WIDEBAND SLOPE OF INTERFERENCE CHANNELS

A DISSERTATION SUBMITTED TO THE GRADUATE DIVISION OF THE UNIVERSITY OF HAWAIʻI AT MĀNOA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

ELECTRICAL ENGINEERING

MAY 2013

By

Minqi Shen

Dissertation Committee:

Anders Høst-Madsen, Chairperson

N. Thomas Gaarder

Aleksandar Kavčić

Narayana P. Santhanam

J. B. Nation

Acknowledgments

I am deeply indebted to many extraordinary individuals for making my graduate studies one of the most rewarding periods of my life. First of all, I would like to express my deepest gratitude to my adviser Dr. Anders Høst-Madsen, for his guidance and support throughout the past few years. I have benefited tremendously from many insightful and inspiring discussions with him on the topics in information theory and multiple user communication systems. This work has been done mostly in our collaboration.

I would like to thank Dr. Thomas Gaarder, Dr. Aleksandar Kavčić, Dr. Narayana Santhanam, and Dr. J. B. Nation for serving on my dissertation committee. I would also like to thank Dr. James Yee for serving on my qualify examination committee. I thank them for their extraordinary teaching, which has always been a source of inspiration.

Further more, a special thanks to my fellow students residing in HOLM 490, from both the present and the past: Da, Ciril, Ying, Sharif, and Navid. I thank them for the enjoyable and informative conversations that lighten up my daily life as a graduate student.

Last but not the least, I would like to thank my parents for their being encouraging and supportive through my life.

ABSTRACT

This dissertation studies the bandwidth-power tradeoff of K-user interference channels in the low-SNR regime, which is mirrored by $R\left(\frac{E_b}{N_0}\right)$: the spectral efficiency as a function of the transmitted energy per information bit. A system working in the low-SNR regime–a concept defined in [1], is characterized by very small R, such that $R\left(\frac{E_b}{N_0}\right)$ can be closely approximated by its first-order expansion determined by two measures: the minimum energy per bit $\left.\frac{E_b}{N_0}\right|_{\min}$ and the wide-band slope $\mathcal{S}_0$, which is the first-order slope of $R\left(\frac{E_b}{N_0}\right)$ as $\frac{E_b}{N_0}$ approaches $\left.\frac{E_b}{N_0}\right|_{\min}$.

The low-SNR regime performance of an interference channel was studied in[2]. It has shown that $\left.\frac{E_b}{N_0}\right|_{\min}$ is not affected by interference while $\mathcal{S}_0$ is reduced by it, implying that interference indeed deteriorates system performance as long as R is non-zero.

[1] indicates that the low-SNR regime performance is approached as SNR tends to zero, where SNR is the signal to noise ratio per second per Hz. This dissertation considers two cases in which the limit SNR $\to 0$ is achieved. One is the large bandwidth case, where the bandwidth $B \to \infty$ while power P is fixed and finite. The other case is the small bandwidth case, where $P \to 0$ and while B is fixed and finite.

The large bandwidth case is characterized by propagation delays considerably larger than the symbol duration. For K-user interference channel working under this case, this dissertation proposes an interference alignment scheme over time domain which achieves $\Delta\mathcal{S}_0 = \frac{1}{2}$ with probability one, independent of the number of users K. It improves the best known performance achieved by TDMA by a

factor of ${}^{K}/_{2}$. Outer bounds on $\mathcal{S}_0$ are also developed, which reveal that this interference alignment scheme is asymptotically optimal as $K \to \infty$.

In the small bandwidth case, propagation delay is assumed to be zero. Quite the contrary to the large bandwidth case, this dissertation shows that there exists a set of channel realizations with non-zero probability, for which the achievable $\mathcal{S}_0$ satisfies $\Delta\mathcal{S}_0 < {}^{1}/_{K} + \delta$, $\forall\delta > 0$, therefore TDMA is almost optimal. For channel realizations not in this set, an interference alignment scheme based on the concept of circularly asymmetric signaling is proposed. For some channel realizations, it gives noticeably better performance comparing with existing achievable schemes.

Contents

List of Figures

Chapter 1

Introduction

This thesis studies the bandwidth-power trade-off of a K-user interference channel in the low-SNR regime[1]. For wireless communication systems, bandwidth and input power are two important design parameters. They are related by the function $R\left(\frac{E_b}{N_0}\right)$, where $\frac{E_b}{N_0}$ is the transmitted energy per bit, and R is the spectral efficiency. Low-SNR regime is a concept established by S. Verdú in the 2002 paper [1]. A system working in this regime is characterized by very small spectral efficiency, such that the $R\left(\frac{E_b}{N_0}\right)$ curve can be closely approximated by its first-order expansion, which is determined by two measures: the minimum energy per bit $\left.\frac{E_b}{N_0}\right|_{\min}$ and the wideband slope $\mathcal{S}_0$. $\left.\frac{E_b}{N_0}\right|_{\min}$ is the minimum transmitted energy per bit required by reliable communication, which can only be achieved at zero spectral efficiency; and $\mathcal{S}_0$ is the first-order slope of $R\left(\frac{E_b}{N_0}\right)$ as $\frac{E_b}{N_0}$ approaches $\left.\frac{E_b}{N_0}\right|_{\min}$.

To evaluate the impact of interference, we will use the performance of its corresponding interference-free channel as the reference point. Here we define the corresponding interference-free channel of an interference channel to be a channel with the same direct-link channel coefficients and same noises as the interference channel, but with no interference links.

In their 2004 work [2], G. Caire et al. have shown that the minimum energy per bit of an interference channel is equal to that of its corresponding interference-free channel. Therefore interference does not affect channel capacity when the spectral efficiency is zero. However, as long as the spectral efficiency is non-zero, even very very small and close to zero, interference starts to hurt system performance: the wideband slope will be deteriorated. For instance, our discussion in chapter 3 will show that $\forall\delta > 0$, there exists a set of 2-user interference channel realizations with non-zero probability, whose the achievable wideband slope of each user must satisfy $\mathcal{S}_0 < 1 + \delta$, while the corresponding interference-free channel can achieve $\mathcal{S}_0$ equal to 2. This result indicates that to achieve the same data rate, these

[1]SNR stands for the signal to noise ratio per second per Hz. It will be formally defined in section 1.1.

2-user interference channel realizations require two times as large the bandwidth required by the corresponding interference-free channel.

The wideband slope region of the interference channel remains an open problem. We will introduce background knowledge of the low-SNR regime in section 1.1 and section 1.2, then briefly review the general Gaussian interference channel capacity and the previous results on its low-SNR regime performance in section 1.3.

1.1 Channel Capacity And Minimum Energy Per Bit

Let us begin from a discrete-time memoryless channel with single input and single output. This channel consists of an input alphabet $\mathcal{X}$, an output alphabet $\mathcal{Y}$, and conditional probability $p\,(y|x)$, which expresses the probability of observing symbol y at the output if symbol x is transmitted at the input. The *channel capacity* C of the discrete-time memoryless channel is defined by C. E. Shannon in [3]. It is the maximum mutual information between the channel input and the channel output, optimized over all the possible input x with cumulative distribution function $F_x\,(x)$, i.e.,

$$C = \max_{F_x(x)} I\,(X\,;\,Y)\,.$$

Now we want to transmit message W over the channel, where W is drawn from the index set $\left\{1, \cdots, 2^{NR}\right\}$. At the transmitter, W is mapped to a length N codeword $X^N\,(W)$. The receiver then decodes W from the channel output Y^N following a proper decoding rule, $\hat{W}\left(Y^N\right)$. An error occurs if $\hat{W} \neq W$. R is defined as the *rate* of the $\left(2^{NR},\, N\right)$ codebook. Shannon's channel coding theorem states that for every $R < C$, there exists a $\left(2^{NR},\, N\right)$ codebook such that the probability of error can be made arbitrarily small as $N \to \infty$. Conversely, any codebook with arbitrarily small probability of error must satisfy $R < C$.

Figure 1.1 depicts a discrete-time Additive White Gaussian Noise (AWGN) channel. At time instance n, the channel output y is

$$y\,[n] \quad = \quad x\,[n] + z\,[n]\,. \tag{1.1}$$

The noise Z is an independent and identically distributed (i.i.d.) circularly symmetric complex Gaussian random variable with variance N_0, whose distribution function is denoted as $\mathcal{CN}\,(0,\, N_0)$. For any length-N input sequence x^N, we require an average energy per channel use constraint

$$\frac{1}{N}\sum_{n=1}^{N}|x[n]|^2 \quad \leq \quad p.$$

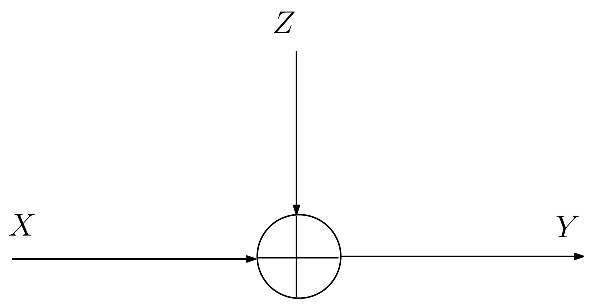

Figure 1.1: AWGN Channel Model

Denote the signal to noise ratio per channel use by $\text{SNR} \triangleq \frac{p}{N_0}$. The capacity of this AWGN channel is

$$C = \log(1+\text{SNR}),$$

achieved using an appropriate $\left(2^{NR}, N\right)$ codebook, whose entries are generated by i.i.d. Gaussian random variable $X \sim \mathcal{CN}(0, p)$.

This discrete-time channel model can represent a bandlimited channel with two-sided bandwidth B and power constraint P, $P = B \cdot p$, being sampled at the Nyquist–Shannon rate B. The achievable rate of the continuous-time channel can be found by [2]

$$\mathsf{R} = B \cdot R.$$

In this continuous-time channel model, R can be interpreted as the spectral efficiency. Given

$$\text{SNR} = \frac{P}{N_0 B}, \tag{1.2}$$

SNR can be interpreted as the signal to noise ratio per second per Hz.

The channel capacity of a continuous-time AWGN channel is

$$\begin{aligned}\mathsf{C} &= B \cdot C \\ &= B \cdot \log(1+\text{SNR}).\end{aligned} \tag{1.3}$$

We can see that C is a function of two design parameters, the power P and the bandwidth B. In his famous 1948 work [3], C. Shannon has noticed that for any

[2]A rigorous proof can be found in [4].

fixed power constraint P, a system with infinite bandwidth B achieves the best possible channel capacity, because C increases monotonically with the bandwidth. This observation is summarized by the following formula

$$\begin{aligned}\mathsf{C}_{\max} &= \lim_{B\to\infty} B\log\left(1+\frac{P}{N_0 B}\right)\\ &= \frac{P}{N_0}\log_2 e.\end{aligned} \tag{1.4}$$

Equivalently, the transmitted energy per bit achieves its minimum value required by reliable communication at infinite B:

$$\left.\frac{E_b}{N_0}\right|_{\min} = \lim_{B\to\infty}\frac{P}{N_0 B\cdot C} \tag{1.5}$$

$$= \log_e 2, \tag{1.6}$$

which is approximately -1.59dB.

The limit in Equation 1.5 can also be achieved in an alternative way. Assume that the system has certain amount of energy E, while the bandwidth $B<\infty$; and the system wants to transmit the maximum amount of data possible given the available amount of energy. To achieve this goal, the available energy shall be spread over the time domain:

$$\begin{aligned}(\mathsf{C}T)_{\max} &= \lim_{T\to\infty} TB\log\left(1+\frac{\mathsf{E}}{N_0 TB}\right) \qquad (1.7)\\ &= TB\cdot\frac{\mathsf{E}}{N_0 TB}\log_2 e\\ &= \frac{\mathsf{E}}{N_0}\log_2 e.\end{aligned}$$

The transmitted energy per bit achieves its minimum value required by reliable communication at $T=\infty$:

$$\left.\frac{E_b}{N_0}\right|_{\min} = \lim_{T\to\infty}\frac{\mathsf{E}}{N_0 TB\cdot C} \tag{1.8}$$

$$= \log_e 2. \tag{1.9}$$

Comparing Equation 1.5 and Equation 1.8, we can see that in general, the minimum energy per bit is achieved as SNR goes to zero, i.e., if a transmission scheme has spectral efficiency $R\,(\mathrm{SNR})$, then its achievable minimum energy per bit can be obtained from

$$\left.\frac{E_b}{N_0}\right|_{\min} = \lim_{\mathrm{SNR}\downarrow 0} \frac{\mathrm{SNR}}{R(\mathrm{SNR})}. \tag{1.10}$$

Given Equation 1.10, further reduction in [1] proves that the achievable minimum energy per bit is determined by the first-order Taylor expansion coefficient of $R(\mathrm{SNR})$ at zero SNR:

$$\left.\frac{E_b}{N_0}\right|_{\min} = \frac{\log_e 2}{\dot{R}(0)}. \tag{1.11}$$

If $R(\mathrm{SNR})$ is differentiable, then

$$\dot{R}(0) = \left.\frac{dR(\mathrm{SNR})}{d\mathrm{SNR}}\right|_{\mathrm{SNR}=0}.$$

Definition 1. [1] We call a transmission scheme *first-order optimal* if it achieves the optimal minimum energy per bit:

$$\dot{R}(0) = \dot{C}(0).$$

Further, [1] shows that the $\dot{C}(0)$ is equal to the channel gain G for a general point-to-point AWGN channel defined as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{z}, \tag{1.12}$$

where the $m \times m$ matrix $\mathbf{H}$ is the channel coefficient matrix, m can be any positive integer. For this channel,

$$\dot{C}(0) = G \tag{1.13}$$

$$G \triangleq \sup_{F_{\mathbf{x}}} \frac{E\left[\|\mathbf{H}\mathbf{x}\|^2\right]}{E\left[\|\mathbf{x}\|^2\right]} \tag{1.14}$$

If $m = 1$,$\mathbf{H}$ becomes a complex scalar h, which gives $\dot{C}(0) = |h|^2$ and

$$\left.\frac{E_b}{N_0}\right|_{\min} = \frac{\log_e 2}{|h|^2}.$$

After the channel coefficient matrix $\mathbf{H}$ is introduced into the channel model, $\dot{C}(0)$ becomes dependent to the channel state information (CSI), i.e., the knowledge of

$\mathbf{H}$, at the transmitters and the receivers. The case where only the transmitters have CSI is discussed in [5] and [6]. It turns out that even if the receivers have no CSI at all, the channel can still achieve optimal $\left.\frac{E_b}{N_0}\right|_{\min}$. Depending on the types of CSI the transmitters have, the optimal channel gain of the general AWGN channel varies. Two extreme cases are:

- The transmitters have perfect CSI. In this case, G is equal to $\lambda_{\max}^2$, where $\lambda_{\max}$ is the largest singular value of $\mathbf{H}$.
- The transmitters have no CSI on the realization and statistic properties of $\mathbf{H}$. In this case, $\sup_{F_{\mathbf{x}}} \frac{E\left[\|\mathbf{Hx}\|^2\right]}{E\left[\|\mathbf{x}\|^2\right]}$ is equal to $\frac{1}{n}E\left[trace\left(\mathbf{H}^\dagger\mathbf{H}\right)\right]$.

In-depth studies on the channel coherence in low-SNR regime include [1], [7] and [8].

It is trivial that Gaussian input, which achieves channel capacity, is first-order optimal. On the other hand, a first-order optimal input is not necessarily capacity achievable. [3] has noticed that BPSK signaling whose input distribution is

$$x = \begin{cases} \sqrt{P}e^{\iota\theta}, & \text{with probability } 1/2 \\ -\sqrt{P}e^{\iota\theta}, & \text{with probability } 1/2 \end{cases}$$

is first-order optimal; and [9] has shown that on-off signaling with vanishing duty cycle whose input distribution is

$$x = \begin{cases} 0, & \text{with probability } 1-\delta \\ \sqrt{A}, & \text{with probability } \delta \end{cases},$$

where $\delta = P/A$, is also first-order optimal.

In summary, at zero SNR, the continuous-time AWGN channel is the most energy efficient, and its performance is determined by the minimum energy per bit. A first-order optimal transmission scheme achieves the optimal minimum energy per bit required for reliable communication. A system with no receiver CSI and small input symbol set can achieve first-order optimal performance. Therefore, working at zero SNR benefits the performance and makes system design much easier.

1.2 Low-SNR Regime And Wideband Slope

Comparing with high SNR, using low SNR is a more energy efficient choice. When the SNR is small, the capacity of a discrete-time point-to-point channel satisfies $\log(1+\mathrm{SNR}) \approx \mathrm{SNR}\log_e 2$, which implies that the spectral efficiency R grows almost linearly with SNR; on the other hand, if the SNR is high, we have $\log(1+\mathrm{SNR}) \approx \log\mathrm{SNR}$, i.e., R grows with SNR in logarithm.

For decades, the low-SNR regime performance of a system is merely measured by its achievable minimum energy per bit, and whether a transmission scheme is first-order optimal is the only criterion of optimality in the low-SNR regime. The main disadvantage of using $\left.\frac{E_b}{N_0}\right|_{\min}$ as the performance measure is that it only characterizes the system performance at zero spectral efficiency, as Verdú has pointed out in [1]. A real world system can neither have infinite bandwidth or zero input power; its spectral efficiency R can be small, but can never be exactly zero. For such systems, the central question which $\left.\frac{E_b}{N_0}\right|_{\min}$ alone can not answer is: how much gain can be obtained in spectral efficiency R, if we are willing to spend more energy on each information bit transmitted than its minimum value $\left.\frac{E_b}{N_0}\right|_{\min}$. To illustrate the trade-off between the spectral efficiency and the energy per bit, we need a new measure which can withstand the test of non-zero spectral efficiency analysis.

The modified concept of low-SNR regime was established by S. Verdú in the 2002 paper [1]. A system working in this regime is characterized by very small but non-zero spectral efficiency, such that the $R\left(\frac{E_b}{N_0}\right)$ curve can be closely approximated by its first-order expansion, which is determined by two measures: the minimum energy per bit $\left.\frac{E_b}{N_0}\right|_{\min}$ and the wideband slope $\mathcal{S}_0$. The wideband slope $\mathcal{S}_0$ is the first-order slope of $R\left(\frac{E_b}{N_0}\right)$ as $\frac{E_b}{N_0}$ approaches $\left.\frac{E_b}{N_0}\right|_{\min}$, which is defined as

$$\mathcal{S}_0 \triangleq \lim_{\frac{E_b}{N_0}\downarrow\frac{E_b}{N_0}_{\min}} \frac{R\left(\frac{E_b}{N_0}\right)}{10\log_{10}\frac{E_b}{N_0} - 10\log_{10}\left.\frac{E_b}{N_0}\right|_{\min}} 10\log_{10}2. \tag{1.15}$$

$\mathcal{S}_0$ is in /s/Hz/3dB. Notice that $\frac{E_b}{N_0}$ is in dB; defining $\mathcal{S}_0$ in this way makes it independent from the channel gain G.

Combining Equation 1.15 and Equation 1.10, further reduction in [1] proves that the wideband slope can be determined by the first and second order Taylor expansion coefficients of $R(\mathrm{SNR})$ at $\mathrm{SNR}=0$:

$$\mathcal{S} = -\frac{2\left(\dot{R}(0)\right)^2}{\ddot{R}(0)}, \tag{1.16}$$

$\ddot{R}(0) = \left.\frac{d^2R(\mathrm{SNR})}{d\mathrm{SNR}^2}\right|_{\mathrm{SNR}=0}$ if $R(\mathrm{SNR})$ is second-order differentiable.

Two observations can be made given Equation 1.15. First, if two systems achieve equal $\left.\frac{E_b}{N_0}\right|_{\min}$ value, the $\frac{E_b}{N_0}$ value of the system with higher wideband slope approaches its minimum value faster, and this system is therefore more spectral efficient. Second, the priority in low-SNR regime is to minimize $\left.\frac{E_b}{N_0}\right|_{\min}$. We say that system A performs better than system B in the low-SNR regime if system A achieves lower $\left.\frac{E_b}{N_0}\right|_{\min}$ value than system B. This is because as long as the wideband slope of system A is non-zero, there always exists $\epsilon > 0$ such that if $\frac{E_b}{N_0} < \epsilon$ then $R_A\left(\frac{E_b}{N_0}\right) > R_B\left(\frac{E_b}{N_0}\right)$. Based on this two observations, we have the following remark:

Remark 2. To make fair comparison of the wideband slopes between different systems, they must have equal $\left.\frac{E_b}{N_0}\right|_{\min}$ in the first place.

In the low-SNR regime, the optimality of a transmission scheme is determined by two measures: the minimum energy per bit and the wideband slope.

Definition 3. [1] A first-order optimal transmission scheme is *second-order optimal* if it achieves optimal $\ddot{R}(0)$, i.e.,

$$\begin{aligned}\dot{R}(0) &= \dot{C}(0)\,,\\ \ddot{R}(0) &= \ddot{C}(0)\,.\end{aligned}$$

It is expected that not all first-order optimal transmission schemes are second-order optimal. In the previous section, we have seen that to achieve first-order optimal performance, the receiver does not need to know the channel state information. However, the wideband slope will be considerably deteriorated if the receiver is deprived of the channel state information. For the AWGN channel defined by Equation 1.12, [1] shows that if the receiver has perfect CSI, the optimal wideband slope is

$$\mathcal{S} = \frac{2l}{m\kappa\left(\sigma_{\max}\right)} \text{ s/Hz/3dB},$$

where $\kappa\left(\right)$ is the kurtosis of a random variable, $\sigma_{\max}$ is the maximal singular value of $\mathbf{H}$, and l is the multiplicity of $\sigma_{\max}$. On the other hand, if the receiver has no CSI, the optimal wideband slope is

$$\mathcal{S} = \frac{2}{m}\frac{\left(\operatorname{Tr}\left(E\left[\mathbf{H}^{\dagger}\mathbf{H}\right]\right)\right)^2}{\operatorname{Tr}\left(E\left[\left(\mathbf{H}^{\dagger}\mathbf{H}\right)^2\right]\right)} \text{ s/Hz/3dB}.$$

Another example is that BPSK signaling is first-order optimal, but it can not withstand the test of non-zero R. [1] shows that the achievable slope of BPSK is only half of the optimal value. It means that to achieve same spectral efficiency in low-SNR regime, BPSK needs approximately twice the bandwidth.

Nevertheless, second-order optimal performance can still be achieved by a small input alphabet. [1] also shows that QPSK signaling is second-order optimal, where the input is draw from:

$$x = \begin{cases} \sqrt{P}e^{\iota\theta}, & \text{with probability } 1/4 \\ -\sqrt{P}e^{\iota\theta}, & \text{with probability } 1/4 \\ \sqrt{P}e^{\iota(\theta+\pi/2)}, & \text{with probability } 1/4 \\ -\sqrt{P}e^{\iota(\theta+\pi/2)}, & \text{with probability } 1/4 \end{cases}.$$

Figure 1.2 compares the $R\left(\frac{E_b}{N_0}\right)$ curve achieved by Gaussian signaling, and its first-order approximation. We could also see that $R\left(\frac{E_b}{N_0}\right)$ is very well approximated by its first-order approximation Equation 1.15 up until R is around 1 bit/s/Hz, which is fairly high.

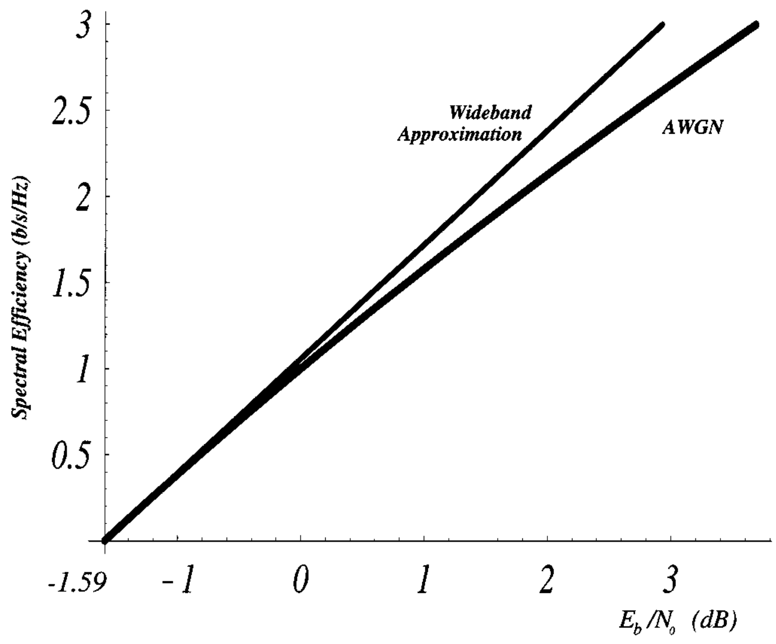

Figure 1.2: Figure 3 in [1]: $R\left(\frac{E_b}{N_0}\right)$ and its first-order approximation.

Figure 1.3 shows the $R\left(\frac{E_b}{N_0}\right)$ curves achieved by BPSK, QPSK and Gaussian signaling of a AWGN channel with scalar channel coefficient, defined as

$$y = hx + z.$$

For this channel, the channel capacity is $C = \log\left(1 + |h|^2 \mathrm{SNR}\right)$ achieved by i.i.d. Gaussian input. $\dot{C}(0) = |h|^2$, $\ddot{C}(0) = -|h|^4$, which gives

$$\begin{aligned} \left.\frac{E_b}{N_0}\right|_{\min} &= \log_e 2 / |h|^2 \\ \mathcal{S}_0 &= 2. \end{aligned}$$

[1] has shown that the performance of QPSK is second-order optimal. In Figure 1.3, we can see that QPSK is almost as good as Gaussian input for R small. On the other hand, the performance of BPSK diverges away from the optimal $R\left(\frac{E_b}{N_0}\right)$ curve as soon as $\frac{E_b}{N_0}$ is greater than $\left.\frac{E_b}{N_0}\right|_{\min}$. Notice that the $\left.\frac{E_b}{N_0}\right|_{\min}$ in the figure is -1.59dB, which is the received minimum energy per bit.

