UNIVERSITY OF CALIFORNIA,
IRVINE

Physical Layer Security with Limited Rate Feedback and Transmitter Cooperation

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Electrical and Computer Engineering

by

Xinjie Yang

Dissertation Committee:
Professor A. Lee Swindlehurst, Chair
Professor Ender Ayanoglu
Professor Ahmed Eltawil

2017

DEDICATION

To
my parents,

Xinxing Liu and Yingjie Yang,

who made all of this possible,
for their endless encouragement and patience.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

I would like to express the deepest appreciation to my advisor, Professor Lee Swindlehurst. He has been very generous with his time and advice over the past years. His thoughtful guidance, insightful vision, and continuing support have lead me to grow immensely. His extensive knowledge is a treasure that I am fortunate to have sampled. I know that what I have learned from him will benefit me eternally.

I am especially thankful to my dissertation committee members, Professor Ender Ayanoglu and Professor Ahmed Eltawil, for their wise insights and precious time in reviewing my dissertation. I also extend gratitude to Professor Syed Jafar and Professor Stanislaw Jarecki for being on my qualifying examination committee.

The best thing about my study at UC Irvine was the opportunity to do research with extremely bright and energetic students and colleagues. I would like to also thank them for their useful comments and discussions.

Many friends have helped me through over the years. Their support and care helped me overcome setbacks. I greatly value their friendship and I deeply appreciate their encouragement.

Most of all, I would like to thank my parents, who always believed in me, and gave me unconditional love, support and understanding at every step of the way. For their love and for everything else, I dedicate this dissertation to them.

CURRICULUM VITAE

Xinjie Yang

EDUCATION

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

Master of Science in Electrical and Computer Engineering **2009**
University of California, Irvine *Irvine, California*

Bachelor of Engineering in Electronic Information Engineering **2007**
Central South University *Changsha, China*

EXPERIENCE

System Engineering Intern **2015–2016**
Broadcom Corporation *Irvine, California*

Firmware Engineering Intern **2011–2012**
Broadcom Corporation *Irvine, California*

Quality Assurance Intern **2010–2011**
Smith Micro Software, Inc. *Aliso Viejo, California*

PUBLICATIONS

X. Yang and A. L. Swindlehurst, "**Limited Rate Feedback for Two-User MISO Interference Channel with and without Secrecy Constraints**," submitted to *IEEE Transactions on Signal Processing*, 2017.

X. Yang and A. L. Swindlehurst, "**Limited Rate Feedback in a MIMO Wiretap Channel With a Cooperative Jammer**," in *IEEE Transactions on Signal Processing*, vol. 64, no. 18, pp. 4695-4706, Sep. 2016.

X. Yang and A. L. Swindlehurst, "**Optimal Bit Allocation of Limited Rate Feedback for Cooperative Jamming**," in *Proc. Signal and Information Processing Assoc. Annual Summit and Conference*, Los Angeles, CA, Dec. 2012.

X. Yang and A. L. Swindlehurst, "**On the Use of Artificial Interference for Secrecy with Imperfect CSI**," in *Proc. IEEE Workshop on Signal Processing Advances in Wireless Communications*, San Francisco, CA, Jun. 2011.

ABSTRACT OF THE DISSERTATION

Physical Layer Security with Limited Rate Feedback and Transmitter Cooperation

By

Xinjie Yang

Doctor of Philosophy in Electrical and Computer Engineering

University of California, Irvine, 2017

Professor A. Lee Swindlehurst, Chair

With the rapid development of wireless communications, security becomes extremely important. In many applications, each transmitter desires to send independent and confidential message to its intended receiver while ensuring mutual information-theoretic secrecy. In this dissertation, I study strategies for enhanced secrecy in wireless communication systems with limited rate feedback and transmitter cooperation. The two-user Gaussian channel model is considered under different scenarios. The transmitters both require channel state information (CSI), which is quantized at the receiver and fed back through the sum-rate-limited feedback channels. The quantization errors reduce the beamforming gain from the direct transmitter, and cause interference leakage from the cross transmitter. In the first scenario, I introduce the wiretap channel model where one receiver is a known eavesdropper, and a second transmitter is used to send a cooperative jamming signal to degrade the eavesdropper's channel. I consider two cases, one where no information about the eavesdroppers is available, and one where statistical CSI is available. With no information about the eavesdroppers, I show how to choose the allocation of feedback bits to the transmitters in order to maximize the amount of jamming power available to interfere with the eavesdroppers, subject to maintaining the lower bound on the rate at a minimum quality-of-service level. For the case of statistical CSI, I derive an approximate lower bound on the average secrecy rate, and optimize the bound to find a suitable bit allocation and the transmit power allocated to

the transmitters. In the second scenario, I consider the interference channel model where the two transmitters are amenable to cooperation for improving the overall secrecy performance of the system. I derive an approximation for the average secrecy rate of each link, and optimize the sum secrecy rate over the transmit power and feedback bits allocated to the transmitters. Interestingly, increasing the transmit power beyond a certain point decreases the secrecy performance. When the transmitters have the same number of antennas, I derive the results in closed form. Simulations validate the theoretical analysis and demonstrate the significant performance gains that result from the use of optimal transmit power control and intelligent feedback bit allocation.

Chapter 1

Introduction

Along with the rapid development of wireless communications, security considerations become paramount. Compared with wired mediums, the broadcast nature of wireless mediums make wireless transmissions extremely convenient, allowing for untethered access to voice, multimedia and data services. However, it also gives rise to a number of security issues. In particular, it is hard to limit access to a wireless medium, as it has no physical boundary. The ease of accessibility makes it easy to eavesdrop on any communication over this medium. Any receivers nearby can hear the transmissions and potentially analyze the transmitted signals. This makes wireless security design a challenging task.

The study of achievable information rates for secure communications is a branch of multi-terminal information theory [1]. As opposed to encryption schemes, provable secrecy can be obtained using a coding approach, where information theoretic analysis provides a proof of secrecy. This form of secrecy is called information theoretic secrecy. It does not require a secure key exchange between the transmitter and the receiver. The secrecy remains intact regardless of the time and computational resources utilized by the eavesdropper in attempting to decode the secret message. In addition, information theoretic schemes can be used in

conjunction with classical symmetric encryption schemes.

Shannon introduced the information theoretic framework for the study of secret communication in [2]. He showed that perfect secrecy is possible if the secret key is at the same or a greater rate than the secret message. However, this result was based on the assumption that the receiver and eavesdropper have access to the same information except the secret key. Later, Wyner developed the concept of the wiretap channel in [3], showing how one could obtain perfect secrecy when the eavesdropper's channel was a degraded version of the receiver's channel. In his model, the transmitter could reliably transmit information to the intended receiver at a positive rate while keeping the eavesdropper in essentially perfect ignorance. Wyner's results for the discrete memoryless wiretap channels were extended to the Gaussian wiretap channel in [4]. Csiszar and Korner generalized [3] by considering the situation where the receiver and the eavesdropper have two different channels with a common input [5]. They showed that secret communication is possible if the eavesdropper's channel is worse than the receiver's channel.

The work in [5] also defined the notion of "secrecy capacity", which essentially is the maximum rate at which the intended receiver's decoding error probability tends to zero, while the eavesdropper's error probability tends to one. The possibility of enhancing secrecy by incorporating common knowledge of the channel impulse response into the data encryption was identified and exploited in [6] and was applied to single antenna mobile radio links in [7] and [8]. A technique for secret communication using channel state information (CSI) as the secret key was described in [8]. In particular, the phase information was used as a secret key and the transmitter compensated for the phase before transmission. The phase of the eavesdropper's channel, being different from that of the receiver's channel, prevented the eavesdropper from decoding the secret message. An abstract characterization of secrecy capacity of the kind discussed by [8] was obtained by [9]. The work in [10] showed an interesting result that perfect secrecy capacity with noisy feedback equals the capacity of the

main channel in the absence of the eavesdropper.

While the above work assumed single antenna nodes, the use of multiple antennas to enhance wireless security via beamforming has recently been a subject of significant interest. Depending on the availability of relevant channel state information (CSI), beamforming techniques can be used to steer information away from eavesdroppers, direct jamming signals directly at them, or fill the spatial modes orthogonal to those of the desired receiver with artificial noise [11–15]. Artificial interference can originate at the information source, or it can be produced by cooperating jammers present in the network [16–23]. In either case, multiple antennas can be used to mitigate the effect of the jamming at the legitimate receiver, provided that accurate CSI is available at both the source of the interference and the receiver.

Assuming perfect CSI at the transmitter is unrealistic in practice due to channel estimation errors in time division duplex (TDD) systems or limited rate feedback and delay in frequency division duplex (FDD) systems. The design and impact of limited rate feedback in FDD systems without secrecy considerations has been studied by a number of researchers; see, for example, [24–28]. The basic approach of these limited rate feedback methods is that, instead of full CSI, only a limited number of feedback bits representing the quantized CSI are fed back to the transmitter from the receiver. Based on the feedback, the transmitter selects the precoding matrix from a pre-designed codebook, and adapts the transmitted signals to the current structure of the channel to achieve an acceptable performance or rate.

There is relatively little work on the effects of limited rate feedback schemes on secrecy at the physical layer. In [29], the authors derive the optimal power allocation between the desired signal and artificial noise to maximize the secrecy rate for a given transmission power and number of feedback bits under quantized channel feedback. Moreover, they derive a scaling law between feedback bits and transmission power to maintain a constant secrecy rate loss compared to the perfect CSI case. The secrecy performance analysis of a codebook-based

beamforming transmission with limited feedback is addressed in [30]. The authors provide an upper bound on the secrecy outage probability as a function of the amount of feedback, and demonstrate that under limited feedback, artificial-noise-aided beamforming does not exhibit any significant advantage over codebook beamforming. The work in [31] considers a multiple-input single-output (MISO) channel scenario with cooperative jamming. The article investigates the impact of quantized channel state information on the secrecy rate, and an adaptive bit allocation strategy is proposed to optimally divide feedback bits between the transmitter and helper channels. In fact, the wiretap channel with a helper can be considered as a special case of the two-user interference channel model. One receiver is a known eavesdropper, and a second transmitter is used to send a jamming signal to enhance secrecy. Secrecy for the two-user interference channel has received considerable attention in recent years, and several approaches have been proposed that discuss how cooperation between the transmitters can be exploited to improve secrecy [32–41].

In this dissertation, I first consider the problem of limited rate feedback design for a wiretap channel with a cooperative jammer, where both the data transmitter and jamming helper require CSI feedback from the receiver. This problem is particularly interesting when the bandwidth available for feedback is limited, and the total number of feedback bits must be properly allocated between the transmitter and helper. The goal is to balance the need to achieve a strong signal from the data transmitter against the need to maximize the impact of the jamming at the eavesdropper and to minimize its impact at the receiver. Unlike [31], I assume that both the legitimate receiver and the eavesdropper possess multiple antennas, and that the transmitter could be sending multiple data streams to the receiver. I assume that both the transmitter and the cooperative jammer employ independent random vector quantization (RVQ) codebooks whose dimensions are to be optimized. Next, I study strategies for transmit power control and feedback bit allocation for the Gaussian two-user MISO interference channel with limited rate feedback. I apply random vector quantization to the CSI of the direct and cross channels, and each receiver sends the indices of the

corresponding codewords to both its own and the interfering transmitter through two sum-rate limited feedback channels. For the given problems, simulations demonstrate a significant gain in performance when the feedback bit and power allocations are chosen according to the proposed algorithms.

The remainder of this dissertation is organized as follows. Chapter 2 introduces the preliminary background. Chapter 3 considers the problem of limited rate feedback design for a wiretap channel with a cooperative jammer. Chapter 4 studies strategies for transmit power control and feedback bit allocation for the Gaussian two-user MISO interference channel with limited rate feedback. Chapter 5 concludes the dissertation.

Throughout the dissertation I use standard lowercase letters to denote scalars, lowercase boldface letters to denote vectors, and uppercase bold letters to denote matrices. The set of n-dimensional complex vectors is denoted by $\mathbb{C}^n$. The space of $m \times n$ complex matrices is denoted by $\mathbb{C}^{m\times n}$. The Hermitian transpose is represented by $(\cdot)^H$, the absolute value $|\cdot|$, the Euclidean norm $\|\cdot\|$, the expectation operator $\mathbb{E}[\cdot]$, the matrix trace $tr(\cdot)$, the zero matrix $\mathbf{0}$ and the $d \times d$ identity matrix $\mathbf{I}_d$. The symbol $\{x\}^+$ denotes $\max\{x, 0\}$. I use $\mathbf{A} \succeq 0$ to denote that $\mathbf{A}$ is positive semidefinite, and write $\mathbf{A} \succeq \mathbf{B}$ if $\mathbf{A} - \mathbf{B} \succeq 0$. If $\mathbf{U}$ is $m \times n$ and satisfies $\mathbf{U}^H\mathbf{U} = \mathbf{I}_n$, then $\mathbf{U}^\perp$ denotes a matrix whose $m - n$ columns are orthogonal vectors that satisfy $\mathbf{U}^H\mathbf{U}^\perp = \mathbf{0}$.

Chapter 2

Preliminary Background

2.1 Basic Wireless Communication System

2.1.1 Fading

Radio waves propagate from a transmit antenna and travel through free space undergoing absorption, reflection, refraction, diffraction, and scattering. They are greatly affected by the ground terrain, the atmosphere, and the objects in their path, such as buildings, bridges, hills, and trees. These multiple physical phenomena are responsible for most of the characteristic features of the received signal.

In most mobile or cellular systems, the height of the mobile antenna may be smaller than the surrounding structures. Thus, the existence of a direct or line-of-sight (LOS) path between the transmitter and the receiver is highly unlikely. In such a case, propagation comes from reflection and scattering from the buildings and diffraction over or around them. Accordingly, the transmitted signal arrives at the receiver via several paths, with different time delays creating a multipath situation, as shown in Figure 2.1.

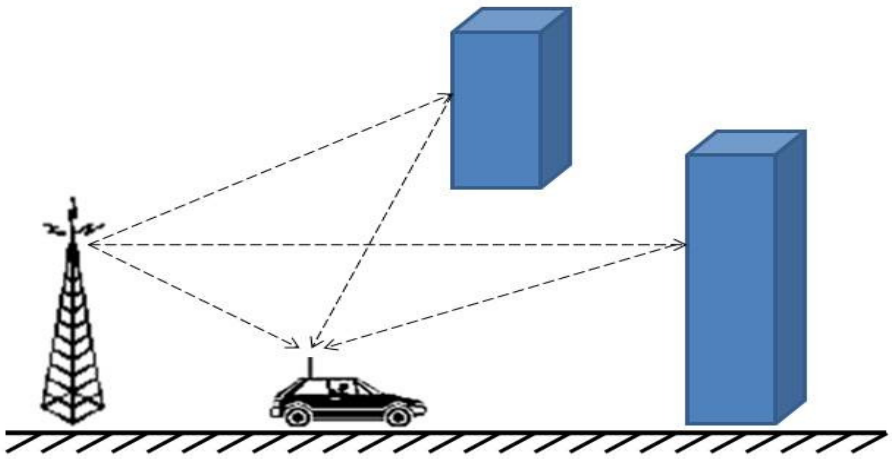

Figure 2.1: Radio propagation in a mobile environment.

At the receiver, these multipath waves with randomly distributed amplitudes and phases combine to give a resultant signal that fluctuates in time and space. Therefore, a receiver at one location may have a signal that differs greatly from the signal at another location only a short distance away because of the change in the phase relationship among the incoming radio waves. This situation causes significant fluctuations in the signal amplitude. This phenomenon of random fluctuations in the received signal level is referred to as fading.

Whereas short-term fluctuations in the signal amplitude caused by local multipath are called small-scale fading, and are observed over distances of about half a wavelength, long-term variations in the mean signal level are called large-scale fading. The latter effect is a result of movement over distances large enough to cause gross variations in the overall path between the transmitter and the receiver. Large-scale fading is also known as shadowing because these variations in the mean signal level are caused by the mobile unit moving into the shadow of surrounding objects, such as buildings and hills. Because of multipath, a moving receiver can experience several fades in a very short duration. In a more serious case, the vehicle may stop at a location where the signal is in a deep fade; in such a situation, maintaining good communication becomes an issue of great concern.

Small-scale fading can be further classified as flat or frequency selective. A received signal is

said to undergo flat fading if the mobile radio channel has a constant gain and the coherent bandwidth larger than the bandwidth of the transmitted signal. Under these conditions, the received signal has amplitude fluctuations as a result of the variations in the channel gain over time caused by multipath. However, the spectral characteristics of the transmitted signal remain intact at the receiver. If the bandwidth of the transmitted signal is larger than the coherence bandwidth of the mobile radio channel, the transmitted signal is said to undergo frequency selective fading. In this case, the received signal is distorted and dispersed because it consists of multiple versions of the transmitted signal, attenuated and delayed in time. The result is time dispersion of the transmitted symbols within the channel arising from these different time delays bringing about intersymbol interference (ISI). The study introduced in this dissertation is confined to the flat fading case.

The mobile antenna, instead of receiving the signal over one LOS path, receives a number of reflected and scattered waves, as shown in Figure 2.1. Because of the varying path lengths, the phases are random and, consequently, the instantaneous received power becomes a random variable. In the case of an unmodulated carrier, the transmitted signal reaches the receiver via a number of paths. If fading is caused by the superposition of a large number of independent scattered components and there is no direct path or LOS component, then the in-phase and quadrature components of the received signal can be assumed to be an independent zero mean Gaussian processes. The probability density function (pdf) of the received signal envelope $f(r)$ can be shown to be Rayleigh distributed, described by

$$f(r) = \frac{r}{\sigma^2} \exp\left\{-\frac{r^2}{2\sigma^2}\right\} , \qquad r \geq 0$$

where $2\sigma^2$ is the average power. In mobile radio channels, the Rayleigh distribution is commonly used to describe the statistical time varying nature of the received envelope of a flat fading signal, or the envelope of an individual multipath component.

2.1.2 Multiple-Antenna Channels

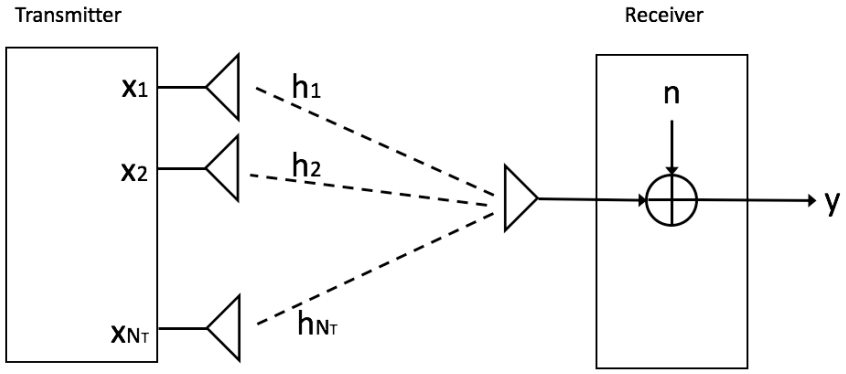

Figure 2.2: MISO channel model.

Consider a multiple input single output (MISO) system with N_T transmit antennas and a single receive antenna in Figure 2.2. Denoting the channel gain between the j-th transmit antenna and the receive antenna by h_j, the MISO channel is represented by an $N_T \times 1$ vector $\mathbf{h}$ given by

$$\mathbf{h} = \begin{bmatrix} h_1 \\ \vdots \\ h_{N_T} \end{bmatrix} .$$

Assume x_j is the signal transmitted from the j-th transmit antenna and y is the received signal. Under the flat fading assumption, the input-output relation for the MISO channel is given by

$$y = \mathbf{h}^H \mathbf{x} + n \ ,$$

where $\mathbf{x} = [x_1 \cdots x_{N_T}]^H$ is an $N_T \times 1$ vector, and n is the corresponding complex Gaussian noise.

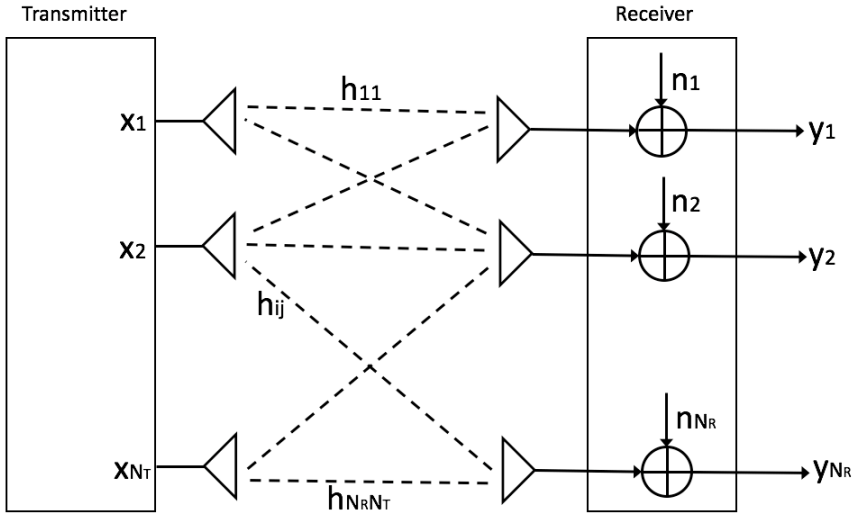

Figure 2.3: MIMO channel model.

A communication system with multiple antennas at both transmitter and receiver is referred to as multiple input multiple output (MIMO) system. Denoting the channel gain between the j-th transmit antenna and the i-th receive antenna by h_{ij}, a MIMO system with N_T transmit and N_R receive antennas is shown in Figure 2.3, and represented by the following model:

$$\begin{bmatrix} y_1 \\ \vdots \\ y_{N_R} \end{bmatrix} = \begin{bmatrix} h_{11} & \cdots & h_{1N_T} \\ \vdots & \ddots & \vdots \\ h_{N_R1} & \cdots & h_{N_RN_T} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_{N_T} \end{bmatrix} + \begin{bmatrix} n_1 \\ \vdots \\ n_{N_R} \end{bmatrix}$$

or simply as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \ .$$

Here $\mathbf{H} \in \mathbb{C}^{N_R \times N_T}$ is a complex Gaussian channel matrix. The components of $\mathbf{n} \in \mathbb{C}^{N_R}$ is the independent and identically distributed (i.i.d.) zero-mean complex Gaussian noise with variance σ_n^2. Denote $\mathbf{Q}_x = E[\mathbf{x}\mathbf{x}^H]$. The transmitter is constrained in its total power to P:

$$tr(\mathbf{Q}_x) = P_x \leq P \ .$$

2.1.3 Parallel Decomposition

Consider a MIMO channel with an $N_R \times N_T$ channel gain matrix $\mathbf{H}$ known to both the transmitter and the receiver. Let R_H denote the rank of $\mathbf{H}$. From matrix theory, for any matrix $\mathbf{H}$ its singular value decomposition (SVD) is

$$\mathbf{H} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^H \ , \tag{2.1}$$

where the $N_R \times N_R$ matrix $\mathbf{U}$ and the $N_T \times N_T$ matrix $\mathbf{V}$ are unitary matrices and $\mathbf{\Lambda}$ is an $N_R \times N_T$ diagonal matrix containing the singular values, denoted by λ_i. These singular values have the property that R_H of these singular values are nonzero. Since R_H cannot exceed the number of columns or rows of $\mathbf{H}$, $R_H \leq \min(N_T, N_R)$. If $\mathbf{H}$ is full rank, as typically occurs in a rich scattering environment, then $R_H = \min(N_T, N_R)$. Other environments may lead to a low rank $\mathbf{H}$.

A parallel decomposition of the channel is obtained by defining a transformation on the channel input $\mathbf{x}$ and output $\mathbf{y}$ through transmit precoding and receiver shaping. In transmit precoding the input to the antennas $\mathbf{x}$ is generated through a linear transformation on input vector $\tilde{\mathbf{x}}$ as $\mathbf{x} = \mathbf{V}\tilde{\mathbf{x}}$. Receiver shaping performs a similar operation at the receiver by multiplying the channel output $\mathbf{y}$ with $\mathbf{U}^H$, as shown in Figure 2.4.

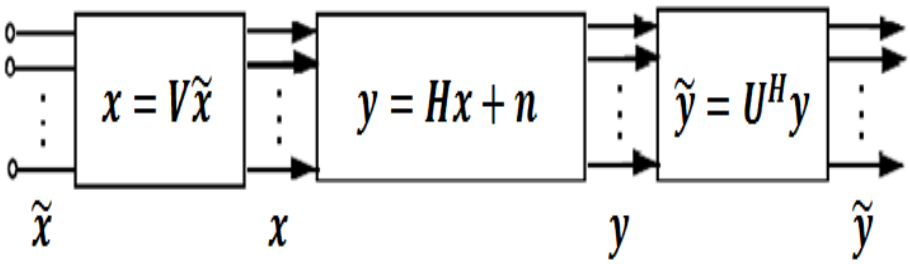

Figure 2.4: Transmit precoding and receiver shaping.

The transmit precoding and receiver shaping transform the MIMO channel into R_H parallel single-input single-output (SISO) channels with input $\tilde{\mathbf{x}}$ and output $\tilde{\mathbf{y}}$. From the singular value decomposition in (2.1),

$$\begin{aligned}\tilde{\mathbf{y}} &= \mathbf{U}^H\mathbf{y} \\ &= \mathbf{U}^H\left(\mathbf{H}\mathbf{x}+\mathbf{n}\right) \\ &= \mathbf{U}^H\left(\mathbf{U}\mathbf{\Lambda}\mathbf{V}^H\mathbf{V}\tilde{\mathbf{x}}+\mathbf{n}\right) \\ &= \mathbf{U}^H\mathbf{U}\mathbf{\Lambda}\mathbf{V}^H\mathbf{V}\tilde{\mathbf{x}}+\mathbf{U}^H\mathbf{n} \\ &= \mathbf{\Lambda}\tilde{\mathbf{x}}+\tilde{\mathbf{n}}\ ,\end{aligned}$$

where $\tilde{\mathbf{n}} = \mathbf{U}^H\mathbf{n}$. Note that multiplication by a unitary matrix does not change the distribution of the noise; *i.e.* $\mathbf{n}$ and $\tilde{\mathbf{n}}$ are identically distributed. Thus, the transmit precoding and receiver shaping transform the MIMO channel into R_H parallel independent channels where the i-th channel has input $\tilde{x}_i$, output $\tilde{y}_i$, noise $\tilde{n}_i$, and channel gain λ_i. The singular values λ_i are related since they are all functions of $\mathbf{H}$, but since the resulting parallel channels do not interfere with each other, the channels with these gains are independent, linked only through the total power constraint. This parallel decomposition is shown in Figure 2.5.

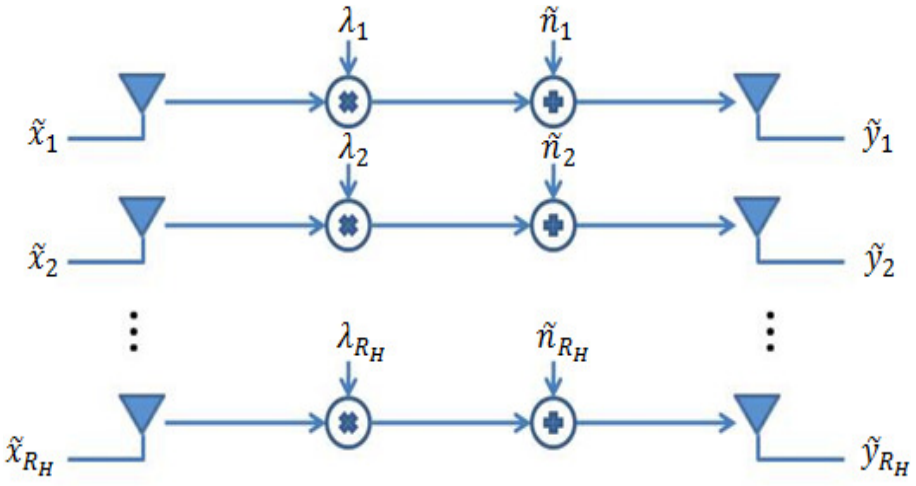

Figure 2.5: Parallel decomposition of the MIMO channel.

2.1.4 MIMO Channel Capacity

The mutual information of the MIMO channel depends on the specific realization of the matrix $\mathbf{H}$, or in particular its singular values λ_i. The average mutual information of a random matrix $\mathbf{H}$, averaged over the matrix distribution, depends on the probability distribution of the singular values of $\mathbf{H}$. In fading channels the transmitter can transmit at a rate equal to this average mutual information and insure correct reception of the data. The transmitter assumes a zero-mean spatially white distribution for $\mathbf{H}$. Ergodic capacity defines the maximum rate, averaged over all channel realizations, that can be transmitted over the channel for a transmission strategy based only on the distribution of $\mathbf{H}$. It leads to a transmitter optimization problem, *i.e.*, finding the optimum input covariance matrix to maximize ergodic capacity subject to the transmit power constraint. Mathematically, the problem is to characterize the optimum $\mathbf{Q}_x$ to maximize

$$\overline{C} = \max_{tr(\mathbf{Q}_x) \leq P} \mathbb{E}_{\mathbf{H}} \left[\log_2 \left| \mathbf{I}_{N_R} + \frac{1}{\sigma_n^2} \mathbf{H} \mathbf{Q}_x \mathbf{H}^H \right| \right] ,$$

where the expectation is with respect to the distribution on the channel matrix $\mathbf{H}$, which for the zero-mean spatially white model is i.i.d. zero-mean circularly symmetric. As in the case of scalar channels, the optimum input covariance matrix that maximizes ergodic capacity for the zero-mean spatially white model is the scaled identity matrix $\mathbf{Q}_x = \frac{P}{N_T} \mathbf{I}_{N_T}$, *i.e.*, the transmit power is divided equally among all the transmit antennas and independent symbols are sent over the different antennas. Thus, the ergodic capacity is given by

$$\overline{C} = \mathbb{E}_{\mathbf{H}} \left[\log_2 \left| \mathbf{I}_{N_R} + \frac{P}{\sigma_n^2 N_T} \mathbf{H} \mathbf{H}^H \right| \right] .$$

In [42] and [43], the authors provide an asymptotic approximation to the average mutual

information of a MIMO channel for a large number of antennas:

$$\mathbb{E}\left[\log_2\left|\mathbf{I}_{N_R}+\frac{P}{\sigma_n^2 N_T}\mathbf{H}\mathbf{H}^H\right|\right]\approx N_R F\left(\frac{N_R}{N_T},\frac{P}{\sigma_n^2}\right)\ , \tag{2.2}$$

where

$$\begin{aligned} F(\beta,\rho) &= \log_2\left(1+\rho\left(\sqrt{\beta}+1\right)^2\right)+(\beta+1)\log_2\left(\frac{1+\sqrt{1-a}}{2}\right) \\ &\quad -(\log_2 e)\sqrt{\beta}\frac{1-\sqrt{1-a}}{1+\sqrt{1-a}}+(\beta-1)\log_2\left(\frac{1+\gamma}{\gamma+\sqrt{1-a}}\right) \\ a &= \frac{4\rho\sqrt{\beta}}{1+\rho\left(\sqrt{\beta}+1\right)^2} \\ \gamma &= \frac{\sqrt{\beta-1}}{\sqrt{\beta+1}}\ . \end{aligned} \tag{2.3}$$

Although (2.2) was originally derived using the central limit theorem under an asymptotic assumption on the number of antennas, the approximation works quite well even for a small number of antennas [43].

