathsf{TS}}$ and $CQ_{MR,\mathsf{CT}}$ can achieve the full-CSI network outage probabilities with a finite average feedback rate for the time-sharing and concurrent transmission strategies, respectively. Therefore, CQ_{MR}^{joint} will achieve $\mathsf{OUT}_{MR}^{\text{joint}}$ at the expense of only a finite average feedback rate. More precisely, the average feedback rate of CQ_{MR}^{joint} will be no larger than $\mathsf{FR}\left(CQ_{MR,\mathsf{TS}}\right) + \mathsf{FR}\left(CQ_{MR,\mathsf{CT}}\right)$.

[5]For the procedure of $CQ_{MR,\mathsf{TS}}$ in Fig. 4.3, it will be shown later that each terminal will be able to decide whether $CQ_{MR,\mathsf{TS}}$ fails (and thus potentially proceed to the concurrent transmission stages) based on the two feedback bits at each round. For example, when $CQ_{MR,\mathsf{TS}}$ continues into Round l with $l \geq 1$, once both receivers feed back "1", $CQ_{MR,\mathsf{TS}}$ will not be able to avoid an outage event with probability one. On the other hand, if both receivers feed back "0", $CQ_{MR,\mathsf{TS}}$ will avoid outage.

4.5 Numerical Simulations

In this section, we present simulations to verify the theoretical results for $CQ_{MR,\mathsf{TS}}$ in time-sharing and $CQ_{MR,\mathsf{CT}}$ in concurrent transmission. For each instance of P and η, a sufficiently large number of channel realizations are generated so as to observe at least 10^4 outage events.

We will compare the performances of $CQ_{MR,\mathsf{TS}}$ and $CQ_{MR,\mathsf{CT}}$ with those of the following schemes: (i) "Open-Loop 1": either transmitter transmits with full power in half of the transmission block and remains silent in the other half; (ii) "Open-Loop 2": both transmitters use full power in the entire transmission block; (iii) the conventional quantizer denoted by CQ_{MR}^{conv}. In CQ_{MR}^{conv}, given the total number of feedback bits per channel state as $B_{\mathtt{tot}}$, Receiver k spends $\frac{B_{\mathtt{tot}}}{4}$ bits to quantize $|h_{1k}|^2$ or $|h_{2k}|^2$ separately based on a scalar codebook generated by the Lloyd Algorithm [33] with a cardinality of $2^{\frac{B_{\mathtt{tot}}}{4}}$ [23]. Then, all terminals decode the feedback bits to reconstruct the quantized $\boldsymbol{H}$ as $\hat{\boldsymbol{H}}$. The time-sharing and/or concurrent transmission pairs are then calculated according to Propositions 4.1 and 4.2, respectively, by treating $\hat{\boldsymbol{H}}$ as $\boldsymbol{H}$.

In Fig. 4.6(a), we show the network outage probabilities of minimum rate for $CQ_{MR,\mathsf{TS}}$, CQ_{MR}^{conv}, "Open-Loop 1" and MR_{opt} in (4.14). We can observe that the network outage probabilities of CQ_{MR}^{conv} and "Open-Loop 1" are much worse than that of $CQ_{MR,\mathsf{TS}}$. Thus, limited feedback is necessary, and the proposed quantizer based on conferencing is superior to the conventional quantizer. In addition, the curves of $CQ_{MR,\mathsf{TS}}$ and MR_{opt} almost coincide in the medium and high-P regimes, which validates that time-sharing is more favorable when P is moderate or large compared with concurrent transmission. In Fig. 4.6(b), we plot the average feedback rate of $CQ_{MR,\mathsf{TS}}$ to show that it is finite and small in the entire interval of P. In particular, when $P \to \infty$ or 0, we observe that the average feedback rate approaches 4 or 2, respectively, which matches our theoretical analysis in Theorem 4.2.

In Fig. 4.7(a), we show the network outage probabilities for minimum rate of $CQ_{MR,\mathsf{CT}}$,

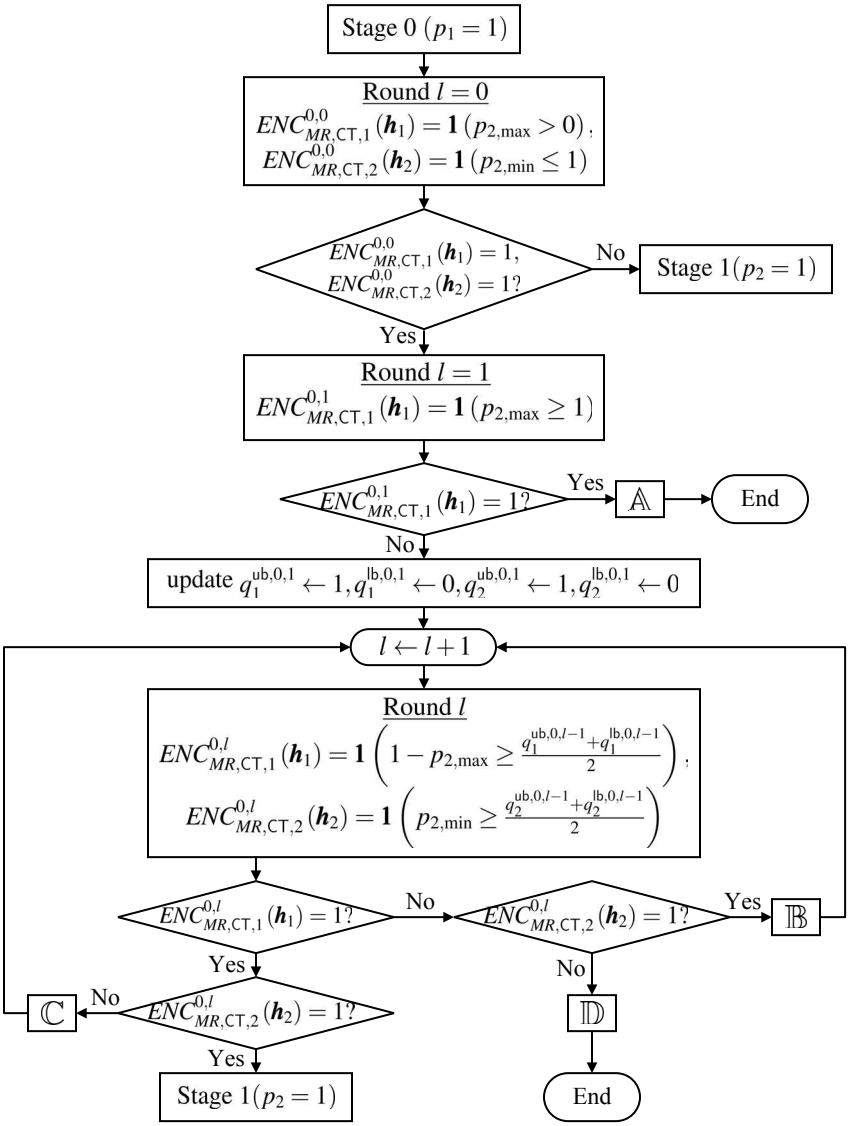
CQ_{MR}^{conv} and "Open-Loop 2". For each $\eta \in \{0.1, 1\}$, the network outage probabilities of $CQ_{MR,\mathsf{CT}}$ are much smaller than those of CQ_{MR}^{conv} and "Open-Loop 2", which verifies that the limited feedback is necessary and the conferencing-based quantizer outperforms the conventional one. In Fig. 4.7(b), we show the average feedback rates of $CQ_{MR,\mathsf{CT}}$, which are finite and small for any P and η, as claimed by Theorem 4.3. For example, as $P \to 0$, the feedback rate converges to 4 bits as can also be inferred from the rate upper bound in Theorem 4.3.

Lastly, we simulate the network outage probability of CQ_{MR}^{joint} in Fig. 4.8 and compare it with those of the maximum minimum rate MR_{opt} in (4.14), CQ_{MR}^{conv} based on MR_{opt}, "Open-Loop 1" and "Open-Loop 2". Additionally, we plot the network outage probabilities of the outer and inner bounds on the information-theoretic capacity region of two-user interference channels given in [29]. It can be observed that the network outage probability of CQ_{MR}^{joint} almost coincides with $\mathsf{OUT}_{MR}^{\text{opt}}$ and greatly exceeds those of the conventional quantizer CQ_{MR}^{conv} and the two schemes with no feedback. Also, as mentioned before, the curve for CQ_{MR}^{joint} is very close to that of the outer bound in [29]. Comparing the performance of CQ_{MR}^{joint} with that of the inner bound, we observe an obvious enhancement. Although we only present the curves for $\eta = 0.1$ here, we have observed that the results for other values of η all exhibit the same behavior.

4.6 Conclusions

We have introduced conferencing-based cooperative quantizers for two-user interference channels where interference signals are treated as additive noise. We have proved that in time-sharing or concurrent transmission, the proposed quantizers are able to achieve the full-CSI network outage probability of sum-rate or the full-CSI network outage probability of minimum rate with finite average feedback rates. Furthermore, a joint strategy of these two quantizers has been proposed to closely approach the network outage probability of maxi-

mum minimum rate with only a negligible gap and finite average feedback rates. Extension of these conferencing-based quantizers to the power-efficient or noisy feedback scenario will be an interesting future research direction.



$\mathbb{A}$	set $CQ_{MR,\mathsf{CT}}(\boldsymbol{H}) = (1, 1)$
$\mathbb{B}$	set $CQ_{MR,\mathsf{CT}}(\boldsymbol{H}) = \left(1, \frac{q_2^{\mathrm{ub},0,l-1} + q_2^{\mathrm{lb},0,l-1}}{2}\right)$, update $q_1^{\mathrm{ub},0,l} \leftarrow \frac{q_1^{\mathrm{ub},0,l-1} + q_1^{\mathrm{lb},0,l-1}}{2}, q_1^{\mathrm{lb},0,l} \leftarrow q_1^{\mathrm{lb},0,l-1}$, $q_2^{\mathrm{ub},0,l} \leftarrow q_2^{\mathrm{ub},0,l-1}, q_2^{\mathrm{lb},0,l} \leftarrow \frac{q_2^{\mathrm{ub},0,l-1} + q_2^{\mathrm{lb},0,l-1}}{2}$
$\mathbb{C}$	set $CQ_{MR,\mathsf{CT}}(\boldsymbol{H}) = \left(1, \frac{q_2^{\mathrm{ub},0,l-1} + q_2^{\mathrm{lb},0,l-1}}{2}\right)$, update $q_1^{\mathrm{ub},0,l} \leftarrow q_1^{\mathrm{ub},0,l-1}, q_1^{\mathrm{lb},0,l} \leftarrow \frac{q_1^{\mathrm{ub},0,l-1} + q_1^{\mathrm{lb},0,l-1}}{2}$, $q_2^{\mathrm{ub},0,l} \leftarrow \frac{q_2^{\mathrm{ub},0,l-1} + q_2^{\mathrm{lb},0,l-1}}{2}, q_2^{\mathrm{lb},0,l} \leftarrow q_2^{\mathrm{lb},0,l-1}$
$\mathbb{D}$	set $CQ_{MR,\mathsf{CT}}(\boldsymbol{H}) = \left(1, \frac{q_2^{\mathrm{ub},0,l-1} + q_2^{\mathrm{lb},0,l-1}}{2}\right)$

Figure 4.4: The flow chart of $CQ_{MR,\mathsf{CT}}$ (Stage 1 with $p_2 = 1$ is similar to Stage 0 with $p_1 = 1$, thus omitted).

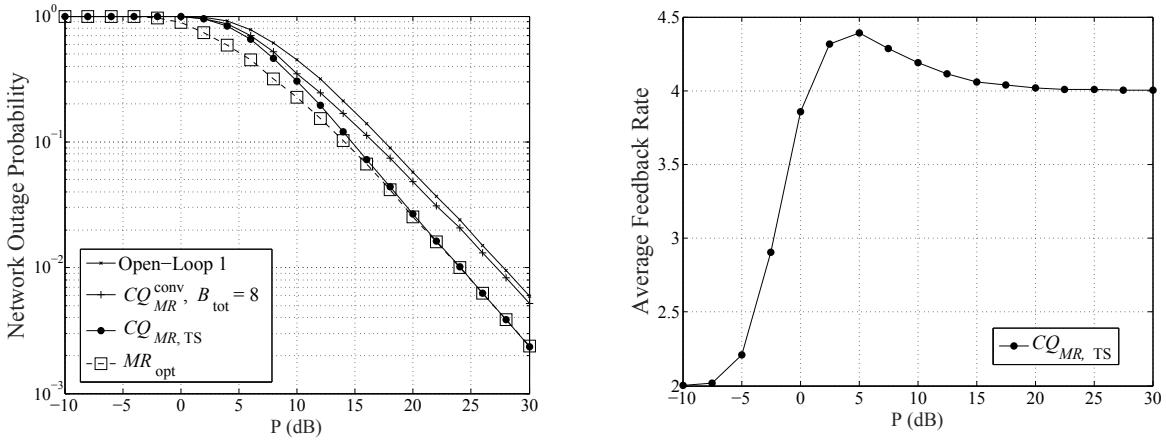


Figure 4.6: Simulated network outage probabilities of minimum rate and average feedback rate for time-sharing.

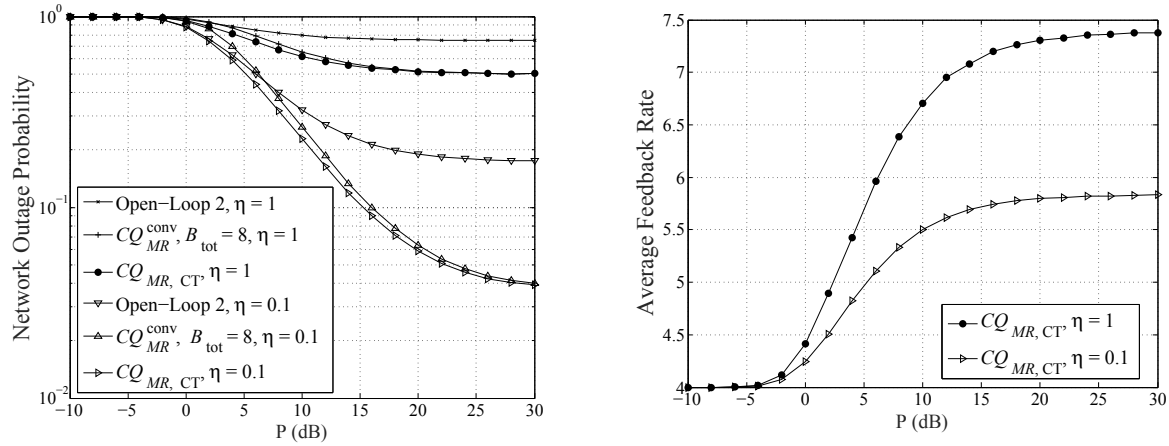


Figure 4.7: Simulated network outage probabilities of minimum rate and average feedback rates for concurrent transmission.

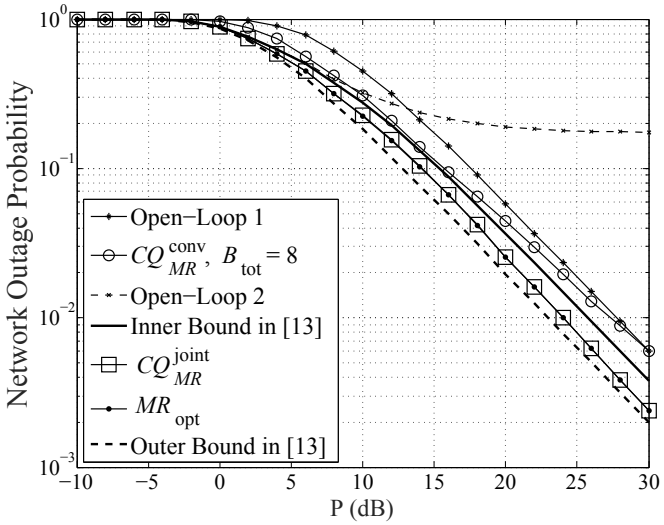


Figure 4.8: Comparison of simulated network outage probabilities of minimum rate for CQ_{MR}^{joint} when $\eta = 0.1$.

Chapter 5

Downlink Non-Orthogonal Multiple Access with Limited Feedback

In this chapter, we analyze downlink non-orthogonal multiple access (NOMA) networks with limited feedback. Our goal is to derive appropriate transmission rates for rate adaptation and minimize outage probability of minimum rate for the constant-rate data service, based on distributed channel feedback information from receivers. We propose an efficient quantizer with variable-length encoding that approaches the best performance of the case where perfect channel state information is available everywhere. We prove that in the typical application with two receivers, the losses in the minimum rate and outage probability decay at least exponentially with the minimum feedback rate. We analyze the diversity gain and provide a sufficient condition for the quantizer to achieve the maximum diversity order. For NOMA with K receivers where $K > 2$, we solve the minimum rate maximization problem within an accuracy of ϵ in time complexity of $O\left(K \log \frac{1}{\epsilon}\right)$, then, we apply the previously proposed quantizers for $K = 2$ to the case of $K > 2$. Numerical simulations are presented to demonstrate the efficiency of our proposed quantizers and the accuracy of the analytical results.

5.1 Introduction

Non-orthogonal multiple access (NOMA) has received significant attention recently for its superior spectral efficiency [34]. It is a promising candidate for mobile communication networks, and has been included in LTE Release 13 for the scenario of two-user downlink transmission under the name of multi-user superposition transmission [35]. The key idea of NOMA is to multiplex multiple users with superposition coding at different power levels, and utilize successive interference cancellation (SIC) at receivers with better channel conditions [36]. Specifically, for NOMA with two receivers, the messages to be sent are superposed with different power allocation coefficients at the BS side. At the receivers' side, the weaker receiver decodes its intended message by treating the other's as noise, while the stronger receiver first decodes the message of the weaker receiver, and then decodes its own by removing the other message from the received signal. In this way, the weaker receiver benefits from larger power, and the stronger receiver is able to decode its own message with no interference. Hence, the overall performance of NOMA is enhanced, compared with traditional orthogonal multiple access schemes. It is shown in [37] that the rate region of NOMA is the same as the capacity region of Gaussian broadcast channels with two receivers, but with an additional constraint that the stronger receiver is assigned less power than the weaker one.

There has been a lot of work on NOMA. In [34] and [37], the authors evaluated the benefits of downlink NOMA from the system and information theoretic perspectives, respectively. The performance of NOMA with randomly deployed users was investigated in [38]. A lot of effort has been put into the power allocation design in NOMA. For example, the authors in [39] and [40] analyzed the necessary conditions for NOMA with two users to beat the performance of time-division-multiple-access (TDMA), and derived closed-form expressions for the expected data rates and outage probabilities. In [41], power allocation based on proportional fairness scheduling was investigated for downlink NOMA. Transmit power minimization subject to rate constraints was discussed in [42]. A joint consideration of dynamic user clustering and

power allocation was studied in [43].

However, all the mentioned papers have assumed a perfect knowledge of the distributed channel state information (CSI) at the BS and all the geographically-distributed receivers, which is difficult to realize in practice. Therefore, we consider the limited feedback scenario wherein each receiver only has access to its own local CSI, from the BS to itself, and then broadcasts its feedback information to the BS and other receivers [44]. Under such settings, interesting problems arise, for example: How to design simple but efficient quantizers for NOMA? What are the performance losses compared with the full-CSI case? A user-selection scheme based on limited feedback was studied in [45]. In [46], the authors derived the outage probability of NOMA based on one-bit feedback of channel quality from each receiver, and performed power allocation to minimize the outage probability. Additionally, the problems of transmit power minimization and user fairness maximization based on statistical CSI subject to outage constraints were studied in [47]. In [48], the authors derived the outage probability and sum rate with fixed power allocation by assuming imperfect and statistical CSI. In [49], the authors solved the sum rate maximization problem for downlink NOMA networks using a minorization-maximization algorithm in statistics. In [50], several antenna selection schemes were proposed for the NOMA systems, and the user fairness was evaluated using the Jain's fairness index.

In this chapter, we focus on the limited feedback design for the typical scenario of downlink NOMA, where a BS communicates with two receivers simultaneously [35]. Based on distributed feedback and in the interest of user fairness, we wish to have the minimum rate of the receivers be as large as possible. Like [51], we also use the minimum achieved rate of all receivers as the performance measure, but moreover, the main focus of our work is to design efficient quantizers for downlink NOMA and analyze the achieved performance. With this goal, to dynamically adjust the transmission rates for better channel utilization, we propose a uniform quantizer which assigns each value to its left boundary point and employs variable-

length encoding (VLE). Then, power allocation is calculated based on the channel feedback. We calculate the transmission rates that can be supported by the current channel states, and analyze the rate loss compared with the full-CSI scenario. The derived upper bound on rate loss shows that it decreases at least exponentially with the minimum of the feedback rates. The feedback rate in this chapter refers to the number of feedback bits each receiver sends for each channel state. where the target data rate needs to be supported and outage probability is the main concern, we conversely propose a uniform quantizer which quantizes each value to its right boundary point.[1] Through the developed upper bound, we show the outage probability loss also decays at least exponentially with the minimum of feedback rate. Additionally, we analyze the achieved diversity gain and provide a sufficient condition on the proposed quantizer in order to achieve the full-CSI diversity order. For the general scenario with K receivers, we solve the minimum rate maximization problem within an accuracy of ϵ in time complexity of $O\left(K \log \frac{1}{\epsilon}\right)$, and apply the previously proposed quantizers for the two-user case here by treating the quantized channels as the perfect ones. We perform Monte Carlo numerical simulations to verify the superiority of our proposed quantizers and the accuracy of the theoretical analysis.

The primary goal of this chapter is to study the impacts of quantization on the performance of NOMA, and provide meaningful insights for practical limited feedback design. To summarize, the main contributions of this chapter are three-fold:

(1) We propose efficient quantizers to maximize the minimum rate in NOMA. The ideas of our proposed quantizers and VLE as well as the designs for rate adaptation and outage probability based on distributed feedback can be generalized to many other scenarios, e.g., NOMA with other performance measures, the more general interference channels, and so on.

[1]For example, in some real-time multimedia service applications, the minimum data rate needs to be supported as often as possible, such that the chance of service outage can be greatly reduced.

(2) Our theoretical analysis serves as a general framework to analyze the performances of such quantizers in NOMA and other scenarios. For instance, it can be easily applied to study the performances of other power allocation schemes in NOMA based on limited feedback, i.e., [39, 40].

(3) We solve the minimum rate maximization problem for any number of receivers with linear time complexity.

Notations: The sets of real and natural numbers are represented by $\mathcal{R}$ and $\mathcal{N}$, respectively. For any $x \in \mathcal{R}$, $\lfloor x \rfloor$ is the largest integer that is less than or equal to x, and $\lceil x \rceil$ is the smallest integer that is larger than or equal to x. $\Pr\{\cdot\}$ and $\mathtt{E}[\cdot]$ represent the probability and expectation, respectively. For a random variable (r.v.) X, $f_X(\cdot)$ is its probability density function (p.d.f.). $\mathbb{CN}(\mu, \lambda)$ represents a circularly symmetric complex Gaussian r.v. with mean μ and variance λ. For a logical statement $\mathtt{ST}$, we let $\mathbf{1}_{\mathtt{ST}} = 1$ when $\mathtt{ST}$ is true, and $\mathbf{1}_{\mathtt{ST}} = 0$ otherwise. Finally, the expression $X \sim_Y Z$ means $0 < \lim_{Y\to\infty} \frac{X}{Z} < \infty$.

5.2 Problem Formulation

5.2.1 System Model

Consider the downlink transmission in Fig. 5.1, where a BS is to transmit a superposition of two symbols to two receivers over the same resource block. Both BS and receivers are equipped with only a single antenna. According to the multiuser superposition transmission scheme [35], the transmitted signal is formed as

$$x = \sqrt{P_1}s_1 + \sqrt{P_2}s_2,$$

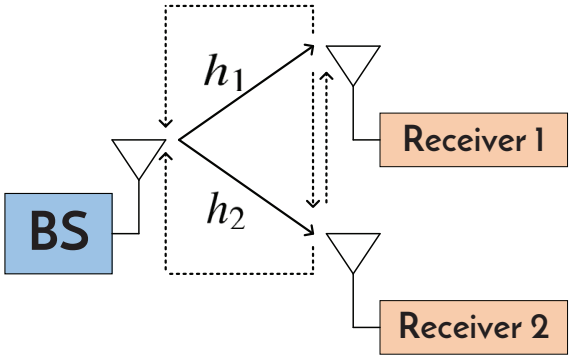


Figure 5.1: Downlink NOMA networks. The solid and dashed lines represent the signal and feedback links, respectively.

where s_i is the information bearing symbol for Receiver i with $\mathtt{E}[s_i] = 0$ and $\mathtt{E}\left[|s_i|^2\right] = 1$ for each channel state (the expectation is over all transmitted symbols); P_i is the average transmit power associated with s_i. Let $P = P_1 + P_2$ be the total transmit power, and $\alpha = \frac{P_1}{P}$ be the power allocation coefficient, then, $P_1 = \alpha P$ and $P_2 = (1-\alpha)P$ with $0 \leq \alpha \leq 1$.

Denote by $h_i \sim \mathbb{CN}(0, \lambda_i)$ the channel coefficient from the BS to Receiver i. Without loss of generality, assume $\lambda_1 \geq \lambda_2$. The received signals at Receivers 1 and 2 are respectively given by

$$y_1 = h_1\sqrt{P_1}s_1 + h_1\sqrt{P_2}s_2 + n_1, \quad y_2 = h_2\sqrt{P_1}s_1 + h_2\sqrt{P_2}s_2 + n_2,$$

where $n_i \sim \mathbb{CN}(0,1)$ represents the background noise. Let $H_i = |h_i|^2$, then, the p.d.f. of H_i is $f_{H_i}(x) = \frac{e^{-\frac{x}{\lambda_i}}}{\lambda_i}$ for $x > 0$.[2] We assume a quasi-static channel model, in which the channels vary independently from one block to another, while remaining constant within each block. Either receiver is assumed to perfectly estimate its local CSI (i.e., H_i), and send the associated quantized local CSI to the other receiver and the BS in a broadcast manner via error-free and delay-free feedback links [4, 30]. In some scenario where the two receivers are far away from each other such that they cannot "talk" directly, the BS can play the role

[2]The results in this chapter can be trivially generalized to other distributions of H_1 and H_2.

of relaying, i.e., forwarding the feedback information received from one receiver to the other.

With SIC, the stronger receiver with better channel condition (i.e., larger H_i) first decodes the message for the weaker receiver, and then decodes its own after removing the message of the weaker one from its received signal; the weaker receiver with poorer channel condition directly decodes its own message by treating the message of the stronger one as noise [41, 52]. Specifically, when $H_1 \geq H_2$, the rate for Receiver 2 (i.e., the weaker one) to decode s_2 by treating s_1 as noise is

$$r_2(\alpha) = \log_2 \left(1 + \frac{PH_2(1-\alpha)}{\alpha H_2 P + 1} \right),$$

which is not larger than the rate for Receiver 1 to decode s_2, given as

$$r_{1\to 2} = \log_2 \left(1 + \frac{PH_1(1-\alpha)}{\alpha H_1 P + 1} \right).$$

If s_2 is transmitted at the rate of $r_2(\alpha)$, Receiver 1 can decode s_2 successfully with an arbitrarily small probability of error [53]. Afterwards, Receiver 1 can remove $h_1\sqrt{P_2}s_2$ from y_1, and achieve a data rate for s_1 as

$$r_1(\alpha) = \log_2 \left(1 + \alpha P H_1 \right).$$

On the other hand, when $H_1 < H_2$, Receiver 2 first decodes s_1, removes $h_2\sqrt{P_1}s_1$ from y_2, and then decodes s_2, while Receiver 1 decodes s_1 directly by treating s_2 as noise.

5.2.2 Maximum Minimum Rate

Our goal is to maximize the minimum of $r_1(\alpha)$ and $r_2(\alpha)$ to ensure fairness between receivers [44, 54]. When perfect CSI is available at the BS and receivers, the optimal power allocation

coefficient $\alpha^\star$ can be found by solving the optimization problem $r_{\max} = \max\limits_{0\le\alpha\le1} \min\{r_1(\alpha), r_2(\alpha)\}$, the solution of which is given in the following theorem.

Theorem 5.1. *When $H_1 \ge H_2$, the solution of $\max\limits_{0\le\alpha\le1} \min\{r_1(\alpha), r_2(\alpha)\}$ is given by*

$$\alpha^\star = \frac{2H_2}{\sqrt{(H_1+H_2)^2 + 4H_1H_2^2P} + (H_1+H_2)}. \tag{5.1}$$

Proof. Notice that with α increasing from 0 to 1, $r_1(\alpha)$ increases from 0 to $\log_2(1+PH_1)$ and $r_2(\alpha)$ decreases from $\log_2(1+PH_2)$ to 0. Since $\log_2(1+PH_1) \ge \log_2(1+PH_2)$, the maximum minimum rate is reached when $r_1(\alpha^\star) = r_2(\alpha^\star)$, from which $\alpha^\star$ in (5.1) is derived. □

The expression of $\alpha^\star$ when $H_1 < H_2$ can be obtained straightforwardly. It is found from (5.1) that: (i) Both messages attain the same rate at optimality, i.e., $r_1(\alpha^\star) = r_2(\alpha^\star) = r_{\max}$. Moreover, it can be verified that the rate pair $(r_1(\alpha^\star), r_2(\alpha^\star))$ is on the rate region boundaries of both NOMA and Gaussian broadcast channels with two receivers [37]. (ii) When $P \to 0$, $\alpha^\star \to \frac{H_2}{H_1+H_2}$, in which case the power assigned to the stronger receiver is in proportion to the channel quality of the weaker one; when $P \to \infty$, $\alpha^\star \to 0$, then, BS should allocate almost all the power to the weaker one. [3] (iii) $\alpha^\star \ge \frac{1}{2}$. Generally, NOMA steers more power towards the weaker receiver to balance their transmissions.

With perfect CSI, the decoding order is determined based on whether $H_1 \ge H_2$ holds. The maximum minimum rate is

$$r_{\max} = \begin{cases} \log_2\left(1 + \frac{2H_1H_2P}{\sqrt{(H_1+H_2)^2+4H_1H_2^2P}+(H_1+H_2)}\right), & H_1 \ge H_2, \\ \log_2\left(1 + \frac{2H_1H_2P}{\sqrt{(H_1+H_2)^2+4H_1^2H_2P}+(H_1+H_2)}\right), & H_1 < H_2, \end{cases} \tag{5.2}$$

[3] Note that $r_1(\alpha^\star) = r_2(\alpha^\star)$ holds for any P. When $P \to \infty$, $\alpha^\star \to 0$, and $r_1(\alpha^\star) = r_2(\alpha^\star) = \log_2\left(1 + \frac{2PH_1H_2}{\sqrt{(H_1+H_2)^2+4H_1H_2^2P}+(H_1+H_2)}\right)$ will approach infinity.

and the outage probability of minimum rate is

$$\mathtt{out}_{\min} = \Pr\left\{r_{\max} < r_{\text{th}}\right\}, \tag{5.3}$$

where r_{th} is the data rate at which the BS will transmit s_1 and s_2 for every channel state.

5.2.3 Limited Feedback

In the limited-feedback scenario, for an arbitrary quantizer $q : \mathcal{R} \to \mathcal{R}$, Receiver i maps H_i to $q(H_i)$, and feeds the index of $q(H_i)$ back to the BS and the other receiver, as shown in Fig. 5.1. The index of $q(H_i)$ is decoded and the value of $q(H_i)$ is recovered. The decoding order will be contingent on whether $q(H_1) \geq q(H_2)$. For instance, when $q(H_1) \geq q(H_2)$, Receiver 1 is considered "stronger", while Receiver 2 is "weaker". In this case, the power allocation coefficient is computed based on (5.1) by treating $q(H_i)$ as H_i, i.e., $\alpha_q = \frac{2q(H_2)}{\sqrt{(q(H_1)+q(H_2))^2+4q(H_1)q^2(H_2)P}+q(H_1)+q(H_2)}$.

For rate adaptation, we shall design appropriate rates $r_{1,q}$ and $r_{2,q}$ for the messages s_1 and s_2 based on limited feedback from the two receivers, such that $r_{1,q}$ and $r_{2,q}$ can be supported and NOMA can be performed. The corresponding rate loss will be $r_{\text{loss}} = r_{\max} - \min\left\{r_{1,q}, r_{2,q}\right\}$, where $r_{\max}$ is given in (5.2).

For a constant-rate service, we care more about whether the current channels are strong enough to support target data rate with the power allocation coefficient computed based on limited feedback. The achieved outage probability is $\mathtt{out}_q = \Pr\left\{r_q < r_{\text{th}}\right\}$, where

$$\begin{aligned} r_q &= \min\left\{r_1(\alpha_q), r_2(\alpha_q)\right\} \\ &= \begin{cases} \min\left\{\log_2\left(1 + P \times \alpha_q \times H_1\right), \log_2\left(1 + \frac{PH_2(1-\alpha_q)}{PH_2\alpha_q+1}\right)\right\}, & q(H_1) \geq q(H_2), \\ \min\left\{\log_2\left(1 + \frac{PH_1(1-\alpha_q)}{PH_1\alpha_q+1}\right), \log_2\left(1 + P \times \alpha_q \times H_2\right)\right\}, & q(H_1) < q(H_2), \end{cases} \end{aligned}$$

The outage probability loss is given as

$$\mathtt{out}_{\mathrm{loss},q} = \mathtt{out}_q - \mathtt{out}_{\min}, \tag{5.4}$$

where $\mathtt{out}_{\min}$ is given in (5.3). In the subsequent sections, we will propose efficient quantizers and investigate the performance losses brought by limited feedback.

5.3 Limited Feedback for Minimum Rate

In this section, we first describe the proposed quantizer when the minimum rate is the concern, then, we show the relationship between the rate loss and the feedback rates.

5.3.1 Proposed Quantizer

We consider a uniform quantizer $q_r : \mathcal{R} \to \mathcal{R}$, given by[4]

$$q_r(x) = \begin{cases} \left\lfloor \frac{x}{\Delta} \right\rfloor \times \Delta, & x \leq T\Delta, \\ T\Delta, & x > T\Delta, \end{cases}$$

where x can be any non-negative real number, and the bin size Δ and the maximum number of bins $T \in \mathcal{N}$ are adjustable parameters. As shown in Fig. 5.2, $q_r(x)$ quantizes x to the left boundary of the interval where x is. For any $x \in [n\Delta, (n+1)\Delta)$ when $0 \leq n \leq T-1$, we have $q_r(x) = n\Delta$ and $x - \Delta \leq q_r(x) \leq x$; for any $x \in [T\Delta, \infty)$, $q_r(x) = T\Delta$ and $q_r(x) \leq x$.

[4]In q_r, "q" stands for quantizer, and the subscript "r" represents rate.

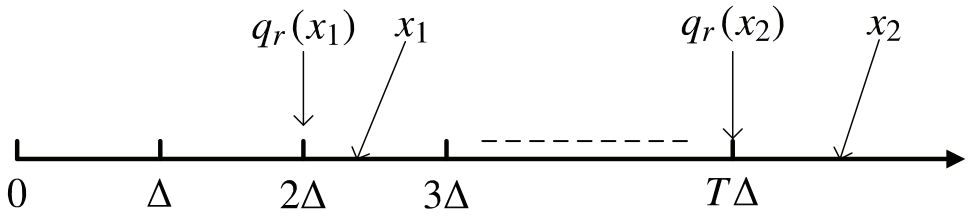


Figure 5.2: A uniform quantizer for minimum rate.

5.3.2 Rate Adaptation and Loss

When $q_r(\cdot)$ is employed, Receiver 2 is viewed as the "weak" receiver if $q_r(H_1) \geq q_r(H_2)$. Then, according to (5.1), the power allocation coefficient α_{q_r} is calculated as

$$\alpha_{q_r} = \begin{cases} \frac{2q_r(H_2)}{\sqrt{[q_r(H_1)+q_r(H_2)]^2+4q_r(H_1)q_r^2(H_2)P}+[q_r(H_1)+q_r(H_2)]}, & q_r(H_1) > 0, q_r(H_2) > 0, \\ 0, & q_r(H_1) = 0 \text{ or } q_r(H_2) = 0, \end{cases}$$

which satisfies $\log_2(1 + P \times \alpha_{q_r} \times q_r(H_1)) = \log_2\left(1 + \frac{q_r(H_2)\times(1-\alpha_{q_r})}{\alpha_{q_r}\times q_r(H_2)+\frac{1}{P}}\right)$ when $\alpha_{q_r} \neq 0$. To exploit the channels as much as possible, we let the BS send messages s_1 and s_2 at rates of

$$r_{1,q_r} = \log_2(1 + P \times \alpha_{q_r} \times q_r(H_1)), r_{2,q_r} = \log_2\left(1 + \frac{P \times q_r(H_2)(1-\alpha_{q_r})}{P \times q_r(H_2)\alpha_{q_r} + 1}\right). \quad (5.5)$$

Lemma 5.1. *When $q_r(H_1) \geq q_r(H_2)$, the rates r_{1,q_r} and r_{2,q_r} in (5.5) can be achieved.*

Proof. Based on the channel coding theorem [53], if we can show the channel capacities for s_1 and s_2 under the settings of NOMA are no smaller than r_{1,q_r} and r_{2,q_r}, the rates r_{1,q_r} and r_{2,q_r} can be achieved with a probability of error that can be made arbitrarily small.

When $q_r(H_1) = 0$ or $q_r(H_2) = 0$, it is trivial to verify that r_{1,q_r} and r_{2,q_r} can be supported. When $q_r(H_1) \geq q_r(H_2) > 0$, the channel capacity for Receiver 2 by treating s_1 as noise is $r_2 = \log_2\left(1 + \frac{H_2(1-\alpha_{q_r})}{\alpha_{q_r}\times H_2+\frac{1}{P}}\right) \geq \log_2\left(1 + \frac{q_r(H_2)\times(1-\alpha_{q_r})}{\alpha_{q_r}\times q_r(H_2)+\frac{1}{P}}\right) = r_{2,q_r}$, since $\log_2\left(1 + \frac{x(1-\alpha)}{x\alpha+\frac{1}{P}}\right)$ is an increasing function of x and $q_r(H_2) \leq H_2$. At the side of Receiver 1, the channel capacity of s_2 with treating s_1 as noise is $r_{1\to2} = \log_2\left(1 + \frac{H_1(1-\alpha_{q_r})}{\alpha_{q_r}\times H_1+\frac{1}{P}}\right) \geq \log_2\left(1 + \frac{q_r(H_1)\times(1-\alpha_{q_r})}{\alpha_{q_r}\times q_r(H_1)+\frac{1}{P}}\right) \geq$

$\log_2\left(1+\frac{q_r(H_2)\times(1-\alpha_{q_r})}{\alpha_{q_r}\times q_r(H_2)+\frac{1}{P}}\right)=r_{2,q_r}$, because $H_1 \geq q_r(H_1) \geq q_r(H_2)$. Hence, s_2 can be decoded at Receiver 1 with an arbitrarily small error and removed from y_1. After that, the channel capacity of s_1 is $r_1 = \log_2(1+P\times\alpha_{q_r}\times H_1) \geq \log_2(1+P\times\alpha_{q_r}\times q_r(H_1)) = r_{1,q_r}$. Therefore, the rates r_{1,q_r} and r_{2,q_r} can be achieved for both s_1 and s_2. □

To sum up, it is the key fact of $q_r(x) \geq x$ that ensures the rates r_{1,q_r} and r_{2,q_r} in (5.5) can be supported. When $q_r(H_1) \geq q_r(H_2)$, the rate loss is $r_{\text{loss}} = r_{\max} - \min\{r_{1,q_r}, r_{2,q_r}\}$.

Lemma 5.2. *The average rate loss of the quantizer $q_r(\cdot)$ is upper-bounded by:*

$$\mathtt{E}[r_{\text{loss}}] \leq \log_2\left(1+C_0\times P\times\max\left\{e^{-\frac{T\Delta}{\lambda_1}},\Delta\right\}\right), \tag{5.6}$$

where C_0 is a positive constant that is independent of P, T and Δ.

Proof. See Appendix D.1. □

We mainly focus on showing how the average rate loss changes with the bin size Δ. It is beyond the scope of this chapter to find the tightest bounds, i.e., the smallest value for C_0. A value for C_0 which is derived from the proof in Appendix A is $C_0 = \max\left\{4+\frac{\lambda_1}{\lambda_2},\lambda_2\right\}$.

It is observed from (5.6) that when $e^{-\frac{T\Delta}{\lambda_1}} > \Delta$, the maximum number of bins, T, can degrade the rate. To eliminate this effect, we choose T such that $e^{-\frac{T\Delta}{\lambda_1}} = \Delta$, which yields $T = \frac{\lambda_1}{\Delta}\log\frac{1}{\Delta}$.[5] With an appropriate value for T, we can make the rate loss decrease at least linearly with Δ.

Corollary 5.1. *When $T = \frac{\lambda_1}{\Delta}\log\frac{1}{\Delta}$, the average rate loss of the quantizer $q_r(\cdot)$ is upper-bounded by:*

$$\mathtt{E}[r_{\text{loss}}] \leq \log_2(1+C_0\times P\times\Delta) \leq C_1\times P\times\Delta, \tag{5.7}$$

[5]Approaching the performance in the full-CSI case generally requires a small value for Δ. We mainly consider the case where $\Delta \leq 1$ in this chapter.

where C_0 and C_1 are positive constants that are independent of P and Δ.

5.3.3 Feedback Rate

Rather than the naive fixed-length encoding (FLE) for feedback information which requires $\lceil \log_2(T+1) \rceil$ bits per receiver per channel state, we consider the more efficient variable-length encoding (VLE) [3, 30].[6] An example of VLE that can be applied here is $b_0 = \{0\}$, $b_1 = \{1\}$, $b_2 = \{00\}$, $b_3 = \{01\}$ and so on, sequentially for all codewords in the set $\{0, 1, 00, 01, 10, 11, \ldots\}$, where b_n is the binary string to be fed back when $q_r(x) = n\Delta$. The length of b_n is $\lfloor \log_2(n+2) \rfloor$. The following theorem derives an upper bound on the rate loss with respect to the feedback rate of Receiver i (denoted by $R_{r,\mathrm{VLE},i}$).

Theorem 5.2. *When variable-length encoding is applied to the quantizer $q_r(\cdot)$, the rate loss decays at least exponentially as:*

$$\begin{aligned} \mathtt{E}\left[r_{\mathrm{loss}}\right] &\leq \log_2\left(1 + C_2 \times P \times 2^{-\min\left\{R_{r,\mathrm{VLE},1}, R_{r,\mathrm{VLE},2}\right\}}\right) \\ &\leq C_3 \times P \times 2^{-\min\left\{R_{r,\mathrm{VLE},1}, R_{r,\mathrm{VLE},2}\right\}}, \end{aligned} \tag{5.8}$$

where C_2 and C_3 are positive constants independent of P and $R_{r,\mathrm{VLE},i}$.