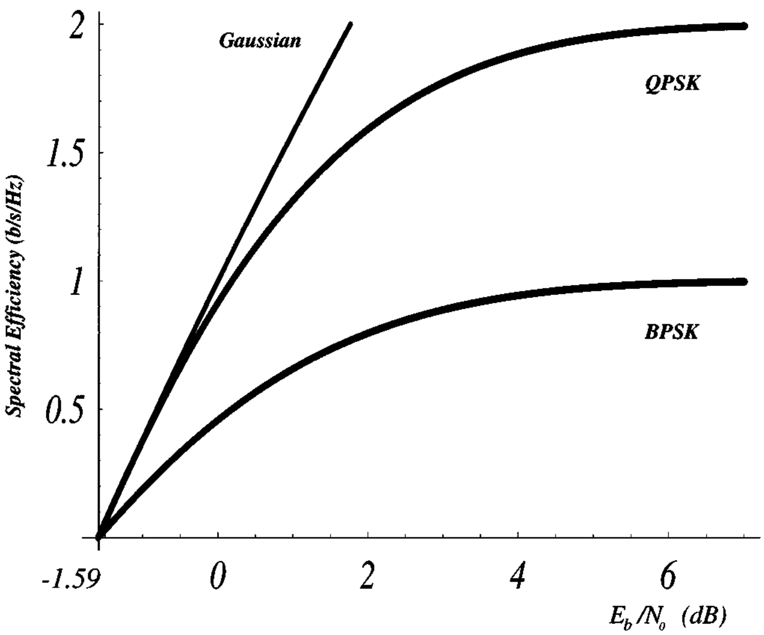

Figure 1.3: Figure 2 in [1]: $R\left(\frac{E_b}{N_0}\right)$ curves of BPSK, QPSK and Gaussian signaling of an AWGN channel.

1.3 Interference Channel Capacity And Its Low-SNR Performance

Multiple-user wireless communication is a research area of great importance from both theoretical and practical aspects. One of the basic channel models is the interference channel. It consists of K pairs of transmitters and receivers; each transmitter only wants to communicate with its corresponding receiver. Due to the broadcast nature of the wireless medium, every receiver can also hear the messages sent by other transmitters, and they will interfere with the message sent by the corresponding receiver.

Figure 1.4 illustrates an AWGN K-user interference channel with scalar channel coefficients. Receiver j wants to decode the message send by transmitter j, but does not need those sent by other transmitters. x_i denotes the input at transmitter i; y_j denotes the output at receiver j; z_j is the additive white Gaussian noise at receiver j; h_{ji} denotes the channel coefficient between transmitter i and receiver j.

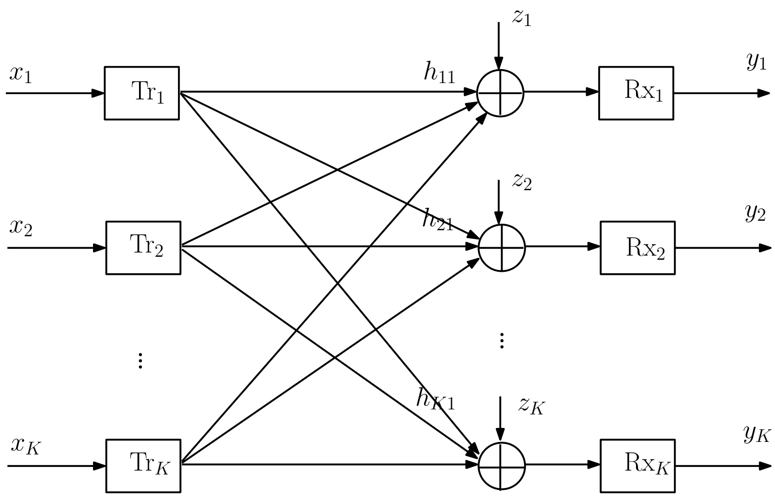

Figure 1.4: K-user AWGN interference with scalar channel coefficients

At receiver j, the received signal y_j is

$$y_j \;=\; h_{jj}x_j + \sum_{i \neq j} h_{ji}x_i + z_j, \tag{1.17}$$

where z_j is an i.i.d. circularly symmetric Gaussian noise drawn from distribution $Z_j \sim \mathcal{CN}(0,\, N_0)$. The capacity region $\mathcal{R}$ of an interference channel is defined as the union of all n-tuple $(\mathsf{R}_1,\, \mathsf{R}_2,\, \cdots,\, \mathsf{R}_K)$ achieved by any input distribution satisfying the power constraints, i.e.,

$$\mathcal{R} \;= \bigcup_{F_{X_1X_2\cdots X_K}(x_1,x_2\cdots,x_K)} \left(\mathsf{R}_1\left(P_1\right),\, \mathsf{R}_2\left(P_2\right),\, \cdots,\, \mathsf{R}_K\left(P_K\right)\right)$$

The capacity region of the interference channel is unknown in general. One better characterized case is the 2-user interference channel with scalar channel coefficients. The capacity region of this channel is known if strong interference condition $|h_{ji}|^2 > |h_{ii}|^2$ is satisfied. [10] and [11] have shown that for a 2-user strong interference channel, the optimal capacity region is equivalent to the intersection of the capacity region of the two embedded multiple-access channel formed by transmitter i, transmitter j, and receiver j; the capacity region of a multiple-access channel is known and has a close form expression. Detailed discussion can be found in [12, Chapter 15].

The capacity region of a 2-user weak interference channels where $|h_{ji}|^2 < |h_{ii}|^2$ is unknown. Research topics are divided into two main categories: good achievable schemes and tight capacity outer bounds. To characterize the capacity region precisely, the goal is to close the gap between the achievable inner bound and the outer bound.

The best achievable scheme is the Han-Kobayashi transmission scheme proposed in [13]. In this scheme, each transmitter divides the input power between two independent messages: the private message and the common message; each receiver decodes the common messages from both transmitters, and only decodes the private message of its corresponding transmitter. Besides allowing each user to split input power between private and common messages, time sharing between different users, and between the two messages of the same user are also allowed. The achievable data rate of Han-Kobayashi scheme is the union of data rate region achieved by all possible combination of the sharing parameters. The scheme itself is very complicated. In depth discussions and interpretations can be find in [14].

There are two main types of outer bounds: the broadcast outer bound and the genie-aided outer bound. The broadcast bound first took form in [11], and was later improved by Kramer in [15]. The basic idea in developing the broadcast bound is that the capacity region of a 2-user interference channel is contained within the capacity region of a 2-user degraded broadcast channel, which is known and has a close form and single-letter expression.

Some of the well known genie-aided outer bounds are [16] and [17]. The basic idea in developing a genie aided outer bound is to let a genie provide the receivers partial knowledge about the input signals. It is intuitively true that the genie-provided side information will not reduce the capacity region, because the receivers can always discard the side information. Thus the capacity region of the interference channel is contained within that of the genie-aided channel. To develop a capacity outer bound, the goal is to design good genie signals so that the capacity region of the genie-aided channel has an close form, single letter expression. None of the capacity outer bounds mentioned above is necessarily tighter than the others, their performances depend on the channel condition.

The major breakthrough in characterizing the 2-user interference channel capacity has been made by [16]. It shows that the gap between the capacity inner bound achieved by a simplified Han-Kobayashi scheme and certain capacity outer bound is within one bit. This result implies that Han-Kobayashi scheme is asymptotically optimal as SNR $\to \infty$, therefore is very powerful when the SNR value is high.

Although the performance of Han-Kobayashi scheme is exceptional, its computational complexity is high. This problem becomes particularly severe as the number of users K increases, since the computational complexity of the generalized Han-Kobayashi scheme grows exponentially with K. From practical perspective, interference management methods with low computational complexity are more attractive, such as orthogonal transmission schemes like time division multiple access (TDMA) and frequency division multiple access (FDMA), or treating interference as noise (TIN).

When it comes to the K-user interference channel, few works discusses the capacity region for the general SNR value case. Much progress has been made recently in characterizing the capacity region in the *high*-SNR *regime*. We know that the capacity of a point-to-point discrete time AWGN channel is $R = \log(1 + \mathrm{SNR})$, and for SNR large, $R \approx \log \mathrm{SNR}$. In other word, the channel capacity grows almost linearly with $\log \mathrm{SNR}$, and the system performance of a high-SNR system is determined by the *degree of freedom*(DoF) defined as

$$DoF = \lim_{SNR \to \infty} \frac{R}{\log SNR}. \tag{1.18}$$

The DoF is 1 for interference-free channel. For interference channel, people are interested in how much decrement will be caused by the interference, which can be measured by

$$\Delta DoF = \frac{DoF \text{ of the interference channel}}{DoF \text{ of the interference} - \text{free channel}}.$$

[18] has shown that any achievable ΔDoF must satisfy $\Delta DoF \leq \frac{1}{2}$. On the other hand, orthogonal transmission schemes such as TDMA only achieve ΔDoF of $\frac{1}{K}$. This gap between the achievable DoF and its outer bound is closed by Cadambe and Jafar in [19]. In their paper, a K-user interference channel with m-dimensional vector channel coefficients is considered. It shows that with probability one ΔDoF of $\frac{1}{2}$ is achievable, independent to the number of users K, which means every user can “get half of the cake”. Similar result is obtained for scalar interference channel in [20].

For K-user channel, the key technique used in achievable scheme design is called *interference alignment*. In [19], [20], and many other noticeable works discussing high-SNR regime performance of the interference channel, the alignment of interference signals is realized by using pre-coding matrices at the transmitters and zero-forcing decoding at the receivers. Under such a transmission scheme, the achievable degree of freedom of user j is equal to the rank of signal subspace, which is spanned by $\mathbf{H}_{jj}x_j$, that is linearly independent of the interference subspace, which is spanned by $\sum_{i\neq j}\mathbf{H}_{ji}x_i$. Notice that the part of the signal subspace overlapping with the interference subspace will be discarded in zero-forcing decoding.

In the high-SNR regime, the signal to noise and interference ratio (SINR) is dominantly determined by the power density of the interference signals, while in the low-SNR regime, SINR is dominantly determined by the power density of the noise. Based on this observation, we may be tempted to conjecture that interference management is not an important issue in the low-SNR regime comparing with the high-SNR regime. This conjecture is true to some extent. The results in [2] reveal that the optimal achievable minimum energy per bit $\left.\frac{E_b}{N_0}\right|_{\min}$ of an interference channel is equal to that of its corresponding interference-free channel. Therefore, at zero spectral efficiency, interference does not affect the channel capacity at all, i.e., user j achieves

$$\mathsf{R}_j \approx \frac{P_j}{N_0} G_{jj} \log_2 e,$$

same as that of the corresponding interference-free channel. Notice that

$$G_{jj} \quad \triangleq \quad \sup_{F_{\mathbf{x}_j}} \frac{E\left[\|\mathbf{H}_{jj}\mathbf{x}_j\|^2\right]}{E\left[\|\mathbf{x}_j\|^2\right]} \tag{1.19}$$

is the channel gain between transmitter j and receiver j. In the scalar interference channel depicted in Figure 1.4, $G_{jj} = |h_{jj}|^2$.

However, [2] shows that as long as R_j is non-zero, interference starts to affect performance: it will deteriorate the wideband slope. Consider the 2-user weak interference channel as an example. Chapter 3 will show that there exists an open set of channel realizations with non-zero probability, for which the wideband slope achieved by TDMA can get arbitrarily close to the wideband slope outer bound. It implies that to achieve equal data rate, the channel realizations in this set require approximately 2 times as large the bandwidth required by the interference-free channel in the low-SNR regime, when the spectral efficiency R is small but non-zero.

The wideband slope region of interference channel has been studied by G. Caire et al. in their 2004 paper [2], which mainly focused on the 2-user case. They have shown that for the 2-user channel, the best known achievable schemes are treating interference as noise and TDMA.

The main goal of this thesis is to characterize the wideband slope region of K-user interference channel, $K > 2$, which is unknown in general. The main differences between the existing discussion on the 2-user case and the discussion on the K-user case are:

- Different performance measures are used.

  In [2], the whole wideband slope region is studied. However, for K large, describing the whole slope region becomes not affordable. Therefore, a new performance measure is needed to characterize the K-user channel performance.

- Different approach to deal the propagation delay.

  In [2], the interference channel is assumed to be delay-free. This approach has no loss of generality in the 2-user case, because our discussion in chapter 3 will show that delay will not make any difference to the performance. However, the results of this thesis will demonstrate that depending on how the low-SNR regime is approached, the K-user channel can have distinct performance. Therefore, a channel model fully characterizing the role that the propagation delay plays is needed.

Existing wideband slope outer bounds developed for 2-user channel can not be easily generalized into the $K > 2$ case; achievable schemes such as treating interference as noise and TDMA which perform fairly well in the 2-user channel are not very satisfying as K increases. New outer bounds and transmission schemes that achieve more promising inner bounds will be developed in this thesis for K-user interference channel.

1.4 Overview of the thesis

The wideband slope region of K-user interference channel, $K > 2$, will be characterized in this thesis. The channel model and performance measures will be defined in chapter 2. In chapter 3, 2-user channel is discussed, the results for which serve as the essential building block for the K-user channel. The main results on K-user channel will be presented in chapter 4 and chapter 5: new transmission schemes and their achievable wideband slope will be discussed in chapter 4; wideband slope outer bounds will be discussed in chapter 5.

Chapter 2

Channel Models And Performance Measures

2.1 Summary

Three main topics will be discussed in this chapter. First, two types of K-user interference channel models will be defined. They are the large-bandwidth case and the small-bandwidth case. Second, the performance measures will be defined. Third, the first-order optimality criteria will be derived: only the achievable schemes and outer bounds satisfying this criteria can provide valid bounds on the wideband slope.

The basic derivation from a continuous-time channel model to its equivalent discrete-time channel model will be discussed in section 2.2. The derivation itself follows the standard procedure used in wireless communication, and the purpose of elaboration is to emphasize the role of propagation delay in the model. As section 1.1 has shown, the low-SNR regime is approached as SNR goes to zero; and this limit can be achieved in two ways, defined by Equation 1.5 and Equation 1.8. In Equation 1.5, the bandwidth goes to infinity while the input power is a constant. For this case, the minimum energy per bit and the wideband slope jointly determine the spectral efficiency of a system with large but finite bandwidth. This is called the large-bandwidth case. In Equation 1.8, the bandwidth is a finite constant while the input power goes to zero. For this case, the minimum energy per bit and the wideband slope jointly determine the spectral efficiency of a system with small but non-zero power. This is called the small-bandwidth case. For the large-bandwidth case, even very small propagation delay can be made relatively large comparing with the symbol duration, while for the small-bandwidth case, the propagation delay can be neglected. The K-user interference channel models under the large-bandwidth case and the small-bandwidth case will be defined in section 2.3.

Performance measures will be discussed in section 2.5. For an interference channel with more than two users it is complicated to characterize the complete slope regions. This thesis will therefore look at a single quantity to characterize performance: *the sum slope*, and this measure will be studied under two different constraints: the equal power constraint and the equal rate constraint.

As we have discussed in chapter 1, the optimal achievable minimum energy per bit $\left.\frac{E_b}{N_0}\right|_{\min}$ of an interference channel is known. Only the achievable schemes and outer bounds that achieve the optimal $\left.\frac{E_b}{N_0}\right|_{\min}$ can provide valid bounds on the wideband slope. In section 2.6, the first-order optimality criteria under equal power and equal rate constraints will be derived.

2.2 Basic K-User Interference Channel Model

This thesis considers the K-user interference channel defined by Figure 1.4. In this section, this channel model will be analyzed in detail. Assume each node is equipped with single antenna, and the channel coefficient h_{ji} between transmitter i and receiver j is a scalar complex number for all $i, j = 1, \cdots, K$. Assume that h_{ji} is known to all transmitters and receivers, and is a constant number throughout the transmission. Further, a line-of-sight channel is considered, where the propagation delay between transmitter i and receiver j is determined by their distance d_{ji}.

First, consider single transmitter-receiver pair. The channel between transmitter i and receiver j is depicted in Figure 2.1.

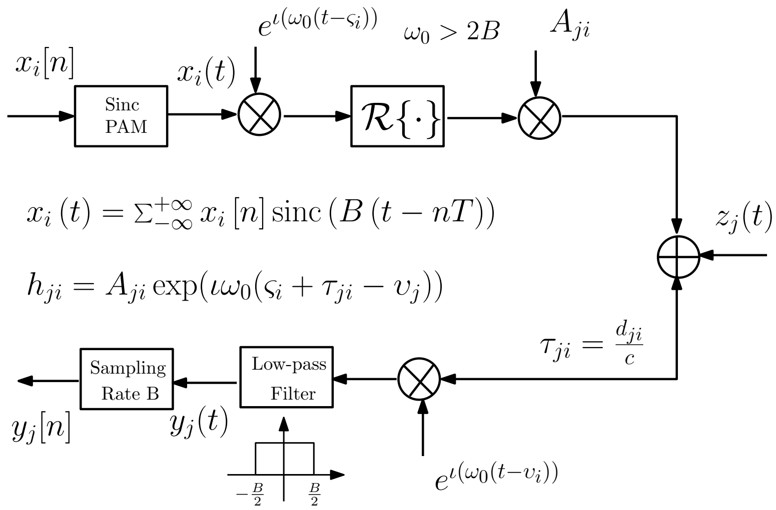

Figure 2.1: Channel between transmitter i and receiver j

Let the discrete-time base-band input signal of transmitter i be x_i. Similar to the point-to-point channel discussed in chapter 1, we require user i to satisfy a power constraint. For any length-N input sequence x_i^N,

$$\frac{1}{N}\sum_{n=1}^{N}|x_i[n]|^2 \quad \leq \quad \mathrm{SNR}_i \tag{2.1}$$

where

$$\mathrm{SNR}_i = \frac{P_i}{N_0 B}. \tag{2.2}$$

Let B be the two-sided bandwidth. The continuous-time base-band signal $x_i(t)$ is obtained from $x_i[n]$ by sinc pulse interpolation, i.e.,

$$x_i(t) \quad = \quad \sum_n x_i[n]\,\mathrm{sinc}\,(Bt-n)\,.$$

The base band signal is then modulated with the carrier signal $c(t) = \exp\iota(\omega_0(t-\varsigma_i))$, where ω_0 is the carrier frequency and ς_i is the delay (phase offset) in the oscillator at transmitter i (and $\iota = \sqrt{-1}$). Transmit the real part $s_i(t)$ of $x_i(t)$:

$$\begin{aligned} s_i(t) &= \Re\{\exp\iota(\omega_0(t-\varsigma_i))x_i(t)\} \\ &= \cos(\omega_0(t-\varsigma_i))\Re\{x_i(t)\} - \sin(\omega_0(t-\varsigma_i))\Im\{x_i(t)\}. \end{aligned}$$

At receiver j, the received signal is modulated to base band by multiplying with $\exp(-\iota\omega_0(t-\upsilon_j))$, where υ_j is the delay in the oscillator at receiver j, then passes a low-pass filter. Now the base band continuous-time signal at receiver j is

$$\begin{aligned} y_j(t) &= A_{ji}\exp(\iota\omega_0(\varsigma_i+\tau_{ji}-\upsilon_j))x_i(t-\tau_{ji}) + z_j(t), \\ \tau_{ji} &= \frac{d_{ji}}{c}, \end{aligned}$$

where A_{ji} is an attenuation factor, c is the speed of light, and τ_{ji} is the propagation delay as a result. Sample $y_j(t)$ at the Nyquist frequency $f_s = B$, the resulting equivalent discrete-time base-band signal is

$$y_j[n] \quad = \quad A_{ji}\exp(\iota\omega_0(\varsigma_i+\tau_{ji}-\upsilon_j))\tilde{x}_i[n-n_{ji}] + z_j[n] \tag{2.3}$$

where

$$\tilde{x}_i[n] \quad = \quad \sum_{m=-\infty}^{\infty} x_i[m]\mathrm{sinc}(n-m+\delta_{ji}). \tag{2.4}$$

We will also occasionally make the dependency on the fractional delay explicit as follows

$$\tilde{x}_i[n, \delta_{ji}] \quad = \sum_{m=-\infty}^{\infty} x_i[m]\mathrm{sinc}(n - m + \delta_{ji}). \tag{2.5}$$

n_{ji} and δ_{ji} are the integer part and the fractional part of the propagation delay respectively,

$$n_{ji} \quad = \quad \left\lfloor \tau_{ji}B + \tfrac{1}{2} \right\rfloor; \tag{2.6}$$

$$\delta_{ji} \quad = \quad \tau_{ji}B - \left\lfloor \tau_{ji}B + \tfrac{1}{2} \right\rfloor. \tag{2.7}$$

$z_j[n]$ is a sequence of i.i.d circularly symmetric Gaussian random variables, $Z_j \sim \mathcal{CN}(0,\ N_0)$.

Now return to the K-user interference channel. The equivalent discrete-time receive signal at receiver j is

$$y_j[n] \quad = h_{jj} \quad x_j[n - n_{jj}] + \sum_{i \neq j} h_{ji}\tilde{x}_i[n - n_{ji}] + z_j[n]. \tag{2.8}$$

As the carrier frequency is large, the phases $\exp(\iota\omega_0(\varsigma_i + \tau_{ji} - \upsilon_j))$ can be reasonably modeled as independent uniform random variables θ_{ji} over the unit circle, $h_{ji} = |h_{ji}| \exp(\iota\omega_0\theta_{ji})$.

2.3 Large Bandwidth Case And Small Bandwidth Case

Remember that the low-SNR regime is approached as SNR goes to zero, where SNR is the signal to noise ratio per second per Hz, defined as

$$\mathrm{SNR} = \frac{P}{BN_0}.$$

In section 1.1, we have presented two different cases in which SNR could tend to zero. The first case is shown in Equation 1.5, where the limit is achieved by letting B tend to infinity, while P is a finite and constant real number. We call this approach the *large bandwidth case*. The second case is shown in Equation 1.7, where the limit is achieved by letting P tend to zero, while B is a finite and constant real number. We call this approach the *small bandwidth case*.

The large bandwidth case and the small bandwidth case have very different effects on τB, which is the propagation delay τ normalized by the symbol duration $1/B$.

Under the large bandwidth case, even very small τ produces arbitrarily large τB as SNR $\to 0$. Under the small bandwidth case, τB is a finite constant and is independent from the SNR value. In a point-to-point channel, these two cases are equivalent, because the receiver can always synchronize with the transmitter and make τ equal to zero. On the other hand, in the interference channel, individuel receiver j can not synchronize with all transmitters simultaneously, as long as the condition $d_{j1} = d_{j2} = \cdots = d_{jK}$ is not true. The results of this thesis will show that when the number of users K is greater than 2, the interference channel has distinct performances under these two cases.

Given the basic channel model Equation 2.8, following assumptions are made.

- Without loose of generality, we assume that the receiver j is time-synchronized with transmitter j. Therefore $n_{jj} = 0$ and $\tau_{jj} = 0$, which implies that $\tilde{x}_j[n, \delta_{jj}] = x_j[n]$.
- Without loose of generality, we assume that the receiver j is phase-synchronized with transmitter j. Therefore $\theta_{jj} = 0$, which implies that $h_{jj} = |h_{jj}|$.
- Because the carrier frequency is large, it is reasonable to assume that the phase of h_{ji} is uniformly distributed over $[0, 2\pi]$.
- All transmitters and receivers are assumed to be placed randomly in a 2 or three dimensional space. Therefore d_{ji} and $\tau_{ji} = {}^{d_{ji}}/_{c}$ are realizations of some continuous random variables.

Definition 4. The K-user interference channel model under the large-bandwidth case is

$$y_j[n] = |h_{jj}| \quad x_j[n] + \sum_{i \neq j} h_{ji}\tilde{x}_i[n - n_{ji}] + z_j[n] \tag{2.9}$$

where $h_{ji} = |h_{ji}| \exp(\iota\omega_0\theta_{ji})$ for $i \neq j$.

In the small-bandwidth case, we will further assume that B is so small that the delays are insignificant, $n_{ji} = 0, \delta_{ji} \approx 0$. Under this assumption, $\tilde{x}_i[n - n_{ji}] \approx x_i[n]$ and the discrete-time channel Equation 2.8 becomes a delay-free channel.

Definition 5. The K-user interference channel model under the small-bandwidth case is

$$y_j[n] = |h_{jj}| \, x_j[n] + \sum_{i \neq j} h_{ji}x_i[n] + z_j[n]. \tag{2.10}$$

2.4 Equivalent 2-dimensional Real Channel

To characterize the Shannon capacity region of the channel defined by Equation 2.10, most works restrict the inputs to be circularly symmetric, i.e., the the real part of the input $\mathrm{Re}\{x_j\}$ and the imaginary part of the input $\mathrm{Im}\{x_j\}$ are i.i.d.. However, [21] shows that circularly asymmetric signaling achieves better performance if the interference channel works in the high-SNR regime. Although the specific interference alignment technique proposed in [21] is not applicable in the low-SNR regime, their result is inspiring nevertheless. In this thesis, an achievable scheme implementing circularly asymmetric signaling will be proposed, and we will see that circularly asymmetric signaling benefits the interference channel working in the low-SNR regime system as well.