2.2 Wiretap Channel

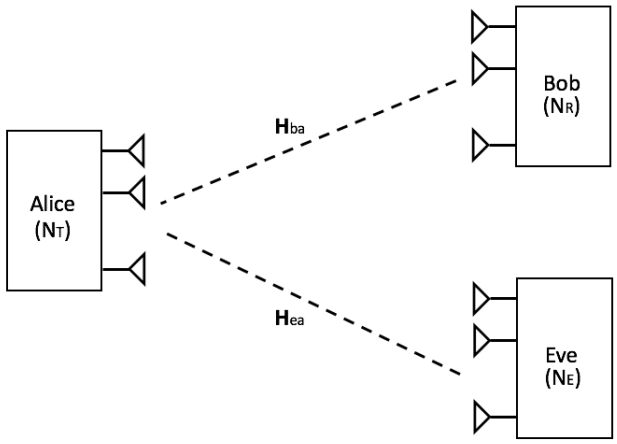

Figure 2.6: MIMO wiretap channel.

Figure 2.6 shows a MIMO wiretap channel model with a transmitter (Alice) with N_T antennas, an intended receiver (Bob) with N_R antennas, and an eavesdropper (Eve) who is an abstraction of multiple colluding eavesdroppers with a total of N_E antennas. $\mathbf{H}_{ba}$ and $\mathbf{H}_{ea}$ denote the channels of the receiver and the eavesdropper respectively. Alice transmits signal $\mathbf{x}$. The signals received by Bob and Eve are

$$\mathbf{y}_b = \mathbf{H}_{ba}\mathbf{x} + \mathbf{n}_b$$
$$\mathbf{y}_e = \mathbf{H}_{ea}\mathbf{x} + \mathbf{n}_e \ ,$$

where the components of $\mathbf{n}_b$ and $\mathbf{n}_e$ are i.i.d. additive white Gaussian noise with variance σ_b^2 and σ_e^2 respectively. For simplicity, $\mathbf{n}_b$ has been normalized so that $\sigma_b^2 = 1$. The elements of $\mathbf{H}_{ba}$ and $\mathbf{H}_{ea}$ are assumed to be i.i.d. and independent of each other, and Bob is able to estimate its channel $\mathbf{H}_{ba}$ perfectly. Block fading is assumed, meaning that the channel gains remain constant long enough so that information theoretic results can be used and that the channel gains in different blocks are independent. The eavesdropper is passive, which means that Eve only listens but does not transmit. Hence, Alice may not know Eve's channel $\mathbf{H}_{ea}$.

The secrecy capacity is considered as the maximum rate at which Alice can transmit, while ensuring that Eve is unable to decode any information. Secrecy capacity is bounded by

$$C_{sec} = \max_{p(\mathbf{x})} \left[I(\mathbf{x}; \mathbf{y}_b) - I(\mathbf{x}; \mathbf{y}_e) \right]^+ \ . \tag{2.4}$$

The significance of (2.4) is that it can be interpreted as the difference in mutual information between the transmitter and receiver and that between the transmitter and eavesdropper. Note that C_{sec} is a random variable because it is a function of $\mathbf{H}_{ba}$ and $\mathbf{H}_{ea}$. The ergodic secrecy capacity of the MIMO wiretap channel is given by

$$\overline{C}_{sec} = \max_{tr(\mathbf{Q}_x) \leq P} \mathbb{E}\left[\log_2 \left| \mathbf{I}_{N_R} + \mathbf{H}_{ba}\mathbf{Q}_x\mathbf{H}_{ba}^H \right| - \log_2 \left| \mathbf{I}_{N_E} + \frac{1}{\sigma_e^2}\mathbf{H}_{ea}\mathbf{Q}_x\mathbf{H}_{ea}^H \right| \right]^+ \ .$$

Assume Alice transmits a d-dimensional data stream $\mathbf{s}$, where $1 \leq d \leq \min\{N_T, N_R\}$, and employs precoder $\mathbf{W} \in \mathbb{C}^{N_T \times d}$, such that $\mathbf{x} = \mathbf{W}\mathbf{s}$. Given that the instantaneous information about Eve's CSI may not be available, a natural approach for the system model is to choose a precoder that provides Bob with a strong signal from Alice. Consequently, the precoder $\mathbf{W}$ is chosen from the first d principle right singular vectors of $\mathbf{H}_{ba}$. Under a relatively high SNR scenario, where Alice uniformly distributes her power across the d signal dimensions, the average achievable secrecy rate is given by

$$\overline{R}_{sec} = \mathbb{E}\left[\log_2\left|\mathbf{I}_{N_R} + \frac{P}{d}\mathbf{H}_{ba}\mathbf{W}\mathbf{W}^H\mathbf{H}_{ba}^H\right| - \log_2\left|\mathbf{I}_{N_E} + \frac{P}{\sigma_e^2 d}\mathbf{H}_{ea}\mathbf{W}\mathbf{W}^H\mathbf{H}_{ea}^H\right|\right]^+ .$$

2.3 Two-User Interference Channel

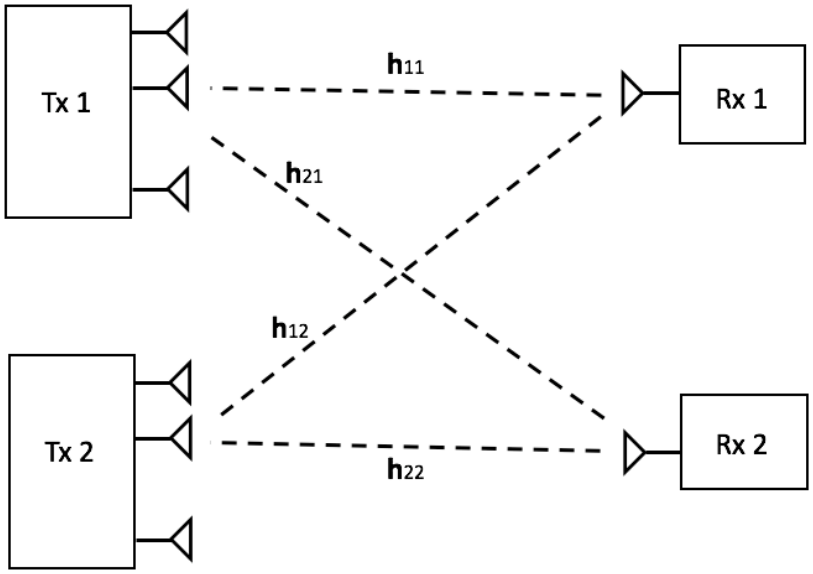

Figure 2.7: Two-user MISO interference channel.

In an interference channel, multiple wireless communication links are simultaneously active in the same time and frequency resource, and hence potentially interfere with each other. I introduce a two-user MISO interference channel model, as shown in Figure 2.7. Assume

transmitter i possesses N_i antennas, while each receiver is equipped with a single antenna. Let $\mathbf{h}_{ii} \in \mathbb{C}^{N_i}$ denote the direct channel from transmitter i to receiver i; and $\mathbf{h}_{ij} \in \mathbb{C}^{N_j}$ denote the cross channel from transmitter j to receiver i ($\forall j \neq i; i, j \in \{1, 2\}$). The elements of these channel vectors are assumed to be independent and identically distributed (i.i.d.), and have zero-mean complex Gaussian distributions with variance σ_{ii}^2 and σ_{ij}^2 respectively. All channels experience independent block fading. Transmitter i employs a unit-norm beamforming vector $\mathbf{w}_i \in \mathbb{C}^{N_i}$ and sends a single data stream s_i to receiver i, which possibly interferes with receiver j. The superposition of the signal received by the i-th receiver is, therefore,

$$y_i = \mathbf{h}_{ii}^H \mathbf{w}_i s_i + \mathbf{h}_{ij}^H \mathbf{w}_j s_j + n_i \ ,$$

where n_i is the corresponding complex Gaussian noise with zero mean and unit variance. Assuming s_i is also Gaussian, define P_i as the actual transmit power used for sending the information signal, which satisfies the power constraint $\mathbb{E}\left[|s_i|^2\right] = P_i \leq P_{max,i}$, where $P_{max,i}$ is the maximum transmit power available at transmitter i.

For the given beamforming vectors, the average achievable transmission rate of the i-th transmitter-receiver link is

$$\overline{R}_i = \mathbb{E}\left[\log_2\left(1 + \frac{P_i \left|\mathbf{h}_{ii}^H \mathbf{w}_i\right|^2}{1 + P_j \left|\mathbf{h}_{ij}^H \mathbf{w}_j\right|^2}\right)\right] \ .$$

From the secrecy point of view, the Gaussian interference channel with confidential messages is also considered, where each transmitter desires to send independent information to its intended receiver while ensuring mutual information-theoretic secrecy. The average achievable secrecy rate of the i-th transmitter-receiver link is defined as in [29, 44]:

$$\overline{R}_{sec,i} = \mathbb{E}\left[\log_2\left(1 + \frac{P_i \left|\mathbf{h}_{ii}^H \mathbf{w}_i\right|^2}{1 + P_j \left|\mathbf{h}_{ij}^H \mathbf{w}_j\right|^2}\right) - \log_2\left(1 + P_i |\mathbf{h}_{ji}^H \mathbf{w}_i|^2\right)\right]^+ \ .$$

In this dissertation, I assume a relatively high SNR scenario such that the zero-forcing (ZF) transmit scheme is sum-rate optimal according to [45]. Ideally, transmitter i chooses a unit-norm beamforming vector $\mathbf{w}_i$, which is orthogonal to $\mathbf{h}_{ji}$ and maximizes $|\mathbf{h}_{ii}^H \mathbf{w}_i|$. This vector is defined as the ZF beamformer and is given by

$$\mathbf{w}_i^{ZF} = \frac{\left(\mathbf{I} - \tilde{\mathbf{h}}_{ji}\tilde{\mathbf{h}}_{ji}^H\right)\mathbf{h}_{ii}}{\left\|\left(\mathbf{I} - \tilde{\mathbf{h}}_{ji}\tilde{\mathbf{h}}_{ji}^H\right)\mathbf{h}_{ii}\right\|},$$

where $\tilde{\mathbf{h}}_{ji}$ is the normalized channel direction

$$\tilde{\mathbf{h}}_{ji} = \frac{\mathbf{h}_{ji}}{\|\mathbf{h}_{ji}\|}$$

and $\left(\mathbf{I} - \tilde{\mathbf{h}}_{ji}\tilde{\mathbf{h}}_{ji}^H\right)$ denotes a projection onto the orthogonal complement of the column space of $\mathbf{h}_{ji}$. The vector $\mathbf{w}_i^{ZF}$ is constructed such that it nulls the interference at receiver j, and the remaining degrees of freedom are used to maximize the transmission rate to receiver i.

2.4 Cooperative Jamming

Early work on secret communication required that the eavesdropper's channel is worse than the receiver's channel. However, in general, there is no guarantee that the receiver will have a better channel than the eavesdropper. For a special case of the two-user interference channel where one receiver is a known eavesdropper, a second transmitter is used to provide artificial interference to degrade the eavesdropper's channel. It is in fact a wiretap channel with a cooperative jammer. The assumed scenario is depicted in Figure 2.8, which features a transmitter (Alice) with N_a antennas, a legitimate receiver (Bob) with N_b antennas, a jamming helper (Hugo) with N_h antennas, and a passive eavesdropper (Eve) with N_e antennas. Eve may be an abstraction of multiple colluding eavesdroppers with a total of N_e

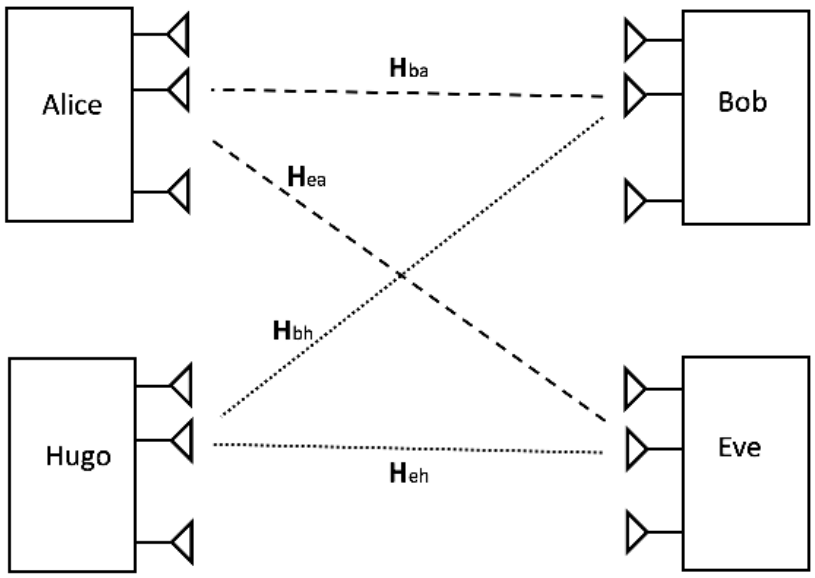

Figure 2.8: Wiretap channel with a cooperative jammer.

antennas, and Hugo is present to provide artificial interference to degrade the channel of any eavesdropper that may be present. The channels from Alice and Hugo to Bob are denoted as $\mathbf{H}_{ba} \in \mathbb{C}^{N_b \times N_a}$ and $\mathbf{H}_{bh} \in \mathbb{C}^{N_b \times N_h}$, and those to Eve are represented as $\mathbf{H}_{ea} \in \mathbb{C}^{N_e \times N_a}$ and $\mathbf{H}_{eh} \in \mathbb{C}^{N_e \times N_h}$. All channels are assumed to experience independent block fading, and the elements of these channel matrices are assumed to be independent and identically distributed (i.i.d.) and have a circularly symmetric complex Gaussian distribution with zero mean and unit variance.

Let $m = \min\{N_a, N_b\}$. Alice transmits a d-dimensional data stream $\mathbf{s}$, where $1 \leq d \leq m$, and that Hugo transmits a p-dimensional jamming signal $\mathbf{v}$. Alice and Hugo employ precoders $\mathbf{W}_a \in \mathbb{C}^{N_a \times d}$ and $\mathbf{W}_h \in \mathbb{C}^{N_h \times p}$, respectively, and Bob uses the beamforming matrix $\mathbf{W}_b \in \mathbb{C}^{N_b \times d}$ to recover the signal of interest. With these assumptions, the signals received

by Bob and Eve are

$$\tilde{\mathbf{y}}_b = \mathbf{W}_b^H \mathbf{y}_b = \mathbf{W}_b^H \mathbf{H}_{ba} \mathbf{W}_a \mathbf{s} + \mathbf{W}_b^H \mathbf{H}_{bh} \mathbf{W}_h \mathbf{v} + \tilde{\mathbf{n}}_b \tag{2.5}$$

$$\mathbf{y}_e = \mathbf{H}_{ea} \mathbf{W}_a \mathbf{s} + \mathbf{H}_{eh} \mathbf{W}_h \mathbf{v} + \mathbf{n}_e \ , \tag{2.6}$$

where $\tilde{\mathbf{n}}_b = \mathbf{W}_b^H \mathbf{n}_b$. The components of $\tilde{\mathbf{n}}_b$ and $\mathbf{n}_e$ are i.i.d. zero-mean complex Gaussian noise with variance σ_b^2 and σ_e^2 respectively. Without loss of generality, $\tilde{\mathbf{n}}_b$ has been normalized so that $\sigma_b^2 = 1$. Define P_s and P_i as the power allocated to the information and jamming signals, which obey the following power constraints:

$$tr(\mathbf{W}_a \mathbf{Q}_s \mathbf{W}_a^H) = P_s \leq P_a$$
$$tr(\mathbf{W}_h \mathbf{Q}_v \mathbf{W}_h^H) = P_i \leq P_h \ ,$$

where

$$\mathbf{Q}_s = \mathbb{E}\left[\mathbf{s}\mathbf{s}^H\right]$$
$$\mathbf{Q}_v = \mathbb{E}\left[\mathbf{v}\mathbf{v}^H\right] \ .$$

Using (2.5) and (2.6), the average achievable secrecy rate is defined as in [29] and [31]:

$$\begin{aligned}\overline{R_{sec}} &= \left[I(S;Y_b|\mathbf{H}_{ba}) - I(S;Y_e|\mathbf{H}_{ba},\mathbf{H}_{ea})\right]^+ \\ &= \mathbb{E}\left[\log_2 \frac{\left|\mathbf{K}_b + \mathbf{W}_b^H \mathbf{H}_{ba} \mathbf{W}_a \mathbf{Q}_s \mathbf{W}_a^H \mathbf{H}_{ba}^H \mathbf{W}_b\right|}{|\mathbf{K}_b|} - \log_2 \frac{\left|\mathbf{K}_e + \mathbf{H}_{ea} \mathbf{W}_a \mathbf{Q}_s \mathbf{W}_a^H \mathbf{H}_{ea}^H\right|}{|\mathbf{K}_e|}\right]^+ ,\end{aligned}$$

where

$$\mathbf{K}_b = \mathbf{W}_b^H \mathbf{W}_b + \mathbf{W}_b^H \mathbf{H}_{bh} \mathbf{W}_h \mathbf{Q}_v \mathbf{W}_h^H \mathbf{H}_{bh}^H \mathbf{W}_b$$
$$\mathbf{K}_e = \sigma_e^2 \mathbf{I}_{N_e} + \mathbf{H}_{eh} \mathbf{W}_h \mathbf{Q}_v \mathbf{W}_h^H \mathbf{H}_{eh}^H \ .$$

Consider the jamming interference term $\mathbf{W}_b^H \mathbf{H}_{bh} \mathbf{W}_h \mathbf{v}$ received by Bob in (2.5). The beamforming strategy can be chosen such that this term is suppressed as much as possible. In the ideal case with perfect CSI at both Hugo and Bob, this interference could be eliminated completely, for example by choosing zero-forcing beamforming. However, Eve's channel is degraded because some components of the jamming interference lie in her signal space. Thus, Bob receives mostly the information signal, while Eve receives both the information signal and the artificial interference. Hence, secrecy is improved.

2.5 Channel Estimation

In communication system, the channel is estimated by the receiver using training signals emitted by the transmitter. The receiver knows the training sequence in advance. In multiple transmit antennas systems, the training signals are mutually orthogonal in some dimension, for instance, time (different time slots) or frequency (different tones) or code (different orthogonal codes). Though orthogonality is not strictly required, orthogonal signals provide the best estimation accuracy for a given transmit power under most circumstances. The training sequence should have good auto and cross correlation properties. Usually the receiver estimates the channel at adequately spaced frequencies or times. The full channel is then determined through interpolation.

Knowledge of the channel at the transmitter is typically used for adapting the modulation rate or for power control. This only needs the gain of the forward channel (downlink). In MISO and MIMO channels, knowledge of the channel can be leveraged in additional ways, such as beamforming or precoding, to provide significant value. Therefore, there is significant motivation for channel knowledge at the transmitter. In multi-user MIMO systems, channel knowledge is necessary to steer signals selectively at users. Depending on the application, differing levels of accuracy in channel information are needed.

Assume a duplex communication scenario. Forward channel estimation at the transmitter is not directly possible as the signal travels through the channel only after leaving the transmitter. Two general techniques are used in channel estimation at the transmitter. In time division duplex (TDD) systems where the reciprocity principle holds (assuming the total slot duration is smaller than the coherence time of the multipath channel), the transmitter first estimates the reverse channel (uplink). This requires the receiver to send a training sequence consisting of pilots. Then, the transmitter performs channel estimation, and uses this estimate to appropriately precode on the forward channel in order to increase the downlink performance. TDD system has the benefit that the uplink channel estimation can be employed also for downlink transmissions, by utilizing the channel reciprocity. This makes it much easier to acquire reliable channel estimates, with little training overhead. However, in the uplink there is a mismatch between the true channel and the channel estimate. By performing the transmit beamforming in the downlink based on the uplink channel estimate, there is also a mismatch with the actual downlink channel due to estimation errors, similarly as in the uplink. Besides this issue, the mismatch between the beamforming and the downlink channel is further increased in the TDD downlink due to the time-varying nature of the channel.

In frequency division duplex (FDD) systems, the channel reciprocity does not hold and there is no connection between the forward channel and the reverse channel. The forward channel information has to be conveyed back to the transmitter from the receiver through a feedback link. First, the receiver obtains an estimation of the forward channel during a downlink training phase. Afterwards, the receiver quantizes the estimated CSI and generates codebooks as a function of the channel combined with the precoder optimization criterion. This codebook is made available at both transmitter and receiver, and only the index of the codebook is sent back over the reverse channel. The transmitter decodes the feedback information and recovers the quantized values corresponding to the forward channel. With a faded and noise-prone uplink, the feedback bits could be received erroneously at the transmitter. This

could lead to further increasing the mismatch in the downlink between the true channel and the available channel knowledge at the transmitter. The mismatch could be additionally increased due to the delay incurred in the feedback process. Thus, the channel knowledge available at the transmitter in the downlink is subject to estimation and quantization errors, and could be outdated and affected by erroneous feedback.

2.6 Limited Rate Feedback

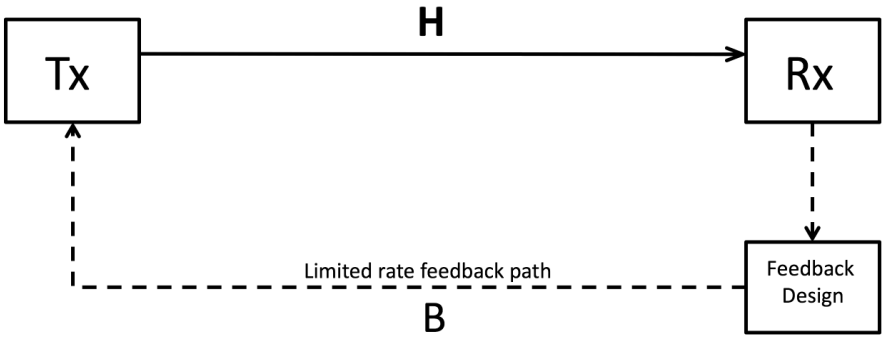

Figure 2.9: Limited rate feedback system.

In the limited rate feedback systems where resources are severely constrained, sending the entire CSI to the transmitter is unrealistic. In practice, allowing the receiver to send a small number of bits about the forward channel conditions to the transmitter can allow near optimal channel adaptation. A block diagram of a limited rate feedback system is depicted in Figure 2.9. For a limited rate feedback approach, this dissertation concentrates on the scenario where the receiver is assumed to have perfect estimation of the CSI, and the feedback channel is zero-delay and error-free. Employing limited feedback requires cooperation between the transmitter and receiver. Given B bits of feedback, the receiver uses intelligent vector quantization techniques to quantize the estimate of the forward channel to one of 2^B integer indices, with each index corresponding to a particular mode of the channel. The transmitter has knowledge of the 2^B-mode codebook, and therefore, it is able to optimize

its transmission strategy based on the feedback information to adapt the transmitted signal to the forward channel. Thus, it is a challenging problem to design optimal quantization schemes and the associated transmission strategies for multiple-antenna systems with limited rate feedback.

One approach to limited feedback is to employ channel quantization. In MISO systems, the channel vector $\mathbf{h} \in \mathbb{C}^{N_T}$ is quantized using a vector quantization (VQ) algorithm. A vector quantizer works by mapping a complex valued vector into one of a finite number of vector realizations. The mapping is usually designed to minimize some sort of distortion function between the input vector and the quantized vector. The codebook can be either fixed or randomly generated. Designing a fixed codebook is a challenging problem that depends on the distribution of the channel. Grassmannian and VQ approaches using the Lloyd algorithm have been shown to efficiently generate codebooks that specifically optimize for both the statistical distribution of the channel as well as the specific performance metric. However, random vector quantization (RVQ) codebooks are more common for most limited rate feedback problems because the optimal vector quantizer is not known in general, and known bounds are rather loose. RVQ is very amenable to analysis and also performs measurably close to optimal quantization. Similar to the standard random coding argument used for channel coding, there always exists at least one quantization codebook that performs at least as well as the ensemble average. In addition, utilizing the statistics of random quantization codebooks, this method simulates the quantization procedure without generating an actual codebook, and reduces the computational complexity as B grows. Assume a RVQ codebook $\mathcal{C}$ consists of 2^B normalized quantization vectors, i.e. $\mathcal{C} = (\mathbf{c}_1, \ . \ . \ . \ , \mathbf{c}_{2^B})$, which are independently chosen from the isotropic distribution on the unit sphere in $\mathbb{C}^{N_T}$, as the normalized channel direction $\tilde{\mathbf{h}} \triangleq \frac{\mathbf{h}}{\|\mathbf{h}\|}$ due to the assumption of i.i.d. Rayleigh fading. Denote $\hat{\mathbf{h}}$ as the quantized version of $\mathbf{h}$, which is chosen from the codebook $\mathcal{C}$ according to the

following rule:

$$\hat{\mathbf{h}} = \arg\max_{\mathbf{c}\in\mathcal{C}} |\tilde{\mathbf{h}}^H \mathbf{c}|^2 \ .$$

The most important quantity of interest is the statistical distribution of the quantization error. According to [46], the average distortion D associated with the given codebook is bounded above by

$$D \triangleq \mathbb{E}\left[1 - |\tilde{\mathbf{h}}^H \hat{\mathbf{h}}|^2\right] \leq \Gamma(\frac{N_T}{N_T - 1}) 2^{-\frac{B}{N_T - 1}} \ ,$$

where $\Gamma(\cdot)$ represents the gamma function. Let the transmit signal be $\mathbf{x} = \mathbf{w}s$, where $\mathbf{w}$ is a unit-norm beamforming vector and s is a single-dimensional complex symbol chosen independently of the channel. The transmitter can then design a beamforming vector $\mathbf{w}$ and the transmit signal power according to the feedback CSI that maximize a curtain optimization criterion.

In MIMO systems, similarly, a random matrix quantization codebook $\mathcal{C}$ consists of 2^B matrices in $\mathbb{C}^{N_R \times N_T}$, i.e. $\mathcal{C} = (\mathbf{C}_1, \ldots, \mathbf{C}_{2^B})$. Each of the 2^B matrices making up the codebook is chosen independently and is isotropically distributed over the $N_R \times N_T$ Grassmann manifold. Let us assume $N_T \geq N_R$. Denote $\widehat{\mathbf{H}}$ as the quantized version of $\mathbf{H}$, which is chosen from the codebook $\mathcal{C}$ according to the following rule:

$$\widehat{\mathbf{H}} = \arg\min_{\mathbf{C}\in\mathcal{C}} d^2\left(\mathbf{H}, \mathbf{C}\right) \ ,$$

where $d\left(\mathbf{H}, \mathbf{C}\right)$ is the chordal distance between $\mathbf{H}$ and $\mathbf{C}$, and is given by [46]

$$d^2\left(\mathbf{H}, \mathbf{C}\right) = \sqrt{\sum_{j=1}^{N_R} \sin^2\theta_j} \ ,$$

where θ_j is the principal angle between the two subspaces spanned by the columns of the

matrices $\mathbf{H}$ and $\mathbf{C}$. As the principal angles depend only on the subspaces spanned by the columns of the matrices, it can be assumed that the elements of $\mathcal{C}$ are semi-unitary matrices. An alternate form for the chordal distance is

$$d^2\left(\mathbf{H},\mathbf{C}\right) = N_R - tr\left(\tilde{\mathbf{H}}^H\mathbf{C}\mathbf{C}^H\tilde{\mathbf{H}}\right) \ ,$$

where $\tilde{\mathbf{H}}$ forms an orthonormal basis for the subspace spanned by $\mathbf{H}$. The distortion or error associated with a given codebook $\mathcal{C}$ for the quantization of $\mathbf{H}$ is defined as:

$$D \triangleq \mathbb{E}\left[d^2\left(\mathbf{H},\widehat{\mathbf{H}}\right)\right] = \mathbb{E}\left[\min_{\mathbf{C}\in\mathcal{C}} \ d^2\left(\mathbf{H},\mathbf{C}\right)\right] \ .$$

It is shown in [46] that the average distortion D associated with the given codebooks $\mathcal{C}$ is bounded above by

$$D \leq \frac{\Gamma(\frac{1}{T})}{T}\Phi^{-\frac{1}{T}}2^{-\frac{B}{T}} + N_R e^{-\left(2^B\Phi\right)^{1-a}} \ , \tag{2.7}$$

for a codebook of size 2^B. Here, $T = N_R(N_T - N_R)$, $\Phi = \frac{1}{\Gamma(T+1)}\prod_{k=1}^{N_R}\frac{\Gamma(N_T-k+1)}{\Gamma(N_R-k+1)}$ and $\Gamma(\cdot)$ represents the gamma function. $a \in (0,1)$ is a real number between 0 and 1 chosen such that $(2^B\Phi)^{-\frac{a}{T}} \leq 1$. The second (exponential) term in (2.7) can be neglected for large B. Thus,

$$D \leq \frac{\Gamma(\frac{1}{T})}{T}\left[\frac{1}{\Gamma(T+1)}\prod_{k=1}^{N_R}\frac{\Gamma(N_T-k+1)}{\Gamma(N_R-k+1)}\right]^{-\frac{1}{T}} 2^{-\frac{B}{T}} \ .$$

Sending a quantized version of the forward link channel from receiver to transmitter gives the transmitter more flexibility to choose among different space-time signaling techniques. However, there can be further increases in performance by focusing on improving the quantized information needed to adapt the transmitted signal to current channel conditions. Notice that, instead of letting the transmitter design precoding matrix on the basis of the received CSI feedback, the receiver can directly design the precoder and send this designed matrix

back to the transmitter. It leads to significant advances in feedback techniques because the receiver will nearly always have higher quality CSI than the transmitter. By restricting $\mathbf{W}$ to lie in a codebook consisting of 2^B possible precoding matrices, the receiver can use its channel knowledge to pick the appropriate precoder from this codebook and control how the signal is adapted to the channel. This is also advantageous in a scenario where the feedback bandwidth is limited, since the dimensionality of the quantization problem is reduced compared with feedback of the quantized channel matrices.