Proof. The feedback rate of Receiver i is derived as

$$\begin{aligned} R_{r,\mathrm{VLE},i} &= \sum_{n=0}^{T-1} \lfloor \log_2(n+2) \rfloor \int_{n\Delta}^{(n+1)\Delta} f_{H_i}(H_i) \mathrm{d}H_i + \lfloor \log_2(T+2) \rfloor \int_{T\Delta}^{\infty} f_{H_i}(H_i) \mathrm{d}H_i \\ &\leq \sum_{n=0}^{\infty} \lfloor \log_2(n+2) \rfloor \int_{n\Delta}^{(n+1)\Delta} f_{H_i}(H_i) \mathrm{d}H_i \leq \sum_{n=0}^{\infty} \underbrace{\log_2(n+2)}_{\leq \log_2(n+1)+1} \int_{n\Delta}^{(n+1)\Delta} \frac{e^{-\frac{H_i}{\lambda_i}}}{\lambda_i} \mathrm{d}H_i \end{aligned}$$

[6] For example, when $\Delta = 0.01$ and $\lambda_1 = 1$, $T = \frac{\lambda_1}{\Delta} \log \frac{1}{\Delta} \approx 460.5$. When FLE is adopted, the feedback rate per receiver will be $\lceil \log_2(T+1) \rceil = 9$ bits per channel state. As shown by the theoretical analysis and numerical simulations later, VLE will cost far fewer bits.

$$\leq \sum_{n=0}^{\infty} e^{-\frac{n\Delta}{\lambda_i}} \left(1 - e^{-\frac{\Delta}{\lambda_i}}\right) \times \log_2(n+1) + \underbrace{\sum_{n=0}^{\infty} 1 \times \int_{n\Delta}^{(n+1)\Delta} \frac{e^{-\frac{H_i}{\lambda_i}}}{\lambda_i} \mathrm{d}H_i}_{=1}$$

$$= 1 + \left(1 - e^{-\frac{\Delta}{\lambda_i}}\right) \sum_{n=0}^{\infty} e^{-\frac{n\Delta}{\lambda_i}} \times \log_2(n+1) \leq 1 + \frac{\Delta}{\lambda_i} \sum_{n=0}^{\infty} e^{-\frac{n\Delta}{\lambda_i}} \times \log_2(n+1).$$

With the help of [30, Eq.(22)]: $\sum_{n=1}^{\infty} e^{-\beta n} \log(n) \leq \frac{e^{-\beta}}{\beta} \left[2 + \log\left(1 + \frac{1}{\beta}\right)\right]$, by letting $\beta = e^{-\frac{\Delta}{\lambda_i}}$, we have

$$\begin{aligned}
\sum_{n=0}^{\infty} e^{-\frac{n\Delta}{\lambda_i}} \times \log_2(n+1) &= \sum_{n=1}^{\infty} e^{-\frac{n\Delta}{\lambda_i}} \times \log_2(n+1) \\
&= \frac{e^{\frac{\Delta}{\lambda_i}}}{\log 2} \sum_{n=2}^{\infty} e^{-\frac{n\Delta}{\lambda_i}} \times \log(n) \\
&\leq \frac{1}{\frac{\Delta}{\lambda_i}} \left[\frac{2}{\log 2} + \log_2\left(1 + \frac{1}{\frac{\Delta}{\lambda_i}}\right)\right].
\end{aligned}$$

Then, $R_{r,\mathrm{VLE},i}$ is upper-bounded by[7]

$$R_{r,\mathrm{VLE},i} \leq \frac{2}{\log 2} + 1 + \log_2\left(1 + \frac{1}{\frac{\Delta}{\lambda_i}}\right), \tag{5.9}$$

or equivalently (when $R_{r,\mathrm{VLE},i}$ is sufficiently large),

$$\Delta \leq \frac{\lambda_i}{2^{R_{r,\mathrm{VLE},i} - 1 - \frac{2}{\log 2}} - 1} \leq \frac{\lambda_i}{2^{R_{r,\mathrm{VLE},i} - 2 - \frac{2}{\log 2}}} = C_4 \times 2^{-R_{r,\mathrm{VLE},i}}. \tag{5.10}$$

Substituting (5.10) into (5.7) proves the theorem. □

Therefore, we can see that appropriate values for T and the use of VLE enable the rate loss to decrease at least exponentially with the feedback rate.

[7] Although it is intractable to derive a closed-form expression for $R_{r,\mathrm{VLE},i}$, the upper bound in (5.9) provides a good estimate on how many feedback bits will be consumed.

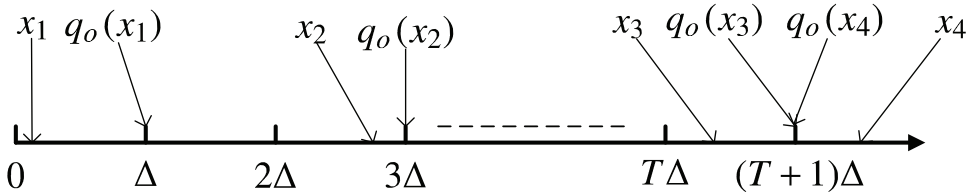


Figure 5.3: A uniform quantizer for outage probability.

5.4 Limited Feedback for Outage Probability

Outage probability is an important performance metric that evaluates the chance that the channels are not strong enough to support the constant-rate data service [55]. An ideal quantizer for outage probability should have at least the following properties: (i) The outage probability loss should decrease toward zero when the feedback rate increases toward infinity. (ii) The outage probability loss should approach zero whenever $P \to 0$ or $P \to \infty$. The intuition of (ii) comes from the fact that when P is adequately small, the outage probabilities of both the full-CSI case and the quantizer should be close to one; when P is significantly large, both outage probabilities should be almost zero. Then, the outage probability losses in both scenarios go to zero.

5.4.1 Proposed Quantizer

As portrayed in Fig. 5.3, the uniform quantizer proposed for outage probability is given by

$$q_o(x) = \begin{cases} \left\lceil \frac{x}{\Delta} \right\rceil \times \Delta, & x \leq T\Delta, \\ (T+1)\Delta, & x > T\Delta. \end{cases} \tag{5.11}$$

The only difference between $q_o(\cdot)$ and $q_r(\cdot)$ lies in whether the left or right boundary of the interval is used as the reconstruction point. The quantizer proposed for rate adaptation cannot be directly inherited because when the channel is very weak (i.e., $H_i < \Delta$), it will

be quantized as zero (i.e., $q_r(H_i) = 0$), which will result in a zero-value power allocation coefficient, i.e.,$\alpha_{q_r} = 0$, and a minimum rate of zero, i.e., $r_1(\alpha_{q_r}) = 0$ or $r_2(\alpha_{q_r})$. In this case, the transmission will surely encounter an outage. However, even a weak channel reserves the possibility of non-outage, so long as the transmit power P is large enough. Therefore, an appropriate quantizer for outage probability should not quantize any value to zero. The quantizer in (5.11) fulfills this requirement.

5.4.2 Outage Probability Loss

Lemma 5.3. *The outage probability loss of the quantizer $q_o(\cdot)$ is upper-bounded by:*

$$\mathtt{out}_{\mathrm{loss},q_o} \leq C_5 \times e^{-\frac{C_6}{P}} \times \frac{1+\sqrt{P}}{P} \times \max\left\{\Delta^{\frac{1}{2}}, \Delta^{\frac{3}{2}}, e^{-\frac{T\Delta}{\lambda_1}}\right\}, \tag{5.12}$$

where C_5 and C_6 are positive constants that are independent of P and Δ.

Proof. See Appendix D.2. □

Different from the rate loss which increases linearly in terms of P, because of the term $e^{-\frac{C_6}{P}} \times \frac{1+\sqrt{P}}{P}$, the upper bound on $\mathtt{out}_{\mathrm{loss},q_o}$ in (5.12) converges to zero either when $P \to 0$ or $P \to \infty$.

To have good performance, we mainly focus on the quantizers with small granularities. When $\Delta \leq 1$, we have $\Delta^{\frac{3}{2}} \leq \Delta^{\frac{1}{2}}$, and the upper bound in (5.12) is restricted by $\max\left\{e^{-\frac{T\Delta}{\lambda_1}}, \Delta^{\frac{1}{2}}\right\}$. For fixed Δ, the optimal choice for T should satisfy $e^{-\frac{T\Delta}{\lambda_1}} = \Delta^{\frac{1}{2}}$, given by $T = \frac{\lambda_1}{2\Delta}\log\frac{1}{\Delta}$.

Corollary 5.2. *When $0 < \Delta \leq 1$ and $T = \frac{\lambda_1}{2\Delta}\log\frac{1}{\Delta}$, the average rate loss of the quantizer $q_o(\cdot)$ is upper-bounded by:*

$$\mathtt{out}_{\mathrm{loss},q_o} \leq C_5 \times e^{-\frac{C_6}{P}} \times \frac{1+\sqrt{P}}{P} \times \Delta^{\frac{1}{2}}, \tag{5.13}$$

where C_5 and C_6 are positive constants independent of P and Δ.

5.4.3 Feedback Rate

The same VLQ for rate adaptation can be applied to $q_o(\cdot)$ for a better utilization of the feedback resource. From (5.9) and (5.10), we obtain $R_{o,\mathrm{VLE},i} \leq \frac{2}{\log 2} + 1 + \log_2\left(1 + \frac{1}{\frac{\Delta}{\lambda_i}}\right)$ and $\Delta \leq C_4 \times 2^{-R_{o,\mathrm{VLE},i}}$. Thus, $\Delta^{\frac{1}{2}} \leq \sqrt{C_4 \times 2^{-R_{o,\mathrm{VLE},i}}} = C_7 \times 2^{-\frac{R_{o,\mathrm{VLE},i}}{2}} \leq C_7 \times 2^{-\frac{\min\left\{R_{o,\mathrm{VLE},1}, R_{o,\mathrm{VLE},2}\right\}}{2}}$. The following theorem states the relationship between the outage probability loss of $q_o(\cdot)$ and the feedback rates.

Theorem 5.3. *When variable-length encoding is applied to the quantizer $q_o(\cdot)$, the rate loss decays at least exponentially as:*

$$\mathtt{out}_{\mathrm{loss},q_o} \leq C_8 \times e^{-\frac{C_6}{P}} \times \frac{1+\sqrt{P}}{P} \times 2^{-\frac{\min\left\{R_{o,\mathrm{VLE},1}, R_{o,\mathrm{VLE},2}\right\}}{2}}, \tag{5.14}$$

where C_6 and C_8 are positive constants independent of P and $R_{o,\mathrm{VLE},i}$.

5.4.4 Diversity Order

With an outage probability $\mathtt{out}$, the achieved diversity order is given as $d = \lim_{P\to\infty} \frac{\log \mathtt{out}}{\log P}$ [55, Section 2.3]. The following lemma shows the achievable diversity order of $q_o(\cdot)$ and a sufficient condition to achieve the maximum diversity order in the full-CSI scenario.

Lemma 5.4. *(1) With $q_o(\cdot)$ and fixed Δ, the diversity orders of $\frac{1}{2}$ and 1 are achievable for Receivers 1 and 2, respectively.*

(2) A sufficient condition for both receivers to achieve the maximum diversity order of 1 is $\Delta \sim_P P^{-\frac{1}{3}}$.

Proof. See Appendix D.3. □

In the full-CSI case, both receivers can achieve the same diversity order of 1 as in the case when no interference exists. In the limited feedback case, it can be found from the proofs in Appendices D.2 and D.3 that the cause of this insufficient diversity order for Receiver 1 comes from the marginal region when $0 < H_1, H_2 \leq \Delta$. Therefore, an adequately small Δ that scales at least in proportion to $P^{-\frac{1}{3}}$ in the high-P region is desired to diminish the probability that H_i falls into that region so as to obtain the maximum diversity gain.

5.5 Extension to More than Two Receivers

5.5.1 Full-CSI Performance

In this section, we consider NOMA with more than two downlink receivers. Assuming perfect CSI universally available and $H_1 \geq H_2 \geq \cdots \geq H_K$, the maximum minimum rate can be obtained by solving the optimization problem:

$$r_{\max} = \max_{\boldsymbol{\alpha}=[\alpha_1,\ldots,\alpha_K]} \min_{k=1,\ldots,K} r_k(\boldsymbol{\alpha}), \text{ subject to } 0 \leq \alpha_k \leq 1, \sum_{k=1}^{K} \alpha_k = 1, \tag{5.15}$$

where K is the number of receivers, and $r_k(\boldsymbol{\alpha}) = \log_2\left(1 + \frac{\alpha_k}{\sum_{i=1}^{k-1}\alpha_i + \frac{1}{PH_k}}\right)$ is the achieved rate for Receiver k under superposition coding and SIC. To the best of our knowledge, no closed-form solution for $r_{\max}$ is available in the literature. We present the following lemma that helps solving the above optimization problem numerically.

Lemma 5.5. *There exists $\boldsymbol{\alpha}^\star = [\alpha_1^\star, \alpha_2^\star, \ldots, \alpha_K^\star]$, such that all receivers achieve the same rate at optimality, i.e., $r_{\max} = r_1(\boldsymbol{\alpha}^\star) = r_2(\boldsymbol{\alpha}^\star) = \cdots = r_K(\boldsymbol{\alpha}^\star)$.*

The proof of Lemma 5.5 is given in Appendix D.4. Since

$$r_{\max} = r_k\left(\boldsymbol{\alpha}^\star\right) = \log_2\left(1 + \frac{\alpha_k^\star}{\sum_{i=1}^{k-1}\alpha_i^\star + \frac{1}{PH_k}}\right)$$

for $k = 1, \ldots, K$, we have $\alpha_k^\star = (2^{r_{\max}} - 1) \times \left(\sum_{i=1}^{k-1}\alpha_i^\star + \frac{1}{PH_k}\right)$, which leads to[8]

$$\alpha_k^\star = (2^{r_{\max}} - 1)\left[\frac{1}{PH_k} + (2^{r_{\max}} - 1)\sum_{i=1}^{k-1}\frac{2^{(k-1-i)r_{\max}}}{PH_i}\right]. \tag{5.16}$$

To find $\alpha_k^\star$, we need to solve for $r_{\max}$ first. Summing both sides from $k = 1, \ldots, K$ and after trivial calculations, we obtain

$$\sum_{k=1}^{K}\alpha_k^\star = 1 = \underbrace{(2^{r_{\max}} - 1)\sum_{i=1}^{K}\frac{2^{(K-i)r_{\max}}}{PH_i}}_{=\varpi(r_{\max})}. \tag{5.17}$$

In other words, $r_{\max}$ satisfies $\varpi\left(r_{\max}\right) = 1$.[9]

Let $r_{\text{ub}} = \log_2\left(1 + \min_{k=1,\ldots,K} PH_k\right) = \log_2(1 + PH_K)$. Since $\varpi\left(x\right)$ is an increasing function of x as well as $\varpi(0) < 1$ and $\varpi(r_{\text{ub}}) \geq 1$, we could use the bisection method to find the root of $\varpi(x) = 1$ in the interval $(0, r_{\text{ub}}]$. The calculation of $\varpi\left(x\right)$ costs $O(K)$, thus, the time complexity of finding $r_{\max}$ within an accuracy of ϵ is $O\left(K \log \frac{1}{\epsilon}\right)$.

5.5.2 Limited Feedback

Under limited feedback, the previously proposed quantizers $q_r\left(\cdot\right)$ and $q_o\left(\cdot\right)$ in Figs. 5.2 and 5.3 can still be applied here for rate adaptation and outage probability, respectively. The maximum minimum rate can be calculated using the bisection method by treating $q_r\left(H_k\right)$

[8] Note that [51] also derives (5.16), but using the tools of convex optimization.

[9] Note that [56] has solved a different optimization problem, i.e. maximizing the sum rate subject to a minimum rate constraint, which satisfies $\sum_{k=1}^{K}\alpha_k^\star = 1$ but results in different $\alpha_k^\star$s.

or $q_o(H_k)$ as H_k, and the corresponding power allocation coefficients can be computed. Although it is non-trivial to derive upper bounds on the losses in rate or outage probability for $K > 2$ theoretically, numerical simulations in the next section show that the relationships between the performance loss and the feedback rate are similar to the case of $K = 2$.

5.6 Numerical Simulations and Discussions

In this section, we perform numerical simulations to validate the effectiveness of our proposed quantizers for rate adaptation and outage probability. In all subsequent simulations for K receivers, we use the channel variances in Table 5.1.

Table 5.1: Channel variances for numerical simulations.

$K = 2$	$\lambda_1 = 1, \lambda_2 = 0.5$
$K > 2$	$\lambda_k = \frac{1}{k}, k = 1, \ldots, K$

Results for other values of channel variances will exhibit similar observations. For outage probability, sufficiently large number of channel realizations are generated to observe at least 10000 outage events.

In Fig. 5.4, we simulated the minimum rates of the full-CSI case, $q_r(\cdot)$ and the TDMA scheme (where each receiver occupies half of the time to transmit). We observe that the proposed quantizer with NOMA outperforms the TDMA scheme when $\Delta = 0.01$ and 0.05. The rate loss between the full-CSI case and $q_r(\cdot)$ with $\Delta = 0.01$ is almost negligible. The corresponding values for $T = \frac{\lambda_1}{\Delta} \log \frac{1}{\Delta}$ and the feedback rates for both receivers (bits/per channel state) are listed in Table 5.2. Compared with FLE which costs $\lceil \log_2(T+1) \rceil$ bits per receiver per channel state, VLE can save almost half of the feedback bits.

In Fig. 5.5, we plot the rate losses of $q_r(\cdot)$ for different values of Δ and the feedback rates

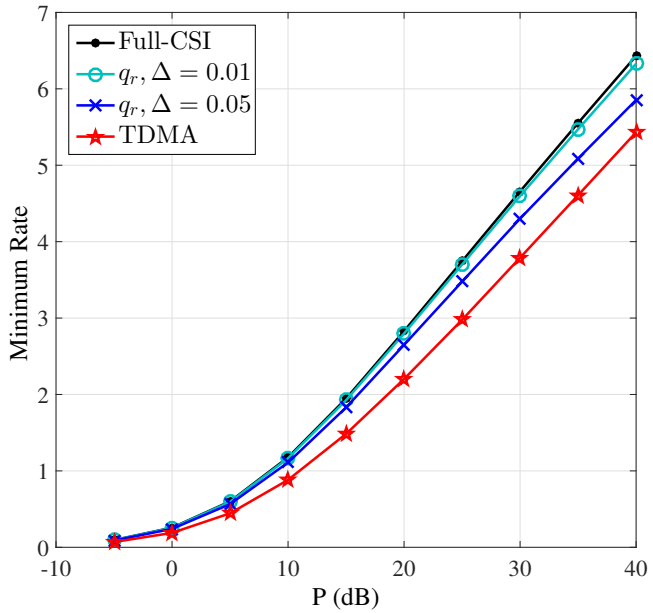


Figure 5.4: Simulated minimum rates of NOMA for $K = 2$.

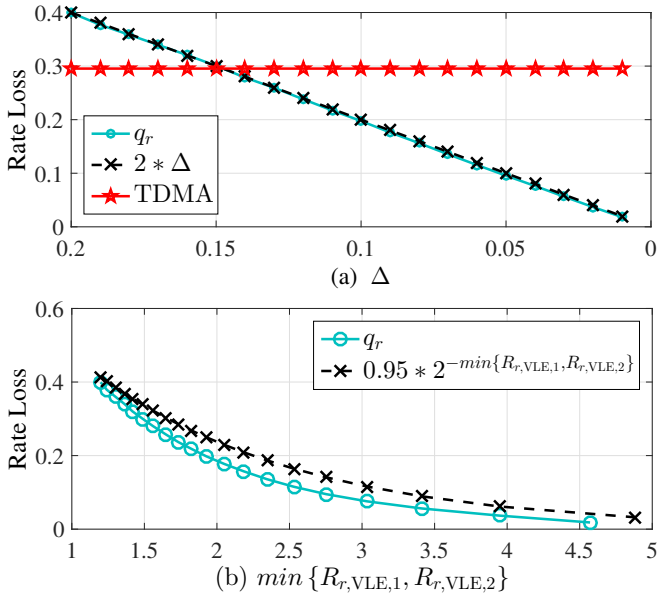


Figure 5.5: Simulated rate losses versus (a) Δ and (b) $\min\{R_{r,\mathrm{VLE},1}, R_{r,\mathrm{VLE},2}\}$ for $K = 2$ and $P = 10$ dB.

Table 5.2: Feedback rate for either receiver.

Δ	T	$\lceil \log_2(T+1) \rceil$	Receiver 1	Receiver 2
0.01	461	9	5.3	4.6
0.05	60	6	3.6	2.7

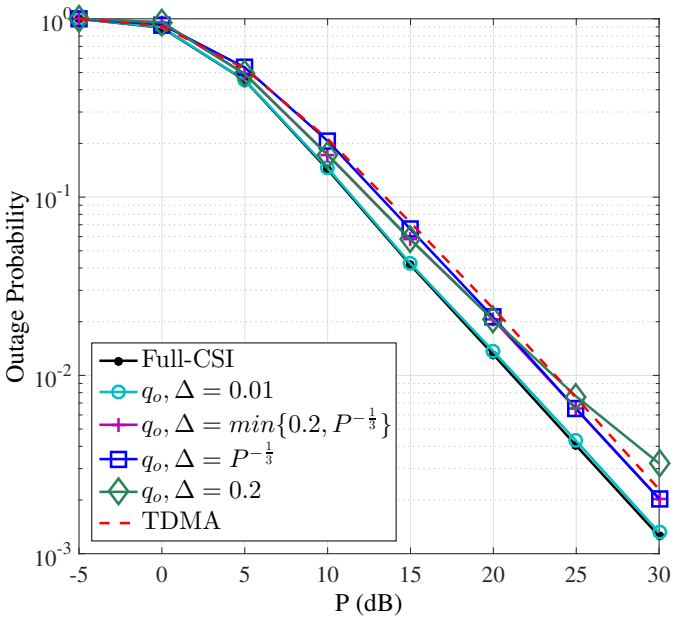


Figure 5.6: Simulated outage probabilities of NOMA for $K = 2$.

$R_{r,\mathrm{VLE},1}$ and $R_{r,\mathrm{VLE},2}$. It shows that the rate loss of $q_r(\cdot)$ decreases at least linearly with respect to Δ and exponentially with $\min\{R_{r,\mathrm{VLE},1}, R_{r,\mathrm{VLE},2}\}$, which validates the accuracy of our derived upper bounds in (5.7) and (5.8). In addition, Fig. 5.5(a) shows that Δ needs to be less than 0.15 such that $q_r(\cdot)$ can obtain a higher rate compared with the TDMA scheme.

In Fig. 5.6, we compare the outage probabilities of the full-CSI case, $q_o(\cdot)$ under various values of Δ and the TDMA scheme. It can be seen that: (i) The curve for $q_o(\cdot)$ with $\Delta = 0.01$ almost coincides with that of the full-CSI case. (ii) When P is large, $q_o(\cdot)$ with $\Delta = 0.2$ suffers from an insufficient diversity gain in the high-P region. According to our analysis in Lemma 4, $\Delta = 0.2$ is large enough not to scale with $P^{-\frac{1}{3}}$.[10] (iii) Although the maximum diversity order is achieved when $\Delta = P^{-\frac{1}{3}}$, much less array gain is obtained in

[10]The value 0.01 for Δ will also exhibit an insufficient diversity order as long as P is large enough, although we might not be able to observe this in the region of $P \leq 30$ dB in Fig. 5.6.

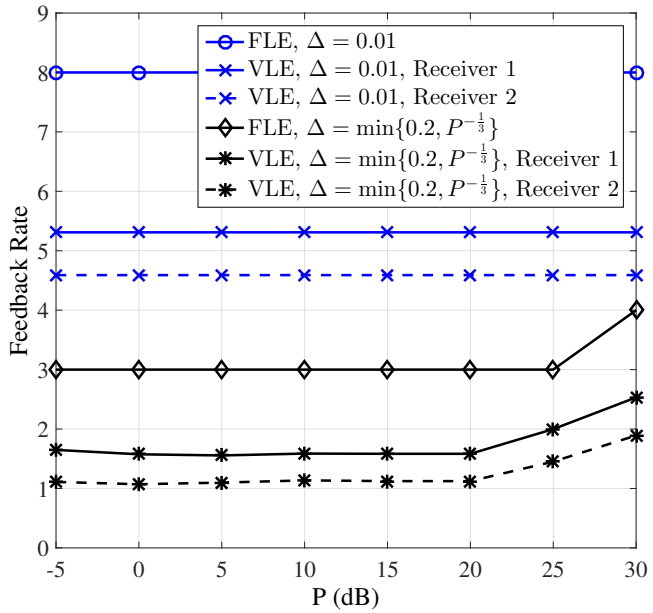


Figure 5.7: Simulated feedback rates versus P for $K = 2$.

the lower and medium-P regions (where Δ is large). Alternatively, $\Delta = \min\left\{0.2, P^{-\frac{1}{3}}\right\}$ will reserve both benefits of the maximum diversity order brought by $P^{-\frac{1}{3}}$ and the higher array gain of $\Delta = 0.2$.[11] The comparison of feedback rates for VLE and FLE (which requires $\lceil \log_2(T+2) \rceil = \left\lceil \log_2\left(\frac{\lambda_1}{2\Delta} \log \frac{1}{\Delta} + 2\right) \right\rceil$ bits per channel state) under different values of Δ and P is shown in Fig. 5.7, which verifies the superiority of VLE. It can be seen that the feedback rates for $\Delta = \min\left\{0.2, P^{-\frac{1}{3}}\right\}$ stay flat in the low and medium-P regions (since $0.2 \leq P^{-\frac{1}{3}}$). When $P^{-\frac{1}{3}} \leq 0.2$ where $P \geq 20.9$ dB, the feedback rates start to increase as Δ gets smaller.

In Fig. 5.8(a), the outage probability loss decays at least linearly with respect to Δ; in Fig. 5.8(b), the outage probability loss approaches zero whenever $P \to 0$ or $P \to \infty$; in Fig. 5.9, the outage probability loss decays at least exponentially with $\frac{\min\left\{R_{0,\mathrm{VLE},1}, R_{0,\mathrm{VLE},2}\right\}}{2}$. All these observations validate our theoretical analysis.

In Figs. 5.10 and 5.11, we simulated the rate and outage probability losses for more than two receivers. For Receiver k, the channel variance is set to be $\lambda_k = \frac{1}{k}$, the maximum number

[11] We also observe a similar effect of Δ on the achieved minimum rates, but we mainly elaborate it on outage probability.

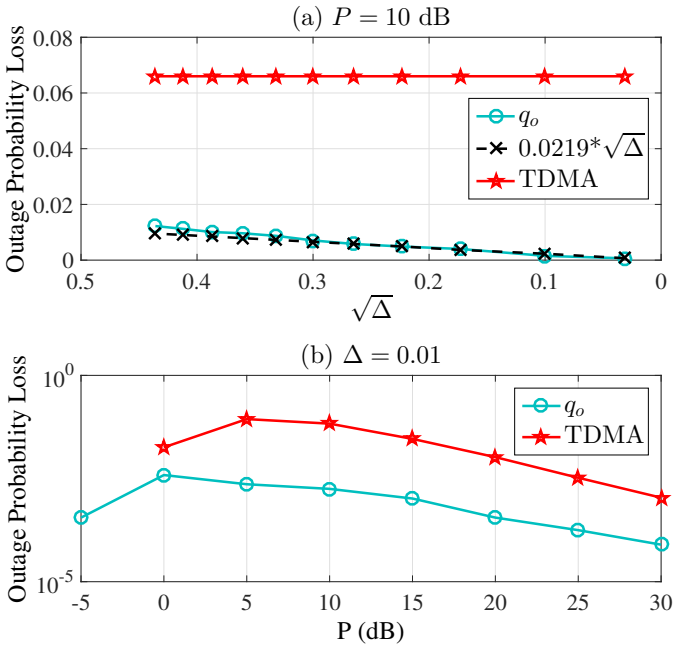


Figure 5.8: Simulated outage probability losses versus $\sqrt{\Delta}$ and P for $K = 2$.

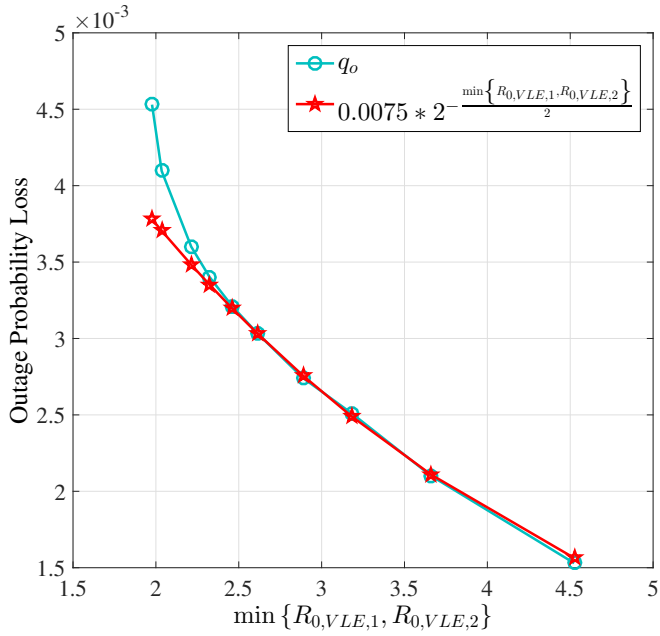


Figure 5.9: Simulated outage probability losses versus $\min\{R_{0,\mathrm{VLE},1}, R_{0,\mathrm{VLE},2}\}$ for $K = 2$.

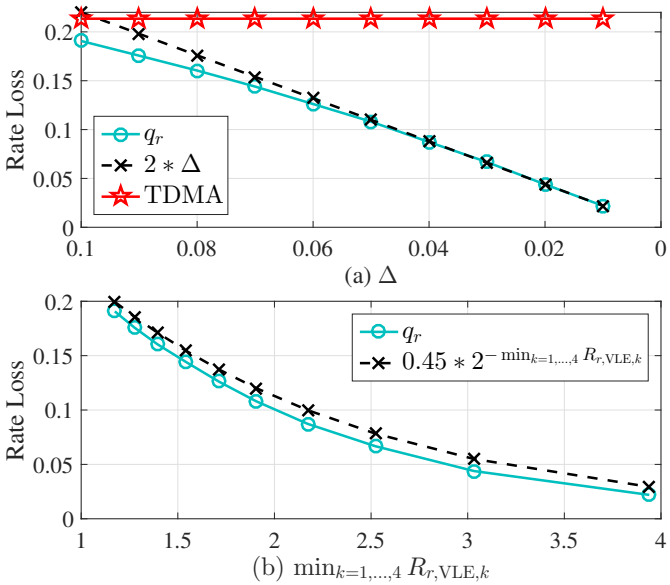


Figure 5.10: Simulated rate losses versus (a) Δ and (b) $\min_{k=1,\ldots,K} R_{r,\mathrm{VLE},k}$ for $K = 4$ and $P = 10$ dB.

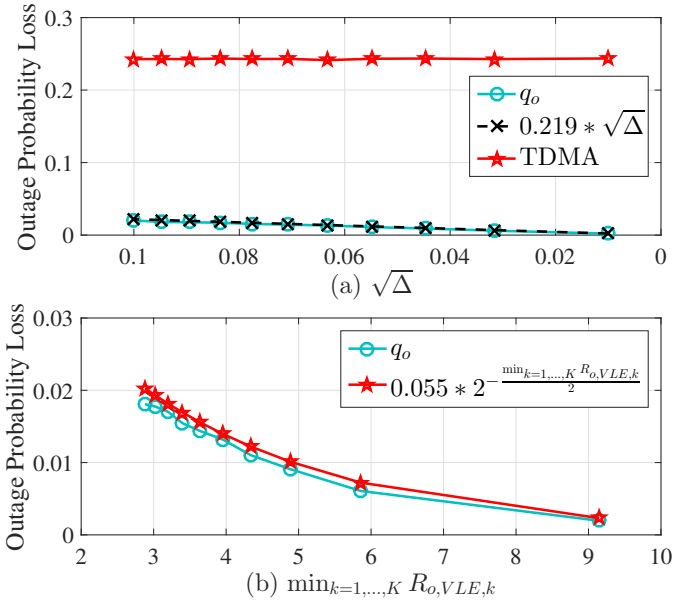


Figure 5.11: Simulated outage probability losses versus (a) $\sqrt{\Delta}$ and (b) $\min_{k=1,\ldots,K} R_{o,\mathrm{VLE},k}$ for $K = 4$ and $P = 10$ dB.

of bins T for $q_r(\cdot)$ and $q_o(\cdot)$ is $T = \frac{\lambda_k}{\Delta} \log \frac{1}{\Delta}$, and the accuracy used by the bisection method is $\epsilon = 10^{-4}$. We simply treat the result of bisection method based on perfect CSI as the "full-CSI" performance. Compared with Figs. 5.5, 5.8 and 5.9 for $K = 2$, Figs. 5.10 and 5.11 exhibit very similar relationships between the losses and Δ or the feedback rates.

5.7 Conclusions and Future Work

We have introduced efficient quantizers for rate adaptation and outage probability of minimum rate in NOMA with two receivers. We have proved that the losses in rate and outage probability both decrease at least exponentially with the minimum of the feedback rates. Furthermore, we generalized the proposed quantizers to NOMA with any number of receivers. The performance of NOMA with noisy quantized feedback and the user scheduling under limited feedback will be interesting future research directions.

Chapter 6

Conclusions and Future Work

In this dissertation, we have applied the idea of variable length encoding and designed efficient distributed quantizers for different kinds of multiuser networks. Compared with traditional FLQs, our proposed VLQs have obtained much less or even zero performance losses with only finite average feedback rates.

For the future work, we would like to extend the proposed VLQs to other wireless networks. For example, the millimeter wave (MmWave) transmission is a promising technology for future cellular systems and has drawn a lot of attention recently [57]. MmWave cellular systems, operating in the 10-300GHz band, brings in more available spectrum and supports multiple gigabit-per-second data rates. However, due to sparse scattering, the characteristics of the MmWave channel are essentially different from those of the wireless channels we have studied, i.e., Rayleigh fading. For a point-to-point MmWave system, the channel matrix can be written as [58]

$$\boldsymbol{H} = \sqrt{\frac{N_{\mathrm{T}} N_{\mathrm{R}}}{L}} \sum_{l=1}^{L} \alpha_l \boldsymbol{\alpha}_{\mathrm{R}}\left(\theta_l\right) \boldsymbol{\alpha}_{\mathrm{T}}^{\dagger}\left(\phi_l\right),$$

where N_{T} and N_{R} are the numbers of antennas at the transmitter and receiver, respectively;

L is the number of scatters; α_l is the complex gain of the l-th path; θ_l and ϕ_l are the l-th path's angles of arrival and departure; $\boldsymbol{\alpha}_{\mathrm{T}}(\phi_l)$ and $\boldsymbol{\alpha}_{\mathrm{R}}(\theta_l)$ are the antenna array response vectors of the transmitter and receiver, respectively. If uniform linear arrays are assumed, $\boldsymbol{\alpha}_{\mathrm{T}}(\phi)$ is defined as:

$$\boldsymbol{\alpha}_{\mathrm{T}}(\phi) = \frac{1}{N_{\mathrm{T}}}\left[1, e^{j\frac{2\pi}{\lambda}d\sin(\phi)}, \ldots, e^{j(N_{\mathrm{T}}-1)\frac{2\pi}{\lambda}d\sin(\phi)}\right]^{\top},$$

where λ is the signal wavelength, and d is the distance between antenna elements. The array response vectors at the receiver, $\boldsymbol{\alpha}_{\mathrm{R}}(\theta)$, can be defined in a similar fashion.

It can be seen that the MmWave channel matrix assumes a complicated composition involving the path-loss coefficients, angles of arrivals/departures and the array form. Therefore, the limited feedback scheme needs to be re-designed to accommodate the MmWave channel.

Bibliography

[1] J. C. Roh and B. D. Rao, "Transmit beamforming in multiple-antenna systems with finite rate feedback: a VQ-based approach," *IEEE Trans. Inf. Theory*, vol. 52, no. 3, pp. 1101–1112, Mar. 2006.

[2] S. Ekbatani and H. Jafarkhani, "Combining beamforming and space-time coding using quantized feedback," *IEEE Trans. Wireless Commun.*, vol. 7, no. 3, pp. 898–908, Mar. 2008.

[3] E. Koyuncu and H. Jafarkhani, "Variable-length limited feedback beamforming in multiple-antenna fading channels," *IEEE Trans. Inf. Theory*, vol. 60, no. 11, pp. 7140–7164, Nov. 2014.

[4] D. J. Love, R. W. Heath, Jr., and T. Strohmer, "Grassmannian beamforming for multiple-input multiple-output wireless systems," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2735–2747, Oct. 2003.

[5] E. Koyuncu and H. Jafarkhani, "Distributed beamforming in wireless multiuser relay-interference networks with quantized feedback," *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4538–4576, July 2012.

[6] K. Anand, E. Gunawan, and Y. L. Guan, "Beamformer design for the MIMO interference channels under limited channel feedback," *IEEE Trans. Commun.*, vol. 61, no. 8, pp. 3246–3258, Aug. 2013.

[7] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3180, Dec. 2004.

[8] M. Xia, Y. C. Wu, and S. Aissa, "Exact outage probability of dual-hop CSI-assisted AF relaying over nakagami-m fading channels," *IEEE Trans. Signal Process.*, vol. 60, no. 10, pp. 5578–5583, Oct. 2012.

[9] N. Ahmed, M. A. A. Khojastepour, A. Sabharwal, and B. Aazhang, "Outage minimization with limited feedback for the fading relay channel," *IEEE Trans. Commun.*, vol. 54, no. 4, pp. 659–669, Apr. 2006.

[10] E. Koyuncu, Y. Jing, and H. Jafarkhani, "Distributed beamforming in wireless relay networks with quantized feedback," *IEEE J. Select. Areas Commun.*, vol. 26, no. 8, pp. 1429–1439, Oct. 2008.

[11] Y. Zhao, R. Adve, and T. J. Lim, "Beamforming with limited feedback in amplify-and-forward cooperative networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 5145–5149, Dec. 2008.

[12] Y. Jing and H. Jafarkhani, "Network beamforming using relays with perfect channel information," *IEEE Trans. Inf. Theory*, vol. 55, no. 6, pp. 2499–2517, Jun. 2009.

[13] ——, "Single and multiple relay selection schemes and their achievable diversity orders," *IEEE Trans. Wireless Commun.*, vol. 8, no. 3, pp. 1414–1423, Mar. 2009.

[14] I. Gradshteyn and I. Ryzhik, *Table of Integrals, Series, and Products*, 7th ed., A. Jeffrey and D. Zwillinger, Eds. Academic Press, Mar. 2007.

[15] C. K. Au-Yeung and D. J. Love, "On the performance of random vector quantization limited feedback beamforming in a MISO system," *IEEE Trans. Wireless Commun.*, vol. 6, no. 2, pp. 458–462, Feb. 2007.

[16] E. Koyuncu and H. Jafarkhani, "Very low-rate variable-length channel quantization for minimum outage probability," in *IEEE Data Compression Conference (DCC)*, pp. 261–270, Mar. 2013.

[17] D. J. Love, R. W. Heath, Jr., V. K. N. Lau, D. Gesbert, B. D. Rao, and M. Andrews, "An overview of limited feedback in wireless communication systems," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 8, pp. 1341–1365, Oct. 2008.

[18] N. Jindal, "MIMO broadcast channels with finite-rate feedback," *IEEE Trans. Inf. Theory*, vol. 52, no. 11, pp. 5045–5060, Nov. 2006.

[19] I. H. Kim, D. J. Love, and S. Y. Park, "Optimal and successive approaches to signal design for multiple antenna physical layer multicasting," *IEEE Trans. Commun.*, vol. 59, no. 8, pp. 2316 –2327, Aug. 2011.

[20] S. X. Wu, W.-K. Ma, and A. M.-C. So, "Physical-layer multicasting by stochastic transmit beamforming and Alamouti space-time coding," *IEEE Trans. Signal Process.*, vol. 61, no. 17, pp. 4230–4245, Sept. 2013.

[21] K. K. Mukkavilli, A. Sabharwal, E. Erkip, and B. Aazhang, "On beamforming with finite rate feedback in multiple-antenna systems," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2562–2579, Oct. 2003.

[22] N. D. Sidiropoulos, T. N. Davidson, and Z.-Q. Luo, "Transmit beamforming for physical-layer multicasting," *IEEE Trans. Signal Process.*, vol. 54, no. 6, pp. 2239–2251, June 2006.

[23] H. Farhadi, C. Wang, and M. Skoglund, "On the throughput of wireless interference networks with limited feedback," in *IEEE International Symposium on Information Theory (ISIT)*, pp. 762–766, July 2011.

[24] R. Bhagavatula and R. W. Heath, Jr., "Adaptive limited feedback for sum-rate maximizing beamforming in cooperative multicell systems," *IEEE Trans. Signal Process.*, vol. 59, no. 2, pp. 800–811, Feb. 2011.

[25] S. Akoum and R. W. Heath Jr., "Interference coordination: Random clustering and adaptive limited feedback," *IEEE Trans. Signal Process.*, vol. 61, no. 7, pp. 1822–1834, Apr. 2013.

[26] V. Shah, N. B. Mehta, and R. Yim, "Splitting algorithms for fast relay selection: Generalizations, analysis, and a unified view," *IEEE Trans. Wireless Commun.*, vol. 9, no. 4, pp. 1525–1535, Apr. 2010.

[27] M. Charafeddine, A. Sezgin, and A. Paulraj, "Rate region frontiers for $n-$user interference channel with interference as noise," in *Annual Allerton Conference on Communication, Control and Computing*, Sept. 2007.

[28] M. Charafeddine and A. Paulraj, "2−sector interference channel communication for sum rates and minimum rate maximization," in *43rd Annual Conference on Information Sciences and Systems (CISS)*, 2009.

[29] R. Etkin, D. Tse, and H. Wang, "Gaussian interference channel capacity to within one bit," *IEEE Trans. Inf. Theory*, vol. 54, no. 12, pp. 5534–5562, Dec. 2008.

[30] X. Liu, E. Koyuncu, and H. Jafarkhani, "Multicast networks with variable-length limited feedback," *IEEE Trans. Wireless Commun.*, vol. 14, no. 1, pp. 252–264, Jan. 2015.

[31] J. Park, Y. Sung, D. Kim, and H. V. Poor, "Outage probability and outage-based robust beamforming for MIMO interference channels with imperfect channel state information," *IEEE Trans. Wireless Commun.*, vol. 11, no. 10, pp. 3561–3573, Oct. 2012.

[32] A. Gjendemsjø, D. Gesbert, G. E. Øien, and S. G. Kiani, "Optimal power allocation and scheduling for two-cell capacity maximization," in *4th International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks*, pp. 1–6, 2006.

[33] Y. Linde, A. Buzo, and R. M. Gray, "An algorithm for vector quantizer design," *IEEE Trans. Commun.*, vol. 28, no. 1, pp. 84–95, Jan. 1980.

[34] Y. Saito, A. Benjebbour, Y. Kishiyama, and T. Nakamura, "System-level performance evaluation of downlink non-orthogonal multiple access (NOMA)," in *IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Sept. 2013, pp. 611–615.

[35] 3rd Generation Partnership Project (3GPP), "Study on downlink multiuser superposition transmission for LTE," Mar. 2015.

[36] K. Hoiguchi and A. Benjebbour, "Non-orthogonal multiple access (NOMA) with successive interference cancellation for future radio access," *IEICE Trans. Commun.*, vol. E98-B, no. 3, pp. 403–414, 2015.

[37] P. Xu, Z. Ding, X. Dai, and H. V. Poor, "A new evaluation criterion for non-orthogonal multiple access in 5G software defined networks," *IEEE Access*, vol. 3, pp. 1633–1639, 2015.

[38] Z. Ding, Z. Yang, P. Fan, and H. V. Poor, "On the performance of non-orthogonal multiple access in 5G systems with randomly deployed users," *IEEE Signal Process. Lett.*, vol. 21, no. 12, pp. 1501–1505, Dec. 2014.

[39] J. A. Oviedo and H. R. Sadjadpour, "A new NOMA approach for fair power allocation," in *IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, Apr. 2016, pp. 843–847.

[40] Z. Yang, Z. Ding, P. Fan, and N. Al-Dhahir, "A general power allocation scheme to guarantee quality of service in downlink and uplink NOMA systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 11, pp. 7244–7257, Nov. 2016.

[41] J. Choi, "Power allocation for max-sum rate and max-min rate proportional fairness in NOMA," *IEEE Commun. Lett.*, vol. 20, no. 10, pp. 2055–2058, Oct. 2016.

[42] L. Lei, D. Yuan, and P. Varbrand, "On power minimization for non-orthogonal multiple access (NOMA)," *IEEE Commun. Lett.*, vol. 20, no. 12, pp. 2458–2461, Dec. 2016.

[43] M. S. Ali, H. Tabassum, and E. Hossain, "Dynamic user clustering and power allocation for uplink and downlink non-orthogonal multiple access (NOMA) systems," *IEEE Access*, vol. 4, pp. 6325–6343, 2016.

[44] X. Liu, E. Koyuncu, and H. Jafarkhani, "Cooperative quantization for two-user interference channels," *IEEE Trans. Commun.*, vol. 63, no. 7, pp. 2698–2712, 2015.

[45] S. Liu and C. Zhang, "Downlink non-orthogonal multiple access system with limited feedback channel," in *International Conference on Wireless Communications Signal Processing (WCSP)*, Oct. 2015, pp. 1–5.

[46] P. Xu, Y. Yuan, Z. Ding, X. Dai, and R. Schober, "On the outage performance of non-orthogonal multiple access with 1-bit feedback," *IEEE Trans. Wireless Commun.*, vol. 15, no. 10, pp. 6716–6730, Oct. 2016.