In circularly asymmetric signaling, the transmitters are allowed to allocate power on real and imaginary dimensions, and the real part of the input $\mathrm{Re}\{x_j\}$ is allowed to be correlated with the imaginary part of the input $\mathrm{Im}\{x_j\}$. To fully describe a scalar complex channel where circularly asymmetric input is considered, we need an equivalent 2-dimensional real channel instead of the complex channel model Equation 2.10.

Following [21], the scalar complex channel defined by Equation 2.10 is equivalent to the two-dimensional real channel defined in the following equation:

$$\mathbf{y}_j = |h_{jj}|\,\mathbf{U}_{jj}\mathbf{x}_j + \sum_{i=1, i\neq j}^{K} |h_{ji}|\,\mathbf{U}_{ji}\mathbf{x}_i + \mathbf{z}_j, \tag{2.11}$$

where $\mathbf{U}_{ji} \triangleq \begin{pmatrix} \cos(\phi_{ji}) & -\sin(\phi_{ji}) \\ \sin(\phi_{ji}) & \cos(\phi_{ji}) \end{pmatrix}$ is a rotation matrix with angle ϕ_{ji}, and $\mathbf{z}_j$ is a 2×1 vector white Gaussian noise, with distribution $\mathcal{N}\left(0, \frac{1}{2}\mathbf{I}_{2\times 2}\right)$. The input signal $\mathbf{x}_j$ is related to the scalar complex model by:

$$\mathbf{x}_j = \begin{pmatrix} \mathrm{Re}\{x_j\} \\ \mathrm{Im}\{x_j\} \end{pmatrix}.$$

Define the average covariance matrix of the jth input as

$$\mathbf{K}_{\mathbf{x}_j} \triangleq \frac{1}{N}\sum_{n=1}^{N} \mathrm{E}\left[\mathbf{x}_j[n]\left(\mathbf{x}_j[n]\right)^{\mathrm{H}}\right]$$

for any length N input sequence $(\mathbf{x}_j[1], \mathbf{x}_j[2], \cdots, \mathbf{x}_j[N])$. Each user should satisfy power constraint $\mathrm{Tr}\left(\mathbf{K}_{\mathbf{x}_j}\right) \le \mathrm{SNR}_j$.

Channel With I.I.D. Gaussian Input

One type of commonly used input is i.i.d. Gaussian. We denote such input by subscript 'G'. Namely, the notation $\mathbf{x}_{jG}$ means that $\mathbf{x}_j[n]$ are independent and identical distributed vector Gaussian random variable, $\mathbf{x}_{jG} \sim \mathcal{N}\left(0, \mathbf{K}_{\mathbf{x}_j}\right)$, where $\mathrm{E}\left[\mathbf{x}_j[n](\mathbf{x}_j[n])^{\mathrm{H}}\right] = \mathbf{K}_{\mathbf{x}_j}$ for any n. And for the case where all users have i.i.d. Gaussian inputs, $\mathbf{y}_j[n]$ will also be i.i.d. Gaussian. Denote $\mathbf{y}_j$ when the inputs are $\mathbf{x}_{1G}, \mathbf{x}_{2G}, \cdots, \mathbf{x}_{KG}$ by $\mathbf{y}_{jG}$.

2.5 Performance Measures

In [2] the whole slope region of the interference channel in the 2-user case was analyzed. However, for interference channel with more than two users, it is complicated to compare complete slope regions, and we are therefore looking at a single quantity to characterize performance: *the sum slope*. It is defined as the first-order slope of the sum spectral efficiency $R_0\left(\frac{E_b}{N_0}\right)$, where $R_0 \triangleq \sum_{j=1}^{K} R_j$ is the sum spectral efficiency. The energy per bit

$$\frac{E_b}{N_0} \triangleq \frac{\mathrm{SNR}_0}{R_0}$$

is the overall energy per bit of the system, where $\mathrm{SNR}_0 = \frac{\sum_{j=1}^{K} P_j}{N_0 B}$. The sum slope $\mathcal{S}_0$ and its corresponding minimum energy per bit $\left.\frac{E_b}{N_0}\right|_{\min}$ are defined as

$$\left.\frac{E_b}{N_0}\right|_{\min} = \lim_{\mathrm{SNR}\to 0} \frac{\mathrm{SNR}_0}{R_0(\mathrm{SNR}_0)} \tag{2.12}$$

$$\mathcal{S}_0 \triangleq \lim_{\frac{E_b}{N_0}\to \frac{E_b}{N_0}_{\min}} \frac{R_0\left(\frac{E_b}{N_0}\right)}{10\log_{10}\frac{E_b}{N_0} - 10\log_{10}\left.\frac{E_b}{N_0}\right|_{\min}} 10\log_{10} 2. \tag{2.13}$$

Similar to Equation 1.11 and Equation 1.16, the sum slope and the corresponding minimum energy per bit can be calculated from

$$\left.\frac{E_b}{N_0}\right|_{\min} = \frac{\log_e 2}{\dot{R}_0(0)} \tag{2.14}$$

$$\mathcal{S}_0 = -\frac{2\left(\dot{R}_0(0)\right)^2}{\ddot{R}_0(0)}. \tag{2.15}$$

From the definitions above, we could see that the goal in system design becomes optimizing the sum capacity under a given sum power constraint.

However, without further constraints, this problem will give trivial solution. Because given Remark 2, the primary parameter to be optimized is $\frac{E_b}{N_0}\Big|_{\min}$. It is easy to see that the optimal $\frac{E_b}{N_0}\Big|_{\min}$ is achieved by locating all transmitted power to the user with the largest direct link gain, while all other users keep silence.

To obtain results providing more insight, two different additional constraints will be considered in this thesis:

- *The equal power constraint.* In this case we maximize the sum spectral efficiency R_0 under the constraint $P_j = {}^{P_0}/_{K}$ for all $j = 1, \cdots, K$.
- *The equal rate constraint.* In this case we minimize the total power P_0 under the constraint $R_j = {}^{R_0}/_{K}$ for all $j = 1, \cdots, K$.

The equal power constraint can be interpreted as a system where each user needs to consume energy at the same rate, e.g., so that batteries last the same duration for all users. The equal rate constraint can be interpreted as a system where we want to minimize total system energy consumption while every user in the system has a well balanced performance.

Another performance measure we will use is

$$\Delta\mathcal{S}_0 = \frac{\mathcal{S}_0}{\mathcal{S}_{0,\text{no interference}}}.$$

The quantity $\mathcal{S}_{0,\text{no inteference}}$ is the sum slope of the corresponding interference-free channel where all interference links are nulled: $h_{ji} = 0$ for all $i \neq j$, and the direct links h_{jj} is unchanged. We can interpret $\Delta\mathcal{S}_0$ as the loss in wideband slope due to interference, or equivalently $(\Delta\mathcal{S}_0)^{-1}$ as (approximately) the additional bandwidth required to overcome interference.

For comparison purposes, the results for the interference-free channel are listed as follows (directly obtained from [1, Theorem 9]):

$$\begin{aligned} \mathcal{S}_{0,\text{no interference}} &= 2\frac{\left(\sum_j |h_{jj}|^2\right)^2}{\sum_j |h_{jj}|^4} && \text{equal power constraint;} \\ \mathcal{S}_{0,\text{no interference}} &= 2K && \text{equal rate constraint.} \end{aligned} \tag{2.16}$$

Alternative Formulas To Calculate $\frac{E_b}{N_0}\Big|_{\min}$ and $\mathcal{S}_0$

Next, we propose an alternative method to calculate $\left.\frac{E_b}{N_0}\right|_{\min}$ and $\mathcal{S}_0$, which will be more convenience for the equal rate constraint.

Lemma 6. *The minimum energy per bit and wideband slope can be calculated from*

$$\begin{aligned} \frac{E_b}{N_0}_{\min} &= \dot{\mathrm{SNR}}\,(0) && (2.17) \\ \mathcal{S} &= \frac{2\dot{\mathrm{SNR}}\,(0)}{\left.\frac{d^2\mathrm{SNR}(R)}{dR^2}\right|_{R=0}} \log 2. && (2.18) \end{aligned}$$

where $\dot{\mathrm{SNR}}\,(0)$ *and* $\ddot{\mathrm{SNR}}\,(0)$ *are the first and second order derivatives at 0 of* $\mathrm{SNR}\,(R)$*,* R *in nats/dimension.*

Proof. This lemma can be proved using a technique similar to (140)~(144) in [1]. Since $\mathrm{SNR}\,(R) = R^{-1}\,(\mathrm{SNR})$ is a monotonically decreasing convex function, we have

$$\begin{aligned} \frac{E_b}{N_0}_{\min} &= \lim_{R\to 0} \frac{\mathrm{SNR}\,(R)}{R} \\ &= \dot{\mathrm{SNR}}\,(0) \log_e 2 && (2.19) \end{aligned}$$

where $\dot{\mathrm{SNR}}\,(0)$ is the derivative at 0 of $\mathrm{SNR}\,(R)$, R in nats/dimension.

$$\begin{aligned} \mathrm{SNR}\,(R) &= \dot{\mathrm{SNR}}\,(0)\,R\log_2 e + \frac{1}{2}\ddot{\mathrm{SNR}}\,(0)\,R^2\log_2 e + o\left(R^2\right) \\ \frac{E_b}{N_0} &= \frac{\mathrm{SNR}\,(R)}{R} \\ &= \dot{\mathrm{SNR}}\,(0)\log_2 e + \frac{1}{2}\ddot{\mathrm{SNR}}\,(0)\,R\log_2 e + o\,(R) && (2.20) \end{aligned}$$

given Equation 2.19 and Equation 2.20, we have

$$\frac{\frac{E_b}{N_0}}{\frac{E_b}{N_0}_{\min}} = 1 + \frac{1}{2}\frac{\ddot{\mathrm{SNR}}\,(0)}{\dot{\mathrm{SNR}}\,(0)}R + o\,(R) \qquad (2.21)$$

combining Equation 2.21 with the definition of $\left.\frac{E_b}{N_0}\right|_{\min}$ and $\mathcal{S}_0$, we have

$$\begin{aligned}
\mathcal{S} &= \lim_{\frac{E_b}{N_0}\downarrow\frac{E_b}{N_0}_{\min}} \frac{\mathcal{C}\left(\frac{E_b}{N_0}\right)}{10\log_{10}\frac{E_b}{N_0}/\frac{E_b}{N_0}_{\min}} \\
&= \lim_{\frac{E_b}{N_0}\downarrow\frac{E_b}{N_0}_{\min}} \frac{R\log_2 e}{10\log_{10}\left(1+\frac{1}{2}\frac{\ddot{\mathrm{SNR}}(0)}{\dot{\mathrm{SNR}}(0)}R+o(R)\right)} \\
&= \frac{R\log_2 e}{10\frac{\ddot{\mathrm{SNR}}(0)}{2\dot{\mathrm{SNR}}(0)}R\log_{10}e+o(R)} \\
&= \frac{2\dot{\mathrm{SNR}}(0)}{\ddot{\mathrm{SNR}}(0)}.
\end{aligned}$$

□

2.6 First-order Optimality Criteria

The results in [2] reveal that the optimal achievable minimum energy per bit $\left.\frac{E_b}{N_0}\right|_{\min}$ of an interference channel is equal to that of its corresponding interference-free channel. Given Remark 2, an capacity bound can provide valid bound on the wideband slope only if it achieves the optimal $\left.\frac{E_b}{N_0}\right|_{\min}$. The first-order optimality criteria are stated in the following theorem.

Theorem 7. *The optimal minimum energy per bit of the interference channel defined by Equation 2.8 is*

$$\frac{E_b}{N_0}_{\min} = \frac{\sum\left(|h_{jj}|^{-2}\right)}{K}\log_e 2 \tag{2.22}$$

under the equal rate constraint; and

$$\frac{E_b}{N_0}_{\min} = \frac{K\log_e 2}{\sum_{j=1}^{K}|h_{jj}|^2} \tag{2.23}$$

under the equal power constraint.

Proof. The minimum energy per bit stated in Equation 2.22 and Equation 2.23 can be achieved by TDMA, whose achievable R_0 is

$$R_j = \frac{1}{K}\log\left(1+|h_{jj}|^2 K\cdot\mathrm{SNR}_j\right).$$

The results are directly obtained from the expression of R_j above combining with Equation 1.11 and Equation 1.16 for the equal power constraint, or Equation 2.17 and Equation 2.18 for the equal power constraint.

Further Equation 2.22 and Equation 2.23 are equal to the $\left.\frac{E_b}{N_0}\right|_{\min}$ of the interference-free channel, where

$$R_j \;=\; \log\left(1 + |h_{jj}|^2\,\mathrm{SNR}_j\right).$$

Therefore we can say these $\left.\frac{E_b}{N_0}\right|_{\min}$ values are optimal, because the capacity region of an interference channel must not be larger than the capacity region of its corresponding interference-free channel. □

Given Remark 1, any achievable scheme or capacity outer bound gives valid bound on the sum slope only if it has correct $\left.\frac{E_b}{N_0}\right|_{\min}$ *values stated in Theorem 7.*

Chapter 3

2-User Channel

3.1 Summary

In this chapter, the wideband slope region of a 2-user interference channel will be studied. It is a special case of the K-user channel defined by Equation 2.8. The channel model is

$$\begin{aligned} y_1(t) &= h_{11}x_1(t) + h_{12}x_2(t-\tau_{12}) + z_1(t) \\ y_2(t) &= h_{21}x_1(t-\tau_{21}) + h_{22}x_2(t) + z_2(t), \end{aligned} \tag{3.1}$$

with equivalent discrete-time channel model

$$y_1[n] = h_{11} \quad x_1[n] + h_{12}\tilde{x}_2[n-n_{12}] + z_1[n] \tag{3.2}$$

$$y_2[n] = h_{21} \quad \tilde{x}_1[n-n_{21}] + h_{22}x_2[n] + z_2[n]. \tag{3.3}$$

The key parameter that determines the performance is $\frac{|h_{ji}|^2}{|h_{ii}|^2}$, which is the interference link gain normalized by the direct link gain for message x_i. As long as $\frac{|h_{ji}|^2}{|h_{ii}|^2} > 1$, the interference x_i can be completely eliminated at receiver j using interference decoding, and the low-SNR performance of this channel is the same as a channel without interference. This is the so called *strong interference channel*, which will be discussed in section 3.2. On the other hand, if $\frac{|h_{ji}|^2}{|h_{ii}|^2}$ is below but close to 1, i.e., $1-\epsilon < \frac{|h_{ji}|^2}{|h_{ii}|^2} < 1$ and ϵ is small, the interference corrupts the channel performance most severely; for these channels, orthogonal transmission such as TDMA is almost optimal. When ϵ increases and $\frac{|h_{ji}|^2}{|h_{ii}|^2}$ becomes farther apart from 1, the achievable performance may be improved. In particular, if the condition $\sqrt{\frac{|h_{12}|^2}{|h_{22}|^2}} + \sqrt{\frac{|h_{21}|^2}{|h_{11}|^2}} \leq 1$ is satisfied, treating interference as noise achieves the optimal wideband slope under equal power constraint. The $\frac{|h_{ji}|^2}{|h_{ii}|^2} < 1$ case is called the weak interference channel, and will be discussed in detail in section 3.3.

The methodology of developing performance bounds used for 2-user channel is applicable to general K-user channel. To obtain either outer bound or achievable inner bound for the sum slope, the first step is to find the corresponding bound on the sum capacity. Next, calculate the minimum energy per bit using Equation 2.14, and check if it satisfies the first-order optimality criteria stated in Theorem 7. Finally, if a capacity bound is first-order optimal, a corresponding bound on the sum slope can be derived using Equation 2.15.

The results on the 2-user interference channel serve as the stepping stones in developing the sum slope region of the K-user interference channel, which will be the main contribution of this thesis. Two types of outer bounds will be studied for the K-user case, both of which are based on the 2-user outer bound discussed in subsection 3.3.2. The first type of K-user outer bound is obtained from dividing the K users into $K/2$ pairs, then applying the 2-user outer bound to each pair of users. This type of pair-wise outer bound will be discussed in chapter 5, section 5.2. It is valid for general K-user channel, and particularly important for the large-bandwidth case. The second type of K-user outer bound is derived from a generalized z-channel, which is particularly useful for the small-bandwidth case. The idea of deriving the capacity outer bound of an interference channel from its corresponding z-channel is first introduced by Kramer in [15] for the 2-user case, and the low-SNR performance of this bound will be discussed in detail in subsection 3.3.2. It is important to understand the 2-user case thoroughly before we generalize this idea into the K-user case in chapter 5, section 5.3.

On the other hand, the nature of a 2-user interference channel and a K-user interference channel can be very different. In particular, the propagation delay affects the low-SNR regime performance of the 2-user channel and that of the K-user channel differently. For the 2-user channel, we will see that the existing results and our contributions are all independent of delay. Quite the contrary to the 2-user channel, our results in chapter 4 and chapter 5 will show that the low-SNR regime performances of the K-user channel under the two different cases are distinct from each other.

3.2 Strong Interference Channel

We say that the interference channel defined by Equation 3.1 is a strong interference channel if the following condition is satisfied for both $i = 2$, $j = 1$ and

$i = 1$, $j = 2$:

$$\frac{|h_{ji}|^2}{|h_{ii}|^2} > 1. \tag{3.4}$$

[10] and [11] have shown that for a 2-user strong interference channel, the optimal capacity region is equivalent to the intersection of the capacity region of the two embedded multiple-access (MAC) channel formed by transmitter i, transmitter j, and receiver j. This result is formally stated in the following theorem.

Theorem 8. *Consider a 2-user interference channel defined by (Equation 3.2). If this channel satisfies* $\frac{|h_{12}|^2}{|h_{22}|^2} > 1$ *and* $\frac{|h_{21}|^2}{|h_{11}|^2} > 1$*. Given power constraints (Equation 2.1), any achievable rate pair* (R_1, R_2) *must satisfy*

$$\begin{aligned} R_1 &\leq \log\left(1 + |h_{11}|^2 \,\mathrm{SNR}_1\right) && (3.5)\\ R_2 &\leq \log\left(1 + |h_{22}|^2 \,\mathrm{SNR}_1\right) && (3.6)\\ R_1 + R_2 &\leq \log\left(1 + |h_{11}|^2 \,\mathrm{SNR}_1 + |h_{12}|^2 \,\mathrm{SNR}_1\right) && (3.7)\\ R_1 + R_2 &\leq \log\left(1 + |h_{21}|^2 \,\mathrm{SNR}_1 + |h_{22}|^2 \,\mathrm{SNR}_1\right). && (3.8) \end{aligned}$$

Further, the optimal capacity region can be achieved by joint decoding with input codebooks generated by i.i.d. Gaussian random variables with distribution $\mathcal{N}(0, \mathrm{SNR}_j)$.

Proof. We briefly outline the proof. Take the MAC channel formed by transmitter 1, transmitter 2, and receiver 1 as an example. For any point (R_1, R_2) inside the capacity region, receiver 2 will be able to decode messages from transmitter 1. This is because if receiver 2 can decode x_2, it can compute

$$\tilde{y}_2 = h_{11}x_1 + h_{12}x_2 + {}^{z_2}\!/\!{\tfrac{|h_{21}|^2}{|h_{11}|^2}}.$$

But $\tilde{y}_2$ is less noisy than y_1 because $\frac{|h_{21}|^2}{|h_{11}|^2} > 1$. Thus, with arbitrarily small error probability, receiver 2 can decode x_1 if receiver 1 can decode x_1. Repeating this argument for $\frac{|h_{12}|^2}{|h_{22}|^2} > 1$, we find that the capacity region of the interference channel is equivalent to that of the compound MAC channel. From the argument above, it is easy to show that the capacity region is independent of delay, details are skipped here. Further discussion on MAC channel capacity can be found in [12, Chapter 15]. □

The following lemma shows that as SNR → 0, the sum capacity outer bounds

Equation 3.7 and Equation 3.8 can always be discarded, and the resulting capacity region of the interference channel is equivalent to that of the corresponding interference-free channel.

Lemma 9. *If the strong interference condition*

$$\frac{|h_{ji}|^2}{|h_{ii}|^2} \;>\; 1 \tag{3.9}$$

is satisfied, there always exists some $\epsilon > 0$*, such that if* $\mathrm{SNR}_j < \epsilon$*, then the optimal achievable capacity region of the interference channel defined by Equation 3.1 is the same as the optimal achievable region of the corresponding interference-free channel, which is*

$$\begin{aligned} R_1 &\leq \log\left(1 + |h_{11}|^2 \mathrm{SNR}_1\right) \\ R_2 &\leq \log\left(1 + |h_{12}|^2 \mathrm{SNR}_2\right) \end{aligned}$$

Proof. Compare the summation of Equation 3.5 and Equation 3.6 with the sum capacity outer bound Equation 3.7 and Equation 3.8, we can see that if the following condition

$$\frac{|h_{ji}|^2}{|h_{ii}|^2} \;\geq\; 1 + |h_{jj}|^2 \mathrm{SNR}_j \tag{3.10}$$

is satisfied for $i, j = 1\, or\, 2$, then Equation 3.7 and Equation 3.8 can be discarded, and the capacity region is only determined by Equation 3.5 and Equation 3.6. Given Equation 3.10, the conclusion in Lemma 9 follows immediately. □

It is obvious that any transmission scheme that achieves optimal capacity region is second-order optimal, i.e., it achieves both optimal minimum energy per bit and wideband slope. Given Lemma 9, we state the low-SNR performance of the strong interference channel in the following theorem.

Theorem 10. *The optimal and achievable wideband slope of the 2-user strong interference channel is*

$$\mathcal{S}_0 \;=\; 2\frac{\left(|C_{11}|^2 + |C_{22}|^2\right)^2}{|C_{11}|^4 + |C_{22}|^4} \tag{3.11}$$

under equal power constraint; and

$$\mathcal{S}_0 \;=\; 4 \tag{3.12}$$

under equal rate constraint. For both constraints,

$$\Delta\mathcal{S}_0 = 1.$$

This result is true for both the large-bandwidth case and the small bandwidth case.

Proof. Equation 3.11 and Equation 3.12 are from the discrete-time capacity of point-to-point channel:

$$R_j \quad \leq \quad \log\left(1 + |C_{jj}|^2 \,\mathrm{SNR}_j\right).$$

Combining Equation 1.11 and Equation 1.16, and the definition of equal power and equal rate constraint, we have (Equation 3.11) and (Equation 3.12). Comparing Equation 3.11 and Equation 3.12 with Equation 2.16, we have $\Delta\mathcal{S}_0 = 1$. □

3.3 Weak Interference Channel

The sum slope region of a 2-user weak interference channels where $|h_{ji}|^2 < |h_{ii}|^2$ is unknown in general. This section will discuss achievable schemes and outer bounds on the sum slope of the 2-user weak interference channel. The achievable performances will be discussed in subsection 3.3.1. For a weak interference channel working in the low-SNR regime, the best known achievable schemes are treating interference as noise and TDMA; the former scheme is better when the interference link gains are low, the later achieves better performance when the interference link gains are relatively high.

A wideband slope outer bound will be developed in subsection 3.3.2, based on the broadcast capacity outer bound developed by G. Kramer in his 2004 paper [15]. Comparing the achievable inner bound and this outer bound, we will see that as the interference link gains tend closer to the direct link gains, the gap between the achievable performance of TDMA and the outer bound can get arbitrarily small. Therefore, there exists a set of channels for whom orthogonal transmission schemes such as TDMA can be almost optimal.

3.3.1 Achievable Schemes

Common achievable schemes include interference decoding, Han-Kobayashi transmission scheme, treating interference as noise, and orthogonal transmission schemes such as TDMA. section 3.2 has shown that interference decoding achieves optimal performance in the strong interference channel; [16] has shown that the gap between the capacity inner bound achieved by a simplified Han-Kobayashi scheme and certain capacity outer bound is within one bit, therefore Han-Kobayashi scheme is asymptotically optimal as SNR $\to \infty$, and is very powerful in high-SNR regime.

However, in both schemes, the receivers need to decode the messages sent by the interfering transmitters, either completely or partially. In a weak interference channel where the interference link gain is lower than the direct link gain, decoding the interference message requires $\frac{E_b}{N_0}$ that is higher than the optimal value. Interference decoding and Han-Kobayashi scheme therefore are not first-order optimal in the weak interference channel, and do not improve the achievable sum slope.

For the 2-user interference channel, the only known first-order optimal achievable schemes are treating interference as noise and TDMA. The channel capacity region they achieved are:

1. *Treating interference as noise.*

   In this transmission scheme, the interference is treated as background noise at each receiver.

   The achievable rate pair $(R_1,\, R_2)$ of treating interference as noise is

$$R_1 \;\leq\; \log\left(1+\frac{|h_{11}|^2\,\mathrm{SNR}_1}{1+|h_{12}|^2\,\mathrm{SNR}_2}\right) \tag{3.13}$$

$$R_2 \;\leq\; \log\left(1+\frac{|h_{22}|^2\,\mathrm{SNR}_2}{1+|h_{21}|^2\,\mathrm{SNR}_1}\right), \tag{3.14}$$

   achieved by Gaussian codebook. Notice that delay does not affect the performance of treating interference as noise, because $\tilde{x}_i[n-n_{ji}]$ has same distribution as $x_i[n]$.