Note that the ideal precoder $\mathbf{W} \in \mathbb{C}^{N_T \times d}$ can be uniquely defined by its orthogonal complement $\mathbf{W}^{\perp} \in \mathbb{C}^{N_T \times (N_T - d)}$. Thus, the feedback from the receiver must be either the precoder or its orthogonal complement, whichever is of smaller dimension and requires fewer bits to encode. Once d is determined, the feedback matrix is drawn from the random quantization codebook $\mathcal{C}$, which is known to the transmitter beforehand. The choice of codebook used depends on whether the precoder or its orthogonal complement is fed back. The transmitter uses the index fed back by the receiver to determine which codebook is to be used for the precoder.

Chapter 3

MIMO Wiretap Channel with a Cooperative Jammer

In this chapter, I study strategies for enhanced secrecy using cooperative jamming in secure communication systems with limited rate feedback. A Gaussian multiple-input multiple-output (MIMO) wiretap channel with a jamming helper is considered. The transmitter and helper both require channel state information (CSI), which is quantized at the receiver and fed back through two sum-rate-limited feedback channels. The quantization errors result in reduced beamforming gain from the transmitter, as well as interference leakage from the helper. First, under the assumption that the eavesdropper's CSI is completely unknown, I derive a lower bound on the average main channel rate and find the feedback bit allocation that maximizes the jamming power under a constraint on the bound. For the case where statistical CSI for the eavesdropper's channel is available, I derive a lower bound on the average secrecy rate, and I optimize the bound to find a suitable bit allocation and the transmit powers allocated to the transmitter and helper. For the case where the transmitter and helper have the same number of antennas, I obtain a closed-form solution for the optimal bit allocation. Simulations verify the theoretical analysis and demonstrate the significant

performance gain that results with intelligent feedback bit allocation and power control.

3.1 Introduction

The use of multiple antennas to enhance wireless security via beamforming has recently been a subject of significant interest. Depending on the availability of relevant channel state information (CSI), beamforming techniques can be used to steer information away from eavesdroppers, direct jamming signals directly at them, or fill the spatial modes orthogonal to those of the desired receiver with artificial noise [11–15]. Artificial interference can originate at the information source, or it can be produced by cooperating jammers present in the network [16–23]. In either case, multiple antennas can be used to mitigate the effect of the jamming at the legitimate receiver, provided that accurate CSI is available at both the source of the interference and the receiver.

Assuming perfect CSI at the transmitter is unrealistic in practice due to channel estimation errors in time division duplex (TDD) systems or limited rate feedback and delay in frequency division duplex (FDD) systems. The design and impact of limited rate feedback in FDD systems without secrecy considerations has been studied by a number of researchers; see, for example, [24–28]. The basic approach of these limited rate feedback methods is that, instead of full CSI, only a limited number of feedback bits representing the quantized CSI are fed back to the transmitter from the receiver. Based on the feedback, the transmitter selects the precoding matrix from a pre-designed codebook, and adapts the transmitted signals to the current eigenstructure of the channel to achieve a certain acceptable performance loss.

There has been relatively little work on the effects of limited rate feedback schemes on secrecy at the physical layer. In [29], the authors derive the optimal power allocation between the desired signal and artificial noise to maximize the secrecy rate for a given transmission power

and number of feedback bits under quantized channel feedback. Moreover, they derive a scaling law between feedback bits and transmission power to maintain a constant secrecy rate loss compared to the perfect CSI case. The secrecy performance analysis of a codebook-based beamforming transmission with limited feedback is addressed in [30]. The authors provide an upper bound on the secrecy outage probability as a function of the amount of feedback, and demonstrate that under limited feedback, artificial-noise-aided beamforming does not exhibit any significant advantage over codebook beamforming. In [31], a multiple-input single-output (MISO) channel scenario with cooperative jamming is considered. The article investigates the impact of quantized channel state information on the secrecy rate, and an adaptive bit allocation strategy is proposed to optimally divide feedback bits between the transmitter and helper channels.

In this chapter, I consider the problem of limited rate feedback design for a wiretap channel with a cooperative jammer, where both the data transmitter and jamming helper require CSI feedback from the receiver. This problem is particularly interesting when the bandwidth available for feedback is limited, and the total number of feedback bits must be properly allocated between the transmitter and helper. The goal is to balance the need to achieve a strong signal from the data transmitter against the need to maximize the impact of the jamming at the eavesdropper and to minimize its impact at the receiver. Unlike [31], I assume that both the legitimate receiver and the eavesdropper possess multiple antennas, and that the transmitter could be sending multiple data streams to the receiver. I assume that both the transmitter and the cooperative jammer employ independent random vector quantization (RVQ) codebooks whose dimensions are to be optimized.

I consider two cases with respect to the eavesdropper CSI. In the first, no CSI is available for the eavesdropper, and in the second, only statistical CSI is available. For the first case, I derive a lower bound on the achievable rate for the primary channel, and following the general approach of [12], I find the feedback bit allocation that allows for maximum jamming from

the helper while still maintaining the lower bound on the rate at a minimum quality-of-service level. No secrecy guarantee is possible since the eavesdropper's CSI is completely unknown, but this approach maximizes the amount of interference available to mask the desired signal from whatever eavesdroppers are present. Next, I consider the second case, in which the statistical CSI for the eavesdropper is available. I derive a closed-form expression for the optimal bit allocation for the transmitter and jammer when they have an equal number of antennas. I then derive a lower bound for the average secrecy rate and optimize it over the power allocated to the transmitter and jammer. For both sets of CSI assumptions, simulations demonstrate a significant gain in performance when the feedback bit and power allocations are chosen according to the proposed algorithms.

The remainder of this chapter is organized as follows. In Section 3.2, I introduce the system model and the necessary assumptions. Section 3.3 provides preliminary background information on beamforming strategies, RVQ codebooks and the calculation of average secrecy rate. In Section 3.4, I analyze the impact of limited rate feedback and discuss the bit allocation and power control algorithms for the two eavesdropper CSI scenarios discussed above. In Section 3.5, I present simulation results to validate my algorithms.

3.2 Signal Modeling Assumptions

I depict the assumed scenario in Figure 3.1, which features a transmitter (Alice) with N_a antennas, a legitimate receiver (Bob) with N_b antennas, a jamming helper (Hugo) with N_h antennas, and a passive eavesdropper (Eve) with N_e antennas. Eve may be an abstraction of multiple colluding eavesdroppers with a total of N_e antennas, and Hugo is present to provide artificial interference to degrade the channel of any eavesdropper that may be present. The channels from Alice and Hugo to Bob are denoted as $\mathbf{H}_{ba} \in \mathbb{C}^{N_b \times N_a}$ and $\mathbf{H}_{bh} \in \mathbb{C}^{N_b \times N_h}$, and those to Eve are represented as $\mathbf{H}_{ea} \in \mathbb{C}^{N_e \times N_a}$ and $\mathbf{H}_{eh} \in \mathbb{C}^{N_e \times N_h}$. All channels experience

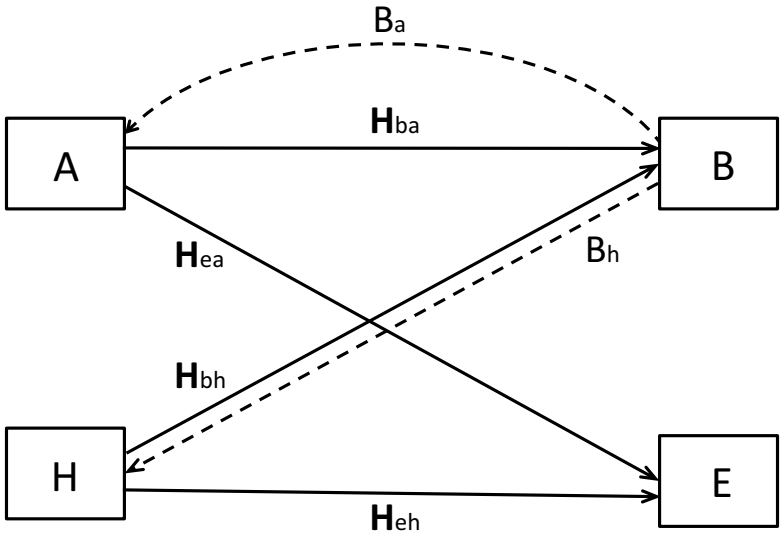

Figure 3.1: System model of a MIMO wiretap channel with a cooperative jammer.

independent block fading, and the elements of these channel matrices are independent and identically distributed (i.i.d.) and have a circularly symmetric complex Gaussian distribution with zero mean and unit variance. Bob has perfect CSI for $\mathbf{H}_{bh}$ and $\mathbf{H}_{ba}$. Alice and Hugo do not know these channels, but can obtain quantized information from Bob through two error-free, zero-delay feedback channels. Bob quantizes the feedback information by selecting the closest codewords from two codebooks respectively containing 2^{B_a} and 2^{B_h} entries. The B_a and B_h index bits corresponding to the chosen codewords are separately fed back to Alice and Hugo. With the limited total rate constraint on the feedback channel, Bob is only able to feed back a fixed total number of bits B, such that

$$B_a + B_h = B \ . \tag{3.1}$$

Let $l = \max\{N_a, N_b\}$ and $m = \min\{N_a, N_b\}$. Alice transmits a d-dimensional data stream $\mathbf{s}$, where $1 \leq d \leq m$, and that Hugo transmits a p-dimensional jamming signal $\mathbf{v}$. Alice and

Hugo employ precoders $\mathbf{W}_a \in \mathbb{C}^{N_a \times d}$ and $\mathbf{W}_h \in \mathbb{C}^{N_h \times p}$, respectively, and Bob uses the beamforming matrix $\mathbf{W}_b \in \mathbb{C}^{N_b \times d}$ to recover the signal of interest. With these assumptions, the signals received by Bob and Eve are:

$$\tilde{\mathbf{y}}_b = \mathbf{W}_b^H \mathbf{y}_b = \mathbf{W}_b^H \mathbf{H}_{ba} \mathbf{W}_a \mathbf{s} + \mathbf{W}_b^H \mathbf{H}_{bh} \mathbf{W}_h \mathbf{v} + \tilde{\mathbf{n}}_b \tag{3.2}$$

$$\mathbf{y}_e = \mathbf{H}_{ea} \mathbf{W}_a \mathbf{s} + \mathbf{H}_{eh} \mathbf{W}_h \mathbf{v} + \mathbf{n}_e \ , \tag{3.3}$$

where $\tilde{\mathbf{n}}_b = \mathbf{W}_b^H \mathbf{n}_b$. As discussed below, the beamforming matrix $\mathbf{W}_b$ is chosen such that $\mathbf{W}_b^H \mathbf{W}_b = \mathbf{I}_d$. The components of $\tilde{\mathbf{n}}_b$ and $\mathbf{n}_e$ are i.i.d. zero-mean complex Gaussian noise with variance σ_b^2 and σ_e^2 respectively. Without loss of generality, $\tilde{\mathbf{n}}_b$ has been normalized so that $\sigma_b^2 = 1$. Define P_s and P_i as the power allocated to the information and jamming signals, which obey the following power constraints:

$$tr(\mathbf{W}_a \mathbf{Q}_s \mathbf{W}_a^H) = P_s \leq P_a \tag{3.4}$$

$$tr(\mathbf{W}_h \mathbf{Q}_v \mathbf{W}_h^H) = P_i \leq P_h \ , \tag{3.5}$$

where

$$\mathbf{Q}_s = \mathbb{E}\left[\mathbf{s}\mathbf{s}^H\right]$$

$$\mathbf{Q}_v = \mathbb{E}\left[\mathbf{v}\mathbf{v}^H\right] \ .$$

3.3 Algorithms and Metrics

3.3.1 Beamforming Design

Based on his knowledge of $\mathbf{H}_{ba}$ and $\mathbf{H}_{bh}$, Bob determines appropriate precoders $\mathbf{W}_a$ and $\mathbf{W}_h$ and feeds them back to Alice and Hugo, respectively. This is advantageous in a scenario where the feedback bandwidth is limited, since the dimensionality of the quantization problem is reduced compared with feedback of the quantized channel matrices [24].

Consider the jamming interference term $\mathbf{W}_b^H \mathbf{H}_{bh} \mathbf{W}_h \mathbf{v}$ received by Bob in (3.2). In the ideal case with perfect CSI at both Hugo and Bob, this interference could be eliminated completely, for example by choosing zero-forcing beamforming. Even in the limited feedback scenario, if $d + p \leq N_b$, Bob has a sufficient number of antennas to cancel the interference. Note that Bob knows the true channel $\mathbf{H}_{bh}$ as well as the quantized precoder $\mathbf{W}_h$, since Bob fed this information back to Hugo. In this case, the beamformer $\mathbf{W}_b$ can be chosen to be orthogonal to $\mathbf{H}_{bh}\mathbf{W}_h$, and thus would completely eliminate the contribution from Hugo; consequently, Hugo could transmit with full power and have no impact on Bob, except to limit the number of data streams available for spatial multiplexing. A bit allocation strategy is unnecessary in this case. Specifically, given an arbitrary fixed precoder $\mathbf{W}_h$, Bob could always design a zero-forcing receive beamformer $\mathbf{W}_b$ that would completely eliminate the interference from Hugo. However, if $N_b < d + p \leq N_h$, Bob has insufficient degrees of freedom for canceling the interference. Therefore, decoupling the links requires $\mathbf{W}_h$ to be chosen orthogonal to $\mathbf{W}_b^H \mathbf{H}_{bh}$. In general, this requires that $N_h > d$, and hence that p, the dimension of the jamming signal, is set to $p = N_h - d$. Since Hugo uses a quantized precoder fed back from Bob, $\mathbf{W}_h$ is no longer orthogonal to $\mathbf{W}_b^H \mathbf{H}_{bh}$ and causes interference leakage that Bob is not able to filter out. I focus on this case throughout the chapter.

Given that the instantaneous information about Eve's CSI is unavailable, a natural approach

for the system model is to choose a precoder that provides Bob with a strong signal from Alice, and a receive beamformer that is focused on the information signal. Consequently, the beamformers $\mathbf{W}_a$ and $\mathbf{W}_b$ are chosen from the principle right and left singular vectors of $\mathbf{H}_{ba}$. Without access to Eve's instantaneous CSI, a reasonable approach is for Hugo to spread the jamming power uniformly across the $N_h - d$ jamming dimensions. Assume a relatively high SNR scenario where errors due to the limited feedback dominate, and hence Alice also uniformly distributes her power across the d signal dimensions.

More specifically, define the singular value decomposition of $\mathbf{H}_{ba}$ and $\mathbf{W}_b^H \mathbf{H}_{bh}$ as follows:

$$\begin{aligned} svd(\mathbf{H}_{ba}) &= \mathbf{U}_a \mathbf{\Lambda}_a \mathbf{V}_a^H \\ &= [\mathbf{U}_{a1}\ \mathbf{U}_{a2}] \begin{bmatrix} \mathbf{\Lambda}_{a1} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_{a2} \end{bmatrix} [\mathbf{V}_{a1}\ \mathbf{V}_{a2}]^H \\ svd(\mathbf{W}_b^H \mathbf{H}_{bh}) &= \mathbf{U}_h \mathbf{\Lambda}_h \mathbf{V}_h^H \\ &= \mathbf{U}_h \left[\mathbf{\Lambda}_{h1}\ \mathbf{0}\right] [\mathbf{V}_{h1}\ \mathbf{V}_{h2}]^H \ , \end{aligned}$$

where $\mathbf{U}_{a1}$, $\mathbf{V}_{a1}$, $\mathbf{\Lambda}_{h1}$ and $\mathbf{V}_{h1}$ denote the first d columns of $\mathbf{U}_a$, $\mathbf{V}_a$, $\mathbf{\Lambda}_h$ and $\mathbf{V}_h$; $\mathbf{U}_{a2}$, $\mathbf{V}_{a2}$ and $\mathbf{V}_{h2}$ contain the remaining columns of $\mathbf{U}_a$, $\mathbf{V}_a$ and $\mathbf{V}_h$; $\mathbf{\Lambda}_{a1}$ denotes the upper left $d \times d$ diagonal submatrix of $\mathbf{\Lambda}_a$, and $\mathbf{\Lambda}_{a2}$ is the lower right diagonal submatrix. In the ideal case without quantization errors, $\mathbf{W}_a = \mathbf{V}_{a1}$ and $\mathbf{W}_b = \mathbf{U}_{a1}$. Furthermore, the zero-forcing precoding matrix for Hugo is $\mathbf{W}_h = \mathbf{V}_{h2}$. Note that the ideal precoders can be uniquely defined by their orthogonal complements, since $\mathbf{V}_{a1} = \mathbf{V}_{a2}^{\perp}$ and $\mathbf{V}_{h2} = \mathbf{V}_{h1}^{\perp}$. Thus, the feedback from Bob must be either the precoder or its orthogonal complement, whichever is

of smaller dimension and requires fewer bits to encode:

$$\mathbf{W}_a = \begin{cases} \widehat{\mathbf{V}}_{a1} & \text{if } d \leq N_a - d \\ \widehat{\mathbf{V}}_{a2}^{\perp} & \text{if } d > N_a - d \end{cases} \tag{3.6}$$

$$\mathbf{W}_h = \begin{cases} \widehat{\mathbf{V}}_{h1}^{\perp} & \text{if } d < N_h - d \\ \widehat{\mathbf{V}}_{h2} & \text{if } d \geq N_h - d \,, \end{cases} \tag{3.7}$$

where the hat indicates a quantized version of the matrix. The size of the codewords for $\mathbf{W}_a$ and $\mathbf{W}_h$ are $N_a \times M_a$ and $N_h \times M_h$, respectively, where $M_a = \min\{d, N_a - d\}$ and $M_h = \min\{d, N_h - d\}$. Once d is determined, Alice and Hugo use the indices fed back by Bob to determine which codebooks are to be used for the precoders.

3.3.2 Random Quantization Codebooks

The precoders $\mathbf{W}_a$ and $\mathbf{W}_h$ are drawn from the random quantization codebooks $\mathcal{C}_a$ and $\mathcal{C}_h$, respectively, which are known to Alice and Hugo beforehand. Assuming the codebooks are indexed by B_a and B_h feedback bits, respectively, the codebooks for Alice and Hugo contain 2^{B_a} and 2^{B_h} entries. Each of the codewords $\mathbf{C}_a \in \mathcal{C}_a$ and $\mathbf{C}_h \in \mathcal{C}_h$ are generated independently and isotropically over the $N_a \times M_a$ and $N_h \times M_h$ Grassmann manifold, and are assumed to be semi-unitary.

The choice of codebook used depends on whether the precoder or its orthogonal complement

is fed back; in particular, I use the notation

$$\mathcal{C}_a \triangleq \begin{cases} \mathcal{C}_{a1} & \text{if } d \leq N_a - d \\ \mathcal{C}_{a2} & \text{if } d > N_a - d \end{cases}$$

$$\mathcal{C}_h \triangleq \begin{cases} \mathcal{C}_{h1} & \text{if } d < N_h - d \\ \mathcal{C}_{h2} & \text{if } d \geq N_h - d \, . \end{cases}$$

The precoder assignments in (3.6) and (3.7) satisfy the following rule:

$$\widehat{\mathbf{V}}_{ij} = \arg \min_{\mathbf{C}_i \in \mathcal{C}_{ij}} d^2 \left(\mathbf{V}_{ij}, \mathbf{C}_i \right) \quad i = a, h; \; j = 1, 2 \, ,$$

where $d \left(\mathbf{V}_{ij}, \mathbf{C}_i \right)$ is the chordal distance between $\mathbf{V}_{ij}$ and $\mathbf{C}_i$, and is given by [46]

$$d^2 \left(\mathbf{V}_{ij}, \mathbf{C}_i \right) = M_i - tr \left(\mathbf{V}_{ij}^H \mathbf{C}_i \mathbf{C}_i^H \mathbf{V}_{ij} \right) \quad i = a, h; \; j = 1, 2 \, . \tag{3.8}$$

It is shown in [46] that the average distortion D_i associated with the given codebooks $\mathcal{C}_i$ is bounded above by

$$D_i \triangleq \mathbb{E} \left[d^2 \left(\mathbf{V}_{ij}, \widehat{\mathbf{V}}_{ij} \right) \right] \leq G_i 2^{-\frac{B_i}{T_i}} \quad i = a, h; \; j = 1, 2 \, , \tag{3.9}$$

where $T_i = d(N_i - d)$ and

$$G_i = \frac{\Gamma(\frac{1}{T_i})}{T_i} \left[\frac{1}{\Gamma(T_i + 1)} \prod_{k=1}^{M_i} \frac{\Gamma(N_i - k + 1)}{\Gamma(M_i - k + 1)} \right]^{-\frac{1}{T_i}} \, ,$$

and $\Gamma(\cdot)$ represents the gamma function.

Since the quantized precoding matrices $\mathbf{W}_a$ and $\mathbf{W}_h$ are both semi-unitary, the power con-

straints in (3.4) and (3.5) become

$$tr(\mathbf{Q}_s) = P_s \leq P_a \tag{3.10}$$

$$tr(\mathbf{Q}_v) = P_i \leq P_h \ . \tag{3.11}$$

3.3.3 Average Secrecy Rate

For the case where the statistics of the eavesdropper channel are available, the metric of interest is the average achievable secrecy rate of the system. Assume $\mathbf{s}$ is Gaussian. Using (3.2) and (3.3), the average achievable secrecy rate is defined as in [29] and [31][1]:

$$\begin{aligned}\overline{R_{sec}} &= [I(S;Y_b|\mathbf{H}_{ba}) - I(S;Y_e|\mathbf{H}_{ba},\mathbf{H}_{ea})]^+ \\ &= \left\{ \mathbb{E}\left[\log_2 \frac{|\mathbf{K}_b + \mathbf{W}_b^H \mathbf{H}_{ba} \mathbf{W}_a \mathbf{Q}_s \mathbf{W}_a^H \mathbf{H}_{ba}^H \mathbf{W}_b|}{|\mathbf{K}_b|}\right] - \mathbb{E}\left[\log_2 \frac{|\mathbf{K}_e + \mathbf{H}_{ea} \mathbf{W}_a \mathbf{Q}_s \mathbf{W}_a^H \mathbf{H}_{ea}^H|}{|\mathbf{K}_e|}\right] \right\}^+ ,\end{aligned} \tag{3.12}$$

where

$$\mathbf{K}_b = \mathbf{W}_b^H \mathbf{W}_b + \mathbf{W}_b^H \mathbf{H}_{bh} \mathbf{W}_h \mathbf{Q}_v \mathbf{W}_h^H \mathbf{H}_{bh}^H \mathbf{W}_b \tag{3.13}$$

$$\mathbf{K}_e = \sigma_e^2 \mathbf{I}_{N_e} + \mathbf{H}_{eh} \mathbf{W}_h \mathbf{Q}_v \mathbf{W}_h^H \mathbf{H}_{eh}^H \ . \tag{3.14}$$

Since a uniform power allocation is assumed for both Alice and Hugo,

$$\mathbf{Q}_s = \frac{P_s}{d} \mathbf{I}_d \tag{3.15}$$

$$\mathbf{Q}_v = \frac{P_i}{N_h - d} \mathbf{I}_{N_h - d} \ . \tag{3.16}$$

[1]Note that there is an alternative definition of average secrecy rate in [47], where the full CSI of the main channel must be used at the transmitter to vary the transmission rate in every channel fading block. This is not possible in my setting since only knowledge of the quantized precoding matrices is available.

Applying the receive beamforming matrix $\mathbf{W}_b$, the average secrecy rate $\overline{R_{sec}}$ can be interpreted as in (3.17) for a given number of data streams d.

$$\overline{R_{sec}}(d) = \left\{ \overline{R_b}(d) - \mathbb{E}_{\mathbf{H}_{ea},\mathbf{H}_{eh},\mathbf{W}_a,\mathbf{W}_h} \left[\log_2 \frac{\left| \sigma_e^2 \mathbf{I}_{N_e} + \frac{P_i}{N_h - d} \mathbf{H}_{eh} \mathbf{W}_h \mathbf{W}_h^H \mathbf{H}_{eh}^H + \frac{P_s}{d} \mathbf{H}_{ea} \mathbf{W}_a \mathbf{W}_a^H \mathbf{H}_{ea}^H \right|}{\left| \sigma_e^2 \mathbf{I}_{N_e} + \frac{P_i}{N_h - d} \mathbf{H}_{eh} \mathbf{W}_h \mathbf{W}_h^H \mathbf{H}_{eh}^H \right|} \right] \right\}^+ , \tag{3.17}$$

where

$$\overline{R_b}(d) \triangleq \mathbb{E}_{\mathbf{H}_{ba},\mathbf{H}_{bh},\mathbf{W}_a,\mathbf{W}_h} \left[\log_2 \frac{\left| \mathbf{W}_b^H \mathbf{W}_b + \frac{P_i}{N_h - d} \mathbf{W}_b^H \mathbf{H}_{bh} \mathbf{W}_h \mathbf{W}_h^H \mathbf{H}_{bh}^H \mathbf{W}_b + \frac{P_s}{d} \mathbf{W}_b^H \mathbf{H}_{ba} \mathbf{W}_a \mathbf{W}_a^H \mathbf{H}_{ba}^H \mathbf{W}_b \right|}{\left| \mathbf{W}_b^H \mathbf{W}_b + \frac{P_i}{N_h - d} \mathbf{W}_b^H \mathbf{H}_{bh} \mathbf{W}_h \mathbf{W}_h^H \mathbf{H}_{bh}^H \mathbf{W}_b \right|} \right] \tag{3.18}$$

$$= \mathbb{E}_{\mathbf{H}_{ba},\mathbf{H}_{bh},\mathbf{W}_a,\mathbf{W}_h} \left[\log_2 \frac{\left| \mathbf{I}_d + \frac{P_i}{N_h - d} \mathbf{U}_h \mathbf{\Lambda}_{h1} \mathbf{V}_{h1}^H \mathbf{W}_h \mathbf{W}_h^H \mathbf{V}_{h1} \mathbf{\Lambda}_{h1}^H \mathbf{U}_h^H + \frac{P_s}{d} \mathbf{\Lambda}_{a1} \mathbf{V}_{a1}^H \mathbf{W}_a \mathbf{W}_a^H \mathbf{V}_{a1} \mathbf{\Lambda}_{a1}^H \right|}{\left| \mathbf{I}_d + \frac{P_i}{N_h - d} \mathbf{U}_h \mathbf{\Lambda}_{h1} \mathbf{V}_{h1}^H \mathbf{W}_h \mathbf{W}_h^H \mathbf{V}_{h1} \mathbf{\Lambda}_{h1}^H \mathbf{U}_h^H \right|} \right] . \tag{3.19}$$

The following lemma, discussed in [42] and [43], provides an asymptotic approximation to the average mutual information of a MIMO channel for a large number of antennas.

Lemma 3.1. *Given an $r \times t$ matrix $\mathbf{H}$ composed of independent, circular complex, zero-mean, unit variance Gaussian elements, then for asymptotically large r and t,*

$$\mathbb{E} \left[\log_2 \left| \mathbf{I}_r + \frac{P}{t} \mathbf{H} \mathbf{H}^H \right| \right] = t F \left(\frac{r}{t}, P \right) , \tag{3.20}$$

where $F(\cdot)$ is given in (2.3).

3.4 Limited Rate Feedback Analysis

3.4.1 Unknown Eavesdropper CSI

To begin my analysis, I investigate the scenario where Eve's channel state information is completely unknown. A formal secrecy metric is impossible to define without any knowledge of Eve, so a reasonable alternative is to maximize the amount of jamming broadcast by Hugo subject to a certain acceptable rate for Bob, i.e., $\overline{R_b}(d) \geq R_t$, taking the effects of the limited feedback quantization error into account. Although the secrecy rate of such a scheme cannot be quantified, this approach aims at making the unintended reception of the signal as difficult as possible [48]. Direct evaluation of (3.19) in terms of the parameters of interest is difficult, so instead I focus on optimizing a lower bound on (3.19) derived below.

Lemma 3.2. *The random quantization codebook model satisfies the following properties:*

$$\mathbb{E}\left[\mathbf{V}_{ij}^H \widehat{\mathbf{V}}_{ij} \widehat{\mathbf{V}}_{ij}^H \mathbf{V}_{ij}\right] = \mathbb{E}\left[\widehat{\mathbf{V}}_{ij}^H \mathbf{V}_{ij} \mathbf{V}_{ij}^H \widehat{\mathbf{V}}_{ij}\right] = \left(1 - \frac{D_i}{M_i}\right) \mathbf{I}_{M_i} \,. \tag{3.21}$$

Proof. See Appendix A.1. □

Theorem 3.1. *The average rate of the main channel using random quantization codebooks of size B_a and B_h can be approximately bounded below by*

$$\overline{R_b}(d) \gtrsim \overline{R_{b,LB}}(d) \triangleq \overline{R}(d) - M_h \log_2 \left(1 + \frac{P_i N_h G_h 2^{-\frac{B_h}{T_h}}}{(N_h - d) M_h}\right) - M_a \log_2 \frac{M_a}{M_a - G_a 2^{-\frac{B_a}{T_a}}} \,, \tag{3.22}$$

where $\overline{R}(d) = dF(\frac{l}{d}, \frac{m}{d} P_s)$ is the ideal rate achieved with perfect CSI.