[47] J. Cui, Z. Ding, and P. Fan, "A novel power allocation scheme under outage constraints in NOMA systems," *IEEE Signal Process. Lett.*, vol. 23, no. 9, pp. 1226–1230, Sept. 2016.

[48] Z. Yang, Z. Ding, P. Fan, and G. K. Karagiannidis, "On the performance of non-orthogonal multiple access systems with partial channel information," *IEEE Trans. Commun.*, vol. 64, no. 2, pp. 654–667, Feb. 2016.

[49] M. F. Hanif, Z. Ding, T. Ratnarajah, and G. K. Karagiannidis, "A minorization-maximization method for optimizing sum rate in the downlink of non-orthogonal multiple access systems," *IEEE Trans. Signal Process.*, vol. 64, no. 1, pp. 76–88, Jan. 2016.

[50] Y. Yu, H. Chen, Y. Li, Z. Ding, and B. Vucetic, "Antenna selection for MIMO non-orthogonal multiple access systems," https://arxiv.org/pdf/1609.07978v1.pdf, 2016.

[51] S. Timotheou and I. Krikidis, "Fairness for non-orthogonal multiple access in 5G systems," *IEEE Signal Process. Lett.*, vol. 22, no. 10, pp. 1647–1651, Oct. 2015.

[52] J. Choi, "On the power allocation for a practical multiuser superposition scheme in NOMA systems," *IEEE Commun. Lett.*, vol. 20, no. 3, pp. 438–441, Mar. 2016.

[53] T. M. Cover and J. A. Thomas, *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*. Wiley-Interscience, 2006.

[54] R. Sun, M. Hong, and Z.-Q. Luo, "Joint downlink base station association and power control for max-min fairness: Computation and complexity," *IEEE J. Select. Areas Commun.*, vol. 33, no. 6, pp. 1040–1054, June 2015.

[55] H. Jafarkhani, *Space-Time Coding: Theory and Practice*, 1st ed. New York, NY, USA: Cambridge University Press, 2005.

[56] Z. Chen, Z. Ding, X. Dai, and R. Zhang, "A mathematical proof of the superiority of NOMA compared to conventional OMA," https://arxiv.org/pdf/1612.01069.pdf, 2016.

[57] A. Alkhateeb and R. W. Heath, "Frequency selective hybrid precoding for limited feedback millimeter wave systems," *IEEE Trans. Commun.*, vol. 64, no. 5, pp. 1801–1818, May 2016.

[58] A. Alkhateeb, G. Leus, and R. W. Heath, "Limited feedback hybrid precoding for multi-user millimeter wave systems," *IEEE Trans. Wireless Commun.*, vol. 14, no. 11, pp. 6481–6494, Nov. 2015.

[59] S. Li, "Concise formulas for the area and volume of a hyperspherical cap," *Asian J. Math. Stat.*, pp. 66 – 77, 2011.

[60] H. Alzer, "Sharp inequalities for the beta function," *Indagationes Mathematicae*, vol. 12, pp. 15–21, 2001.

[61] M. Abramowitz and I. A. Stegun, *Handbook of mathematical functions*, 1964.

[62] X. Liu and H. Jafarkhani, "Downlink non-orthogonal multiple access with limited feedback," https://arxiv.org/pdf/1701.05247.pdf, 2017.

Appendices

A Supplementary Proofs for Chapter 2

A.1 Proof of Theorem 2.1

Proof. We need the following lemmas to prove Theorem 2.1.

Lemma A.1. *Given* $\boldsymbol{H} \in \mathsf{C}^{2N\times 1}$ *with* $\Gamma\left(\boldsymbol{w}^{\star}_{\mathtt{SUM}}, \boldsymbol{H}\right) > \alpha$*, let*

$$\Pi \triangleq \frac{\Gamma\left(\boldsymbol{w}^{\star}_{\mathtt{SUM}}, \boldsymbol{H}\right) - \alpha}{2\sqrt{N}P\left(\sum_{n=1}^{N}|g_n|\right)^2\left(1+\sum_{n=1}^{N}\frac{|g_n|^2}{|f_n|^2+\frac{1}{P}}\right)}. \tag{A.1}$$

Then, for any $\boldsymbol{w} \in \mathcal{W}_{\mathtt{SUM}}$ *with* $||\boldsymbol{w} - \boldsymbol{w}^{\star}_{\mathtt{SUM}}|| \leq \Pi$*, we have* $\Gamma\left(\boldsymbol{w}, \boldsymbol{H}\right) \geq \alpha$*.*

Proof. We first find an upper bound on $\Gamma\left(\boldsymbol{w}^{\star}_{\mathtt{SUM}}, \boldsymbol{H}\right) - \Gamma\left(\boldsymbol{w}, \boldsymbol{H}\right)$ for any $\boldsymbol{w} \in \mathcal{W}_{\mathtt{SUM}}$. Based on the bound, given $\boldsymbol{H}$ with $\Gamma\left(\boldsymbol{w}^{\star}_{\mathtt{SUM}}, \boldsymbol{H}\right) > \alpha$, we find conditions on $\boldsymbol{w}$ that guarantee $\Gamma\left(\boldsymbol{w}, \boldsymbol{H}\right) \geq \alpha$.

In order to bound $\Gamma\left(\boldsymbol{w}^{\star}_{\mathtt{SUM}}, \boldsymbol{H}\right) - \Gamma\left(\boldsymbol{w}, \boldsymbol{H}\right)$, we successively alter each component of $\boldsymbol{w}^{\star}_{\mathtt{SUM}}$ until we reach $\boldsymbol{w}$, while keeping track of the SNR variation at each step of the alteration [10, Appendix B]. Thus, $\overline{\Gamma\left(\boldsymbol{H}\right)} \triangleq \Gamma\left(\boldsymbol{w}^{\star}_{\mathtt{SUM}}, \boldsymbol{H}\right) - \Gamma\left(\boldsymbol{w}, \boldsymbol{H}\right)$ is decomposed as

$$\overline{\Gamma\left(\boldsymbol{H}\right)} = \sum_{k=1}^{N}\overline{\Gamma_k\left(\boldsymbol{H}\right)} = \sum_{k=1}^{N}\left[\Gamma\left(\boldsymbol{w}^{(k-1)}, \boldsymbol{H}\right) - \Gamma\left(\boldsymbol{w}^{(k)}, \boldsymbol{H}\right)\right], \tag{A.2}$$

where $\boldsymbol{w}^{(0)} \triangleq \boldsymbol{w}^{\star}_{\mathtt{SUM}}$, $\boldsymbol{w}^{(k)} \triangleq \left[[\boldsymbol{w}]_1, \ldots, [\boldsymbol{w}]_k, [\boldsymbol{w}^{\star}_{\mathtt{SUM}}]_{k+1}, \ldots, [\boldsymbol{w}^{\star}_{\mathtt{SUM}}]_N\right]^{\top}$ and $\boldsymbol{w}^{(N)} \triangleq \boldsymbol{w}$. Let

$$
\begin{aligned}
\tilde{f}_n &\triangleq \frac{1}{|f_n|^2 + \frac{1}{P}}, \\
A_k &\triangleq \sum_{n=1}^{N} \left[\boldsymbol{w}^{(k-1)}\right]_n^* f_n g_n \sqrt{\tilde{f}_n}, \\
B_k &\triangleq 1 + \sum_{n=1}^{N} \left|\left[\boldsymbol{w}^{(k-1)}\right]_n\right|^2 |g_n|^2 \tilde{f}_n, \\
\hat{A}_k &\triangleq \sum_{n=1}^{N} \left[\boldsymbol{w}^{(k)}\right]_n^* f_n g_n \sqrt{\tilde{f}_n} = A_k - \left([\boldsymbol{w}^{\star}_{\mathtt{SUM}}]_k^* - [\boldsymbol{w}]_k^*\right) f_k g_k \sqrt{\tilde{f}_k}, \\
\hat{B}_k &\triangleq 1 + \sum_{n=1}^{N} \left|\left[\boldsymbol{w}^{(k)}\right]_n\right|^2 |g_n|^2 \tilde{f}_n = B_k - \left(\left|[\boldsymbol{w}^{\star}_{\mathtt{SUM}}]_k\right|^2 - \left|[\boldsymbol{w}]_k\right|^2\right) |g_k|^2 \tilde{f}_k.
\end{aligned}
$$

According to (2.2), we have

$$
\overline{\Gamma_k(\boldsymbol{H})} = P\frac{|A_k|^2}{B_k} - P\frac{\left|\hat{A}_k\right|^2}{\hat{B}_k} = P\frac{|A_k|^2}{B_k} - P\frac{\left|A_k - \left([\boldsymbol{w}^{\star}_{\mathtt{SUM}}]_k^* - [\boldsymbol{w}]_k^*\right) f_k g_k \sqrt{\tilde{f}_k}\right|^2}{\hat{B}_k}.
$$

An upper bound on $\overline{\Gamma_k(\boldsymbol{H})}$ can then be derived as

$$
\begin{aligned}
\overline{\Gamma_k(\boldsymbol{H})} &= P\frac{|A_k|^2}{B_k} - P\frac{|A_k|^2}{\hat{B}_k} - \underbrace{P\frac{\left|\left([\boldsymbol{w}^{\star}_{\mathtt{SUM}}]_k^* - [\boldsymbol{w}]_k^*\right) f_k g_k \sqrt{\tilde{f}_k}\right|^2}{\hat{B}_k}}_{\geq 0} + 2P\frac{\mathrm{Re}\left\{A_k^*\left([\boldsymbol{w}^{\star}_{\mathtt{SUM}}]_k^* - [\boldsymbol{w}]_k^*\right) f_k g_k \sqrt{\tilde{f}_k}\right\}}{\hat{B}_k} \\
&\leq P\frac{|A_k|^2\left(\hat{B}_k - B_k\right)}{B_k \hat{B}_k} + 2P\frac{|A_k| \cdot \left|[\boldsymbol{w}^{\star}_{\mathtt{SUM}}]_k - [\boldsymbol{w}]_k\right| \cdot |f_k| \cdot |g_k| \sqrt{\tilde{f}_k}}{\hat{B}_k} \\
&= P\frac{-|A_k|^2\left(\left|[\boldsymbol{w}^{\star}_{\mathtt{SUM}}]_k\right|^2 - \left|[\boldsymbol{w}]_k\right|^2\right)|g_k|^2 \tilde{f}_k}{B_k \hat{B}_k} + 2P\frac{|A_k| \cdot \left|[\boldsymbol{w}^{\star}_{\mathtt{SUM}}]_k - [\boldsymbol{w}]_k\right| \cdot |f_k| \cdot |g_k| \sqrt{\tilde{f}_k}}{\hat{B}_k} \\
&\leq P\frac{|A_k|^2\left|\left|[\boldsymbol{w}^{\star}_{\mathtt{SUM}}]_k\right|^2 - \left|[\boldsymbol{w}]_k\right|^2\right| \cdot |g_k|^2 \tilde{f}_k}{B_k \hat{B}_k} + 2P\frac{|A_k| \cdot \left|[\boldsymbol{w}^{\star}_{\mathtt{SUM}}]_k - [\boldsymbol{w}]_k\right| \cdot |f_k| \cdot |g_k| \sqrt{\tilde{f}_k}}{\hat{B}_k} \\
&\leq P|A_k|^2 \left|[\boldsymbol{w}^{\star}_{\mathtt{SUM}}]_k - [\boldsymbol{w}]_k\right| \cdot \left|[\boldsymbol{w}^{\star}_{\mathtt{SUM}}]_k + [\boldsymbol{w}]_k\right| \cdot |g_k|^2 \tilde{f}_k \\
&+ 2P|A_k| \cdot \left|[\boldsymbol{w}^{\star}_{\mathtt{SUM}}]_k - [\boldsymbol{w}]_k\right| \cdot |f_k| \cdot |g_k| \sqrt{\tilde{f}_k}, \\
&\leq 2P|A_k|^2 \left|[\boldsymbol{w}^{\star}_{\mathtt{SUM}}]_k - [\boldsymbol{w}]_k\right| \cdot |g_k|^2 \tilde{f}_k + 2P|A_k| \cdot \left|[\boldsymbol{w}^{\star}_{\mathtt{SUM}}]_k - [\boldsymbol{w}]_k\right| \cdot |f_k| \cdot |g_k| \sqrt{\tilde{f}_k}.
\end{aligned}
\tag{A.3}
$$

For the second last inequality, we have used the bounds $\left||c_1|^2 - |c_2|^2\right| \leq |c_1 - c_2| \times |c_1 + c_2|$ for

$c_1, c_2 \in \mathsf{C}$, and $B_k, \hat{B}_k \geq 1$. Also, (A.3) follows since $|[\boldsymbol{w}^{\star}_{\mathtt{SUM}}]_k + [\boldsymbol{w}]_k| \leq |[\boldsymbol{w}^{\star}_{\mathtt{SUM}}]_k| + |[\boldsymbol{w}]_k| \leq 2$. Now, using the fact that $|f_k| \sqrt{\tilde{f}_k} \leq 1$, we can find an upper bound on the term $|A_k|$ in (A.3) as

$$|A_k| \leq \sum_{n=1}^{N} \left| \left[\boldsymbol{w}^{(k-1)} \right]_n \right| \times |f_n| \, |g_n| \sqrt{\tilde{f}_n} \leq \sum_{n=1}^{N} |f_n| \, |g_n| \sqrt{\tilde{f}_n} \leq \sum_{n=1}^{N} |g_n| \, .$$

Substituting to (A.3), we obtain

$$\begin{aligned}
\overline{\Gamma_k(\boldsymbol{H})} &\leq 2P \left(\sum\nolimits_{n=1}^{N} |g_n| \right)^2 |[\boldsymbol{w}^{\star}_{\mathtt{SUM}}]_k - [\boldsymbol{w}]_k| \times |g_k|^2 \tilde{f}_k \\
&\quad + 2P \left(\sum\nolimits_{n=1}^{N} |g_n| \right) \times |[\boldsymbol{w}^{\star}_{\mathtt{SUM}}]_k - [\boldsymbol{w}]_k| \times |g_k| \\
&\leq 2P \left(\sum\nolimits_{n=1}^{N} |g_n| \right)^2 |[\boldsymbol{w}^{\star}_{\mathtt{SUM}}]_k - [\boldsymbol{w}]_k| \times \sum\nolimits_{n=1}^{N} |g_n|^2 \tilde{f}_n \\
&\quad + 2P \left(\sum\nolimits_{n=1}^{N} |g_n| \right) \times |[\boldsymbol{w}^{\star}_{\mathtt{SUM}}]_k - [\boldsymbol{w}]_k| \times \left(\sum\nolimits_{n=1}^{N} |g_n| \right) \\
&= 2P \left(\sum\nolimits_{n=1}^{N} |g_n| \right)^2 |[\boldsymbol{w}^{\star}_{\mathtt{SUM}}]_k - [\boldsymbol{w}]_k| \left(1 + \sum_{n=1}^{N} \frac{|g_n|^2}{|f_n|^2 + \frac{1}{P}} \right).
\end{aligned}$$

Applying the final inequality to (A.2), we have

$$\begin{aligned}
\overline{\Gamma(\boldsymbol{H})} &\leq 2P \left(\sum_{n=1}^{N} |g_n| \right)^2 \left(1 + \sum_{n=1}^{N} \frac{|g_n|^2}{|f_n|^2 + \frac{1}{P}} \right) \sum_{k=1}^{N} |[\boldsymbol{w}^{\star}_{\mathtt{SUM}}]_k - [\boldsymbol{w}]_k| \\
&\leq 2P \left(\sum_{n=1}^{N} |g_n| \right)^2 \left(1 + \sum_{n=1}^{N} \frac{|g_n|^2}{|f_n|^2 + \frac{1}{P}} \right) \sqrt{N \sum_{k=1}^{N} |[\boldsymbol{w}^{\star}_{\mathtt{SUM}}]_k - [\boldsymbol{w}]_k|^2} \\
&= \underbrace{2\sqrt{N} P \left(\sum_{n=1}^{N} |g_n| \right)^2 \left(1 + \sum_{n=1}^{N} \frac{|g_n|^2}{|f_n|^2 + \frac{1}{P}} \right)}_{\triangleq \Xi} \times ||\boldsymbol{w}^{\star}_{\mathtt{SUM}} - \boldsymbol{w}|| \, . \qquad \text{(A.4)}
\end{aligned}$$

Now, given $\Gamma(\boldsymbol{w}^{\star}_{\mathtt{SUM}}, \boldsymbol{H}) - \alpha > 0$, let $\Pi \triangleq \frac{\Gamma(\boldsymbol{w}^{\star}_{\mathtt{SUM}}, \boldsymbol{H}) - \alpha}{\Xi}$. Then, if $||\boldsymbol{w}^{\star}_{\mathtt{SUM}} - \boldsymbol{w}|| \leq \Pi$, we have

$$\begin{aligned}
\Gamma(\boldsymbol{w}, \boldsymbol{H}) &= \Gamma(\boldsymbol{w}^{\star}_{\mathtt{SUM}}, \boldsymbol{H}) - \overline{\Gamma(\boldsymbol{H})} \geq \Gamma(\boldsymbol{w}^{\star}_{\mathtt{SUM}}, \boldsymbol{H}) - \Xi \times ||\boldsymbol{w}^{\star}_{\mathtt{SUM}} - \boldsymbol{w}|| \\
&\geq \Gamma(\boldsymbol{w}^{\star}_{\mathtt{SUM}}, \boldsymbol{H}) - \Xi \times \Pi = \alpha.
\end{aligned}$$

as desired. To complete the proof, let us verify that $0 < \Pi < 1$. Since $\Gamma\left(\boldsymbol{w}_{\mathtt{SUM}}^{\star}, \boldsymbol{H}\right) - \alpha > 0$ and $\Xi > 0$, we have $\Pi > 0$. Moreover, since $\Gamma\left(\boldsymbol{w}_{\mathtt{SUM}}^{\star}, \boldsymbol{H}\right) = P\sum_{n=1}^{N} \frac{|f_n|^2|g_n|^2}{|f_n|^2+|g_n|^2+\frac{1}{P}} < P\sum_{n=1}^{N}|g_n|^2 < P\left(\sum_{n=1}^{N}|g_n|\right)^2$, we have $\Pi < \frac{\Gamma(\boldsymbol{w}_{\mathtt{SUM}}^{\star}, \boldsymbol{H})}{\Xi} < \frac{1}{2\sqrt{N}\left(1+\sum_{n=1}^{N}\frac{|g_n|^2}{|f_n|^2+\frac{1}{P}}\right)} < 1$. □

Lemma A.2. *Let* $\mathcal{W}_{\mathtt{R}} \triangleq \left\{\boldsymbol{w}_{\mathtt{R}} : \boldsymbol{w}_{\mathtt{R}} \in \mathsf{R}^{2N\times 1}, ||\boldsymbol{w}_{\mathtt{R}}|| = 1\right\}$. *For a fixed* $\boldsymbol{u} \in \mathcal{W}_{\mathtt{R}}$, *a real number* $0 \le t \le 1$, *and a random real vector* $\boldsymbol{v}$ *which is uniformly distributed on the real unit sphere* $\mathcal{W}_{\mathtt{R}}$, *we have*

$$Pr\left\{\boldsymbol{u}^{\top}\boldsymbol{v} \ge t\right\} = \frac{1}{2} I_{1-t^2}\left(\frac{2N-1}{2}, \frac{1}{2}\right),$$

where $I_z(a,b) = \frac{1}{\beta(a,b)}\int_0^z x^{a-1}(1-x)^{b-1}\mathrm{d}x$ *is the regularized incomplete beta function,* $\beta(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1}\mathrm{d}x$ *is the beta function [14].*

Proof. Similar to [21, Eqs.(23)-(24)], we have $\Pr\left\{\boldsymbol{u}^{\top}\boldsymbol{v} \ge t\right\} = \frac{S_{2N,t,\mathrm{cap}}}{S_{2N}}$, where $S_{2N,t,\mathrm{cap}}$ is the surface area of the spherical cap formed by the intersection of the subspace $\boldsymbol{u}^{\top}\boldsymbol{v} \ge t$ and the real unit hyper-sphere $\mathcal{W}_{\mathtt{R}}$. From [59], we obtain $S_{2N} = \frac{2\pi^N}{(N-1)!}$ and $S_{2N,t,\mathrm{cap}} = \frac{\pi^N}{(N-1)!} I_{1-t^2}\left(\frac{2N-1}{2}, \frac{1}{2}\right)$. Then, Lemma A.2 is obtained by dividing $S_{2N,t,\mathrm{cap}}$ by S_{2N}. □

We are now ready to prove the theorem. Since a full-CSI system provides the minimum possible outage probability, the inequality

$$\mathtt{Out}\left(\mathtt{Full}_{\mathtt{SUM}}\right) \le \mathtt{Out}\left(\mathrm{VLQ}_{\mathtt{SUM}}\right) \tag{A.5}$$

holds. It is thus sufficient to prove the reverse inequality $\mathtt{Out}\left(\mathrm{VLQ}_{\mathtt{SUM}}\right) \le \mathtt{Out}\left(\mathtt{Full}_{\mathtt{SUM}}\right)$. For this purpose, we need the following definitions. Let

$$\mathcal{H}_1 \triangleq \left\{\boldsymbol{H} \in \mathsf{C}^{2N\times 1} : \Gamma\left(\boldsymbol{w}_{\mathtt{SUM}}^{\star}, \boldsymbol{H}\right) < \alpha\right\}, \tag{A.6}$$

$$\mathcal{H}_2 \triangleq \left\{\boldsymbol{H} \in \mathsf{C}^{2N\times 1} : \Gamma\left(\boldsymbol{w}_{\mathtt{SUM}}^{\star}, \boldsymbol{H}\right) = \alpha\right\}, \tag{A.7}$$

$$\mathcal{H}_3 \triangleq \left\{\boldsymbol{H} \in \mathsf{C}^{2N\times 1} : \Gamma\left(\boldsymbol{w}_{\mathtt{SUM}}^{\star}, \boldsymbol{H}\right) > \alpha\right\}. \tag{A.8}$$

According to (2.5), we have $\mathtt{Out}\left(\mathtt{Full}_{\mathtt{SUM}}\right) = \Pr\left\{\boldsymbol{H} \in \mathcal{H}_1\right\}$. Also, let

$$\mathcal{H} \triangleq \left\{\boldsymbol{H} \in \mathsf{C}^{2N\times 1} : \Gamma\left(\boldsymbol{w}_i, \boldsymbol{H}\right) < \alpha, \forall i \in \mathsf{N}\right\}, \tag{A.9}$$

Note that we have omitted the dependency of $\mathcal{H}$ on the realization of the random codebook $\{\boldsymbol{w}_i\}_{\mathsf{N}}$ for brevity. Using the definition in (2.11), we have

$$\begin{aligned}
\mathtt{Out}\left(\mathrm{VLQ}_{\mathtt{SUM}}\right) &= \mathtt{E}_{\boldsymbol{H}}\mathtt{E}_{\{\boldsymbol{w}_i\}_{\mathsf{N}}}\mathbf{1}\left\{\boldsymbol{H} \in \mathcal{H}\right\} && \text{(A.10)}\\
&= \mathtt{E}_{\boldsymbol{H}}\mathtt{E}_{\{\boldsymbol{w}_i\}_{\mathsf{N}}}\big(\mathbf{1}\left\{\boldsymbol{H} \in \mathcal{H} \cap \mathcal{H}_1\right\} + \mathbf{1}\left\{\boldsymbol{H} \in \mathcal{H} \cap \mathcal{H}_2\right\} + \mathbf{1}\left\{\boldsymbol{H} \in \mathcal{H} \cap \mathcal{H}_3\right\}\big) && \text{(A.11)}\\
&\leq \mathtt{E}_{\boldsymbol{H}}\mathtt{E}_{\{\boldsymbol{w}_i\}_{\mathsf{N}}}\big(\mathbf{1}\left\{\boldsymbol{H} \in \mathcal{H}_1\right\} + \mathbf{1}\left\{\boldsymbol{H} \in \mathcal{H}_2\right\} + \mathbf{1}\left\{\boldsymbol{H} \in \mathcal{H} \cap \mathcal{H}_3\right\}\big) && \text{(A.12)}\\
&= \Pr\left\{\boldsymbol{H} \in \mathcal{H}_1\right\} + \Pr\left\{\boldsymbol{H} \in \mathcal{H}_2\right\} + \mathtt{E}_{\boldsymbol{H}}\mathtt{E}_{\{\boldsymbol{w}_i\}_{\mathsf{N}}}\mathbf{1}\left\{\boldsymbol{H} \in \mathcal{H} \cap \mathcal{H}_3\right\} && \text{(A.13)}\\
&= \mathtt{Out}\left(\mathtt{Full}_{\mathtt{SUM}}\right) + \Pr\left\{\Gamma(\boldsymbol{w}^{\star}_{\mathtt{SUM}}, \boldsymbol{H}) = \alpha\right\} + \mathtt{E}_{\boldsymbol{H}}\mathtt{E}_{\{\boldsymbol{w}_i\}_{\mathsf{N}}}\mathbf{1}\left\{\boldsymbol{H} \in \mathcal{H} \cap \mathcal{H}_3\right\} && \text{(A.14)}\\
&= \mathtt{Out}\left(\mathtt{Full}_{\mathtt{SUM}}\right) + \mathtt{E}_{\boldsymbol{H}}\mathtt{E}_{\{\boldsymbol{w}_i\}_{\mathsf{N}}}\mathbf{1}\left\{\boldsymbol{H} \in \mathcal{H} \cap \mathcal{H}_3\right\}, && \text{(A.15)}
\end{aligned}$$

where (A.11) follows since $\mathcal{H}_1, \mathcal{H}_2, \mathcal{H}_3$ are disjoint sets that cover $\mathsf{C}^{2N\times 1}$, (A.12) follows as $\mathbf{1}\{\mathbf{H} \in A \cap B\} \leq \mathbf{1}\{\mathbf{H} \in A\}$ for any $A, B \subset \mathsf{C}^{2N\times 1}$, and (A.13) follows since

$$\mathtt{E}_{\boldsymbol{H}}\mathtt{E}_{\{\boldsymbol{w}_i\}_{\mathsf{N}}}\mathbf{1}\left\{\boldsymbol{H} \in A\right\} = \mathtt{E}_{\{\boldsymbol{w}_i\}_{\mathsf{N}}}\mathtt{E}_{\boldsymbol{H}}\mathbf{1}\left\{\boldsymbol{H} \in A\right\} = \mathtt{E}_{\{\boldsymbol{w}_i\}_{\mathsf{N}}}\Pr\left\{\boldsymbol{H} \in A\right\} = \Pr\left\{\boldsymbol{H} \in A\right\},$$

for every $A \subset \mathsf{C}^{2N\times 1}$. Equality in (A.14) is by the definitions of $\mathcal{H}_1$ and $\mathcal{H}_2$. Finally, (A.15) follows since $\Gamma(\boldsymbol{w}^{\star}_{\mathtt{SUM}}, \boldsymbol{H})$ is a continuous random variable, and the probability that a continuous random variable assumes a specific real value is zero.

In light of (A.5) and (A.15), in order to conclude the proof of the theorem, we need to prove that the equality $\mathtt{E}_{\boldsymbol{H}}\mathtt{E}_{\{\boldsymbol{w}_i\}_{\mathsf{N}}}\mathbf{1}\left\{\boldsymbol{H} \in \mathcal{H} \cap \mathcal{H}_3\right\} = 0$ holds. In fact, we shall prove the stronger statement that $\mathtt{E}_{\{\boldsymbol{w}_i\}_{\mathsf{N}}}\mathbf{1}\left\{\boldsymbol{H} \in \mathcal{H}\right\} = 0$ for every $\boldsymbol{H} \in \mathcal{H}_3$. Assume, for the sake of

contradiction, that there exists $\tilde{\boldsymbol{H}} \in \mathcal{H}_3$, such that $\mathtt{E}_{\{\boldsymbol{w}_i\}_{\mathsf{N}}} \mathbf{1}\left\{\tilde{\boldsymbol{H}} \in \mathcal{H}\right\} = \varepsilon > 0$. We have

$$\begin{aligned}\mathrm{E}_{\{\boldsymbol{w}_i\}_{\mathsf{N}}} \mathbf{1}\left\{\tilde{\boldsymbol{H}} \in \mathcal{H}\right\} &= \Pr\left\{\Gamma\left(\boldsymbol{w}_i, \tilde{\boldsymbol{H}}\right) < \alpha, \forall i \in \mathsf{N}\right\} \\ &\leq \Pr\left\{\Gamma\left(\boldsymbol{w}_i, \tilde{\boldsymbol{H}}\right) < \alpha, 0 \leq i \leq K-1\right\} = \left[\Pr\left\{\Gamma\left(\boldsymbol{w}_i, \tilde{\boldsymbol{H}}\right) < \alpha\right\}\right]^K,\end{aligned} \tag{A.16}$$

where $K \geq 1$ is an arbitrary natural number. For the equality in (A.16), note that $\boldsymbol{w}_i$, $i \in \mathsf{N}$ are mutually independent and $\Gamma(\cdot,\cdot)$ is a deterministic function. Therefore, $\Gamma\left(\boldsymbol{w}_0, \tilde{\boldsymbol{H}}\right), \ldots, \Gamma\left(\boldsymbol{w}_{K-1}, \tilde{\boldsymbol{H}}\right)$ are mutually independent given $\tilde{\boldsymbol{H}}$, which proves the equality in (A.16).

Now, using Lemma A.1, for any given $\tilde{\boldsymbol{H}}$, we obtain

$$\Pr\left\{\Gamma\left(\boldsymbol{w}_i, \tilde{\boldsymbol{H}}\right) \geq \alpha\right\} \geq \Pr\left\{||\boldsymbol{w} - \boldsymbol{w}^{\star}_{\mathtt{SUM}}|| \leq \Pi\right\} = \Pr\left\{\mathrm{Real}\left\{\boldsymbol{w}^{\dagger}\boldsymbol{w}^{\star}_{\mathtt{SUM}}\right\} \geq 1 - \frac{\Pi^2}{2}\right\}. \tag{A.17}$$

Note that $\mathrm{Real}\left\{\boldsymbol{w}^{\dagger}\boldsymbol{w}^{\star}_{\mathtt{SUM}}\right\} = \left[\mathrm{Real}\left\{\boldsymbol{w}\right\}; \mathrm{Imag}\left\{\boldsymbol{w}\right\}\right]^{\top} \left[\mathrm{Real}\left\{\boldsymbol{w}^{\star}_{\mathtt{SUM}}\right\}; \mathrm{Imag}\left\{\boldsymbol{w}^{\star}_{\mathtt{SUM}}\right\}\right]$. Since $\boldsymbol{w}$ is uniformly distributed on $\mathcal{W}_{\mathtt{SUM}}$, the random vector $\left[\mathrm{Real}\left\{\boldsymbol{w}\right\}; \mathrm{Imag}\left\{\boldsymbol{w}\right\}\right] \in \mathsf{R}^{2N\times 1}$ is uniformly distributed on the unit real sphere $\mathcal{W}_{\mathtt{R}}$ (For a proof, note that $\boldsymbol{w}$ is equal in distribution to $\frac{\boldsymbol{w}_{\mathsf{R}} + j * \boldsymbol{w}_{\mathsf{I}}}{\sqrt{||\boldsymbol{w}_{\mathsf{R}}||^2 + ||\boldsymbol{w}_{\mathsf{I}}||^2}}$, where $\boldsymbol{w}_{\mathsf{R}}, \boldsymbol{w}_{\mathsf{I}} \in \mathbb{N}\left(\mathbf{0}_{N\times 1}, \frac{1}{2}\boldsymbol{I}_{N\times N}\right)$ are independent[21]. Thus, $\left[\mathrm{Real}\left\{\boldsymbol{w}\right\}; \mathrm{Imag}\left\{\boldsymbol{w}\right\}\right] = \frac{[\boldsymbol{w}_{\mathsf{R}}; \boldsymbol{w}_{\mathsf{I}}]}{\sqrt{||\boldsymbol{w}_{\mathsf{R}}||^2 + ||\boldsymbol{w}_{\mathsf{I}}||^2}}$ is uniformly distributed on $\mathcal{W}_{\mathtt{R}}$.). Applying Lemma A.2 to (A.17), we have

$$\begin{aligned}\Pr\left\{\Gamma\left(\boldsymbol{w}_i, \tilde{\boldsymbol{H}}\right) < \alpha\right\} &\leq 1 - \Pr\left\{\mathrm{Real}\left\{\boldsymbol{w}^{\dagger}\boldsymbol{w}^{\star}_{\mathtt{SUM}}\right\} \geq 1 - \frac{\Pi^2}{2}\right\} \\ &= 1 - \frac{1}{2} I_{1-\left(1-\frac{\Pi^2}{2}\right)^2}\left(\frac{2N-1}{2}, \frac{1}{2}\right) \\ &= 1 - \frac{\int_0^{1-\left(1-\frac{\Pi^2}{2}\right)^2} x^{\frac{2N-3}{2}} \overbrace{(1-x)^{-\frac{1}{2}}}^{\geq 1} \mathrm{d}x}{2 \times \beta\left(\frac{2N-1}{2}, \frac{1}{2}\right)}\end{aligned}$$

$$\leq 1 - \frac{\int_0^{1-\left(1-\frac{\Pi^2}{2}\right)^2} x^{\frac{2N-3}{2}} \mathrm{d}x}{2 \times \beta\left(\frac{2N-1}{2}, \frac{1}{2}\right)}$$
$$= 1 - \frac{\left(1-\left(1-\frac{\Pi^2}{2}\right)^2\right)^{\frac{2N-1}{2}}}{(2N-1)\times\beta\left(\frac{2N-1}{2},\frac{1}{2}\right)} \triangleq \Phi < 1. \tag{A.18}$$

Letting $K \triangleq \lceil \log_\Phi \varepsilon \rceil + 1$ in (A.16), we have $\mathtt{E}_{\{\boldsymbol{w}_i\}_\mathtt{N}} \mathbf{1}\left\{\tilde{\boldsymbol{H}} \in \mathcal{H}\right\} \leq \Phi^K = \Phi^{\lceil \log_\Phi \varepsilon \rceil + 1} < \Phi^{\log_\Phi \varepsilon} = \varepsilon$, which contradicts the assumption that $\mathtt{E}_{\{\boldsymbol{w}_i\}_\mathtt{N}} \mathbf{1}\left\{\tilde{\boldsymbol{H}} \in \mathcal{H}\right\} = \varepsilon$. Thus, $\mathtt{E}_{\{\boldsymbol{w}_i\}_\mathtt{N}} \mathbf{1}\left\{\boldsymbol{H} \in \mathcal{H}\right\} = 0$ for any $\boldsymbol{H} \in \mathcal{H}_3$, and this concludes the proof. □

A.2 Proof of Theorem 2.2

Proof. We need the following lemma to prove Theorem 2.2.

Lemma A.3. *Let* $\Gamma_n \triangleq \frac{|f_n|^2|g_n|^2}{|f_n|^2+|g_n|^2+\frac{1}{P}}$. *Then, for* $N \geq 2$, *the probability density function (PDF) of* $\frac{\Gamma(\boldsymbol{w}^\star,\boldsymbol{H})}{P} = \sum_{n=1}^N \Gamma_n$ *is upper-bounded by*

$$f_{\sum_{n=1}^N \Gamma_n}(x) \leq$$
$$e^{-\frac{x}{\sigma_{g_N}^2}} D_N \left[x^{N-1} + \frac{1}{P^{N-1}} + \frac{1}{P^N} + \mathbf{1}\{N \geq 3\} \times \sum_{m=1}^{N-2}\left(\frac{x^m}{P^{N-m-1}} + \frac{x^m}{P^{N-m}}\right)\right], \tag{A.19}$$

where $D_N > 0$ *is a constant that is independent of* P. □

Proof. According to [8, (5)], the cumulative distribution function (CDF) of Γ_n is $\Pr\{\Gamma_n < x\} = 1 - \frac{e^{-\frac{x}{\sigma_{g_n}^2}-\frac{x}{\sigma_{f_n}^2}}}{\sigma_{g_n}^2}\int_0^\infty e^{-\frac{y}{\sigma_{g_n}^2}-\frac{x^2+\frac{x}{P}}{y\times\sigma_{f_n}^2}}\mathrm{d}y$. By taking the derivative, the PDF of Γ_n is therefore

$$f_{\Gamma_n}(x) = e^{-\frac{x}{\sigma_{g_n}^2}-\frac{x}{\sigma_{f_n}^2}} \frac{\frac{1}{\sigma_{g_n}^2}+\frac{1}{\sigma_{f_n}^2}}{\sigma_{g_n}^2} \int_0^\infty e^{-\frac{y}{\sigma_{g_n}^2}-\frac{x^2+\frac{x}{P}}{y\times\sigma_{f_n}^2}}\mathrm{d}y$$
$$+ \frac{e^{-\frac{x}{\sigma_{g_n}^2}-\frac{x}{\sigma_{f_n}^2}}}{\sigma_{g_n}^2\sigma_{f_n}^2}\left(2x+\frac{1}{P}\right)\int_0^\infty \frac{1}{y} e^{-\frac{y}{\sigma_{g_n}^2}-\frac{x^2+\frac{x}{P}}{y\times\sigma_{f_n}^2}}\mathrm{d}y \tag{A.20}$$

$$= 2e^{-\frac{x}{\sigma_{g_n}^2}-\frac{x}{\sigma_{f_n}^2}} \frac{\frac{1}{\sigma_{g_n}^2}+\frac{1}{\sigma_{f_n}^2}}{\sigma_{g_n}^2} \sqrt{\frac{\sigma_{g_n}^2\left(x^2+\frac{x}{P}\right)}{\sigma_{f_n}^2}} \boldsymbol{K}_1\left(2\sqrt{\frac{x^2+\frac{x}{P}}{\sigma_{g_n}^2\sigma_{f_n}^2}}\right)$$

$$+ 2\frac{e^{-\frac{x}{\sigma_{g_n}^2}-\frac{x}{\sigma_{f_n}^2}}}{\sigma_{g_n}^2\sigma_{f_n}^2}\left(2x+\frac{1}{P}\right)\boldsymbol{K}_0\left(2\sqrt{\frac{x^2+\frac{x}{P}}{\sigma_{g_n}^2\sigma_{f_n}^2}}\right) \tag{A.21}$$

$$\leq C_2 \times e^{-\frac{x}{\sigma_{g_N}^2}}\left(\sqrt{x^2+\frac{x}{P}}\boldsymbol{K}_1\left(2\sqrt{\frac{x^2+\frac{x}{P}}{\sigma_{g_n}^2\sigma_{f_n}^2}}\right)+\left(x+\frac{1}{P}\right)\boldsymbol{K}_0\left(2\sqrt{\frac{x^2+\frac{x}{P}}{\sigma_{g_n}^2\sigma_{f_n}^2}}\right)\right), \tag{A.22}$$

$$\leq C_3 \times e^{-\frac{x}{\sigma_{g_N}^2}}\left(1+\left(x+\frac{1}{P}\right)\boldsymbol{K}_0\left(\frac{2x}{\sqrt{\sigma_{g_n}^2\sigma_{f_n}^2}}\right)\right), \tag{A.23}$$

where C_2 and C_3 are constants that only depends on $\sigma_{f_n}^2$ and $\sigma_{g_n}^2$. Moreover, (A.21) follows since $\int_0^\infty x^{v-1}e^{-\frac{\beta}{x}-\gamma x}\mathrm{d}x = 2\left(\frac{\beta}{\gamma}\right)^{\frac{v}{2}}\boldsymbol{K}_v\left(2\sqrt{\beta\gamma}\right)$ [14, (3.471.9)], (A.22) follows since $e^{-\frac{x}{\sigma_{f_n}^2}} \leq 1$ and $e^{-\frac{x}{\sigma_{g_n}^2}} \leq e^{-\frac{x}{\sigma_{g_N}^2}}$ due to $\sigma_{g_1}^2 \leq \sigma_{g_2}^2 \leq \ldots \leq \sigma_{g_N}^2$ in the assumption, and (A.23) follows from since $\boldsymbol{K}_1(x) \leq \frac{1}{x}$ [10, (25)] and $\boldsymbol{K}_0(\cdot)$ is a decreasing function [14, (3.471.9)].