2. Time-Division Multiple Access.

   In this transmission scheme, the transmitters use orthogonal time slots. The achievable rate pair $(R_1,\, R_2)$ of TDMA is

$$R_1 \;\leq\; \frac{1}{2}\log\left(1+2\,|h_{11}|^2\,\mathrm{SNR}_1\right) \tag{3.15}$$

$$R_2 \;\leq\; \frac{1}{2}\log\left(1+2\,|h_{22}|^2\,\mathrm{SNR}_2\right). \tag{3.16}$$

   With delay, some care has to be taken in allocating time slots, but this does not affect performance in the limit of large code length. To see this, denote the synchronization error between transmitter j and receiver i normalized by symbol duration as m_{ji}. Let the receiver j be symbol synchronized with transmitter j, then $m_{ji} = 0\, for\, i = j$. We further assume that for $i \neq j$, $|m_{ji}| < M$ where $0 \leq M < \infty$, and both the transmitters and receivers know the value of M. Now let the cycle length of TDMA be $K \cdot N$ and assign each user one TDMA slot of N symbols. To avoid interferences in the presence of synchronization error, choose $N > 2M$ and to let each transmitter pads M dummy symbols at the beginning and M at the end of its assigned slot. We could see that $\lim_{N\to\infty} \frac{N-2M}{N} = 1$; i.e., as the block length increases, the effect of synchronization error vanish.

The minimum energy per bit achieved by treating interference as noise and TDMA can be obtained given their data rate expressions and Equation 1.11. Comparing the results with Theorem 7, it is easy to see that they are first-order optimal under both the equal power constraint and the equal rate constraint. Therefore, they provide valid inner bounds on the wideband slope. For general channel, the expressions are too complex to give much insight. We only show the wideband slope inner bound for a 2-user channel with unit direct link gains and symmetric interference link gains.

Theorem 11. *Consider a 2-user weak interference channel where* $|C_{jj}|^2 = 1$ *and* $|C_{ji}|^2 = a,\ i \neq j$. *The optimal sum slope must satisfy*

$$\mathcal{S}_0 \geq \begin{cases} 2 & \frac{1}{2} < a < 1 \\ \frac{4}{1+2a} & a \leq \frac{1}{2} \end{cases} \tag{3.17}$$

under both the equal power constraint and the equal rate constraint.

Proof. The first line in Equation 3.17 is the sum slope achieved by TDMA, and the second line is the $\mathcal{S}_0$ achieved by treating interference as noise. The calculation is straight forward given $R_j\left(\mathrm{SNR}_j\right)$ and Equation 1.16. □

We can see that when the interference link gains are low, treating interference as noise gives better performance. However, as the interference link gains increase, the performance achieved by treating interference as noise deteriorates. The critical point is $a = 1/2$. Beyond this point, the performance of TDMA becomes better than treating interference as noise. Further, on the contrary of treating interference as noise, the sum slope achieved by TDMA is constant, $\mathcal{S}_{0,\text{TDMA}} = 2$, independent from the interference link gain. Also notice that for the symmetric channel, the achievable sum slope regions are identical under equal power and equal rate constraint, but this is not the case in general.

3.3.2 Sum Slope Outer Bounds

There are two main types of outer bounds for the capacity region of interference channel: the broadcast outer bound and the genie-aided outer bound. The basic idea in developing a genie-aided outer bound is to let a genie provide one or several receivers partial knowledge about the input signals of their corresponding transmitters. Important genie-aided capacity outer bounds include [16] and [17]. In particular, [16] has developed a type of genie-aided capacity outer bound, such that the gap between this outer bound and the capacity inner bound achieved by a simplified Han-Kobayashi scheme is within one bit. Therefore this bound depicts the capacity region especially well in the high-SNR regime. However, most genie-aided outer bounds are not first-order optimal. This is because the genie-provided side-information will increase the channel gain, and reduce the transmitted minimum energy per bit as a result. Therefore they do not provide valid outer bounds on the wideband slope.

The broadcast outer bound first took form in M. Costa's 1985 paper [11], and was later improved by G. Kramer in the 2004 paper [15]. The first step in developing the broadcast outer bound is to eliminate one interference link, changing the original 2-user interference channel into an one-sided interference channel, which is called a Z-channel. See Figure 3.1.

Since eliminating one interference link enlarges the capacity region of the interference channel, the capacity region of the Z-channel is an outer bound for the capacity region of the interference channel. A delay-free channel has been discussed in [11] and [15]. They have shown that the capacity region of the Z-channel is equivalent to the capacity region of a 2-user degraded broadcast channel, which is known and has a close form expression. In the following theorem, this result will be extended into a channel with non-zero propagation delay.

Theorem 12 (Kramer's bound)**.** *Consider the 2-user interference channel defined by (Equation 3.2). Suppose that* $|h_{21}| < |h_{11}|$*. Then any achievable rate pair*

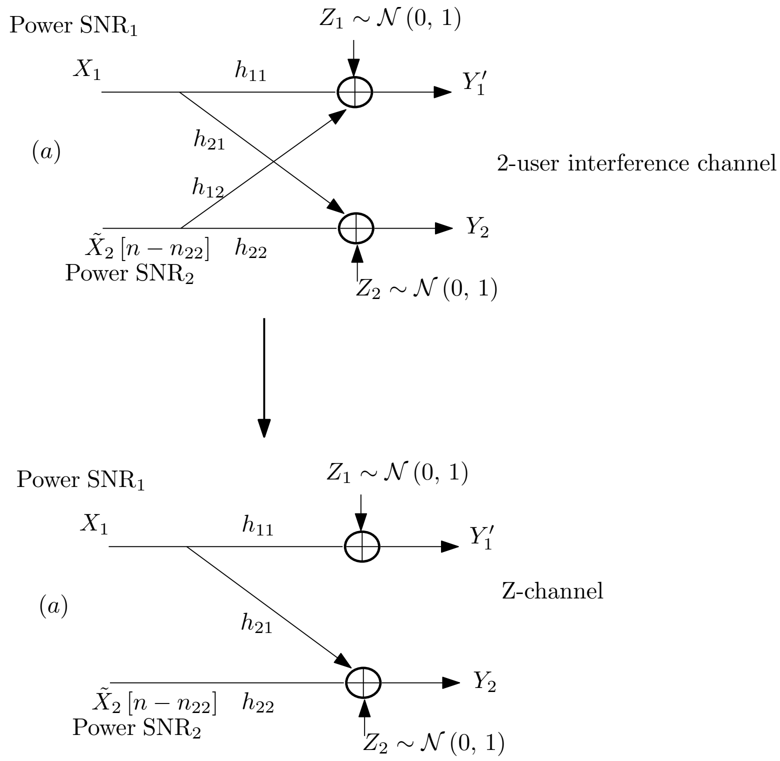

Figure 3.1: Construct Z-channel from its corresponding 2-user interference channel

(R_1, R_2) *must satisfy*

$$\begin{aligned} R_1 &\leq \log\left(1+|h_{11}|^2\,\mathrm{SNR}_1\right) && (3.18)\\ R_2 &\leq \log\left(\frac{1+|h_{22}|^2\,\mathrm{SNR}_2+|h_{21}|^2\,\mathrm{SNR}_1}{\frac{|h_{21}|^2}{|h_{11}|^2}2^{R_1}+1-\frac{|h_{21}|^2}{|h_{11}|^2}}\right) && (3.19)\\ R_1+R_2 &\leq \log\left(1+|h_{11}|^2\,\mathrm{SNR}_1\right) && (3.20)\\ &\quad+\log\left(1+\frac{|h_{22}|^2\,\mathrm{SNR}_2}{1+|h_{21}|^2\,\mathrm{SNR}_1}\right) && (3.21) \end{aligned}$$

independent of delay.

Proof. Put $h_{12}=0$ to enlarge the capacity region. Now assume that, different from the system model Equation 2.9, receiver 2 also samples the received signal synchronously with the transmitted signal of user 1. A Z-channel with delay is formed:

$$\begin{aligned} y_1'[n] &= h_{11}x_1[n]+z_1[n]\\ y_2[n] &= h_{22}\tilde{x}_2[n-n_{22}]+h_{21}x_1[n]+z_2[n] && (3.22) \end{aligned}$$

where $\tilde{x}_2[n]$ is defined by Equation 2.4.

Next, we show that the capacity region of Equation 3.22 is independent of delay. The channel (a) and (b) illustrated in Figure 3.2 have identical capacity region because $p\left(\check{y}_2 | x_1, \tilde{x}_2\right)$ and $p\left(y_2 | x_1, \tilde{x}_2\right)$ have the same distribution. $\check{z}_2[n]$ is i.i.d Gaussian noise independent of $z_1[n]$ and the input signals, with power $\left(1 - \frac{|h_{21}|^2}{|h_{11}|^2}\right) N_0 B$. Because $\frac{|h_{21}|^2}{|h_{11}|^2} < 1$, such $\check{z}_2[n]$ is guaranteed to exist. The argument is identical to (a)~(c) of Figure 6 in [11]. Details are skipped here.

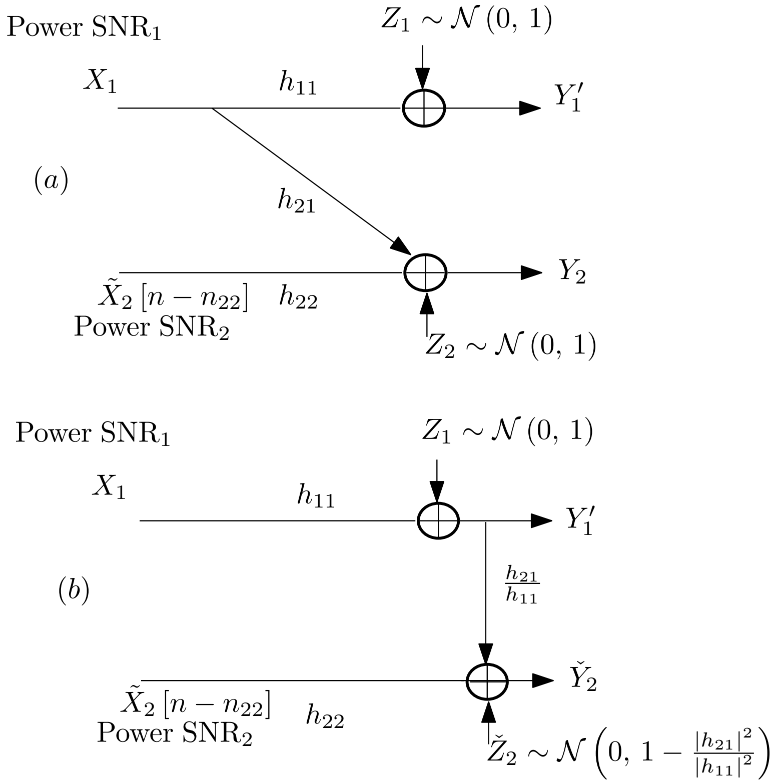

Figure 3.2: Channels with Equivalent Capacity Region

The channel (b) has the form

$$
\begin{aligned}
y_1'[n] &= h_{11} x_1[n] + z_1[n] && (3.23) \\
\check{y}_2[n] &= h_{22} \tilde{x}_2[n - n_{22}] + \frac{h_{21}}{h_{11}} y_1'[n] + \check{z}_2[n]. && (3.24)
\end{aligned}
$$

Using Fano's inequality, we can now bound the capacity of this channel by

$$\begin{aligned}
nR_2 - n\epsilon_n &\leq h(\check{y}_2^n) - h(\check{y}_2^n|\tilde{x}_2) && (3.25)\\
&\leq h(\check{y}_2^n) - h(\check{y}_2^n|w_2) && (3.26)\\
&= h(\check{y}_2^n) - h\left(\frac{h_{21}}{h_{11}}y_1'^n + \check{z}_2^n|w_2\right) && (3.27)
\end{aligned}$$

where w_2 is the message sent by transmitter 2. Equation 3.25 to Equation 3.24 is from data processing inequality, as $\tilde{x}_2$ is a function of the transmitted codeword x_2, which is a function of w_2. The second term in Equation 3.27 is independent of delay, and can be lower bounded by the entropy power inequality [22]. The first term can be upper bounded by the delay-free case. Therefore, the capacity region of Equation 3.22 is identical to that of the channel without delay.[11] and [15] show that the capacity region of delay-free channel can be derived from an equivalent degraded broadcast channel. Given Theorem 1 in [18] its rate region has upper bound Equation 3.18. Finally, it is easy to see that the capacity region of Equation 3.2 is contained within that of the Z-channel. Equation 3.21 is a restatement of (47) in [15]. □

Results similar to Theorem 12 can be obtained for the $|h_{12}|^2 < |h_{22}|^2$ case by interchanging the indices $'1'$ and $'2'$. Next, we derive the sum slope outer bound based on the capacity outer bound defined by Theorem 12.

Sum Slope Outer Bound Under Equal Power Constraint

Before deriving the outer bound on the wideband slope, we need to prove the following lemma first.

Lemma 13. *For the 2-user interference channel defined by Equation 3.1, if $|h_{ji}|^2 < |h_{ii}|^2$ for $j = 1, i = 2$ or $j = 2, i = 1$, then the capacity outer bound defined in Theorem 12 is first-order optimal under equal power constraint.*

Proof. Because Theorem 12 is a capacity outer bound, it can provide a lower bound on the $\frac{E_b}{N_0}\Big|_{\min}$ of the interference channel. Given Theorem 12, the minimum energy per bit of the interference channel must satisfy

$$\frac{E_b}{N_0}_{\min} \geq \frac{K \log_e 2}{|C_{11}|^2 + |C_{22}|^2} \tag{3.28}$$

under the equal power constraint. Comparing the right hand side of Equation 3.28 and Theorem 7, we can see that the outer bound defined by Theorem 12 is first-order optimal. □

Given Lemma 13, a wideband slope outer bound can be derived from the capacity outer bound defined in Theorem 12.

Corollary 14. *Consider the 2-user interference channel satisfying* $|h_{21}|^2 < |h_{11}|^2$*. Under the equal power constraint, any achievable sum slope must satisfy*

$$\mathcal{S}_0 \leq 2\frac{(|h_{11}|^2 + |h_{22}|^2)^2}{2|h_{21}|^2|h_{22}|^2 + |h_{11}|^4 + |h_{22}|^4} \tag{3.29}$$

$$\Delta\mathcal{S}_0 \leq \frac{1}{2\frac{|h_{21}|^2|h_{22}|^2}{|h_{11}|^4+|h_{22}|^4} + 1}, \tag{3.30}$$

independent of delay, valid for both the small-bandwidth case and the large-bandwidth case. Furthermore, if additionally $|h_{12}|^2 > |h_{22}|^2$ *this bound is achieved by treating interference as noise.*

Proof. This result can be easily shown combining Equation 3.21 and the formulas Equation 2.14 and Equation 2.15. Details are skipped here. Because the outer bound is independent from delay, it is valid for both the small-bandwidth case and the large-bandwidth case, given the definitions of these two cases in chapter 2. □

Results similar to Corollary 14 can be obtained for $|h_{12}|^2 < |h_{22}|^2$ case by interchanging the indices $'1'$ and $'2'$.

Sum Slope Outer Bound Under Equal Rate Constraint

Next, we consider the equal rate case. Different from the equal power constraint, for the equal-rate constraint if only one interference link is weak, bound Equation 3.18 is not first-order optimal, *therefore cannot be used for bounding the wideband slope.* Only when $|h_{ji}|^2 < |h_{ii}|^2$ for both $j = 1$, $i = 2$ and $j = 1$, $i = 2$, the outer bound defined in Theorem 12 is first-order optimal. When both interference links are weak, we have following corollary.

Theorem 15 (Kramer's bound)**.** *Suppose that* $|h_{ji}|^2 < |h_{ii}|^2$ *for both* $j = 1$, $i = 2$ *and* $j = 1$, $i = 2$*. Then*

$$R_j \leq \min\Big\{\log\left(1 + |h_{jj}|^2\,\mathrm{SNR}_j\right), \tag{3.31}$$

$$\log\left(\frac{1 + |h_{jj}|^2\,\mathrm{SNR}_j + |h_{ji}|^2\,\mathrm{SNR}_i}{\frac{|h_{ji}|^2}{|h_{ii}|^2}2^{R_i} + 1 - \frac{|h_{ji}|^2}{|h_{ii}|^2}}\right)\Big\}, \tag{3.32}$$

$$R_j + R_i \leq \log\left(1 + |h_{jj}|^2\,\mathrm{SNR}_i\right) \tag{3.33}$$

$$+\log\left(1 + \frac{|h_{ii}|^2\,\mathrm{SNR}_i}{1 + |h_{ij}|^2\,\mathrm{SNR}_j}\right) \tag{3.34}$$

for both $j = 1, i = 2$ and $j = 2, i = 1$, independent of delay.

Proof. This result is directly from Theorem 12. □

Next, we are going to show that if $|h_{ji}|^2 < |h_{ii}|^2$ for $j = 1$, $i = 2$ and $j = 2$, $i = 1$, then the capacity outer bound defined in Theorem 15 is first-order optimal under equal rate constraint.

Lemma 16. *For the 2-user interference channel defined by Equation 3.1, the minimum energy per bit of its capacity outer bound defined in Theorem 15 is*

$$\frac{E_b}{N_0}_{\min} = \frac{|C_{11}|^{-2} + |C_{22}|^{-2}}{K} \log_e 2 \tag{3.35}$$

under the equal rate constraint.

Proof. This result can be easily shown combining Equation 3.21 and Equation 2.14. Details are skipped here. □

Remark 17. The capacity outer bound defined in Theorem 15 is first-order optimal under equal rate constraint.

Proof. Equation 3.35 is identical to the optimal and achievable minimum energy per bit of the 2-user interference channel stated in Theorem 7. □

Given Remark 1 and Remark 6, we know that a wideband slope outer bound can be derived from the capacity outer bound defined in Theorem 15.

Corollary 18. *Suppose that $|h_{21}|^2 < |h_{11}|^2$ and $|h_{12}|^2 < |h_{22}|^2$. Under the equal rate constraint, any achievable sum slope must satisfy*

$$\begin{aligned} \mathcal{S}_0 \leq\ & 4 \cdot \left(|h_{11}|^2 + |h_{22}|^2\right)\left(1 - \frac{|h_{12}|^2}{|h_{22}|^2}\frac{|h_{21}|^2}{|h_{11}|^2}\right) \cdot \Bigg(|h_{11}|^2 + |h_{22}|^2 + \\ & |h_{21}|^2\left(2 - 3\frac{|h_{12}|^2}{|h_{22}|^2}\right) + |h_{12}|^2\left(2 - 3\frac{|h_{21}|^2}{|h_{11}|^2}\right)\Bigg)^{-1} \\ \Delta\mathcal{S}_0 \leq\ & \left(|h_{11}|^2 + |h_{22}|^2\right)\left(1 - \frac{|h_{12}|^2}{|h_{22}|^2}\frac{|h_{21}|^2}{|h_{11}|^2}\right) \cdot \Bigg(|h_{11}|^2 + |h_{22}|^2 + \\ & |h_{21}|^2\left(2 - 3\frac{|h_{12}|^2}{|h_{22}|^2}\right) + |h_{12}|^2\left(2 - 3\frac{|h_{21}|^2}{|h_{11}|^2}\right)\Bigg)^{-1} \end{aligned}$$

independent of delay, valid for both the small-bandwidth case and the large-bandwidth case.

Proof. Equation 3.18 gives

$$|h_{21}|^2\,\mathrm{SNR}_1 + |h_{22}|^2\,\mathrm{SNR}_2 \;\geq\; 2^{R_2}\left(\frac{|h_{21}|^2}{|h_{11}|^2}2^{R_1} - \frac{|h_{21}|^2}{|h_{11}|^2} + 1\right) - 1 \quad (3.36)$$

$$|h_{11}|^2\,\mathrm{SNR}_1 + |h_{12}|^2\,\mathrm{SNR}_2 \;\geq\; 2^{R_1}\left(\frac{|h_{12}|^2}{|h_{22}|^2}2^{R_2} - \frac{|h_{12}|^2}{|h_{22}|^2} + 1\right) - 1. \quad (3.37)$$

Under the equal rate constraint, $R_1 = R_2 = \frac{R_s}{2}$, and our objective is to minimize $\mathrm{SNR}_1 + \mathrm{SNR}_2$. We construct following optimization problem

$$\begin{aligned} \min \quad & \mathrm{SNR}_1 + \mathrm{SNR}_2 \\ \text{s.t.} \quad & \mathbf{A}\begin{pmatrix} \mathrm{SNR}_1 \\ \mathrm{SNR}_2 \end{pmatrix} \geq \mathrm{b} \\ & P_j \geq 0 \end{aligned}$$

where $\mathbf{A} = \begin{pmatrix} |h_{21}|^2 & |h_{22}|^2 \\ |h_{11}|^2 & |h_{12}|^2 \end{pmatrix}$ and $\mathrm{b} = \begin{pmatrix} 2^{R_s/2}\left(\frac{|h_{21}|^2}{|h_{11}|^2}2^{R_s/2} - \frac{|h_{21}|^2}{|h_{11}|^2} + 1\right) - 1 \\ 2^{R_s/2}\left(\frac{|h_{12}|^2}{|h_{22}|^2}2^{R_s/2} - \frac{|h_{12}|^2}{|h_{22}|^2} + 1\right) - 1 \end{pmatrix}$. Using simple linear programming principles, one optimal solution can be found at the vertex of the feasible region. That is, $\mathrm{SNR}_1 + \mathrm{SNR}_2|_{\min} = \mathrm{SNR}_{1o} + \mathrm{SNR}_{2o}$ where $\begin{pmatrix} \mathrm{SNR}_{1o} \\ \mathrm{SNR}_{2o} \end{pmatrix} = \mathbf{A}^{-1}\mathrm{b} > 0$. We solve this simple linear system and get

$$\mathrm{SNR}_{1o} \quad (3.38)$$

$$= |h_{11}|^{-2} \cdot \frac{2^{\frac{R_s}{K}}\left(2^{\frac{R_s}{K}} - 1\right)\frac{|h_{12}|^2}{|h_{22}|^2}\left(1 - \frac{|h_{21}|^2}{|h_{11}|^2}\right) + \left(1 - \frac{|h_{21}|^2}{|h_{11}|^2}\right)\left(2^{\frac{R_s}{K}} - 1\right)}{1 - \frac{|h_{12}|^2}{|h_{22}|^2}\frac{|h_{21}|^2}{|h_{11}|^2}} \quad (3.39)$$

$$\mathrm{SNR}_{2o} \quad (3.40)$$

$$= |h_{22}|^{-2} \cdot \frac{2^{\frac{R_s}{K}}\left(2^{\frac{R_s}{K}} - 1\right)\frac{|h_{21}|^2}{|h_{11}|^2}\left(1 - \frac{|h_{12}|^2}{|h_{22}|^2}\right) + \left(1 - \frac{|h_{12}|^2}{|h_{22}|^2}\right)\left(2^{\frac{R_s}{K}} - 1\right)}{1 - \frac{|h_{12}|^2}{|h_{22}|^2}\frac{|h_{21}|^2}{|h_{11}|^2}}. \quad (3.41)$$

Now we have the expression of sum power as a function of sum rate. They can be proved using the technique similar to (140)~(144) in [1]. Details are skipped here. Combining Equation 3.38, Equation 3.40 and Equation 2.17 and Equation 2.18, we have

$$
\begin{aligned}
\frac{E_b}{N_0}_{\min} &= \frac{\left(|h_{11}|^{-2} + |h_{22}|^{-2}\right)}{2} \log_e 2 \\
\mathcal{S}_0 &= 4 \cdot \left(|h_{11}|^2 + |h_{22}|^2\right) \left(1 - \frac{|h_{12}|^2}{|h_{22}|^2} \frac{|h_{21}|^2}{|h_{11}|^2}\right) \cdot \left(|h_{11}|^2 + |h_{22}|^2 \right. \\
&\quad \left. + |h_{21}|^2 \left(2 - 3\frac{|h_{12}|^2}{|h_{22}|^2}\right) + |h_{12}|^2 \left(2 - 3\frac{|h_{21}|^2}{|h_{11}|^2}\right)\right)^{-1}
\end{aligned}
$$

□

3.4 Noisy Interference Channel

Beside the strong interference channel, there is another type of interference channel for whom the wideband slope region is known. It is the noisy interference channel under equal power constraint, and the optimal wideband slope is achieved by treating interference as noise.

Intuitively, if the interference link gains are very weak, then the system can obtain good performance by treating interference as background noise at the receivers. [17], [23] and [24] show that there exists a class of channels whose optimal sum capacity can be achieved by using input codebooks generated from i.i.d. Gaussian distribution at the transmitters, and treating interference as noise at the receivers. This class of channels is called the noisy interference channel. In this section, we discuss the sum slope of the noisy interference channel. The channel model in [17], [23] and [24] is delay-free. In the following theorem, we will show that their results still hold for the more general channel model defined in Equation 3.2. Thus the results in this section are valid for both the large bandwidth case and the small bandwidth case.