Proof. See Appendix A.2. □

The use of limited rate feedback produces the two negative terms in (3.22). The first term is due to interference leakage from Hugo, and the second is due to mismatch with the desired beamforming for Alice. Together they constitute the throughput loss caused by quantization errors. It is clear from (3.22) that $\overline{R_b}(d) = \overline{R}(d)$ and the rate loss is 0 if the feedback rate is infinite, *i.e.,* if $B_a, B_h \to \infty$.

The lower bound is tight when the number of feedback bits is sufficiently large and properly allocated, so I use it as an approximation to the average link rate. Since the codebook size is fixed for all the channel realizations, the optimization problem is to find a bit allocation strategy for Alice and Hugo that maximizes the jamming power subject to the constraint that the lower bound on the average rate in (3.22) is above the target value R_t. In addition to optimal values for B_a and B_h, there is a best choice for d that maximizes the jamming power. In practical applications, the possible number of values for d is limited ($d \leq m$), so my approach is to repeat the optimization over B_a and B_h for each possible d, and then choose the value for d whose optimal bit allocation provides the largest jamming power. Since Alice is unable to exploit rate or power adaptation, she always transmits with full power in this scenario, i.e., $P_s = P_a$. Consequently, I choose B_a and B_h to maximize the jamming power:

$$
\begin{aligned}
& P_i^*(d) = \max_{B_a, B_h} \; P_i(B_a, B_h) \qquad (3.23) \\
\text{s.t.} \quad & \overline{R_{b,LB}}(d) = R_t \\
& B_a + B_h = B \\
& B_a, B_h \in \mathbb{Z}^+ \\
& 0 \leq P_i \leq P_h \ ,
\end{aligned}
$$

where $\mathbb{Z}^+$ is the set of non-negative integers.

In general, the required optimization is an integer programming problem. However, if the

integer constraint is relaxed, the following closed-form expression for the optimal solution can be obtained when Alice and Hugo have the same number of antennas.

Theorem 3.2. *When $N_a = N_h$, the optimal solution to (3.23) for B_h without the integer constraint, and the corresponding maximum jamming power is shown in (3.24a)-(3.24c) for a given number of data streams d, where $r = 2^{\frac{\overline{R}(d)-R_t}{M_h}}$ and $r_0 = 2^{\frac{\overline{R}(d)}{M_h}}$.*

1. $2G_h < M_h(1-r_0^{-1})$ *and* $P_h \geq \frac{(N_h-d)M_h}{N_hG_h}\left(\frac{M_h-G_h}{M_h}r_0-1\right)2^{\frac{B}{T_h}}$

$$\begin{cases} B_h^*(d) = B \\ P_i^*(d) = \dfrac{(N_h-d)M_h}{N_hG_h}\left(\dfrac{M_h-G_h}{M_h}r-1\right)2^{\frac{B}{T_h}} \end{cases} \quad \text{if } 0 < R_t < \overline{R}(d) - M_h\log_2\frac{M_h}{M_h-2G_h}$$

$$\begin{cases} B_h^*(d) = B - T_h\log_2\dfrac{2G_hr}{M_h(r-1)} \\ P_i^*(d) = \dfrac{(N_h-d)M_h^2(r-1)^2}{4N_hG_h^2r}2^{\frac{B}{T_h}} \end{cases} \quad \text{if } \overline{R}(d) - M_h\log_2\frac{M_h}{M_h-2G_h} \leq R_t \leq \overline{R}(d) - M_h\log_2\frac{M_h}{M_h-2G_h2^{-\frac{B}{T_h}}}$$

$$\begin{cases} B_h^*(d) = 0 \\ P_i^*(d) = \dfrac{(N_h-d)M_h}{N_hG_h}\left(\dfrac{M_h-G_h2^{-\frac{B}{T_h}}}{M_h}r-1\right) \end{cases} \quad \text{if } \overline{R}(d) - M_h\log_2\frac{M_h}{M_h-2G_h2^{-\frac{B}{T_h}}} < R_t \leq \overline{R}(d) - M_h\log_2\frac{M_h}{M_h-G_h2^{-\frac{B}{T_h}}}$$

(3.24a)

2. $M_h(1-r_0^{-1}) \leq 2G_h < M_h2^{\frac{B}{T_h}}(1-r_0^{-1})$ *and* $P_h \geq \frac{(N_h-d)M_h^2(r_0-1)^2}{4N_hG_h^2r_0}2^{\frac{B}{T_h}}$

$$\begin{cases} B_h^*(d) = B - T_h\log_2\dfrac{2G_hr}{M_h(r-1)} \\ P_i^*(d) = \dfrac{(N_h-d)M_h^2(r-1)^2}{4N_hG_h^2r}2^{\frac{B}{T_h}} \end{cases} \quad \text{if } 0 < R_t \leq \overline{R}(d) - M_h\log_2\frac{M_h}{M_h-2G_h2^{-\frac{B}{T_h}}}$$

$$\begin{cases} B_h^*(d) = 0 \\ P_i^*(d) = \dfrac{(N_h-d)M_h}{N_hG_h}\left(\dfrac{M_h-G_h2^{-\frac{B}{T_h}}}{M_h}r-1\right) \end{cases} \quad \text{if } \overline{R}(d) - M_h\log_2\frac{M_h}{M_h-2G_h2^{-\frac{B}{T_h}}} < R_t \leq \overline{R}(d) - M_h\log_2\frac{M_h}{M_h-G_h2^{-\frac{B}{T_h}}}$$

(3.24b)

3. $M_h2^{\frac{B}{T_h}}(1-r_0^{-1}) \leq 2G_h < 2M_h2^{\frac{B}{T_h}}(1-r_0^{-1})$ *and* $P_h \geq \frac{(N_h-d)M_h}{N_hG_h}\left(\frac{M_h-G_h2^{-\frac{B}{T_h}}}{M_h}r_0-1\right)$

$$\begin{cases} B_h^*(d) = 0 \\ P_i^*(d) = \dfrac{(N_h-d)M_h}{N_hG_h}\left(\dfrac{M_h-G_h2^{-\frac{B}{T_h}}}{M_h}r-1\right) \end{cases} \quad \text{if } 0 < R_t \leq \overline{R}(d) - M_h\log_2\frac{M_h}{M_h-G_h2^{-\frac{B}{T_h}}} \tag{3.24c}$$

Proof. By setting the lower bound in (3.22) equal to R_t and replacing B_a by $B - B_h$, the jamming power can be expressed as

$$P_i(B_h) = \frac{(N_h - d)M_h}{N_h G_h 2^{-\frac{B_h}{T_h}}} \left(\frac{M_h - G_h 2^{-\frac{B-B_h}{T_h}}}{M_h} r - 1 \right) . \tag{3.25}$$

After relaxing the integer constraints, the optimization problem in (3.23) can be rewritten in standard form as

$$\begin{aligned} P_i^*(d) = \max_{B_h} \; & P_i(B_h) \\ \text{s.t.} \quad & B - B_h \geq 0 \\ & B_h \geq 0 \\ & P_i(B_h) \geq 0 \\ & P_h - P_i(B_h) \geq 0 \, . \end{aligned}$$

I solve this problem in general, but, due to space limitations, only present the results for P_h sufficiently large. For each case, I present the threshold. The closed-form solution is given in (3.24a)-(3.24c). □

The optimal d^* is taken to be the one that leads to the maximum jamming power $P_i^*(d)$. For the actual feedback link, search the integer values above and below $B_h^*(d)$ in (3.25) to determine the integer bit allocation $\widetilde{B}_h^*(d^*)$ and $\widetilde{B}_a^*(d^*) \triangleq B - \widetilde{B}_h^*(d^*)$ and the actual jamming power $\widetilde{P}_i^*(d^*)$. If $R_t > \overline{R}(d) - M_h \log_2 \frac{M_h}{M_h - G_h 2^{-\frac{B}{T_h}}}$, the target rate R_t cannot be achieved and the link is assumed to be in outage.

3.4.2 Statistical Eavesdropper CSI

In the last subsection, I developed a feedback bit allocation strategy for situations with no eavesdropper CSI that provided the maximum possible jamming power to disrupt potential eavesdroppers, subject to the constraint that a lower bound on the average rate of the desired link is above a target level R_t. However, the secrecy rate of such a scheme cannot in general be guaranteed; a well-endowed eavesdropper in the right location could end up with a better quality signal, and secrecy would be lost.

Here I assume that statistical information about Eve's channel is available. In particular, I assume isotropic distributions for $\mathbf{H}_{ea}$ and $\mathbf{H}_{eh}$, and investigate maximizing the average secrecy rate in (3.17) by adjusting both the feedback bit allocation and the amount of transmission power at Alice and Hugo:

$$\begin{aligned} \max_{B_a,B_h,P_s,P_i} \quad & \overline{R_{sec}}(d) \\ \text{s.t.} \quad & B_a + B_h = B \\ & B_a, B_h \in \mathbb{Z}^+ \\ & 0 \leq P_s \leq P_a \\ & 0 \leq P_i \leq P_h \ . \end{aligned} \tag{3.26}$$

My approach first optimizes the secrecy rate over B_a, B_h for fixed P_s and P_i, and then addresses the power allocation.

Theorem 3.3. *When $N_a = N_h$, the optimal solution for B_h without the integer constraint is shown in (3.27a)-(3.27c) for a given number of data streams d and fixed jamming power P_i.*

1. $2G_h < M_h$

$$B_h^*(d,P_i) = \begin{cases} B & \text{if } P_i > \frac{(N_h-d)M_h 2^{\frac{B}{T_h}}}{N_h(M_h-2G_h)} \\ T_h \log_2 \frac{-P_i N_h G_h + \sqrt{(P_i N_h G_h)^2 + P_i N_h (N_h-d) M_h^2 2^{\frac{B}{T_h}}}}{(N_h-d)M_h} & \text{if } \frac{(N_h-d)M_h}{N_h(M_h 2^{\frac{B}{T_h}}-2G_h)} \le P_i \le \frac{(N_h-d)M_h 2^{\frac{B}{T_h}}}{N_h(M_h-2G_h)} \\ 0 & \text{if } P_i < \frac{(N_h-d)M_h}{N_h(M_h 2^{\frac{B}{T_h}}-2G_h)} \end{cases} \tag{3.27a}$$

2. $M_h \le 2G_h < M_h 2^{\frac{B}{T_h}}$

$$B_h^*(d,P_i) = \begin{cases} T_h \log_2 \frac{-P_i N_h G_h + \sqrt{(P_i N_h G_h)^2 + P_i N_h (N_h-d) M_h^2 2^{\frac{B}{T_h}}}}{(N_h-d)M_h} & \text{if } P_i \ge \frac{(N_h-d)M_h}{N_h(M_h 2^{\frac{B}{T_h}}-2G_h)} \\ 0 & \text{if } P_i < \frac{(N_h-d)M_h}{N_h(M_h 2^{\frac{B}{T_h}}-2G_h)} \end{cases} \tag{3.27b}$$

3. $2G_h \ge M_h 2^{\frac{B}{T_h}}$

$$B_h^*(d,P_i) = 0 \tag{3.27c}$$

Proof. Define $\widetilde{\mathbf{H}}_{ea} \triangleq \mathbf{H}_{ea}\mathbf{W}_a$ and $\widetilde{\mathbf{H}}_{eh} \triangleq \mathbf{H}_{eh}\mathbf{W}_h$. The average secrecy rate is given by

$$\overline{R_{sec}}(d) = \left\{ \overline{R_b}(d) - \mathbb{E}_{\widetilde{\mathbf{H}}_{ea},\widetilde{\mathbf{H}}_{eh}} \left[R_e(d) \right] \right\}^+ , \tag{3.28}$$

where

$$R_e(d) \triangleq \log_2 \frac{\left| \sigma_e^2 \mathbf{I}_{N_e} + \frac{P_i}{N_h-d} \widetilde{\mathbf{H}}_{eh} \widetilde{\mathbf{H}}_{eh}^H + \frac{P_s}{d} \widetilde{\mathbf{H}}_{ea} \widetilde{\mathbf{H}}_{ea}^H \right|}{\left| \sigma_e^2 \mathbf{I}_{N_e} + \frac{P_i}{N_h-d} \widetilde{\mathbf{H}}_{eh} \widetilde{\mathbf{H}}_{eh}^H \right|} .$$

Because each quantized precoding matrix is isotropically distributed over the Grassmann manifold, changing the number of codebook entries does not affect its distribution. In addition, the quantized precoding matrices are independent of Eve's channel. Varying the bit allocations does not affect the probability distribution of $\widetilde{\mathbf{H}}_{ea}$ and $\widetilde{\mathbf{H}}_{eh}$, since it is only determined by their dimensions. Thus, the second term in (3.28) does not depend on B_a and B_h, and hence the initial optimization problem over B_a and B_h is given by

$$\begin{aligned} \max_{B_a, B_h} \quad & \overline{R_b}(d) \\ \text{s.t.} \quad & B_a + B_h = B \\ & B_a, B_h \in \mathbb{Z}^+ \ . \end{aligned}$$

To solve this problem, I first relax the integer constraint. Then, I use the average link rate lower bound in (3.22) as an approximation to the above objective function. Since $\overline{R}(d)$ is independent of B_a and B_h, substituting $B - B_h$ for B_a, the optimal bit allocation problem is converted to

$$\begin{aligned} \min_{B_h} \quad & \left(1 + \frac{P_i N_h G_h 2^{-\frac{B_h}{T_h}}}{(N_h - d)M_h}\right) \frac{M_a}{M_a - G_a 2^{-\frac{B-B_h}{T_a}}} \\ \text{s.t.} \quad & B_h - B \leq 0 \\ & -B_h \leq 0 \ . \end{aligned} \tag{3.29}$$

This formulation gives rise to a convex optimization problem with standard solution methods. As in the previous case, a closed-form solution can be obtained in (3.27a)-(3.27c) when Alice and Hugo have the same number of antennas. □

With the integer bit allocation $\widetilde{B}_h^*$ determined ($B_h^*(d, P_i)$ rounded to the nearest integer), the idea is to substitute it back into the expression for the average secrecy rate so that

it can be optimized over the power allocation. This is prohibitively complicated for the general expression of (3.17), so instead I proceed by substituting the expressions for $\widetilde{B}_h^*$ of Theorem 3.3 in the lower bound of (3.22). This leads to the following result.

Theorem 3.4. *The average secrecy rate under the optimal bit allocation strategy can be approximately bounded below by*

$$\begin{aligned}
&\overline{R_{sec}}(d) \gtrsim \overline{R_{sec,LB}}(d) \\
&\triangleq \left\{ dF\left(\frac{l}{d}, \frac{m}{d}P_s\right) - M_h \log_2\left(1 + \frac{P_i N_h G_h 2^{-\frac{\widetilde{B}_h^*}{T_h}}}{(N_h - d)M_h}\right) - M_h \log_2 \frac{M_h}{M_h - G_h 2^{-\frac{B-\widetilde{B}_h^*}{T_h}}} \right. \\
&\left. -(N_h - d)\left[F\left(\frac{N_e}{N_h - d}, \frac{P_i}{\sigma_e^2 + P_s}\right) - F\left(\frac{N_e}{N_h - d}, \frac{P_i}{\sigma_e^2}\right)\right] - N_e \log_2(1 + \frac{P_s}{\sigma_e^2}) \right\}^+ .
\end{aligned} \tag{3.30}$$

Proof. See Appendix A.3. □

Assume Bob chooses values for P_s, P_i and d to maximize the average secrecy rate lower bound given in (3.30):

$$\max_{\substack{0 \le P_s \le P_a \\ 0 \le P_i \le P_h}} \overline{R_{sec,LB}}(d) \ . \tag{3.31}$$

This can be done by performing a 2-dimensional line search for P_s and P_i for each candidate d, and then choosing the value for d that provides the largest value for the lower bound.

3.5 Simulation Results

In this section, I verify the validity of the analytical results through Monte Carlo simulations. For a given channel realization, I use the numerical results in [49] to randomly generate the associated quantized feedback. Utilizing the statistics of random quantization codebooks,

this method simulates the quantization procedure without generating an actual codebook, and reduces the computational complexity as B grows. I consider a case where Alice and Hugo have an equal number of antennas, i.e., $N_a = N_h = 4$, and set $N_b = 2$. The limited feedback bandwidth B is 20 bits, and all results are based on averages obtained over 1000 independent channel realizations. While conventional commercial systems employ a coarser feedback quantization, note that the application considered here is significantly different and much more advanced than those addressed in current systems. Very accurate CSI is required if one is attempting to null a strong interferer and obtain a reasonable secrecy rate, as I am trying to do in this case. In addition, the benefit of the proposed optimization is less apparent for small numbers of feedback bits.

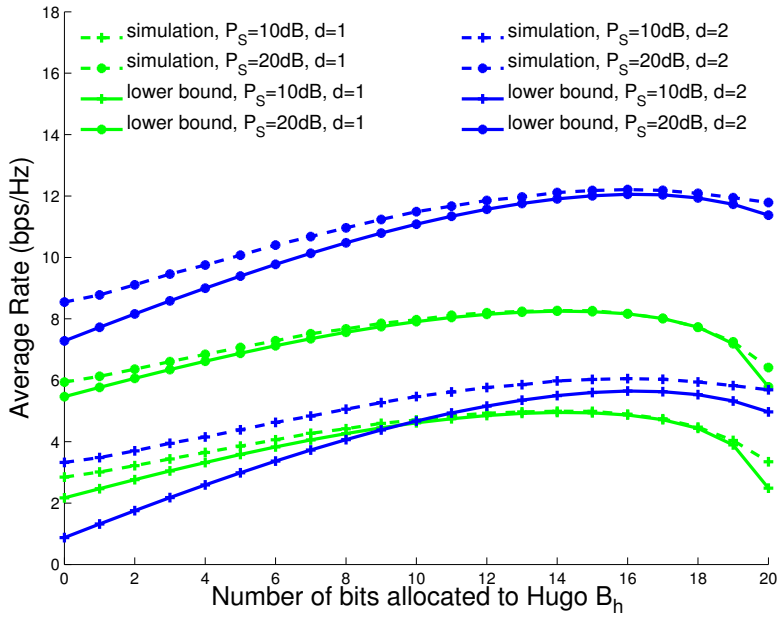

Figure 3.2: Accuracy of the average rate lower bound with $P_s = 10$ dB, 20 dB and $P_i = 10$ dB.

Figure 3.2 compares the numerical evaluation of (3.18) with the approximate lower bound in Theorem 3.1 as a function of the number of bits allocated to Hugo (B_h). The figure shows the results for the two possible values of $d = 1$ and $d = 2$, with transmit powers fixed at $P_s = 10$ dB, 20 dB and $P_i = 10$ dB. This figure verifies the accuracy of the approximate lower bound especially at the maximum rate where B_a and B_h are properly allocated; most

importantly, the peaks of the two curves coincide exactly.

In the next few examples, I assume no CSI is available for the eavesdropper, and thus I focus on maximizing the power available for jamming. Figure 3.3 shows the optimal bit allocation versus target rate with $P_s = 20$ dB for different d. In most cases, the interference leakage term in (3.22) dominates, and thus Hugo receives a higher allocation of bits than Alice. Note that even for very small target rates, the optimal solution sometimes still allocates a small number of bits for feedback to Alice, since the loss in beamforming gain cannot be compensated for by reduced interference (i.e., when $d = 1$, those 2 remaining bits are more valuable to Alice than they are to Hugo). Figure 3.4 shows the optimal jamming power versus target rate that is achieved with $P_s = 20$ dB. When the target rate for the main channel is low, a single data stream is transmitted and the rest of the dimensions are used to interfere with the eavesdropper at higher jamming power. As the target rate increases, a single data stream can no longer meet the rate requirement in the presence of any significant jamming. Thus, at this point it is better to switch to $d = 2$ and achieve higher jamming power over the two remaining spatial dimensions. The available jamming power eventually decays to zero as the constraint on the quality of the main link becomes more stringent. Figures 3.5 and 3.6 plot the optimal jamming power as a function of the number of bits allocated to Hugo for small and large values of R_t respectively. The proper choice of B_h and d make a significant difference in the amount of jamming power that is available for interfering with any eavesdroppers that are present. Figures 3.5 and 3.6 illustrate the advantage of my results over a non-optimized feedback allocation. Figure 3.7 plots the average rate for the main channel (dashed line) together with the target lower bound on the rate (solid line), indicating that the constraint on the lower bound is met in all instances. Also plotted is the rate achieved in the ideal case (dash-dot lines) where there are no quantization errors for $d = 1$ (about 9 bps/Hz) and $d = 2$ (about 14 bps/Hz). The fact that the target rate approaches the ideal rate of 14 bps/Hz when $d = 2$ is a reflection of the fact that the available jamming power is essentially zero.

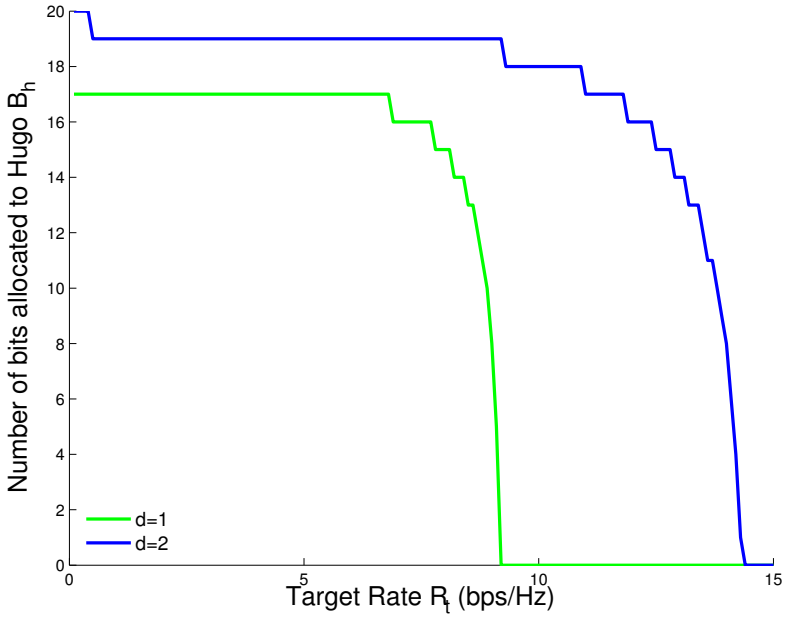

Figure 3.3: Optimal bit allocation versus target rate with $P_s = 20$ dB.

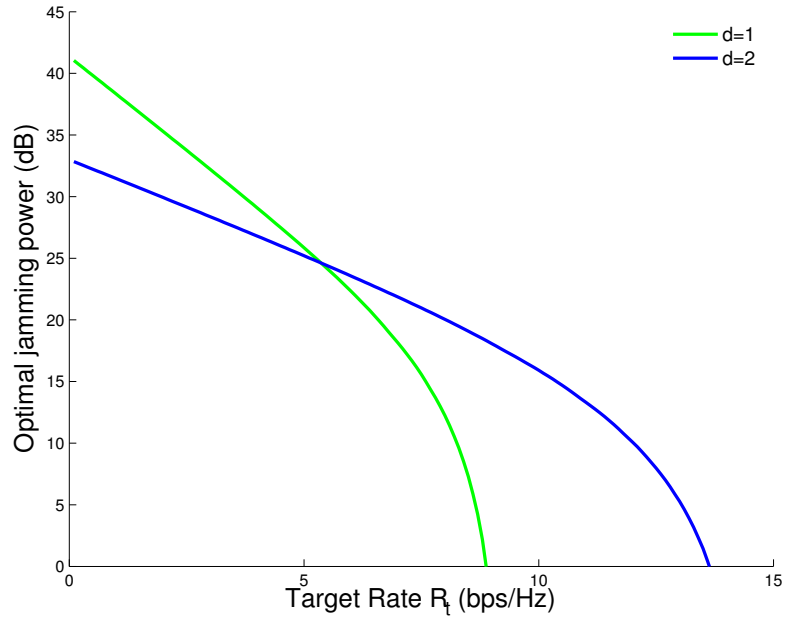

Figure 3.4: Optimal jamming power versus target rate with $P_s = 20$ dB.

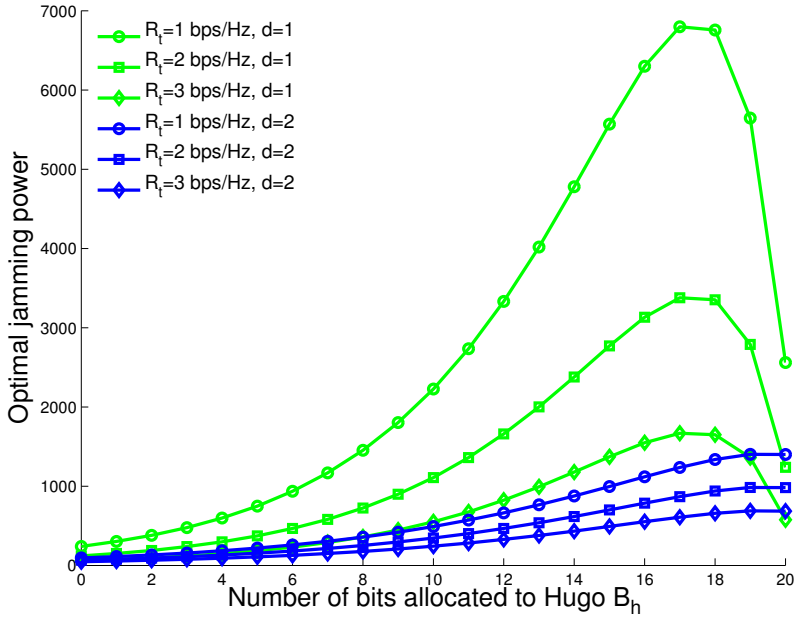

Figure 3.5: Optimal jamming power versus bit allocation for low target rates.

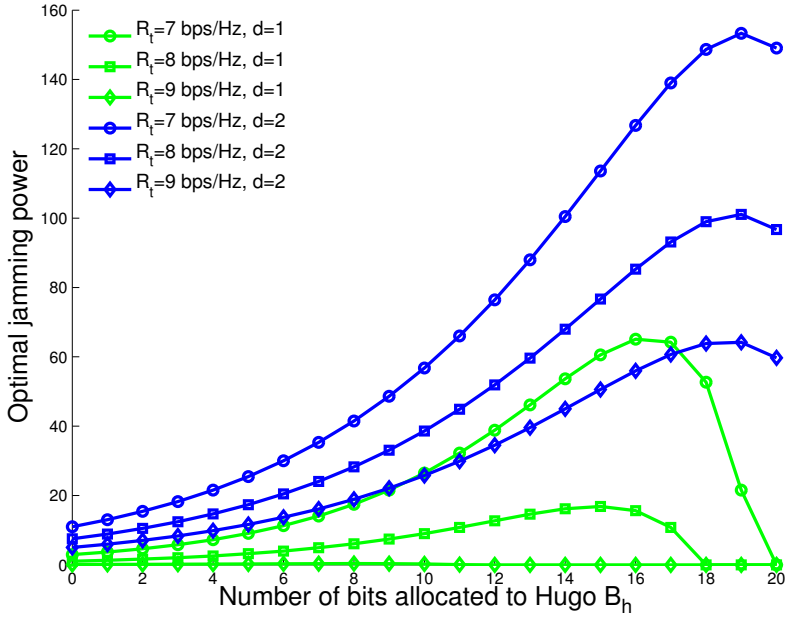

Figure 3.6: Optimal jamming power versus bit allocation for high target rates.

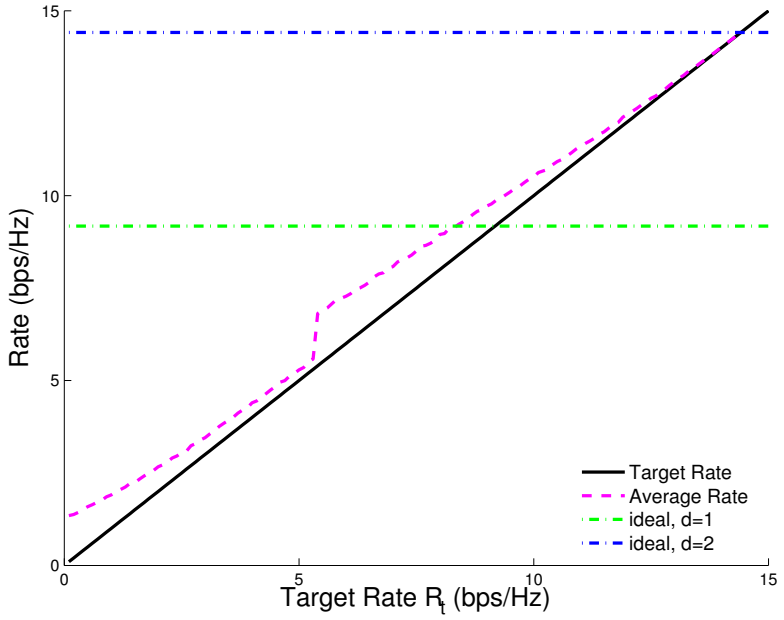

Figure 3.7: The average rate using optimal jamming power and optimal bit allocation.