We now proceed by induction. For $N = 2$, the PDF of $\Gamma_1 + \Gamma_2$ is the convolution of $f_{\Gamma_1}(x)$ and $f_{\Gamma_2}(x)$. With the bounds on $f_{\Gamma_1}(x)$ and $f_{\Gamma_2}(x)$ in (A.23), the upper bound on $f_{\Gamma_1+\Gamma_2}(x)$ is given by

$$\begin{aligned} f_{\Gamma_1+\Gamma_2}(x) &= \int_0^x f_{\Gamma_1}(r)f_{\Gamma_2}(x-r)\mathrm{d}r \\ &\leq \int_0^x C_3 e^{-\frac{r}{\sigma_{g_N}^2}}\left(1+\left(r+\frac{1}{P}\right)\boldsymbol{K}_0\left(\frac{2r}{\sqrt{\sigma_{g_n}^2\sigma_{f_n}^2}}\right)\right) \\ &\quad \times C_3 e^{-\frac{x-r}{\sigma_{g_N}^2}}\left(1+\left(x-r+\frac{1}{P}\right)\boldsymbol{K}_0\left(\frac{2x-2r}{\sqrt{\sigma_{g_n}^2\sigma_{f_n}^2}}\right)\right)\mathrm{d}r. \end{aligned}$$

Using $\int_0^\infty \boldsymbol{K}_0(ax)\mathrm{d}x = \frac{\pi}{2a}$, $\int_0^\infty \boldsymbol{K}_0^2(ax)\mathrm{d}x = \frac{\pi^2}{4a}$ for $a > 0$ [14, (6.511.12)-(6.511.13)], $\boldsymbol{K}_0(x) \leq$

$\frac{2}{x}$ for $x > 0$ [10, (27)] and after some calculations, we can obtain

$$f_{\Gamma_1+\Gamma_2}(x) \leq C_4 \left(x e^{-\frac{x}{\sigma^2_{g_N}}} + \frac{e^{-\frac{x}{\sigma^2_{g_N}}}}{P} + \frac{e^{-\frac{x}{\sigma^2_{g_N}}}}{P^2} \right),$$

for some constant C_4 that depends only on the channel variances. In the inductive step, by substituting the upper bound on $f_{\Gamma_{k+1}}(x)$ in (A.23) and the upper bound on $f_{\sum_{i=1}^{k}\Gamma_i}(x)$ in (A.19) when $N = k$ into $f_{\sum_{i=1}^{k+1}\Gamma_i}(x) = \int_0^x f_{\sum_{i=1}^{k}\Gamma_i}(r) f_{\Gamma_{k+1}}(x-r)\mathrm{d}r$, we obtain an upper bound on $f_{\sum_{i=1}^{k+1}\Gamma_i}(x)$ as

$$\begin{aligned}
&f_{\sum_{i=1}^{k+1}\Gamma_i}(x) \\
&\leq \int_0^x e^{-\frac{r}{\sigma^2_{g_N}}} D_k \left[\underbrace{r^{k-1}}_{\leq x^{k-1}} + \underbrace{\mathbf{1}\{k \geq 3\}}_{\leq \mathbf{1}\{k+1\geq 3\}} \times \sum_{m=1}^{k-2} \left(\frac{\overbrace{r^m}^{\leq x^m}}{P^{k-m-1}} + \frac{\overbrace{r^m}^{\leq x^m}}{P^{k-m}} \right) + \left(\frac{1}{P^{k-1}} + \frac{1}{P^k} \right) \right] \\
&\times C_3 \times e^{-\frac{x-r}{\sigma^2_{g_N}}} \left(1 + \left(\underbrace{x-r}_{\leq x} + \frac{1}{P} \right) \boldsymbol{K}_0 \left(\frac{2x-2r}{\sqrt{\sigma^2_{g_{k+1}}\sigma^2_{f_{k+1}}}} \right) \right) \\
&\leq e^{-\frac{x}{\sigma^2_{g_N}}} C_5 \left[x^{k-1} + \mathbf{1}\{k+1 \geq 3\} \times \sum_{m=1}^{k-2} \left(\frac{x^m}{P^{k-m-1}} + \frac{x^m}{P^{k-m}} \right) + \left(\frac{1}{P^{k-1}} + \frac{1}{P^k} \right) \right] \\
&\times \int_0^x \left[1 + \left(x + \frac{1}{P} \right) \boldsymbol{K}_0 \left(\frac{2x-2r}{\sqrt{\sigma_{g^2_{k+1}}\sigma_{f^2_{k+1}}}} \right) \right] \mathrm{d}r \\
&= e^{-\frac{x}{\sigma^2_{g_N}}} C_5 \left[x^{k-1} + \mathbf{1}\{k+1 \geq 3\} \times \sum_{m=1}^{k-2} \left(\frac{x^m}{P^{k-m-1}} + \frac{x^m}{P^{k-m}} \right) + \left(\frac{1}{P^{k-1}} + \frac{1}{P^k} \right) \right] \\
&\times \left[x + C_6 \left(x + \frac{1}{P} \right) \right],
\end{aligned}$$

where C_5 and C_6 are constants depending only on the channel variances, and the last equality follows since $\int_0^\infty \boldsymbol{K}_0(ax)\mathrm{d}x = \frac{\pi}{2a}$ for $a > 0$. After trivial mathematical manipulations, we can obtain the upper bound on $f_{\sum_{i=1}^{k+1}\Gamma_i}(x)$ in (A.19) where $N = k+1$. By induction, Lemma A.3 stands for any $N \geq 2$. □

We are now ready to prove the theorem. First, we recall from (2.12) that

$$\mathtt{FR}\left(\mathrm{VLQ}_{\mathtt{SUM}}\right)=\sum_{i=0}^{\infty}\lfloor\log_2(i+2)\rfloor\times\mathtt{E}_{\boldsymbol{H}}\mathtt{E}_{\{\boldsymbol{w}_i\}_{\mathsf{N}}}\mathbf{1}\left\{\boldsymbol{H}\in\mathcal{S}_i\right\},\tag{A.24}$$

where, as defined in (2.10),

$$\mathcal{S}_i\triangleq\begin{cases}\left\{\boldsymbol{H}:\Gamma\left(\boldsymbol{w}_0,\boldsymbol{H}\right)\geq\alpha\right\}\cup\bigcap_{i\in\mathsf{N}}\left\{\boldsymbol{H}:\Gamma\left(\boldsymbol{w}_i,\boldsymbol{H}\right)<\alpha\right\}, & i=0,\\ \left\{\boldsymbol{H}:\Gamma\left(\boldsymbol{w}_i,\boldsymbol{H}\right)\geq\alpha\right\}\cap\bigcap_{k=0}^{i-1}\left\{\boldsymbol{H}:\Gamma\left(\boldsymbol{w}_k,\boldsymbol{H}\right)<\alpha\right\}, & i\in\mathsf{N}-\{0\}.\end{cases}\tag{A.25}$$

Let $p\triangleq\mathtt{E}_{\boldsymbol{w}}\mathbf{1}\left\{\Gamma\left(\boldsymbol{w},\boldsymbol{H}\right)<\alpha\right\}$, where $\boldsymbol{w}$ is uniformly distributed on the complex unit sphere $\mathcal{W}_{\mathtt{SUM}}$. We have omitted the dependency of p on the channel state $\boldsymbol{H}$ for brevity. For $i\geq1$, we have $\mathtt{E}_{\{\boldsymbol{w}_i\}_{\mathsf{N}}}\mathbf{1}\left\{\boldsymbol{H}\in\mathcal{S}_i\right\}=p^i(1-p)$ if $\Gamma\left(\boldsymbol{w}^{\star}_{\mathtt{SUM}},\boldsymbol{H}\right)\geq\alpha$, and $\mathtt{E}_{\{\boldsymbol{w}_i\}_{\mathsf{N}}}\mathbf{1}\left\{\boldsymbol{H}\in\mathcal{S}_i\right\}=0$ if $\Gamma\left(\boldsymbol{w}^{\star}_{\mathtt{SUM}},\boldsymbol{H}\right)<\alpha$. Also, since

$$\lfloor\log_2(i+2)\rfloor\leq\log_2(i+2)\leq\log_2(2i+2)=1+\log_2(i+1),$$

the quantity $\mathtt{FR}\left(\mathrm{VLQ}_{\mathtt{SUM}}\right)$ can be upper-bounded by

$$\begin{aligned}\mathtt{FR}\left(\mathrm{VLQ}_{\mathtt{SUM}}\right)&\leq\sum_{i=0}^{\infty}\mathtt{E}_{\boldsymbol{H}}\mathtt{E}_{\{\boldsymbol{w}_i\}_{\mathsf{N}}}\left[\mathbf{1}\left\{\boldsymbol{H}\in\mathcal{S}_i\right\}\right]+\sum_{i=0}^{\infty}\log_2(i+1)\times\mathtt{E}_{\boldsymbol{H}}\mathtt{E}_{\{\boldsymbol{w}_i\}_{\mathsf{N}}}\left[\mathbf{1}\left\{\boldsymbol{H}\in\mathcal{S}_i\right\}\right]\\&=1+\sum_{i=1}^{\infty}\log_2(i+1)\times\mathtt{E}_{\boldsymbol{H}}\mathtt{E}_{\{\boldsymbol{w}_i\}_{\mathsf{N}}}\left[\mathbf{1}\left\{\boldsymbol{H}\in\mathcal{S}_i\right\}\right]\\&=1+\int_{\hat{\mathcal{H}}}\left[\sum_{i=1}^{\infty}p^i(1-p)\log_2(i+1)\right]f_{\boldsymbol{H}}\left(\boldsymbol{H}\right)\mathrm{d}\boldsymbol{H},\end{aligned}\tag{A.26}$$

where $\hat{\mathcal{H}}\triangleq\left\{\boldsymbol{H}\in\mathsf{C}^{2N\times1}:\Gamma\left(\boldsymbol{w}^{\star}_{\mathtt{SUM}},\boldsymbol{H}\right)\geq\alpha\right\}$. For the sum inside the square brackets, we have

$$\begin{aligned}\sum_{i=1}^{\infty}p^i(1-p)\log_2(i+1)&\leq p(1-p)+\left(\frac{6}{\log2}+2\right)p^2+\frac{2}{\log2}p^2\log\frac{1}{1-p}\\&\leq C_7+C_8\log\frac{1}{1-p},\end{aligned}\tag{A.27}$$

where $C_7 \triangleq \frac{6}{\log 2} + 3$, $C_8 \triangleq \frac{2}{\log 2}$, and the first inequality follows from [30, Lemma 1]. Thus,

$$\begin{aligned}
\mathtt{FR}\left(\mathrm{VLQ}_{\mathtt{SUM}}\right) &\leq 1 + \int_{\hat{\mathcal{H}}} \left(C_7 + C_8 \log \frac{1}{1-p} \right) f_{\boldsymbol{H}}\left(\boldsymbol{H}\right) \mathrm{d}\boldsymbol{H} && \text{(A.28)} \\
&= 1 + C_7 \underbrace{\int_{\hat{\mathcal{H}}} f_{\boldsymbol{H}}\left(\boldsymbol{H}\right) \mathrm{d}\boldsymbol{H}}_{\leq 1} + C_8 \int_{\hat{\mathcal{H}}} \log \frac{1}{1-p} f_{\boldsymbol{H}}\left(\boldsymbol{H}\right) \mathrm{d}\boldsymbol{H} \\
&\leq C_9 + C_8 \underbrace{\int_{\hat{\mathcal{H}}} \log \frac{1}{1-p} f_{\boldsymbol{H}}\left(\boldsymbol{H}\right) \mathrm{d}\boldsymbol{H}}_{\triangleq I}, && \text{(A.29)}
\end{aligned}$$

where $C_9 \triangleq 1 + C_7$. We now find an upper bound on the integral I in the final inequality. First, according to (A.18) in the proof of Theorem 2.1 in Appendix A.1, we have

$$p \leq 1 - \frac{\left(1 - \left(1 - \frac{\Pi^2}{2}\right)^2\right)^{\frac{2N-1}{2}}}{(2N-1)\beta\left(\frac{2N-1}{2}, \frac{1}{2}\right)}, \tag{A.30}$$

where Π is as defined in (A.1). Substituting to (A.29), we obtain

$$\begin{aligned}
I \leq & \int_{\hat{\mathcal{H}}} \log \left((2N-1)\beta\left(\frac{2N-1}{2}, \frac{1}{2}\right) \right) f_{\boldsymbol{H}}\left(\boldsymbol{H}\right) \mathrm{d}\boldsymbol{H} \\
& + \frac{2N-1}{2} \int_{\hat{\mathcal{H}}} \log \frac{1}{1 - \left(1 - \frac{\Pi^2}{2}\right)^2} f_{\boldsymbol{H}}\left(\boldsymbol{H}\right) \mathrm{d}\boldsymbol{H}. && \text{(A.31)}
\end{aligned}$$

As shown in [60], we have $\beta\left(\frac{2N-1}{2}, \frac{1}{2}\right) \leq \frac{4}{2N-1} + 1$. Moreover, since $1 - \left(1 - \frac{\Pi^2}{2}\right)^2 \geq \frac{\Pi^2}{2}$ and $\int_{\hat{\mathcal{H}}} f_{\boldsymbol{H}}\left(\boldsymbol{H}\right) \mathrm{d}\boldsymbol{H} \leq 1$, the integral I can be further bounded by

$$\begin{aligned}
I &\leq \log(2N+3) + \frac{2N-1}{2} \int_{\hat{\mathcal{H}}} \overbrace{\log \frac{2}{\Pi^2}}^{=\log 2 + 2\log\frac{1}{\Pi}} f_{\boldsymbol{H}}\left(\boldsymbol{H}\right) \mathrm{d}\boldsymbol{H} \\
&\leq C_{10} + C_{11} \int_{\hat{\mathcal{H}}} \left(\log \frac{1}{\Pi} \right) f_{\boldsymbol{H}}\left(\boldsymbol{H}\right) \mathrm{d}\boldsymbol{H}, && \text{(A.32)}
\end{aligned}$$

where $C_{10} \triangleq \log(2N+3) + (2N-1)\frac{\log 2}{2}$ and $C_{11} \triangleq 2N - 1$. Now, using the inequalities $\left(\sum_{n=1}^{N} |g_n|\right)^2 \leq N\left(\sum_{n=1}^{N} |g_n|^2\right) \leq N \exp\left(\sum_{n=1}^{N} |g_n|^2\right)$ and $\sum_{n=1}^{N} \frac{|g_n|^2}{|f_n|^2 + \frac{1}{P}} \leq \sum_{n=1}^{N} \frac{|g_n|^2}{|f_n|^2} \leq$

$N \max_n \frac{|g_n|^2}{|f_n|^2}$ on (A.1), we obtain

$$\Pi \geq \frac{\Gamma\left(\boldsymbol{w}^{\star}_{\mathtt{SUM}}, \boldsymbol{H}\right) - \alpha}{2N^{\frac{3}{2}} P \exp\left(\sum_{n=1}^{N} |g_n|^2\right) \left(1 + N \max_n \frac{|g_n|^2}{|f_n|^2}\right)}. \tag{A.33}$$

Substituting to (A.32) and then using the inequality $\int_{\hat{\mathcal{H}}} f_{\boldsymbol{H}}(\boldsymbol{H})\, \mathrm{d}\boldsymbol{H} \leq 1$ one more time, we obtain

$$\begin{aligned} I \leq C_{12} + C_{11} \underbrace{\int_{\hat{\mathcal{H}}} \left(\sum_{n=1}^{N} |g_n|^2\right) f_{\boldsymbol{H}}(\boldsymbol{H})\, \mathrm{d}\boldsymbol{H}}_{=I_1} + C_{11} \underbrace{\int_{\hat{\mathcal{H}}} \log\left(1 + N \max_n \frac{|g_n|^2}{|f_n|^2}\right) f_{\boldsymbol{H}}(\boldsymbol{H})\, \mathrm{d}\boldsymbol{H}}_{=I_2} \\ + C_{11} \underbrace{\int_{\hat{\mathcal{H}}} \log \frac{P}{\Gamma\left(\boldsymbol{w}^{\star}_{\mathtt{SUM}}, \boldsymbol{H}\right) - \alpha} f_{\boldsymbol{H}}(\boldsymbol{H})\, \mathrm{d}\boldsymbol{H}}_{=I_3}, \end{aligned} \tag{A.34}$$

where $C_{12} \triangleq C_{10} + C_{11} \log\left(2N^{\frac{3}{2}}\right)$. We now find upper bounds on I_k for $k = 1, \ldots, 3$. For I_1, we have

$$I_1 \leq \sum_{n=1}^{N} \int_0^{\infty} x f_{|g_n|^2}(x)\, \mathrm{d}x = \sum_{n=1}^{N} \int_0^{\infty} \frac{x}{\sigma^2_{g_n}} e^{-\frac{x}{\sigma^2_{g_n}}}\, \mathrm{d}x = \sum_{n=1}^{N} \sigma^2_{g_n} = C_{13}. \tag{A.35}$$

Regarding I_2, first note that the CDF of $\frac{|g_n|^2}{|f_n|^2}$ is $\Pr\left\{\frac{|g_n|^2}{|f_n|^2} < x\right\} = \frac{x}{x + \sigma^2_{g_n}/\sigma^2_{f_n}}$. The CDF of $\Upsilon = \max_n \frac{|g_n|^2}{|f_n|^2}$ is thus $\Pr\{\Upsilon < x\} = \prod_{n=1}^{N} \frac{x}{x + \sigma^2_{g_n}/\sigma^2_{f_n}}$. The PDF Υ is therefore

$$\begin{aligned} f_{\Upsilon}(x) &= \sum_{n=1}^{N} \frac{\sigma^2_{g_n}/\sigma^2_{f_n}}{\left(x + \sigma^2_{g_n}/\sigma^2_{f_n}\right)^2} \prod_{n_1=1, n_1 \neq n}^{N} \underbrace{\frac{x}{x + \sigma^2_{g_{n_1}}/\sigma^2_{f_{n_1}}}}_{\leq 1} \\ &\leq \sum_{n=1}^{N} \frac{\sigma^2_{g_n}/\sigma^2_{f_n}}{\left(x + \sigma^2_{g_n}/\sigma^2_{f_n}\right)^2} \leq \left(\sum_{n=1}^{N} \frac{\sigma^2_{g_n}}{\sigma^2_{f_n}}\right) \frac{1}{\left(x + \min\limits_{n=1,\ldots,N} \frac{\sigma^2_{g_n}}{\sigma^2_{f_n}}\right)^2}. \end{aligned}$$

We can then obtain an upper bound on I_2 as

$$\begin{aligned}
I_2 &\leq \int_{\boldsymbol{H}\in\mathbb{C}^{2N\times 1}} \log\left(1+N\Upsilon\right) f_{\boldsymbol{H}}\left(\boldsymbol{H}\right)\mathrm{d}\boldsymbol{H} \\
&= \int_{0\leq N\Upsilon\leq 1} \underbrace{\log\left(1+N\Upsilon\right)}_{\leq \log 2} f_{\Upsilon}\left(\Upsilon\right)\mathrm{d}\Upsilon + \int_{N\Upsilon>1} \underbrace{\log\left(1+N\Upsilon\right)}_{\leq \log(2N\Upsilon)=\log(2N)+\log\Upsilon} f_{\Upsilon}\left(\Upsilon\right)\mathrm{d}\Upsilon \\
&\leq \log 2 + \log(2N) + \sum_{n=1}^{N}\frac{\sigma_{g_n}^2}{\sigma_{f_n}^2}\int_{\frac{1}{N}}^{\infty}\frac{\log\Upsilon}{\left(\Upsilon+\min\limits_{n=1,\ldots,N}\sigma_{g_n}^2/\sigma_{f_n}^2\right)^2}\mathrm{d}\Upsilon = C_{14},
\end{aligned} \tag{A.36}$$

where $C_{14} \triangleq \log 2 + \log(2N) + \sum_{n=1}^{N}\sigma_{g_n}^2/\sigma_{f_n}^2\left(\frac{\log N}{\frac{1}{N}+\min_{n=1}^{N}\sigma_{g_n}^2/\sigma_{f_n}^2} + \frac{\log\left(N\min_{n=1}^{N}\sigma_{g_n}^2/\sigma_{f_n}^2+1\right)}{\min_{n=1}^{N}\sigma_{g_n}^2/\sigma_{f_n}^2}\right)$.

Applying Lemma A.3 to (A.34), we can find an upper bound on I_3 as

$$\begin{aligned}
I_3 &= \int_{\frac{\alpha}{P}}^{\infty}\left(\log\frac{1}{x-\frac{\alpha}{P}}\right) f_{\sum_{n=1}^{N}\Gamma_n}(x)\mathrm{d}x = \int_{0}^{\infty}\left(\log\frac{1}{y}\right) f_{\sum_{i=1}^{N}\Gamma_i}\left(y+\frac{\alpha}{P}\right)\mathrm{d}y \\
&\leq D_N e^{-\frac{\alpha}{P\times\sigma_{g_N}^2}}\int_{0}^{\infty} e^{-\frac{y}{\sigma_{g_N}^2}}\left(\log\frac{1}{y}\right)\left(y+\frac{\alpha}{P}\right)^{N-1}\mathrm{d}x \\
&+ D_N e^{-\frac{\alpha}{P\times\sigma_{g_N}^2}}\left(\frac{1}{P^{N-1}}+\frac{1}{P^N}\right)\int_{0}^{\infty} e^{-\frac{y}{\sigma_{g_N}^2}}\left(\log\frac{1}{y}\right)\mathrm{d}y \\
&+ \mathbf{1}\left\{N\geq 3\right\}\times D_N e^{-\frac{\alpha}{P\times\sigma_{g_N}^2}}\sum_{m=1}^{N-2}\left(\frac{1}{P^{N-m-1}}+\frac{1}{P^{N-m}}\right) \\
&\times \int_{0}^{\infty} e^{-\frac{y}{\sigma_{g_N}^2}}\left(\log\frac{1}{y}\right)\left(y+\frac{\alpha}{P}\right)^{m}\mathrm{d}x.
\end{aligned} \tag{A.37}$$

The integral $\int_0^\infty e^{-\frac{y}{\sigma_{g_N}^2}}\left(\log\frac{1}{y}\right)\mathrm{d}y$ in (A.37) is computed as

$$\begin{aligned}
\int_{0}^{\infty} e^{-\frac{y}{\sigma_{g_N}^2}}\left(\log\frac{1}{y}\right)\mathrm{d}y &\overset{z=\log\frac{1}{y}}{=} \int_{-\infty}^{\infty} e^{-\frac{e^{-z}}{\sigma_{g_N}^2}} z e^{-z}\mathrm{d}z \\
&\leq \int_{0}^{\infty} e^{-\frac{e^{-z}}{\sigma_{g_N}^2}} z e^{-z}\mathrm{d}z \leq \int_{0}^{\infty} z e^{-z}\mathrm{d}z = 1.
\end{aligned} \tag{A.38}$$

Similarly, the integral $\int_0^\infty e^{-\frac{y}{\sigma_{g_N}^2}} \left(\log\frac{1}{y}\right)\left(y+\frac{\alpha}{P}\right)^n \mathrm{d}y$ for $n \geq 1$ is bounded by

$$\int_0^\infty e^{-\frac{y}{\sigma_{g_N}^2}} \left(\log\frac{1}{y}\right)\left(y+\frac{\alpha}{P}\right)^n \mathrm{d}y \leq 3 \times 2^n n! \times \left(1+\sigma_{g_N}^{2n}\right) \times \left(1+\left(\frac{\alpha}{P}\right)^n\right). \tag{A.39}$$

Applying (A.38) and (A.39) to (A.37), we obtain

$$\begin{aligned}
I_3 &\leq 3 \times 2^{N-1}(N-1)! \times \left(1+\sigma_{g_N}^{2(N-1)}\right) D_N e^{-\frac{\alpha}{P\times\sigma_{g_N}^2}} \left[1+\left(\frac{\alpha}{P}\right)^{N-1}\right] \\
&\quad + D_N e^{-\frac{\alpha}{P\times\sigma_{g_N}^2}} \underbrace{\left(\frac{1}{P^{N-1}}+\frac{1}{P^N}\right)}_{\leq 2\left[\frac{1}{P}+\frac{1}{P^N}\right]} \\
&\quad + \mathbf{1}\{N\geq 3\} \times D_N e^{-\frac{\alpha}{P\times\sigma_{g_N}^2}} \sum_{m=1}^{N-2} \underbrace{\left(\frac{1}{P^{N-m-1}}+\frac{1}{P^{N-m}}\right)}_{2\left[\frac{1}{P}+\frac{1}{P^N}\right]} \\
&\quad \times 3 \times \underbrace{\left(1+\sigma_{g_N}^{2m}\right)}_{\leq 2\left(1+\sigma_{g_N}^{2N}\right)} \times \underbrace{2^m m!}_{\leq 2^{N-2}(N-2)!} \underbrace{\left(1+\left(\frac{\alpha}{P}\right)^m\right)}_{\leq 2\left[1+\left(\frac{\alpha}{P}\right)^N\right]} \\
&\leq C_{15} \underbrace{e^{-\frac{\alpha}{P\times\sigma_{g_N}^2}} \left[1+\left(\frac{\alpha}{P}\right)^{N-1}\right]}_{\leq 1+e^{-(N-1)}(N-1)^{N-1}\sigma_{g_N}^{2(N-1)}} + C_{16} e^{-\frac{\alpha}{P\times\sigma_{g_N}^2}} \left[\frac{1}{P}+\frac{1}{P^N}\right] \\
&\quad + C_{17} e^{-\frac{\alpha}{P\times\sigma_{g_N}^2}} \left[\frac{1}{P}+\frac{1}{P^N}\right]\left[1+\left(\frac{\alpha}{P}\right)^N\right] \\
&\leq C_{18} + C_{19} e^{-\frac{\alpha}{P\times\sigma_{g_N}^2}} \left[\frac{1}{P}+\frac{1}{P^N}\right]\left[1+\left(\frac{\alpha}{P}\right)^N\right],
\end{aligned} \tag{A.40}$$

where $C_{15} \triangleq 3 \times 2^{N-1}(N-1)! \times \left(1+\sigma_{g_N}^{2(N-1)}\right) D_N$, $C_{16} \triangleq 2D_N$, $C_{17} \triangleq \mathbf{1}\{N\geq 3\} \times D_2 \times 3 \times (N-1)! 2^{N+1}\left(1+\sigma_{g_N}^{2N}\right)$, $C_{18} \triangleq C_{15} + C_{15} e^{-(N-1)}(N-1)^{N-1}\sigma_{g_N}^{2(N-1)}$ and $C_{19} \triangleq C_{16} + C_{17}$. Substituting (A.35), (A.36) and (A.40) into (A.34) and (A.29) completes the proof of Theorem 2.2. □

A.3 Proof of Theorem 2.3

Proof. Following the same derivations in Appendix A.1, it is sufficient to show that: (i) $\Pr\{\Gamma(\boldsymbol{\mu}^{\star}_{\mathtt{IND}}, \boldsymbol{H}) = \alpha\} = 0$; ii$\mathtt{E}_{\{\boldsymbol{\mu}_i\}_{\mathsf{N}}}[\mathbf{1}\{\boldsymbol{H} \in \mathcal{H}^{\mathtt{IND}}\}] = 0$, $\forall \boldsymbol{H} \in \mathcal{H}_3^{\mathtt{IND}}$, where

$$\mathcal{H}^{\mathtt{IND}} \triangleq \left\{\boldsymbol{H} \in \mathtt{C}^{2N\times 1} : \Gamma(\boldsymbol{\mu}_i, \boldsymbol{H}) < \alpha, \forall i \in \mathsf{N}\right\},$$
$$\mathcal{H}_3^{\mathtt{IND}} \triangleq \left\{\boldsymbol{H} \in \mathtt{C}^{2N\times 1} : \Gamma(\boldsymbol{\mu}^{\star}_{\mathtt{IND}}, \boldsymbol{H}) > \alpha\right\},$$

are analogues to the sets $\mathcal{H}, \mathcal{H}_3$ in Appendix A.1.

We first show $\Pr\{\Gamma(\boldsymbol{\mu}^{\star}_{\mathtt{IND}}, \boldsymbol{H}) = \alpha\} = 0$. Let $2^{\{1,\ldots,N\}}$ be the set of all subsets of $\{1,\ldots,N\}$. Based on the expression of $\boldsymbol{\mu}^{\star}_{\mathtt{IND}}$ in (2.7), for every $\boldsymbol{H}$, we have $\Gamma(\boldsymbol{\mu}^{\star}_{\mathtt{IND}}, \boldsymbol{H}) \in \{\Gamma_A(\boldsymbol{H}) : A \in 2^{\{1,\ldots,N\}}\}$, where

$$\Gamma_A(\boldsymbol{H}) \triangleq P\frac{\left|\sum_{n\in A}\frac{|f_n g_n|}{\sqrt{|f_n|^2+\frac{1}{P}}} + \frac{1+\sum_{n\in A}\frac{|g_n|^2}{|f_n|^2+\frac{1}{P}}}{\sum_{n\in A}\frac{|f_n g_n|}{\sqrt{|f_n|^2+\frac{1}{P}}}}\sum_{n\notin A}|f_n|^2\right|^2}{1+\sum_{n\in A}\frac{|g_n|^2}{|f_n|^2+\frac{1}{P}} + \left(\frac{1+\sum_{n\in A}\frac{|g_n|^2}{|f_n|^2+\frac{1}{P}}}{\sum_{n\in A}\frac{|f_n g_n|}{\sqrt{|f_n|^2+\frac{1}{P}}}}\right)^2\sum_{n\notin A}|f_n|^2}.$$

Since $\Gamma_A(\boldsymbol{H})$ is a continuous function of continuous random variables, it is a continuous random variable. Therefore, $\Pr\{\Gamma_A(\boldsymbol{H}) = \alpha\} = 0$, and

$$\Pr\{\Gamma(\boldsymbol{\mu}^{\star}_{\mathtt{IND}}, \boldsymbol{H}) = \alpha\} \leq \sum_{A\in 2^{\{1,\ldots,N\}}} \Pr\{\Gamma_A(\boldsymbol{H}) = \alpha\} = 0.$$

We now show $\mathtt{E}_{\{\boldsymbol{\mu}_i\}_{\mathsf{N}}}[\mathbf{1}\{\boldsymbol{H} \in \mathcal{H}^{\mathtt{IND}}\}] = 0$, $\forall \boldsymbol{H} \in \mathcal{H}_3^{\mathtt{IND}}$ by contradiction. Suppose that there is an $\tilde{\boldsymbol{H}} \in \mathcal{H}_3^{\mathtt{IND}}$, such that $\mathtt{E}_{\{\boldsymbol{\mu}_i\}_{\mathsf{N}}}\left[\mathbf{1}\left\{\tilde{\boldsymbol{H}} \in \mathcal{H}^{\mathtt{IND}}\right\}\right] = \varepsilon > 0$. Then, we have

$$\mathrm{E}_{\{\boldsymbol{\mu}_i\}_{\mathsf{N}}}\left[\mathbf{1}\left\{\tilde{\boldsymbol{H}} \in \mathcal{H}^{\mathtt{IND}}\right\}\right] = \Pr\left\{\Gamma\left(\boldsymbol{\mu}_i, \tilde{\boldsymbol{H}}\right) < \alpha, \forall i \in \mathsf{N}\right\}$$

$$\leq \Pr\left\{\Gamma\left(\boldsymbol{\mu}_i, \tilde{\boldsymbol{H}}\right) < \alpha, 0 \leq i \leq K-1\right\}$$
$$= \left[\Pr\left\{\Gamma\left(\boldsymbol{\mu}_i, \tilde{\boldsymbol{H}}\right) < \alpha\right\}\right]^K, \tag{A.41}$$

where $K \geq 1$ is an arbitrary finite natural number. Using the upper bound derived in (A.4), for any $\boldsymbol{\mu}_i = [\mu_{i,1}, \ldots, \mu_{i,N}]^\top$ and $\boldsymbol{\mu}^\star_{\texttt{IND}} = \left[\mu^\star_{i,1}, \ldots, \mu^\star_{i,N}\right]^\top$, we obtain

$$\begin{aligned}
&\Gamma\left(\boldsymbol{\mu}^\star_{\texttt{IND}}, \boldsymbol{H}\right) - \Gamma\left(\boldsymbol{\mu}_i, \boldsymbol{H}\right) \\
&\leq \hat{\Xi} \times \sum_{k=1}^{N} \left|\mu^\star_{i,k} - \mu_{i,k}\right| = \hat{\Xi} \times \sum_{k=1}^{N} \left|\left|\mu^\star_{i,k}\right| - |\mu_{i,k}| \times e^{j\left[\arg\left(\mu_{i,k}\right) - \arg\left(\mu^\star_{i,k}\right)\right]}\right| \\
&= \hat{\Xi} \times \sum_{k=1}^{N} \sqrt{\left(\left|\mu^\star_{i,k}\right| - |\mu_{i,k}|\right)^2 + 2\underbrace{\left|\mu^\star_{i,k}\right| \cdot |\mu_{i,k}|}_{\leq 1} \cdot \underbrace{\left[1 - \cos\left(\arg\left(\mu_{i,k}\right) - \arg\left(\mu^\star_{i,k}\right)\right)\right]}_{\leq \frac{1}{2}\left|\arg\left(\mu_{i,k}\right) - \arg\left(\mu^\star_{i,k}\right)\right|^2}},
\end{aligned}$$

where $\hat{\Xi} \triangleq 2P\left(\sum_{n=1}^{N} |g_n|\right)^2 \left(1 + \sum_{n=1}^{N} \frac{|g_n|^2}{|f_n|^2 + \frac{1}{P}}\right)$. For $\Gamma\left(\boldsymbol{\mu}^\star_{\texttt{IND}}, \boldsymbol{H}\right) > \alpha$, let $\delta \triangleq \frac{\Gamma\left(\boldsymbol{\mu}^\star_{\texttt{IND}}, \boldsymbol{H}\right) - \alpha}{\hat{\Xi}\sqrt{1+4\pi^2 N}} > 0$. Then, for any $\boldsymbol{\mu}_i$ with $\left|\left|\mu^\star_{i,k}\right| - |\mu_{i,k}|\right| \leq \delta$ and $\left|\arg\left(\mu_{i,k}\right) - \arg\left(\mu^\star_{i,k}\right)\right| \leq 2\pi \times \delta$, we have $\Gamma\left(\boldsymbol{\mu}^\star_{\texttt{IND}}, \boldsymbol{H}\right) - \Gamma\left(\boldsymbol{\mu}_i, \boldsymbol{H}\right) \leq \hat{\Xi}\sqrt{1+4\pi^2}N \times \delta = \Gamma\left(\boldsymbol{\mu}^\star_{\texttt{IND}}, \boldsymbol{H}\right) - \alpha$, which implies $\Gamma\left(\boldsymbol{\mu}_i, \boldsymbol{H}\right) \geq \alpha$. For $\tilde{\boldsymbol{H}}$, it follows that

$$\begin{aligned}
&\Pr\left\{\Gamma\left(\boldsymbol{\mu}_i, \tilde{\boldsymbol{H}}\right) < \alpha\right\} = 1 - \Pr\left\{\Gamma\left(\boldsymbol{\mu}_i, \tilde{\boldsymbol{H}}\right) \geq \alpha\right\} \\
&\leq 1 - \Pr\left\{\left|\left|\mu^\star_{i,k}\right| - |\mu_{i,k}|\right| \leq \delta, \left|\arg\left(\mu_{i,k}\right) - \arg\left(\mu^\star_{i,k}\right)\right| \leq 2\pi \times \delta, k = 1, \ldots, N\right\} \\
&= 1 - \prod_{k=1}^{N} \Pr\left\{\left|\left|\mu^\star_{i,k}\right| - |\mu_{i,k}|\right| \leq \delta\right\} \times \Pr\left\{\left|\arg\left(\mu_{i,k}\right) - \arg\left(\mu^\star_{i,k}\right)\right| \leq 2\pi \times \delta\right\} \\
&= 1 - \prod_{k=1}^{N} \Delta_{1,k}\Delta_{2,k} \triangleq \Delta,
\end{aligned} \tag{A.42}$$

where

$$\begin{aligned}
\Delta_{1,k} &\triangleq \left[\min\left(1, \left|\mu^\star_{i,k}\right| + \delta\right) - \max\left(0, \left|\mu^\star_{i,k}\right| - \delta\right)\right], \\
\Delta_{2,k} &\triangleq \left[\min\left(1, \tfrac{\arg\left(\mu^\star_{i,k}\right)}{2\pi} + \delta\right) - \max\left(0, \tfrac{\arg\left(\mu^\star_{i,k}\right)}{2\pi} - \delta\right)\right].
\end{aligned}$$

The last equality in (A.42) is derived from the fact that $|\mu_{i,k}|$ and $\arg(\mu_{i,k})$ are uniformly distributed in $(0,1]$ and $(0,2\pi]$, respectively, as defined in (2.15). It can be readily observed that $0 < \Delta_{1,k}, \Delta_{2,k} \leq 1$, and therefore, $0 \leq \Delta < 1$. Substituting (A.42) into (A.41), we have $\mathrm{E}_{\{\boldsymbol{\mu}_i\}_{\mathsf{N}}}\left[\mathbf{1}\left\{\tilde{\boldsymbol{H}} \in \mathcal{H}^{\mathtt{IND}}\right\}\right] \leq \Delta^K$. When $\Delta = 0$, we have $\mathrm{E}_{\{\boldsymbol{\mu}_i\}_{\mathsf{N}}}\left[\mathbf{1}\left\{\tilde{\boldsymbol{H}} \in \mathcal{H}^{\mathtt{IND}}\right\}\right] = 0 < \varepsilon$. Moreover, when $\Delta > 0$, letting $K \triangleq \lceil \log_{\Delta}^{\varepsilon} \rceil + 1$, we obtain $\mathrm{E}_{\{\boldsymbol{\mu}_i\}_{\mathsf{N}}}\left[\mathbf{1}\left\{\tilde{\boldsymbol{H}} \in \mathcal{H}^{\mathtt{IND}}\right\}\right] \leq \Delta^{\lceil \log_{\Delta}^{\varepsilon} \rceil + 1} < \Delta^{\log_{\Delta}^{\varepsilon}} = \varepsilon$. Both cases contradict the assumption that $\mathtt{E}_{\{\boldsymbol{\mu}_i\}_{\mathsf{N}}}\left[\mathbf{1}\left\{\tilde{\boldsymbol{H}} \in \mathcal{H}^{\mathtt{IND}}\right\}\right] = \varepsilon$. □

B Supplementary Proofs for Chapter 3

B.1 Proof of Theorem 3.1

Proof. Before showing the detailed proof, let us summarize the main idea behind the proof first. Based on (3.1) and (3.6), to show $\mathtt{Out}(\mathsf{Q}_{\mathrm{VLQ}}) = \mathtt{Out}(\mathtt{Full})$, it is equivalent to prove:

1. For any $\boldsymbol{H}$ satisfying $\gamma(\mathtt{Full}(\boldsymbol{H}), \boldsymbol{H}) < \frac{1}{P}$, $\mathbf{1}_{\gamma(\mathtt{Full}(\boldsymbol{H}),\boldsymbol{H})<\frac{1}{P}} = \mathrm{E}_{\{\boldsymbol{x}_i\}_{i\in\mathbb{N}}}\mathbf{1}_{\gamma(\boldsymbol{x}_i,\boldsymbol{H})<\frac{1}{P},\forall i\in\mathbb{N}} = 1$;

2. For any $\boldsymbol{H}$ satisfying $\gamma(\mathtt{Full}(\boldsymbol{H}), \boldsymbol{H}) = \frac{1}{P}$,

$$\int_{\boldsymbol{H}} \mathbf{1}_{\gamma(\mathtt{Full}(\boldsymbol{H}),\boldsymbol{H})<\frac{1}{P}} f_{\boldsymbol{H}}(\boldsymbol{H})\,\mathrm{d}\boldsymbol{H} = \int_{\boldsymbol{H}} \mathrm{E}_{\{\boldsymbol{x}_i\}_{i\in\mathbb{N}}}\mathbf{1}_{\gamma(\boldsymbol{x}_i,\boldsymbol{H})<\frac{1}{P},\forall i\in\mathbb{N}} f_{\boldsymbol{H}}(\boldsymbol{H})\,\mathrm{d}\boldsymbol{H} = 0;$$

3. For any $\boldsymbol{H}$ satisfying $\gamma(\mathtt{Full}(\boldsymbol{H}), \boldsymbol{H}) > \frac{1}{P}$, $\mathbf{1}_{\gamma(\mathtt{Full}(\boldsymbol{H}),\boldsymbol{H})>\frac{1}{P}} = \mathrm{E}_{\{\boldsymbol{x}_i\}_{i\in\mathbb{N}}}\mathbf{1}_{\gamma(\boldsymbol{x}_i,\boldsymbol{H})>\frac{1}{P},\forall i\in\mathbb{N}} = 0$.

Define

$$\mathcal{S}_1 = \left\{\boldsymbol{H} : \boldsymbol{H} \in \mathbb{C}^{t\times 2}, \gamma(\mathtt{Full}(\boldsymbol{H}), \boldsymbol{H}) < \frac{1}{P}\right\}.$$

For any realization of $\{\boldsymbol{x}_i\}_{i\in\mathbb{N}}$, define

$$\mathcal{S}_2\left(\{\boldsymbol{x}_i\}_{i\in\mathbb{N}}\right)=\left\{\boldsymbol{H}:\boldsymbol{H}\in\mathbb{C}^{t\times 2},\gamma\left(\boldsymbol{x}_i,\boldsymbol{H}\right)<\frac{1}{P},\forall i\in\mathbb{N}\right\}.$$

For brevity, we omit the dependency of $\mathcal{S}_2\left(\{\boldsymbol{x}_i\}_{i\in\mathbb{N}}\right)$ on $\{\boldsymbol{x}_i\}_{i\in\mathbb{N}}$ and simply use $\mathcal{S}_2$. From (3.1) and (3.6), $\mathtt{Out}\left(\mathtt{Full}\right)$ and $\mathtt{Out}(\mathsf{Q}_{\mathrm{VLQ}})$ can be rewritten as

$$\mathtt{Out}\left(\mathtt{Full}\right) = \mathrm{E}_{\boldsymbol{H}}\mathbf{1}_{\boldsymbol{H}\in\mathcal{S}_1}, \tag{B.43}$$

$$\mathtt{Out}(\mathsf{Q}_{\mathrm{VLQ}}) = \mathrm{E}_{\{\boldsymbol{x}_i\}_{i\in\mathbb{N}}}\mathrm{E}_{\boldsymbol{H}}\mathbf{1}_{\boldsymbol{H}\in\mathcal{S}_2}. \tag{B.44}$$

For convenience, we define

$$\begin{aligned}
\mathcal{S}_{21} &= \left\{\boldsymbol{H}:\boldsymbol{H}\in\mathcal{S}_2,\gamma\left(\mathtt{Full}\left(\boldsymbol{H}\right),\boldsymbol{H}\right)<\tfrac{1}{P}\right\},\\
\mathcal{S}_{22} &= \left\{\boldsymbol{H}:\boldsymbol{H}\in\mathcal{S}_2,\gamma\left(\mathtt{Full}\left(\boldsymbol{H}\right),\boldsymbol{H}\right)=\tfrac{1}{P}\right\},\\
\mathcal{S}_{23} &= \left\{\boldsymbol{H}:\boldsymbol{H}\in\mathcal{S}_2,\gamma\left(\mathtt{Full}\left(\boldsymbol{H}\right),\boldsymbol{H}\right)>\tfrac{1}{P}\right\}.
\end{aligned}$$

Since $\mathcal{S}_2=\mathcal{S}_{21}\cup\mathcal{S}_{22}\cup\mathcal{S}_{23}$ and $\mathcal{S}_{21}$, $\mathcal{S}_{22}$, $\mathcal{S}_{23}$ are mutually exclusive, $\mathtt{Out}\left(\mathsf{Q}_{\mathrm{VLQ}}\right)$ in (B.44) is rewritten as

$$\mathtt{Out}\left(\mathsf{Q}_{\mathrm{VLQ}}\right)=\sum_{l=1}^{3}\mathrm{E}_{\{\boldsymbol{x}_i\}_{i\in\mathbb{N}}}\mathrm{E}_{\boldsymbol{H}}\mathbf{1}_{\boldsymbol{H}\in\mathcal{S}_{2l}}. \tag{B.45}$$

In order to prove $\mathtt{Out}(\mathsf{Q}_{\mathrm{VLQ}})=\mathtt{Out}(\mathtt{Full})$, according to (B.43) and (B.45), we will show $\mathrm{E}_{\{\boldsymbol{x}_i\}_{i\in\mathbb{N}}}\mathrm{E}_{\boldsymbol{H}}\mathbf{1}_{\boldsymbol{H}\in\mathcal{S}_{21}}=\mathrm{E}_{\boldsymbol{H}}\mathbf{1}_{\boldsymbol{H}\in\mathcal{S}_1}$, $\mathrm{E}_{\{\boldsymbol{x}_i\}_{i\in\mathbb{N}}}\mathrm{E}_{\boldsymbol{H}}\mathbf{1}_{\boldsymbol{H}\in\mathcal{S}_{22}}=0$ and $\mathrm{E}_{\{\boldsymbol{x}_i\}_{i\in\mathbb{N}}}\mathrm{E}_{\boldsymbol{H}}\mathbf{1}_{\boldsymbol{H}\in\mathcal{S}_{23}}=0$.

First, to prove $\mathrm{E}_{\{\boldsymbol{x}_i\}_{i\in\mathbb{N}}}\mathrm{E}_{\boldsymbol{H}}\mathbf{1}_{\boldsymbol{H}\in\mathcal{S}_{21}}=\mathrm{E}_{\boldsymbol{H}}\mathbf{1}_{\boldsymbol{H}\in\mathcal{S}_1}$, it is sufficient to prove $\mathbf{1}_{\boldsymbol{H}\in\mathcal{S}_1}=\mathbf{1}_{\boldsymbol{H}\in\mathcal{S}_{21}}$ for any given $\boldsymbol{H}$ and $\{\boldsymbol{x}_i\}_{i\in\mathbb{N}}$. When $\mathbf{1}_{\boldsymbol{H}\in\mathcal{S}_1}=0$, based on the definition of $\mathcal{S}_1$, it means $\boldsymbol{H}\notin\mathcal{S}_1$ and $\gamma\left(\mathtt{Full}\left(\boldsymbol{H}\right),\boldsymbol{H}\right)\geq\frac{1}{P}$. From the definition of $\mathcal{S}_{21}$, we have $\boldsymbol{H}\notin\mathcal{S}_{21}$ and $\mathbf{1}_{\boldsymbol{H}\in\mathcal{S}_{21}}=0$. When

$\mathbf{1}_{\boldsymbol{H}\in\mathcal{S}_1}=1, \boldsymbol{H}\in\mathcal{S}_1$. By the optimality of $\texttt{Full}\left(\boldsymbol{H}\right)$, it must have $\boldsymbol{H}\in\mathcal{S}_2$. Since $\mathcal{S}_{21}=\mathcal{S}_1\cap\mathcal{S}_2$, $\boldsymbol{H}\in\mathcal{S}_{21}$ and $\mathbf{1}_{\boldsymbol{H}\in\mathcal{S}_{21}}=1$. Therefore, $\mathbf{1}_{\boldsymbol{H}\in\mathcal{S}_{21}}=\mathbf{1}_{\boldsymbol{H}\in\mathcal{S}_1}$ and $\mathrm{E}_{\{\boldsymbol{x}_i\}_{i\in\mathbb{N}}}\mathrm{E}_{\boldsymbol{H}}\mathbf{1}_{\boldsymbol{H}\in\mathcal{S}_{21}}=\mathrm{E}_{\boldsymbol{H}}\mathbf{1}_{\boldsymbol{H}\in\mathcal{S}_1}$.