Theorem 19. *For a 2-user interference channel defined by Equation 3.2, if there* *exist complex numbers ρ_1,ρ_2 and positive real numbers σ_1^2, σ_2^2 such that,*

$$|\rho_1|^2 \le \sigma_1^2 \le 1 - \frac{|\rho_2|^2}{\sigma_2^2} \tag{3.42}$$

$$|\rho_2|^2 \le \sigma_2^2 \le 1 - \frac{|\rho_1|^2}{\sigma_1^2} \tag{3.43}$$

$$h_{21} = \frac{\rho_1 h_{11}}{|h_{12}|^2 \,\mathrm{SNR}_2 + 1} \tag{3.44}$$

$$h_{12} = \frac{\rho_1 h_{11}}{|h_{12}|^2 \,\mathrm{SNR}_2 + 1}, \tag{3.45}$$

then the optimal sum capacity is

$$R_1 + R_2 \quad \le \quad \log\left(1 + \frac{|h_{11}|^2\mathrm{SNR}_1^2}{1+|h_{12}|^2\mathrm{SNR}_2}\right) + \log\left(1 + \frac{|h_{22}|^2\mathrm{SNR}_2^2}{1+|h_{21}|^2\mathrm{SNR}_1}\right) \tag{3.46}$$

which is achievable by i.i.d. Gaussian input and treating interference as noise at the receivers. Further, Equation 3.42 ~ Equation 3.45 can be satisfied as long as the inequality below holds

$$\sqrt{\frac{|h_{12}|^2}{|h_{22}|^2}}\left(1+|h_{21}|^2\mathrm{SNR}_1\right) + \sqrt{\frac{|h_{21}|^2}{|h_{11}|^2}} \;\; (1+|h_{12}|^2\mathrm{SNR}_2) \le \;\; 1. \tag{3.47}$$

.

Proof. Please see the Appendix section A.1. □

Theorem 19 is identical to the case where the delays are strictly zero, which is discussed in [25, Theorem 6]. Next, we derive noisy interference condition for the low-SNR regime under equal power constraint.

Corollary 20. *Consider the 2-user interference channel defined by Equation 2.9. Under equal power constraint, if the channel coefficients satisfy*

$$\sqrt{\frac{|h_{12}|^2}{|h_{22}|^2}} + \sqrt{\frac{|h_{21}|^2}{|h_{11}|^2}} \;\; < \;\; 1,$$

then i.i.d. Gaussian inputs and treating interference as noise achieve optimal sum *slope $\mathcal{S}_0$, which is*

$$\mathcal{S}_0 = \frac{2\left(|h_{11}|^2 + |h_{22}|^2\right)^2}{|h_{11}|^4 + |h_{22}|^4 + 2\left(|h_{11}|^2|h_{12}|^2 + |h_{21}|^2|h_{22}|^2\right)} \tag{3.48}$$

$$\Delta\mathcal{S}_0 = 1 + \frac{2\left(|h_{11}|^2|h_{12}|^2 + |h_{21}|^2|h_{22}|^2\right)}{|h_{11}|^4 + |h_{22}|^4}. \tag{3.49}$$

Proof. Under the equal power constraint where $\mathrm{SNR}_i = \frac{\mathrm{SNR}_s}{2}$ there must exist some $\epsilon > 0$, such that if $\mathrm{SNR} < \epsilon$ then Equation 3.47 can be satisfied. Because the low-SNR regime is approached as $\mathrm{SNR} \to 0$, the noisy interference condition given by Equation 3.47 in Theorem 19

$$\sqrt{\frac{|h_{12}|^2}{|h_{22}|^2}} + \sqrt{\frac{|h_{21}|^2}{|h_{11}|^2}} < 1$$

Given Equation 3.46, under equal power constraint the sum rate achieved by treating interference as noise is

$$R_s \leq \log\left(1 + \frac{|h_{11}|^2 \frac{\mathrm{SNR}}{2}}{1 + |h_{12}|^2 \frac{\mathrm{SNR}}{2}}\right) + \log\left(1 + \frac{|h_{22}|^2 \frac{\mathrm{SNR}}{2}}{1 + |h_{21}|^2 \frac{\mathrm{SNR}}{2}}\right). \tag{3.50}$$

Combining Equation 3.50 with Equation A.8 and Equation A.9 we have Equation 3.48.

□

3.5 Conclusion

The key parameter that determines the performance of the 2-user interference channel is $\frac{|h_{ji}|^2}{|h_{ii}|^2}$, which is the interference link gain normalized by the direct link gain for message x_i. If $\frac{|h_{ji}|^2}{|h_{ii}|^2} > 1$, the channel is a strong interference channel. section 3.2 shows that the low-SNR performance of a strong interference channel channel is the same as the corresponding interference-free channel; a formal description of its sum slope region can be found in Theorem 10.

If $\frac{|h_{ji}|^2}{|h_{ii}|^2} < 1$, the channel is a weak interference channel. The sum slope region of a 2-user weak interference channel is unknown in general. The sum slope inner bound and outer bound of the weak interference channel is discussed in subsection 3.3.1

and subsection 3.3.2, respectively. Comparing these bounds, we can see that the sum slope inner bound achieved by treating interference as noise can get arbitrarily close to the outer bound as the interference link gains tend to zero. The sum slope inner bound achieved by TDMA can get arbitrarily close to the outer bound as the interference link gains tend to their corresponding direct link gains. Based on this observation, we have the following remark.

Remark 21. For any $\delta > 0$, there exists a non empty set of 2-user interference channels for whom the gap between the sum slope inner bound achieved by either treating interference as noise or TDMA and the sum slope outer bound is less than δ. Further, if the channel coefficients are drawn from some continuous random distribution, then the probability that a channel realization falls inside this set is greater than zero.

Further, the noisy interference condition under equal power constraint has been developed in section 3.4. When the channel realization satisfies this condition, treating interference as noise achieves optimal wideband slope.

Figure 3.3 illustrates the sum slope region of a 2-user interference channel with unit direct link gain, and symmetric cross link gain, that is, $|h_{11}|^2 = |h_{22}|^2 = 1$, and $|h_{12}|^2 = |h_{21}|^2 = a$. In this figure, the inner bound is given by Theorem 11, and the outer bound is derived from Corollary 14 . Given Corollary 20, if $a \leq \frac{1}{4}$, treating interference as noise achieves optimal sum slope, i.e., the inner bound is tight.

Notice that whether the interference link gain a is below or above 1 is critical. If $a > 1$, interference does not affect the performance, given the discusssion in section 3.2; if a slightly below 1, the channel will be affected by interference most severely. One could wonder if, for the K-user case, the former fact could be used effectively. It turns out that is not the case. In the next chapter, we will show that in a K-user interference channel when K is large, with high probability each user will form a 2-user weak interference pair where a is just below 1 with some other user. And the performance of the K users is limited by $\frac{K}{2}$ such 2-user weak interference channels.

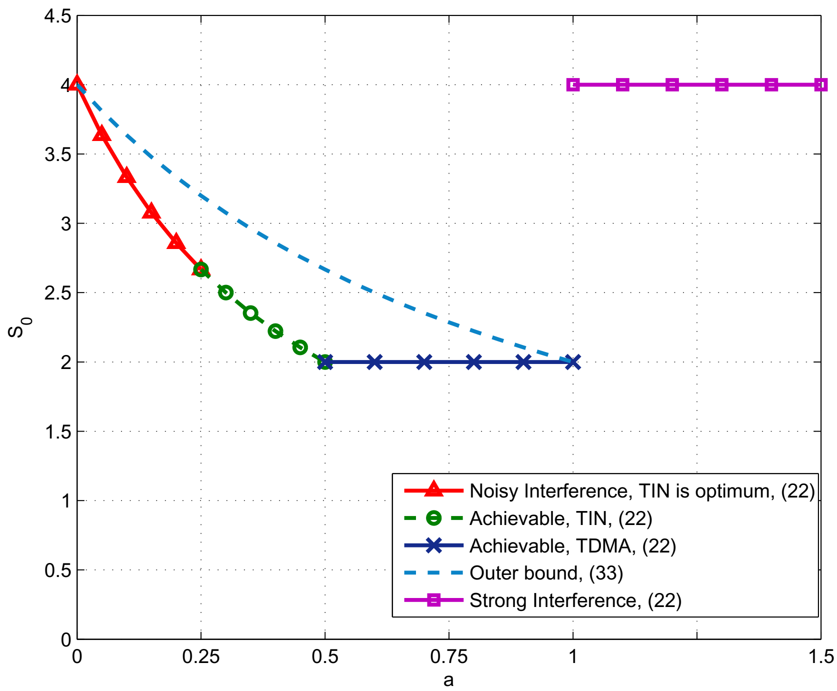

Figure 3.3: Wideband slope region of a symmetric 2-user interference channel versus $\frac{|h_{ji}|^2}{|h_{ii}|^2}$.

Chapter 4

K-User Interference Channel: Interference Alignment Transmission

4.1 Summary

This chapter focuses on designing achievable schemes for scalar AWGN K-user interference channel working in the low-SNR regime. Most of the existing works about the K-user channel focus on its high-SNR regime performance or its capacity region for general SNR value. Important works discussing its high-SNR regime performance include [19], [21], [26] and [27]. For K-user channel with general SNR value, the capacity bound for certain channel realizations is discussed in [28]. [29] considers deterministic channel where there is no background noise and interference from non-corresponding users is the only factor that deteriorates the performance. Deterministic channel can be seen as an extreme case of the high-SNR regime channel where the noise can be neglected as the SNR grows large.

However, when it comes to the low-SNR regime, there is few existing works, and the best known achievable schemes are treating interference as noise and TDMA. Let us demonstrate the performance of treating interference as noise and TDMA in the K-user case with an example.

Consider a K-user weak interference channel with unit direct link gains and symmetric interference link gains, i.e., $|h_{jj}|^2 = 1$ and $|h_{ji}|^2 = a$ for $i \neq j$, $a < 1$ as an example. For this channel, equal rate constraint and equal power constraint are equivalent for TIN and TDMA. Remember that we denote the sum spectral efficiency by R_0 and denote the sum signal to noise ratio per second per Hz by SNR_0. The $R_0\left(\mathrm{SNR}_0\right)$ achieved by TIN is

$$R_0\left(\mathrm{SNR}_0\right) = K\log\left(1+\frac{\mathrm{SNR}_0}{K+a\cdot(K-1)\,\mathrm{SNR}_0}\right).$$

Combining with Equation 2.15, TIN achieves

$$\mathcal{S}_0 = \frac{2K}{1+2a\,(K-1)}; \tag{4.1}$$

$$\Delta\mathcal{S}_0 = \frac{1}{1+2a\,(K-1)} \tag{4.2}$$

The $R_0\,(\mathrm{SNR}_0)$ achieved by TDMA is

$$R_0\,(\mathrm{SNR}_0) = \log\left(1+\mathrm{SNR}_0\right).$$

Combining with Equation 2.15, TDMA achieves

$$\mathcal{S}_0 = 2 \tag{4.3}$$

$$\Delta\mathcal{S}_0 = \frac{1}{K} \tag{4.4}$$

It is obvious that in a small neighborhood around $a = 0$, treating interference as noise is almost optimal, i.e. close to the performance of an interference-free channel. However, when the number of user K is large, its performance deteriorates very fast. Comparing Equation 4.1 and Equation 4.3, we can see that for all K value, the wideband slope achieved by treating interference as noise and TDMA meet at $a = \frac{1}{2}$. When $1/2 < a < 1$, TDMA has better performance. Therefore when the interference link gains are moderate, to achieve the same spectral efficiency, the system requires K times the bandwidth required by the interference-free channel.

For 2-user case, comparing with a wideband slope outer bound, we can tell that treating interference as noise achieves almost optimal performance if a is close to zero, and TDMA achieves almost optimal performance if a is close to 1. Is this also true for the K-user case?

In high-SNR regime, many noticeable works show that interference alignment can achieve scalable performance. In particular, [19] has designed an pre-coding and zero-forcing transmission scheme for a K-user MIMO interference channel, which achieves ΔDoF of $\frac{1}{2}$ for all $K > 2$, i.e., the performance of the K-user channel is the same as a 2-user channel. Although existing interference alignment algorithms are not applicable in the low-SNR regime, it still is tempting to conjecture that there exists first-order optimal interference alignment schemes that perform better than treating interference as noise and TDMA. This chapter will propose achievable schemes for both the large bandwidth case and the small bandwidth case. Their performances support this conjecture for almost all channel realizations under the large bandwidth case, and for a set of channel realizations with non-zero probability under the small bandwidth case.

In section 4.2, we will briefly review the concept of interference alignment, which is an important interference management mechanism for K-user interference channel. Common interference alignment techniques used in moderate to high SNR interference channel will be introduced. Then the challenges and opportunities in designing interference alignment algorithms for low-SNR regime systems will be discussed.

In section 4.3, we will discuss the large bandwidth case, where the low-SNR regime is approached by letting the bandwidth B go to infinity while the power P be fixed and finite. In this case, even very small propagation delay can become arbitrarily large comparing with symbol duration B^{-1}. Contrary to the 2-user case whose performance is independent of delay, K-user channel can use delay to its advantage: when K is greater than 2, extra freedom for system design comes into existence. A type of interference alignment scheme over time domain will be proposed, which achieves $\Delta\mathcal{S}_0 = \frac{1}{2}$ with probability one under the assumption that all transmitters and receivers are placed randomly in a 2 or 3 dimensional space. Comparing with TDMA which achieves $\Delta\mathcal{S}_0 = \frac{1}{K}$, the achievable scheme we proposed improves the performance by a factor of $K/2$.

In section 4.4, we will study the small bandwidth case, where the low-SNR regime is approached by letting the input power P goes to zero while the bandwidth is fixed and finite. We further assume that B is so small as to render propagation delays negligible comparing with the symbol duration. A phase interference alignment scheme will be proposed. In this scheme, transmitter j only transmits in the direction $e^{\iota\theta_j}$, instead of using the complete 2-dimensional complex plane. Simulation results will show that when the interference link gains are moderate, this interference alignment scheme performs better than existing transmission schemes such as treating interference as noise and TDMA with high probability, under the assumption that the phase of the channel coefficient h_{ji} is drawn from a continuous random distribution over $[-\pi, \pi]$.

4.2 Interference Management For K-user Interference Channel

From chapter 1 and chapter 3, we have seen that for 2-user interference channel, the best known low-SNR regime performance is achieved by traditional transmission schemes such as treating interference as noise and TDMA. The sophisticated Han-Kobayashi transmission scheme is known to be asymptotically optimal in high-SNR regime [16]. However, in the low-SNR regime, it is not first-order optimal, i.e., the minimum energy per bit achieved by the Han-Kobayashi scheme is higher than the optimal value. Therefore it does not improve the wideband slope.

System design faces similar challenge when the number of of users is greater than 2. Most of the efficient interference alignment transmission schemes which have good performance in the high-SNR regime are not first-order optimal, therefore do not improve the wideband slope. The difference in the performance of the same transmission scheme in the high-SNR regime, which is measured by degree of freedom, and in the low-SNR regime, which is measured by the wideband slope, is expected, since given the definitions of these two measures in Equation 1.18 and Equation 1.15, we can see that the degree of freedom and the wideband slope have distinct physical meanings.

The challenges in interference management have very different natures for the K-user interference channel working in the high-SNR regime and that working in the low-SNR regime. In the high-SNR regime, the dominant factor affecting the performance is the interference from non-corresponding users. Respectively, the performance measure *degree of freedom* (DoF) is defined by letting the SNR tend to infinity, and the noise power can be neglected. Therefore, a transmission scheme achieving good performance in the high-SNR regime usually requires each user to make compromise between obtaining higher channel gain at its corresponding receiver and causing less interference at the non-corresponding receivers. Such transmission schemes may not achieve the optimal channel gain defined in Equation 1.14. On the other hand, in the low-SNR regime, the SNR tends to zero and the noise affects the performance tremendously. The primary goal in system design is to achieve the largest received signal to noise ratio, i.e., the transmission scheme must be first-order optimal, as that has been discussed in detail in Section section 2.6.

In high-SNR regime, the key interference management method is called *interfer-*

ence alignment. This concept was first introduced by [30] and [31] in the context of the multiple-input-multiple-output (MIMO) X channel and MIMO compound broadcast channel respectively. Cadambe and Jafar introduced interference alignment to K-user interference vector channel in [21], where the channel model is defined as

$$\mathbf{y}_j = \mathbf{H}_{jj}\mathbf{x}_j + \sum_{i \neq j} \mathbf{H}_{ji}\mathbf{x}_i + \mathbf{z}_j. \tag{4.5}$$

Common seen vector channels include MIMO channel and frequency selective channel. For a MIMO channel, the element on the pth row and qth column of the channel coefficient matrix $\mathbf{H}_{ji}$ represent the channel coefficient between the pth antenna of transmitter i and the $q$the antenna of receiver j; for an frequency selective channel, $\mathbf{H}_{ji}$ is a diagonal matrix where the pth element on the diagonal is the channel coefficient of the pth frequency slot between transmitter i and receiver j. Both MIMO channel and frequency selective channel are discussed in [21]. Following the works on vector channel, high-SNR regime interference channel with scalar channel coefficients has also been studied. Important works include [32, 33, 34].

Next, we use MIMO K-user interference channel as an example to demonstrate how the interference alignment works in the high-SNR regime. Assume that each transmitter and each receiver has n antennas. Let each user use zero-forcing at the receivers, then DoF achieved by each user is equal to the dimension of the linear subspace spanned by its received signal $\mathbf{H}_{jj}\mathbf{x}_j$ that is linearly independent from the subspace spanned by the received interference $\sum_{i \neq j} \mathbf{H}_{ji}\mathbf{x}_i$. Here $\mathbf{H}_{ji}$ is $n \times n$ channel matrix and $\mathbf{x}_i$ is length n transmitted vector. [21] has proposed a type of pre-coding algorithm, such that if the elements in the channel matrices are i.i.d., then the following conditions

$$rank\left(\mathbf{H}_{jj}\mathbf{x}_j\right) = {}^{n}\!/_{2} \tag{4.6}$$

$$rank\left(\left\langle \mathbf{H}_{jj}\mathbf{x}_j, \sum_{i \neq j} \mathbf{H}_{ji}\mathbf{x}_i \right\rangle\right) = n \tag{4.7}$$

can be satisfied for all i, $j = 1, \cdots, K$ with probability 1. Using zero-forcing at the receiver, this transmission scheme achieves

$$\Delta Dof = \frac{1}{2},$$

where

$$\Delta DoF = \frac{DoF \,\text{of the interference channel}}{DoF \,\text{of the interference} - \text{free channel}}.$$

Therefore, as SNR $\to \infty$, the performance of this K-user interference channel is only reduced by a factor of 2 at the presence of $K-1$ interference signals.

On the other hand, low-SNR regime interference channel faces a unique challenge: any interference management scheme must achieve the optimal minimum energy per bit defined in Theorem 7. Again, let us look at the MIMO channel defined by Equation 4.5. To obtain the correct $\frac{E_b}{N_0}\Big|_{\min}$, the input signal $\mathbf{x}_j$ must achieve the channel gain G_{jj},

$$G_{jj} \triangleq \sup_{F_{\mathbf{x}_j}} \frac{E\left[\|\mathbf{H}_{jj}\mathbf{x}_j\|^2\right]}{E\left[\|\mathbf{x}_j\|^2\right]}. \tag{4.8}$$

To achieve the optimal G_{jj}, each user need to transmit along the eigenvector corresponding to the maximal eigenvalue, even if this first-order optimal transmitting direction may generate relatively strong interference at non-corresponding receivers and heavily deteriorate the wideband slope.

Notice that in the high-SNR case we have just discussed, the design philosophy is quite different, where each user compromises between achieving optimal gain at the corresponding receiver and causing less interference at the non-corresponding receivers. With high probability $\mathbf{x}_j$ satisfying Equation 4.6 and Equation 4.7 does not satisfy Equation 4.8. Such a $\mathbf{x}_j$ is not first-order optimal as a result, and will not improve the wideband slope. This observation is true in general for existing achievable schemes focusing on improving the high-SNR performance of the interference channel with scalar or vector channel coefficients. Low-SNR interference channel is asking for new interference management technique.

4.3 Large Bandwidth Case: Interference Alignment Over Time Domain

In this section, we focus on the large bandwidth case, where the signal to noise ratio per second per Hz tends to zero by letting the bandwidth B go to infinity, while let the power P be a finite and constant real number. A type of interference alignment scheme over time domain will be proposed. Theoretical analysis will show that it is first order optimal and achieves $\Delta\mathcal{S} = \frac{1}{2}$. The main results in this section has been partially published in the author's work [35], in collaboration with A. Høst-Madsen.

The channel model is restated here for the reader's convenience. The received baseband discrete time signal at receiver j is

$$y_j[n] \quad = |h_{jj}| \quad x_j[n] + \sum_{i \neq j} h_{ji} \tilde{x}_i[n - n_{ji}] + z_j[n] \tag{4.9}$$

where

$$\tilde{x}_i[n] \quad = \quad \sum_{m=-\infty}^{\infty} x_i[m] \mathrm{sinc}(n - m + \delta_{ji}), \tag{4.10}$$

and n_{ji} and δ_{ji} are the integer part and the fractional part of the propagation delay respectively,

$$n_{ji} \quad = \quad \left\lfloor \tau_{ji} B + \tfrac{1}{2} \right\rfloor; \tag{4.11}$$

$$\delta_{ji} \quad = \quad \tau_{ji} B - \left\lfloor \tau_{ji} B + \tfrac{1}{2} \right\rfloor. \tag{4.12}$$

Any length-N input sequence x_i^N must satisfy power constraint $\frac{1}{N} \sum_{n=1}^{N} |x_i\,[n]|^2 \leq \mathrm{SNR}_i$, where $\mathrm{SNR}_i = \frac{P_i}{N_0 B}$.

The key idea in the design is to utilize the differences between propagation delays τ_{ji} for different $(j,\, i)$ pairs. We show that for any set of τ_{ji}, $i, j = 1, \cdots, K$ that are linearly independent over rational numbers, there exists arbitrarily large B that could make the propagation delay arbitrarily close to some odd integer. As a result, if we let each user use even time instance in the discrete time baseband channel model, then at the receiver, the desired message and the interference signal are almost orthogonal over time domain. Therefore, during half of the time, every user can enjoy almost interference-free transmission, and the interference channel can achieve $\Delta\mathcal{S}_0 = \frac{1}{2}$. Remember that transmitters and receivers are assumed to be placed in a 2 or 3 dimensional space randomly. Therefore with probability one τ_{ji} will be linear independent , and this transmission scheme will achieve good performance.

The idea of interference alignment over time domain is also used in [36] and [37]. Similar to the scheme proposed in this thesis, both works acknowledge that the interference and the message can not be completely orthogonal at the transmitter, and there always will be a part of the interference that leaks onto the signal space. The main difference lies in how this leaked interference is managed. [36] applies rectangular window transmission in time domain, which increase bandwidth requirement. [37] uses a channel model assuming the fractional delay δ is zero; and they also assume that the number of users is asymptotically large.

The transmission scheme proposed here will simply treat interference as noise. No approximation and assumption of window transmission is made. We will see that in the large bandwidth regime, interference alignment over time domain becomes a more natural choice.

Here, we refine the definition of minimum energy per bit and wideband slope in [1, (34), (29)] as

$$\begin{aligned} \frac{E_b}{N_{0\,\min}} &= \liminf_{\frac{P}{N_0 B}\to 0} \tfrac{P}{N_0 B}/R\left(\tfrac{P}{N_0 B}\right) && (4.13) \\ \mathcal{S}_0 &\triangleq \limsup_{\frac{E_b}{N_0}\downarrow \frac{E_b}{N_{0\,\min}}} \frac{R\left(\frac{E_b}{N_0}\right)}{10\log_{10}\frac{E_b}{N_0} - 10\log_{10}\frac{E_b}{N_{0\,\min}}} 10\log_{10} 2 && (4.14) \end{aligned}$$

The original definitions do not define the case where the limit in Equation 4.13 and Equation 4.14 do not exist for all infinite non increasing sequences $\left.\frac{P}{N_0 B}\right|_1, \left.\frac{P}{N_0 B}\right|_2 \cdots$ with $\lim_{i\to\infty} \left.\frac{P}{N_0 B}\right|_i = 0$. In the modified definitions $\left.\frac{E_b}{N_0}\right|_{\min}$ and $\mathcal{S}$ are well defined if there is lat least one sequence of $\left.\frac{P}{N_0 B}\right|_n$ such that the limits Equation 4.13 and Equation 4.14 exist. The necessity of this modification will be illustrated in subsection 4.3.3 by an example.

4.3.1 An Interference Alignment Scheme

In this section, we propose an interference alignment scheme over time domain which utilizes the propagation delay.

Definition 22. The transmission scheme is designed as:

1. At transmitter j

    - Let the discrete baseband signal be

$$x_j[n] \sim \begin{cases} \mathcal{N}\left(0,\ \frac{2P_j}{B_o}\right), & if\, n = 2k \\ 0, & if\, n = 2k+1 \end{cases} \quad \text{for all integer } k$$

        – Construct the continuous baseband signal by sinc pulse train

$$x_j(t) = \sum_{-\infty}^{+\infty} x_j[n] \operatorname{sinc}(B_o(t - nT))$$

a) At receiver j:

- Sample the received continuous time baseband signal with rate B_o and time synchronize with $x_j(t)$;
- Decode $x_j[2m]$ from $y_j[2m]$;
- The interferences are treated as noise.

B_o is the working bandwidth. The criterion used to choose B_o is

$$|\tau_{ji} B_o - 2k_{ji} - 1| \leq \delta \tag{4.15}$$

for all i, $j = 1, \cdots, K$. δ is some arbitrarily chosen design specification.

Next, we show that it is guaranteed to find an infinite sequence of $\{B_{o1}, B_{o2}, \cdots\}$, $B_{o(k+1)} > B_{ok}$, such that B_{ok}, $k = 1, 2, \cdots$ satisfies Equation 4.15. The concept of *linearly independent over rational number* will be introduced first.