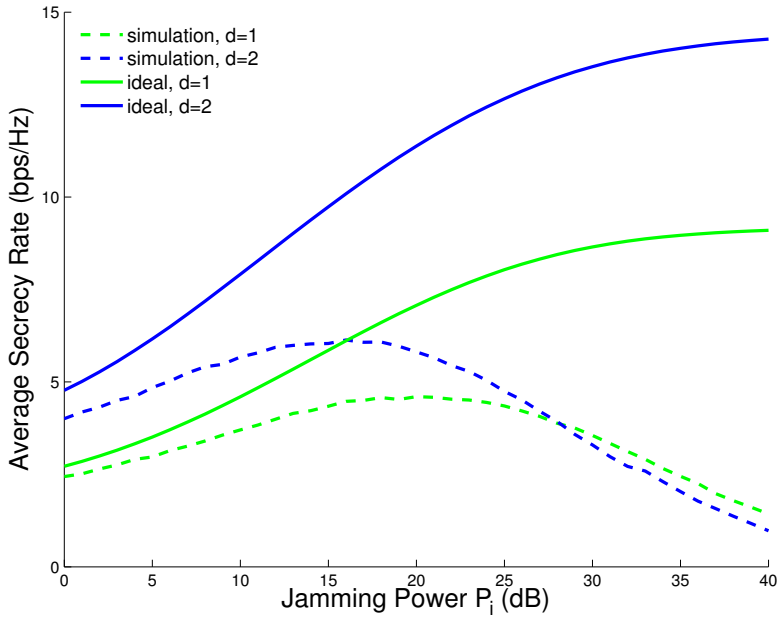

Figure 3.8: Average secrecy rate versus jamming power with $P_s = 20$ dB.

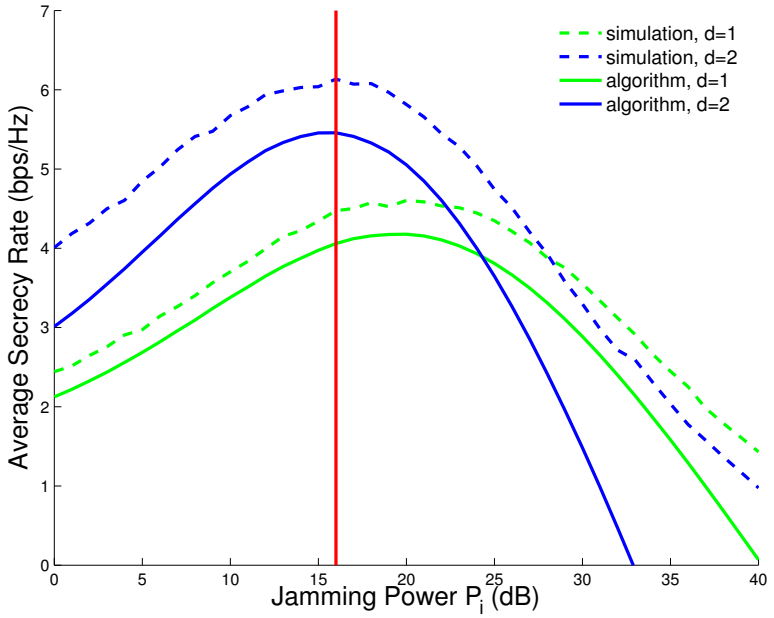

Figure 3.9: Accuracy of the average secrecy rate lower bound and optimal jamming power with $P_s = 20$ dB.

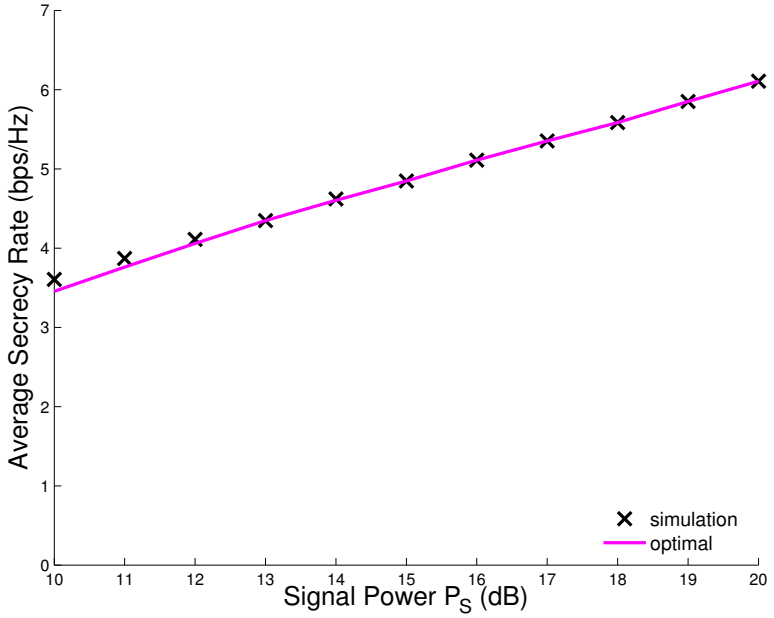

Figure 3.10: Accuracy of the average secrecy rate versus transmit power.

For the case where statistical information about Eve's CSI is available, I show the average secrecy rate as a function of jamming power with $N_e = 2$ in Figure 3.8, both for the case with no feedback quantization error (solid lines) and with feedback quantization (dashed lines) assuming the optimal feedback allocation. In this example, the power constraints are 20 dB and 40 dB at Alice and Hugo respectively. This figure illustrates the trade-off associated with the use of a cooperative jammer in a scenario with limited feedback and hence inaccurate CSI; at a certain point the jamming hurts the desired receiver more than it confuses the eavesdropper. In Figure 3.9, I add curves that represent the lower bound of the average secrecy rate in Theorem 3.4, which are used to identify the optimal bit allocation for each value of d. The peak values of the lower bound accurately coincide with the peak values of the actual secrecy rate, so use of the lower bound allows the optimal system parameters to be found in the absence of the exact rate. In this case, the peak of the lower bound for $d = 2$ exceeds that for $d = 1$, and leads to the choice of an optimal jamming power of $P_i^* = 16$ dB, indicated by the vertical red line. Figure 3.10 demonstrates the accuracy of the average achievable secrecy rate expression with respect to transmit power P_S. The plot shows that the average secrecy rate achieved by using the optimal results (P_i^*, d^* and B_h^*) obtained from (3.31) is essentially identical to the best possible average secrecy rate obtained from Monte Carlo simulations according to (3.26). This figure also illustrates that transmitting with full power at Alice achieves the maximum average secrecy rate.

3.6 Summary

This chapter has considered power and bit allocation strategies for enhanced secrecy in a limited rate feedback MIMO wiretap channel involving a cooperative jammer. I examined two cases, one where no information about the eavesdroppers is available, and one where statistical channel state information is available. With no information about the eavesdrop-

pers, I showed how to choose the allocation of feedback bits to the transmitter and helper in order to maximize the amount of jamming power available to interfere with the eavesdroppers, subject to maintaining a lower bound on the target rate for the desired link. A closed-form solution was found for the special case where the transmitter and jammer have the same number of antennas. For the case of statistical CSI, I derived an approximate lower bound on the average secrecy rate, and again found a closed-form solution for the feedback bit allocation that maximizes this lower bound for an equal number of transmit antennas. Optimization of the transmit power in this case requires a 2-dimensional numerical search. Simulation results indicate the accuracy of the approximations used in this chapter, and demonstrate how proper choice of the feedback bit allocation can dramatically enhance the security provided by the cooperative jammer.

Chapter 4

Two-User MISO Interference Channel

In this chapter, I study enhanced wireless communication in a two-user multiple-input single-output (MISO) interference channel with limited rate feedback and transmitter cooperation. Each receiver quantizes the channel state information (CSI) of the direct and cross channels, and sends the codebook indices back to the transmitters through two sum-rate-limited feedback channels. The quantization errors reduce the beamforming gain from the direct transmitter, and cause interference leakage from the cross transmitter. First, I approximate the average transmission rate of each link, and use the sum rate to find the optimal transmit power and corresponding feedback bit allocation. I show that the maximum sum throughput is achieved using full transmit power, and the achievable sum rate under limited feedback is bounded above by a constant. I then extend the results to the case where secrecy is desired. In contrast to the first problem, increasing the transmit power beyond a certain point decreases the secrecy performance. I derive all the results in closed form. Simulations validate the theoretical analysis and demonstrate the significant performance gains that result from the use of optimal transmit power control and intelligent feedback bit allocation.

4.1 Introduction

In an interference channel, multiple wireless communication links are simultaneously active in the same time and frequency resource, and hence potentially interfere with each other. A long history of studying the interference channel has provided various achievable rate regions [50–54]. Lately, extending the interfence channel results to cases involving multiple antenna transceivers has drawn significant interest [55–58]. Employing multiple antennas increases the diversity gain and can help mitigate the interference in the system, provided that accurate channel state information (CSI) is available.

When CSI feedback is used, perfect and global CSI at the transmitters is unrealistic in practice due to the limited feedback bandwidth, and this issue has recently been a subject of substantial research [59–65]. The methods considered in these papers quantize the CSI at the receivers and then feed the quantized information back to the transmitters. Based on the feedback, the transmitters design precoding vectors, and adapt the signals to the current structure of the channels to achieve an acceptable performance or rate. More similar to the topic of this chapter are the studies conducted in [66] and [67], which examine the feedback bit partition between desired and interfering channels in a multicell system. However, there is relatively little work on feedback bit allocation in the interference channel subject to a constrained total feedback bandwidth.

Interference is not only an issue for communication rates, but also for security since information can be extracted from the "interference" by users who are not the intended recipients. In such scenarios it is desirable to minimize the leakage of information to those unintended receivers [68]. In Chapter 3, I studied the impact of limited rate feedback on the wiretap channel with a helper, which can be considered as a special case of the two-user interference channel model. One receiver is a known eavesdropper, and a second transmitter is used to send a jamming signal to enhance secrecy. Secrecy for the two-user interference channel

has received considerable attention in recent years, and several approaches have been proposed that discuss how cooperation between the transmitters can be exploited to improve secrecy [32–41].

In this chapter, I study strategies for transmit power control and feedback bit allocation for the Gaussian two-user MISO interference channel with limited rate feedback. Random vector quantization (RVQ) is applied to the CSI of the direct and cross channels, and each receiver sends the indices of the corresponding codewords to both its own and the interfering transmitter through two sum-rate limited feedback channels. I consider two cases with respect to system performance. In the first, I assume that the receivers only decode their own messages, and treat the interference as Gaussian noise. I derive an approximation for the average transmission rate of each link, and find the transmit power control and feedback bit allocation that maximize the system sum throughput. In the second case, I consider the limited rate feedback interference channel with confidential messages, where each transmitter desires to send independent information to its intended receiver while ensuring mutual information-theoretic secrecy. In this case I assume the two transmitters are amenable to cooperation for improving the overall secrecy performance of the system. I derive an approximation for the average secrecy rate of each link, and optimize the sum secrecy rate over the transmit power and feedback bits allocated to the transmitters. Simulations demonstrate a significant gain for each performance metric.

The remainder of this chapter is organized as follows. In Section 4.2, I introduce the system model and assumptions, and provide preliminary background information. In Section 4.3 and 4.4, I discuss the power control and bit allocation algorithms for the two scenarios discussed above, and analyze the impact of limited rate feedback on the interference channel. In Section 4.5, I present simulation results to validate my algorithms.

Throughout the chapter, the indices i, j will be used as subscripts to denote one of the two transmitters, receivers, or channel links, and since I am dealing exclusively with the two-user

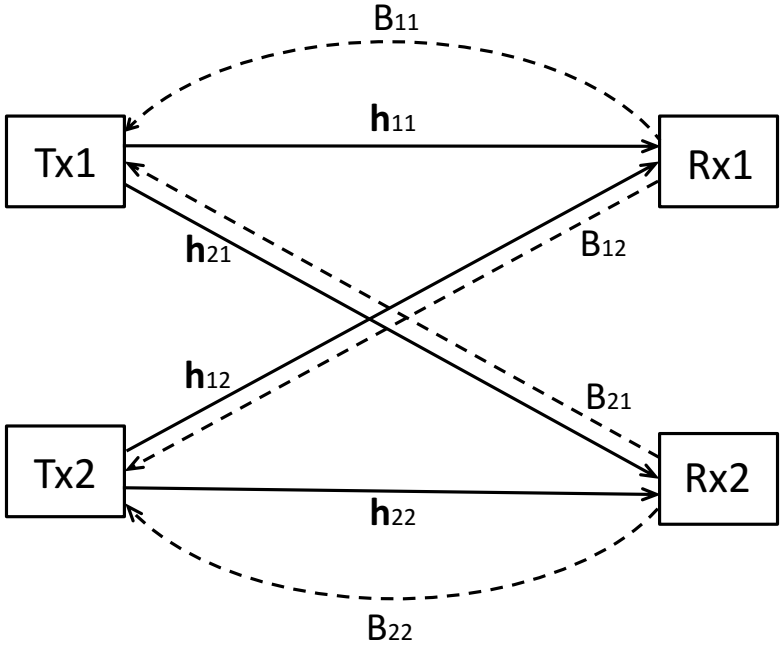

Figure 4.1: System model of a two-user MISO interference channel.

interference channel, i and j can only take on values in the set $\{1, 2\}$. If i and j are used together in an expression or as a subscript, they represent different values $i \neq j$.

4.2 Assumptions and Preliminaries

4.2.1 Signal Modeling

I consider the two-user MISO interference channel in Figure 4.1. Assume transmitter i possesses N_i antennas, while each receiver is equipped with a single antenna. I use $\mathbf{h}_{ii} \in \mathbb{C}^{N_i}$ to denote the direct channel from transmitter i to receiver i; and $\mathbf{h}_{ij} \in \mathbb{C}^{N_j}$ to denote the cross channel from transmitter j to receiver i ($\forall j \neq i; i, j \in \{1, 2\}$). The elements of these channel vectors are assumed to be independent and identically distributed (i.i.d.), and have zero-mean complex Gaussian distributions with variance σ_{ii}^2 and σ_{ij}^2 respectively. All channels experience independent block fading. The receivers are assumed to have perfect CSI for their

own channels. None of the transmitters know their instantaneous channels, but they obtain quantized CSI from the receivers through two error-free, zero-delay feedback channels. The i-th receiver quantizes the CSI of its direct and cross links respectively over each channel realization by selecting the closest codeword from two independent codebooks containing $2^{B_{ii}}$ and $2^{B_{ij}}$ entries. The B_{ii} and B_{ij} index bits corresponding to the chosen codewords are separately fed back to the direct and interfering transmitters. With the limited total rate constraint on the feedback channel, the i-th receiver is only able to feed back a fixed total number of bits b_i, such that

$$B_{ii} + B_{ij} = b_i \ .$$

Transmitter i employs a unit-norm beamforming vector $\mathbf{w}_i \in \mathbb{C}^{N_i}$ and sends a single data stream s_i to receiver i, which possibly interferes with receiver j. The superposition of the signal received by the i-th receiver is, therefore,

$$y_i = \mathbf{h}_{ii}^H \mathbf{w}_i s_i + \mathbf{h}_{ij}^H \mathbf{w}_j s_j + n_i \ ,$$

where n_i is the corresponding complex Gaussian noise with zero mean and unit variance. Assuming s_i is also Gaussian, define P_i as the actual transmit power used for sending the information signal, which satisfies the power constraint $\mathbb{E}\left[|s_i|^2\right] = P_i \leq P_{max,i}$, where $P_{max,i}$ is the maximum transmit power available at transmitter i.

While the transmitters are not aware of the instantaneous CSI of their channels, I assume they do know their corresponding channel statistics; *i.e.,* in particular, transmitter i has knowledge of σ_{ii}^2 and σ_{ji}^2. As we will see later, in order to compute the optimal feedback bit allocation, the transmitters will need to cooperate by sharing their maximum transmit power $P_{max,i}$, their actual transmit power P_i (for the case of secrecy) and the cross-channel statistics σ_{ji}^2.

4.2.2 Random Quantization Codebooks

For the analysis I assume random vector quantization (RVQ) codebooks. At receiver i, the quantized direct and cross channels are drawn from two independent random quantization codebooks $\mathcal{C}_{ii}$ and $\mathcal{C}_{ij}$, which are known to the corresponding transmitters beforehand. Assuming these codebooks are indexed by B_{ii} and B_{ij} feedback bits, they respectively contain $2^{B_{ii}}$ and $2^{B_{ij}}$ entries. Each of the codewords $\mathbf{c}_{ii} \in \mathcal{C}_{ii}$ and $\mathbf{c}_{ij} \in \mathcal{C}_{ij}$ are generated independently and isotropically over the N_i-dimensional and N_j-dimensional unit spheres. The i-th receiver selects the quantized versions of the channels according to the following rule:

$$\begin{aligned}\hat{\mathbf{h}}_{ii} &= \arg\max_{\mathbf{c}_{ii}\in\mathcal{C}_{ii}} |\tilde{\mathbf{h}}_{ii}^H \mathbf{c}_{ii}|^2 \\ \hat{\mathbf{h}}_{ij} &= \arg\max_{\mathbf{c}_{ij}\in\mathcal{C}_{ij}} |\tilde{\mathbf{h}}_{ij}^H \mathbf{c}_{ij}|^2 \ ,\end{aligned}$$

where $\tilde{\mathbf{h}}_{ii}$ and $\tilde{\mathbf{h}}_{ij}$ are the normalized channel directions:

$$\begin{aligned}\tilde{\mathbf{h}}_{ii} &= \frac{\mathbf{h}_{ii}}{\|\mathbf{h}_{ii}\|} \\ \tilde{\mathbf{h}}_{ij} &= \frac{\mathbf{h}_{ij}}{\|\mathbf{h}_{ij}\|} \ .\end{aligned}$$

According to [46], the average distortions associated with the given codebooks are bounded above by

$$\begin{aligned}D_{ii} &\triangleq \mathbb{E}\left[1 - |\tilde{\mathbf{h}}_{ii}^H \hat{\mathbf{h}}_{ii}|^2\right] \leq \Gamma(\frac{N_i}{N_i - 1}) 2^{-\frac{B_{ii}}{N_i - 1}} \\ D_{ij} &\triangleq \mathbb{E}\left[1 - |\tilde{\mathbf{h}}_{ij}^H \hat{\mathbf{h}}_{ij}|^2\right] \leq \Gamma(\frac{N_j}{N_j - 1}) 2^{-\frac{B_{ij}}{N_j - 1}} \ ,\end{aligned}$$

where $\Gamma(\cdot)$ represents the gamma function. Since the upper bounds are tight, I use them to

approximate the distortion rates. Then, the random quantization codebook model satisfies

$$\mathbb{E}\left[|\tilde{\mathbf{h}}_{ii}^H \hat{\mathbf{h}}_{ii}|^2\right] \approx 1 - \gamma_i 2^{-\frac{B_{ii}}{N_i - 1}} \tag{4.1}$$

$$\mathbb{E}\left[|\tilde{\mathbf{h}}_{ij}^H \hat{\mathbf{h}}_{ij}|^2\right] \approx 1 - \gamma_j 2^{-\frac{B_{ij}}{N_j - 1}} \ , \tag{4.2}$$

where $\gamma_i \triangleq \Gamma\left(\frac{N_i}{N_i - 1}\right)$ and $\gamma_j \triangleq \Gamma\left(\frac{N_j}{N_j - 1}\right)$.

4.2.3 Beamforming Design

Assume a relatively high SNR scenario such that the zero-forcing (ZF) transmit scheme is sum-rate optimal according to [45]. Under the assumption of perfect feedback, transmitter i chooses a unit-norm beamforming vector $\mathbf{w}_i$, which is orthogonal to $\mathbf{h}_{ji}$ and maximizes $|\mathbf{h}_{ii}^H \mathbf{w}_i|$. This vector is defined as the ZF beamformer and is given by

$$\mathbf{w}_i^{ZF} = \frac{\left(\mathbf{I} - \tilde{\mathbf{h}}_{ji}\tilde{\mathbf{h}}_{ji}^H\right)\mathbf{h}_{ii}}{\left\|\left(\mathbf{I} - \tilde{\mathbf{h}}_{ji}\tilde{\mathbf{h}}_{ji}^H\right)\mathbf{h}_{ii}\right\|} \ , \tag{4.3}$$

where $\left(\mathbf{I} - \tilde{\mathbf{h}}_{ji}\tilde{\mathbf{h}}_{ji}^H\right)$ denotes a projection onto the orthogonal complement of the column space of $\mathbf{h}_{ji}$. The vector $\mathbf{w}_i^{ZF}$ is constructed such that it nulls the interference at receiver j, and the remaining degrees of freedom are used to maximize the transmission rate to receiver i.

Under the limited rate feedback scenario, however, the transmitters only have access to the quantized version of their channels. Thus, the ZF beamforming vector at transmitter i becomes

$$\mathbf{w}_i = \frac{\left(\mathbf{I} - \hat{\mathbf{h}}_{ji}\hat{\mathbf{h}}_{ji}^H\right)\hat{\mathbf{h}}_{ii}}{\left\|\left(\mathbf{I} - \hat{\mathbf{h}}_{ji}\hat{\mathbf{h}}_{ji}^H\right)\hat{\mathbf{h}}_{ii}\right\|} \ . \tag{4.4}$$

4.2.4 Average Transmission Rate with Perfect CSI

The following lemmas provide useful preliminaries.

Lemma 4.1. *Consider a t-dimensional vector $\mathbf{g} \in \mathbb{C}^t$ with zero-mean and independent complex Gaussian entries,* i.e., $\mathbf{g} \sim \mathcal{CN}(\mathbf{0}, \sigma_g^2 \mathbf{I})$. *The expected value of the logarithm of $\|\mathbf{g}\|^2$ is given by*

$$\mathbb{E}\left[\ln \|\mathbf{g}\|^2\right] = \psi(t) + \ln \sigma_g^2 \,, \tag{4.5}$$

where $\psi(\cdot)$ is the digamma function.

Proof. See Appendix B.1. □

Lemma 4.2. *For any two i.i.d. unit-norm vectors $\mathbf{u}, \mathbf{v} \in \mathbb{C}^t$, the following holds:*

$$\mathbb{E}\left[\ln \left|\mathbf{u}^H \mathbf{v}\right|^2\right] = \psi(1) - \psi(t) \tag{4.6}$$

$$\mathbb{E}\left[\ln \left(1 - \left|\mathbf{u}^H \mathbf{v}\right|^2\right)\right] = \psi(t-1) - \psi(t) \,. \tag{4.7}$$

Proof. See Appendix B.2. □

Lemma 4.3. *Let ρ be a fixed constant and let Q be a random variable. The following approximation holds for large ρ:*

$$\mathbb{E}\left[\log_2 (1 + \rho Q)\right] \approx \log_2 \left(1 + \rho \, e^{\mathbb{E}[\ln Q]}\right) \,. \tag{4.8}$$

Proof. See Appendix B.3. □

Under perfect feedback, the transmitters can suppress all interference using the ZF beamforming vectors in (4.3). The average transmission rate of the i-th transmitter-receiver link

with perfect CSI can be expressed as

$$R_i^{ZF} = \mathbb{E}\left[\log_2\left(1 + P_i\left|\mathbf{h}_{ii}^H \mathbf{w}_i^{ZF}\right|^2\right)\right] \ . \tag{4.9}$$

Proposition 4.1. *The average transmission rate of the i-th transmitter-receiver link with perfect feedback satisfies the approximation:*

$$R_i^{ZF} \approx \log_2\left(1 + P_i\sigma_{ii}^2 e^{\psi(N_i-1)}\right) \ . \tag{4.10}$$

Proof. Plugging the ZF beamforming vector (4.3) into (4.9),

$$\begin{aligned}
R_i^{ZF} &= \mathbb{E}\left[\log_2\left(1 + \frac{P_i\left|\mathbf{h}_{ii}^H\left(\mathbf{I} - \tilde{\mathbf{h}}_{ji}\tilde{\mathbf{h}}_{ji}^H\right)\mathbf{h}_{ii}\right|^2}{\left\|\left(\mathbf{I} - \tilde{\mathbf{h}}_{ji}\tilde{\mathbf{h}}_{ji}^H\right)\mathbf{h}_{ii}\right\|^2}\right)\right] \\
&= \mathbb{E}\left[\log_2\left(1 + P_i\left\|\left(\mathbf{I} - \tilde{\mathbf{h}}_{ji}\tilde{\mathbf{h}}_{ji}^H\right)\mathbf{h}_{ii}\right\|^2\right)\right] \\
&= \mathbb{E}\left[\log_2\left(1 + P_i\left(1 - \left|\tilde{\mathbf{h}}_{ji}^H\tilde{\mathbf{h}}_{ii}\right|^2\right)\|\mathbf{h}_{ii}\|^2\right)\right] \\
&\approx \log_2\left(1 + P_i e^{\mathbb{E}\left[\ln\left(\left(1-\left|\tilde{\mathbf{h}}_{ji}^H\tilde{\mathbf{h}}_{ii}\right|^2\right)\|\mathbf{h}_{ii}\|^2\right)\right]}\right) \qquad (4.11) \\
&= \log_2\left(1 + P_i e^{\mathbb{E}\left[\ln\left(1-\left|\tilde{\mathbf{h}}_{ji}^H\tilde{\mathbf{h}}_{ii}\right|^2\right)\right]+\mathbb{E}\left[\ln\|\mathbf{h}_{ii}\|^2\right]}\right) \\
&= \log_2\left(1 + P_i\sigma_{ii}^2 e^{\psi(N_i-1)}\right) \ . \qquad (4.12)
\end{aligned}$$

Under the high SNR assumption, I use the approximation in (4.11) following from (4.8) in Lemma 4.3. Given that $\tilde{\mathbf{h}}_{ii}$ is independent of $\tilde{\mathbf{h}}_{ji}$, (4.12) is obtained from (4.5) in Lemma 4.1 and (4.7) in Lemma 4.2. □

4.3 Limited Rate Feedback Analysis

Here I assume that each receiver only intends to decode the information signal from its paired transmitter, treating the message of the other transmitter as interference. There is no attempt to decode and subtract the interfering message. For the given beamforming vectors in (4.4), the average achievable transmission rate of the i-th transmitter-receiver link under limited rate feedback is:

$$R_i = \mathbb{E}\left[\log_2\left(1 + \frac{P_i \left|\mathbf{h}_{ii}^H \mathbf{w}_i\right|^2}{1 + P_j \left|\mathbf{h}_{ij}^H \mathbf{w}_j\right|^2}\right)\right] . \tag{4.13}$$

In order to evaluate the performance metric, I use the following lemmas.

Lemma 4.4. *The expected value of the signal power term can be approximated as*

$$\mathbb{E}\left[\log_2\left(1 + P_i \left|\mathbf{h}_{ii}^H \mathbf{w}_i\right|^2\right)\right] \approx \log_2\left(1 + P_i \sigma_{ii}^2 e^{\psi(N_i - 1)}\right) + \log_2\left(1 - \gamma_i 2^{-\frac{B_{ii}}{N_i - 1}}\right) . \tag{4.14}$$

Proof. See Appendix B.4. □

Lemma 4.5. *The expected value of the noise power term can be approximated as*

$$\mathbb{E}\left[\log_2\left(1 + P_j \left|\mathbf{h}_{ij}^H \mathbf{w}_j\right|^2\right)\right] \approx \log_2\left(1 + P_j \sigma_{ij}^2 e^{\psi(N_j) - \psi(N_j - 1) + \psi(1)} \gamma_j 2^{-\frac{B_{ij}}{N_j - 1}}\right) . \tag{4.15}$$

Proof. See Appendix B.5. □

To characterize the effect of limited feedback, I derive an approximation of the average transmission rate of the i-th transmitter-receiver link:

Proposition 4.2. *The average transmission rate of the i-th transmitter-receiver link is approximately given by*

$$R_i \approx \log_2\left(1+P_i\sigma_{ii}^2 e^{\psi(N_i-1)}\right)+\log_2\left(1-\gamma_i 2^{-\frac{B_{ii}}{N_i-1}}\right) - \log_2\left(1+P_j\sigma_{ij}^2 e^{\psi(N_j)-\psi(N_j-1)+\psi(1)}\gamma_j 2^{-\frac{B_{ij}}{N_j-1}}\right) . \tag{4.16}$$

Proof. The average transmission rate of the link from transmitter i to receiver i in (4.13) can be approximated as below:

$$\begin{aligned} R_i \approx & \mathbb{E}\left[\log_2\left(\frac{1+P_i\left|\mathbf{h}_{ii}^H\mathbf{w}_i\right|^2}{1+P_j\left|\mathbf{h}_{ij}^H\mathbf{w}_j\right|^2}\right)\right] && (4.17)\\ = & \mathbb{E}\left[\log_2\left(1+P_i\left|\mathbf{h}_{ii}^H\mathbf{w}_i\right|^2\right)\right]-\mathbb{E}\left[\log_2\left(1+P_j\left|\mathbf{h}_{ij}^H\mathbf{w}_j\right|^2\right)\right] \\ \approx & \log_2\left(1+P_i\sigma_{ii}^2 e^{\psi(N_i-1)}\right)+\log_2\left(1-\gamma_i 2^{-\frac{B_{ii}}{N_i-1}}\right) \\ & -\log_2\left(1+P_j\sigma_{ij}^2 e^{\psi(N_j)-\psi(N_j-1)+\psi(1)}\gamma_j 2^{-\frac{B_{ij}}{N_j-1}}\right) . && (4.18)\end{aligned}$$

The approximation in (4.17) holds by eliminating the interference term $P_j\left|\mathbf{h}_{ij}^H\mathbf{w}_j\right|^2$ in the numerator under the high SNR assumption. Equation (4.18) follows from (4.14) and (4.15) in Lemma 4.4 and Lemma 4.5. □

We observe from (4.16) that the first term is the approximate average transmission rate under perfect feedback obtained from (4.10) in Proposition 4.1. Furthermore, the use of limited rate feedback produces two additional negative terms. The second term is due to the mismatch with the ideal beamformer for the direct transmitter, and the third term is due to interference leakage from the cross transmitter. Together, they constitute the throughput loss due to quantization errors. It is clear that $R_i \approx R_i^{ZF}$ and the rate loss is 0 if the feedback rate is infinite, *i.e.,* if $B_{ii}, B_{ij} \to \infty$.