Second, we will prove $\mathrm{E}_{\{\boldsymbol{x}_i\}_{i\in\mathbb{N}}}\mathrm{E}_{\boldsymbol{H}}\mathbf{1}_{\boldsymbol{H}\in\mathcal{S}_{22}}=0$. Define $\mathcal{S}_3=\left\{\boldsymbol{H}:\boldsymbol{H}\in\mathbb{C}^{t\times 2},\gamma\left(\texttt{Full}\left(\boldsymbol{H}\right),\boldsymbol{H}\right)=\frac{1}{P}\right\}$. By definition, $\mathcal{S}_{22}=\mathcal{S}_2\cap\mathcal{S}_3\subseteq\mathcal{S}_3$. Then,

$$\mathrm{E}_{\{\boldsymbol{x}_i\}_{i\in\mathbb{N}}}\mathrm{E}_{\boldsymbol{H}}\mathbf{1}_{\boldsymbol{H}\in\mathcal{S}_{22}}\leq\mathrm{E}_{\boldsymbol{H}}\mathbf{1}_{\boldsymbol{H}\in\mathcal{S}_3}=\mathrm{Prob}\left\{\gamma\left(\texttt{Full}\left(\boldsymbol{H}\right),\boldsymbol{H}\right)=\frac{1}{P}\right\}.$$

Let $m_{\min}=\underset{m=1,2}{\operatorname{argmin}}\,\chi_m$, $m_{\max}=\underset{m=1,2}{\operatorname{argmax}}\,\chi_m$ and $\theta=\frac{\left|\boldsymbol{h}_1^{\dagger}\boldsymbol{h}_2\right|^2}{\chi_1\chi_2}$. According to [19, Theorem 2],

$$\gamma\left(\texttt{Full}\left(\boldsymbol{H}\right),\boldsymbol{H}\right)=\begin{cases}\chi_{m_{\min}}, & \theta\geq\frac{\chi_{m_{\min}}}{\chi_{m_{\max}}},\\ \frac{\chi_{m_{\min}}}{1+\beta^2}, & \theta<\frac{\chi_{m_{\min}}}{\chi_{m_{\max}}},\end{cases}\tag{B.46}$$

where $\beta=\frac{\sqrt{\chi_{m_{\min}}}-\sqrt{\chi_{m_{\max}}\theta}}{\sqrt{\chi_{m_{\max}}-\chi_{m_{\max}}\theta}}$. Since θ, χ_1, and χ_2 are mutually independent, θ and $\chi_{m_{\min}}$ and $\chi_{m_{\max}}$ are also mutually independent [21]. With (B.46), it is straightforward to show that $\mathrm{Prob}\left\{\gamma\left(\texttt{Full}\left(\boldsymbol{H}\right),\boldsymbol{H}\right)=\frac{1}{P}\right\}=0$ by fixing $\chi_{m_{\min}},\chi_{m_{\max}}$ or θ as well as using the fact that the probability of a continuous r.v. assuming a specific value is zero. Therefore, $\mathrm{E}_{\{\boldsymbol{x}_i\}_{i\in\mathbb{N}}}\mathrm{E}_{\boldsymbol{H}}\mathbf{1}_{\boldsymbol{H}\in\mathcal{S}_{22}}\leq 0$. Since the probability is non-negative, $\mathrm{E}_{\{\boldsymbol{x}_i\}_{i\in\mathbb{N}}}\mathrm{E}_{\boldsymbol{H}}\mathbf{1}_{\boldsymbol{H}\in\mathcal{S}_{22}}=0$.

Finally, we will prove $\mathrm{E}_{\{\boldsymbol{x}_i\}_{i\in\mathbb{N}}}\mathrm{E}_{\boldsymbol{H}}\mathbf{1}_{\boldsymbol{H}\in\mathcal{S}_{23}}=0$. Define

$$\mathcal{S}_4=\left\{\boldsymbol{H}:\boldsymbol{H}\in\mathbb{C}^{t\times 2},\gamma\left(\texttt{Full}\left(\boldsymbol{H}\right),\boldsymbol{H}\right)>\frac{1}{P}\right\}.$$

Since $\mathcal{S}_{23}=\mathcal{S}_2\cap\mathcal{S}_4$, we obtain $\mathrm{E}_{\{\boldsymbol{x}_i\}_{i\in\mathbb{N}}}\mathrm{E}_{\boldsymbol{H}}\mathbf{1}_{\boldsymbol{H}\in\mathcal{S}_{23}}=\int_{\boldsymbol{H}\in\mathcal{S}_4}f_{\boldsymbol{H}}(\boldsymbol{H})\mathrm{E}_{\{\boldsymbol{x}_i\}_{i\in\mathbb{N}}}\mathbf{1}_{\boldsymbol{H}\in\mathcal{S}_2}\mathrm{d}\boldsymbol{H}$. To prove $\mathrm{E}_{\{\boldsymbol{x}_i\}_{i\in\mathbb{N}}}\mathrm{E}_{\boldsymbol{H}}\mathbf{1}_{\boldsymbol{H}\in\mathcal{S}_{23}}=0$, it is sufficicent to show $\mathrm{E}_{\{\boldsymbol{x}_i\}_{i\in\mathbb{N}}}\mathbf{1}_{\boldsymbol{H}\in\mathcal{S}_2}=0$ for any $\boldsymbol{H}\in\mathcal{S}_4$. By contradiction, assume $\exists\tilde{\boldsymbol{H}}\in\mathcal{S}_4$, s.t. $\mathrm{E}_{\{\boldsymbol{x}_i\}_{i\in\mathbb{N}}}\mathbf{1}_{\tilde{\boldsymbol{H}}\in\mathcal{S}_2}=\varepsilon>0$. Then

$$\mathrm{E}_{\{\boldsymbol{x}_i\}_{i\in\mathbb{N}}}\mathbf{1}_{\tilde{\boldsymbol{H}}\in\mathcal{S}_2}=\mathrm{Prob}\left\{\gamma\left(\boldsymbol{x}_i,\tilde{\boldsymbol{H}}\right)<\frac{1}{P},\forall i\in\mathbb{N}\right\}$$

$$\leq \text{Prob}\left\{\gamma\left(\boldsymbol{x}_i, \tilde{\boldsymbol{H}}\right) < \frac{1}{P}, \forall i \in \{0, 1, \ldots, K-1\}\right\}$$
$$= \left[\text{Prob}\left\{\gamma\left(\boldsymbol{x}_i, \tilde{\boldsymbol{H}}\right) < \frac{1}{P}\right\}\right]^K, \tag{B.47}$$

where $K \geq 1$ can be any finite natural number, and the last equality is because for a given $\tilde{\boldsymbol{H}}$, $\gamma\left(\boldsymbol{x}_i, \tilde{\boldsymbol{H}}\right)$ for $i = 1, \ldots, K$ are mutually independent due to the independence of $\boldsymbol{x}_i$ for $i = 1, \ldots, K$. We shall use the following lemma, the proof of which is in Appendix B.4.

Lemma B.4. *If $\gamma\left(\texttt{Full}\left(\boldsymbol{H}\right), \boldsymbol{H}\right) > \frac{1}{P}$, there exists $\Pi \in (0,1)$ such that for any $\boldsymbol{x} \in \mathcal{X}$ with $\left|\boldsymbol{x}^\dagger \texttt{Full}\left(\boldsymbol{H}\right)\right|^2 \geq \Pi$, $\gamma\left(\boldsymbol{x}, \boldsymbol{H}\right) \geq \frac{1}{P}$ holds.*

From Lemma 3.1, for a given $\tilde{\boldsymbol{H}}$, we have

$$\text{Prob}\left\{\gamma\left(\boldsymbol{x}_i, \tilde{\boldsymbol{H}}\right) \geq \frac{1}{P}\right\} \geq \text{Prob}\left\{\left|\boldsymbol{x}_i^\dagger \texttt{Full}\left(\tilde{\boldsymbol{H}}\right)\right|^2 \geq \Pi\right\} = (1-\Pi)^{t-1} > 0.$$

Therefore, $\text{Prob}\left\{\gamma\left(\boldsymbol{x}_i, \tilde{\boldsymbol{H}}\right) < \frac{1}{P}\right\} \leq 1 - (1-\Pi)^{t-1} < 1$. By (B.47), it can be derived that $\text{E}_{\{\boldsymbol{x}_i\}_{i\in\mathbb{N}}} \mathbf{1}_{\tilde{\boldsymbol{H}} \in \mathcal{S}_2} \leq \left[1 - (1-\Pi)^{t-1}\right]^K$. Let $K = \left\lceil \log_{(1-(1-\Pi)^{t-1})} \varepsilon \right\rceil + 1$, then $\text{E}_{\{\boldsymbol{x}_i\}_{i\in\mathbb{N}}} \mathbf{1}_{\tilde{\boldsymbol{H}} \in \mathcal{S}_2} \leq \left[1 - (1-\Pi)^{t-1}\right]^{\left\lceil \log_{(1-(1-\Pi)^{t-1})} \varepsilon \right\rceil + 1} < \left[1 - (1-\Pi)^{t-1}\right]^{\log_{(1-(1-\Pi)^{t-1})} \varepsilon} = \varepsilon$, which contradicts the assumption that $\text{E}_{\{\boldsymbol{x}_i\}_{i\in\mathbb{N}}} \mathbf{1}_{\tilde{\boldsymbol{H}} \in \mathcal{S}_2} = \varepsilon$. Thus $\text{E}_{\{\boldsymbol{x}_i\}_{i\in\mathbb{N}}} \mathbf{1}_{\tilde{\boldsymbol{H}} \in \mathcal{S}_2} = 0$ and $\text{E}_{\{\boldsymbol{x}_i\}_{i\in\mathbb{N}}} \text{E}_{\boldsymbol{H}} \mathbf{1}_{\boldsymbol{H} \in \mathcal{S}_{23}} = 0$, which completes the proof. □

Remark 3: It follows from (3.7) and (B.45) that $\texttt{Out}\left(\mathsf{Q}_{\text{VLQ}}\right) = \text{E}_{\{\boldsymbol{x}_i\}_{i\in\mathbb{N}}} \text{E}_{\boldsymbol{H}} \mathbf{1}_{\mathcal{S}_2} = \texttt{Out}\left(\texttt{Full}\right) = \text{E}_{\boldsymbol{H}} \mathbf{1}_{\mathcal{S}_1}$, thus $\text{E}_{\boldsymbol{H}}\left[\mathbf{1}_{\mathcal{S}_1} - \text{E}_{\{\boldsymbol{x}_i\}_{i\in\mathbb{N}}} \mathbf{1}_{\mathcal{S}_2}\right] = 0$. Based on the definitions of $\mathcal{S}_1$ and $\mathcal{S}_2$, $\mathbf{1}_{\mathcal{S}_1} - \text{E}_{\{\boldsymbol{x}_i\}_{i\in\mathbb{N}}} \mathbf{1}_{\mathcal{S}_2}$ is always non-positive for any $\boldsymbol{H}$. Therefore, $\mathbf{1}_{\mathcal{S}_1} - \text{E}_{\{\boldsymbol{x}_i\}_{i\in\mathbb{N}}} \mathbf{1}_{\mathcal{S}_2} = 0$ for any $\boldsymbol{H}$ with probability one. In other words, $\mathcal{R}_0$ in (3.3) is equal to the expectation of

$$\left\{\boldsymbol{H} : \gamma\left(\boldsymbol{x}_0, \boldsymbol{H}\right) \geq \frac{1}{P}\right\} \bigcup \left\{\boldsymbol{H} : \gamma\left(\texttt{Full}\left(\boldsymbol{H}\right), \boldsymbol{H}\right) < \frac{1}{P}\right\}$$

with regard to $\{\boldsymbol{x}_i\}_{i\in\mathbb{N}}$ with probability one.

B.2 Proof of Lemma 3.1

Proof. For $p = 0$, $\Phi = 0$, then the upper bound in (3.10) holds. Hence, suppose that $0 < p < 1$. Then

$$\begin{aligned}
\Phi &= \sum_{i=1}^{\infty} p^i (1-p) \lfloor \log_2(i+1) \rfloor \\
&= p(1-p) + \sum_{i=2}^{\infty} p^i (1-p) \lfloor \log_2(i+1) \rfloor \\
&\le p(1-p) + \sum_{i=2}^{\infty} p^i (1-p) \log_2(i+1) \\
&= p(1-p) + p(1-p) \sum_{i=1}^{\infty} p^i \log_2(i+2) \\
&= p(1-p) + p(1-p) \left[p \log_2 3 + \sum_{i=2}^{\infty} p^i \log_2(i+2) \right] \\
&\le p(1-p) + p(1-p) \left[p \log_2 3 + \frac{2}{\log 2} \sum_{i=1}^{\infty} p^i \log i \right]. \qquad \text{(B.48)}
\end{aligned}$$

We estimate the sum $\sum_{i=1}^{\infty} p^i \log i$ via the integral of the function $f(x) = e^{-\beta x} \log x$, where $0 < \beta \triangleq -\log p < \infty$. We calculate $f'(x) = e^{-\beta x} \left(\frac{1}{x} - \beta \log x \right)$, where f' represents the derivative of f. For $y \log y = \frac{1}{\beta}$, $f'(x) > 0$ for $1 \le x < y$, $f'(x) = 0$ for $x = y$, and $f'(x) < 0$ for $x > y$. The global maximum of f is thus $f(y)$. Since $y \log y = \frac{1}{\beta} > 0$, $y \ge 1$ must hold, which implies $f(y) = e^{-\beta y} \log y \le e^{-\beta} \log y \le e^{-\beta} y \log y = \frac{e^{-\beta}}{\beta}$. Let $j = \lfloor y \rfloor$. Then $1 \le j \le y < j+1$, and

$$\begin{aligned}
\sum_{i=1}^{\infty} f(i) &= \mathbf{1}_{j \ge 2} \sum_{i=1}^{j-1} f(i) + f(j) + f(j+1) + \sum_{i=j+2}^{\infty} f(i) \\
&= \mathbf{1}_{j \ge 2} \sum_{i=1}^{j-1} \int_i^{i+1} f(i) \mathrm{d}x + f(j) + f(j+1) + \sum_{i=j+2}^{\infty} \int_{i-1}^{i} f(i) \mathrm{d}x \\
&\le \mathbf{1}_{j \ge 2} \sum_{i=1}^{j-1} \int_i^{i+1} f(x) \mathrm{d}x + f(y) + f(y) + \sum_{i=j+2}^{\infty} \int_{i-1}^{i} f(x) \mathrm{d}x
\end{aligned}$$

$$= \mathbf{1}_{j\geq 2} \int_1^j f(x)\mathrm{d}x + 2f(y) + \int_{j+1}^{\infty} f(x)\mathrm{d}x$$
$$< 2f(y) + \int_1^{\infty} f(x)\mathrm{d}x \leq \frac{2e^{-\beta}}{\beta} + \int_1^{\infty} f(x)\mathrm{d}x, \tag{B.49}$$

where the first inequality follows since f is increasing on $(1, j)$ and decreasing on $(j+1, \infty)$. We now estimate the integral. With a change of variables $u = \log x$, $\mathrm{d}v = e^{-\beta x}\mathrm{d}x$, we obtain

$$\int_1^{\infty} f(x)\mathrm{d}x = \left(-\frac{1}{\beta}\log x e^{-\beta x}\right)\Big|_1^{\infty} + \frac{1}{\beta}\int_1^{\infty} \frac{1}{x} e^{-\beta x} dx = \frac{1}{\beta}\mathtt{E}_1(\beta) < \frac{e^{-\beta}}{\beta}\log\left(1 + \frac{1}{\beta}\right).$$

Combining with (B.49) and subsituting $\beta = -\log p$, it follows that

$$\begin{aligned}\sum_{i=1}^{\infty} f(i) &< \frac{p}{-\log p}\left[2 + \log\left(1 + \frac{1}{-\log p}\right)\right] \\ &< \frac{p}{1-p}\left[2 + \log\left(1 + \frac{1}{1-p}\right)\right] \\ &< \frac{p}{1-p}\left[2 + \log\frac{2}{1-p}\right] < \frac{p}{1-p}\left[3 + \log\frac{1}{1-p}\right],\end{aligned} \tag{B.50}$$

where the second inequality is because $-\log p > 1 - p$ for $0 < p < 1$. Substituting (B.50) into (B.48) yields that

$$\begin{aligned}\Phi &\leq p(1-p) + p^2(1-p)\log_2 3 + \frac{2p^2}{\log 2}\left(3 + \log\frac{1}{1-p}\right) \\ &\leq p(1-p) + 2p^2 + \frac{6p^2}{\log 2} + \frac{2p^2}{\log 2}\log\frac{1}{1-p} \\ &= p(1-p) + \left(\frac{6}{\log 2} + 2\right)p^2 + \frac{2}{\log 2}p^2\log\frac{1}{1-p}.\end{aligned}$$

This concludes the proof. □

B.3 Proof of Theorem 3.2

Proof. Based on (3.11), we will derive upper bounds on I_1, I_2 and I_3, separately. First, since

$\mathcal{H}_3 \subseteq \mathcal{H}_0$, we get

$$I_1 \leq C_1 \int_{\boldsymbol{H} \in \mathcal{H}_0} p(1-p) f_{\boldsymbol{H}}(\boldsymbol{H}) \mathrm{d}\boldsymbol{H}.$$

Substituting the upper bounds in (3.12) and (3.14) into I_1, it is derived that

$$I_1 \leq \frac{C_4}{P} \sum_{m=1}^{2} \int_{\boldsymbol{H} \in \mathcal{H}_0} \frac{1}{\chi_m} \left(1 - \frac{1}{P\chi_1}\right)^{t-1} f_{\boldsymbol{H}}(\boldsymbol{H}) \mathrm{d}\boldsymbol{H},$$

where $C_4 = (t-1)C_1$. Since χ_m is chi-squared distributed, the PDF of χ_m is

$$f_{\chi_m}(\chi_m) = \frac{\chi_m^{t-1} e^{-\chi_m}}{(t-1)!},$$

for $m = 1, 2$ [61]. Then we obtain

$$\begin{aligned} I_1 &\leq \frac{C_4}{P} \sum_{m=1}^{2} \int_{\frac{1}{P}}^{\infty} \int_{\frac{1}{P}}^{\infty} \frac{1}{\chi_m} \left(1 - \frac{1}{P\chi_1}\right)^{t-1} \left[\frac{\chi_1^{t-1} e^{-\chi_1}}{(t-1)!}\right] \left[\frac{\chi_2^{t-1} e^{-\chi_2}}{(t-1)!}\right] \mathrm{d}\chi_1 \mathrm{d}\chi_2 \\ &= \frac{C_5}{P} \int_{\frac{1}{P}}^{\infty} \left(1 - \frac{1}{P\chi_1}\right)^{t-1} \chi_1^{t-2} e^{-\chi_1} \mathrm{d}\chi_1 \int_{\frac{1}{P}}^{\infty} \frac{\chi_2^{t-1} e^{-\chi_2}}{(t-1)!} \mathrm{d}\chi_2 \\ &+ \frac{C_5}{P} \int_{\frac{1}{P}}^{\infty} \left(1 - \frac{1}{P\chi_1}\right)^{t-1} \chi_1^{t-1} e^{-\chi_1} \mathrm{d}\chi_1 \int_{\frac{1}{P}}^{\infty} \frac{\chi_2^{t-2} e^{-\chi_2}}{(t-1)!} \mathrm{d}\chi_2, \end{aligned}$$

where $C_5 = \frac{C_4}{(t-1)!}$. Noting that $\int_0^{\infty} x^{n-1} e^{-x} \mathrm{d}x = (n-1)!$ for $n \geq 1$ and $n \in \mathbb{N}$ [61], I_1 is bounded by

$$\begin{aligned} I_1 &\leq \frac{C_5}{P} \int_{\frac{1}{P}}^{\infty} \left(1 - \frac{1}{P\chi_1}\right)^{t-1} \chi_1^{t-2} e^{-\chi_1} \mathrm{d}\chi_1 \int_{0}^{\infty} \frac{\chi_2^{t-1} e^{-\chi_2}}{(t-1)!} \mathrm{d}\chi_2 \\ &+ \frac{C_5}{P} \int_{\frac{1}{P}}^{\infty} \left(1 - \frac{1}{P\chi_1}\right)^{t-1} \chi_1^{t-1} e^{-\chi_1} \mathrm{d}\chi_1 \int_{0}^{\infty} \frac{\chi_2^{t-2} e^{-\chi_2}}{(t-1)!} \mathrm{d}\chi_2 \\ &\leq \frac{C_5}{P} \int_{\frac{1}{P}}^{\infty} \left(1 - \frac{1}{P\chi_1}\right)^{t-1} \chi_1^{t-2} e^{-\chi_1} \mathrm{d}\chi_1 + \frac{C_6}{P} \int_{\frac{1}{P}}^{\infty} \left(1 - \frac{1}{P\chi_1}\right)^{t-1} \chi_1^{t-1} e^{-\chi_1} \mathrm{d}\chi_1, \end{aligned}$$

where $C_6 = \frac{C_5}{t-1}$. Letting $\chi_1 - \frac{1}{P} = \lambda_1$, the bound is derived as

$$\begin{aligned} I_1 &\leq C_5 \frac{e^{-\frac{1}{P}}}{P} \int_0^\infty \frac{\lambda_1}{\lambda_1 + \frac{1}{P}} \lambda_1^{t-2} e^{-\lambda_1} \mathrm{d}\lambda_1 + C_6 \frac{e^{-\frac{1}{P}}}{P} \int_0^\infty \lambda_1^{t-1} e^{-\lambda_1} \mathrm{d}\lambda_1 \\ &\leq C_5 \frac{e^{-\frac{1}{P}}}{P} \int_0^\infty \lambda_1^{t-2} e^{-\lambda_1} \mathrm{d}\lambda_1 + C_6 (t-1)! \frac{e^{-\frac{1}{P}}}{P} \\ &= C_5 (t-2)! \frac{e^{-\frac{1}{P}}}{P} + C_6 (t-1)! \frac{e^{-\frac{1}{P}}}{P} = C_7 \frac{e^{-\frac{1}{P}}}{P}, \end{aligned} \tag{B.51}$$

where $C_7 = (t-2)!C_5 + (t-1)!C_6$.

To derive I_2, applying the upper bound in (3.12) and based on the fact that $\mathcal{H}_1 \subseteq \mathcal{H}_0$, we obtain

$$\begin{aligned} I_2 &\leq \frac{C_8}{P^2} \int_{\boldsymbol{H} \in \mathcal{H}_0} \left[\frac{1}{\chi_1} + \frac{1}{\chi_2} \right]^2 f_{\boldsymbol{H}}(\boldsymbol{H}) \mathrm{d}\boldsymbol{H} \\ &= \frac{C_8}{P^2} \int_{\frac{1}{P}}^\infty \int_{\frac{1}{P}}^\infty \left[\frac{1}{\chi_1^2} + \frac{1}{\chi_2^2} + \frac{2}{\chi_1 \chi_2} \right] \left[\frac{\chi_1^{t-1} e^{-\chi_1}}{(t-1)!} \right] \left[\frac{\chi_2^{t-1} e^{-\chi_2}}{(t-1)!} \right] \mathrm{d}\chi_1 \mathrm{d}\chi_2 \\ &= \frac{2C_8}{(t-1)!P^2} \int_{\frac{1}{P}}^\infty \chi_1^{t-3} e^{-\chi_1} \mathrm{d}\chi_1 \int_{\frac{1}{P}}^\infty \frac{\chi_2^{t-1} e^{-\chi_2}}{(t-1)!} \mathrm{d}\chi_2 \\ &+ \frac{2C_8}{(t-1)!P^2} \int_{\frac{1}{P}}^\infty \chi_1^{t-2} e^{-\chi_1} \mathrm{d}\chi_1 \int_{\frac{1}{P}}^\infty \frac{\chi_2^{t-2} e^{-\chi_2}}{(t-1)!} \mathrm{d}\chi_2 \\ &\leq \frac{C_9}{P^2} \int_{\frac{1}{P}}^\infty \chi_1^{t-3} e^{-\chi_1} \mathrm{d}\chi_1 \int_0^\infty \frac{e^{-\chi_2} \chi_2^{t-1}}{(t-1)!} \mathrm{d}\chi_2 \\ &+ \frac{C_9}{P^2} \int_{\frac{1}{P}}^\infty \chi_1^{t-2} e^{-\chi_1} \mathrm{d}\chi_1 \int_0^\infty \frac{\chi_2^{t-2} e^{-\chi_2}}{(t-1)!} \mathrm{d}\chi_2 \\ &= \frac{C_9}{P^2} \int_{\frac{1}{P}}^\infty \chi_1^{t-3} e^{-\chi_1} \mathrm{d}\chi_1 + \frac{C_{10}}{P^2} \int_{\frac{1}{P}}^\infty \chi_1^{t-2} e^{-\chi_1} \mathrm{d}\chi_1, \end{aligned}$$

where $C_8 = (t-1)^2 C_2$, $C_9 = \frac{2C_8}{(t-1)!}$ and $C_{10} = \frac{C_9}{t-1}$. When $t \geq 3$, I_2 is upper-bounded by

$$I_2 \leq \frac{C_9}{P^2} \Gamma\left(t-2, \frac{1}{P}\right) + \frac{C_{10}}{P^2} \Gamma\left(t-1, \frac{1}{P}\right), \tag{B.52}$$

where $\Gamma(n,a) = \int_a^\infty x^{n-1} e^{-x} \mathrm{d}x$ is the incomplete gamma function for $n > 0, a > 0$. The following lemma shows an upper bound on the incomplete gamma function, the proof of

which is in Appendix B.5.

Lemma B.5. *For $n > 0, n \in \mathbb{N}$ and $a > 0$, we have*

$$\Gamma(n, a) \leq n! e^{-a} \left(1 + \alpha^{n-1}\right). \tag{B.53}$$

Applying (B.53) to (B.52) yields

$$\begin{aligned} I_2 &\leq \frac{C_9}{P^2}(t-2)! e^{-\frac{1}{P}} \left(1 + \frac{1}{P^{t-3}}\right) + \frac{C_{10}}{P^2}(t-1)! e^{-\frac{1}{P}} \left(1 + \frac{1}{P^{t-2}}\right) \\ &= C_{11}\frac{e^{-\frac{1}{P}}}{P^2} + C_{12}\frac{e^{-\frac{1}{P}}}{P^{t-1}} + C_{13}\frac{e^{-\frac{1}{P}}}{P^t}, \end{aligned} \tag{B.54}$$

where $C_{11} = C_9(t-2)! + C_{10}(t-1)!$, $C_{12} = C_9(t-2)!$ and $C_{13} = C_{10}(t-1)!$. When $t = 2$, the upper bound on I_2 is

$$\begin{aligned} I_2 &\leq \frac{C_9}{P^2}\int_{\frac{1}{P}}^{\infty} \frac{e^{-\chi_1}}{\chi_1} \mathrm{d}\chi_1 + \frac{C_{10}}{P^2}\int_{\frac{1}{P}}^{\infty} e^{-\chi_1} \mathrm{d}\chi_1 \\ &= \frac{C_9}{P^2} \mathtt{E}_1\left(\frac{1}{P}\right) + C_{10}\frac{e^{-\frac{1}{P}}}{P^2} \\ &\leq C_9 \frac{e^{-\frac{1}{P}}}{P^2} \log\left(1 + P\right) + C_{10}\frac{e^{-\frac{1}{P}}}{P^2}, \end{aligned} \tag{B.55}$$

where $\mathtt{E}_1(z) = \int_z^{\infty} \frac{e^{-z}}{z} \mathrm{d}z$ is the exponential integral with an upper bound as $\mathtt{E}_1(z) \leq e^{-z} \log\left(1 + \frac{1}{z}\right)$ [61]. From (B.54) and (B.55), the upper bound on I_2 for any $t \geq 2$ can be obtained as

$$\begin{aligned} I_2 &\leq \left[C_{11}\frac{e^{-\frac{1}{P}}}{P^2} + C_{12}\frac{e^{-\frac{1}{P}}}{P^{t-1}} + C_{13}\frac{e^{-\frac{1}{P}}}{P^t}\right] \times \mathbf{1}_{t \geq 3} + \left[C_9 \frac{e^{-\frac{1}{P}}}{P^2} \log\left(1 + P\right) + C_{10}\frac{e^{-\frac{1}{P}}}{P^2}\right] \times \mathbf{1}_{t=2} \\ &\leq C_{14} e^{-\frac{1}{P}} \left[\frac{1}{P} + \frac{1}{P^{2t}} + \frac{\log\left(1 + P\right)}{P}\right], \end{aligned} \tag{B.56}$$

where $C_{14} = [C_{11} + C_{12} + C_{13}] \times \mathbf{1}_{t \geq 3} + [C_9 + C_{10}] \times \mathbf{1}_{t=2}$. The last inequality is obtained by comparing both cases where $0 < P \leq 1$ and $P > 1$.

To derive I_3, we need an upper bound on $\log\frac{1}{1-p}$ first. By applying (3.13), we obtain

$$\log\frac{1}{1-p}\leq 2(t-1)\log\frac{1}{\min\limits_{m=1,2}\frac{\left|[\mathtt{Full}(\boldsymbol{H})]^{\dagger}\boldsymbol{h}_m\right|^2-\frac{1}{P}}{\chi_m}}.$$

From (B.46), it is found when $\theta\geq\frac{\chi_{m_{\min}}}{\chi_{m_{\max}}}$, $\min\limits_{m=1,2}\frac{\left|[\mathtt{Full}(\boldsymbol{H})]^{\dagger}\boldsymbol{h}_m\right|^2-\frac{1}{P}}{\chi_m}\geq\frac{\chi_{m_{\min}}-\frac{1}{P}}{\chi_{m_{\max}}}=\frac{\gamma(\mathtt{Full}(\boldsymbol{H}),\boldsymbol{H})-\frac{1}{P}}{\chi_{m_{\max}}}$; when $\theta<\frac{\chi_{m_{\min}}}{\chi_{m_{\max}}}$, $\min\limits_{m=1,2}\frac{\left|[\mathtt{Full}(\boldsymbol{H})]^{\dagger}\boldsymbol{h}_m\right|^2-\frac{1}{P}}{\chi_m}=\frac{\gamma(\mathtt{Full}(\boldsymbol{H}),\boldsymbol{H})-\frac{1}{P}}{\chi_{m_{\max}}}$. Therefore,

$$\min_{m=1,2}\frac{\left|\left[\mathtt{Full}\left(\boldsymbol{H}\right)\right]^{\dagger}\boldsymbol{h}_m\right|^2-\frac{1}{P}}{\chi_m}\geq\frac{\gamma\left(\mathtt{Full}\left(\boldsymbol{H}\right),\boldsymbol{H}\right)-\frac{1}{P}}{\chi_{m_{\max}}}.$$

and

$$\log\frac{1}{1-p}\leq 2(t-1)\log\frac{\max\limits_{m=1,2}\chi_m}{\gamma\left(\mathtt{Full}\left(\boldsymbol{H}\right),\boldsymbol{H}\right)-\frac{1}{P}}. \tag{B.57}$$

Define $\mathcal{H}_4=\{\boldsymbol{H}:\boldsymbol{H}\in\mathbb{C}^{t\times 2},\chi_1\geq\chi_2\}$ and $\mathcal{H}_5=\{\boldsymbol{H}:\boldsymbol{H}\in\mathbb{C}^{t\times 2},\chi_1<\chi_2\}$. Substituting (B.57) and (3.12) into I_3 yields

$$\begin{aligned}
I_3&\leq\frac{C_{15}}{P^2}\int_{\boldsymbol{H}\in\mathcal{H}_1\cap\mathcal{H}_4}\left[\frac{1}{\chi_1}+\frac{1}{\chi_2}\right]^2\log\frac{\chi_1}{\gamma\left(\mathtt{Full}\left(\boldsymbol{H}\right),\boldsymbol{H}\right)-\frac{1}{P}}f_{\boldsymbol{H}}\left(\boldsymbol{H}\right)\mathrm{d}\boldsymbol{H}\\
&\quad+\frac{C_{15}}{P^2}\int_{\boldsymbol{H}\in\mathcal{H}_1\cap\mathcal{H}_5}\left[\frac{1}{\chi_1}+\frac{1}{\chi_2}\right]^2\log\frac{\chi_2}{\gamma\left(\mathtt{Full}\left(\boldsymbol{H}\right),\boldsymbol{H}\right)-\frac{1}{P}}f_{\boldsymbol{H}}\left(\boldsymbol{H}\right)\mathrm{d}\boldsymbol{H}\\
&=\frac{2C_{15}}{P^2}\int_{\boldsymbol{H}\in\mathcal{H}_1\cap\mathcal{H}_4}\left[\frac{1}{\chi_1}+\frac{1}{\chi_2}\right]^2\log\frac{\chi_1}{\gamma\left(\mathtt{Full}\left(\boldsymbol{H}\right),\boldsymbol{H}\right)-\frac{1}{P}}f_{\boldsymbol{H}}\left(\boldsymbol{H}\right)\mathrm{d}\boldsymbol{H},
\end{aligned}$$

where $C_{15}=2(t-1)^3C_3$. For any $\boldsymbol{H}\in\mathcal{H}_1\cap\mathcal{H}_4$, $\left[\frac{1}{\chi_1}+\frac{1}{\chi_2}\right]^2\leq\left[\frac{1}{\chi_2}+\frac{1}{\chi_2}\right]^2=\frac{4}{\chi_2^2}$. Therefore, it follows that

$$I_3\leq\frac{C_{16}}{P^2}\int_{\boldsymbol{H}\in\mathcal{H}_1\cap\mathcal{H}_4}\frac{1}{\chi_2^2}\log\frac{\chi_1}{\gamma\left(\mathtt{Full}\left(\boldsymbol{H}\right),\boldsymbol{H}\right)-\frac{1}{P}}f_{\boldsymbol{H}}\left(\boldsymbol{H}\right)\mathrm{d}\boldsymbol{H}, \tag{B.58}$$

where $C_{16}=8C_{15}$.

Define $\mathcal{H}_6 = \left\{ \boldsymbol{H} : \boldsymbol{H} \in \mathcal{H}_1 \bigcap \mathcal{H}_4, \chi_2 \leq \left| \boldsymbol{h}_1^\dagger \boldsymbol{h}_2 \right| \right\}$ and $\mathcal{H}_7 = \left\{ \boldsymbol{H} : \boldsymbol{H} \in \mathcal{H}_1 \bigcap \mathcal{H}_4, \chi_2 > \left| \boldsymbol{h}_1^\dagger \boldsymbol{h}_2 \right| \right\}$. With such notations, $\gamma\left(\texttt{Full}\left(\boldsymbol{H}\right), \boldsymbol{H}\right)$ in [19, Theorem 2] can be rewritten as

$$\gamma\left(\texttt{Full}\left(\boldsymbol{H}\right), \boldsymbol{H}\right) = \begin{cases} \chi_2, & \boldsymbol{H} \in \mathcal{H}_6, \\ \frac{\chi_2}{1+\beta^2}, & \boldsymbol{H} \in \mathcal{H}_7, \end{cases}$$

where $\beta = \frac{\sqrt{\chi_2} - \sqrt{\chi_1 \theta}}{\sqrt{\chi_1 - \chi_1 \theta}}$ and $\theta = \frac{\left| \boldsymbol{h}_1^\dagger \boldsymbol{h}_2 \right|^2}{\chi_1 \chi_2}$. Then the upper bound on I_3 in (B.58) can be further deduced as

$$\begin{aligned} I_3 &\leq \frac{C_{16}}{P^2} \int_{\boldsymbol{H} \in \mathcal{H}_6} \frac{1}{\chi_2^2} \log \frac{\chi_1}{\chi_2 - \frac{1}{P}} f_{\boldsymbol{H}}(\boldsymbol{H}) \mathrm{d}\boldsymbol{H} + \frac{C_{16}}{P^2} \int_{\boldsymbol{H} \in \mathcal{H}_7} \frac{1}{\chi_2^2} \log \frac{\chi_1}{\frac{\chi_2}{1+\beta^2} - \frac{1}{P}} f_{\boldsymbol{H}}(\boldsymbol{H}) \mathrm{d}\boldsymbol{H} \\ &= \underbrace{\frac{C_{16}}{P^2} \int_{\boldsymbol{H} \in \mathcal{H}_6 \cup \mathcal{H}_7} \frac{1}{\chi_2^2} \overbrace{\left(\log \chi_1\right)}^{\leq \chi_1} f_{\boldsymbol{H}}(\boldsymbol{H}) \mathrm{d}\boldsymbol{H}}_{= I_{3,1}} + \underbrace{\frac{C_{16}}{P^2} \int_{\boldsymbol{H} \in \mathcal{H}_6} \frac{1}{\chi_2^2} \log \frac{1}{\chi_2 - \frac{1}{P}} f_{\boldsymbol{H}}(\boldsymbol{H}) \mathrm{d}\boldsymbol{H}}_{= I_{3,2}} \\ &\quad + \underbrace{\frac{C_{16}}{P^2} \int_{\boldsymbol{H} \in \mathcal{H}_7} \frac{1}{\chi_2^2} \log\left(1 + \beta^2\right) f_{\boldsymbol{H}}(\boldsymbol{H}) \mathrm{d}\boldsymbol{H}}_{= I_{3,3}} + \underbrace{\frac{C_{16}}{P^2} \int_{\boldsymbol{H} \in \mathcal{H}_7} \frac{1}{\chi_2^2} \log \frac{1}{\chi_2 - \frac{1+\beta^2}{P}} f_{\boldsymbol{H}}(\boldsymbol{H}) \mathrm{d}\boldsymbol{H}}_{= I_{3,4}}. \end{aligned}$$

Based on the fact that $\{\mathcal{H}_6 \cup \mathcal{H}_7\} = \{\mathcal{H}_1 \cap \mathcal{H}_4\} \subseteq \mathcal{H}_1 \subseteq \mathcal{H}_0$ and using a similar mathematical derivation for the upper bounds on I_1, I_2, the upper bound on $I_{3,1}$ is derived as

$$I_{3,1} \leq C_{17} e^{-\frac{1}{P}} \left[\frac{1}{P} + \frac{1}{P^{2t}} + \frac{\log(1+P)}{P} \right], \tag{B.59}$$

where $C_{17} = \frac{2t}{t-1} C_{16} \times \mathbf{1}_{t \geq 3} + 2 C_{16} \times \mathbf{1}_{t=2}$.

For $I_{3,2}$, since $\mathcal{H}_6 \subseteq \mathcal{H}_0$, its upper bound can be

$$\begin{aligned} I_{3,2} &\leq \frac{C_{16}}{P^2} \int_{\frac{1}{P}}^{\infty} \left[\frac{\chi_1^{t-1} e^{-\chi_1}}{(t-1)!} \right] \mathrm{d}\chi_1 \int_{\frac{1}{P}}^{\infty} \frac{1}{\chi_2^2} \log \frac{1}{\chi_2 - \frac{1}{P}} \left[\frac{\chi_2^{t-1} e^{-\chi_2}}{(t-1)!} \right] \mathrm{d}\chi_2 \\ &\leq \frac{C_{16}}{(t-1)! P^2} \int_0^{\infty} \frac{\chi_1^{t-1} e^{-\chi_1}}{(t-1)!} \mathrm{d}\chi_1 \int_{\frac{1}{P}}^{\infty} \chi_2^{t-3} e^{-\chi_2} \log \frac{1}{\chi_2 - \frac{1}{P}} \mathrm{d}\chi_2 \end{aligned}$$

$$= C_{18}\frac{e^{-\frac{1}{P}}}{P^2}\int_0^{\infty}\left(\log\frac{1}{\lambda_2}\right)\left(\lambda_2+\frac{1}{P}\right)^{t-3}e^{-\lambda_2}\mathrm{d}\lambda_2, \tag{B.60}$$

where $C_{18} = \frac{C_{16}}{(t-1)!}$ and the last equality is obtained by replacing $\chi_2 - \frac{1}{P}$ with λ_2. When $t \geq 4$, with the help of (B.53), we obtain

$$\begin{aligned}
I_{3,2} &\leq C_{18}\frac{e^{-\frac{1}{P}}}{P^2}\int_0^{\frac{1}{P}}\left(\log\frac{1}{\lambda_2}\right)\left(\lambda_2+\frac{1}{P}\right)^{t-3}e^{-\lambda_2}\mathrm{d}\lambda_2 \\
&+ C_{18}\frac{e^{-\frac{1}{P}}}{P^2}\int_{\frac{1}{P}}^{\infty}\left(\log\frac{1}{\lambda_2}\right)\left(\lambda_2+\frac{1}{P}\right)^{t-3}e^{-\lambda_2}\mathrm{d}\lambda_2 \\
&\leq C_{18}\frac{e^{-\frac{1}{P}}}{P^2}\int_0^{\frac{1}{P}}\left(\log\frac{1}{\lambda_2}\right)\left(\frac{1}{P}+\frac{1}{P}\right)^{t-3}\mathrm{d}\lambda_2 \\
&+ C_{18}\frac{e^{-\frac{1}{P}}}{P^2}\int_{\frac{1}{P}}^{\infty}\frac{1}{\lambda_2}\left(\lambda_2+\lambda_2\right)^{t-3}e^{-\lambda_2}\mathrm{d}\lambda_2 \\
&\leq \frac{2^{t-3}C_{18}e^{-\frac{1}{P}}}{P^{t-1}}\int_0^{\frac{1}{P}}\log\frac{1}{\lambda_2}\mathrm{d}\lambda_2 + \frac{2^{t-3}C_{18}}{P^2}\int_{\frac{1}{P}}^{\infty}\lambda_2^{t-4}e^{-\lambda_2}\mathrm{d}\lambda_2 \\
&= \frac{C_{19}e^{-\frac{1}{P}}}{P^{t-1}}\left[\frac{1}{P}+\frac{\log P}{P}\right] + \frac{C_{19}}{P^2}\Gamma\left(t-3,\frac{1}{P}\right) \\
&\leq \frac{C_{19}e^{-\frac{1}{P}}}{P^{t-1}}\left[\frac{1}{P}+1\right] + \frac{(t-3)!C_{19}e^{-\frac{1}{P}}}{P^2}\left[1+\frac{1}{P^{t-4}}\right] \\
&\leq C_{20}e^{-\frac{1}{P}}\left[\frac{1}{P}+\frac{1}{P^{2t}}\right],
\end{aligned} \tag{B.61}$$

where $C_{19} = 2^{t-3}C_{18}$ and $C_{20} = 2\times(t-3)!C_{19} + 2C_{19}$. When $t = 3$, (B.60) becomes

$$\begin{aligned}
I_{3,2} &\leq C_{18}\frac{e^{-\frac{1}{P}}}{P^2}\int_0^{\frac{1}{P}}\left(\log\frac{1}{\lambda_2}\right)e^{-\lambda_2}\mathrm{d}\lambda_2 + C_{18}\frac{e^{-\frac{1}{P}}}{P^2}\int_{\frac{1}{P}}^{\infty}\left(\log\frac{1}{\lambda_2}\right)e^{-\lambda_2}\mathrm{d}\lambda_2 \\
&\leq C_{18}\frac{e^{-\frac{1}{P}}}{P^2}\int_0^{\frac{1}{P}}\log\frac{1}{\lambda_2}\mathrm{d}\lambda_2 + \frac{C_{18}}{P^2}\int_{\frac{1}{P}}^{\infty}\frac{e^{-\lambda_2}}{\lambda_2}\mathrm{d}\lambda_2 \\
&\leq C_{18}\frac{e^{-\frac{1}{P}}}{P^2}\left[\frac{1}{P}+\frac{\log P}{P}\right] + \frac{C_{18}}{P^2}\mathtt{E}_1\left(\frac{1}{P}\right) \\
&\leq C_{18}\frac{e^{-\frac{1}{P}}}{P^2}\left[\frac{1}{P}+1\right] + \frac{C_{18}}{P^2}e^{-\frac{1}{P}}\log\left(1+P\right) \\
&\leq C_{21}e^{-\frac{1}{P}}\left[\frac{1}{P}+\frac{1}{P^{2t}}+\frac{\log(1+P)}{P}\right],
\end{aligned} \tag{B.62}$$

where $C_{21} = 3C_{18}$. When $t = 2$, since $\frac{1}{\lambda_2 + \frac{1}{P}} \leq P$, $I_{3,2} \leq C_{18} \frac{e^{-\frac{1}{P}}}{P} \int_0^\infty \log \frac{1}{\lambda_2} e^{-\lambda_2} \mathrm{d}\lambda_2$. Following the same steps in (B.62), $I_{3,2}$ can be bounded by

$$I_{3,2} \leq C_{22} e^{-\frac{1}{P}} \left[\frac{1}{P} + \frac{1}{P^{2t}} + \frac{\log(1+P)}{P} \right], \tag{B.63}$$

where $C_{22} = 3C_{18}$. Based on (B.61), (B.62) and (B.63), the upper bound on $I_{3,2}$ for any $t \geq 2$ is

$$I_{3,2} \leq C_{23} e^{-\frac{1}{P}} \left[\frac{1}{P} + \frac{1}{P^{2t}} + \frac{\log(1+P)}{P} \right], \tag{B.64}$$

where $C_{23} = C_{20} \times \mathbf{1}_{t \geq 4} + C_{21} \times \mathbf{1}_{t=3} + C_{22} \times \mathbf{1}_{t=2}$.