Definition 23. [38]A set of real numbers $\theta = \{\theta_1, \theta_2, \cdots, \theta_n\}$ are linearly independent over rational number if $\sum a_i \theta_i = 0$ only if $a_i = 0$ for all $a_i \in \mathbb{Z}$.

Lemma 24. *If τ_{ji}, i, $j \in \{1, \cdots, K\}$ are linearly independent over the rational numbers, then for any $\delta > 0$, there exist arbitrarily large real number B , such that*

$$|\tau_{ji} B - 2k_{ji} - 1| \leq \delta \tag{4.16}$$

for some integers k_{ji}, j, $i \in \{1, \cdots, K\}$.

The proof of Lemma 24 is based on the following fundamental approximation results in number theory.

Theorem 25. *[38, Theorem 7.9, First Form of Kronecker's Theorem]If $\alpha_1, \cdots, \alpha_n$ are arbitrary real numbers, if $\theta_1, \cdots, \theta_n$ are linearly independent real numbers over the rational numbers, and if $\epsilon > 0$ is arbitrary, then there exists an real number t and integers $h_1, \cdots, h_n$ such that*

$$|t\theta_i - h_i - \alpha_i| < \epsilon \tag{4.17}$$

$\forall i \in \{1, 2, \ldots, n\}$

We also have

Lemma 26. *[38, Exercise 7.7, page 160]Under the hypotheses of Theorem 25, if $T > 0$ is given, there exists a real number $t > T$ satisfying the n inequalities Equation 4.17.*

Now let us prove Lemma 24.

Proof of Lemma 24. Let $\alpha_1, \cdots, \alpha_n = 0.5$, $\epsilon = \frac{\delta}{2}$. According to Theorem 25, there exist arbitrarily large real number $\hat{B}$ and some integers n_{ji}, $i, j \in \{1, \cdots, K\}$ such that

$$\left|\tau_{ji}\hat{B} - k_{ji} - 0.5\right| \leq \frac{\delta}{2}.$$

Let $B = 2\hat{B}$, we have

$$|\tau_{ji}B - 2k_{ji} - 1| \leq \delta.$$

Combining the inequality above with Lemma 26, Lemma 24 is proved. $\square$

4.3.2 The Achievable Sum Slope

In this section, we show that the interference alignment scheme proposed in Definition 22 achieves $\Delta\mathcal{S}_0 = \frac{1}{2}$.

Lemma 24 shows that using this transmission scheme, the desired signal is almost orthogonal with the interference signal in time domain. However, this is not sufficient. Since the energy of interference signal that leaks onto the signal space will be treated as noise, we also need to show that as $\delta_{ji} \to 0$, the energy of interference signal falls onto signal space will also converge to zero. This is stated in the following lemma.

Lemma 27. *Under the assumptions $x_j[2m]$ are i.i.d. Gaussian random variable with distribution $\mathcal{N}(0, 2P_j)$ and $x_j[2m+1] = 0$ for all j and m, $\mathrm{E}[\tilde{x}_i[n_1, \delta_{ji}]\tilde{x}_i^*[n_2, \delta_{ji}]]$ is a continuous function of δ_{ji} which satisfies*

$$\lim_{\delta_{ji}\downarrow 0} \mathrm{E}[\tilde{x}_i^*[n_1, \delta_{ji}]\tilde{x}_i[n_2, \delta_{ji}]] = \begin{cases} 2P_i & \text{if } n_1 = n_2 = 2k, \\ & \text{for some integer } k \\ 0 & o.w. \end{cases}$$

Proof. Please see the Appendix section A.2. □

Equipped with the interference alignment scheme in definition 22, Lemma 24, and Lemma 27, we proceed to show our the main results on the achievable sum slope of K-user interference channel.

Theorem 28. *Suppose that the set of delays τ_{ji}, i, $j \in \{1, \cdots, K\}$ are linearly independent over the rational numbers. Then the following wideband slope is achievable*

$$\begin{aligned} \mathcal{S}_0 &= \frac{\left(\sum_j |h_{jj}|^2\right)^2}{\sum_j |h_{jj}|^4} && \textit{equal power constraint;} \\ \mathcal{S}_0 &= K && \textit{equal rate constraint.} \end{aligned}$$

Under both constraints,

$$\Delta\mathcal{S}_0 = \frac{1}{2}$$

is achievable.

Proof. Assume that the system uses the transmission scheme proposed in Definition 22. Let

$$\epsilon_j(B) = \sum_{i \neq j}^{K} \left| \mathrm{E}\left[(\tilde{x}_i[2n])^2 \right] \right|$$

denote the power of the leaked interference.

The best rate with this scheme is clearly achieved if the leaked interference power is zero; in that case the channel is an interference-free channel where half the symbols are not used. We can therefore conclude

$$\Delta\mathcal{S}_0 \leq \frac{1}{2} \tag{4.18}$$

On the other hand, taking into account the leaked interference, the achievable rate at receiver j is

$$
\begin{aligned}
& R_j & (4.19)\\
= \;& \frac{1}{2}\log\left(1+\frac{|h_{jj}|^2\frac{2P_j}{BN_0}}{1+\frac{\epsilon_j(B)}{BN_0}}\right) & (4.20)\\
= \;& |h_{jj}|^2\frac{P_j}{BN_0}-\left(\epsilon_j(B)\,|h_{jj}|^2 P_j+|h_{jj}|^4 P_j^2\right)\left(\frac{1}{BN_0}\right)^2 & (4.21)\\
& +o\left(\left(\frac{1}{BN_0}\right)^2\right). & (4.22)
\end{aligned}
$$

The wideband slope is a continuous function of the coefficients in the first two terms in the Taylor series of R_j in $\frac{1}{B}$. According to Lemma 24 for any $\delta > 0$ there exists some B_δ and a set of integers k_{ji} such that $n_{ji} = 2k_{ji} + 1$, i.e., the integer part of the delay is an odd number, and the fractional part of the delay satisfies $|\delta_{ji}| \leq \delta$.

From Lemma 24 and Lemma 27, we can then conclude that there exists a sequence of real numbers $\{B_{o1}, B_{o2}, \cdots\}$, $B_{o(k+1)} > B_{ok}$, so that $k \to \infty$ and $\epsilon_j(B_{ok}) \to 0$ for all $j = 1, \cdots, K$. This means that $\Delta\mathcal{S}_0 = \frac{1}{2}$ is a limit point, and together with Equation 4.18 this shows that $\Delta\mathcal{S}_0 = \frac{1}{2}$ is the limit superior. □

Now we can easily conclude the following corollary.

Corollary 29. *Suppose that all nodes have independent positions and each node position has a continuous distribution. Then the propagation delays τ_{ji}, $i, j \in \{1, \cdots, K\}$, are linearly independent over rational numbers with high probability, and*

$$\Delta\mathcal{S}_0 = \frac{1}{2}$$

is achievable.

Corollary 29 implies that in practice $\Delta\mathcal{S}_0 = \frac{1}{2}$ is achievable with probability one, since in our channel model, transmitters and receivers are placed randomly, and the probability that they are positioned accurately in a grid is zero.

4.3.3 Practical Implementation and Simulation Results

In this section we will show that the interference alignment ideas of the previous section can be used in a practical system, and show some simulation results. This will also make it clear why the modified definition Equation 4.14 is needed.

We can see that one key question concerning the transmission scheme defined by Definition 22 is: how to find B_o? Here we propose an algorithm, stated in the following proposition.

Proposition 30. *Assume that both the transmitters and the receivers have perfect channel knowledge. With probability one, the following searching method terminates after a finite number of iterations:*
Initialize B to be any positive integer. Proceed to the following while loop:
While $(|\tau_{ji}B - 2k_{ji} - 1| > \delta)$:
Increase B by 1, *i.e.*, $B = B + 1$.

The output B of this while loop satisfies $|\tau_{ji}B - 2k_{ji} - 1| \leq \delta$, therefore can be chosen as B_o.

Lemma 31 guarantees that the searching algorithm defined in the proposition above terminates. The proof is almost identical to the proof of Lemma 24. However, the essential difference is that while Lemma 24 only shows the existence of B satisfying Equation 4.16 over the set of positive real numbers $\mathbb{R}^+$, the results in this section ensure that such B can be found even if we restrict B to be integer.

Lemma 31. *If τ_{ji} are linearly independent over the rational numbers, then for any $\delta > 0$, there exist an integer B , such that*

$$|\tau_{ji}B - 2k_{ji} - 1| \;\leq\; \delta \tag{4.23}$$

for some integers k_{ji}. Further, B can be made arbitrarily large.

The proof of Lemma 31 is based on the second form of Kronecker's theorem[38, Theorem 7.10, Second Form of Kronecker's Theorem], which shows that Theorem 25 still holds even if we require t to be an integer. Details are skipped here.

We can see that the brute force algorithm of searching through all integer B is guaranteed to find good operating bandwidths. Figure 4.2 shows the performance of the proposed achievable scheme when the system operating at a sequence of B_δ, $\delta = 0.2$. However, designing more efficient B_o-searching algorithm could be a subject of further research.

In the simulation, we consider a 3-user channel with symmetric channel gain: $|h_{jj}|^2 = 1$, $|h_{ji}|^2 = 0.8$. Notice that for channels with symmetric link gains, equal power and equal rate constraints are equivalent. The delays τ_{ji} are chosen such that they are linearly independent over the rational numbers.

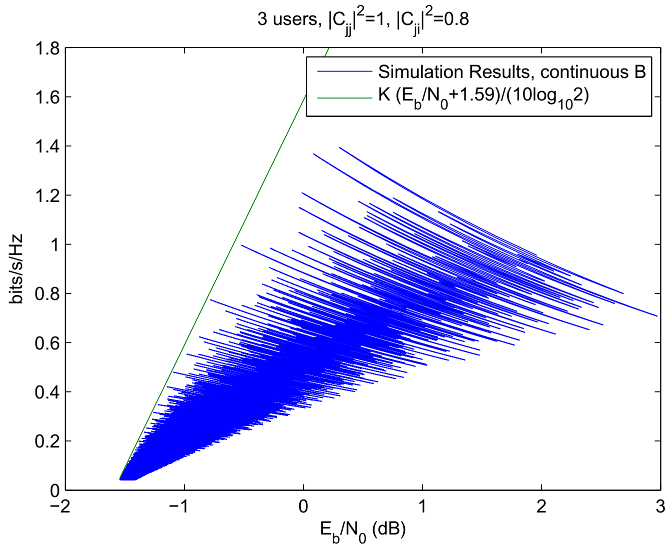

Figure 4.1: Performance for arbitrary bandwidth.

Figure 4.1 shows the simulation results of the case where bandwidth B increases continuously. The system performance shows a noticeable oscillating behavior. This phenomenon can be explained as followed. At receiver j the interference caused by user i is an increasing function of δ_{ji}; and δ_{ji} is a periodic function of B, oscillating between 0 and 1. It can be proved that the cumulative effect of leaked interferences from all other users has same (almost) periodic behavior. The proof is similar to that of Lemmas Equation 5.14 and 26; details are skipped here.

Figure 4.1 also shows why we need the modified definition Equation 4.14. We could see that the limits defined by Equation 2.12 and Equation 2.13 do not exists: if we take one infinite sequence of bandwidth values from the upper envelope of the $R\left(\frac{E_b}{N_0}\right)$ curve, and take another infinite sequence of bandwidth values from its lower envelope, they will give different wideband slope. However, the $\limsup$ always exists. More importantly, this modified definition is *operational*: given the channel state information, the system can find a large but finite B which achieves good low-SNR regime performance.

4.4 Small Bandwidth Case: Circularly Asymmetric Signaling

In this section, a type of interference alignment scheme will be proposed for the small bandwidth case. This achievable scheme applies the concept of *circularly asymmetric signaling,* which has been introduced in section 2.4. To fully describe the circularly asymmetric input, instead of the scalar complex channel model, a

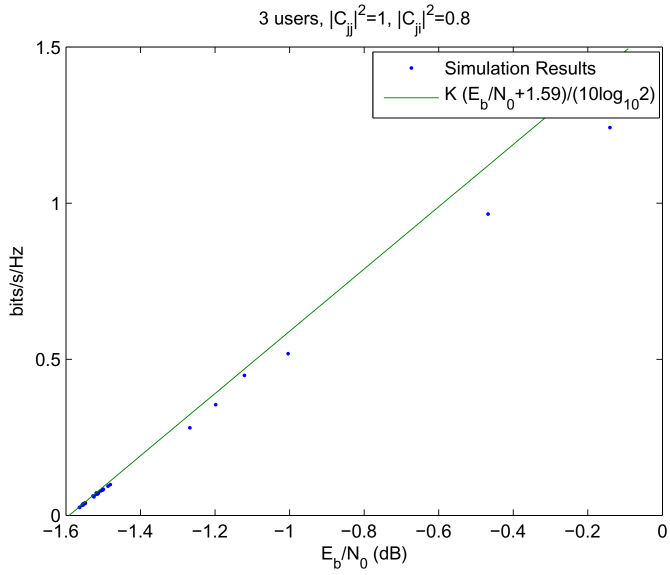

Figure 4.2: Peak points of performance.

2-dimensional real channel model equivalent to the scalar complex channel model may be more convenience.

The channel model is restated here for the reader's convenience. The received baseband discrete time signal at receiver j is

$$y_j[n] \quad = \quad h_{jj}x_i[n] + \sum_{i \neq j} h_{ji}x_i[n] + z_j[n]. \tag{4.24}$$

The channel coefficient h_{ji} is complex scalar number, and the noise z_j is i.i.d. circularly symmetric complex random variable. For any length-N input sequence x_i^N, power constraint $\frac{1}{N}\sum_{n=1}^{N}|x_i[n]|^2 \leq \mathrm{SNR}_i$ is satisfied, where $\mathrm{SNR}_i = \frac{P_i}{N_0 B}$. Its equivalent two-dimensional real channel is

$$\mathbf{y}_j \quad = \quad |h_{jj}|\,\mathbf{U}_{jj}\mathbf{x}_j + \sum_{i=1, i\neq j}^{K} |h_{ji}|\,\mathbf{U}_{ji}\mathbf{x}_i + \mathbf{z}_j \tag{4.25}$$

where $\mathbf{U}_{ji} \triangleq \begin{pmatrix} \cos(\phi_{ji}) & -\sin(\phi_{ji}) \\ \sin(\phi_{ji}) & \cos(\phi_{ji}) \end{pmatrix}$ is the rotation matrix with angle ϕ_{ji}. And the 2×1 vector white Gaussian noise is $\mathbf{z}_j \sim \mathcal{N}\left(0, \frac{1}{2}\mathbf{I}_{2\times 2}\right)$; the input signal $\mathbf{x}_j$ is related to the scalar complex model by: $\mathbf{x}_j = \begin{pmatrix} \mathrm{Re}\{x_j\} \\ \mathrm{Im}\{x_j\} \end{pmatrix}$, with the average covariance matrix of the jth input as $\mathbf{K}_{\mathbf{x}_j} \triangleq \frac{1}{N}\sum_{n=1}^{N} \mathrm{E}\left[\mathbf{x}_j[n](\mathbf{x}_j[n])^{\mathrm{H}}\right]$ satisfying power constraint $\mathrm{Tr}\left(\mathbf{K}_{\mathbf{x}_j}\right) \leq \mathrm{SNR}_j$. Denote covariant matrix normalized by SNR as $\hat{\mathbf{K}}_j = \frac{\mathbf{K}_j}{\mathrm{SNR}_j}$.

Many previous works on the Shannon capacity region of a interference channel only consider circularly symmetric inputs. However, [21] shows that K-user interference channel using circularly asymmetric signaling achieves better performance in high-SNR regime. Their result is one of our main inspirations, although the specific interference alignment technique they proposed is not applicable to the low-SNR regime.

In the example follows, we can see that for certain channel realizations, circularly asymmetric indeed can improve the low-SNR regime performance. Take a channel whose direct link coefficients have zero phase, i.e., $\phi_{jj} = 0$, and interference link coefficients have phase $\pi/2$ as an example. If we let every transmitter only transmits over the real axis, then at every receiver, all interferences will be aligned to the imaginary axis, orthogonal to the desired *signal*. This transmission scheme achieves $\Delta\mathcal{S}_0 = 1/2$ regardless of the link gains $|h_{ji}|^2$. Of course, since we assume ϕ_{ji} is drawn from a uniform distribution over $(-\pi, \pi)$ such channel will occur with zero probability. Nevertheless, the idea of interference alignment over phase domain is inspiring. Here we are going to generalize this idea into more general cases. The main results in this section has been partially published in the author's works [39, 40], in collaboration with A. Høst-Madsen.

4.4.1 One-Dimensional Gaussian Signaling

In this section, an phase interference alignment scheme will be proposed. In this scheme, transmitter j only transmits in the direction $e^{\iota\theta_j}$, instead of using the whole complex plane. Simulation results will be shown in the next section, we will see that when the interference link gains are moderate, with high probability, this interference alignment scheme performs better than existing transmission schemes: treating interference as noise and TDMA.

Consider a simple transmission scheme with the following properties:

- the input sequence $x_j[n]$ is i.i.d. distributed Gaussian complex random variable for all n.
- $E[\mathrm{Re}\{x_j\}] = E[\mathrm{Im}\{x_j\}] = 0$, $E\left[x_j x_j^*\right] = \mathrm{SNR}_j$ and the correlation between $\mathrm{Re}\{x_j\}$ and $\mathrm{Im}\{x_j\}$ is 1.
- Receivers treat interference as noise.

The following definition is equivalent to the description above, but provide a more clear geometric interpretation.

Definition 32. One-dimensional Gaussian signaling transmission scheme

- At transmitter j, let input sequence be $x_j[n] = w_j[n]\, e^{j\theta_j}$, where $w_j[n]$ is drawn from i.i.d real Gaussian random variable with distribution $\mathcal{N}(0, \mathrm{SNR}_j)$, and the phase θ_j is a priory chosen design parameter, unchanged for all n during the transmission.
- At receiver j, interference is treated as noise.

We call this one-dimensional because every transmitter only transmit along $e^{j\theta_j}$, therefore only one dimension is used out of the two-dimensional signal space.

Our objective is to find the set of phases $\underline{\theta} = \{\theta_1, \cdots, \theta_K\}$ that maximize the achievable wideband slope $\mathcal{S}_0$.

The achievable $\mathcal{S}_0$ for any $\underline{\theta}$ is stated in the next lemma. For computational convenience, we return to the equivalent two-dimensional real channel model. In the equivalent 2-dimensional real channel model, the input $\mathbf{x}_j$ has covariant matrix

$$\mathbf{K}_j = \mathrm{SNR}_j \begin{pmatrix} \cos^2\theta & \frac{\sin 2\theta}{2} \\ \frac{\sin 2\theta}{2} & \sin^2\theta \end{pmatrix},$$

$rank\,(\mathbf{K}_j) = 1$.

Lemma 33. *For the 2-dimensional real channel defined by Equation 4.25, the sum slope achieved by the one-dimensional Gaussian signaling is*

$$\mathcal{S}_0 = \frac{\left(\sum_{j=1}^{K} |h_{jj}|^2\right)^2}{\sum_{j=1}^{K} |h_{jj}|^4 + \sum_{j=1}^{K}\sum_{i\neq j}^{K} |h_{jj}|^2 |h_{ji}|^2 + f(\underline{\theta})}, \tag{4.26}$$

where

$$f(\underline{\theta}) \triangleq \sum_{j=1}^{K}\sum_{i\neq j}^{K} |h_{jj}|^2 |h_{ji}|^2 \cos 2(\phi_{ji} - \theta_j + \theta_i). \tag{4.27}$$

Proof. Using treating interference as noise at the receiver, the achievable sum rate

32 is

$$R_s = \textstyle\sum_{j=1}^{K} \left(\frac{1}{2} \log \left| \mathbf{I}_2 + \frac{2}{K} \mathrm{SNR}_s \left(\left| h_{jj} \right|^2 \mathbf{U}(\phi_{jj}) \hat{\mathbf{K}}_j \mathbf{U}_2(-\phi_{jj}) + \sum_{i=1, i \neq j}^{K} \left| h_{ji} \right|^2 \mathbf{U}(\phi_{ji}) \hat{\mathbf{K}}_i \mathbf{U}(-\phi_{ji}) \right) \right| - \frac{1}{2} \cdot \right. \tag{4.28}$$

$$\left. \log \left| \mathbf{I}_2 + \frac{2}{K} \mathrm{SNR}_s \sum_{i=1, i \neq j}^{K} \left| h_{ji} \right|^2 \mathbf{U}(\phi_{ji}) \hat{\mathbf{K}}_i \mathbf{U}(-\phi_{ji}) \right| \right) \tag{4.29}$$

under equal power constraint where $\mathrm{SNR}_j = \frac{\mathrm{SNR}_s}{K}$. Combining Equation 2.14, Equation 1.16 and Equation 4.28, we have

$$\dot{R}_s(0) = \frac{\sum_{j=1}^{K} \left| h_{jj} \right|^2}{K} \tag{4.30}$$

$$-\ddot{R}_s(0) = \frac{2 \sum_{j=1}^{K} \left| h_{jj} \right|^4}{K^2} + \frac{2 \sum_{j=1}^{K} \sum_{i \neq j}^{K} \left| h_{jj} \right|^2 \left| h_{ji} \right|^2}{K^2} + \frac{2}{K^2} \sum_{j=1}^{K} \sum_{i \neq j}^{K} \left(\left| h_{jj} \right|^2 \left| h_{ji} \right|^2 \cdot \cos 2\left(\phi_{ji} - \theta_j + \theta_i \right) \right). \tag{4.31}$$

Given $\mathcal{S}_0 = \frac{2\dot{R}_s^2(0)}{-\ddot{R}_s(0)}$, Equation 4.26 followed. □

Given Equation 4.26, to maximizes $\mathcal{S}_0$ is equivalent to find the set of θ_j that minimizes $f(\underline{\theta})$.

Denote this optimization problem by $P(\underline{\theta})$, which is defined as

$$\begin{array}{ll} \min & f(\underline{\theta}) \\ \text{subject to} & \theta_j \in [-\pi, \pi]. \end{array}$$

Notice that $\theta_j \bmod 2\pi$ will not affect the value of $f(\underline{\theta})$. Therefore, $P(\underline{\theta})$ can be solved using standard numerical methods for unconstrained optimization problems.

4.4.2 Simulation Results and Discussions

In this section, we simulate the performance of the one-dimensional signaling scheme in 10-user interference channel with unit direct link gains and symmetric weak interference link gains, i.e., $\left| h_{jj} \right|^2 = 1$ and $\left| h_{ji} \right|^2 = a < 1$ for all

$i, j = 1, \cdots, 10$. the phases ϕ_{ji} is drawn from $U[-\pi, \pi]$ in each channel realization. This performance will be compared with existing achievable schemes: treating interference as noise and TDMA.

Notice that to simplify the analysis, we only consider weak interference channel. However, when K is large the event that all interference links are weak may only happens with very low probability, especially when a is close to 1. For a mixed channel, a better choice for achievable schemes may be a combination between interference decoding and the transmission schemes that suit weak interference case.

The simulation results are presented below. We can see that when a, the ratio between the direct link gain and the interference link gain, is close to 1, then with non-zero probability the one dimensional Gaussian signaling transmission scheme performs better than TDMA.

Figure 4.3 illustrates the empirical cumulative distribution functions of the sum slope achieved by the one-dimensional interference alignment scheme at different a values. For comparison, $\mathcal{S}_0$ achieved by treating interference as noise are also shown, and TDMA always achieves$\mathcal{S}_0 = 2$ for all a value.

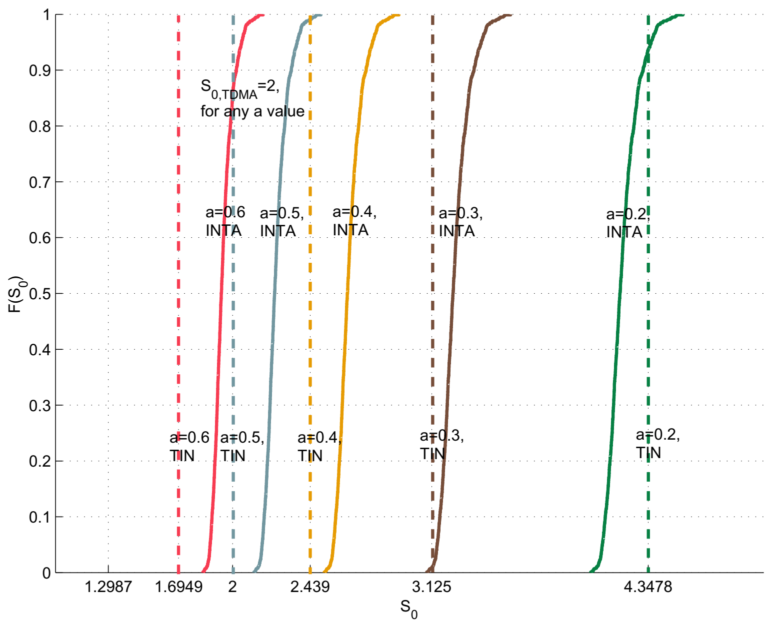

Figure 4.3: Empirical cumulative distribution functions of $\mathcal{S}_0$ achieved by treating interference as noise (TIN), interference alignment (INTA) and TDMA under different a values.

In Figure 4.4, we compare the median value of $\mathcal{S}_0$ achieved by one-dimensional interference alignment scheme with the performance of treating interference as noise and TDMA.