The optimization problem is to maximize the average sum transmission rate by adjusting the

amount of transmit power and the feedback bit allocation for the direct and cross channels:

$$\max_{P_i,B_{ii},B_{ij}} R = \sum_{i=1}^{2} R_i \quad i \neq j;\ i,j \in \{1,2\} \tag{4.19}$$

$$\text{s.t.} \quad 0 < P_i \leq P_{max,i}$$

$$B_{ii} + B_{ij} = b_i$$

$$B_{ii}, B_{ij} \in \mathbb{Z}^+ \ ,$$

where $\mathbb{Z}^+$ is the set of non-negative integers. Relaxing the integer constraint, the following closed-form expression for the optimal solution holds when the transmitters have the same number of antennas:

Theorem 4.1. *Suppose the transmitters have the same number of antennas N. Then the optimal solution to (4.19) without the integer constraint, and the corresponding maximum average sum transmission rate is shown in (4.20)-(4.22), where $\gamma \triangleq \Gamma\left(\frac{N}{N-1}\right)$, $i \neq j$ and $i,j \in \{1,2\}$.*

$$P_i^* = P_{max,i} \tag{4.20}$$

$$B_{ii}^* = \begin{cases} (N-1)\log_2\left(\sqrt{\gamma^2 + \frac{2^{\frac{b_i}{N-1}}}{P_j^*\sigma_{ij}^2 e^{\psi(N)-\psi(N-1)+\psi(1)}}} + \gamma\right) & \text{if } P_{max,j} \geq \frac{1}{\sigma_{ij}^2 e^{\psi(N)-\psi(N-1)+\psi(1)}\left(2^{\frac{b_i}{N-1}}-2\gamma\right)} \\ b_i & \text{if } P_{max,j} < \frac{1}{\sigma_{ij}^2 e^{\psi(N)-\psi(N-1)+\psi(1)}\left(2^{\frac{b_i}{N-1}}-2\gamma\right)} \end{cases} \tag{4.21}$$

$$R^* \approx \begin{cases} \sum_{i=1}^{2} \log_2 \frac{\left(1+P_i^*\sigma_{ii}^2 e^{\psi(N-1)}\right)\frac{2^{\frac{b_i}{N-1}}}{P_j^*\sigma_{ij}^2 e^{\psi(N)-\psi(N-1)+\psi(1)}}}{\left(\sqrt{\gamma^2 + \frac{2^{\frac{b_i}{N-1}}}{P_j^*\sigma_{ij}^2 e^{\psi(N)-\psi(N-1)+\psi(1)}}} + \gamma\right)^2} & \text{if } P_{max,j} \geq \frac{1}{\sigma_{ij}^2 e^{\psi(N)-\psi(N-1)+\psi(1)}\left(2^{\frac{b_i}{N-1}}-2\gamma\right)} \\ \sum_{i=1}^{2} \log_2 \frac{\left(1+P_i^*\sigma_{ii}^2 e^{\psi(N-1)}\right)\left(1-\gamma 2^{-\frac{b_i}{N-1}}\right)}{\left(1+P_j^*\sigma_{ij}^2 e^{\psi(N)-\psi(N-1)+\psi(1)}\gamma\right)} & \text{if } P_{max,j} < \frac{1}{\sigma_{ij}^2 e^{\psi(N)-\psi(N-1)+\psi(1)}\left(2^{\frac{b_i}{N-1}}-2\gamma\right)} \end{cases} \tag{4.22}$$

Proof. To solve the optimization problem, first relax the integer constraint. Then, I use the sum of the average transmission rate in (4.16) as an approximation to the objective function.

Substituting $b_i - B_{ii}$ for B_{ij}, the optimization problem in (4.19) can be rewritten in standard form as

$$\begin{aligned}
\max_{P_i, B_{ii}} \quad & \sum_{i=1}^{2} \log_2 \frac{\left(1 + P_i \sigma_{ii}^2 e^{\psi(N_i - 1)}\right)\left(1 - \gamma_i 2^{-\frac{B_{ii}}{N_i - 1}}\right)}{1 + P_j \sigma_{ij}^2 e^{\psi(N_j) - \psi(N_j - 1) + \psi(1)} \gamma_j 2^{-\frac{b_i - B_{ii}}{N_j - 1}}} \\
\text{s.t.} \quad & -P_i < 0 \\
& P_i - P_{max,i} \leq 0 \\
& -B_{ii} \leq 0 \\
& B_{ii} - b_i \leq 0 \ .
\end{aligned}$$

This formulation leads to a convex optimization problem that can be solved by standard methods. Formulate the Lagrangian

$$\begin{aligned}
\mathcal{L} = \sum_{i=1}^{2} \Bigg[& \log_2 \frac{\left(1 + P_i \sigma_{ii}^2 e^{\psi(N_i - 1)}\right)\left(1 - \gamma_i 2^{-\frac{B_{ii}}{N_i - 1}}\right)}{1 + P_j \sigma_{ij}^2 e^{\psi(N_j) - \psi(N_j - 1) + \psi(1)} \gamma_j 2^{-\frac{b_i - B_{ii}}{N_j - 1}}} \\
& + \mu_{i1} P_i - \mu_{i2}(P_i - P_{max,i}) + \mu_{i3} B_{ii} - \mu_{i4}(B_{ii} - b_i)\Bigg] \ ,
\end{aligned} \tag{4.23}$$

where $\mu_{ik}, k = 1, ..., 4$ is the Lagrange multiplier. By taking the derivative of (4.23) with respect to P_i and B_{ii}, and applying Karush-Kuhn-Tucker (KKT) conditions, the closed-form solution is given in (4.20)-(4.22) when the transmitters have the same number of antennas.

□

Remark 4.1. *We see from Theorem 4.1 that transmitting with full power is the optimal transmission strategy. Note that the need for cooperation is evident since the cross channel variance σ_{ij}^2 and the maximum power of the other transmitter $P_{max,j}$ play a large role in determining the optimal feedback bit allocation B_{ii}^*.*

In practice, $P_{max,j}$ will exceed the threshold condition specified in Equations (4.21)-(4.22) when b_i is sufficiently large. Consequently, the optimization often produces a non-integer bit

allocation that is less than b_i. For the actual feedback link, search the integer values above and below B_{ii}^* to determine the integer bit allocation $\tilde{B}_{ii}^*$ and $\tilde{B}_{ij}^* = b_i - \tilde{B}_{ii}^*$ and the actual sum rate $\tilde{R}^*$ according to the expression in (4.16):

$$\tilde{R}^* \approx \sum_{i=1}^{2} \left[\log_2\left(1+P_i^* \sigma_{ii}^2 e^{\psi(N-1)}\right) + \log_2\left(1-\gamma 2^{-\frac{\tilde{B}_{ii}^*}{N-1}}\right) - \log_2\left(1 + P_j^* \sigma_{ij}^2 e^{\psi(N)-\psi(N-1)+\psi(1)} \gamma 2^{-\frac{\tilde{B}_{ij}^*}{N-1}}\right)\right] . \tag{4.24}$$

Remark 4.2. *Due to the relaxation of the integer constraint, $\tilde{R}^*$ is bounded above by R^*. As the transmit power grows to infinity, the average sum transmission rate with a fixed number of feedback bits is bounded above by*

$$\tilde{R}^* \leq \lim_{P_{max,i}\to\infty} R^* \approx \sum_{i=1}^{2} \log_2 \frac{\sigma_{ii}^2 2^{\frac{b_i}{N-1}}}{4\gamma^2 \sigma_{ij}^2 e^{\psi(N)-2\psi(N-1)+\psi(1)}} . \tag{4.25}$$

Unlike the perfect CSI scenario, where the average achievable transmission rate with the ZF beamformer can be made arbitrarily large by increasing the transmit power, the system performance of the limited rate feedback interference channel converges to a constant at high SNR.

4.4 Secrecy Rate Analysis

Next I consider the Gaussian interference channel where the two transmitters wish to send independent and confidential messages to their respective receivers. From the secrecy point of view, the metric of interest is the average achievable secrecy rate defined as in [29, 44]:

$$R_{sec,i} = \left\{ \mathbb{E}\left[\log_2\left(1 + \frac{P_i \left|\mathbf{h}_{ii}^H \mathbf{w}_i\right|^2}{1 + P_j \left|\mathbf{h}_{ij}^H \mathbf{w}_j\right|^2}\right) - \log_2\left(1 + P_i |\mathbf{h}_{ji}^H \mathbf{w}_i|^2\right)\right]\right\}^+ . \tag{4.26}$$

I approximate the average secrecy rate of the i-th transmitter-receiver link, similarly to Proposition 4.2.

Proposition 4.3. *The average secrecy rate of the i-th transmitter-receiver link is approximately given by*

$$\begin{aligned} R_{sec,i} \approx \Big\{ & \log_2\big(1 + P_i \sigma_{ii}^2 e^{\psi(N_i - 1)}\big) + \log_2\Big(1 - \gamma_i 2^{-\frac{B_{ii}}{N_i - 1}}\Big) \\ & - \log_2\left(1 + P_j \sigma_{ij}^2 e^{\psi(N_j) - \psi(N_j - 1) + \psi(1)} \gamma_j 2^{-\frac{B_{ij}}{N_j - 1}}\right) \\ & - \log_2\left(1 + P_i \sigma_{ji}^2 e^{\psi(N_i) - \psi(N_i - 1) + \psi(1)} \gamma_i 2^{-\frac{B_{ji}}{N_i - 1}}\right)\Big\}^+ . \end{aligned} \tag{4.27}$$

Proof. Similar to the proof of Proposition 4.2, the average secrecy rate of the i-th transmitter-receiver link in (4.26) can be approximated as

$$\begin{aligned} R_{sec,i} &= \left\{ R_i - \mathbb{E}\left[\log_2\left(1 + P_i |\mathbf{h}_{ji}^H \mathbf{w}_i|^2\right)\right]\right\}^+ \\ &\approx \Big\{ \log_2\left(1 + P_i \sigma_{ii}^2 e^{\psi(N_i - 1)}\right) + \log_2\left(1 - \gamma_i 2^{-\frac{B_{ii}}{N_i - 1}}\right) \\ &\quad - \log_2\left(1 + P_j \sigma_{ij}^2 e^{\psi(N_j) - \psi(N_j - 1) + \psi(1)} \gamma_j 2^{-\frac{B_{ij}}{N_j - 1}}\right) \\ &\quad - \log_2\left(1 + P_i \sigma_{ji}^2 e^{\psi(N_i) - \psi(N_i - 1) + \psi(1)} \gamma_i 2^{-\frac{B_{ji}}{N_i - 1}}\right)\Big\}^+ . \end{aligned}$$

The last step is based on the approximation in Proposition 4.2 and Lemma 4.5. □

The goal is to maximize the average sum secrecy rate by adjusting the amount of transmit power and the feedback bit allocation:

$$\begin{aligned} \max_{P_i, B_{ii}, B_{ij}} \quad & R_{sec} = \sum_{i=1}^{2} R_{sec,i} \qquad i \neq j;\ i, j \in \{1, 2\} \\ \text{s.t.} \quad & 0 < P_i \leq P_{max,i} \\ & B_{ii} + B_{ij} = b_i \\ & B_{ii}, B_{ij} \in \mathbb{Z}^+ . \end{aligned} \tag{4.28}$$

Theorem 4.2. *When the transmitters have the same number of antennas N, the optimal solution to (4.28) without the integer constraint, and the corresponding maximum average sum secrecy rate is characterized by (4.29)-(4.31), where $i \neq j$ and $i, j \in \{1, 2\}$.*

$$P_i^* = \min\left(P_{max,i}\ ,\ \frac{1}{\sigma_{ii}^2 e^{\psi(N-1)}}\left(\frac{\sigma_{ii}^2 2^{\frac{b_j}{N-1}}}{4\gamma^2\sigma_{ji}^2 e^{\psi(N)-2\psi(N-1)+\psi(1)}} - \frac{5}{4} + \sqrt{\left(\frac{\sigma_{ii}^2 2^{\frac{b_j}{N-1}}}{4\gamma^2\sigma_{ji}^2 e^{\psi(N)-2\psi(N-1)+\psi(1)}} - \frac{5}{4}\right)^2 - 1}\right)\right) \tag{4.29}$$

$$B_{ii}^* = (N-1)\log_2\left(\sqrt{\left(\frac{3}{4}\gamma\right)^2 + \frac{2^{\frac{b_i}{N-1}}}{2P_j^*\sigma_{ij}^2 e^{\psi(N)-\psi(N-1)+\psi(1)}}} + \frac{3}{4}\gamma\right) \tag{4.30}$$

$$R_{sec}^* \approx \sum_{i=1}^{2}\log_2 \frac{\left(1 + P_i^*\sigma_{ii}^2 e^{\psi(N-1)}\right)\left(\frac{2^{\frac{b_i}{N-1}}}{2P_j^*\sigma_{ij}^2 e^{\psi(N)-\psi(N-1)+\psi(1)}}\right)^2}{\left(\sqrt{\left(\frac{3}{4}\gamma\right)^2 + \frac{2^{\frac{b_i}{N-1}}}{2P_j^*\sigma_{ij}^2 e^{\psi(N)-\psi(N-1)+\psi(1)}}} - \frac{1}{4}\gamma\right)\left(\sqrt{\left(\frac{3}{4}\gamma\right)^2 + \frac{2^{\frac{b_i}{N-1}}}{2P_j^*\sigma_{ij}^2 e^{\psi(N)-\psi(N-1)+\psi(1)}}} + \frac{3}{4}\gamma\right)^3}\ . \tag{4.31}$$

Proof. After relaxing the integer constraint as in Theorem 4.1, I use the sum of the average secrecy rate in (4.27) as an approximation to the objective function. Substituting $b_i - B_{ii}$ for B_{ij}, the optimization problem in (4.28) becomes

$$\begin{aligned}
\max_{P_i, B_{ii}}\ & \sum_{i=1}^{2}\log_2 \frac{\left(1 + P_i\sigma_{ii}^2 e^{\psi(N_i-1)}\right)\left(1 - \gamma_i 2^{-\frac{B_{ii}}{N_i-1}}\right)}{\left(1 + P_j\sigma_{ij}^2 e^{\psi(N_j)-\psi(N_j-1)+\psi(1)}\gamma_j 2^{-\frac{b_i-B_{ii}}{N_j-1}}\right)^2} \\
\text{s.t.}\quad & -P_i < 0 \\
& P_i - P_{max,i} \leq 0 \\
& -B_{ii} \leq 0 \\
& B_{ii} - b_i \leq 0\ .
\end{aligned}$$

As before, consider the Lagrangian

$$\mathcal{L} = \sum_{i=1}^{2} \Bigg[\log_2 \frac{\left(1 + P_i \sigma_{ii}^2 e^{\psi(N_i-1)}\right)\left(1 - \gamma_i 2^{-\frac{B_{ii}}{N_i-1}}\right)}{\left(1 + P_j \sigma_{ij}^2 e^{\psi(N_j)-\psi(N_j-1)+\psi(1)} \gamma_j 2^{-\frac{b_i - B_{ii}}{N_j-1}}\right)^2}$$
$$+ \mu_{i1} P_i - \mu_{i2}(P_i - P_{max,i}) + \mu_{i3} B_{ii} - \mu_{i4}(B_{ii} - b_i) \Bigg] \ . \tag{4.32}$$

By equating the derivative of (4.32) to zero with respect to P_i and B_{ii}, and applying KKT conditions, I present the closed-form results in (4.29)-(4.31) when the transmitters have the same number of antennas. □

For the actual feedback link, search the integer values above and below B_{ii}^* in (4.30) to determine the integer bit allocation $\tilde{B}_{ii}^*$ and $\tilde{B}_{ij}^* = b_i - \tilde{B}_{ii}^*$. According to (4.27), the actual sum secrecy rate $\tilde{R}_{sec}^*$ is:

$$\tilde{R}_{sec}^* \approx \sum_{i=1}^{2} \Bigg\{ \log_2\left(1 + P_i^* \sigma_{ii}^2 e^{\psi(N-1)}\right) + \log_2\left(1 - \gamma 2^{-\frac{\tilde{B}_{ii}^*}{N-1}}\right)$$
$$- \log_2\left(1 + P_j^* \sigma_{ij}^2 e^{\psi(N)-\psi(N-1)+\psi(1)} \gamma 2^{-\frac{\tilde{B}_{ij}^*}{N-1}}\right)$$
$$- \log_2\left(1 + P_i^* \sigma_{ji}^2 e^{\psi(N)-\psi(N-1)+\psi(1)} \gamma 2^{-\frac{\tilde{B}_{ji}^*}{N-1}}\right) \Bigg\}^+ \ . \tag{4.33}$$

Remark 4.3. *The need for cooperation is again clear since the optimal B_{ii}^* requires knowledge of the power employed by the other transmitter P_j^* and the cross channel statistics σ_{ij}^2, but with a minor difference compared to (4.21) in Theorem 4.1. However, the transmission strategy behaves differently when secrecy is taken into consideration. Beyond a certain threshold, increasing transmit power hurts the secrecy performance due to the limited feedback. For sufficiently large b_j, (4.29) can be approximated as*

$$P_i^* \approx \min\left(P_{max,i} \ , \ \frac{2^{\frac{b_j}{N-1}}}{2\gamma^2 \sigma_{ji}^2 e^{\psi(N)-\psi(N-1)+\psi(1)}} \right) \ ,$$

where the transmit power control mainly depends on the cross channel statistics σ_{ji}^2.

4.5 Simulation Results

In this section, I validate the analytical results through Monte Carlo simulations. For a given channel realization, I use the numerical results in [49] to randomly generate the associated quantized feedback. Utilizing the statistics of RVQ codebooks, this method simulates the quantization procedure without generating an actual codebook, and reduces the computational complexity as the number of feedback bits grows. To facilitate exposition, I assume the system is symmetric. Transmitters have the same number of antennas and power constraints, *i.e.,* $N_1 = N_2 = 4$ and $P_{max,1} = P_{max,2} = P_{max}$. Unless stated otherwise, the limited feedback bandwidth for both receivers is set to 20 bits. The i.i.d. entries of the direct and cross channels are distributed as $\mathcal{CN}(0, \sigma_d^2)$ and $\mathcal{CN}(0, \sigma_c^2)$, *i.e.,* $\sigma_{11}^2 = \sigma_{22}^2 = \sigma_d^2$ and $\sigma_{12}^2 = \sigma_{21}^2 = \sigma_c^2$. In each figure, the values of σ_d^2 and σ_c^2 will be depicted. All results are based on averages obtained over 1000 independent channel realizations.

Figure 4.2 compares the numerical evaluation of the average sum transmission rate according to (4.13) with the approximate average sum transmission rate based on (4.16) as a function of the number of bits allocated to the direct channel (B_{ii}). The figure shows the results when the transmit power is fixed at $P_i = 10$ dB and 20 dB and for various channel conditions. The approximate average sum transmission rate is especially accurate at the maximum, where B_{ii} and B_{ij} are properly allocated. Most importantly, the peaks of the two curves coincide exactly.

I show the average sum transmission rate as a function of transmit power in Figure 4.3. The dashed lines indicate the numerical simulations with perfect CSI and with quantization under the optimal feedback allocation. The solid lines represent the respective approximate sum

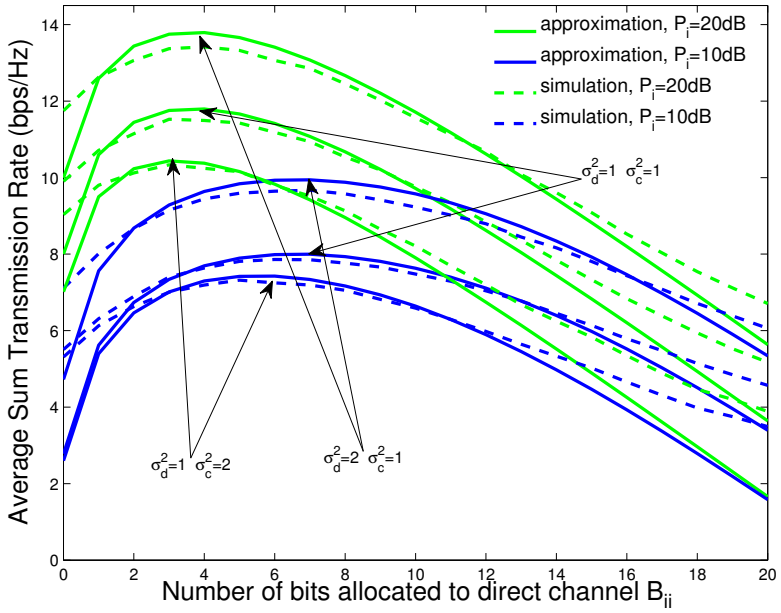

Figure 4.2: Accuracy of the approximate average sum transmission rate for the fixed transmit power $P_i = 10$ dB and 20 dB.

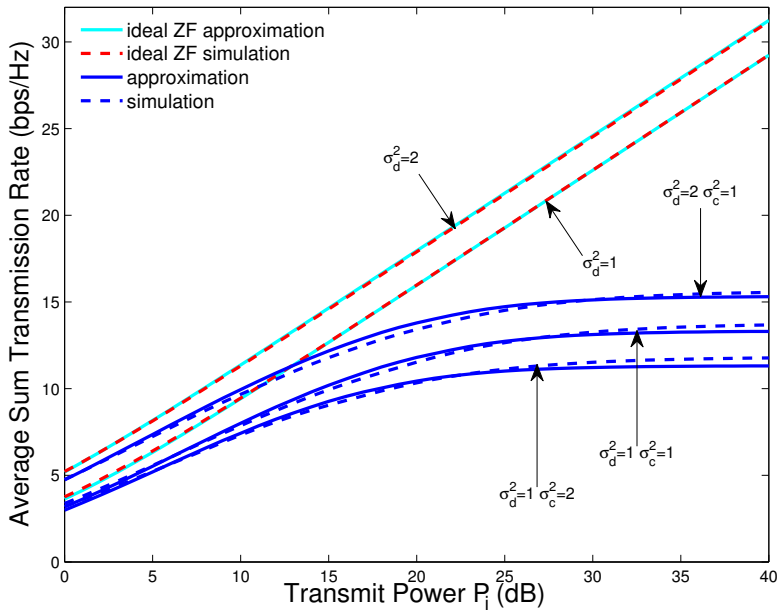

Figure 4.3: Approximate average sum transmission rate versus transmit power.

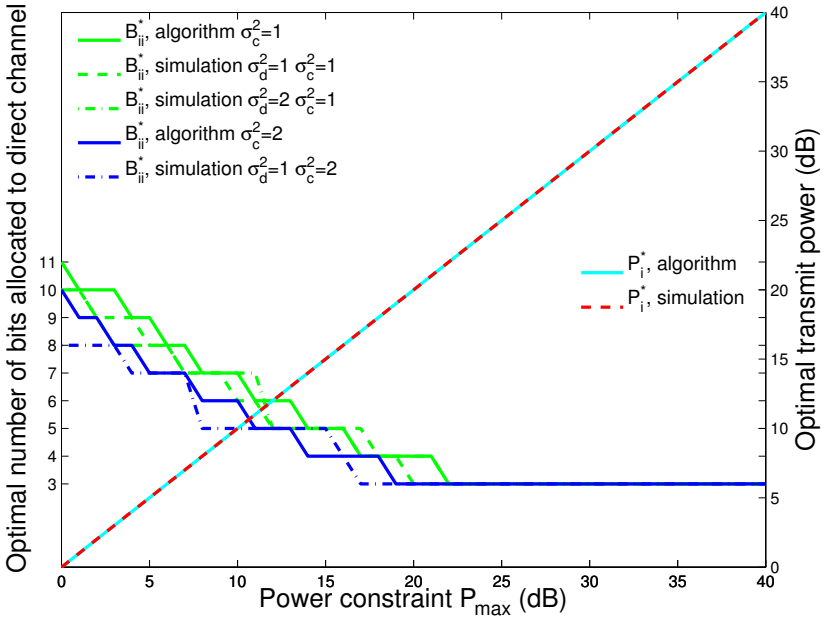

Figure 4.4: Accuracy of the transmit power control and feedback bit allocation algorithm versus transmit power constraint.

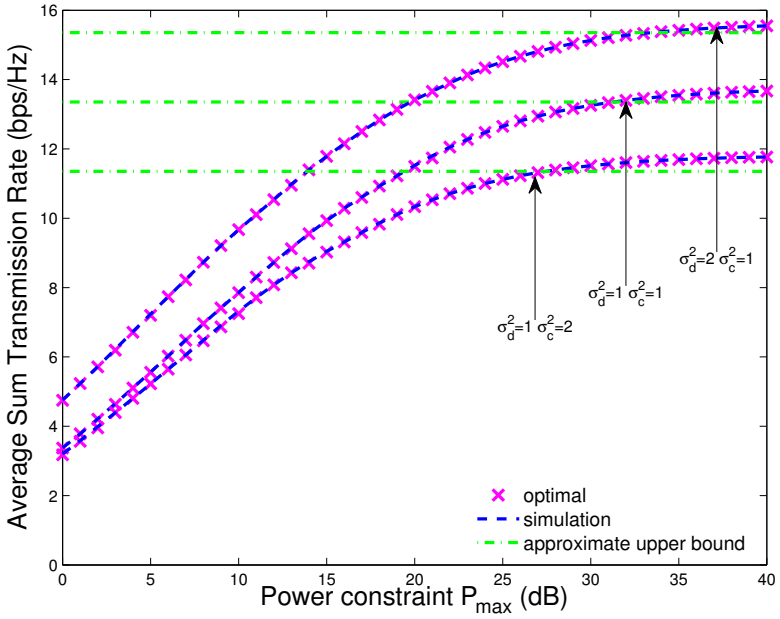

Figure 4.5: Accuracy of the actual average sum transmission rate versus transmit power constraint.

rates based on (4.10) and (4.16). The figure further verifies the accuracy of the approximate sum rates in terms of the transmit power. In addition, it indicates that the system throughput using ideal ZF beamforming grows without bound as SNR increases. However, the limited feedback system is interference-limited and the sum rate converges to an upper limit.

Figure 4.4 compares the optimal transmit power and feedback bit allocation results (solid lines) in (4.20) and (4.21) to the optimal results obtained from a grid search method (dashed lines). Figure 4.5 demonstrates the accuracy of the average sum transmission rate with respect to the transmit power constraint P_{max}. The actual average sum transmission rate based on the optimal results (P_i^* and $\tilde{B}_{ii}^*$) in Figure 4.4 is essentially identical to the best possible average sum transmission rate obtained from the simulations. Figures 4.4 and 4.5 illustrate that transmitting with full power achieves the maximum average sum transmission rate. Additionally, I depict the sum rate achieved using infinite transmit power (dash-dot line) given by (4.25). We see that the approximate upper bound in (4.25) roughly predicts the limiting throughput.

Figures 4.6 and 4.7 plot the numerical evaluation of the average sum secrecy rate according to (4.26) along with the approximate average sum secrecy rate based on (4.27) as a function of B_{ii} and P_i respectively. Both figures verify the accuracy of the approximate average sum secrecy rate, which is used to identify the optimal results. The peak values of the approximate rate coincide with the peak values of the actual average sum secrecy rate. By using the approximation, we can find the optimal system parameters without knowing the exact rate. Furthermore, Figure 4.7 illustrates the trade-off associated with increasing transmit power in a scenario with limited feedback and hence inaccurate CSI; beyond a certain point increasing transmit power decreases the secrecy performance.

Similarly, Figures 4.8 and 4.9 present the accuracy of the optimal parameters obtained from Theorem 4.2 with respect to P_{max}. The figures indicate that the transmit power of $P_i = 16$ dB and 19 dB (vertical lines) derived in (4.29) attains the maximum of the average sum secrecy

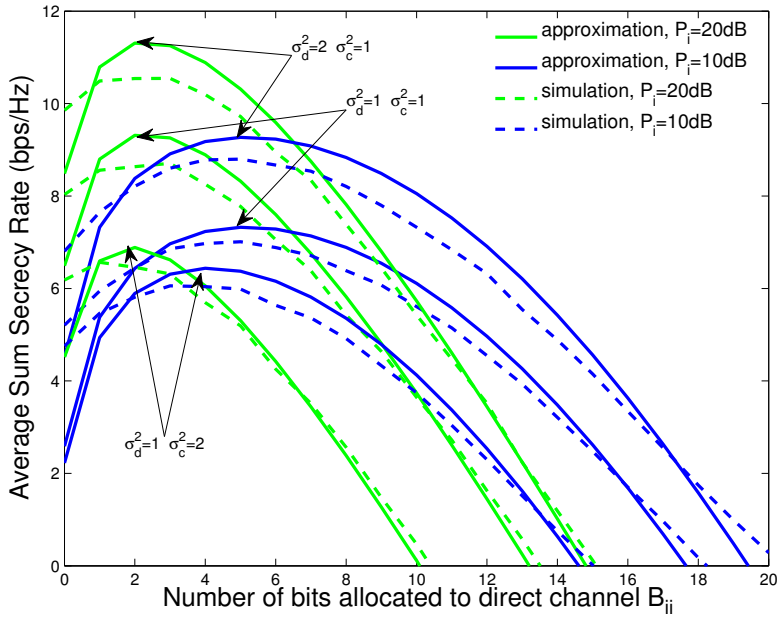

Figure 4.6: Accuracy of the approximate average sum secrecy rate for the fixed transmit power $P_i = 10$ dB and 20 dB.

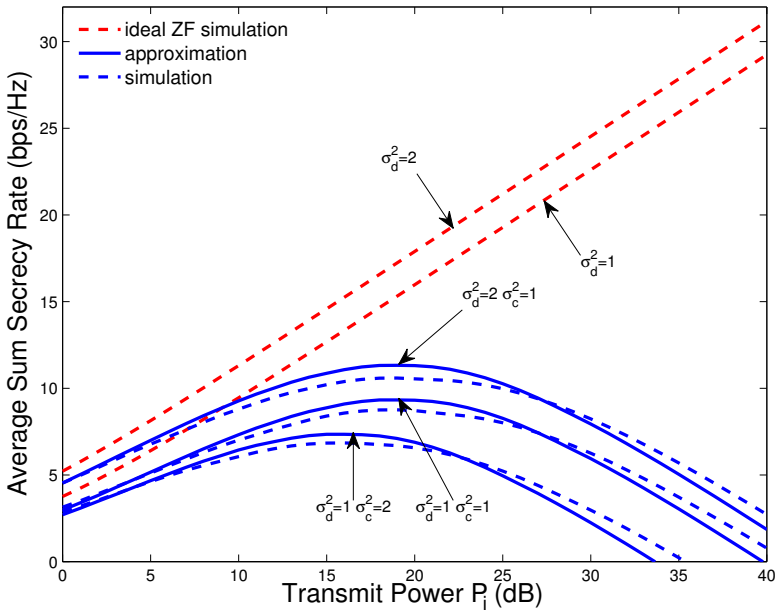

Figure 4.7: Approximate average sum secrecy rate versus transmit power.