In $I_{3,3}$, since $0 \leq \beta \leq 1$, $\log(1 + \beta^2) \leq \log 2 < 1$. Similarly, the bound on $I_{3,3}$ is obtained as

$$I_{3,3} \leq C_{24} e^{-\frac{1}{P}} \left[\frac{1}{P} + \frac{1}{P^{2t}} + \frac{\log(1+P)}{P} \right], \tag{B.65}$$

where $C_{24} = \frac{2C_{16}}{t-1} \times \mathbf{1}_{t \geq 3} + \frac{C_{16}}{(t-1)!} \times \mathbf{1}_{t=2}$.

For $I_{3,4}$, since χ_1, χ_2 and θ are mutually independent, its upper bound can be derived as

$$\begin{aligned} I_{3,4} &\leq \frac{C_{16}}{P^2} \int_{\{\chi_1, \chi_2, \theta\} \in \mathcal{H}_7'} \frac{1}{\chi_2^2} \left(\log \frac{1}{\chi_2 - \frac{1+\beta^2}{P}} \right) f_{\chi_1}(\chi_1) f_{\chi_2}(\chi_2) f_\theta(\theta) \mathrm{d}\chi_1 \mathrm{d}\chi_2 \mathrm{d}\theta \\ &= \frac{C_{25}}{P^2} \int_{\{\chi_1, \chi_2, \theta\} \in \mathcal{H}_7'} \left(\log \frac{1}{\chi_2 - \frac{1+\beta^2}{P}} \right) \chi_1^{t-1} e^{-\chi_1} \chi_2^{t-3} e^{-\chi_2} (1-\theta)^{t-2} \mathrm{d}\chi_1 \mathrm{d}\chi_2 \mathrm{d}\theta, \end{aligned}$$

where $C_{25} = \frac{C_{16}}{(t-1)!(t-2)!}$ and $\mathcal{H}_7'$ is a transformed version of the pre-defined $\mathcal{H}_7$ with respect to χ_1, χ_2 and θ. The PDF of θ is given by $f_\theta(\theta) = (t-1)(1-\theta)^{t-2}$ for $0 \leq \theta \leq 1$[21]. By changing the integration variables from (χ_1, χ_2, θ) into (β, χ_2, θ), we obtain the Jacobian of the transformation as $\left| \frac{\partial(\chi_1, \chi_2, \theta)}{\partial(\beta, \chi_2, \theta)} \right| = \left| \frac{\partial \chi_1}{\partial \beta} \right|$. For any $\boldsymbol{H} \in \mathcal{H}_7'$, $\beta = \frac{\sqrt{\chi_2} - \sqrt{\chi_1 \theta}}{\sqrt{\chi_1 - \chi_1 \theta}}$, $\chi_1 = \frac{\chi_2}{\left(\sqrt{\theta} + \beta\sqrt{1-\theta}\right)^2}$

and $\left|\frac{\partial\chi_1}{\partial\beta}\right| = \frac{2\sqrt{1-\theta}\chi_2}{\left(\sqrt{\theta}+\beta\sqrt{1-\theta}\right)^3}$. Therefore, $I_{3,4}$ can be bounded by

$$\begin{aligned}
I_{3,4} &\leq \frac{C_{26}}{P^2}\int_{\{\beta,\chi_2,\theta\}\in\mathcal{H}_7''}\left(\log\frac{1}{\chi_2-\frac{1+\beta^2}{P}}\right)\left[\frac{\chi_2}{(\sqrt{\theta}+\beta\sqrt{1-\theta})^2}\right]^{t-1}e^{-\frac{\chi_2}{(\sqrt{\theta}+\beta\sqrt{1-\theta})^2}}\chi_2^{t-3}e^{-\chi_2}\\
&\quad\times(1-\theta)^{t-2}\frac{\sqrt{1-\theta}\chi_2}{\left(\sqrt{\theta}+\beta\sqrt{1-\theta}\right)^3}\mathrm{d}\chi_2\mathrm{d}\beta\mathrm{d}\theta\\
&= \frac{C_{26}}{P^2}\int_{\{\beta,\chi_2,\theta\}\in\mathcal{H}_7''}\left(\log\frac{1}{\chi_2-\frac{1+\beta^2}{P}}\right)\frac{\chi_2^{2t-3}e^{-\chi_2}(1-\theta)^{t-\frac{3}{2}}}{\left(\sqrt{\theta}+\beta\sqrt{1-\theta}\right)^{2t+1}}e^{-\frac{\chi_2}{(\sqrt{\theta}+\beta\sqrt{1-\theta})^2}}\mathrm{d}\chi_2\mathrm{d}\beta\mathrm{d}\theta\\
&\leq \frac{C_{26}}{P^2}\int_{\{\beta,\chi_2,\theta\}\in\mathcal{H}_7''}\left(\log\frac{1}{\chi_2-\frac{1+\beta^2}{P}}\right)\frac{\chi_2^{2t-3}e^{-\chi_2}}{\left(\sqrt{\theta}+\beta\sqrt{1-\theta}\right)^{2t+1}}e^{-\frac{\chi_2}{(\sqrt{\theta}+\beta\sqrt{1-\theta})^2}}\mathrm{d}\chi_2\mathrm{d}\beta\mathrm{d}\theta,
\end{aligned}$$

where $C_{26}=2C_{25}$ and $\mathcal{H}_7''$ is a transformed version of $\mathcal{H}_7'$ with respect to β,χ_2 and θ. By replacing $\chi_2-\frac{1+\beta^2}{P}$ by χ, $\left|\frac{\partial(\beta,\chi_2,\theta)}{\partial(\beta,\chi,\theta)}\right| = \left|\frac{\partial\chi_2}{\partial\chi}\right| = 1$, then $I_{3,4}$ is further bounded by

$$\begin{aligned}
I_{3,4} &\leq \frac{C_{26}}{P^2}\int_{\{\beta,\chi,\theta\}\in\mathcal{H}_7'''}\left(\log\frac{1}{\chi}\right)\frac{e^{-\frac{\chi+\frac{1+\beta^2}{P}}{(\sqrt{\theta}+\beta\sqrt{1-\theta})^2}}}{\left(\sqrt{\theta}+\beta\sqrt{1-\theta}\right)^{2t+1}}\\
&\quad\times\left[\chi+\frac{1+\beta^2}{P}\right]^{2t-3}e^{-\chi-\frac{1+\beta^2}{P}}\mathrm{d}\chi\mathrm{d}\beta\mathrm{d}\theta\\
&\leq \frac{C_{26}}{P^2}\int_{\{\beta,\chi,\theta\}\in\mathcal{H}_7'''}\left(\log\frac{1}{\chi}\right)\frac{e^{-\frac{\chi+\frac{1+\beta^2}{P}}{(\sqrt{\theta}+\beta\sqrt{1-\theta})^2}}}{\left(\sqrt{\theta}+\beta\sqrt{1-\theta}\right)^{2t+1}}\\
&\quad\times\left[\chi+\frac{1+\beta^2}{P}\right]^{2t-3}e^{-\chi-\frac{1}{P}}\mathrm{d}\chi\mathrm{d}\beta\mathrm{d}\theta,
\end{aligned}\tag{B.66}$$

where $\mathcal{H}_7'''$ is a transformed version of $\boldsymbol{H}_7''$ with respect to β,χ and θ. Letting $\phi=\frac{\chi+\frac{1+\beta^2}{P}}{\left(\sqrt{\theta}+\beta\sqrt{1-\theta}\right)^2}$, $\left|\frac{\partial(\chi,\beta,\theta)}{\partial(\chi,\beta,\phi)}\right| = \left|\frac{\partial\theta}{\partial\phi}\right|$. Since $\frac{\sqrt{\chi+\frac{1+\beta^2}{P}}}{\sqrt{\phi}} = \sqrt{\theta}+\beta\sqrt{1-\theta}$, $\left|\frac{\partial\theta}{\partial\phi}\right| = \frac{\phi^{-\frac{3}{2}}\sqrt{\chi+\frac{1+\beta^2}{P}}}{\left|\frac{1}{\sqrt{\theta}}-\frac{\beta}{\sqrt{1-\theta}}\right|}$. For any $\boldsymbol{H}\in\mathcal{H}_7'''$, $\chi_1\geq\chi_2$, thus $\phi=\frac{\chi_1}{\chi_2}\geq 1$ and $0\leq\sqrt{\theta}+\beta\sqrt{1-\theta}\leq 1$. Then $0\leq\beta\leq\frac{1-\sqrt{\theta}}{\sqrt{1-\theta}}$. Hence, $\frac{1}{\sqrt{\theta}}-\frac{\beta}{\sqrt{1-\theta}}\geq\frac{1}{\sqrt{\theta}}-\frac{1}{\sqrt{1-\theta}}\times\frac{1-\sqrt{\theta}}{\sqrt{1-\theta}} = \frac{1}{(1+\sqrt{\theta})\sqrt{\theta}} > 0$. Therefore, $\left|\frac{\partial\theta}{\partial\phi}\right| \leq \phi^{-\frac{3}{2}}\sqrt{\chi+\frac{1+\beta^2}{P}}(1+$

$\sqrt{\theta})\sqrt{\theta} \leq 2\phi^{-\frac{3}{2}}\sqrt{\chi + \frac{1+\beta^2}{P}}$ due to $0 \leq \theta \leq 1$. Moreover, since

$$\mathcal{H}_7''' \subseteq \{(\beta, \chi, \phi) : 0 \leq \beta \leq 1, \chi > 0, \phi > 0\},$$

the upper bound in (B.66) becomes

$$\begin{aligned}
I_{3,1,4} &\leq 2C_{26}\frac{e^{-\frac{1}{P}}}{P^2}\int_{\{\beta,\chi,\phi\}\in\mathcal{H}_7'''}\left(\log\frac{1}{\chi}\right)e^{-\phi}\left[\chi + \frac{1+\beta^2}{P}\right]^{t-3}\phi^{t-1}e^{-\chi}\mathrm{d}\chi\mathrm{d}\beta\mathrm{d}\phi \\
&\leq 2C_{26}\frac{e^{-\frac{1}{P}}}{P^2}\int_0^\infty\int_0^1\int_0^\infty\left(\log\frac{1}{\chi}\right)e^{-\phi}\left[\chi + \frac{1+\beta^2}{P}\right]^{t-3}\phi^{t-1}e^{-\chi}\mathrm{d}\chi\mathrm{d}\beta\mathrm{d}\phi \\
&= 2C_{26}\frac{e^{-\frac{1}{P}}}{P^2}\left[\int_0^\infty\phi^{t-1}e^{-\phi}\mathrm{d}\phi\right]\int_0^\infty\int_0^1\left(\log\frac{1}{\chi}\right)\left[\chi + \frac{1+\beta^2}{P}\right]^{t-3}e^{-\chi}\mathrm{d}\chi\mathrm{d}\beta \\
&\leq C_{27}\frac{e^{-\frac{1}{P}}}{P^2}\int_0^\infty\int_0^1\left(\log\frac{1}{\chi}\right)\left[\chi + \frac{1+\beta^2}{P}\right]^{t-3}e^{-\chi}\mathrm{d}\chi\mathrm{d}\beta,
\end{aligned}$$

where $C_{27} = 2(t-1)!C_{26}$. When $t \geq 4$, $\left[\chi + \frac{1+\beta^2}{P}\right]^{t-3} \leq \left[\chi + \frac{2}{P}\right]^{t-3}$ due to $0 \leq \beta \leq 1$. Similar to (B.61), an upper bound on $I_{3,1,4}$ is derived as

$$I_{3,4} \leq C_{28}e^{-\frac{1}{P}}\left[\frac{1}{P} + \frac{1}{P^{2t}}\right], \tag{B.67}$$

where $C_{28} = 2^{2t-5}C_{27} + (t-4)!2^{t-2}C_{27}$. When $t = 3$, similar to (B.62), the upper bound on $I_{3,4}$ is

$$I_{3,4} \leq C_{29}e^{-\frac{1}{P}}\left[\frac{1}{P} + \frac{1}{P^{2t}} + \frac{\log(1+P)}{P}\right], \tag{B.68}$$

where $C_{29} = 3C_{27}$. When $t = 2$, since $\frac{1}{\chi+\frac{1+\beta^2}{P}} \leq \frac{P}{1+\beta^2} \leq P$, $I_{3,4} \leq C_{27}\frac{e^{-\frac{1}{P}}}{P}\int_0^\infty\left(\log\frac{1}{\chi}\right)e^{-\chi}\mathrm{d}\chi$. Similar to (B.62), we obtain

$$I_{3,4} \leq C_{30}e^{-\frac{1}{P}}\left[\frac{1}{P} + \frac{1}{P^{2t}} + \frac{\log(1+P)}{P}\right], \tag{B.69}$$

where $C_{30} = 3C_{27}$. Combining bounds in (B.67), (B.68) and (B.69), the upper bound on $I_{3,4}$

for any $t \geq 2$ is

$$I_{3,4} \leq C_{31} e^{-\frac{1}{P}} \left[\frac{1}{P} + \frac{1}{P^{2t}} + \frac{\log(1+P)}{P} \right], \tag{B.70}$$

where $C_{31} = C_{28} \times \mathbf{1}_{t \geq 4} + C_{29} \times \mathbf{1}_{t=3} + C_{30} \times \mathbf{1}_{t=2}$. Based on (B.59), (B.64), (B.65) and (B.70), I_3 is upper-bounded by

$$I_3 \leq C_{32} e^{-\frac{1}{P}} \left[\frac{1}{P} + \frac{1}{P^{2t}} + \frac{\log(1+P)}{P} \right], \tag{B.71}$$

where $C_{32} = C_{19} + C_{23} + C_{26} + C_{31}$. From (B.51), (B.56) and (B.71), we finally get the upper bound in (3.15), where $C_0 = C_7 + C_{14} + C_{32}$. □

B.4 Proof of Lemma B.4

Proof. We use the following lemma, the proof of which is given in Appendix B.6.

Lemma B.6. *For unit-normal complex vectors $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \mathbb{C}^{t \times 1}$, we have*

$$\left| |\boldsymbol{u}^\dagger \boldsymbol{v}|^2 - |\boldsymbol{u}^\dagger \boldsymbol{w}|^2 \right| \leq \sqrt{1 - |\boldsymbol{v}^\dagger \boldsymbol{w}|^2}. \tag{B.72}$$

For any $\boldsymbol{H}$ satisfying $\gamma\left(\texttt{Full}\left(\boldsymbol{H}\right), \boldsymbol{H}\right) > \frac{1}{P}$, let $\Delta_m = \left| \left[\texttt{Full}\left(\boldsymbol{H}\right)\right]^\dagger \boldsymbol{h}_m \right|^2 - \frac{1}{P}$, where $0 < \frac{\Delta_m}{\chi_m} < 1$ for $m = 1, 2$. If $|\boldsymbol{x}^\dagger \texttt{Full}\left(\boldsymbol{H}\right)|^2 \geq \Pi = 1 - \min\limits_{m=1,2} \left[\frac{\Delta_m}{\chi_m}\right]^2$, by applying (B.72) and letting $\boldsymbol{u} = \frac{\boldsymbol{h}_m}{|\boldsymbol{h}_m|}$, $\boldsymbol{v} = \boldsymbol{x}$, $\boldsymbol{w} = \texttt{Full}\left(\boldsymbol{H}\right)$, we derive that

$$\begin{aligned} &\left| \left| \frac{\boldsymbol{h}_m^\dagger}{|\boldsymbol{h}_m|} \boldsymbol{x} \right|^2 - \left| \frac{\boldsymbol{h}_m^\dagger}{|\boldsymbol{h}_m|} \texttt{Full}\left(\boldsymbol{H}\right) \right|^2 \right| \leq \sqrt{1 - |\boldsymbol{x}^\dagger \texttt{Full}\left(\boldsymbol{H}\right)|^2} \\ \Longrightarrow &\left| \frac{\boldsymbol{h}_m^\dagger}{|\boldsymbol{h}_m|} \boldsymbol{x} \right|^2 \geq \left| \frac{\boldsymbol{h}_m^\dagger}{|\boldsymbol{h}_m|} \texttt{Full}\left(\boldsymbol{H}\right) \right|^2 - \sqrt{1 - |\boldsymbol{x}^\dagger \texttt{Full}\left(\boldsymbol{H}\right)|^2} \end{aligned}$$

$$\Longrightarrow \left| \frac{\boldsymbol{h}_m^\dagger}{|\boldsymbol{h}_m|} \boldsymbol{x} \right|^2 \geq \frac{1}{P\chi_m} + \frac{\Delta_m}{\chi_m} - \sqrt{1-\Pi} \geq \frac{1}{P\chi_m}$$
$$\Longrightarrow \left| \boldsymbol{h}_m^\dagger \boldsymbol{x} \right|^2 \geq \frac{1}{P},$$

where "$\Longrightarrow$" represents "it follows that". Since $0 < \Pi < 1$, the proof is complete. □

B.5 Proof of Lemma B.5

Proof. $\Gamma(n,a)$ can be expanded as $\Gamma(n,a) = (n-1)!e^{-a}\sum_{k=0}^{n-1}\frac{a^k}{k!}$ [61]. When $0 < a \leq 1$, $\Gamma(n,a) \leq (n-1)!e^{-a}\sum_{k=0}^{n-1}\frac{1}{k!} \leq n!e^{-a}$; when $\alpha > 1$, $\Gamma(n,a) \leq (n-1)!e^{-a}\sum_{k=0}^{n-1}\alpha^k \leq (n-1)!e^{-a}\sum_{k=0}^{n-1}\alpha^{n-1} = n!e^{-a}\alpha^{n-1}$. Therefore, $\Gamma(n,a) \leq \max\left\{n!e^{-a}, n!e^{-a}\alpha^{n-1}\right\} \leq n!e^{-a} + n!e^{-a}\alpha^{n-1} = n!e^{-a}\left(1+\alpha^{n-1}\right)$. □

B.6 Proof of Lemma B.6

Proof. Let $\boldsymbol{G} \triangleq \boldsymbol{v}\boldsymbol{v}^\dagger - \boldsymbol{w}\boldsymbol{w}^\dagger$ and $z \triangleq \boldsymbol{v}^\dagger\boldsymbol{w}$. It can be verified (after some tedious but straightforward calculations) that $\boldsymbol{G}$ admits the decomposition

$$\boldsymbol{G} = \sqrt{1-|z|^2}\left(\boldsymbol{u}_1\boldsymbol{u}_1^\dagger - \boldsymbol{u}_2\boldsymbol{u}_2^\dagger\right),$$

where

$$\boldsymbol{u}_1 = \alpha\boldsymbol{v} - \beta\boldsymbol{v}_0\exp\left(-j\angle z\right),$$
$$\boldsymbol{u}_2 = \beta\boldsymbol{v} + \alpha\boldsymbol{v}_0\exp\left(-j\angle z\right)$$

are orthonormal vectors with

$$\boldsymbol{v}_0 = \frac{\boldsymbol{w} - \boldsymbol{v}\boldsymbol{v}^\dagger \boldsymbol{w}}{\sqrt{1-|z|^2}},$$
$$(\alpha, \beta) = \left(\sqrt{\frac{1+\sqrt{1-|z|^2}}{2}}, \sqrt{\frac{1-\sqrt{1-|z|^2}}{2}} \right).$$

We can then obtain

$$\begin{aligned}
\left| |\boldsymbol{u}^\dagger \boldsymbol{v}|^2 - |\boldsymbol{u}^\dagger \boldsymbol{w}|^2 \right| &= \left| \boldsymbol{u}^\dagger \boldsymbol{G} \boldsymbol{u} \right| \\
&= \sqrt{1-|z|^2} \left| |\boldsymbol{u}^\dagger \boldsymbol{u}_1|^2 - |\boldsymbol{u}^\dagger \boldsymbol{u}_2|^2 \right| \\
&\leq \sqrt{1-|z|^2} \left(|\boldsymbol{u}^\dagger \boldsymbol{u}_1|^2 + |\boldsymbol{u}^\dagger \boldsymbol{u}_2|^2 \right) \\
&\leq \sqrt{1-|z|^2} \, ||\boldsymbol{u}||^2 \\
&= \sqrt{1-|z|^2}.
\end{aligned}$$

This concludes the proof. □

C Supplementary Proofs for Chapter 4

C.1 Proof of Propositions 4.1 and 4.2

Proof. We first prove Proposition 4.1. Since the optimal time-sharing pair that minimizes $\mathsf{OUT}_{MR,\mathsf{TS}}$ should maximize $MR_{\mathsf{TS}}(t_1, t_2)$, we have $(t_1^\star, t_2^\star) = \arg\max_{t_1+t_2=1, t_1, t_2 \geq 0} MR_{\mathsf{TS}}(t_1, t_2)$. Clearly, $t_2^\star = 1 - t_1^\star$, and therefore, it suffices to determine $t_1^\star = \arg\max_{0 \leq t_1 \leq 1} MR_{\mathsf{TS}}(t_1, 1-t_1)$, where

$$MR_{\mathsf{TS}}(t_1, 1-t_1) = \min \left\{ t_1 \log_2 \left(1 + P|h_{11}|^2\right), (1-t_1) \log_2 \left(1 + P|h_{22}|^2\right) \right\}.$$

As t_1 increases from 0 to 1, the term $t_1 \log_2\left(1+P|h_{11}|^2\right)$ increases from 0 to $\log_2\left(1+P|h_{11}|^2\right)$; the term $(1-t_1)\log_2\left(1+P|h_{22}|^2\right)$ decreases from $\log_2\left(1+P|h_{22}|^2\right)$ to 0. Since either term is also a continous function of t_1, it follows that the optimal solution $t_1^\star$ should make the two terms equal. In other words, $t_1^\star$ should satisfy $t_1^\star \log_2\left(1+P|h_{11}|^2\right) = (1-t_1^\star)\log_2\left(1+P|h_{22}|^2\right)$. This leads to $t_1^\star$ and $t_2^\star$ as given by (4.3), and thus concludes the proof of Proposition 4.1.

We now move to the proof of Proposition 4.2. Similarly, the optimal power pair that minimizes $\mathsf{OUT}_{MR,\mathsf{CT}}$ should maximize $MR_{\mathsf{CT}}(p_1,p_2)$. We first show $p_1^\star = 1$ or $p_2^\star = 1$ by contradiction. Assume $0 < p_1^\star, p_2^\star < 1$. Letting $\beta \triangleq \min\left\{\frac{1}{p_1^\star}, \frac{1}{p_2^\star}\right\} > 1$, we have

$$\begin{aligned} MR_{\mathsf{CT}}\left(\beta p_1^\star, \beta p_2^\star\right) &= \min\left\{\log_2\left(1+\frac{p_1^\star|h_{11}|^2}{p_2^\star|h_{21}|^2+\frac{1}{\beta P}}\right), \log_2\left(1+\frac{p_2^\star|h_{22}|^2}{p_1^\star|h_{12}|^2+\frac{1}{\beta P}}\right)\right\} \\ &> \min\left\{\log_2\left(1+\frac{p_1^\star|h_{11}|^2}{p_2^\star|h_{21}|^2+\frac{1}{P}}\right), \log_2\left(1+\frac{p_2^\star|h_{22}|^2}{p_1^\star|h_{12}|^2+\frac{1}{P}}\right)\right\} \\ &= MR_{\mathsf{CT}}\left(p_1^\star, p_2^\star\right), \end{aligned} \tag{C.73}$$

which contradicts the fact that $(p_1^\star, p_2^\star)$ is optimal. Therefore, $p_1^\star = 1$ or $p_2^\star = 1$.

Now, let $\alpha(p_1,p_2) \triangleq \frac{p_1|h_{11}|^2}{p_2|h_{21}|^2+\frac{1}{P}}$ and $\beta(p_1,p_2) \triangleq \frac{p_2|h_{22}|^2}{p_1|h_{12}|^2+\frac{1}{P}}$. We also define the variables

$$\widetilde{p}_1 \triangleq \arg\max_{0\le p_1\le 1} MR_{\mathsf{CT}}(p_1,1) = \arg\max_{0\le p_1\le 1} \min\left\{\alpha(p_1,1), \beta(p_1,1)\right\}, \tag{C.74}$$

$$\widetilde{p}_2 \triangleq \arg\max_{0\le p_2\le 1} MR_{\mathsf{CT}}(1,p_2) \arg\max_{0\le p_2\le 1} \min\left\{\alpha(1,p_2), \beta(1,p_2)\right\}. \tag{C.75}$$

Since either $p_1^\star = 1$ or $p_2^\star = 1$, we have $(p_1^\star, p_2^\star) = (\widetilde{p}_1, 1)$ if $MR_{\mathsf{CT}}(\widetilde{p}_1,1) > MR_{\mathsf{CT}}(1,\widetilde{p}_2)$, and otherwise, $(p_1^\star, p_2^\star) = (1, \widetilde{p}_2)$ if $MR_{\mathsf{CT}}(\widetilde{p}_1,1) \le MR_{\mathsf{CT}}(1,\widetilde{p}_2)$.

Now, suppose $\alpha(1,1) \le \beta(1,1)$. We will evaluate the corresponding $\widetilde{p}_1$ and $\widetilde{p}_2$, and show that $MR_{\mathsf{CT}}(\widetilde{p}_1,1) \le MR_{\mathsf{CT}}(1,\widetilde{p}_2)$. This will imply (according to our discussion above) that $(p_1^\star, p_2^\star) = (1, \widetilde{p}_2)$. Let us first evaluate $\widetilde{p}_1$ as defined in (C.74). Note that as p_1 increases from 0 to 1, the quantity $\alpha(p_1,1)$ increases from $\alpha(0,1) = 0$ to $\alpha(1,1)$, and $\beta(p_1,1)$ decreases from

$\beta(0,1)$ to $\beta(1,1)$. Hence, if $\alpha(1,1) \leq \beta(1,1)$, we have $\alpha(p_1,1) \leq \beta(p_1,1)$, $\forall p_1 \in [0,1]$, and thus $\widetilde{p}_1 = 1$ with $MR_{\mathsf{CT}}(\widetilde{p}_1,1) = \alpha(1,1)$ whenever $\alpha(1,1) \leq \beta(1,1)$. We now evaluate $\widetilde{p}_2$ as defined in (C.75). As p_2 increases from 0 to 1, the quantity $\alpha(1,p_2)$ decreases from $\alpha(1,0)$ to $\alpha(1,1) > 0$, while $\beta(1,p_2)$ increases from $\beta(1,0) = 0$ to $\beta(1,1) > 0$. Hence, when $\alpha(1,1) \leq \beta(1,1)$, the solution $\widetilde{p}_2$ of the optimization problem in (C.75) should satisfy $\alpha(1,\widetilde{p}_2) = \beta(1,\widetilde{p}_2)$ with $MR_{\mathsf{CT}}(1,\widetilde{p}_2) \geq \alpha(1,1)$. Substituting the explicit expressions for $\alpha(1,\widetilde{p}_2)$ and $\beta(1,\widetilde{p}_2)$, we obtain $\widetilde{p}_2 = \frac{1}{2P|h_{21}|^2}\left(\sqrt{\frac{4P^2|h_{11}|^2|h_{12}|^2|h_{21}|^2+4P|h_{11}|^2|h_{21}|^2}{|h_{22}|^2}+1}-1\right)$. Hence, when $\alpha(1,1) \leq \beta(1,1)$, we have $MR_{\mathsf{CT}}(1,\widetilde{p}_2) \geq \alpha(1,1) = MR_{\mathsf{CT}}(\widetilde{p}_1,1)$, which implies $(p_1^\star, p_2^\star) = (1,\widetilde{p}_2)$, and thus proves (4.5). The proof of (4.4) can be accomplished in the same manner and is thus omitted for brevity. This concludes the proof of Proposition 4.2. □

C.2 Proof of Theorem 4.2

Proof. To prove Theorem 4.2, we will need the following lemma.

Lemma C.7. *When $0 < t_{k,\min} < 1$ for $k = 1,2$, let $t_{k,\min} = [0.b_{k,1}b_{k,2}\ldots]_2$. Then, at Round l of $CQ_{MR,\mathsf{TS}}$ where $l \geq 1$, we have $ENC^l_{MR,\mathsf{TS},k}(\boldsymbol{h}_k) = b_{k,l}$, $t^{\mathsf{lb},l}_{k,\min} = [0.b_{k,1}b_{k,2}\ldots b_{k,l}]_2$ and $t^{\mathsf{ub},l}_{k,\min} = t^{\mathsf{lb},l}_{k,\min} + 2^{-l}$ for $k = 1,2$.*

Proof. We will prove Lemma C.7 by induction. It is straightforward to verify Lemma C.7 when $l = 1$. Suppose it also holds when $l \leq m$ where $m \geq 1$. Thus, $t^{\mathsf{lb},m}_{k,\min} = [0.b_{k,1}b_{k,2}\ldots b_{k,m}]_2$ and $t^{\mathsf{ub},m}_{k,\min} = t^{\mathsf{lb},m}_{k,\min} + 2^{-m}$. When $l = m+1$, we have $ENC^{m+1}_{MR,\mathsf{TS},k}(\boldsymbol{h}_k) = \mathbf{1}\left(t_{k,\min} \geq \frac{t^{\mathsf{ub},m}_{k,\min}+t^{\mathsf{lb},m}_{k,\min}}{2}\right)$, where

$$\frac{t^{\mathsf{ub},m}_{k,\min}+t^{\mathsf{lb},m}_{k,\min}}{2} = \frac{2\times[0.b_{k,1}b_{k,2}\ldots b_{k,m}]_2 + 2^{-m}}{2} = [0.b_{k,1}b_{k,2}\ldots b_{k,m}1]_2\,.$$

If $ENC^{m+1}_{MR,\mathsf{TS},k}(\boldsymbol{h}_k) = 1$, $t_{k,\min} \geq \frac{t^{\mathsf{ub},m}_{k,\min}+t^{\mathsf{lb},m}_{k,\min}}{2}$, then, $t_{k,\min} = [0.b_{k,1}b_{k,2}\ldots b_{k,m}b_{k,m+1}\ldots]_2 \geq$

$[0.b_{k,1}b_{k,2}\dots b_{k,m}1]_2$, thus, $b_{k,m+1}=1=ENC^{m+1}_{MR,\mathsf{TS},k}(\boldsymbol{h}_k)$. In addition, $t^{\mathsf{lb},m+1}_{k,\min}$ and $t^{\mathsf{ub},m+1}_{k,\min}$ are updated as $t^{\mathsf{lb},m+1}_{k,\min}=\frac{t^{\mathsf{ub},m}_{k,\min}+t^{\mathsf{lb},m}_{k,\min}}{2}=[0.b_{k,1}b_{k,2}\dots b_{k,m}1]_2$ and $t^{\mathsf{ub},m+1}_{k,\min}=t^{\mathsf{ub},m}_{k,\min}$. Since $t^{\mathsf{lb},m}_{k,\min}=[0.b_{k,1}b_{k,2}\dots b_{k,m}]_2$, $t^{\mathsf{lb},m+1}_{k,\min}=t^{\mathsf{lb},m}_{k,\min}+2^{-(m+1)}$. Then, $t^{\mathsf{ub},m+1}_{k,\min}=t^{\mathsf{ub},m}_{k,\min}=t^{\mathsf{lb},m}_{k,\min}+2^{-m}=t^{\mathsf{lb},m+1}_{k,\min}-2^{-(m+1)}+2^{-m}=t^{\mathsf{lb},m+1}_{k,\min}+2^{-(m+1)}$. Similarly, it can be shown that Lemma C.7 is valid if $ENC^{m+1}_{MR,\mathsf{TS},k}(\boldsymbol{h}_k)=0$. Therefore, Lemma C.7 stands when $l=m+1$. By induction, Lemma C.7 holds for any $l\geq 1$. □

A. Proof of Theorem 4.2.(a)

From the procedure of $CQ_{MR,\mathsf{TS}}$ in Fig. 4.3, $CQ_{MR,\mathsf{TS}}(\boldsymbol{H})=\left(\frac{1}{2},\frac{1}{2}\right)$ when Situation $\mathbb{A}$ happens. Thus, $\hat{t}_1=\hat{t}_2=\frac{1}{2}$ and $\hat{t}_1+\hat{t}_2=1$. When Situation $\mathbb{B}$, $\mathbb{C}$, or $\mathbb{D}$ happens at Round $\hat{l}$, it must have $\hat{l}\geq 1$ and $0<t_{k,\min}<1$ for $k=1,2$. Then, $\left(ENC^{l}_{MR,\mathsf{TS},1}(\boldsymbol{h}_1),ENC^{l}_{MR,\mathsf{TS},2}(\boldsymbol{h}_2)\right)=(0,1)$ or $(1,0)$ for $1\leq l\leq\hat{l}-1$. From Lemma C.7, we obtain (i) $(b_{1,l},b_{2,l})=(0,1)$ or $(1,0)$ for $1\leq l\leq\hat{l}-1$; (ii) $t^{\mathsf{lb},\hat{l}-1}_{k,\min}=\left[0.b_{k,1}b_{k,2}\dots b_{k,\hat{l}-1}\right]_2$ and $t^{\mathsf{ub},\hat{l}-1}_{k,\min}=t^{\mathsf{lb},\hat{l}-1}_{k,\min}+2^{-\left(\hat{l}-1\right)}$ for $k=1,2$. Therefore, $CQ_{MR,\mathsf{TS}}(\boldsymbol{H})=\left(\frac{t^{\mathsf{ub},\hat{l}-1}_{1,\min}+t^{\mathsf{lb},\hat{l}-1}_{1,\min}}{2},\frac{t^{\mathsf{ub},\hat{l}-1}_{2,\min}+t^{\mathsf{lb},\hat{l}-1}_{2,\min}}{2}\right)$ is derived as

$$CQ_{MR,\mathsf{TS}}(\boldsymbol{H})=\left(\underbrace{\left[0.b_{1,1}b_{1,2}\dots b_{1,\hat{l}-1}1\right]_2}_{=\hat{t}_1},\underbrace{\left[0.b_{2,1}b_{2,2}\dots b_{2,\hat{l}-1}1\right]_2}_{=\hat{t}_2}\right). \tag{C.76}$$

Since $b_{1,l}+b_{2,l}=1$ when $1\leq l\leq\hat{l}-1$, we obtain $0<\hat{t}_1,\hat{t}_2<1$ and $\hat{t}_1+\hat{t}_2=1$. This completes the proof of Theorem 4.2.(a).

B. Proof of Theorem 4.2.(b)

Let's prove $\mathsf{OUT}\left(CQ_{MR,\mathsf{TS}}\right) = \mathsf{OUT}^{\text{opt}}_{MR,\mathsf{TS}}$ first. Recall that $t_{k,\min} = \frac{1}{\log_2\left(1+P|h_{kk}|^2\right)}$, and let

$$\mathcal{H}_1 = \{\boldsymbol{H} : t_{1,\min} + t_{2,\min} > 1, t_{1,\min} > 0, t_{2,\min} > 0\},$$
$$\mathcal{H}_2 = \{\boldsymbol{H} : t_{1,\min} + t_{2,\min} = 1, t_{1,\min} > 0, t_{2,\min} > 0\},$$
$$\mathcal{H}_3 = \{\boldsymbol{H} : t_{1,\min} + t_{2,\min} < 1, t_{1,\min} > 0, t_{2,\min} > 0\}.$$

Since $t_{1,\min} + t_{2,\min} = \frac{1}{MR_{\mathsf{TS}}\left(t_1^\star, t_2^\star\right)}$, we have $\mathsf{OUT}^{\text{opt}}_{MR,\mathsf{TS}} = \Pr\{\boldsymbol{H} \in \mathcal{H}_1\}$. Let

$$\mathsf{OUT}_i \triangleq \Pr\left\{\boldsymbol{H} \in \mathcal{H}_i,\ MR_{\mathsf{TS}}\left(CQ_{MR,\mathsf{TS}}\left(\boldsymbol{H}\right)\right) < 1\right\}.$$

To prove $\mathsf{OUT}\left(CQ_{MR,\mathsf{TS}}\right) = \mathsf{OUT}^{\text{opt}}_{MR,\mathsf{TS}}$, it suffices to show that $\mathsf{OUT}_1 = \Pr\{\boldsymbol{H} \in \mathcal{H}_1\}$ and $\mathsf{OUT}_2 = \mathsf{OUT}_3 = 0$.

We first show $\mathsf{OUT}_1 = \Pr\{\boldsymbol{H} \in \mathcal{H}_1\}$. Obviously, $\mathsf{OUT}_1 \le \Pr\{\boldsymbol{H} \in \mathcal{H}_1\}$. On the other hand, for $\boldsymbol{H} \in \mathcal{H}_1$, $t_{1,\min} + t_{2,\min} > 1$, and $MR_{\mathsf{TS}}\left(CQ_{MR,\mathsf{TS}}\left(\boldsymbol{H}\right)\right) \le MR_{\mathsf{TS}}\left(t_1^\star, t_2^\star\right) = \frac{1}{t_{1,\min}+t_{2,\min}} < 1$. It follows that $\mathsf{OUT}_1 \ge \Pr\{\boldsymbol{H} \in \mathcal{H}_1\}$. Thus, $\mathsf{OUT}_1 = \Pr\{\boldsymbol{H} \in \mathcal{H}_1\}$.

For OUT_2, we obtain $\mathsf{OUT}_2 \le \Pr\{t_{1,\min} + t_{2,\min} = 1\} = 0$ as the probability of the continuous r.v. $t_{1,\min} + t_{2,\min} = \sum_{k=1}^{2} \frac{1}{\log_2\left(1+P|h_{kk}|^2\right)}$ assuming a specific value is zero. Since $\mathsf{OUT}_2 \ge 0$, $\mathsf{OUT}_2 = 0$.

To prove $\mathsf{OUT}_3 = 0$, it is sufficient to show for any $\boldsymbol{H} \in \mathcal{H}_3$, $MR_{\mathsf{TS}}\left(CQ_{MR,\mathsf{TS}}\left(\boldsymbol{H}\right)\right) \ge 1$. When $\boldsymbol{H} \in \mathcal{H}_3$, it implies $t_{1,\min} + t_{2,\min} < 1$, then, there must exist $\tilde{l} \in \mathtt{N}$ such that $t_{1,\min} + t_{2,\min} \le 1 - 2^{-\tilde{l}}$. Let $t_{k,\min} = [0.b_{k,1}b_{k,2}\ldots]_2$ for $k = 1, 2$. Then, $t_{1,\min} + t_{2,\min} \le 1 - 2^{-\tilde{l}}$ is equivalent to

$$\left[0.b_{1,1}b_{1,2}\ldots b_{1,\tilde{l}}\ldots\right]_2 + \left[0.b_{2,1}b_{2,2}\ldots b_{2,\tilde{l}}\ldots\right]_2 \le [0.\overbrace{11\ldots1}^{\tilde{l}}]_2. \tag{C.77}$$

Any $(t_{1,\min}, t_{2,\min})$ satisfying (C.77) will fall into one of the following two categories:

① $\exists 1 \leq l^{'} \leq \tilde{l}$ such that $\left(b_{1,l^{'}}, b_{2,l^{'}}\right) = (0,0)$ and $(b_{1,l}, b_{2,l}) = (0,1)$ or $(1,0)$ for $1 \leq l \leq l^{'}-1$.

② $(b_{1,l}, b_{2,l}) = (0,1)$ or $(1,0)$ for $1 \leq l \leq \tilde{l}$ and $(b_1, b_2) = (0,0)$ for $l \geq \tilde{l}+1$.

For example, $(t_{1,\min}, t_{2,\min}) = (0.5625, 0.25)$ belongs to ①, because $0.5625 = [0.1001]_2$ and $0.25 = [0.0100]_2$ with $\tilde{l} = l^{'} = 3$.

For any $(t_{1,\min}, t_{2,\min})$ in ①, according to Lemma C.7, we have $ENC^{l^{'}}_{MR,\mathsf{TS},1}(\boldsymbol{h}_1) = ENC^{l^{'}}_{MR,\mathsf{TS},2}(\boldsymbol{h}_2) = 0$ and $\left(ENC^{l}_{MR,\mathsf{TS},1}(\boldsymbol{h}_1), ENC^{l}_{MR,\mathsf{TS},2}(\boldsymbol{h}_2)\right) = (0,1)$ or $(1,0)$ for $1 \leq l \leq l^{'}-1$. Based on the procedure of $CQ_{MR,\mathsf{TS}}$, Situation $\mathbb{D}$ will happen at Round $l^{'}$. Similar to (C.76), the time-sharing pair will be determined as

$$CQ_{MR,\mathsf{TS}}(\boldsymbol{H}) = \left(\hat{t}_1, \hat{t}_2\right) = \left(\left[0.b_{1,1}b_{1,2}\ldots b_{1,l^{'}-1}1\right]_2, \left[0.b_{2,1}b_{2,2}\ldots b_{2,l^{'}-1}1\right]_2\right).$$

Since $t_{k,\min} = \left[0.b_{k,1}b_{k,2}\ldots b_{k,l^{'}-1}0\ldots\right]_2 < \left[0.b_{k,1}b_{k,2}\ldots b_{k,l^{'}-1}1\right]_2 = \hat{t}_k$, $\hat{t}_k \log_2\left(1+P\left|h_{kk}\right|^2\right) > t_{k,\min}\log_2\left(1+P\left|h_{kk}\right|^2\right) = 1$, then, $MR_{\mathsf{TS}}\left(\hat{t}_1, \hat{t}_2\right) \geq 1$.

Similarly, we can prove $MR_{\mathsf{TS}}\left(CQ_{MR,\mathsf{TS}}(\boldsymbol{H})\right) \geq 1$ also holds for ②. In summary, for any $\boldsymbol{H} \in \mathcal{H}_3$, $MR_{\mathsf{TS}}\left(CQ_{MR,\mathsf{TS}}(\boldsymbol{H})\right) \geq 1$. This implies $\mathsf{OUT}_3 = 0$, and thus concludes the proof of the claim $\mathsf{OUT}\left(CQ_{MR,\mathsf{TS}}\right) = \mathsf{OUT}^{\text{opt}}_{MR,\mathsf{TS}}$.