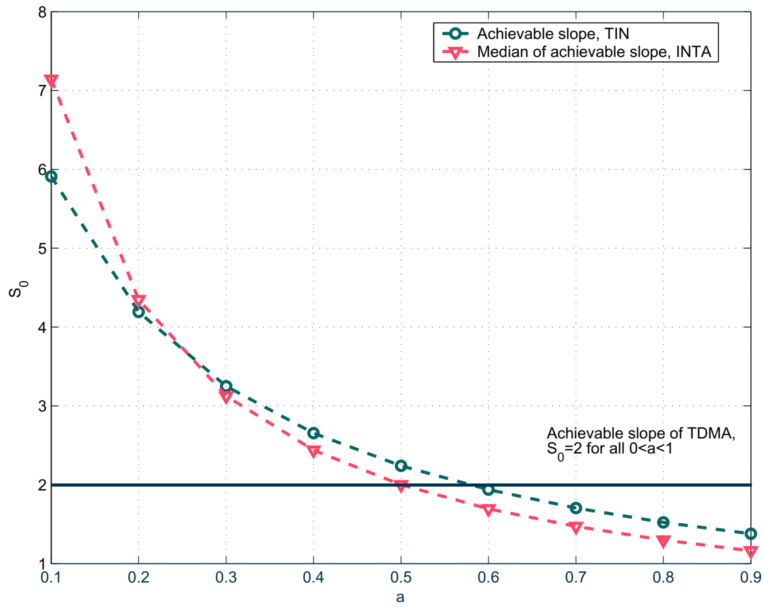

Figure 4.4: the median value of $\mathcal{S}_0$ achieved by INTA, and the achievable $\mathcal{S}_0$ of TIN and TDMA as a function of a

For the canonical symmetric channel considered, we could see that the performance of one-dimensional interference alignment is more likely to be better than TIN and TDMA when the interference link gain a is moderate. If a is small, treating interference as noise achieves better performance; when a grows closer to 1, TDMA becomes more likely to have better performance.

Chapter 5

K-User Interference Channel: Wideband Slope Outer Bounds

5.1 Summary

Few works discusses the low-SNR regime performance of K-user interference channel, $K > 2$. Noticeable works on the performance outer bounds of K-user interference channel include [18], [28], and [41]. [18] considers channel working in the high-SNR regime; [41] discusses partially connected K-user channel. The latest work [28] considers fully connected K-user interference channel with general SNR value, and develops capacity outer bound for certain channel realizations. However, if we assume that the channel coefficients are continuous random variable, then these channel realizations occur with zero probability. In this chapter, first-order optimal outer bounds on the wideband slope of the K-user interference channel will be developed, for both the large bandwidth case and the small bandwidth case. These bounds work for a set of channel realizations with non-zero probability. Together with the optimal and achievable minimum energy per bit defined in Theorem 7, the wideband slope outer bounds provide outer bounds on the spectral efficiency of the K-user interference channel working in the low-SNR regime.

First, the large bandwidth case will be discussed. In chapter 4 section 4.3, we have seen that under the large bandwidth case, benefiting from the propagation delay, which can grow arbitrarily large as $B \to \infty$, the K-user interference channel achieves $\Delta\mathcal{S}_0 = 1/2$ with probability one. Of course, if the channel condition is amiable, better achievable transmission scheme exists. For instance, if all of the interference link gains are higher than the direct link gains, then interference decoding achieves higher wideband slope; and if all of the interference link gains are very weak, then treating interference as noise may achieve better performance. However, our results in section 5.2 implies that under certain conditions, channel realizations with very strong or very weak interference link gains rarely occur.

section 5.2 will show that if channel coefficients are drawn from continuously distributed random variable, and the number of users K is large, then with high probability, the K-user interference channel can be discomposed into $K/2$ compounded 2-user interference channels, such that

1. each 2-user component forms a 2-user weak interference channel;
2. In each 2-user component, the interference link gains are close to the direct link gains,

i.e. each compounded 2-user interference channel satisfies $(1-\epsilon) \leq |h_{ji}|^2/|h_{ii}|^2 < 1$, where ϵ is an arbitrary positive real number.

Applying the outer bound for the 2-user interference channel developed in chapter 3, section 3.3 to the compounded 2-user interference channels, we obtain a wideband slope outer bound for the K-user interference channel. It shows that

1. Under equal rate constraint, the achievable sum slope of a channel realization must satisfy $\Delta\mathcal{S}_0 < 1/2 + \delta$ for all $\delta > 0$, as $K \to \infty$.
2. Under equal power constraint, the achievable sum slope of a channel realization must satisfy $\Delta\mathcal{S}_0 < 1/2 + \hat{\delta} + \delta$ for all $\delta > 0$, as $K \to \infty$, where $\hat{\delta}$ is a positive real number depending on the distribution function of the channel coefficients.

This outer bound holds for both the large bandwidth case and the small bandwidth case.

Comparing with the achievable wideband slope region discussed in the previous chapter, we can see that for the large bandwidth case, the interference alignment transmission scheme proposed in chapter 4 section 4.3, which achieves $\Delta\mathcal{S}_0 = 1/2$, is almost optimal.

For the small bandwidth case, the gap between this slope outer bound and the inner bound achieved by the phase alignment transmission scheme proposed in chapter 4 section 5.3 becomes larger. Of course, there exists certain channel realization for which the phase alignment scheme can achieve $\Delta\mathcal{S}_0 = 1/2$. For instance, if a channel realization has $\phi_{jj} = 0$ and $\phi_{ji} = \pi/2$ for all $i,\, j = 1, \cdots, K$, then $\Delta\mathcal{S}_0 = 1/2$ is achievable if we let $\theta_j = 0$. However, from the simulation results shown in section 5.3, we can see that with non-zero probability, TDMA has better performance as the interference link gains grows closer to the direct link gains, which can only achieve $\Delta\mathcal{S}_0 = 1/K$.

In section 5.3, a tighter outer bound will be derived, which implies that for a set of small bandwidth K-user interference channel realizations, the achievable sum slope of a channel realization must satisfy $\Delta\mathcal{S}_0 < 1/K + \delta$ for all $\delta > 0$, and the probability that a channel realization falls into this set is strictly greater than zero. This result implies that for this set of channels, TDMA is almost optimal.

5.2 Large Bandwidth Case

Chapter 4 section 4.3 shows that using interference alignment and treating interference as noise, $\Delta\mathcal{S}_0 = \frac{1}{2}$ is achievable. Is it possible to obtain better performance? For certain channels, the answer is yes. For instance, if all interference link gains are greater than the direct link gains, i.e., $\frac{|h_{ji}|^2}{|h_{ii}|^2} > 1$, then interference decoding can completely eliminate the interference, and the system achieves $\Delta\mathcal{S}_0 = 1$. Another example is a channel whose interference link gains are very small, for this channel treating interference as noise may achieves better performance.

However, if $|h_{ji}|^2$ is the realization of a continuous random variable and the number of users is large, then the examples we just mentioned only happen with very low probability. Our questions are

- How does a typical channel look like?
- How can its wideband slope be bounded?

By typical channel, we mean that it happens with high probability.

Let us first define $(1-\epsilon)$-interference pair and weak $(1-\epsilon)$-interference pair.

Definition 34. We say that users i and j form an $(1-\epsilon)$-*interference pair* if

$$1-\epsilon \leq \frac{|h_{ji}|^2}{|h_{ii}|^2}, \frac{|h_{ij}|^2}{|h_{jj}|^2} < 1,$$

and form a *weak* $(1-\epsilon)$-*interference* pair if

$$1-\epsilon \leq \frac{|h_{ji}|^2}{|h_{ii}|^2} < 1 \text{ or } 1-\epsilon \leq \frac{|h_{ij}|^2}{|h_{jj}|^2} < 1.$$

The outer bounds in this section are proven under the assumption that the channel coefficients h_{ji} for all $i,\ j \in \{1, \cdots, K\}$ are i.i.d. random variables. However, this is not a necessary condition, only a convenient condition to simplify proofs; later in the section we will comment more on this. The main results in this section have been partially published in the author's work [35], in collaboration with A. Høst-Madsen.

5.2.1 The Equal Rate Constraint

The equal rate constraint will be discussed in this section. Our main results will show that as the number of users $K \to \infty$, the event

$$\{\text{user } j, \forall j \in \{1, \cdots, K\}, \text{ forms a } (1-\epsilon) - \text{interference pair} \\ \text{with at least one other user}\}$$

happens with high probability, which gives $\Delta\mathcal{S}_0 \leq \frac{1}{2} + \delta$, $\forall \delta > 0$ under the equal rate constraint.

First consider the equal rate constraint. We assume that the channel coefficients h_{ji} are i.i.d. and $E\left[|h_{ii}|^{-2}\right] < \infty$; if the later assumption were not satisfied, $\lim_{K\to\infty} \frac{1}{K}\sum_{i=1}^{K} P_i = \infty$ even if the channel is interference-free(see Equation 2.22), so the energy per bit and wideband slope would not be well-defined for large K (see also the comment at the top of page 1325 in [1] about Rayleigh fading).

For $\forall\epsilon$, $\hat{\epsilon} > 0$, define two sets

$$\begin{aligned} R_{\hat{\epsilon}} &= \{x \in \mathbb{R} : F_{|h_{ii}|^2}(x) - F_{|h_{ii}|^2}((1-\epsilon)x) < \hat{\epsilon}\}; \\ D_{\hat{\epsilon}} &= \left\{all\ i \in \{1, \dots K\} : |h_{ii}|^2 \in R_{\hat{\epsilon}}\right\}. \end{aligned}$$

The following lemma shows that as the number of users $K \to \infty$, conditioned on $i \in D_{\hat{\epsilon}}^c$, with high probability user i forms a $(1-\epsilon)$-interference pair with at least one other user.

Lemma 35. *Given* $\forall\epsilon$, $\hat{\epsilon} > 0$, *denote*

$$\begin{aligned} B_{\epsilon,\hat{\epsilon}} \triangleq &\left\{all\, i \in \{1, \dots K\} \bigcap i \in D_{\hat{\epsilon}}^c : \right. \\ &\left. \text{user } i \text{ does not form an } (1-\epsilon) - \text{inteference pair with any other user}\right\}. \end{aligned}$$

Then

$$\lim_{K\to\infty} \Pr\left(B_{\epsilon,\hat{\epsilon}} = \emptyset\right) = 1.$$

Proof. Please see the Appendix section A.3. □

On the other hand, conditioned on $i \in D_{\hat{\epsilon}}$, the chance that $|h_{ji}|^2$ falls on the interval $\left((1-\epsilon)|h_{ii}|^2, |h_{ii}|^2\right)$ is low for $\hat{\epsilon}$ small, and probability that user i can form a $(1-\epsilon)$-interference pair is low as a result. However, in the following lemma, we show that the sequence $\Pr(R_{\hat{\epsilon}})$ can be made arbitrarily small.

Lemma 36. *Given any infinite sequence $\hat{\epsilon}_n > 0$ satisfying $\hat{\epsilon}_n > \hat{\epsilon}_{n+1}$ and $\hat{\epsilon}_n \to 0$, the corresponding sequence of $R_{\hat{\epsilon}_n}$ satisfies*

1) $R_{\hat{\epsilon}_{n+1}} \subseteq R_{\hat{\epsilon}_n}$*;*

2) $\Pr(R_{\hat{\epsilon}_n}) \to 0$.

Proof. Please see the Appendix section A.4. □

Regarding to the sum slope, if all users are in $D^c_{\hat{\epsilon}}$, then it is easy to show that there exists a sum slope outer bound which can be arbitrarily close to the inner bound given in Theorem 28. On the other hand, for users in $D_{\hat{\epsilon}}$ better performance may be achieved, and the over all sum slope of the system may be improved as a result. However, given the following lemma, we will see that the impact on system performance of users in $D_{\hat{\epsilon}}$ can be made arbitrarily small.

Lemma 37. *Let X be a positive random variable with $E[X] < \infty$, and let μ_X be the measure induced by the cumulative distribution function of X. Let $G_i \subset \mathbb{R}$ be a sequence of measurable sets with $G_{i+1} \subseteq G_i$ and $\lim_{i\to\infty} \mu_X(G_i) = 0$. Define*

$$X_i = \begin{cases} X & X \in G_i \\ 0 & X \notin G_i \end{cases}.$$

Then

$$\lim_{i\to\infty} E[X_i] = 0.$$

Proof. Please see the Appendix section A.5. □

Our main result is stated in the following theorem.

Theorem 38. *Suppose that the channel coefficient h_{ij} are i.i.d.. Under the equal rate constraint*

$$\forall\delta > 0 : \lim_{K\to\infty} \Pr\left(\Delta\mathcal{S}_0 \le \frac{1}{2} + \delta\right) = 1. \tag{5.1}$$

Proof. We discuss users in the set $D_{\hat{\epsilon}}$ and those in the set $D_{\hat{\epsilon}}^c$ separately.

First, let us look at user j, $j \in D_{\hat{\epsilon}}^c$. We assume that each user $j \in D_{\hat{\epsilon}}^c$ forms a $(1-\epsilon)$-interference pair with some user $i_{(j)}$. Given Lemma 35, this happens with high probability. Consider a single $(1-\epsilon)$-interference pair $(j, i_{(j)})$. We can get an upper bound on the spectral efficiency, by eliminating all interference links except the links between users j and $i_{(j)}$, so that the received signal is

$$\begin{aligned} y_j &= h_{jj}x_j + h_{ji_{(j)}}x_{i_{(j)}} + z_j \\ y_{i_{(j)}} &= h_{i_{(j)}j}x_j + h_{i_{(j)}i_{(j)}}x_{i_{(j)}} + z_{i_{(j)}}. \end{aligned}$$

Let $\frac{\left|h_{ji_{(j)}}\right|^2}{\left|h_{i_{(j)}i_{(j)}}\right|^2} = 1-\epsilon_{ji_{(j)}}, \frac{\left|h_{i_{(j)}j}\right|^2}{|h_{jj}|^2} = 1-\epsilon_{i_{(j)}j}$. Since $\left\{y_j, y_{i_{(j)}}\right\}$ is a $(1-\epsilon)$-interference pair, we have

$$0 \leq \epsilon_{ji_{(j)}}, \epsilon_{i_{(j)}j} < \epsilon. \tag{5.2}$$

Applying Equation 3.38 and Equation 3.40 to $\left\{y_j, y_{i_{(j)}}\right\}$, we have the optimum solution

$$\begin{aligned} \mathrm{SNR}_{i_{(j)}o} &= \left|h_{i_{(j)}i_{(j)}}\right|^{-2} \cdot \frac{2^{\frac{R_s}{K}}\left(2^{\frac{R_s}{K}}-1\right)\left(1-\epsilon_{i_{(j)}j}\right)\epsilon_{ji_{(j)}} + \epsilon_{i_{(j)}j}\left(2^{\frac{R_s}{K}}-1\right)}{1-\left(1-\epsilon_{ji_{(j)}}\right)\left(1-\epsilon_{i_{(j)}j}\right)} &(5.3)\\ \mathrm{SNR}_{jo} &= |h_{jj}|^{-2} \cdot \frac{2^{\frac{R_s}{K}}\left(2^{\frac{R_s}{K}}-1\right)\left(1-\epsilon_{ji_{(j)}}\right)\epsilon_{i_{(j)}j} + \epsilon_{ji_{(j)}}\left(2^{\frac{R_s}{K}}-1\right)}{1-\left(1-\epsilon_{ji_{(j)}}\right)\left(1-\epsilon_{i_{(j)}j}\right)}. &(5.4) \end{aligned}$$

And $\mathrm{SNR}_{i_{(j)}} + \mathrm{SNR}_j \geq \mathrm{SNR}_{i_{(j)}o} + \mathrm{SNR}_{jo}$. Notice that the right hand side of Equation 5.5 and Equation 5.6 are monotonically decreasing function of either $\epsilon_{ji_{(j)}}$ or $\epsilon_{i_{(j)}j}$. Thus, given the condition Equation 5.2, we can relax Equation 5.5 and Equation 5.6 by substituting $\epsilon_{ji_{(j)}}$ and $\epsilon_{i_{(j)}j}$ by ϵ,

$$\begin{aligned} \mathrm{SNR}_{i_{(j)}o} &\geq \left|h_{i_{(j)}i_{(j)}}\right|^{-2} \cdot \frac{2^{\frac{R_s}{K}}\left((1-\epsilon)\,2^{\frac{R_s}{K}}+\epsilon\right)-1}{2-\epsilon} &(5.5)\\ \mathrm{SNR}_{jo} &\geq |h_{jj}|^{-2} \cdot \frac{2^{\frac{R_s}{K}}\left((1-\epsilon)\,2^{\frac{R_s}{K}}+\epsilon\right)-1}{2-\epsilon}. &(5.6) \end{aligned}$$

Thus

$$\mathrm{SNR}_{jo} = \frac{2^{\frac{R_s}{K}}\left((1-\epsilon)\,2^{\frac{R_s}{K}}+\epsilon\right)}{2-\epsilon}|h_{jj}|^{-2}, \text{if } j \in D_{\hat{\epsilon}}^c. \tag{5.7}$$

Second, for user k, $k \in D_{\hat{\epsilon}}$, we treat them as being interference-free. In this case, we have

$$\mathrm{SNR}_k \geq \left(2^{\frac{R_s}{K}} - 1\right) |h_{kk}|^{-2}, \text{if } k \in D_{\hat{\epsilon}}. \tag{5.8}$$

Combining Equation 5.6 and Equation 5.8, the minimum sum power required for an equal rate system with sum spectral efficiency R_s is lower bounded by

$$\begin{aligned} \mathrm{SNR}_s \geq & \frac{2^{\frac{R_s}{K}}\left((1-\epsilon)\,2^{\frac{R_s}{K}} + \epsilon\right)}{2-\epsilon} \sum_{j \in D_{\hat{\epsilon}}^c} |h_{jj}|^{-2} \\ & + \left(2^{\frac{R_s}{K}} - 1\right) \sum_{k \in D_{\hat{\epsilon}}} |h_{kk}|^{-2}. \end{aligned} \tag{5.9}$$

Using Equation 2.18 on Equation 5.9 we get

$$\frac{E_b}{N_0}_{\min} = \frac{\sum\left(|h_{jj}|^{-2}\right)}{K} \log 2 \tag{5.10}$$

$$\Delta \mathcal{S}_0 = \frac{(2-\epsilon)}{(4-3\epsilon)(1-\theta) + (2-\epsilon)\,\theta} \tag{5.11}$$

where

$$\theta \triangleq \frac{\sum_{k \in D_{\hat{\epsilon}}} |h_{kk}|^{-2}}{\sum_{j=1}^{K} |h_{jj}|^{-2}} = \frac{\frac{1}{K}\sum_{k \in D_{\hat{\epsilon}}} |h_{kk}|^{-2}}{\frac{1}{K}\sum_{j=1}^{K} |h_{jj}|^{-2}}.$$

Notice that the outer bound converges to the correct $\left.\frac{E_b}{N_0}\right|_{\min}$, and Equation 5.11 can therefore be used as an outer bound on the slope.

Now, we want to show that $\forall \epsilon > 0$, θ can be made arbitrarily small. Define random variable $H_{j,\hat{\epsilon}}$ as

$$H_{j,\hat{\epsilon}} = \begin{cases} |h_{jj}|^{-2} & j \in D_{\hat{\epsilon}} \\ 0 & j \notin D_{\hat{\epsilon}} \end{cases}.$$

Given the fact that $H_{j,\hat{\epsilon}}$ and $H_{i,\hat{\epsilon}}$, $i \neq j$ are independent, and $\sum_{k \in D_{\hat{\epsilon}}} |h_{kk}|^{-2} = \sum_{j=1}^{K} H_{j,\hat{\epsilon}}$, we can apply the law of large number to θ, which gives

$$P\left(\lim_{K \to \infty} \theta = \frac{E\left(H_{j,\hat{\epsilon}}\right)}{E\left(|h_{jj}|^{-2}\right)}\right) = 1. \tag{5.12}$$

Combining Lemma 36and Lemma 37, we have

$$\lim_{\hat{\epsilon}\downarrow 0} E\left(H_{j,\hat{\epsilon}}\right) = 0. \tag{5.13}$$

This proves Equation 5.1 explicitly as follows. For any $\delta > 0$ we can choose $\epsilon, \theta > 0$ sufficiently small to make Equation 5.11 less than $\frac{1}{2} + \delta$. We can choose $\hat{\epsilon} > 0$ sufficiently small to make $\frac{E\left(H_{j,\hat{\epsilon}}\right)}{E\left(|h_{jj}|^{-2}\right)}$ smaller than θ. Finally we can choose K large enough to make $\frac{\sum_{k\in D_{\hat{\epsilon}}}|h_{kk}|^{-2}}{\sum_{j=1}^{K}|h_{jj}|^{-2}}$ smaller than θ with high probability and $\Pr\left(B_{\epsilon,\hat{\epsilon}} = \emptyset\right)$ close to 1. □

5.2.2 The Equal Power Constraint

The equal power constraint will be discussed in this section. Our main results will show that as $K \to \infty$, the event

$$\left\{K \text{ users form } \frac{K}{2} \text{ disjoint } (1-\epsilon) - \text{interference pairs}\right\}$$

happens with high probability, which gives $\Delta\mathcal{S}_0 < 1/2 + \hat{\delta} + \delta$ for all $\delta > 0$, as $K \to \infty$, where $\hat{\delta}$ is a positive real number depending on the distribution function of the channel coefficients.

Assume that the number of users K is an even integer, $K = 2M$. For $\forall\epsilon > 0$, we define event $A_\epsilon \triangleq \{K$ users can form M *disjoint* weak $(1-\epsilon) -$ pairs$\}$, and denote the indices of users belong to the same weak $(1-\epsilon)$-pairs as $\{m_1, m_2\}$.

Let the channel coefficients $h_{ij}, i, j = 1, \ldots, K$ be random variables with a distribution that could depend on K. We consider the following properties of this sequence of distributions

Proposition 39. $\Pr\left(A_\epsilon\right) \to 1$ *as* $K \to \infty$.

If the channel gains h_{ij} *are i.i.d (independent of* K*) with continuous distribution, Property 39 is satisfied.*

Proof. Please see the Appendix section A.6. □

Theorem 40. *If property 39 is satisfied and the* direct *channel gains* h_{jj} *are i.i.d* *with finite 4th order moments , then under the equal power constraint*

$$\forall\delta>0:\lim_{K\to\infty}\Pr\left(\Delta\mathcal{S}_0\le\frac{1}{\frac{\left(E\left[|h_{jj}|^2\right]\right)^2}{E\left[|h_{jj}|^4\right]}+1}+\delta\right)\;=\;1. \tag{5.14}$$

Proof. For the equal power constraint where $\mathrm{SNR}_j=\frac{\mathrm{SNR}_s}{K}$, if property 39 is satisfied, then for $K=2M$ users, M *disjoint* weak $(1-\epsilon)$-pairs $\{m_1,\,m_2\}\,,m=1,\cdots,M$ can be formed with high probability, and we will assume this is the case. Applying Kramer's bound Theorem 12 on each pair, we have

$$\begin{aligned}&R_{m_1}+R_{m_2}\\ \le\;&\min\left(\log\left(1+|h_{m_1m_1}|^2\frac{\mathrm{SNR}_s}{K}\right)+\log\left(1+\frac{|h_{m_2m_2}|^2\,\mathrm{SNR}_s/K}{1+|h_{m_2m_1}|^2\frac{\mathrm{SNR}_s}{K}}\right),\right. &(5.15)\\ &\left.\log\left(1+|h_{m_2m_2}|^2\frac{\mathrm{SNR}_s}{K}\right)+\log\left(1+\frac{|h_{m_1m_1}|^2\,\mathrm{SNR}_s/K}{1+|h_{m_1m_2}|^2\frac{\mathrm{SNR}_s}{K}}\right)\right) &(5.16)\end{aligned}$$

in nats/s. For each weak $(1-\epsilon)$-pair, Equation 5.16 gives

$$\begin{aligned}&\left.\frac{d\left(R_{m_1}+R_{m_2}\right)}{dP_s}\right|_{P_s=0}\\ =\;&\frac{|h_{m_1m_1}|^2+|h_{m_2m_2}|^2}{K}\\ &-\left.\frac{d^2\left(R_{m_1}+R_{m_2}\right)}{d\mathrm{SNR}_s^2}\right|_{P_s=0}\\ \ge\;&\frac{|h_{m_1m_1}|^2+|h_{m_2m_2}|^2}{K^2}+\\ &\frac{2\min\left\{|h_{m_1m_2}|^2\,|h_{m_1m_1}|^2,|h_{m_2m_1}|^2\,|h_{m_2m_2}|^2\right\}}{K^2}\\ \ge\;&\frac{|h_{m_1m_1}|^2+|h_{m_2m_2}|^2}{K^2}+\\ &\frac{2\min\left\{(1-\epsilon_1)\,|h_{m_2m_2}|^2\,|h_{m_1m_1}|^2,(1-\epsilon_2)\,|h_{m_1m_1}|^2\,|h_{m_2m_2}|^2\right\}}{K^2}\\ \ge\;&\frac{|h_{m_1m_1}|^2+|h_{m_2m_2}|^2+2\left(1-\epsilon\right)|h_{m_1m_1}|^2\,|h_{m_2m_2}|^2}{K^2}\end{aligned}$$

since the M pairs are disjoint and using the linearity of derivatives, we have

$$\begin{aligned}
\left.\frac{dR_s}{d\mathrm{SNR}_s}\right|_{P_s=0} &= \sum_{m=1}^{M} \left.\frac{d\left(R_{m_1}+R_{m_2}\right)}{d\mathrm{SNR}_s}\right|_{P_s=0} \\
&= \frac{\sum_{j=1}^{K}\left|h_{jj}\right|^2}{K} \\
-\left.\frac{d^2R_s}{d\mathrm{SNR}_s^2}\right|_{P_s=0} &= \sum_{m=1}^{M}\left(-\left.\frac{d^2\left(R_{m_1}+R_{m_2}\right)}{d\mathrm{SNR}_s^2}\right|_{P_s=0}\right) \\
&\geq \frac{\sum_{m=1}^{M}\left(\left|h_{m_1m_1}\right|^4+\left|h_{m_2m_2}\right|^4+2\left(1-\epsilon\right)\left|h_{m_1m_1}\right|^2\left|h_{m_2m_2}\right|^2\right)}{K^2}
\end{aligned}$$

therefore

$$\begin{aligned}
\frac{E_b}{N_0}_{\min} &= \frac{K\log_e 2}{\sum_{j=1}^{K}\left|h_{jj}\right|^2} \\
\mathcal{S}_s &\leq 2\frac{\left(\sum_{j=1}^{K}\left|h_{jj}\right|^2\right)^2}{\sum_{m=1}^{M}\left(\left|h_{m_1m_1}\right|^4+\left|h_{m_2m_2}\right|^4+2\left(1-\epsilon\right)\left|h_{m_1m_1}\right|^2\left|h_{m_2m_2}\right|^2\right)} \\
&= 2K\frac{\left(\frac{1}{K}\sum_{j=1}^{K}\left|h_{jj}\right|^2\right)^2}{\frac{1}{K}\sum_{j=1}^{K}\left|h_{jj}\right|^4+\left(1-\epsilon\right)\frac{1}{M}\sum_{m=1}^{M}\left|h_{m_1m_1}\right|^2\left|h_{m_2m_2}\right|^2}.
\end{aligned}$$

Now

$$\begin{aligned}
\frac{1}{K}\sum_{j=1}^{K}\left|h_{jj}\right|^2 &\xrightarrow{P} \mathrm{E}\left[\left|h_{jj}\right|^2\right] \\
\frac{1}{K}\sum_{j=1}^{K}\left|h_{jj}\right|^4 &\xrightarrow{P} \mathrm{E}\left[\left|h_{jj}\right|^4\right] \\
\frac{1}{M}\sum_{m=1}^{M}\left|h_{m_1m_1}\right|^2\left|h_{m_2m_2}\right|^2 &\xrightarrow{P} \mathrm{E}\left[\left|h_{jj}\right|^2\right]^2
\end{aligned}$$

as $K \to \infty$, where $\xrightarrow{P}$ stands for convergence in probability since all random variables are positive and the moments are assumed to exist. Using standard rules for convergence of transformation, we then obtain Equation 5.14. □

We will discuss some implications of these theorems. For the equal rate constraint, Equation 5.1 essentially states that the wideband slope is bounded by $\frac{1}{2}$ of that of no interference for large K. Since this is also achievable by Theorem 28, this is indeed the wideband slope, and delay-based interference alignment is optimum. The bound for the equal power constraint is slightly weaker. If the channel coefficients are i.i.d. circularly Gaussian the bound Equation 5.14 for the equal power constraint also gives $\frac{1}{2}$; for other distributions we get a slightly weaker bound.