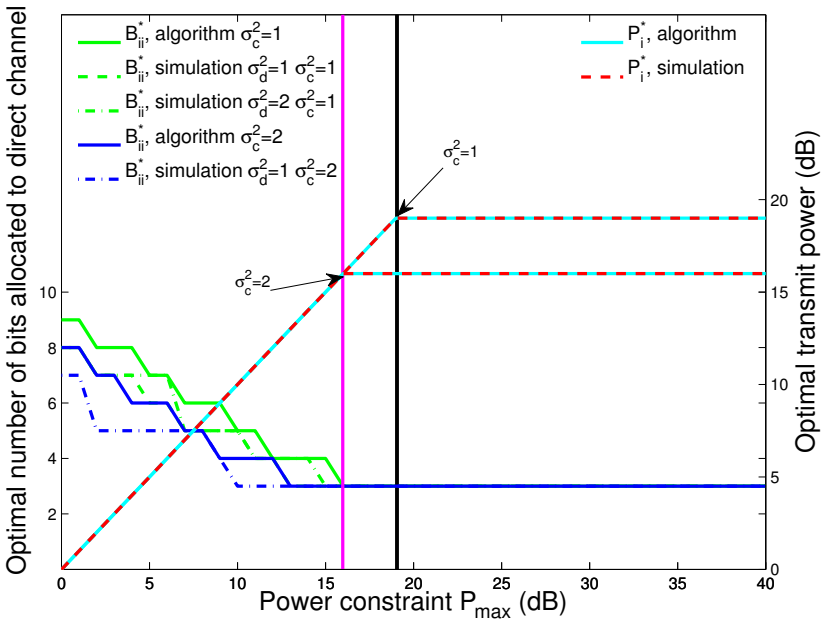

Figure 4.8: Accuracy of the transmit power control and feedback bit allocation algorithm versus transmit power constraint.

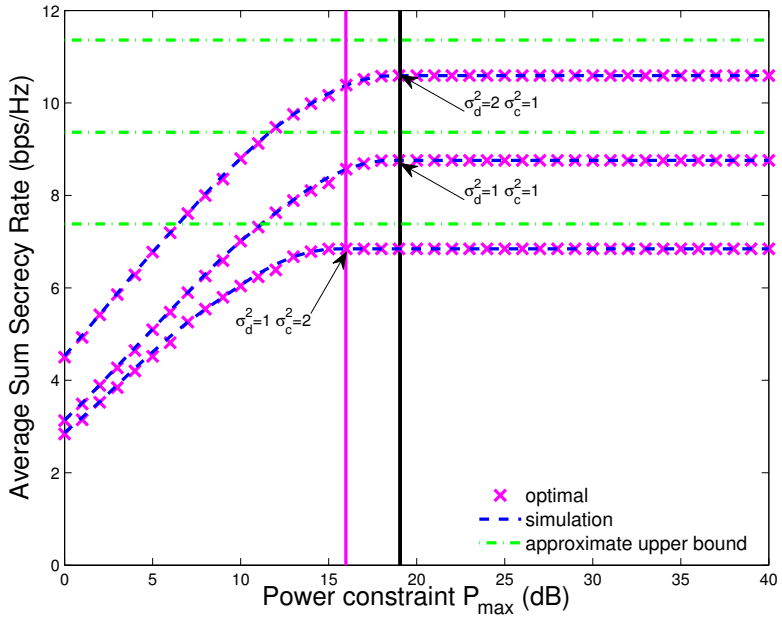

Figure 4.9: Accuracy of the actual average sum secrecy rate versus transmit power constraint.

rate for the corresponding channel conditions. Increasing the power constraint beyond this threshold does not improve the secrecy performance of the system. In Figure 4.9, I also illustrate the approximate upper bound given by (4.31).

4.6 Summary

This chapter has considered the allocation of transmit power and feedback bits in a co-operative two-user MISO interference channel with limited rate feedback. I analyzed two scenarios regarding the system throughput and secrecy performance. For each case, I derived an approximate rate of the transmission link and maximized the sum rate to find closed-form expressions for the optimal transmit power and feedback bit allocation. Moreover, I showed that the system throughput is interference-limited at high SNR under the assumption of limited feedback. There exists a critical value of the power constraint above which increasing transmit power reduces the sum secrecy rate. Simulation results justify the approximations used in this chapter, and demonstrate how the proper choice of the transmit power and feedback bit allocation can dramatically enhance the system performance.

Chapter 5

Conclusion

In this dissertation, I study strategies for enhanced secrecy in wireless communication systems with limited rate feedback and transmitter cooperation. The two-user Gaussian channel model is considered under different scenarios.

The first scenario has considered power and bit allocation strategies for enhanced secrecy in a limited rate feedback MIMO wiretap channel involving a cooperative jammer. Two cases are studied, one where no information about the eavesdroppers is available, and one where statistical channel state information is available. With no information about the eavesdroppers, I showed how to choose the allocation of feedback bits to the transmitter and helper in order to maximize the amount of jamming power available to interfere with the eavesdroppers, subject to maintaining a lower bound on the target rate for the desired link. A closed-form solution was found for the special case where the transmitter and jammer have the same number of antennas. For the case of statistical CSI, I derived an approximate lower bound on the average secrecy rate, and again found a closed-form solution for the feedback bit allocation that maximizes this lower bound for an equal number of transmit antennas. Optimization of the transmit power in this case requires a 2-dimensional numerical search.

Simulation results indicate the accuracy of the approximations used in this chapter and demonstrate how proper choice of the feedback bit allocation can dramatically enhance the security provided by the cooperative jammer.

The second scenario has considered the allocation of transmit power and feedback bits in a cooperative two-user MISO interference channel with limited rate feedback. I analyzed two cases regarding the system throughput and secrecy performance. For each case, I derived an approximate rate of the transmission link and maximized the sum rate to find closed-form expressions for the optimal transmit power and feedback bit allocation. Moreover, I showed that the system throughput is interference-limited at high SNR under the assumption of limited feedback. There exists a critical value of the power constraint above which increasing transmit power reduces the sum secrecy rate. Again, simulation results justify the approximations adopted in this chapter, and demonstrate how the proper choice of the transmit power and feedback bit allocation can dramatically enhance the system performance.

Besides the scenarios discussed in this dissertation, a natural direction to generalize my research is to consider the two-user Gaussian channel with an untrusted user. A trust degree $\alpha \in (0, 1)$ can be assigned to this user who is eavesdropping with probability α. For two extreme cases, if $\alpha = 0$, I studied in Chapter 3 that one receiver is a known eavesdropper, and a second transmitter is used to send a cooperative jamming signal. If $\alpha = 1$, both receivers are intended, and each transmitter sends data signal to its paired user as discussed in Chapter 4. It is insightful to analyze the transmission strategies for this generalized model, *i.e.* allocating feedback bits, designing beamformers, and incorporating cooperative jamming techniques. An extension of the two-user MISO interference channel to MIMO scenario is also of great interest.

Bibliography

[1] I. Csiszar and J. Korner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Orlando, FL, 1981.

[2] C. E. Shannon, "Communication theory of secrecy systems," *Bell Syst. Tech. J.*, vol. 28, pp. 656–715, 1949.

[3] A. D. Wyner, "The Wiretap Channel," *The Bell System Technical Journal*, vol. 54, pp. 1355–1387, 1975.

[4] S. Leung-Yan-Cheong and M. Hellman, "The gaussian wire-tap channel," *IEEE Trans. Info. Theory*, vol. 24, no. 4, pp. 451–456, Jul. 1978.

[5] I. Csiszar and J. Korner, "Broadcast channels with confidential messages," *IEEE Trans. Info. Theory*, vol. 24, no. 3, pp. 339–348, May 1978.

[6] A. A. H. J. E. Hershey and R. Yarlagadda, "Unconventional cryptographic keying variable management," *IEEE Trans. Commun.*, vol. 43, pp. 3–6, Jan. 1995.

[7] J. E. H. A. A. Hassan, W. E. Stark and S. Chennakeshu, "Cryptographic key agreement for mobile radio," in *Digital Signal Processing*, vol. 6, San Diego, CA, 1996, pp. 207–212.

[8] H. Koorapaty, A. A. Hassan, and S. Chennakeshu, "Secure information transmission for mobile radio," *IEEE J. COML*, vol. 4, no. 2, pp. 52–55, Feb. 2000.

[9] U. M. Maurer and S. Wolf, "Unconditionally secure key agreement and the intrinsic conditional information," *IEEE Trans. Info. Theory*, vol. 45, no. 2, pp. 499–514, Mar. 1999.

[10] L. Lai, H. Gamal, and H. Poor, "The wiretap channel with feedback: Encryption over the channel," *IEEE Trans. Info. Theory*, vol. 54, no. 11, pp. 5059–5067, Nov. 2008.

[11] S. Goel and R. Negi, "Guaranteeing secrecy using artificial noise," *IEEE Trans. Info. Theory*, vol. 7, no. 6, pp. 2180–2189, Jun. 2008.

[12] A. Swindlehurst, "Fixed SINR solutions for the MIMO wiretap channel," in *Proc. 2009 IEEE ICASSP*, Apr. 2009, pp. 2437 –2440.

[13] A. Khisti and G. W. Wornell, "Secure transmission with multiple antennas I: The MISOME wiretap channel," *IEEE Trans. Info. Theory*, vol. 56, no. 7, pp. 3088–3104, Jul. 2010.

[14] A. Mukherjee, A. Fakoorian, J. Huang, and A. Swindlehurst, "MIMO signal processing algorithms for enhanced physical-layer security," in *Physical Layer Security in Wireless Communications*, X. Zhou, L. Song, and Y. Zhang, Eds. CRC Press, 2013, ch. 6.

[15] Y.-W. Hong, P.-C. Lan, and C.-C. Kuo, "Enhancing physical-layer secrecy in multi-antenna wireless systems: An overview of signal processing approaches," *Signal Processing Magazine, IEEE*, vol. 30, no. 5, pp. 29–40, Sep. 2013.

[16] M. Jorgensen, B. Yanakiev, G. Kirkelund, P. Popovski, H. Yomo, and T. Larsen, "Shout to secure: Physical-layer wireless security with known interference," in *IEEE GlobeCom*, Nov. 2007, pp. 33–38.

[17] P. Popovski and O. Simeone, "Wireless secrecy with infrastructure-aided cooperation," in *IEEE Info. Theory Workshop*, May 2008, pp. 159–163.

[18] X. Tang, R. Liu, P. Spasojevic, and H. Poor, "The Gaussian wiretap channel with a helping interferer," in *IEEE Int'l Symp. on Info. Theory*, Jul. 2008, pp. 389–393.

[19] L. Lai and H. El Gamal, "The relay-eavesdropper channel: Cooperation for secrecy," *IEEE Trans. Info. Theory*, vol. 54, no. 9, pp. 4005–4019, Sep. 2008.

[20] E. Tekin and A. Yener, "The general Gaussian multiple-access and two-way wiretap channels: Achievable rates and cooperative jamming," *IEEE Trans. Info. Theory*, vol. 54, no. 6, pp. 2735–2751, Jun. 2008.

[21] J. Huang and A. Swindlehurst, "Cooperative jamming for secure communications in MIMO relay networks," *IEEE Trans. on Signal Processing*, vol. 59, no. 10, pp. 4871–4884, 2011.

[22] J. Wang and A. Swindlehurst, "Cooperative jamming in MIMO ad-hoc networks," in *Proc. 43rd Asilomar Conf. on Signals, Systems and Computers*, 2009, pp. 1719–1723.

[23] A. Fakoorian and A. Swindlehurst, "Solutions for the MIMO gaussian wiretap channel with a cooperative jammer," *IEEE Trans. on Signal Processing*, vol. 59, no. 10, pp. 5013–5022, 2011.

[24] D. Love, J. Heath, R.W., W. Santipach, and M. Honig, "What is the value of limited feedback for MIMO channels?" *Communications Magazine, IEEE*, vol. 42, no. 10, pp. 54 – 59, Oct. 2004.

[25] D. Love and R. Heath, "Limited feedback unitary precoding for spatial multiplexing systems," *IEEE Trans. Info. Theory*, vol. 51, no. 8, pp. 2967–2976, 2005.

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

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

[28] W. Santipach and M. L. Honig, "Capacity of a multiple-antenna fading channel with a quantized precoding matrix," *IEEE Trans. Info. Theory*, vol. 55, no. 3, pp. 1218–1234, Mar. 2009.

[29] S.-C. Lin, T.-H. Chang, Y.-L. Liang, Y. Hong, and C.-Y. Chi, "On the impact of quantized channel feedback in guaranteeing secrecy with artificial noise: The noise leakage problem," *IEEE Trans. Wireless Commun.*, vol. 10, no. 3, pp. 901 –915, Mar. 2011.

[30] S. Bashar, Z. Ding, and G. Li, "On secrecy of codebook-based transmission beamforming under receiver limited feedback," *IEEE Trans. Wireless Comm.*, vol. 10, no. 4, pp. 1212 –1223, Apr. 2011.

[31] M. Pei, A. Swindlehurst, D. Ma, and J. Wei, "Adaptive limited feedback for MISO wiretap channels with cooperative jamming," *Signal Processing, IEEE Transactions on*, vol. 62, no. 4, pp. 993–1004, Feb. 2014.

[32] R. Liu, I. Maric, P. Spasojevic, and R. D. Yates, "Discrete memoryless interference and broadcast channels with confidential messages: Secrecy rate regions," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2493–2507, Jun. 2008.

[33] Y. Liang, A. Somekh-Baruch, H. V. Poor, S. Shamai, and S. Verdu, "Capacity of cognitive interference channels with and without secrecy," *IEEE Trans. Inf. Theory*, vol. 55, no. 2, pp. 604–619, Feb. 2009.

[34] O. O. Koyluoglu, H. E. Gamal, L. Lai, and H. V. Poor, "Interference alignment for secrecy," *IEEE Trans. Inf. Theory*, vol. 57, no. 6, pp. 3323–3332, Jun. 2011.

[35] S. A. A. Fakoorian and A. L. Swindlehurst, "MIMO interference channel with confidential messages: Achievable secrecy rates and precoder design," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 3, pp. 640–649, Sep. 2011.

[36] S. I. Bross, Y. Steinberg, and S. Tinguely, "The discrete memoryless interference channel with one-sided generalized feedback," *IEEE Trans. Inf. Theory*, vol. 59, no. 7, pp. 4171–4191, Jul. 2013.

[37] S. I. Bross, "The discrete memoryless interference channel with one-sided generalized feedback and secrecy," *IEEE Trans. Inf. Theory*, vol. 63, no. 5, pp. 2710–2725, May 2017.

[38] J. Zhu, J. Mo, and M. Tao, "Cooperative secret communication with artificial noise in symmetric interference channel," *IEEE Commun. Lett.*, vol. 14, no. 10, pp. 885–887, Oct. 2010.

[39] J. Ni, K. K. Wong, Z. Fei, C. Xing, H. Chen, K. F. Tong, and J. Kuang, "Secrecy-rate balancing for two-user MISO interference channels," *IEEE Wireless Commun. Lett.*, vol. 3, no. 1, pp. 6–9, Feb. 2014.

[40] O. O. Koyluoglu and H. E. Gamal, "Cooperative encoding for secrecy in interference channels," *IEEE Trans. Inf. Theory*, vol. 57, no. 9, pp. 5682–5694, Sep. 2011.

[41] P. Mohapatra and C. R. Murthy, "On the capacity of the two-user symmetric interference channel with transmitter cooperation and secrecy constraints," *IEEE Trans. Inf. Theory*, vol. 62, no. 10, pp. 5664–5689, Oct. 2016.

[42] B. M. Hochwald, T. L. Marzetta, and B. Hassibi, "Space-time autocoding," *IEEE Trans. Info. Theory*, vol. 47, pp. 2761–2781, 1999.

[43] B. Hochwald, T. Marzetta, and V. Tarokh, "Multiple-antenna channel hardening and its implications for rate feedback and scheduling," *IEEE Trans. Info. Theory*, vol. 50, no. 9, pp. 1893–1909, 2004.

[44] X. Yang and A. L. Swindlehurst, "Limited rate feedback in a MIMO wiretap channel with a cooperative jammer," *IEEE Trans. Signal Process.*, vol. 64, no. 18, pp. 4695–4706, Sep. 2016.

[45] E. G. Larsson, D. Danev, and E. A. Jorswieck, "Asymptotically optimal transmit strategies for the multiple antenna interference channel," in *Proc. IEEE 46th Ann. Allerton Conf. Commun., Contr., Comput.*, Sep. 2008, pp. 708–714.

[46] W. Dai, Y. Liu, and B. Rider, "Quantization bounds on Grassmann manifolds and applications to MIMO communications," *IEEE Trans. Inf. Theory*, vol. 54, no. 3, pp. 1108 –1123, Mar. 2008.

[47] P. K. Gopala, L. Lai, and H. El-Gamal, "On the secrecy capacity of fading channels," *IEEE Trans. Info. Theory*, vol. 54, no. 10, pp. 4687–4698, Oct. 2008.

[48] A. Mukherjee and A. Swindlehurst, "Fixed-rate power allocation strategies for enhanced secrecy in MIMO wiretap channels," in *Proc. Workshop on Signal Processing Advances in Wireless Communications*, Jun. 2009, pp. 344–348.

[49] N. Ravindran and N. Jindal, "Limited feedback-based block diagonalization for the MIMO broadcast channel," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 8, pp. 1473 –1482, Oct. 2008.

[50] R. Ahlswede, "The capacity region of a channel with two senders and two receivers," *Ann. Prob.*, vol. 2, no. 5, pp. 805–814, Oct. 1974.

[51] A. Carleial, "Interference channels," *IEEE Trans. Inf. Theory*, vol. 24, no. 1, pp. 60–70, Jan. 1978.

[52] T. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Trans. Inf. Theory*, vol. 27, no. 1, pp. 49–60, Jan. 1981.

[53] M. Costa, "On the Gaussian interference channel," *IEEE Trans. Inf. Theory*, vol. 31, no. 5, pp. 607–615, Sep. 1985.

[54] G. Kramer, "Outer bounds on the capacity of Gaussian interference channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 3, pp. 581–586, Mar. 2004.

[55] X. Shang, B. Chen, and M. J. Gans, "On the achievable sum rate for MIMO interference channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 4313–4320, Sep. 2006.

[56] S. A. Jafar and M. J. Fakhereddin, "Degrees of freedom for the MIMO interference channel," *IEEE Trans. Inf. Theory*, vol. 53, no. 7, pp. 2637–2642, Jul. 2007.

[57] V. S. Annapureddy and V. V. Veeravalli, "Sum capacity of MIMO interference channels in the low interference regime," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 2565–2581, May 2011.

[58] S. W. Peters and R. W. Heath, "Cooperative algorithms for MIMO interference channels," *IEEE Trans. Veh. Technol.*, vol. 60, no. 1, pp. 206–218, Jan. 2011.

[59] J. S. Kim, S. H. Moon, S. R. Lee, and I. Lee, "A new channel quantization strategy for MIMO interference alignment with limited feedback," *IEEE Trans. Wireless Commun.*, vol. 11, no. 1, pp. 358–366, Jan. 2012.

[60] K. Huang and R. Zhang, "Cooperative precoding with limited feedback for MIMO interference channels," *IEEE Trans. Wireless Commun.*, vol. 11, no. 3, pp. 1012–1021, Mar. 2012.

[61] S. Cho, K. Huang, D. K. Kim, V. K. N. Lau, H. Chae, H. Seo, and B. H. Kim, "Feedback-topology designs for interference alignment in MIMO interference channels," *IEEE Trans. Signal Process.*, vol. 60, no. 12, pp. 6561–6575, Dec. 2012.

[62] S. Cho, K. Huang, D. Kim, and H. Seo, "Interference alignment for uplink cellular systems with limited feedback," *IEEE Commun. Lett.*, vol. 16, no. 7, pp. 960–963, Jul. 2012.

[63] Y. Zhang and R. S. Cheng, "High SNR performance of antenna selection in MISO interference channel with zero-forcing and limited feedback," in *Proc. IEEE Wireless Commun. Networking Conf. (WCNC)*, Apr. 2012, pp. 858–862.

[64] R. T. Krishnamachari and M. K. Varanasi, "Interference alignment under limited feedback for MIMO interference channels," *IEEE Trans. Signal Process.*, vol. 61, no. 15, pp. 3908–3917, Aug. 2013.

[65] Y. Zhang and R. S. Cheng, "On the design of interference alignment scheme for multi-user MIMO with limited feedback," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2013, pp. 5646–5650.

[66] J. Zhang and J. G. Andrews, "Adaptive spatial intercell interference cancellation in multicell wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 9, pp. 1455–1468, Dec. 2010.

[67] R. Bhagavatula and R. W. Heath, "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.

[68] A. Mukherjee, S. A. A. Fakoorian, J. Huang, and A. L. Swindlehurst, "Principles of physical layer security in multiuser wireless networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 3, pp. 1550–1573, 3rd Quart. 2014.

[69] E. Bjornson, D. Hammarwall, and B. Ottersten, "Exploiting quantized channel norm feedback through conditional statistics in arbitrarily correlated MIMO systems," *IEEE Trans. Signal Process.*, vol. 57, no. 10, pp. 4027–4041, Oct. 2009.

[70] I. S. Gradshteyn, I. M. Ryzhik, and A. Jeffrey, *Table of Integrals, Series, and Products.* Academic Press, 1980.

[71] T. Yoo, N. Jindal, and A. Goldsmith, "Multi-antenna downlink channels with limited feedback and user selection," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 7, pp. 1478–1491, Sep. 2007.

Appendix A

Appendix for Chapter 3

A.1 Proof of Lemma 3.2

Using the definition of chordal distance in (3.8), the average distortion associated with the codebooks $\mathcal{C}_i$ is given by

$$\begin{aligned} D_i &= \mathbb{E}\left[d^2\left(\mathbf{V}_{ij}, \widehat{\mathbf{V}}_{ij}\right)\right] \\ &= \mathbb{E}\left[M_i - tr\left(\mathbf{V}_{ij}^H \widehat{\mathbf{V}}_{ij} \widehat{\mathbf{V}}_{ij}^H \mathbf{V}_{ij}\right)\right] \\ &= M_i - \mathbb{E}\left[tr\left(\mathbf{V}_{ij}^H \widehat{\mathbf{V}}_{ij} \widehat{\mathbf{V}}_{ij}^H \mathbf{V}_{ij}\right)\right] \\ &= M_i - \mathbb{E}\left[tr\left(\widehat{\mathbf{V}}_{ij}^H \mathbf{V}_{ij} \mathbf{V}_{ij}^H \widehat{\mathbf{V}}_{ij}\right)\right] \ . \end{aligned}$$

From the analysis in [49] we know that $\mathbb{E}\left[\mathbf{V}_{ij}^H\widehat{\mathbf{V}}_{ij}\widehat{\mathbf{V}}_{ij}^H\mathbf{V}_{ij}\right]$ and $\mathbb{E}\left[\widehat{\mathbf{V}}_{ij}^H\mathbf{V}_{ij}\mathbf{V}_{ij}^H\widehat{\mathbf{V}}_{ij}\right]$ are multiples of the identity matrix. Therefore,

$$\begin{aligned}\mathbb{E}\left[\mathbf{V}_{ij}^H\widehat{\mathbf{V}}_{ij}\widehat{\mathbf{V}}_{ij}^H\mathbf{V}_{ij}\right] &= \mathbb{E}\left[\widehat{\mathbf{V}}_{ij}^H\mathbf{V}_{ij}\mathbf{V}_{ij}^H\widehat{\mathbf{V}}_{ij}\right] \\ &= \frac{\mathbb{E}\left[tr\left(\mathbf{V}_{ij}^H\widehat{\mathbf{V}}_{ij}\widehat{\mathbf{V}}_{ij}^H\mathbf{V}_{ij}\right)\right]}{M_i}\mathbf{I}_{M_i} \\ &= \left(1-\frac{D_i}{M_i}\right)\mathbf{I}_{M_i}\ .\end{aligned}$$

A.2 Proof of Theorem 3.1

First, the average rate of the main channel $\overline{R_b}(d)$ in (3.19) can be bounded below by (A.1)-(A.3).

$$\overline{R_b}(d) \geq \mathbb{E}_{\mathbf{H}_{ba},\mathbf{H}_{bh},\mathbf{W}_a,\mathbf{W}_h}\left[\log_2\frac{\left|\mathbf{I}_d+\frac{P_s}{d}\mathbf{\Lambda}_{a1}\mathbf{V}_{a1}^H\mathbf{W}_a\mathbf{W}_a^H\mathbf{V}_{a1}\mathbf{\Lambda}_{a1}^H\right|}{\left|\mathbf{I}_d+\frac{P_i}{N_h-d}\mathbf{U}_h\mathbf{\Lambda}_{h1}\mathbf{V}_{h1}^H\mathbf{W}_h\mathbf{W}_h^H\mathbf{V}_{h1}\mathbf{\Lambda}_{h1}^H\mathbf{U}_h^H\right|}\right] \tag{A.1}$$

$$= \mathbb{E}_{\mathbf{H}_{ba},\mathbf{H}_{bh},\mathbf{W}_a,\mathbf{W}_h}\left[\log_2\frac{\left|\mathbf{I}_d+\frac{P_s}{d}\mathbf{\Lambda}_{a1}^2\mathbf{V}_{a1}^H\mathbf{W}_a\mathbf{W}_a^H\mathbf{V}_{a1}\right|}{\left|\mathbf{I}_d+\frac{P_i}{N_h-d}\mathbf{\Lambda}_{h1}^2\mathbf{V}_{h1}^H\mathbf{W}_h\mathbf{W}_h^H\mathbf{V}_{h1}\right|}\right] \tag{A.2}$$

$$\geq \mathbb{E}_{\mathbf{H}_{ba},\mathbf{H}_{bh},\mathbf{W}_a,\mathbf{W}_h}\left[\log_2\frac{\left|\mathbf{I}_d+\frac{P_s}{d}\mathbf{\Lambda}_{a1}^2\right|\left|\mathbf{V}_{a1}^H\mathbf{W}_a\mathbf{W}_a^H\mathbf{V}_{a1}\right|}{\left|\mathbf{I}_d+\frac{P_i}{N_h-d}\mathbf{\Lambda}_{h1}^2\mathbf{V}_{h1}^H\mathbf{W}_h\mathbf{W}_h^H\mathbf{V}_{h1}\right|}\right]\ . \tag{A.3}$$

Equation (A.1) holds by eliminating the term $\frac{P_i}{N_h-d}\mathbf{U}_h\mathbf{\Lambda}_{h1}\mathbf{V}_{h1}^H\mathbf{W}_h\mathbf{W}_h^H\mathbf{V}_{h1}\mathbf{\Lambda}_{h1}^H\mathbf{U}_h^H$ which is positive semidefinite in the numerator. Equation (A.2) follows from Sylvester's determinant theorem, and Equation (A.3) holds due to the fact that $\mathbf{I}_d \succeq \mathbf{V}_{a1}^H\mathbf{W}_a\mathbf{W}_a^H\mathbf{V}_{a1}$. Note that this lower bound is asymptotically tight under the high SNR assumption. Then, Equation (A.3)

can be rewritten as

$$\begin{aligned}\overline{R_b}(d) \geq & \mathbb{E}_{\mathbf{H}_{ba}}\left[\log_2\left|\mathbf{I}_d + \frac{P_s}{d}\mathbf{\Lambda}_{a1}^2\right|\right] \\ & - \mathbb{E}_{\mathbf{H}_{bh},\mathbf{W}_h}\left[\log_2\left|\mathbf{I}_d + \frac{P_i}{N_h - d}\mathbf{\Lambda}_{h1}^2\mathbf{V}_{h1}^H\mathbf{W}_h\mathbf{W}_h^H\mathbf{V}_{h1}\right|\right] \\ & + \mathbb{E}_{\mathbf{H}_{ba},\mathbf{W}_a}\left[\log_2\left|\mathbf{V}_{a1}^H\mathbf{W}_a\mathbf{W}_a^H\mathbf{V}_{a1}\right|\right] \ . \end{aligned} \tag{A.4}$$

The first term in (A.4) is the average achievable rate using ideal beamforming. Define the effective channel from Alice to Bob as $\widetilde{\mathbf{H}}_{ba} \triangleq \mathbf{W}_b^H\mathbf{H}_{ba}$. Since $\mathbf{W}_b$ is a submatrix of a unitary matrix, $\widetilde{\mathbf{H}}_{ba}$ is also a zero-mean complex Gaussian matrix with i.i.d. elements. However, the elements of $\widetilde{\mathbf{H}}_{ba}$ have a variance greater than unity due to the truncation of the $m-d$ smallest eigenvalues. In order to apply the random matrix result stated in (3.20), it is necessary to normalize the effective channel $\widetilde{\mathbf{H}}_{ba}$ to obtain unit variance elements. Because the exact normalization constant is difficult to obtain analytically, I scale $\widetilde{\mathbf{H}}_{ba}$ by an approximate factor $\sqrt{\frac{d}{m}}$ [11]. In the sequel, this normalization factor is absorbed into the transmit power. Since the quantity $\mathbb{E}_{\mathbf{H}_{ba}}\left[\log_2\left|\mathbf{I}_d + \frac{P_s}{d}\mathbf{\Lambda}_{a1}^2\right|\right]$ is approximately given by $dF(\frac{l}{d}, \frac{m}{d}P_s)$, it can be represented as $\overline{R}(d)$.