We now prove the upper bound on $\mathsf{FR}\left(CQ_{MR,\mathsf{TS}}\right)$ in (4.7). For $l \in \mathtt{N}$, let

$$\mathcal{R}_l = \left\{\boldsymbol{H} : CQ_{MR,\mathsf{TS}} \text{ ends after Round } l\right\}.$$

Based on the procedure of $CQ_{MR,\mathsf{TS}}$ in Fig. 4.3, a geometric representation of $\mathcal{R}_l$ is shown in Fig. C.1 for clarity of exposition. Since the number of feedback bits is $2(l+1)$ after Round l, the average feedback rate of $CQ_{MR,\mathsf{TS}}$ is calculated as

$$\mathsf{FR}\left(CQ_{MR,\mathsf{TS}}\right) = \sum_{l=0}^{\infty} 2(l+1) \times \Pr\left\{\boldsymbol{H} \in \mathcal{R}_l\right\}$$

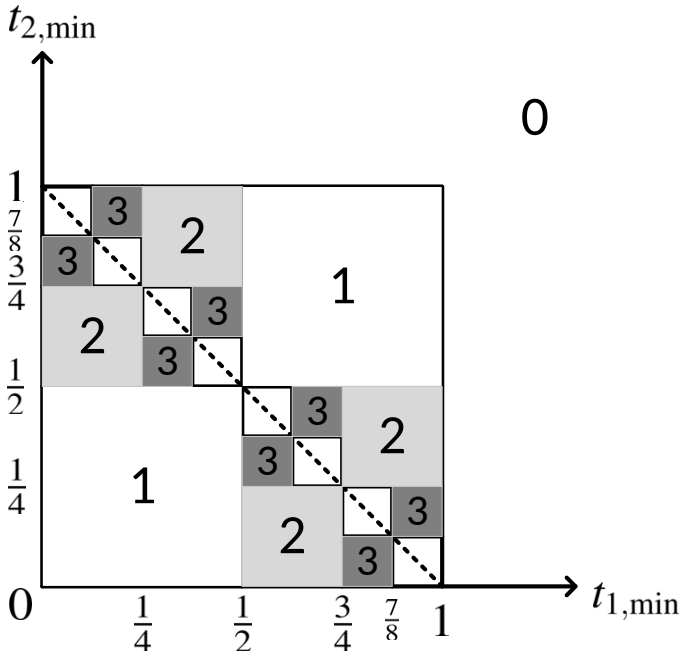


Figure C.1: A geometric representation of the regions $\mathcal{R}_l$ with l being the number in the corresponding area. For example, $\mathcal{R}_1$ is the union of the two "squares" on the figure that are marked with a "1," and $\mathcal{R}_0$ consists of the points that do not belong to the unit square. Only the four regions $\mathcal{R}_0, \ldots, \mathcal{R}_3$ are shown explicitly.

$$
\begin{aligned}
&= 2 + 2 \times \Pr\{\boldsymbol{H} \in \mathcal{R}_1\} + 2 \sum_{l=2}^{\infty} l \times \Pr\{\boldsymbol{H} \in \mathcal{R}_l\} \\
&\leq 2 + 2 \times \Pr\{\boldsymbol{H} \in \mathcal{R}_1\} + 2 \sum_{l=2}^{\infty} l \times \Pr\left\{\boldsymbol{H} \in \bigcup_{w=l}^{\infty} \mathcal{R}_w\right\}, \qquad \text{(C.78)}
\end{aligned}
$$

where the second equality is from $\sum_{l=0}^{\infty} \Pr\{\boldsymbol{H} \in \mathcal{R}_l\} = 1$; the last inequality is due to the fact that $\mathcal{R}_l \subseteq \cup_{w=l}^{\infty} \mathcal{R}_w$. Now, since

$$
\mathcal{R}_1 \subseteq \left\{\boldsymbol{H} : 0 \leq t_{1,\min}, t_{2,\min} \leq \frac{1}{2}\right\} \cup \left\{\boldsymbol{H} : \frac{1}{2} \leq t_{1,\min}, t_{2,\min} \leq 1\right\},
$$

we have

$$
\Pr\{\boldsymbol{H} \in \mathcal{R}_1\} \leq \int_0^{\frac{1}{2}} f_{t_{1,\min}}(x_1)\mathrm{d}x_1 \int_0^{\frac{1}{2}} f_{t_{2,\min}}(x_2)\mathrm{d}x_2 + \int_{\frac{1}{2}}^{1} f_{t_{1,\min}}(x_1)\mathrm{d}x_1 \int_{\frac{1}{2}}^{1} f_{t_{2,\min}}(x_2)\mathrm{d}x_2
$$

$$\leq \int_0^{\frac{1}{2}} f_{t_{1,\min}}(x_1)\mathrm{d}x_1 + \int_{\frac{1}{2}}^1 f_{t_{1,\min}}(x_1)\mathrm{d}x_1 = \int_0^1 f_{t_{1,\min}}(x_1)\mathrm{d}x_1 = e^{-\frac{1}{P}}, \tag{C.79}$$

where the last equality is from the explicit form $f_{t_{k,\min}}(x) = \frac{\log 2}{Px^2} e^{-\frac{e^{\frac{\log 2}{x}}-1}{P}} e^{\frac{\log 2}{x}}$, $x > 0$ of the pdf of $t_{k,\min}$. According to the procedure of $CQ_{MR,\mathsf{TS}}$ in Fig. 4.3, when $l \geq 2$, we obtain

$$\mathcal{R}_l \subseteq \bigcup_{q=0}^{2^{l-1}-1} \left\{ \mathcal{R}_{l,q}^{(1)} \cup \mathcal{R}_{l,q}^{(2)} \right\}, \tag{C.80}$$

where

$$\mathcal{R}_{l,q}^{(1)} = \left\{ \begin{array}{ll} & \frac{2q}{2^l} \leq t_{1,\min} \leq \frac{2q+1}{2^l}, \\ \boldsymbol{H}: & 1 - \frac{2q+2}{2^l} \leq t_{2,\min} \leq 1 - \frac{2q+1}{2^l}, \\ & 0 < t_{1,\min}, t_{2,\min} < 1 \end{array} \right\},$$

$$\mathcal{R}_{l,q}^{(2)} = \left\{ \begin{array}{ll} & \frac{2q+1}{2^l} \leq t_{1,\min} \leq \frac{2q+2}{2^l}, \\ \boldsymbol{H}: & 1 - \frac{2q+1}{2^l} \leq t_{2,\min} \leq 1 - \frac{2q}{2^l}, \\ & 0 < t_{1,\min}, t_{2,\min} < 1 \end{array} \right\}. \tag{C.81}$$

It follows that

$$\bigcup_{w=l}^{\infty} \mathcal{R}_w \subseteq \left\{ \begin{array}{l} \bigcup_{u=0}^{2^{l-2}-1} \left\{ \begin{array}{ll} \boldsymbol{H}: & \frac{1}{2} - \frac{u+1}{2^{l-1}} \leq t_{1,\min} \leq \frac{1}{2} - \frac{u}{2^{l-1}}, \\ & \frac{1}{2} + \frac{u}{2^{l-1}} \leq t_{2,\min} \leq \frac{1}{2} + \frac{u+1}{2^{l-1}} \end{array} \right\} \\ \cup \bigcup_{u=0}^{2^{l-2}-1} \left\{ \begin{array}{ll} \boldsymbol{H}: & \frac{1}{2} + \frac{u}{2^{l-1}} \leq t_{1,\min} \leq \frac{1}{2} + \frac{u+1}{2^{l-1}}, \\ & \frac{1}{2} - \frac{u+1}{2^{l-1}} \leq t_{2,\min} \leq \frac{1}{2} - \frac{u}{2^{l-1}} \end{array} \right\} \end{array} \right\}. \tag{C.82}$$

Thus, for $l \geq 2$, an upper bound on $\Pr\{\boldsymbol{H} \in \cup_{w=l}^{\infty} \mathcal{R}_w\}$ is deduced as

$$\Pr\left\{ \boldsymbol{H} \in \bigcup_{w=l}^{\infty} \mathcal{R}_w \right\} \leq \sum_{u=0}^{2^{l-2}-1} \int_{\frac{1}{2}-\frac{u+1}{2^{l-1}}}^{\frac{1}{2}-\frac{u}{2^{l-1}}} f_{t_{1,\min}}(x_1)\mathrm{d}x_1 \int_{\frac{1}{2}+\frac{u}{2^{l-1}}}^{\frac{1}{2}+\frac{u+1}{2^{l-1}}} f_{t_{2,\min}}(x_2)\mathrm{d}x_2$$

$$+\sum_{u=0}^{2^{l-2}-1}\int_{\frac{1}{2}+\frac{u}{2^{l-1}}}^{\frac{1}{2}+\frac{u+1}{2^{l-1}}} f_{t_{1,\min}}(x_1)\mathrm{d}x_1\int_{\frac{1}{2}-\frac{u+1}{2^{l-1}}}^{\frac{1}{2}-\frac{u}{2^{l-1}}} f_{t_{2,\min}}(x_2)\mathrm{d}x_2$$
$$=2\sum_{u=0}^{2^{l-2}-1}\int_{\frac{1}{2}-\frac{u}{2^{l-1}}}^{\frac{1}{2}-\frac{u}{2^{l-1}}} f_{t_{1,\min}}(x_1)\mathrm{d}x_1\int_{\frac{1}{2}-\frac{u+1}{2^{l-1}}}^{\frac{1}{2}+\frac{u+1}{2^{l-1}}} f_{t_{2,\min}}(x_2)\mathrm{d}x_2.$$

When $\frac{1}{2}+\frac{u}{2^{l-1}}\le x_2\le\frac{1}{2}+\frac{u+1}{2^{l-1}}$, since $\frac{1}{2}\le x_2\le 1$, we have $f_{t_{2,\min}}(x_2)=\frac{\log 2}{Px_2^2}e^{-\frac{e^{\frac{\log 2}{x_2}}-1}{P}}e^{\frac{\log 2}{x_2}}\le \frac{4\log 2}{P}e^{-\frac{1}{P}}e^{2\log 2}=C_2\frac{e^{-\frac{1}{P}}}{P}$, where $C_2=4\log 2e^{2\log 2}$. Hence,

$$\begin{aligned}\Pr\left\{\boldsymbol{H}\in\bigcup_{w=l}^{\infty}\mathcal{R}_w\right\}&\le 2\sum_{u=0}^{2^{l-2}-1}\int_{\frac{1}{2}-\frac{u}{2^{l-1}}}^{\frac{1}{2}-\frac{u}{2^{l-1}}} f_{t_{1,\min}}(x_1)\mathrm{d}x_1\int_{\frac{1}{2}+\frac{u}{2^{l-1}}}^{\frac{1}{2}+\frac{u+1}{2^{l-1}}}\left(C_2\frac{e^{-\frac{1}{P}}}{P}\right)\mathrm{d}x_2\\&=2C_2\frac{e^{-\frac{1}{P}}}{P}\times\frac{1}{2^{l-1}}\sum_{u=0}^{2^{l-2}-1}\int_{\frac{1}{2}-\frac{u+1}{2^{l-1}}}^{\frac{1}{2}-\frac{u}{2^{l-1}}} f_{t_{1,\min}}(x_1)\mathrm{d}x_1\\&=C_2\frac{e^{-\frac{1}{P}}}{P}\times\frac{1}{2^{l-2}}\int_0^{\frac{1}{2}} f_{t_{1,\min}}(x_1)\mathrm{d}x_1\le C_2\frac{e^{-\frac{1}{P}}}{P}\times\frac{1}{2^{l-2}}.\end{aligned}\tag{C.83}$$

Subsituting (C.79) and (C.83) into (C.78) and using the fact that $\sum_{l=2}^{\infty}\frac{l}{2^{l-2}}$ is finite yields (4.7). This concludes the proof of Theorem 4.2. □

C.3 Proof of Theorem 4.3

Proof. To prove Theorem 4.3, we need the following lemma.

Lemma C.8. *In Stage* 0 *of* $CQ_{MR,\mathsf{CT}}$ *where* $p_1=1$, *when* $0<1-p_{2,\max}<1$ *and* $0<p_{2,\min}<1$, *let* $1-p_{2,\max}=[0.b_{1,1}b_{1,2}\ldots]_2$ *and* $p_{2,\min}=[0.b_{2,1}b_{2,2}\ldots]_2$. *Then, at Round* l *in Stage* 0 *for* $l\ge 2$, *we have* $ENC^{0,l}_{MR,\mathsf{CT},1}(\boldsymbol{h}_1)=b_{1,l-1}$, $ENC^{0,l}_{MR,\mathsf{CT},2}(\boldsymbol{h}_2)=b_{2,l-1}$, $q_1^{\mathsf{lb},0,l}=[0.b_{1,1}b_{1,2}\ldots b_{1,l-1}]_2$, $q_1^{\mathsf{ub},0,l}=q_1^{\mathsf{lb},0,l}+2^{-(l-1)}$, $q_2^{\mathsf{lb},0,l}=[0.b_{2,1}b_{2,2}\ldots b_{2,l-1}]_2$ *and* $q_2^{\mathsf{ub},0,l}=q_2^{\mathsf{lb},0,l}+2^{-(l-1)}$.

The proof of Lemma C.8 is the same as that of Lemma C.7 in Appendix C.2. Similar results holds when $0<1-p_{1,\max}<1$ and $0<p_{1,\min}<1$ in Stage 1 where $p_2=1$.

A. Proof of Theorem 4.3.(a)

From the procedure of $CQ_{MR,\mathsf{CT}}$ in Fig. 4.4, we have $CQ_{MR,\mathsf{CT}}(\boldsymbol{H}) = (1,1)$ under Situation $\mathbb{A}$ in Stage 0. Thus, $\hat{p}_1 = \hat{p}_2 = 1$ satisfies $0 \leq \hat{p}_1, \hat{p}_2 \leq 1$.

When Situation $\mathbb{B}$, $\mathbb{C}$, or $\mathbb{D}$ happens at Round $\hat{l}$ in Stage 0, it must have $\hat{l} \geq 2$ and $0 < 1 - p_{2,\max} < 1, 0 < p_{2,\min} \leq 1$. When $0 < 1 - p_{2,\max} < 1$ and $0 < p_{2,\min} < 1$, based on Lemma C.8 and following the same steps in the proof of Theorem 4.2.(a), we will obtain $CQ_{MR,\mathsf{CT}}(\boldsymbol{H}) = (\hat{p}_1, \hat{p}_2) = \left(1, \left[0.b_{2,1}b_{2,2}\ldots b_{2,\hat{l}-2}1\right]_2\right)$, thus, $0 \leq \hat{p}_1, \hat{p}_2 \leq 1$. Trivially, when $0 < 1 - p_{2,\max} < 1$ and $p_{2,\min} = 1$, we can show that $0 \leq q_2^{\mathsf{ub},0,l}, q_2^{\mathsf{lb},0,l} \leq 1$ for $l \geq 1$. Since $CQ_{MR,\mathsf{CT}}(\boldsymbol{H}) = (\hat{p}_1, \hat{p}_2) = \left(1, \frac{q_2^{\mathsf{ub},0,l-1} + q_2^{\mathsf{lb},0,l-1}}{2}\right)$, thus, $0 \leq \hat{p}_1, \hat{p}_2 \leq 1$. The same result can be obtained for Stage 1 where $p_2 = 1$. Thus, the proof of Theorem 4.3.(a) is complete.

B. Proof of Theorem 4.3.(b)

We first prove $\mathsf{OUT}\left(CQ_{MR,\mathsf{CT}}\right) = \mathsf{OUT}^{\text{opt}}_{MR,\mathsf{CT}}$. It is obvious that $\mathsf{OUT}\left(CQ_{MR,\mathsf{CT}}\right) \geq \mathsf{OUT}^{\text{opt}}_{MR,\mathsf{CT}}$, thus, it is sufficient to show $\mathsf{OUT}\left(CQ_{MR,\mathsf{CT}}\right) \leq \mathsf{OUT}^{\text{opt}}_{MR,\mathsf{CT}}$. Since Proposition 4.2 shows $p_1^\star = 1$ or $p_2^\star = 1$, $\mathsf{OUT}^{\text{opt}}_{MR,\mathsf{CT}} = \Pr\left\{MR_{\mathsf{CT}}(p_1^\star, p_2^\star) < 1\right\}$ can be alternatively expressed as $\mathsf{OUT}^{\text{opt}}_{MR,\mathsf{CT}} = \Pr\left\{\boldsymbol{H} \in \widetilde{\mathcal{H}_1}\right\}$, where

$$\begin{aligned}\widetilde{\mathcal{H}_1} = &\left\{\boldsymbol{H} : MR_{\mathsf{CT}}(p_1 = 1, p_2) < 1 \text{ for } 0 \leq p_2 \leq 1\right\} \\ &\cap \left\{\boldsymbol{H} : MR_{\mathsf{CT}}(p_1, p_2 = 1) < 1 \text{ for } 0 \leq p_1 \leq 1\right\}.\end{aligned}$$

After mathematical calculations, $\widetilde{\mathcal{H}_1}$ is equivalent to

$$\widetilde{\mathcal{H}_1} = \left\{ \boldsymbol{H} : \begin{array}{c} \{p_{2,\max} \leq 0 \text{ or } p_{2,\min} > 1\} \\ \text{or } \{p_{2,\min} \leq 1 \text{ and } p_{2,\min} > p_{2,\max}\} \end{array} \right\} \cap \left\{ \boldsymbol{H} : \begin{array}{c} \{p_{1,\max} \leq 0 \text{ or } p_{1,\min} > 1\} \\ \text{or } \{p_{1,\min} \leq 1 \text{ and } p_{1,\min} > p_{1,\max}\} \end{array} \right\}.$$

Let

$$\widetilde{\mathcal{H}_2} = \bigcup_{k=1}^{2} \left\{ \boldsymbol{H} : \begin{array}{c} 0 < p_{k,\max} < 1, 0 < p_{k,\min} < 1, \\ (1 - p_{k,\max}) + p_{k,\min} = 1 \end{array} \right\} \cup \bigcup_{k=1}^{2} \{\boldsymbol{H} : p_{k,\max} \geq 1, p_{k,\min} = 1\},$$

$$\widetilde{\mathcal{H}_3} = \bigcup_{k=1}^{2} \left\{ \boldsymbol{H} : \begin{array}{c} 0 < p_{k,\max} < 1, 0 < p_{k,\min} < 1, \\ (1 - p_{k,\max}) + p_{k,\min} < 1 \end{array} \right\} \cup \bigcup_{k=1}^{2} \{\boldsymbol{H} : p_{k,\max} \geq 1, 0 < p_{k,\min} < 1\}.$$

It can be verified that $\bigcup_{i=1}^{3} \widetilde{\mathcal{H}_i} = \mathbb{C}^{2\times 2}$ and $\widetilde{\mathcal{H}_1} \cap \left\{ \widetilde{\mathcal{H}_2} \cup \widetilde{\mathcal{H}_3} \right\} = \emptyset$. Thus, $\mathsf{OUT}\left(CQ_{MR,\mathsf{CT}}\right) \leq \sum_{i=1}^{3} \widetilde{\mathsf{OUT}_i}$, where

$$\widetilde{\mathsf{OUT}_i} \triangleq \Pr\left\{ \boldsymbol{H} \in \widetilde{\mathcal{H}_i}, MR_{\mathsf{CT}}\left(CQ_{MR,\mathsf{CT}}\left(\boldsymbol{H}\right)\right) < 1 \right\}.$$

To prove $\mathsf{OUT}\left(CQ_{MR,\mathsf{CT}}\right) \leq \mathsf{OUT}^{\text{opt}}_{MR,\mathsf{CT}}$, it is sufficient to show $\widetilde{\mathsf{OUT}_1} = \Pr\left\{ \boldsymbol{H} \in \widetilde{\mathcal{H}_1} \right\}$ and $\widetilde{\mathsf{OUT}_2} = \widetilde{\mathsf{OUT}_3} = 0$.

The proofs of $\widetilde{\mathsf{OUT}_1} = \Pr\left\{ \boldsymbol{H} \in \widetilde{\mathcal{H}_1} \right\}$ and $\widetilde{\mathsf{OUT}_2} = 0$ are similar to those of $\mathsf{OUT}_1 = \Pr\{\boldsymbol{H} \in \mathcal{H}_1\}$ and $\mathsf{OUT}_2 = 0$ in Appendix C.2, thus omitted.

To prove $\widetilde{\mathsf{OUT}_3} = 0$, it is sufficient to prove that for any $\boldsymbol{H} \in \widetilde{\mathcal{H}_3}$, $MR_{\mathsf{CT}}\left(CQ_{MR,\mathsf{CT}}\left(\boldsymbol{H}\right)\right) \geq 1$. When $\boldsymbol{H} \in \left\{\boldsymbol{H} : p_{k,\max} \geq 1, 0 < p_{k,\min} < 1\right\}$ for $k = 1$ or 2, the power pair is chosen to be $CQ_{MR,\mathsf{CT}}\left(\boldsymbol{H}\right) = (1,1)$. Since $p_{k,\max} \geq 1 \geq p_{k,\min}$, we have $MR_{\mathsf{CT}}\left(1,1\right) \geq 1$. When $\boldsymbol{H} \in \left\{\boldsymbol{H} : 0 < p_{2,\max} < 1, 0 < p_{2,\min} < 1, (1-p_{2,\max}) + p_{2,\min} < 1\right\}$, similar to the proof of $\mathsf{OUT}_3 = 0$ in Appendix C.2, we obtain: (i) Situation $\mathbb{D}$ in Stage 0 happens at Round $\bar{l} \geq 2$ with the power pair chosen as $CQ_{MR,\mathsf{CT}}\left(\boldsymbol{H}\right) = \left(1, \frac{q_2^{\mathsf{ub},0,\bar{l}-1} + q_2^{\mathsf{lb},0,\bar{l}-1}}{2}\right)$; (ii) $\frac{q_1^{\mathsf{ub},0,\bar{l}-1} + q_1^{\mathsf{lb},0,\bar{l}-1}}{2} = 1 - \frac{q_2^{\mathsf{ub},0,\bar{l}-1} + q_2^{\mathsf{lb},0,\bar{l}-1}}{2} > 1 - p_{2,\max}$ and $\frac{q_2^{\mathsf{ub},0,\bar{l}-1} + q_2^{\mathsf{lb},0,\bar{l}-1}}{2} > p_{2,\min}$. Then, $p_{2,\min} < \frac{q_2^{\mathsf{ub},0,\bar{l}-1} + q_2^{\mathsf{lb},0,\bar{l}-1}}{2} < p_{2,\max} < 1$, and thus $MR_{\mathsf{CT}}\left(1, \frac{q_2^{\mathsf{ub},0,\bar{l}-1} + q_2^{\mathsf{lb},0,\bar{l}-1}}{2}\right) \geq 1$. Similarly, we prove $MR_{\mathsf{CT}}\left(CQ_{MR,\mathsf{CT}}\left(\boldsymbol{H}\right)\right) \geq 1$ when $\boldsymbol{H} \in \left\{\boldsymbol{H} : 0 < p_{1,\max} < 1, 0 < p_{1,\min} < 1, (1-p_{1,\max}) + p_{1,\min} < 1\right\}$. Hence, we obtain $MR_{\mathsf{CT}}\left(CQ_{MR,\mathsf{CT}}\left(\boldsymbol{H}\right)\right) \geq 1$ for any $\boldsymbol{H} \in \widetilde{\mathcal{H}_3}$, implying $\widetilde{\mathsf{OUT}_3} = 0$. The proof of $\mathsf{OUT}\left(CQ_{MR,\mathsf{CT}}\right) = \mathsf{OUT}^{\mathrm{opt}}_{MR,\mathsf{CT}}$ is complete.

We now prove the upper bound on $\mathsf{FR}\left(CQ_{MR,\mathsf{CT}}\right)$ in (4.12). For $l_1, l_2 \in \mathbb{N}$, let

$$\widetilde{\mathcal{R}_{0,l_1}} = \left\{\boldsymbol{H} : CQ_{MR,\mathsf{CT}} \text{ ends after Round } l_1 \text{ in Stage } 0\right\},$$
$$\widetilde{\mathcal{R}_{1,l_1,l_2}} = \left\{\boldsymbol{H} : \begin{array}{c} \text{after } l_1 \text{ rounds in Stage 0, } CQ_{MR,\mathsf{CT}} \\ \text{ends after Round } l_2 \text{ in Stage 1} \end{array}\right\}.$$

There will be no more than $2 \times (l_1 + 1)$ feedback bits for $\boldsymbol{H} \in \widetilde{\mathcal{R}_{0,l_1}}$ and no more than $2(l_1+1) + 2(l_2+1)$ bits for $\boldsymbol{H} \in \widetilde{\mathcal{R}_{1,l_1,l_2}}$, thus, an upper bound on $\mathsf{FR}\left(CQ_{MR,\mathsf{CT}}\right)$ can be

$$\begin{aligned}
&\mathsf{FR}\left(CQ_{MR,\mathsf{CT}}\right) \\
&\leq \sum_{l_1=0}^{\infty} 2(l_1+1) \times \Pr\left\{\boldsymbol{H} \in \widetilde{\mathcal{R}_{0,l_1}}\right\} + \sum_{l_1=0}^{\infty}\sum_{l_2=0}^{\infty} [2(l_1+1) + 2(l_2+1)] \times \Pr\left\{\boldsymbol{H} \in \widetilde{\mathcal{R}_{1,l_1,l_2}}\right\} \\
&= 2\sum_{l_1=0}^{\infty} (l_1+1) \times P_{0,l_1} + 2\sum_{l_2=0}^{\infty} (l_2+1) \times P_{1,l_2}, \qquad \text{(C.84)}
\end{aligned}$$

where $P_{0,l_1} = \Pr\left\{\boldsymbol{H} \in \widetilde{\mathcal{R}_{0,l_1}}\right\} + \sum_{l_2=0}^{\infty} \Pr\left\{\boldsymbol{H} \in \widetilde{\mathcal{R}_{1,l_1,l_2}}\right\}$ and $P_{1,l_2} = \sum_{l_1=0}^{\infty} \Pr\left\{\boldsymbol{H} \in \widetilde{\mathcal{R}_{1,l_1,l_2}}\right\}$. Since $\widetilde{\mathcal{R}_{0,l_1}} \cap \widetilde{\mathcal{R}_{1,l_1,l_2}} = \emptyset$ and $\widetilde{\mathcal{R}_{1,l_1,l_2}} \cap \widetilde{\mathcal{R}_{1,l_1,l_2'}} = \emptyset$ for any $l_1, l_2, l_2' \in \mathbb{N}$ and $l_2 \neq l_2'$, $P_{0,l_1} =$

$\Pr\left\{\boldsymbol{H} \in \widehat{\mathcal{R}_{0,l_1}}\right\}$, where $\widehat{\mathcal{R}_{0,l_1}} = \widetilde{\mathcal{R}_{0,l_1}} \cup \bigcup_{l_2=0}^{\infty} \widetilde{\mathcal{R}_{1,l_1,l_2}}$. Intuitively, $\widehat{\mathcal{R}_{0,l_1}}$ is the set of channel states for which after Round l_1 in Stage 0, $CQ_{MR,\mathsf{CT}}$ will not continue into Round l_1+1 in Stage 0 (either end or continue into Stage 1).

When $l_1 = 0$, from the procedure of $CQ_{MR,\mathsf{CT}}$ in Fig. 4.4, we obtain $\widehat{\mathcal{R}_{0,0}} = \{\boldsymbol{H} : p_{2,\max} \leq 0\} \cup \{\boldsymbol{H} : p_{2,\min} > 1\}$, then,

$$\begin{aligned} P_{0,0} &= \Pr\{p_{2,\max} \leq 0\} + \Pr\{p_{2,\min} > 1\} - \Pr\{p_{2,\max} \leq 0\} \Pr\{p_{2,\min} > 1\} \\ &= \Pr\left\{|h_{11}|^2 \leq \frac{1}{P}\right\} + \Pr\left\{|h_{22}|^2 < |h_{12}|^2 + \frac{1}{P}\right\} \\ &\quad - \Pr\left\{|h_{11}|^2 \leq \frac{1}{P}\right\} \Pr\left\{|h_{22}|^2 < |h_{12}|^2 + \frac{1}{P}\right\}. \end{aligned}$$

After simple mathematical calculations, $P_{0,0}$ is derived as

$$P_{0,0} = 1 - \frac{e^{-\frac{2}{P}}}{1+\eta}. \tag{C.85}$$

When $l_1 = 1$ and 2, we obtain

$$\begin{aligned} &\widehat{\mathcal{R}_{0,1}} = \{\boldsymbol{H} : p_{2,\max} \geq 1, p_{2,\min} \leq 1\}, \\ &\widehat{\mathcal{R}_{0,2}} \subseteq \{\boldsymbol{H} : 0 < p_{2,\max} < 1, p_{2,\min} \leq 1\}. \end{aligned}$$

Then,

$$\begin{aligned} &P_{0,1} = \Pr\{p_{2,\max} \geq 1\} \Pr\{p_{2,\min} \leq 1\} = \frac{e^{-\frac{2}{P}}}{(1+\eta)^2}, \\ &P_{0,2} \leq \Pr\{0 < p_{2,\max} < 1\} \Pr\{p_{2,\min} \leq 1\} = \frac{\eta e^{-\frac{2}{P}}}{(1+\eta)^2}. \end{aligned} \tag{C.86}$$

When $l_1 \geq 3$, based on Fig. 4.4, $\widehat{\mathcal{R}_{0,l_1}}$ is composed of channel states for which $ENC^{0,l_1}_{MR,\mathsf{CT},1}(\boldsymbol{h}_1) = ENC^{0,l_1}_{MR,\mathsf{CT},2}(\boldsymbol{h}_2) = 0$ or 1, and $\left(ENC^{0,l}_{MR,\mathsf{CT},1}(\boldsymbol{h}_1), ENC^{0,l}_{MR,\mathsf{CT},2}(\boldsymbol{h}_2)\right) = (0,1)$ or $(1,0)$ for

$2 \le l \le l_1 - 1$. The geometric representation for $\widehat{\mathcal{R}_{0,l_1}}$ is similar to that of $\mathcal{R}_l$ in Fig. C.1. Similar to (C.80), we obtain $\widehat{\mathcal{R}_{0,l_1}} \subseteq \widehat{\mathcal{R}_{0,l_1}^{(0)}} \cup \widehat{\mathcal{R}_{0,l_1}^{(1)}}$, where $\widehat{\mathcal{R}_{0,l_1}^{(0)}} = \{\boldsymbol{H} : p_{2,\min} = 1\}$, $\widehat{\mathcal{R}_{0,l_1}^{(1)}} = \bigcup_{q=0}^{2^{l_1-2}-1} \left\{ \widehat{\mathcal{R}_{0,l_1,q}^{(1,1)}} \cup \widehat{\mathcal{R}_{0,l_1,q}^{(1,2)}} \right\}$ and

$$\widehat{\mathcal{R}_{0,l_1,q}^{(1,1)}} = \left\{ \boldsymbol{H} : \begin{array}{l} \frac{2q}{2^{l_1-1}} \le 1 - p_{2,\max} \le \frac{2q+1}{2^{l_1-1}}, \\ 1 - \frac{2q+2}{2^{l_1-1}} \le p_{2,\min} \le 1 - \frac{2q+1}{2^{l_1-1}}, \\ 0 < 1 - p_{2,\max}, p_{2,\min} < 1 \end{array} \right\}, \tag{C.87}$$

$$\widehat{\mathcal{R}_{0,l_1,q}^{(1,2)}} = \left\{ \boldsymbol{H} : \begin{array}{l} \frac{2q+1}{2^{l_1-1}} \le 1 - p_{2,\max} \le \frac{2q+2}{2^{l_1-1}}, \\ 1 - \frac{2q+1}{2^{l-1-1}} \le p_{2,\min} \le 1 - \frac{2q}{2^{l_1-1}}, \\ 0 < 1 - p_{2,\max}, p_{2,\min} < 1 \end{array} \right\}.$$

Here, $\widehat{\mathcal{R}_{0,l_1}^{(0)}}$ embraces the zero-probabilistic cases where $0 < 1 - p_{2,\max} < 1$ and $p_{2,\min} = 1$ after Round 1; $\widehat{\mathcal{R}_{0,l_1}^{(1)}}$ is the set of channel states for which conferencing ends after Round l_1 with $0 < 1 - p_{2,\max} < 1$ and $0 < p_{2,\min} < 1$. Similar to (C.82), it follows that

$$\widehat{\mathcal{R}_{0,l_1}^{(1)}} \subseteq \bigcup_{w=l_1}^{\infty} \widehat{\mathcal{R}_{0,l_1}^{(1)}} \subseteq \left\{ \begin{array}{l} \bigcup_{u=0}^{2^{l_1-3}-1} \left\{ \boldsymbol{H} : \begin{array}{l} \frac{1}{2} - \frac{u+1}{2^{l_1-2}} \le 1 - p_{2,\max} \le \frac{1}{2} - \frac{u}{2^{l_1-2}}, \\ \frac{1}{2} + \frac{u}{2^{l_1-2}} \le p_{2,\min} \le \frac{1}{2} + \frac{u+1}{2^{l_1-2}} \end{array} \right\} \\ \cup \bigcup_{u=0}^{2^{l_1-3}-1} \left\{ \boldsymbol{H} : \begin{array}{l} \frac{1}{2} + \frac{u}{2^{l_1-2}} \le 1 - p_{2,\max} \le \frac{1}{2} + \frac{u+1}{2^{l_1-2}}, \\ \frac{1}{2} - \frac{u+1}{2^{l_1-2}} \le p_{2,\min} \le \frac{1}{2} - \frac{u}{2^{l_1-2}} \end{array} \right\} \end{array} \right\}.$$

Then, an upper bound on P_{0,l_1} for $l_1 \ge 3$ can be

$$\begin{aligned} P_{0,l_1} &\le \Pr\left\{ \boldsymbol{H} \in \widehat{\mathcal{R}_{0,l_1}^{(0)}} \right\} + \Pr\left\{ \boldsymbol{H} \in \widehat{\mathcal{R}_{0,l_1}^{(1)}} \right\} \\ &= \Pr\left\{ \boldsymbol{H} \in \widehat{\mathcal{R}_{0,l_1}^{(1)}} \right\} \le \Pr\left\{ \boldsymbol{H} \in \cup_{w=l_1}^{\infty} \widehat{\mathcal{R}_{0,l_1}^{(1)}} \right\} \\ &\le \sum_{u=0}^{2^{l_1-3}-1} \Pr\left\{ \frac{1}{2} - \frac{u+1}{2^{l_1-2}} \le 1 - p_{2,\max} \le \frac{1}{2} - \frac{u}{2^{l_1-2}} \right\} \\ &\qquad \times \Pr\left\{ \frac{1}{2} + \frac{u}{2^{l_1-2}} \le p_{2,\min} \le \frac{1}{2} + \frac{u+1}{2^{l_1-2}} \right\} \end{aligned}$$

$$+\sum_{u=0}^{2^{l_1-3}-1}\Pr\left\{\frac{1}{2}+\frac{u}{2^{l_1-2}}\le 1-p_{2,\max}\le\frac{1}{2}+\frac{u+1}{2^{l_1-2}}\right\}$$
$$\times\Pr\left\{\frac{1}{2}-\frac{u+1}{2^{l_1-2}}\le p_{2,\min}\le\frac{1}{2}-\frac{u}{2^{l_1-2}}\right\}. \tag{C.88}$$

For any a satisfying $0\le a\le 1$ and $0\le a+\frac{1}{2^{l_1-2}}\le 1$, $\Pr\left\{a\le 1-p_{2,\max}\le a+\frac{1}{2^{l_1-2}}\right\}$ can be deduced as

$$\begin{aligned}
&\Pr\left\{a\le 1-p_{2,\max}\le a+\frac{1}{2^{l_1-2}}\right\}\\
&=\Pr\left\{1-a-\frac{1}{2^{l_1-2}}\le p_{2,\max}=\frac{|h_{11}|^2-\frac{1}{P}}{|h_{21}|^2}\le 1-a\right\}\\
&=\int_0^\infty\frac{e^{-\frac{|h_{21}|^2}{\eta}}}{\eta}\mathrm{d}\,|h_{21}|^2\int_{\left(1-a-\frac{1}{2^{l_1-2}}\right)|h_{21}|^2+\frac{1}{P}}^{(1-a)|h_{21}|^2+\frac{1}{P}}e^{-|h_{11}|^2}\mathrm{d}\,|h_{11}|^2\\
&\le\int_0^\infty\frac{e^{-\frac{|h_{21}|^2}{\eta}}}{\eta}\mathrm{d}\,|h_{21}|^2\,e^{-\left(1-a-\frac{1}{2^{l_1-2}}\right)|h_{21}|^2-\frac{1}{P}}\times\frac{|h_{21}|^2}{2^{l_1-2}}\\
&\le\frac{e^{-\frac{1}{P}}}{2^{l_1-2}}\int_0^\infty|h_{21}|^2\frac{e^{-\frac{|h_{21}|^2}{\eta}}}{\eta}\mathrm{d}\,|h_{21}|^2=\frac{\eta e^{-\frac{1}{P}}}{2^{l_1-2}}.
\end{aligned}\tag{C.89}$$

Letting $a=\frac{1}{2}-\frac{u+1}{2^{l_1-2}}$ or $\frac{1}{2}+\frac{u}{2^{l_1-2}}$ for $0\le u\le 2^{l_1-3}-1$ in (C.88) and using (C.89), the upper bound on P_{0,l_1} is further obtained as

$$\begin{aligned}
P_{0,l_1}&\le\frac{\eta e^{-\frac{1}{P}}}{2^{l_1-2}}\sum_{u=0}^{2^{l_1-3}-1}\Pr\left\{\frac{1}{2}+\frac{u}{2^{l_1-2}}\le p_{2,\min}\le\frac{1}{2}+\frac{u+1}{2^{l_1-2}}\right\}\\
&\quad+\frac{\eta e^{-\frac{1}{P}}}{2^{l_1-2}}\sum_{u=0}^{2^{l_1-3}-1}\Pr\left\{\frac{1}{2}-\frac{u+1}{2^{l_1-2}}\le p_{2,\min}\le\frac{1}{2}-\frac{u}{2^{l_1-2}}\right\}\\
&=\frac{\eta e^{-\frac{1}{P}}}{2^{l_1-2}}\Pr\left\{\frac{1}{2}\le p_{2,\min}\le 1\right\}+\frac{\eta e^{-\frac{1}{P}}}{2^{l_1-2}}\Pr\left\{0\le p_{2,\min}\le\frac{1}{2}\right\}\\
&=\frac{\eta e^{-\frac{1}{P}}}{2^{l_1-2}}\Pr\left\{0\le p_{2,\min}\le 1\right\}\le\frac{\eta e^{-\frac{1}{P}}}{2^{l_1-2}}.
\end{aligned}\tag{C.90}$$

For P_{1,l_2} in (C.84), since $\widetilde{\mathcal{R}_{1,l_1,l_2}}\cap\widetilde{\mathcal{R}_{1,l_1',l_2}}=\emptyset$ for $l_1\ne l_1'$, we have $P_{1,l_2}=\Pr\left\{\boldsymbol{H}\in\widehat{\mathcal{R}_{1,l_2}}\right\}$, where $\widehat{\mathcal{R}_{1,l_2}}=\bigcup_{l_1=0}^{\infty}\widetilde{\mathcal{R}_{1,l_1,l_2}}$. Define a cooperative quantizer $CQ_{MR,\mathsf{CT}}^{\aleph}$ that only executes Stage

1 of $CQ_{MR,\mathsf{CT}}$ (Stage 0 is skipped). Letting $\widehat{\mathcal{R}^{\aleph}_{1,l_2}} = \left\{\boldsymbol{H} : CQ^{\aleph}_{MR,\mathsf{CT}} \text{ ends after Round } l_2\right\}$, we obtain $\widehat{\mathcal{R}_{1,l_2}} \subseteq \widehat{\mathcal{R}^{\aleph}_{1,l_2}}$, thus, $P_{1,l_2} \leq \Pr\left\{\boldsymbol{H} \in \widehat{\mathcal{R}^{\aleph}_{1,l_2}}\right\}$. An upper bound on $\Pr\left\{\boldsymbol{H} \in \widehat{\mathcal{R}^{\aleph}_{1,l_2}}\right\}$ can be derived via using the same methodology for $\Pr\left\{\boldsymbol{H} \in \widehat{\mathcal{R}_{0,l_1}}\right\}$. Substituting the upper bounds in (C.85), (C.86) and (C.90) into (C.84) and applying the fact that $\sum_{l=3}^{\infty} \frac{l+1}{2^{l-2}}$ is finite yield (4.12) and conclude the proof of Theorem 4.3. □

D Supplementary Proofs for Chapter 5

D.1 Proof of Lemma 5.2

Proof. To clarify, the notation D_i for $i \in \mathbb{N}$ represents a positive constant independent of P, T and Δ. The average rate loss of $q_r(\cdot)$ can be expressed as

$$\mathtt{E}\left[r_{\text{loss}}\right] = \underbrace{\int_{\mathcal{H}_{0,\geq}} r_{\text{loss}} \prod_{i=1}^{2} f_{H_i}(H_i) \mathrm{d}H_i}_{=\mathtt{E}_{\geq}[r_{\text{loss}}]} + \underbrace{\int_{\mathcal{H}_{0,<}} r_{\text{loss}} \prod_{i=1}^{2} f_{H_i}(H_i) \mathrm{d}H_i}_{=\mathtt{E}_{<}[r_{\text{loss}}]},$$

where $\mathcal{H}_{0,\geq} = \{(H_1, H_2) : q_r(H_1) \geq q_r(H_2)\}$ and $\mathcal{H}_{0,<} = \{(H_1, H_2) : q_r(H_1) < q_r(H_2)\}$. We will only show $\mathtt{E}_{\geq}\left[r_{\text{loss}}\right] \leq \log_2\left(1 + D_0 \times P \times \max\left\{e^{-\frac{T\Delta}{\lambda_1}}, \Delta\right\}\right)$, and skip the proof for $\mathtt{E}_{<}\left[r_{\text{loss}}\right]$ due to similarity. Note that $q_r(H_1) \geq q_r(H_2)$ does not necessarily mean $H_1 \geq H_2$, since it is possible that $q_r(H_1) = q_r(H_2)$ and $H_1 < H_2$. When $q_r(H_1) \geq q_r(H_2)$, define

$$\mathsf{snr}_{\max} = \begin{cases} \alpha^{\star} H_1 = g_{\geq}(H_1, H_2), & \text{if } H_1 \geq H_2, \\ \alpha^{\star} H_2 = g_{<}(H_1, H_2), & \text{if } H_1 < H_2, \end{cases} \tag{D.91}$$

$$\mathsf{snr}_{q_r} = \alpha_{q_r} \times q_r(H_1) = g_{\geq}\left(q_r(H_1), q_r(H_2)\right), \mathsf{snr}_{\text{loss}} = \mathsf{snr}_{\max} - \mathsf{snr}_{q_r}.$$

where $g_{\geq}(x,y) = \frac{2xy}{\sqrt{(x+y)^2+4xy^2P}+x+y}$ and $g_{<}(x,y) = \frac{2xy}{\sqrt{(x+y)^2+4x^2yP}+x+y}$. Then, we have $r_{\text{loss}} = \log_2(1+P\times \mathsf{snr}_{\max}) - \log_2(1+P\times \mathsf{snr}_{q_r}) = \log_2\left(1+P\frac{\mathsf{snr}_{\text{loss}}}{1+P\times \mathsf{snr}_{q_r}}\right) \leq \log_2(1+P\times \mathsf{snr}_{\text{loss}})$. Grounded on this, the main steps of the proof are listed as follows:

(1) Partition $\mathcal{H}_{0,\geq}$ into the following mutually disjoint sub-regions $\mathcal{H}_1,\ldots,\mathcal{H}_4$:

$$\begin{aligned}
\mathcal{H}_1 &= \{(H_1,H_2): q_r(H_1)\geq q_r(H_2), H_1<T\Delta, H_2<T\Delta, H_1<\Delta \text{ or } H_2<\Delta\},\\
\mathcal{H}_2 &= \{(H_1,H_2): q_r(H_1)\geq q_r(H_2), H_1\geq H_2, \Delta\leq H_1<T\Delta, \Delta\leq H_2<T\Delta\}\\
\mathcal{H}_3 &= \{(H_1,H_2): q_r(H_1)= q_r(H_2), H_1< H_2, \Delta\leq H_1<T\Delta, \Delta\leq H_2<T\Delta\}\\
\mathcal{H}_4 &= \{(H_1,H_2): q_r(H_1)\geq q_r(H_2), H_1\geq T\Delta \text{ or } H_2\geq T\Delta\}.
\end{aligned}$$

Here, $\mathcal{H}_1$ and $\mathcal{H}_4$ are edge regions where $H_i<\Delta$ or $H_i\geq T\Delta$; $\mathcal{H}_2$ and $\mathcal{H}_3$ are the dominant regions where $\Delta\leq H_i<T\Delta$. It can be verified that $\mathcal{H}_i\cap\mathcal{H}_j=\emptyset$ for $i\neq j$, and $\mathcal{H}_{0,\geq}=\bigcup_{i=1}^4\mathcal{H}_i$.

(2) Let $\mathscr{E}_i = \int_{\mathcal{H}_i}\mathsf{snr}_{\text{loss}}\prod_{i=1}^2 f_{H_i}(H_i)\mathrm{d}H_i$. Then, $\mathtt{E}_{\geq}[\mathsf{snr}_{\text{loss}}] = \sum_{i=1}^4\mathscr{E}_i$. Prove $\mathscr{E}_i\leq D_i\times\max\left\{e^{-\frac{T\Delta}{\lambda_1}},\Delta\right\}$ for $i=1,\ldots,4$.

(3) After Steps (1) and (2), we obtain $\mathtt{E}_{\geq}[\mathsf{snr}_{\text{loss}}]\leq D_0\times\max\left\{e^{-\frac{T\Delta}{\lambda_1}},\Delta\right\}$. Based on Jensen's inequality, we have

$$\begin{aligned}
\mathtt{E}_{\geq}[r_{\text{loss}}] &\leq \mathtt{E}_{\geq}[\log_2(1+P\times\mathsf{snr}_{\text{loss}})] \leq \log_2(1+P\times\mathtt{E}_{\geq}[\mathsf{snr}_{\text{loss}}])\\
&\leq \log_2\left(1+D_0\times P\times\max\left\{e^{-\frac{T\Delta}{\lambda_1}},\Delta\right\}\right).
\end{aligned}$$

Now, we only need to show the upper bound on $\mathscr{E}_i$ in Step (2).

For $\mathscr{E}_1$, since $\mathcal{H}_1\subseteq\{(H_1,H_2): H_2\leq\Delta\}$ and $\mathsf{snr}_{\text{loss}}\leq\mathsf{snr}_{\max}\leq H_1$, we obtain

$$\mathscr{E}_1\leq\int_0^\infty H_1\frac{e^{-\frac{H_1}{\lambda_1}}}{\lambda_1}\mathrm{d}H_1\int_0^\Delta\frac{e^{-\frac{H_2}{\lambda_2}}}{\lambda_2}\mathrm{d}H_2 = \lambda_1\left(1-e^{-\frac{\Delta}{\lambda_2}}\right)\leq\lambda_1\times\frac{\Delta}{\lambda_2} = D_1\times\Delta,$$

where the last inequality follows since $1-e^{-x} \leq x$ for $x \geq 0$.

For $\mathscr{E}_2$, since $H_1 \geq H_2$ and $q_r(H_i) \leq H_i \leq q_r(H_i) + \Delta$ for $H_i \leq T\Delta$, we upper-bound $\mathsf{snr}_{\text{loss}}$ by

$$\begin{aligned}
\mathsf{snr}_{\text{loss}} &= \frac{2H_1H_2}{\underbrace{\sqrt{(H_1+H_2)^2+4H_1H_2^2P}}_{=\Upsilon}+(H_1+H_2)} \\
&\quad - \frac{2q_r(H_1)\,q_r(H_2)}{\underbrace{\sqrt{[q_r(H_1)+q_r(H_2)]^2+4q_r(H_1)\,q_r^2(H_2)\,P}+[q_r(H_1)+q_r(H_2)]}_{\leq \Upsilon+H_1+H_2}} \\
&\leq 2\frac{H_1H_2-q_r(H_1)\,q_r(H_2)}{\Upsilon+H_1+H_2} \leq 2\frac{H_1H_2-(H_1-\Delta)(H_2-\Delta)}{\Upsilon+H_1+H_2} \\
&= 2\Delta\frac{H_1+H_2-\Delta}{\Upsilon+H_1+H_2} \leq 2\Delta. \qquad \text{(D.92)}
\end{aligned}$$

Then, an upper bound on $\mathscr{E}_2$ can be $\mathscr{E}_2 \leq 2\Delta \int_{\mathcal{H}_2} \prod_{i=1}^{2} f_{H_i}(H_i)\mathrm{d}H_i \leq 2\Delta = D_2 \times \Delta$.

For $\mathscr{E}_3$, we have $q_r(H_1) = q_r(H_2) \leq H_1 < H_2$ and $q_r(H_i) \leq H_i \leq q_r(H_i)+\Delta$ hold for $(H_1,H_2) \in \mathcal{H}_3$. Similar to (D.92), we can also obtain $\mathsf{snr}_{\text{loss}} \leq 2\Delta$ and $\mathscr{E}_3 \leq D_3 \times \Delta$.

For $\mathscr{E}_4$, since $\mathcal{H}_4 \subseteq \{(H_1,H_2): H_1 > T\Delta\}$ and $\mathsf{snr}_{\text{loss}} \leq \mathsf{snr}_{\max} \leq H_2$, the upper-bound on $\mathscr{E}_4$ can be $\mathscr{E}_4 \leq \int_{T\Delta}^{\infty} f_{H_1}(H_1)\mathrm{d}H_1 \int_0^{\infty} H_2 f_{H_2}(H_2)\mathrm{d}H_2 = \int_{T\Delta}^{\infty} \frac{e^{-\frac{H_1}{\lambda_1}}}{\lambda_1}\mathrm{d}H_1 \int_0^{\infty} H_2 \frac{e^{-\frac{H_2}{\lambda_2}}}{\lambda_2}\mathrm{d}H_2 = \lambda_2 e^{-\frac{T\Delta}{\lambda_1}} = D_4 \times e^{-\frac{T\Delta}{\lambda_1}}$. We have accomplished Step (2) and the proof of (5.6) is complete. □

D.2 Proof of Lemma 5.3

Proof. When the uniform quantizer $q_o(\cdot)$ is applied, the outage probability loss in (5.4) is rewritten as

$$\mathtt{out}_{\text{loss},q_o} = \underbrace{\int_{I_{0,\geq}} \mathbf{1}_{\min\{r_1(\alpha_{q_o}),r_2(\alpha_{q_o})\}<r_{\text{th}}} \prod_{i=1}^{2} f_{H_i}(H_i)\mathrm{d}H_i}_{=\mathtt{out}_{\geq,\text{loss},q_o}}$$

$$+\underbrace{\int_{I_{0,<}} \mathbf{1}_{\min\{r_1(\alpha_{q_o}),r_2(\alpha_{q_o})\}<r_{\rm th}} \prod_{i=1}^{2} f_{H_i}(H_i)\mathrm{d}H_i}_{=\mathtt{out}_{<,\mathrm{loss},q_o}}.$$

where

$$\begin{aligned}
I_{0,\geq} &= \left\{(H_1,H_2): q_r(H_1)\geq q_r(H_2), r_{\max}=\log_2\left(1+P\times \mathsf{snr}_{\max}\right)\geq r_{\rm th}\right\}\\
&= \left\{(H_1,H_2): q_r(H_1)\geq q_r(H_2), \mathsf{snr}_{\max}\geq \tfrac{\beta}{P}=\tfrac{2^{r_{\rm th}}-1}{P}\right\},\\
I_{0,<} &= \left\{(H_1,H_2): q_r(H_1)< q_r(H_2), \mathsf{snr}_{\max}< \tfrac{\beta}{P}\right\}.
\end{aligned}$$

and $\mathsf{snr}_{\max}$ is defined in (D.91). We show

$$\mathtt{out}_{\geq,\mathrm{loss},q_o}\leq D_5\times e^{-\frac{D_6}{P}}\times\frac{1+\sqrt{P}}{P}\times\max\left\{e^{-\frac{T\Delta}{\lambda_1}},\Delta^{\frac{1}{2}},\Delta^{\frac{3}{2}}\right\},$$

and skip the proof for $\mathtt{out}_{<,\mathrm{loss},q_o}$ due to similarity. The main steps of the proof are:

(1) Partition $I_{0,\geq}$ into the following mutually disjoint sub-regions:

$$\begin{aligned}
I_1 &= \left\{(H_1,H_2): q_r(H_1)\geq q_r(H_2), \mathsf{snr}_{\max}\geq\tfrac{\beta}{P}, H_1\leq\Delta, H_2\leq\Delta\right\},\\
I_2 &= \left\{(H_1,H_2): q_r(H_1)\geq q_r(H_2), \mathsf{snr}_{\max}=g_{\geq}(H_1,H_2)\geq\tfrac{\beta}{P},\right.\\
&\qquad \left.\Delta<H_1\leq T\Delta, H_2\leq\Delta\right\},\\
I_3 &= \left\{(H_1,H_2): q_r(H_1)\geq q_r(H_2), H_1\geq H_2, g_{\geq}(H_1,H_2)\geq\tfrac{\beta}{P},\right.\\
&\qquad \left.\Delta<H_1\leq T\Delta, \Delta<H_2\leq T\Delta\right\},\\
I_4 &= \left\{(H_1,H_2): q_r(H_1)= q_r(H_2), H_1< H_2, g_{<}(H_1,H_2)\geq\tfrac{\beta}{P},\right.\\
&\qquad \left.\Delta<H_1\leq T\Delta, \Delta<H_2\leq T\Delta\right\},\\
I_5 &= \left\{(H_1,H_2): q_r(H_1)\geq q_r(H_2), \mathsf{snr}_{\max}\geq\tfrac{\beta}{P}, H_1>T\Delta \text{ or } H_2>T\Delta\right\}.
\end{aligned}$$

Here, I_1, I_2 and I_5 are the marginal regions where $H_i\leq\Delta$ or $H_i>T\Delta$; I_3 and I_4 are the main regions where $\Delta<H_i\leq T\Delta$. It can be verified that $I_i\cap I_j=\emptyset$ for $i\neq j$,

and $I_{0,\geq} = \bigcup_{i=1}^{5} I_i$.

(2) Let $\mathscr{F}_i = \int_{I_i} \mathbf{1}_{\min\{r_1(\alpha_{q_o}), r_2(\alpha_{q_o})\} < r_{\text{th}}} \prod_{i=1}^{2} f_{H_i}(H_i) \mathrm{d}H_i$. Then, $\mathtt{out}_{\geq,\text{loss},q_o} = \sum_{i=1}^{5} \mathscr{F}_i$. Prove $\mathscr{F}_i \leq D_{2i+5} \times e^{-\frac{D_{2i+6}}{P}} \times \frac{1+\sqrt{P}}{P} \times \max\left\{e^{-\frac{T\Delta}{\lambda_1}}, \Delta^{\frac{1}{2}}, \Delta^{\frac{3}{2}}\right\}$ for $i = 1, \ldots, 5$.

Now, we need to show the upper bound on $\mathscr{F}_i$ in Step (2).

For $\mathscr{F}_1$, we have $q_o(H_1) = q_o(H_2) = \Delta \geq H_2$, and thus, $\alpha_{q_o} = \frac{1}{\sqrt{P\Delta+1}+1} \leq \frac{1}{\sqrt{PH_2+1}+1}$. For any $(H_1, H_2) \in I_1$, since $g_{\geq}(x,y) \leq \min\{x,y\}$ and $g_{<}(x,y) \leq \min\{x,y\}$, it must have $\frac{\beta}{P} \leq \mathsf{snr}_{\max} \leq \min\{H_1, H_2\}$. Moreover, we obtain $\mathbf{1}_{\min\{r_1(\alpha_{q_o}), r_2(\alpha_{q_o})\} < r_{\text{th}}} \leq \mathbf{1}_{r_1(\alpha_{q_o}) < r_{\text{th}}} + \mathbf{1}_{r_2(\alpha_{q_o}) < r_{\text{th}}}$, and

$$\mathbf{1}_{r_1(\alpha_{q_o}) < r_{\text{th}}} = \mathbf{1}_{H_1 \times \alpha_{q_o} < \frac{\beta}{P}} = \mathbf{1}_{H_1 < \beta \frac{\sqrt{P\Delta+1}+1}{P}},$$
$$\mathbf{1}_{r_2(\alpha_{q_o}) < r_{\text{th}}} = \mathbf{1}_{\frac{H_2(1-\alpha_{q_o})}{PH_2\alpha_{q_o}+1} < \frac{\beta}{P}} \leq \mathbf{1}_{\frac{H_2\left(1-\frac{1}{\sqrt{PH_2+1}+1}\right)}{PH_2 \times \frac{1}{\sqrt{PH_2+1}+1}+1} < \frac{\beta}{P}} = \mathbf{1}_{H_2 < \frac{\beta^2+2\beta}{P}}.$$

Thus, an upper bound on $\mathscr{F}_1$ is

$$\begin{aligned}
\mathscr{F}_1 &\leq \int_{I_1} \mathbf{1}_{H_1 < \beta \frac{\sqrt{P\Delta+1}+1}{P}} \prod_{i=1}^{2} f_{H_i}(H_i) \mathrm{d}H_i + \int_{I_1} \mathbf{1}_{H_2 < \frac{\beta^2+2\beta}{P}} \prod_{i=1}^{2} f_{H_i}(H_i) \mathrm{d}H_i \\
&\leq \int_{\frac{\beta}{P}}^{\beta \frac{\sqrt{P\Delta+1}+1}{P}} \frac{e^{-\frac{H_1}{\lambda_1}}}{\lambda_1} \int_{\frac{\beta}{P}}^{\Delta} \frac{e^{-\frac{H_2}{\lambda_2}}}{\lambda_2} \mathrm{d}H_1 \mathrm{d}H_2 + \int_{\frac{\beta}{P}}^{\Delta} \frac{e^{-\frac{H_1}{\lambda_1}}}{\lambda_1} \int_{\frac{\beta}{P}}^{\frac{\beta^2+2\beta}{P}} \frac{e^{-\frac{H_2}{\lambda_2}}}{\lambda_2} \mathrm{d}H_1 \mathrm{d}H_2 \\
&\leq \frac{e^{-\frac{\frac{\beta}{P}}{\lambda_1}}}{\lambda_1} \times \left[\beta \frac{\sqrt{P\Delta+1}+1}{P} - \frac{\beta}{P}\right] \times \frac{1}{\lambda_2} \times \left[\Delta - \frac{\beta}{P}\right] \\
&\quad + \frac{1}{\lambda_1} \times \left[\Delta - \frac{\beta}{P}\right] \times \frac{e^{-\frac{\frac{\beta}{P}}{\lambda_2}}}{\lambda_2} \times \left[\frac{\beta^2+2\beta}{P} - \frac{\beta}{P}\right] \\
&\leq \frac{e^{-\frac{\frac{\beta}{P}}{\lambda_1}}}{\lambda_1} \times \beta \times \frac{\overbrace{\sqrt{P\Delta+1}}^{\leq \sqrt{P\Delta}+1}}{P} \times \frac{1}{\lambda_2} \times \Delta + \frac{1}{\lambda_1} \times \Delta \times \frac{e^{-\frac{\frac{\beta}{P}}{\lambda_2}}}{\lambda_2} \times \frac{\beta^2+\beta}{P} \\
&\leq D_{17} \times e^{-\frac{D_{18}}{P}} \times \frac{\sqrt{P\Delta}+1}{P} \times \Delta + D_{19} \times e^{-\frac{D_{20}}{P}} \times \frac{\Delta}{P} \\
&\leq D_7 \times e^{-\frac{D_8}{P}} \times \frac{1+\sqrt{P}}{P} \times \max\left\{e^{-\frac{T\Delta}{\lambda_1}}, \Delta^{\frac{1}{2}}, \Delta^{\frac{3}{2}}\right\}. \qquad \text{(D.93)}
\end{aligned}$$

For $\mathscr{F}_2$, let $\mathscr{F}_{2,i} = \int_{I_2} \mathbf{1}_{r_i(\alpha_{q_o})<r_{\rm th}} \prod_{i=1}^{2} f_{H_i}(H_i)\mathrm{d}H_i$ for $i = 1, 2$. Then, $\mathscr{F}_2 \leq \mathscr{F}_{2,1} + \mathscr{F}_{2,2}$. For $\mathscr{F}_{2,1}$, since $H_1 > H_2$ for $(H_1, H_2) \in I_2$ and $g_{\geq}(x, y)$ is increasing on x and y, we have

$$\begin{aligned}
\mathbf{1}_{r_1(\alpha_{q_o})<r_{\rm th}} &= \mathbf{1}_{\frac{2H_1\times q_o(H_2)}{\sqrt{[q_o(H_1)+q_o(H_2)]^2+4q_o(H_1)q_o^2(H_2)P}+[q_o(H_1)+q_o(H_2)]}<\frac{\beta}{P}} \\
&\leq \mathbf{1}_{\frac{2(q_o(H_1)-\Delta)\times q_o(H_2)}{\sqrt{[q_o(H_1)+q_o(H_2)]^2+4q_o(H_1)q_o^2(H_2)P}+[q_o(H_1)+q_o(H_2)]}<\frac{\beta}{P}} = \mathbf{1}_{g_{\geq}(q_o(H_1),q_o(H_2))<\frac{\beta}{P}\times\frac{1}{1-\frac{\Delta}{q_o(H_1)}}} \qquad \text{(D.94)} \\
&\leq \mathbf{1}_{g_{\geq}(q_o(H_1),q_o(H_2))<\frac{\beta}{P}\times\left(1+\frac{2\Delta}{q_o(H_1)}\right)} \leq \mathbf{1}_{g_{\geq}(q_o(H_1),q_o(H_2))<\frac{\beta}{P}\times\left(1+\frac{2\Delta}{q_o(H_2)}\right)} \qquad \text{(D.95)} \\
&\leq \mathbf{1}_{g_{\geq}(q_o(H_1),q_o(H_2))<\frac{\beta}{P}\times\left(1+\frac{2\Delta}{H_2}\right)} \leq \mathbf{1}_{g_{\geq}(H_1,H_2)<\frac{\beta}{P}\times\left(1+\frac{2\Delta}{H_2}\right)}, \qquad \text{(D.96)}
\end{aligned}$$

where (D.94) follows from $q_o(H_1) \leq H_1+\Delta$, (D.95) follows from $\left(1-\frac{\Delta}{q_o(H_1)}\right)\times\left(1+\frac{2\Delta}{q_o(H_1)}\right) \geq 1$ because $q_o(H_1) \geq 2\Delta > q_o(H_2) = \Delta$, and (D.96) follows from $q_o(H_2) \geq H_2$ as well as $g_{\geq}\left(q_o(H_1), q_o(H_2)\right) \geq g_{\geq}\left(H_1, H_2\right)$. Then, we obtain

$$\mathscr{F}_{2,1} \leq \int_{I_2'=I_2\cap\left\{(H_1,H_2):g_{\geq}(H_1,H_2)<\frac{\beta}{P}\times\left(1+\frac{2\Delta}{H_2}\right)\right\}} \prod_{i=1}^{2} f_{H_i}(H_i)\mathrm{d}H_i.$$

We change the integration variables from (H_1, H_2) to (ϕ, H_2) where $\phi = g_{\geq}(H_1, H_2)$. Then, $H_1 = \frac{\phi^2P+\phi}{H_2-\phi}\times H_2$, and the Jacobian matrix is $\left|\frac{\mathrm{d}H_1}{\mathrm{d}\phi}\right| = \frac{2\phi PH_2+H_2-\phi^2P}{(H_2-\phi)^2}\times H_2 \leq \frac{2\phi PH_2+H_2}{(H_2-\phi)^2}\times H_2 \leq \frac{2\phi PH_2+2H_2}{(H_2-\phi)^2}\times H_2 = \frac{2(\phi P+1)}{(H_2-\phi)^2}\times H_2^2$. For any $(H_1, H_2) \in I_2'$, we have: (i) $\frac{\beta}{P} \leq \phi = g_{\geq}(H_1, H_2) \leq H_2$ and $\phi < \frac{\beta}{P}\times\left(1+\frac{2\Delta}{H_2}\right)$; (ii) since $H_1 \geq H_2$, $H_1 = \frac{\phi^2P+\phi}{H_2-\phi}\times H_2 \geq H_2$, then, $H_2 \leq \phi^2P+2\phi$. Therefore, $\mathscr{F}_{2,1}$ is derived as

$$\mathscr{F}_{2,1} \leq \int_{I_2''=\left\{(H_1,H_2):\frac{\beta}{P}\leq H_2\leq\phi^2P+2\phi,\frac{\beta}{P}\leq\phi\leq\min\left\{H_2,\frac{\beta}{P}\left(1+\frac{2\Delta}{H_2}\right)\right\}\right\}} \prod_{i=1}^{2} f_{H_i}(H_i)\mathrm{d}H_i.$$

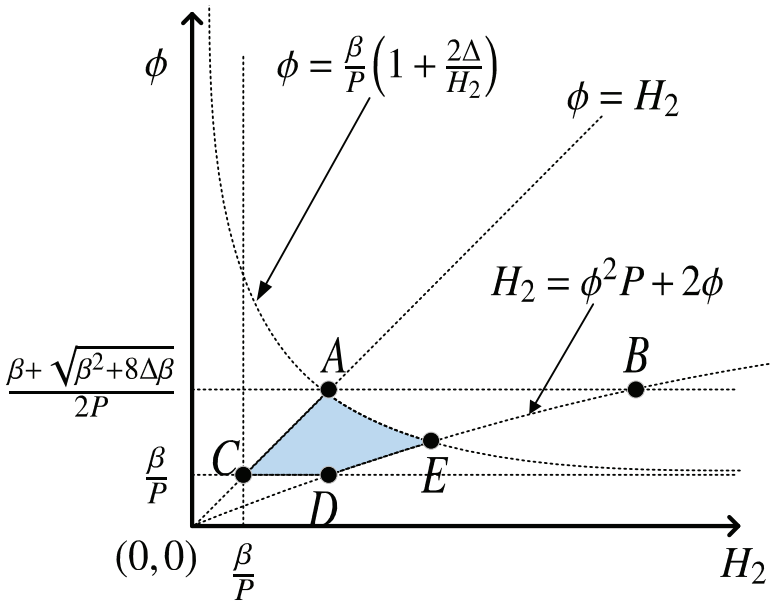


Figure D.2: The integration region $I_2^{''}$.

The integration region $I_2^{''}$ is demonstrated in Fig. D.2 as the shaded area surrounded by the points A, E, D and C. It can be strictly proven that $I_2^{''}$ is within the region surrounded the points A, B, D and C. Recall that $H_1 = \frac{\phi^2 P+\phi}{H_2-\phi} \times H_2$ and $\left|\frac{\mathrm{d}H_1}{\mathrm{d}\phi}\right| \leq \frac{2(\phi P+1)}{(H_2-\phi)^2} \times H_2^2$. Then, we have

$$
\begin{aligned}
\mathscr{F}_{2,1} &\leq \int_{\frac{\beta}{P}}^{\frac{\beta+\sqrt{\beta^2+8\Delta\beta}}{2P}} \int_{\phi}^{\phi^2 P+2\phi} \frac{e^{-\frac{H_2}{\lambda_2}}}{\lambda_2} \times \frac{e^{-\frac{1}{\lambda_1}\times\frac{\phi^2 P+\phi}{H_2-\phi}\times H_2}}{\lambda_1} \times \frac{2\left(\phi P+1\right)}{\left(H_2-\phi\right)^2} \times H_2^2 \mathrm{d}\phi \mathrm{d}H_2 \\
&\overset{z=H_2-\phi}{=} D_{21} \int_{\frac{\beta}{P}}^{\frac{\beta+\sqrt{\beta^2+8\Delta\beta}}{2P}} \int_{0}^{\phi^2 P+\phi} \underbrace{e^{-\frac{z}{\lambda_2}-\frac{\phi}{\lambda_2}}}_{\leq e^{-\frac{z}{\lambda_2}} \times e^{-\frac{\frac{\beta}{P}}{\lambda_2}}} \times \underbrace{e^{-\frac{1}{\lambda_1}\times\frac{\phi^2 P+\phi}{z}\times(z+\phi)}}_{\leq e^{-\frac{\phi^2(\phi P+1)}{\lambda_1 z}}} \\
&\quad \times \frac{\phi P+1}{z^2} \times (z+\phi)^2 \mathrm{d}\phi \mathrm{d}z \\
&\leq D_{21} \times e^{-\frac{\beta}{\lambda_2 P}} \int_{\frac{\beta}{P}}^{\frac{\beta+\sqrt{\beta^2+8\Delta\beta}}{2P}} \int_{0}^{\phi^2 P+\phi} e^{-\frac{z}{\lambda_2}} e^{-\frac{\phi^2(\phi P+1)}{\lambda_1 z}} \\
&\quad \times (\phi P+1) \times \left[1+\frac{2\phi}{z}+\frac{\phi^2}{z^2}\right] \mathrm{d}\phi \mathrm{d}z. \qquad \text{(D.97)}
\end{aligned}
$$

Using the inequalities: (i) $\int_0^\infty x^{v-1} e^{-\frac{\beta}{x}-\gamma x} \mathrm{d}x = 2\left(\frac{\beta}{\gamma}\right)^{\frac{v}{2}} \mathcal{K}_v\left(2\sqrt{\beta\gamma}\right)$ [14, Eq. (3.471.9)] with

$\mathcal{K}_v(z)$ being the modified bessel function of the second kind, (ii) $\mathcal{K}_0(x) \leq \frac{2}{x}$ and $\mathcal{K}_{-1}(x) = \mathcal{K}_1(x) \leq \frac{1}{x}$ for $x > 0$ [10, Eq. (27)], after lengthy but basic calculations, we obtain $\mathscr{F}_{2,1} \leq D_{22} \times e^{-\frac{D_{23}}{P}} \times \frac{\Delta+\sqrt{\Delta}}{P}$. Detailed calculations for (D.97) can be found in Appendix B of [62].

For $\mathscr{F}_{2,2}$, because $H_1 > H_2$ and $q_o(H_1) > q_o(H_2) = \Delta \geq H_2$, we have

$$\begin{aligned}\alpha_{q_o} &\leq \frac{2q_o(H_2)}{\sqrt{\left[q_o(H_2)+q_o(H_2)\right]^2 + 4q_o(H_2)q_o^2(H_2)P} + q_o(H_2) + q_o(H_2)} \\ &= \frac{1}{\sqrt{q_o(H_2)P+1}+1} = \frac{1}{\sqrt{P\Delta+1}+1}. \end{aligned} \tag{D.98}$$

Since $r_2(\alpha_{q_o})$ is decreasing on α_{q_o}, we obtain $r_2(\alpha_{q_o}) \geq r_2\left(\frac{1}{\sqrt{P\Delta+1}+1}\right)$ and $\mathbf{1}_{r_2(\alpha_{q_o})<r_{\text{th}}} \leq \mathbf{1}_{r_2\left(\frac{1}{\sqrt{P\Delta+1}+1}\right)<r_{\text{th}}} = \mathbf{1}_{\frac{H_2\left(1-\frac{1}{\sqrt{P\Delta+1}+1}\right)}{PH_2\frac{1}{\sqrt{P\Delta+1}+1}+1}<\frac{\beta}{P}} = \mathbf{1}_{\frac{H_2\sqrt{P\Delta+1}}{PH_2+1+\sqrt{P\Delta+1}}<\frac{\beta}{P}} \leq \mathbf{1}_{\frac{H_2\sqrt{P\Delta+1}}{P\Delta+1+\sqrt{P\Delta+1}}<\frac{\beta}{P}} = \mathbf{1}_{H_2\leq\frac{\beta(\sqrt{P\Delta+1}+1)}{P}}$.
Similar to (D.93), we will have $\mathscr{F}_{2,2} \leq \int_{I_2} \mathbf{1}_{H_2<\beta\frac{\sqrt{P\Delta+1}+1}{P}} \prod_{i=1}^2 f_{H_i}(H_i)\mathrm{d}H_i \leq D_{24} \times e^{-\frac{D_{25}}{P}} \times \frac{\sqrt{P\Delta}+1}{P} \times \Delta$. Together with the upper bound on $\mathscr{F}_{2,1}$, we obtain $\mathscr{F}_2 \leq \mathscr{F}_{2,1} + \mathscr{F}_{2,2} \leq D_{22} \times e^{-\frac{D_{23}}{P}} \times \frac{\Delta+\sqrt{\Delta}}{P} + D_{24} \times e^{-\frac{D_{25}}{P}} \times \frac{\sqrt{P\Delta}+1}{P} \times \Delta \leq D_9 \times e^{-\frac{D_{10}}{P}} \times \frac{1+\sqrt{P}}{P} \times \max\left\{e^{-\frac{T\Delta}{\lambda_1}}, \Delta^{\frac{1}{2}}, \Delta^{\frac{3}{2}}\right\}$.

For $\mathscr{F}_3$, since $q_o(H_1) \geq q_o(H_2)$ and $q_o(H_i) - \Delta \leq H_i \leq q_o(H_i)$ for $i = 1, 2$, we obtain

$$\begin{aligned} r_1(\alpha_{q_o}) &= \log_2\left(1 + PH_1 \times \alpha_{q_o}\right) \geq \log_2\left(1 + P \times \left(q_o(H_1) - \Delta\right) \times \alpha_{q_o}\right) \\ &= \log_2\left(1 + P \times q_o(H_1) \times \alpha_{q_o} - P \times \Delta \times \alpha_{q_o}\right) \\ &= \log_2\left(1 + P \times g_{\geq}(q_o(H_1), q_o(H_2)) - P \times g_{\geq}(q_o(H_1), q_o(H_2)) \times \frac{\Delta}{q_o(H_1)}\right) \\ &= \log_2\left(1 + P \times g_{\geq}(q_o(H_1), q_o(H_2)) \times \left(1 - \frac{\Delta}{q_o(H_1)}\right)\right) \\ &\geq \log_2\left(1 + P \times g_{\geq}(q_o(H_1), q_o(H_2)) \times \left(1 - \frac{\Delta}{q_o(H_2)}\right)\right) \\ &\geq \log_2\left(1 + P \times g_{\geq}(H_1, H_2) \times \left(1 - \frac{\Delta}{q_o(H_2)}\right)\right), \\ r_2(\alpha_{q_o}) &= \log_2\left(1 + \frac{H_2\left(1-\alpha_{q_o}\right)}{H_2\alpha_{q_o} + \frac{1}{P}}\right) = \log_2\left(1 + \frac{\left(q_o(H_2) - \Delta\right) \times \left(1-\alpha_{q_o}\right)}{\left(q_o(H_2) - \Delta\right) \times \alpha_{q_o} + \frac{1}{P}}\right) \\ &\geq \log_2\left(1 + \frac{\left(q_o(H_2) - \Delta\right) \times \left(1-\alpha_{q_o}\right)}{q_o(H_2) \times \alpha_{q_o} + \frac{1}{P}}\right) \end{aligned} \tag{D.99}$$

$$
\begin{aligned}
&= \log_2\left(1+\frac{q_o(H_2)\times(1-\alpha_{q_o})}{q_o(H_2)\times\alpha_{q_o}+\frac{1}{P}}-\frac{\Delta\times(1-\alpha_{q_o})}{q_o(H_2)\times\alpha_{q_o}+\frac{1}{P}}\right)\\
&= \log_2\left(1+P\times g_{\geq}(q_o(H_1),q_o(H_2))\times\left(1-\frac{\Delta}{q_o(H_2)}\right)\right)\\
&\geq \log_2\left(1+P\times g_{\geq}(H_1,H_2)\times\left(1-\frac{\Delta}{q_o(H_2)}\right)\right),
\end{aligned}
$$

Therefore, we have

$$
\begin{aligned}
\mathbf{1}_{\min\{r_1(\alpha_{q_o}),r_2(\alpha_{q_o})\}<r_{\rm th}} &\leq \mathbf{1}_{\log_2\left(1+P\times g_{\geq}(H_1,H_2)\times\left(1-\frac{\Delta}{q_o(H_2)}\right)\right)<r_{\rm th}} = \mathbf{1}_{g_{\geq}(H_1,H_2)<\frac{\beta}{P\left(1-\frac{\Delta}{q_o(H_2)}\right)}}\\
&\leq \mathbf{1}_{g_{\geq}(H_1,H_2)<\frac{\beta}{P}\left(1+\frac{2\Delta}{q_o(H_2)}\right)} \leq \mathbf{1}_{g_{\geq}(H_1,H_2)<\frac{\beta}{P}\left(1+\frac{2\Delta}{H_2}\right)}, \qquad \text{(D.100)}
\end{aligned}
$$

where (D.100) is because $\left(1-\frac{\Delta}{q_o(H_2)}\right)\times\left(1+\frac{2\Delta}{q_o(H_2)}\right) = 1+\frac{\Delta}{q_o(H_2)}-2\left(\frac{\Delta}{q_o(H_2)}\right)^2 \geq 1$ since $q_o(H_2)\geq 2\Delta$ for $(H_1,H_2)\in I_3$, and $q_o(H_2)\geq H_2$. Similar to (D.96) and (D.97), we can obtain an upper bound on $\mathscr{F}_3$ (the detailed derivation is omitted due to similarity). For $\mathscr{F}_4$, its upper bound can be developed in the same way as the upper bound on $\mathscr{F}_3$.

For $\mathscr{F}_5$, when $H_1\geq H_2\geq\Delta$, since $g_{\geq}(H_1,H_2)\geq\frac{2H_1H_2}{\sqrt{(H_1+H_1)^2+4H_1^2H_2P}+H_1+H_1} = \frac{H_2}{\sqrt{PH_2+1}+1}$, we obtain from (D.100) that

$$
\begin{aligned}
\mathbf{1}_{\min\{r_1(\alpha_{q_o}),r_2(\alpha_{q_o})\}<r_{\rm th}} &\leq \mathbf{1}_{g_{\geq}(H_1,H_2)<\frac{\beta}{P}\left(1+\frac{2\Delta}{H_2}\right)} \leq \mathbf{1}_{g_{\geq}(H_1,H_2)<\frac{\beta}{P}\left(1+\frac{2\Delta}{\Delta}\right)=\frac{3\beta}{P}}\\
&\leq \mathbf{1}_{\frac{H_2}{\sqrt{1+H_2P}+1}<\frac{3\beta}{P}} = \mathbf{1}_{H_2<\frac{D_{26}}{P}}, \qquad \text{(D.101)}
\end{aligned}
$$

where $D_{26} = (3\beta+1)^2-1$. Similarly, when $H_1<H_2$, we have $\mathbf{1}_{\min\{r_1(\alpha_{q_o}),r_2(\alpha_{q_o})\}<r_{\rm th}} \leq \mathbf{1}_{H_1<\frac{D_{26}}{P}}$. Therefore, an upper bound on $\mathscr{F}_5$ is

$$
\begin{aligned}
\mathscr{F}_5 \leq &\int_{I_4\cap\{(H_1,H_2):H_1\geq H_2\}} \mathbf{1}_{H_2<\frac{D_{26}}{P}}\times\prod_{i=1}^{2} f_{H_i}(H_i)\mathrm{d}H_i\\
&+\int_{I_4\cap\{(H_1,H_2):H_1<H_2\}} \mathbf{1}_{H_1<\frac{D_{26}}{P}}\times\prod_{i=1}^{2} f_{H_i}(H_i)\mathrm{d}H_i
\end{aligned}
$$

$$
\begin{aligned}
&\leq \underbrace{\int_{T\Delta}^{\infty} \frac{1}{\lambda_1} e^{-\frac{H_1}{\lambda_1}} \mathrm{d}H_1}_{=e^{-\frac{T\Delta}{\lambda_1}}} \int_{\frac{\beta}{P}}^{\frac{D_{26}}{P}} \frac{1}{\lambda_2} \underbrace{e^{-\frac{H_2}{\lambda_2}}}_{\leq e^{-\frac{\beta}{P\lambda_2}} \leq e^{-\frac{\beta}{P\lambda_1}}} \mathrm{d}H_2 \\
&+ \underbrace{\int_{T\Delta}^{\infty} \frac{1}{\lambda_2} e^{-\frac{H_2}{\lambda_2}} \mathrm{d}H_2}_{=e^{-\frac{T\Delta}{\lambda_2}} \leq e^{-\frac{T\Delta}{\lambda_1}}} \int_{\frac{\beta}{P}}^{\frac{D_{26}}{P}} \frac{1}{\lambda_1} \underbrace{e^{-\frac{H_1}{\lambda_1}}}_{\leq e^{-\frac{\beta}{P\lambda_1}}} \mathrm{d}H_1 \qquad \text{(D.102)} \\
&\leq e^{-\frac{T\Delta}{\lambda_1}} \times \frac{1}{\lambda_2} \times e^{-\frac{\beta}{P\lambda_1}} \times \frac{D_{26}-\beta}{P} + e^{-\frac{T\Delta}{\lambda_1}} \times \frac{1}{\lambda_1} \times e^{-\frac{\beta}{P\lambda_1}} \times \frac{D_{26}-\beta}{P} \\
&\leq D_{27} \times e^{-\frac{D_{28}}{P}} \times \frac{1}{P} \times e^{-\frac{T\Delta}{\lambda_1}} \\
&\leq D_{15} \times e^{-\frac{D_{16}}{P}} \times \frac{1+\sqrt{P}}{P} \times \max\left\{e^{-\frac{T\Delta}{\lambda_1}}, \Delta^{\frac{1}{2}}, \Delta^{\frac{3}{2}}\right\},
\end{aligned}
$$

where (D.102) is based on the assumption that $\lambda_1 \geq \lambda_2$. This completes the proof of the upper bound on $\mathtt{out}_{\mathrm{loss},q_o}$ in (5.12). □

D.3 Proof of Lemma 5.4

Proof. It is trivial to obtain the maximum diversity order for both receivers is 1 in the full-CSI case.[1] When $q_o(\cdot)$ is employed, the outage probability of Receiver i is $\mathtt{out}_{q_o,i} = \int \mathbf{1}_{r_i(\alpha_{q_o})<r_{\mathrm{th}}} \prod_{i=1}^{2} f_{H_i}(H_i)\mathrm{d}H_i$ for $i = 1, 2$. Following the derivations of $\mathscr{F}_i$ for $i = 1, \ldots, 5$ in Appendix D.2, we will obtain $\mathtt{out}_{q_o,1} \leq \mathtt{out}_{\min} + D_{29} \times e^{-\frac{D_{30}}{P}} \times \left[\frac{\sqrt{\Delta}+e^{-\frac{T\Delta}{\lambda_1}}}{P} + \frac{\Delta^{\frac{3}{2}}}{\sqrt{P}}\right]$ and $\mathtt{out}_{q_o,2} \leq \mathtt{out}_{\min} + D_{31} \times e^{-\frac{D_{32}}{P}} \times \frac{D_{33}+\Delta+e^{-\frac{T\Delta}{\lambda_1}}}{P}$.[2] Therefore, for fixed Δ, the diversity orders of $\frac{1}{2}$ and 1 are achievable for Receivers 1 and 2, respectively.

For Receiver 1, $\frac{\Delta^{\frac{3}{2}}}{\sqrt{P}}$ in the upper bound on $\mathtt{out}_{q_o,1}$ is the bottleneck for diversity gains. If we scale Δ as $\Delta^{\frac{3}{2}} \sim_P \frac{1}{\sqrt{P}}$, i.e., $\Delta \sim_P P^{-\frac{1}{3}}$, the diversity order of 1 is also achievable for Receiver

[1] Detailed derivations for the maximum diversity order can be found in Appendix C of [62].

[2] Note that when we derive the diversity order for $\mathscr{F}_{2,2}$, we will not use its upper bound here. From (D.98), we obtain $\alpha_{q_o} \leq \frac{1}{\sqrt{P\Delta+1}+1} \leq \frac{1}{\sqrt{PH_2+1}+1}$, and $\mathbf{1}_{r_2(\alpha_{q_o})<r_{\mathrm{th}}} \leq \mathbf{1}_{r_2\left(\frac{1}{\sqrt{PH_2+1}+1}\right)<r_{\mathrm{th}}} = \mathbf{1}_{H_2<\frac{\beta^2+\beta}{P}}$, then, it is trivial to obtain that $\mathscr{F}_{2,2} \leq D_{34} \times \frac{e^{-\frac{D_{35}}{P}}}{P}$.

1. □

D.4 Proof of Lemma 5.5

Proof. Given K and $\beta > 0$, define the following two optimization problems:

(**P1**) $r^{\star}_{\max}(K,\beta) = \max\limits_{\boldsymbol{\alpha}=[\alpha_1,\ldots,\alpha_K]} \min\limits_{k=1,\ldots,K} r_k(\boldsymbol{\alpha})$, subject to $0 \le \alpha_k \le \beta$ and $\sum_{k=1}^{K} \alpha_k = \beta$.

(**P2**) $r^{\dagger}_{\max}(K,\beta) = \max\limits_{\boldsymbol{\alpha}=[\alpha_1,\ldots,\alpha_K]} \min\limits_{k=1,\ldots,K} r_k(\boldsymbol{\alpha})$, subject to $r_1(\boldsymbol{\alpha}) = \cdots = r_K(\boldsymbol{\alpha})$, $0 \le \alpha_k \le \beta$, and $\sum_{k=1}^{K} \alpha_k = \beta$,

where (**P1**) is the original optimization problem in (5.15) when $\beta = 1$. We will show that the maximum minimum rates of (**P1**) and (**P2**) are the same, i.e., $r^{\star}_{\max}(K,\beta) = r^{\dagger}_{\max}(K,\beta)$, which proves the lemma.

Denote the optimal power allocations for (**P1**) and (**P2**) by $\boldsymbol{\alpha}^{\star}_K(\beta) = \left[\alpha^{\star}_{1,K}(\beta),\ldots,\alpha^{\star}_{K,K}(\beta)\right]$ and $\boldsymbol{\alpha}^{\dagger}_K(\beta) = \left[\alpha^{\dagger}_{1,K}(\beta),\ldots,\alpha^{\dagger}_{K,K}(\beta)\right]$, respectively. Since $r^{\star}_{\max}(K,\beta) \ge r^{\dagger}_{\max}(K,\beta)$, it is sufficient to prove that $r^{\star}_{\max}(K,\beta) \le r^{\dagger}_{\max}(K,\beta)$.

The proof for $K = 2$ is provided in the proof of Theorem 1. By induction, assume $r^{\star}_{\max}(K,\beta) = r^{\dagger}_{\max}(K,\beta)$ holds for $K = K_1$. When $K = K_1 + 1$, there are two possibilities:

(i) If $r_{K_1+1}\left(\boldsymbol{\alpha}^{\star}_{K_1+1}(\beta)\right) \ge r_{K_1+1}\left(\boldsymbol{\alpha}^{\dagger}_{K_1+1}(\beta)\right)$, since $r_{K_1+1}(\boldsymbol{\alpha}) = \log_2\left(1 + \frac{\alpha_{K_1+1}}{\sum_{i=1}^{K_1}\alpha_i + \frac{1}{PH_{K_1+1}}}\right) = \log_2\left(1 + \frac{\alpha_{K_1+1}}{\beta - \alpha_{K_1+1} + \frac{1}{PH_{K_1+1}}}\right)$ for any $\boldsymbol{\alpha}$ satisfying $\sum_{i=1}^{K_1+1}\alpha_i = \beta$, it must have

$$\alpha^{\star}_{K_1+1,K_1+1}(\beta) \ge \alpha^{\dagger}_{K_1+1,K_1+1}(\beta),$$

then,

$$\begin{aligned}\beta_1 &= \sum_{k=1}^{K_1} \alpha^{\star}_{k,K_1+1}(\beta) = \beta - \alpha^{\star}_{K_1+1,K_1+1}(\beta) \\ &\leq \beta - \alpha^{\dagger}_{K_1+1,K_1+1}(\beta) = \sum_{k=1}^{K_1} \alpha^{\dagger}_{k,K_1+1}(\beta) = \beta_2.\end{aligned}$$

Next, we obtain

$$\begin{aligned}r^{\star}_{\max}(K_1+1,\beta) &= \min\left\{\left\{\min_{k=1,\ldots,K_1} r_k\left(\boldsymbol{\alpha}^{\star}_{K_1+1}(\beta)\right)\right\}, r_{K_1+1}\left(\boldsymbol{\alpha}^{\star}_{K_1+1}(\beta)\right)\right\} \\ &\leq \min\left\{r^{\star}_{\max}\left(K_1,\beta_1\right), r_{K_1+1}\left(\boldsymbol{\alpha}^{\star}_{K_1+1}(\beta)\right)\right\} &\text{(D.103)} \\ &= \min\left\{r^{\dagger}_{\max}\left(K_1,\beta_1\right), r_{K_1+1}\left(\boldsymbol{\alpha}^{\star}_{K_1+1}(\beta)\right)\right\} &\text{(D.104)} \\ &\leq \min\left\{r^{\dagger}_{\max}\left(K_1,\beta_2\right), r_{K_1+1}\left(\boldsymbol{\alpha}^{\star}_{K_1+1}(\beta)\right)\right\} &\text{(D.105)} \\ &= \min\left\{r^{\dagger}_{\max}\left(K_1+1,\beta\right), r_{K_1+1}\left(\boldsymbol{\alpha}^{\star}_{K_1+1}(\beta)\right)\right\} &\text{(D.106)} \\ &= \min\left\{r_{K_1+1}\left(\boldsymbol{\alpha}^{\dagger}_{K_1+1}(\beta)\right), r_{K_1+1}\left(\boldsymbol{\alpha}^{\star}_{K_1+1}(\beta)\right)\right\} \\ &= r_{K_1+1}\left(\boldsymbol{\alpha}^{\dagger}_{K_1+1}(\beta)\right) = r^{\dagger}_{\max}\left(K_1+1,\beta\right).\end{aligned}$$

Thus, $r^{\star}_{\max}(K_1+1,\beta) \leq r^{\dagger}_{\max}(K_1+1,\beta)$. The inequality (D.103) is due to the optimality of $r^{\star}_{\max}\left(K_1,\beta_1\right)$; (D.104) arises from the assumption that $r^{\star}_{\max}(K,\beta_1) = r^{\dagger}_{\max}(K,\beta_1)$ when $K = K_1$; (D.105) is because $r^{\dagger}_{\max}(K,\beta)$ is non-decreasing on β; (D.106) holds since $r^{\dagger}_{\max}\left(K_1,\beta_2\right) = r^{\dagger}_{\max}\left(K_1+1,\beta\right)$.

(ii) If $r_{K_1+1}\left(\boldsymbol{\alpha}^{\star}_{K_1+1}(\beta)\right) < r_{K_1+1}\left(\boldsymbol{\alpha}^{\dagger}_{K_1+1}(\beta)\right)$, we have $r^{\star}_{\max}(K_1+1,\beta) \leq r_{K_1+1}\left(\boldsymbol{\alpha}^{\star}_{K_1+1}(\beta)\right) < r_{K_1+1}\left(\boldsymbol{\alpha}^{\dagger}_{K_1+1}(\beta)\right) = r^{\dagger}_{\max}(K_1+1,\beta)$, which completes the proof of Lemma 5.5.

□