The second term in (A.4) is due to interference leakage from Hugo, and

$$\begin{aligned} & - \mathbb{E}\left[\log_2\left|\mathbf{I}_d + \frac{P_i}{N_h - d}\mathbf{\Lambda}_{h1}^2\mathbf{V}_{h1}^H\mathbf{W}_h\mathbf{W}_h^H\mathbf{V}_{h1}\right|\right] \\ \geq & - \mathbb{E}\left[\log_2\left|\mathbf{I}_d + \frac{P_i}{N_h - d}\mathbb{E}\left[\mathbf{\Lambda}_{h1}^2\right]\mathbf{V}_{h1}^H\mathbf{W}_h\mathbf{W}_h^H\mathbf{V}_{h1}\right|\right] \qquad \text{(A.5)} \\ = & - \mathbb{E}\left[\log_2\left|\mathbf{I}_d + \frac{P_iN_h}{N_h - d}\mathbf{V}_{h1}^H\mathbf{W}_h\mathbf{W}_h^H\mathbf{V}_{h1}\right|\right] \ , \qquad \text{(A.6)} \end{aligned}$$

where (A.5) follows from Jensen's inequality due to the concavity of the log-determinant function and the statistical independence of $\mathbf{\Lambda}_{h1}$ and $\mathbf{V}_{h1}^H\mathbf{W}_h\mathbf{W}_h^H\mathbf{V}_{h1}$, and (A.6) holds since $\mathbb{E}\left[\mathbf{\Lambda}_{h1}^2\right] = N_h\mathbf{I}_d$ [29].

1. If $M_h = d$, (A.6) becomes

$$
\begin{aligned}
&- \mathbb{E}\left[\log_2\left|\mathbf{I}_d + \frac{P_i N_h}{N_h - d}\mathbf{V}_{h1}^H \widehat{\mathbf{V}}_{h1}^{\perp}\widehat{\mathbf{V}}_{h1}^{\perp H}\mathbf{V}_{h1}\right|\right] \\
&= - \mathbb{E}\left[\log_2\left|\mathbf{I}_{M_h} + \frac{P_i N_h}{N_h - d}\left(\mathbf{I}_{M_h} - \mathbf{V}_{h1}^H \widehat{\mathbf{V}}_{h1}\widehat{\mathbf{V}}_{h1}^H \mathbf{V}_{h1}\right)\right|\right] \\
&\geq - \log_2\left|\mathbf{I}_{M_h} + \frac{P_i N_h}{N_h - d}\left(\mathbf{I}_{M_h} - \mathbb{E}\left[\mathbf{V}_{h1}^H \widehat{\mathbf{V}}_{h1}\widehat{\mathbf{V}}_{h1}^H \mathbf{V}_{h1}\right]\right)\right| \qquad \text{(A.7)} \\
&= -M_h \log_2\left(1 + \frac{P_i N_h D_h}{(N_h - d)\, M_h}\right) \ ; \qquad \text{(A.8)}
\end{aligned}
$$

2. if $M_h = N_h - d$, (A.6) becomes

$$
\begin{aligned}
&- \mathbb{E}\left[\log_2\left|\mathbf{I}_d + \frac{P_i N_h}{N_h - d}\mathbf{V}_{h1}^H \widehat{\mathbf{V}}_{h2}\widehat{\mathbf{V}}_{h2}^H \mathbf{V}_{h1}\right|\right] \\
&= - \mathbb{E}\left[\log_2\left|\mathbf{I}_{M_h} + \frac{P_i N_h}{N_h - d}\widehat{\mathbf{V}}_{h2}^H \mathbf{V}_{h1}\mathbf{V}_{h1}^H \widehat{\mathbf{V}}_{h2}\right|\right] \qquad \text{(A.9)} \\
&= - \mathbb{E}\left[\log_2\left|\mathbf{I}_{M_h} + \frac{P_i N_h}{N_h - d}\left(\mathbf{I}_{M_h} - \widehat{\mathbf{V}}_{h2}^H \mathbf{V}_{h2}\mathbf{V}_{h2}^H \widehat{\mathbf{V}}_{h2}\right)\right|\right] \\
&\geq - \log_2\left|\mathbf{I}_{M_h} + \frac{P_i N_h}{N_h - d}\left(\mathbf{I}_{M_h} - \mathbb{E}\left[\widehat{\mathbf{V}}_{h2}^H \mathbf{V}_{h2}\mathbf{V}_{h2}^H \widehat{\mathbf{V}}_{h2}\right]\right)\right| \qquad \text{(A.10)} \\
&= -M_h \log_2\left(1 + \frac{P_i N_h D_h}{(N_h - d)\, M_h}\right) \ . \qquad \text{(A.11)}
\end{aligned}
$$

Equation (A.9) follows from Sylvester's determinant theorem. Equations (A.7) and (A.10) are derived from Jensen's inequality, and (A.8) and (A.11) are due to the result of Lemma 3.2 in (3.21). Based on (3.9) which indicates that $D_h \leq G_h 2^{-\frac{B_h}{T_h}}$ and the equivalence of (A.8) and (A.11), the second term in (A.4) is bounded below by

$$
-\mathbb{E}\left[\log_2\left|\mathbf{I}_d + \frac{P_i}{N_h - d}\mathbf{\Lambda}_{h1}^2 \mathbf{V}_{h1}^H \mathbf{W}_h \mathbf{W}_h^H \mathbf{V}_{h1}\right|\right] \geq -M_h \log_2\left(1 + \frac{P_i N_h G_h 2^{-\frac{B_h}{T_h}}}{(N_h - d)\, M_h}\right) \ . \qquad \text{(A.12)}
$$

The third term in (A.4) is due to mismatch with the ideal beamforming gain for Alice.

1. If $M_a = d$, the third term in (A.4) becomes

$$\mathbb{E}\left[\log_2\left|\mathbf{V}_{a1}^H\widehat{\mathbf{V}}_{a1}\widehat{\mathbf{V}}_{a1}^H\mathbf{V}_{a1}\right|\right] \approx M_a\log_2\left(1-\frac{D_a}{M_a}\right)\ ; \tag{A.13}$$

2. if $M_a = N_a - d$, the third term in (A.4) becomes

$$\begin{aligned}
&\mathbb{E}\left[\log_2\left|\mathbf{V}_{a1}^H\widehat{\mathbf{V}}_{a2}^{\perp}\widehat{\mathbf{V}}_{a2}^{\perp\,H}\mathbf{V}_{a1}\right|\right]\\
&=\mathbb{E}\left[\log_2\left|\mathbf{I}_d-\mathbf{V}_{a1}^H\widehat{\mathbf{V}}_{a2}\widehat{\mathbf{V}}_{a2}^H\mathbf{V}_{a1}\right|\right]\\
&=\mathbb{E}\left[\log_2\left|\mathbf{I}_{M_a}-\widehat{\mathbf{V}}_{a2}^H\mathbf{V}_{a1}\mathbf{V}_{a1}^H\widehat{\mathbf{V}}_{a2}\right|\right]\\
&=\mathbb{E}\left[\log_2\left|\widehat{\mathbf{V}}_{a2}^H\mathbf{V}_{a2}\mathbf{V}_{a2}^H\widehat{\mathbf{V}}_{a2}\right|\right]\\
&\approx M_a\log_2\left(1-\frac{D_a}{M_a}\right)\ .
\end{aligned} \tag{A.14}$$

For B_a sufficiently large, the second term in (A.4) dominates the third term, and thus I use the approximation in (A.13) and (A.14) following from (3.21) in Lemma 3.2. Since from (3.9) we have that $D_a \le G_a 2^{-\frac{B_a}{T_a}}$ and the equivalence of (A.13) and (A.14), the third term in (A.4) is approximately bounded below by

$$\mathbb{E}\left[\log_2\left|\mathbf{V}_{a1}^H\mathbf{W}_a\mathbf{W}_a^H\mathbf{V}_{a1}\right|\right] \gtrsim M_a\log_2\left(1-\frac{G_a2^{-\frac{B_a}{T_a}}}{M_a}\right) = -M_a\log_2\frac{M_a}{M_a-G_a2^{-\frac{B_a}{T_a}}}\ . \tag{A.15}$$

In summary, the approximate lower bound for the average rate of the main channel is given by

$$\overline{R_b}(d) \gtrsim dF(\frac{l}{d},\frac{m}{d}P_s) - M_h\log_2\left(1+\frac{P_iN_hG_h2^{-\frac{B_h}{T_h}}}{(N_h-d)\,M_h}\right) - M_a\log_2\frac{M_a}{M_a-G_a2^{-\frac{B_a}{T_a}}}\ .$$

A.3 Proof of Theorem 3.4

The approximate lower bound in the first term of (3.17) is given by (3.22) in Theorem 3.1. Here I apply the optimal bit allocation result of (3.27a)-(3.27c) in Theorem 3.3 that maximizes this lower bound:

$$\overline{R_b}(d) \gtrsim dF\left(\frac{l}{d}, \frac{m}{d}P_s\right) - M_h \log_2\left(1 + \frac{P_i N_h G_h 2^{-\frac{\widetilde{B}_h^*}{T_h}}}{(N_h - d)\, M_h}\right) - M_h \log_2 \frac{M_h}{M_h - G_h 2^{-\frac{B - \widetilde{B}_h^*}{T_h}}} \,.$$

With knowledge of the the statistics of Eve's channel, the second term in (3.17) becomes

$$\begin{aligned}
&- \mathbb{E}_{\widetilde{\mathbf{H}}_{ea}, \widetilde{\mathbf{H}}_{eh}}\left[\log_2 \frac{\left|\sigma_e^2 \mathbf{I}_{N_e} + \frac{P_i}{N_h - d}\widetilde{\mathbf{H}}_{eh}\widetilde{\mathbf{H}}_{eh}^H + \frac{P_s}{d}\widetilde{\mathbf{H}}_{ea}\widetilde{\mathbf{H}}_{ea}^H\right|}{\left|\sigma_e^2 \mathbf{I}_{N_e} + \frac{P_i}{N_h - d}\widetilde{\mathbf{H}}_{eh}\widetilde{\mathbf{H}}_{eh}^H\right|}\right] \\
\geq &- \mathbb{E}_{\widetilde{\mathbf{H}}_{eh}}\left[\log_2 \frac{\left|\sigma_e^2 \mathbf{I}_{N_e} + \frac{P_i}{N_h - d}\widetilde{\mathbf{H}}_{eh}\widetilde{\mathbf{H}}_{eh}^H + \frac{P_s}{d}\mathbb{E}\left[\widetilde{\mathbf{H}}_{ea}\widetilde{\mathbf{H}}_{ea}^H\right]\right|}{\left|\sigma_e^2 \mathbf{I}_{N_e} + \frac{P_i}{N_h - d}\widetilde{\mathbf{H}}_{eh}\widetilde{\mathbf{H}}_{eh}^H\right|}\right] \qquad \text{(A.16)} \\
= &- \mathbb{E}_{\widetilde{\mathbf{H}}_{eh}}\left[\log_2 \left|(\sigma_e^2 + P_s)\mathbf{I}_{N_e} + \frac{P_i}{N_h - d}\widetilde{\mathbf{H}}_{eh}\widetilde{\mathbf{H}}_{eh}^H\right| - \log_2 \left|\sigma_e^2 \mathbf{I}_{N_e} + \frac{P_i}{N_h - d}\widetilde{\mathbf{H}}_{eh}\widetilde{\mathbf{H}}_{eh}^H\right|\right] \qquad \text{(A.17)} \\
= &- \mathbb{E}_{\widetilde{\mathbf{H}}_{eh}}\left[\log_2 \left|\mathbf{I}_{N_e} + \frac{P_i}{(\sigma_e^2 + P_s)(N_h - d)}\widetilde{\mathbf{H}}_{eh}\widetilde{\mathbf{H}}_{eh}^H\right|\right] \\
&+ \mathbb{E}_{\widetilde{\mathbf{H}}_{eh}}\left[\log_2 \left|\mathbf{I}_{N_e} + \frac{P_i}{\sigma_e^2(N_h - d)}\widetilde{\mathbf{H}}_{eh}\widetilde{\mathbf{H}}_{eh}^H\right|\right] - N_e \log_2(1 + \frac{P_s}{\sigma_e^2}) \\
= &- (N_h - d)\left[F\left(\frac{N_e}{N_h - d}, \frac{P_i}{\sigma_e^2 + P_s}\right) - F\left(\frac{N_e}{N_h - d}, \frac{P_i}{\sigma_e^2}\right)\right] - N_e \log_2(1 + \frac{P_s}{\sigma_e^2}) \,, \qquad \text{(A.18)}
\end{aligned}$$

where (A.16) holds due to the concavity of the log-determinant and (A.17) follows from the fact that $\mathbb{E}\left[\widetilde{\mathbf{H}}_{ea}\widetilde{\mathbf{H}}_{ea}^H\right] = d\mathbf{I}_{N_e}$. Equation (A.18) is obtained from (3.20) in Lemma 3.1 because the distribution of $\widetilde{\mathbf{H}}_{eh}$ is also known. Putting this all together, the average secrecy rate using the optimal bit allocation strategy can be approximately bounded below as in (3.30).

Appendix B

Appendix for Chapter 4

B.1 Proof of Lemma 4.1

Let $z = \|\mathbf{g}\|^2$. The pdf of z is given by [69]:

$$f(z) = \begin{cases} \frac{z^{t-1}e^{-z/\sigma_g^2}}{\Gamma(t)\sigma_g^{2t}} & z > 0 \\ 0 & \text{otherwise .} \end{cases}$$

The expected value of the logarithm of z is:

$$\begin{aligned} \mathbb{E}\left[\ln z\right] &= \int_0^\infty \ln z \ \frac{z^{t-1}e^{-z/\sigma_g^2}}{\Gamma(t)\sigma_g^{2t}} \ dz \\ &= \frac{1}{\Gamma(t)\sigma_g^{2t}} \int_0^\infty z^{t-1}e^{-z/\sigma_g^2} \ln z \ dz \\ &= \frac{1}{\Gamma(t)\sigma_g^{2t}} \ \sigma_g^{2t}\Gamma(t) \left[\psi(t) + \ln \sigma_g^2\right] \\ &= \psi(t) + \ln \sigma_g^2 \ , \end{aligned} \tag{B.1}$$

where [70] applies (B.1). Therefore,

$$\mathbb{E}\left[\ln \|\mathbf{g}\|^2\right] = \psi(t) + \ln \sigma_g^2 \ .$$

B.2 Proof of Lemma 4.2

The squared inner product of any two t-dimensional independent and isotropically distributed unit-norm vectors is Beta-distributed [71]:

$$\left|\mathbf{u}^H \mathbf{v}\right|^2 \sim \beta\left(1, t-1\right) \ .$$

Due to the properties of the beta distribution,

$$1 - \left|\mathbf{u}^H \mathbf{v}\right|^2 \sim \beta\left(t-1, 1\right) \ ,$$

and the expected values of the logarithm of the beta random variables are

$$\begin{aligned} \mathbb{E}\left[\ln \left|\mathbf{u}^H \mathbf{v}\right|^2\right] &= \psi(1) - \psi(t) \\ \mathbb{E}\left[\ln\left(1 - \left|\mathbf{u}^H \mathbf{v}\right|^2\right)\right] &= \psi(t-1) - \psi(t) \ . \end{aligned}$$

B.3 Proof of Lemma 4.3

Let q_i be a realization of the random variable Q,

$$
\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n}\log_2\left(1+\rho q_i\right)-\log_2\left(1+\rho\ e^{\frac{1}{n}\sum\limits_{i=1}^{n}\ln q_i}\right)\\
=&\frac{1}{n}\left[\sum_{i=1}^{n}\log_2\left(1+\rho q_i\right)-\log_2\left(1+\rho\ e^{\frac{1}{n}\sum\limits_{i=1}^{n}\ln q_i}\right)^n\right]\\
=&\frac{1}{n}\left[\log_2\prod_{i=1}^{n}\left(1+\rho q_i\right)-\log_2\left(1+\rho\left(\prod_{i=1}^{n}q_i\right)^{\frac{1}{n}}\right)^n\right]\\
=&\frac{1}{n}\log_2\prod_{i=1}^{n}\frac{1+\rho q_i}{1+\rho\left(\prod\limits_{i=1}^{n}q_i\right)^{\frac{1}{n}}}\\
\approx&\frac{1}{n}\log_2\prod_{i=1}^{n}\frac{q_i}{\left(\prod\limits_{i=1}^{n}q_i\right)^{\frac{1}{n}}} \qquad \text{(B.2)}\\
=&\frac{1}{n}\log_2\frac{\prod\limits_{i=1}^{n}q_i}{\prod\limits_{i=1}^{n}q_i}=0\ .
\end{aligned}
$$

The approximation in (B.2) is based on the assumption that ρ is sufficiently large. Therefore,

$$
\frac{1}{n}\sum_{i=1}^{n}\log_2\left(1+\rho q_i\right)\approx\log_2\left(1+\rho\ e^{\frac{1}{n}\sum\limits_{i=1}^{n}\ln q_i}\right)\ .
$$

As $n\to\infty$, by the law of large numbers,

$$
\begin{aligned}
\frac{1}{n}\sum_{i=1}^{n}\log_2\left(1+\rho q_i\right)&\to\mathbb{E}\left[\log_2\left(1+\rho Q\right)\right]\\
\frac{1}{n}\sum_{i=1}^{n}\ln q_i&\to\mathbb{E}\left[\ln Q\right]\ .
\end{aligned}
$$

Hence, the following approximation holds:

$$\mathbb{E}\left[\log_2\left(1+\rho Q\right)\right] \approx \log_2\left(1+\rho\, e^{\mathbb{E}[\ln Q]}\right) \ .$$

B.4 Proof of Lemma 4.4

Using the beamforming vectors in (4.4), the expected value of the received signal power term under the limited rate feedback scheme can be derived as follows:

$$
\begin{aligned}
&\mathbb{E}\left[\log_2\left(1+P_i\left|\mathbf{h}_{ii}^H\mathbf{w}_i\right|^2\right)\right] \\
=&\mathbb{E}\left[\log_2\left(1+\frac{P_i\left|\mathbf{h}_{ii}^H\left(\mathbf{I}-\hat{\mathbf{h}}_{ji}\hat{\mathbf{h}}_{ji}^H\right)\hat{\mathbf{h}}_{ii}\right|^2}{\left\|\left(\mathbf{I}-\hat{\mathbf{h}}_{ji}\hat{\mathbf{h}}_{ji}^H\right)\hat{\mathbf{h}}_{ii}\right\|^2}\right)\right] \\
\approx&\mathbb{E}\left[\log_2\left(1+\frac{P_i\left|\mathbf{h}_{ii}^H\mathbb{E}\left[\mathbf{I}-\hat{\mathbf{h}}_{ji}\hat{\mathbf{h}}_{ji}^H\right]\hat{\mathbf{h}}_{ii}\right|^2}{\left\|\left(\mathbf{I}-\hat{\mathbf{h}}_{ji}\hat{\mathbf{h}}_{ji}^H\right)\hat{\mathbf{h}}_{ii}\right\|^2}\right)\right] &\text{(B.3)}\\
\approx&\mathbb{E}\left[\log_2\left(1+\frac{P_i e^{2\psi(N_i-1)-2\psi(N_i)}\left|\mathbf{h}_{ii}^H\hat{\mathbf{h}}_{ii}\right|^2}{\left\|\left(\mathbf{I}-\hat{\mathbf{h}}_{ji}\hat{\mathbf{h}}_{ji}^H\right)\hat{\mathbf{h}}_{ii}\right\|^2}\right)\right] &\text{(B.4)}\\
=&\mathbb{E}\left[\log_2\left(1+\frac{P_i e^{2\psi(N_i-1)-2\psi(N_i)}\left\|\mathbf{h}_{ii}\right\|^2\left|\tilde{\mathbf{h}}_{ii}^H\hat{\mathbf{h}}_{ii}\right|^2}{1-\left|\hat{\mathbf{h}}_{ji}^H\hat{\mathbf{h}}_{ii}\right|^2}\right)\right] \\
\approx&\mathbb{E}\left[\log_2\left(1+\frac{P_i e^{2\psi(N_i-1)-2\psi(N_i)}\left\|\mathbf{h}_{ii}\right\|^2\mathbb{E}\left[\left|\tilde{\mathbf{h}}_{ii}^H\hat{\mathbf{h}}_{ii}\right|^2\right]}{1-\left|\hat{\mathbf{h}}_{ji}^H\hat{\mathbf{h}}_{ii}\right|^2}\right)\right] &\text{(B.5)}\\
\approx&\mathbb{E}\left[\log_2\left(1+\frac{P_i e^{2\psi(N_i-1)-2\psi(N_i)}\left\|\mathbf{h}_{ii}\right\|^2\left(1-\gamma_i 2^{-\frac{B_{ii}}{N_i-1}}\right)}{1-\left|\hat{\mathbf{h}}_{ji}^H\hat{\mathbf{h}}_{ii}\right|^2}\right)\right] &\text{(B.6)}\\
\approx&\mathbb{E}\left[\log_2\left(1+\frac{P_i e^{2\psi(N_i-1)-2\psi(N_i)}\left\|\mathbf{h}_{ii}\right\|^2}{1-\left|\hat{\mathbf{h}}_{ji}^H\hat{\mathbf{h}}_{ii}\right|^2}\right)\right]+\log_2\left(1-\gamma_i 2^{-\frac{B_{ii}}{N_i-1}}\right) &\text{(B.7)}\\
\approx&\log_2\left(1+P_i e^{\mathbb{E}\left[\ln\frac{e^{2\psi(N_i-1)-2\psi(N_i)}\left\|\mathbf{h}_{ii}\right\|^2}{1-\left|\hat{\mathbf{h}}_{ji}^H\hat{\mathbf{h}}_{ii}\right|^2}\right]}\right)+\log_2\left(1-\gamma_i 2^{-\frac{B_{ii}}{N_i-1}}\right) &\text{(B.8)}\\
=&\log_2\left(1+P_i e^{2\psi(N_i-1)-2\psi(N_i)+\mathbb{E}\left[\ln\left\|\mathbf{h}_{ii}\right\|^2\right]-\mathbb{E}\left[\ln\left(1-\left|\hat{\mathbf{h}}_{ji}^H\hat{\mathbf{h}}_{ii}\right|^2\right)\right]}\right)+\log_2\left(1-\gamma_i 2^{-\frac{B_{ii}}{N_i-1}}\right) \\
=&\log_2\left(1+P_i\sigma_{ii}^2 e^{\psi(N_i-1)}\right)+\log_2\left(1-\gamma_i 2^{-\frac{B_{ii}}{N_i-1}}\right)\ . &\text{(B.9)}
\end{aligned}
$$

I replace $\left(\mathbf{I} - \hat{\mathbf{h}}_{ji}\hat{\mathbf{h}}_{ji}^{H}\right)$ and $\left|\tilde{\mathbf{h}}_{ii}^{H}\hat{\mathbf{h}}_{ii}\right|^{2}$ with their expected values in (B.3) and (B.5). The proof of (B.4) is given in Appendix B.6. Equation (B.6) is obtained from (4.1). The approximation in (B.7) is due to the high SNR assumption, (B.8) is based on (4.8) in Lemma 4.3, and (B.9) follows from (4.5) in Lemma 4.1 and (4.7) in Lemma 4.2.

B.5 Proof of Lemma 4.5

Using the beamforming vectors in (4.4) and defining $\tilde{\mathbf{v}}_{ij}^H \triangleq \frac{\mathbf{h}_{ij}^H\left(\mathbf{I}-\hat{\mathbf{h}}_{ij}\hat{\mathbf{h}}_{ij}^H\right)}{\left\|\mathbf{h}_{ij}^H\left(\mathbf{I}-\hat{\mathbf{h}}_{ij}\hat{\mathbf{h}}_{ij}^H\right)\right\|}$, the average received noise power under the limited rate feedback scheme can be derived as follows:

$$
\begin{aligned}
&\mathbb{E}\left[\log_2\left(1+P_j\left|\mathbf{h}_{ij}^H\mathbf{w}_j\right|^2\right)\right]\\
=&\mathbb{E}\left[\log_2\left(1+\frac{P_j\left|\mathbf{h}_{ij}^H\left(\mathbf{I}-\hat{\mathbf{h}}_{ij}\hat{\mathbf{h}}_{ij}^H\right)\hat{\mathbf{h}}_{jj}\right|^2}{\left\|\left(\mathbf{I}-\hat{\mathbf{h}}_{ij}\hat{\mathbf{h}}_{ij}^H\right)\hat{\mathbf{h}}_{jj}\right\|^2}\right)\right]\\
=&\mathbb{E}\left[\log_2\left(1+\frac{P_j\left|\tilde{\mathbf{v}}_{ij}^H\hat{\mathbf{h}}_{jj}\right|^2\left\|\mathbf{h}_{ij}\right\|^2\left\|\tilde{\mathbf{h}}_{ij}^H\left(\mathbf{I}-\hat{\mathbf{h}}_{ij}\hat{\mathbf{h}}_{ij}^H\right)\right\|^2}{\left\|\left(\mathbf{I}-\hat{\mathbf{h}}_{ij}\hat{\mathbf{h}}_{ij}^H\right)\hat{\mathbf{h}}_{jj}\right\|^2}\right)\right]\\
=&\mathbb{E}\left[\log_2\left(1+\frac{P_j\left|\tilde{\mathbf{v}}_{ij}^H\hat{\mathbf{h}}_{jj}\right|^2\left\|\mathbf{h}_{ij}\right\|^2\left(1-\left|\tilde{\mathbf{h}}_{ij}^H\hat{\mathbf{h}}_{ij}\right|^2\right)}{1-\left|\hat{\mathbf{h}}_{ij}^H\hat{\mathbf{h}}_{jj}\right|^2}\right)\right]\\
\approx&\mathbb{E}\left[\log_2\left(1+\frac{P_j\left|\tilde{\mathbf{v}}_{ij}^H\hat{\mathbf{h}}_{jj}\right|^2\left\|\mathbf{h}_{ij}\right\|^2\left(1-\mathbb{E}\left[\left|\tilde{\mathbf{h}}_{ij}^H\hat{\mathbf{h}}_{ij}\right|^2\right]\right)}{1-\left|\hat{\mathbf{h}}_{ij}^H\hat{\mathbf{h}}_{jj}\right|^2}\right)\right] &\text{(B.10)}\\
\approx&\mathbb{E}\left[\log_2\left(1+\frac{P_j\left|\tilde{\mathbf{v}}_{ij}^H\hat{\mathbf{h}}_{jj}\right|^2\left\|\mathbf{h}_{ij}\right\|^2\gamma_j 2^{-\frac{B_{ij}}{N_j-1}}}{1-\left|\hat{\mathbf{h}}_{ij}^H\hat{\mathbf{h}}_{jj}\right|^2}\right)\right] &\text{(B.11)}\\
\approx&\log_2\left(1+P_j e^{\mathbb{E}\left[\ln\frac{\left|\tilde{\mathbf{v}}_{ij}^H\hat{\mathbf{h}}_{jj}\right|^2\left\|\mathbf{h}_{ij}\right\|^2\gamma_j 2^{-\frac{B_{ij}}{N_j-1}}}{1-\left|\hat{\mathbf{h}}_{ij}^H\hat{\mathbf{h}}_{jj}\right|^2}\right]}\right) &\text{(B.12)}\\
=&\log_2\left(1+P_j e^{\mathbb{E}\left[\ln\left|\tilde{\mathbf{v}}_{ij}^H\hat{\mathbf{h}}_{jj}\right|^2\right]+\mathbb{E}\left[\ln\|\mathbf{h}_{ij}\|^2\right]-\mathbb{E}\left[\ln\left(1-\left|\hat{\mathbf{h}}_{ij}^H\hat{\mathbf{h}}_{jj}\right|^2\right)\right]}\gamma_j 2^{-\frac{B_{ij}}{N_j-1}}\right)\\
=&\log_2\left(1+P_j\sigma_{ij}^2 e^{\psi(N_j)-\psi(N_j-1)+\psi(1)}\gamma_j 2^{-\frac{B_{ij}}{N_j-1}}\right)\ . &\text{(B.13)}
\end{aligned}
$$

$\left|\tilde{\mathbf{h}}_{ij}^H\hat{\mathbf{h}}_{ij}\right|^2$ is replaced with its expected value in (B.10). Equation (B.11) is obtained from (4.2). Equation (B.12) is based on (4.8) in Lemma 4.3. Equation (B.13) follows from (4.5) in

Lemma 4.1 and (4.6)-(4.7) in Lemma 4.2.

B.6 Proof of Equation (B.4)

The expected value of matrix term $\left(\mathbf{I} - \hat{\mathbf{h}}_{ji}\hat{\mathbf{h}}_{ji}^H\right)$ can be approximated as

$$\begin{aligned}
\mathbb{E}\left[\mathbf{I} - \hat{\mathbf{h}}_{ji}\hat{\mathbf{h}}_{ji}^H\right] &= \mathbf{I} - \mathbb{E}\left[\hat{\mathbf{h}}_{ji}\hat{\mathbf{h}}_{ji}^H\right] \\
&= \left(1 - \frac{1}{N_i}\right)\mathbf{I} && \text{(B.14)} \\
&= \left(1 + \frac{1}{N_i - 1}\right)^{-1}\mathbf{I} \\
&\approx \left(e^{\frac{1}{N_i - 1}}\right)^{-1}\mathbf{I} && \text{(B.15)} \\
&= \left(e^{\psi(N_i) - \psi(N_i - 1)}\right)^{-1}\mathbf{I} && \text{(B.16)} \\
&= \left(e^{\psi(N_i - 1) - \psi(N_i)}\right)\mathbf{I}\,,
\end{aligned}$$

where (B.14) holds since $\mathbb{E}\left[\hat{\mathbf{h}}_{ji}\hat{\mathbf{h}}_{ji}^H\right] = \frac{1}{N_i}\mathbf{I}$. Equation (B.15) is from the first order Taylor approximation, and (B.16) is due to the recurrence relation of digamma function.