Theorem 38 and 40 have been proven under an i.i.d. assumption on all channel coefficients. This can seem restrictive and not that realistic in a line of sight model. However, the i.i.d. assumption is not essential. In Theorem 38 it is used to prove that every user has at least one other user with which it forms an $(1-\epsilon)$-pair with high probability. This might be true under many other model assumptions. It is also used to invoke the law of large numbers, which has a wide range of generalizations. In Theorem 40 the i.i.d. assumption is used to prove that users form disjoint weak $(1-\epsilon)$-pairs, and again for invoking the law of large numbers.

What can be concluded is that for small special examples it is possible to find a better wideband slope by optimizing a combination of interference alignment, interference decoding, and treating interference as noise. However, it probably does not pay off to try to find a general algorithm for optimizing wideband slope: comparing the achievable sum slope given by Theorem 28 and the upper bounds provided by Theorem 38 and 40, we could see that as the number of users K grows large, the gap between the upper bounds and the inner bounds achieved by the interference alignment scheme defined by Definition 22 could become arbitrarily small.

Another interesting observation is that the outer bounds do not depend on delay, only on the channel gains. Thus, the outer bound depends on the macroscopic location of nodes (e.g., if gain is proportional to d_{ji}^{α} for some $\alpha > 0$), while the inner bounds depend on the microscopic location (i.e., fractional delay differences).

5.3 Small Bandwidth Case: Generalized Z-Channel Outer Bound

In this section, we develop a new outer bound on the wideband slope for a set of the 2-dimensional vector channels defined by Equation 2.11, under the equal power constraint. The outer bound is specific to the low-rate regime.

The outer bound is derived from the sum Shannon capacity of a type of generalized Z-channel, which is constructed by eliminating a subset of the interference links. In subsection 5.3.1, we show that for a subset of channels $\mathcal{C}$, the optimal sum capacity of their corresponding Z-channels can be achieved by i.i.d. 2-dimensional vector Gaussian inputs. Further, assuming that channel coefficients h_{ji} is drawn from i.i.d. continuous distribution, the set $\mathcal{C}$ has non-zero probability.

In subsection 5.3.2, the Z-channel outer bound is used to derive an outer bound on wideband slope . As a corollary, this outer bound reveals that for any $\delta > 0$, there exists $\mathcal{C}_\delta \subseteq \mathcal{C}$, such that if a channel realization is in $\mathcal{C}_\delta$, then its achievable wideband slope must satisfy $\Delta\mathcal{S}_0 < \frac{1}{K} + \delta$. $\mathcal{C}_\delta$ also has non-zero probability. Further, from previous discussion we know that $\Delta\mathcal{S}_0 = \frac{1}{K}$ is achievable by TDMA. Therefore for the channels in $\mathcal{C}_\delta$, TDMA is near optimal and the pay-off of finding optimal transmission schemes is low. The main results in this section have been partially published in the author's works [39, 40], in collaboration with A. Høst-Madsen.

5.3.1 Generalized Z-Channel And Its Sum Capacity

We define the generalized Z-channel corresponding to the interference channel Equation 2.11 as

$$\underline{\hat{Y}}_j = |h_{jj}|\,\underline{X}_j + \sum_{i=j+1}^{K} |h_{ji}|\,\mathbf{U}_{ji}\underline{X}_i + \underline{Z}_j. \tag{5.17}$$

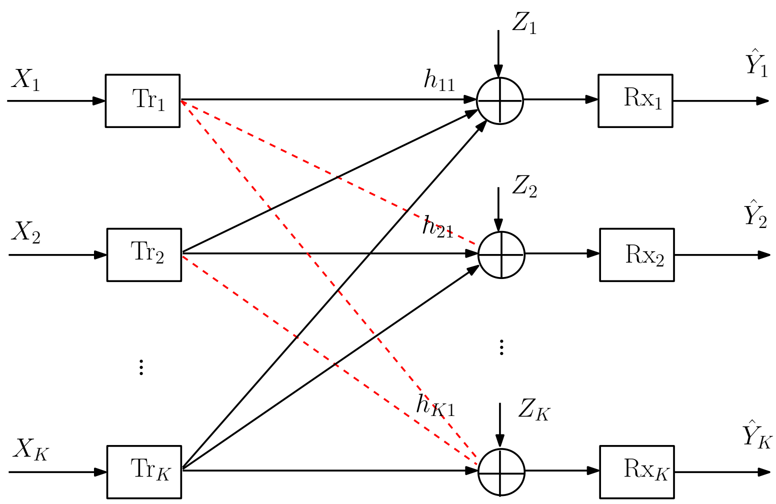

Figure 5.1: Generalized Z-channel

Eliminating a subset of interference links will not reduce channel capacity and therefore, the sum capacity outer bound for the generalized Z-channel is also a sum capacity outer bound for the interference channel.

To derive the Z-channel sum capacity, we provide receiver j, $j = 2, \cdots, K$ with side information $\underline{\mathbf{S}}_j^n = \left(\underline{S}_{j1}^n, \cdots, \underline{S}_{j(j-1)}^n\right)^T$, where

$$\underline{S}_{jp}^n \;=\; |h_{pj}|\,\mathbf{U}_{pj}\underline{X}_j^n + \sum_{i=j+1}^{K} |h_{pi}|\,\mathbf{U}_{pi}\underline{X}_i^n + \underline{W}_{jp}^n \tag{5.18}$$

$p = 1, \cdots, j-1$. . The entries in the length n noise vector $\underline{W}_{jp}^n$ are i.i.d 2×1 vector Gaussian noise with the same marginal distribution as Z_j. Further, they satisfy the following properties

- $\underline{W}_{j(j-1)}^n, \cdots, \underline{W}_{j1}^n$ are independent of all input length n codewords $\underline{X}_i^n$, $i = 1, \cdots, K$;
- $\left(\underline{Z}_j, \underline{W}_{j(j-1)}, \cdots, \underline{W}_{j1}\right)$ are jointly Gaussian random variables, with zero mean and covariance matrix

$$\mathbf{K}_{S_j} = \begin{pmatrix} \mathbf{I} & \mathbf{A}_{j(j-1)} & \cdots & \mathbf{A}_{j1} & \mathbf{A}_{j1} \\ \mathbf{A}_{j(j-1)}^T & \mathbf{I} & \mathbf{A}_{(j-1)(j-2)} & \cdots & \mathbf{A}_{(j-1)1} \\ \vdots & & \ddots & \ddots & \vdots \\ \mathbf{A}_{j1}^T & & & \mathbf{I} & \mathbf{A}_{21} \\ \mathbf{A}_{j1}^T & \mathbf{A}_{(j-1)1}^T & \cdots & \mathbf{A}_{21}^T & \mathbf{I} \end{pmatrix} \tag{5.19}$$

To guarantee such multivariate Gaussian random variable exists, $\mathbf{A}_{jk}$ should be chosen such that for all $j = 1, \cdots, K$

$$\mathbf{K}_{S_j} \;\succeq\; 0 \tag{5.20}$$

We emphasize the following property of $\mathbf{K}_{S_j}$, which will play a key role in the proof of the main result.

Lemma 41. *The distributions of*

$$\underline{S}_{(j-1)p}^n \Big| \underline{S}_{(j-1)(p-1)}^n, \cdots, \underline{S}_{(j-1)1}^n, \underline{X}_{(j-1)}^n$$

and

$$\underline{S}_{jp}^n \Big| \underline{S}_{j(p-1)}^n, \cdots, \underline{S}_{j1}^n$$

are equal.

Proof. Please see the Appendix section A.7. □

Lemma 42. *The distributions of*

$$\underline{\hat{Y}}_{j-1}^{n}\Big|\,\underline{S}_{(j-1)(j-2)}^{n},\cdots,\underline{S}_{(j-1)1}^{n},\underline{X}_{(j-1)}^{n}$$

and

$$\underline{S}_{j(j-1)}^{n}\Big|\,\underline{S}_{j(j-2)}^{n},\cdots,\underline{S}_{j1}^{n}$$

are equal.

Proof. Please see the Appendix section A.8. □

Define the average covariance matrix of the input at transmitter as

$$\tilde{\mathbf{V}}_j \triangleq \frac{1}{n}\sum_{i=1}^{n}\mathrm{E}\left[\underline{X}_j^{(i)}\left(\underline{X}_j^{(i)}\right)^T\right]$$

for any length n input sequence $\underline{X}_j^n$. It must satisfy the power constraint $\tilde{\mathbf{V}}_j \preceq \mathbf{V}_j$. The next lemma states how to choose $\mathbf{A}_{jk}$.

Lemma 43. *Let* $\mathbf{A}_{jp}$*,* $j=2,\cdots,K$ *and* $p=1,\cdots,j-1$ *be*

$$\begin{aligned}\mathbf{A}_{jp} &= \frac{|h_{pj}|^2}{|h_{jj}|^2}\mathbf{U}\left(-\phi_{pj}\right) \\ &+\frac{|h_{pj}|^2}{|h_{jj}|^2}\sum_{i=j+1}^{K}|h_{ji}|^2\,\mathbf{U}\left(\phi_{ji}\right)\mathbf{V}_i\mathbf{U}\left(-\phi_{pj}-\phi_{ji}\right) \\ &-\sum_{i=j+1}^{K}|h_{ji}|^2\,|h_{pi}|^2\,\mathbf{U}\left(\phi_{ji}\right)\mathbf{V}_i\mathbf{U}\left(-\phi_{pi}\right)\end{aligned} \qquad (5.21)$$

If$\mathbf{A}_{jp}$ *defined by Equation 5.21 satisfy* $\mathbf{K}_{S_j}\succeq 0$*, then*

$$\underline{X}_{jG}\rightarrow\ \underline{\hat{Y}}_{jG}\rightarrow\left(\underline{S}_{j1G},\cdots,\underline{S}_{j(j-1)G}\right)^T \qquad (5.22)$$

forms a Markov chain for all $j=2,\cdots,K$*.*

Proof. Please see the Appendix section A.9. □

For a channel realization, denote its channel coefficients by $\underline{h}\triangleq\{h_{ji};\, i,\, j=1,\cdots,K\}$. In the following lemma, we state a sufficient condition on $\underline{h}$ so that $\mathbf{K}_{S_j}\succeq 0$ if $\mathbf{A}_{jp}$ is chosen according to Equation 5.21.

Lemma 44. *For any* $0 < \alpha < 1$ *there exist some* $\epsilon_\alpha, \epsilon'_\alpha > 0$ *and* $\epsilon''_\alpha(\underline{h}) > 0$ *so that if*

$$\begin{aligned}
\underline{h} &\in \mathcal{H}_\alpha \triangleq \left\{ h_{ij} : \left| \frac{|h_{ij}|^2}{|h_{jj}|^2} - \alpha \right| < \epsilon_\alpha, \right. \\
&\qquad |\phi_{ji}| < \epsilon'_\alpha \} && (5.23) \\
P_j &< \epsilon''_\alpha(\underline{h}) && (5.24)
\end{aligned}$$

then $\mathbf{K}_{S_j} \succeq 0$ *for* $\mathbf{A}_{jp}$ *chosen according to Equation 5.21.*

Proof. Proof of Lemma 44 is in Appendix section A.10. □

Our main result of this section is stated in the following theorem.

Theorem 45. *For every interference channel realization* $\underline{h} \in \mathcal{H} = \bigcup_{\alpha \in (0,1)} \mathcal{H}_\alpha$ *defined by Equation 5.23 there exists an* $\epsilon''_\alpha(\underline{h}) > 0$ *so that if* $P_j < \epsilon''_\alpha(\underline{h})$ *the sum capacity of its corresponding Z-channel is given by*

$$\sum_{j=1}^{K} R_j \leq C_{\text{sum}}, \tag{5.25}$$

$$C_{\text{sum}} \tag{5.26}$$

$$= \max_{\substack{\operatorname{Tr}(\mathbf{V}_j) \leq P_j \\ \mathbf{V}_j \succeq \mathbf{0},\, j=1,\cdots,K}} \sum_{j=1}^{K} I\left(\underline{X}_{jG}; \underline{\hat{Y}}_{jG}\right) \tag{5.27}$$

$$= \max_{\substack{\operatorname{Tr}(\mathbf{V}_j) \leq P_j \\ \mathbf{V}_j \succeq \mathbf{0},\, j=1,\cdots,K}} \sum_{j=1}^{K} \log \left| \left(\mathbf{I} + \sum_{i=j}^{K} |h_{ji}|^2 \mathbf{V}_i\right) \left(\mathbf{I} + \sum_{i=j+1}^{K} |h_{ji}|^2 \mathbf{V}_i\right)^{-1} \right| \tag{5.28}$$

Because the sum capacity of the interference channel is outer bounded by the sum capacity of the generalized Z-channel, Equation 5.25 is an outer bound for the sum capacity of the interference channel.

Proof. Proof of Theorem 45 is in Appendix section A.11. □

Note that the bound in Theorem 45 is valid for $P_j < \epsilon''_\alpha(\underline{h})$, and it therefore bounds the actual capacity for suitably low SNR. However, we will mainly use it to bound the wideband slope, a weaker result.

5.3.2 Sum Slope Outer Bound for the Interference Channel

Given the capacity in Theorem 45, we have following result on the low-rate performance of the interference channel.

Theorem 46. *For the interference channel Equation 2.10, the sum capacity is outer bounded by Equation 5.25 for low* SNR. *Under the equal power constraint, the minimum energy per bit of this upper bound satisfy the requirement imposed by Remark 2, which is*

$$\left.\frac{E_b}{N_0}\right|_{\min} = \frac{K \log 2}{\sum_{j=1}^{K} |h_{jj}|^2} \tag{5.29}$$

For channel realizations $\underline{h} \in \mathcal{H} = \bigcup_{\alpha\in(0,1)} \mathcal{H}_\alpha$ defined as Equation 5.23 it therefore gives the following valid upper bound on the sum slope:

$$\mathcal{S}_0 \leq \left(\sum_{j=1}^{K} |h_{jj}|^2\right)^2 \tag{5.30}$$

$$\times \max_{\substack{\mathrm{Tr}\left(\hat{\mathbf{V}}_j\right) \leq 1 \\ \hat{\mathbf{V}}_j \succeq \mathbf{0}}} \left(\sum_{j=1}^{K} |h_{jj}|^4 \mathrm{Tr}\left(\hat{\mathbf{V}}_j^2\right)\right. \tag{5.31}$$

$$\left. +2\sum_{j=1}^{K-1} \sum_{i=j+1}^{K} |h_{jj}|^2 |h_{ji}|^2 \mathrm{Tr}\left(\hat{\mathbf{V}}_j \mathbf{U}_{ji} \hat{\mathbf{V}}_i \mathbf{U}_{ji}^\dagger\right)\right)^{-1} \tag{5.32}$$

Proof. Proof of Theorem 46 is in Appendix section A.12. □

Theorem 47. *For the symmetric channel where $h_{jj} = 1$, $h_{ji} = \alpha \in (0,1)$, the sum slope is bounded by*

$$\mathcal{S}_0 \leq \frac{2K}{\alpha K + (1-\alpha)}$$

Proof. Proof of Theorem 47 is in Appendix section A.13. □

As discussed in the introduction, the wideband slope in the point $\underline{h} = \mathbf{1}$ is $\frac{2}{K}$ per user, achievable by TDMA. Theorem 47 shows that the point $\underline{h} = \mathbf{1}$ is not exceptional in the low-rate regime: for α close to 1 (from below) the channel with $h_{jj} = 1, h_{ji} = \alpha$ has slope close to $\frac{2}{K}$. However, the set of channels $h_{jj} = 1, h_{ji} = \alpha$ still has Lebesgue measure zero, i.e., if the channel coefficients are drawn from a continuous distribution, this set has probability zero. The main result of the paper is the following theorem that shows that the set of channels with slope close to $\frac{2}{K}$ can be be extended to a set of non-zero measure.

Theorem 48. *For all $\sigma > 0$, there exists an open set $\tilde{\mathcal{H}}_\sigma \subset \mathbb{C}^{K(K-1)}$ with $\mathbf{1} \in cl\left(\tilde{\mathcal{H}}_\sigma\right)$, so that for $\underline{h} \in \tilde{\mathcal{H}}_\sigma$*

$$\mathcal{S}_0 \leq 2 + \sigma, \tag{5.33}$$

If the magnitude and phase of the channel coefficients are drawn from continuous random distribution, $Pr\left(\mathcal{H}_\sigma\right) > 0$.

And as $\sigma \to 0$,

$$\lim_{\sigma \to 0} \Delta\mathcal{S}_0 = \frac{1}{K}$$

Because $\Delta\mathcal{S}_0$ achieved by TDMA is $\frac{1}{K}$, when σ is small, TDMA transmission scheme is almost optimal for channels in $\mathcal{H}_\sigma$.

Proof. Proof of Theorem 48 is in Appendix section A.14. □

Chapter 6

Conclusion

This thesis has studied the low-SNR regime power-bandwidth tradeoff of a K-user interference channel with complex scalar channel coefficients, and perfect channel state information at every transmitter and receiver. The power-bandwidth tradeoff is represented by the $R\left(\frac{E_b}{N_0}\right)$ curve, where R is the spectral efficiency and $\frac{E_b}{N_0}$ is the transmitted energy per bit. The low-SNR regime is a concept defined by [1]. A systems working in the low-SNR regime is characterized by small but non-zero spectral efficiency. [1] has shown that in the low-SNR regime, the $R\left(\frac{E_b}{N_0}\right)$ curve is well approximated by its first-order approximation, which is determined by two measures: the minimum energy per bit $\left.\frac{E_b}{N_0}\right|_{\min}$ and the wideband slope $\mathcal{S}_0$. $\left.\frac{E_b}{N_0}\right|_{\min}$ is the optimal achievable transmitted energy per information bit required by reliable communication. $\mathcal{S}_0$ is the first-order slope of the $R\left(\frac{E_b}{N_0}\right)$ curve as $\frac{E_b}{N_0}$ approaching $\left.\frac{E_b}{N_0}\right|_{\min}$.

The 2-user interference channel has been studied by S. Verdú in [1]. Its main results are:

- The optimal $\left.\frac{E_b}{N_0}\right|_{\min}$ of the interference channel is equal to the $\left.\frac{E_b}{N_0}\right|_{\min}$ of the corresponding interference free channel, achievable by treating interference as noise or TDMA. This is also true if $K > 2$.
- For a strong interference channel where ${|h_{ji}|^2}/{|h_{ii}|^2} > 1 \, for \, j = 1,\, 2,\, i \neq j$, its wideband slope region is equivalent to the slope region of the corresponding interference-free channel. The optimal performance can be achieved by joint decoding.
- For a weak interference channel where ${|h_{ji}|^2}/{|h_{ii}|^2} < 1 \, for \, j = 1,\, 2,\, i \neq j$, the best known achievable schemes are treating interference as noise and TDMA.

The main contributions of this thesis are:

Channel Models: The Large Bandwidth Case And The Small Bandwidth Case

[1] has shown that $\left.\frac{E_b}{N_0}\right|_{\min}$ and $\mathcal{S}_0$ can be determined by the first and second order derivatives of $R\,(\text{SNR})$ at $\text{SNR} = 0$, where $\text{SNR} \triangleq \frac{P}{BN_0}$ is the signal to noise ratio per second per Hz. [1] also notes that the low-SNR regime performance can be achieved by letting SNR tends to zero in general, although it has only considered the case where zero SNR is approached by letting B go to infinity.

This thesis has defined two types of low-SNR regime systems in chapter 2. One is the large bandwidth case, where zero SNR is approached as $B \to \infty$ while P is a finite constant number. The other case is the small bandwidth case, where zero SNR is approached as $P \to 0$ while B is a finite constant number. The key difference between these two cases is the behavior of propagation delay. In the large bandwidth case, even very small propagation delay can become arbitrarily large as $\text{SNR} \to 0$, after normalized by the symbol duration B^{-1}; in the small bandwidth case, the propagation is a finite constant as $\text{SNR} \to 0$.

While chapter 3 shows that the performance bounds of a 2-user channel will not be affected by the propagation delay, our results in chapter 4 and chapter 5 imply that if the number of users K is greater than 2, then the interference channel has distinct performances under these two different cases. Therefore, discussing these two $\text{SNR} \to 0$ approaches separately is necessary.

The* 2*-user Channel

In this thesis, the performance of a 2-user interference channel is discussed in chapter 3. First, in section 3.2 and subsection 3.3.1, our discussion shows that the main results of in [1] still hold in the presence of the propagation delay, therefore is applicable to both the large bandwidth case and the small bandwidth case.

Second, for the interference channel with weak or mixed interference link gains, a wideband slope outer bound is developed in subsection 3.3.2, Corollary 14. This bound is derived from the Kramer's capacity outer bound for the discrete-time channel [15]. While Kramer's bound considers delay-free channel model, we have proved that this bound holds if the propagation delay is non-zero in Theorem 12.

Third, in section 3.4, Corollary 20, we derive the noisy interference condition for the low-SNR regime under the equal power constraint. When the channel realization satisfies this condition, treating interference as noise achieves optimal wideband slope. Corollary 20 is based on the noisy interference condition developed by [17], [23] and [24] for the discrete-time channel capacity. While they consider delay-free channel model, we have proved that this condition holds if the propagation delay is non-zero in Theorem 19.

Finally, comparing the achievable wideband slope and its outer bound, we make the observation that for this 2-user interference channel, although the optimal achievable wideband slope region is unknown general, there exists a set of channel realizations with non-zero probability, for whom treating interference as noise or TDMA can achieve fairly good performance, in Remark 21.

K-user Channel: The Large Bandwidth Case

In the large bandwidth case, the low-SNR regime is approached by letting the bandwidth B go to infinity while the power P be fixed and finite. In this case, even very small propagation delay can become arbitrarily large comparing with symbol duration B^{-1}. Contrary to the 2-user case whose performance is independent of delay, K-user channel can use delay to its advantage: when K is greater than 2, extra freedom for system design comes into existence, and the propagation delay will benefit the performance. A type of interference alignment scheme over time domain has been proposed in chapter 4, section 4.3. This transmission scheme achieves $\Delta\mathcal{S}_0 = \frac{1}{2}$ with probability one under the assumption that the delay between transmitter j and receiver i is drawn from a continuous random distribution. Comparing with TDMA which achieves $\Delta\mathcal{S}_0 = \frac{1}{K}$, the achievable scheme we proposed improves the performance by a factor of $K/2$.

The wideband slope outer bounds of K-user interference channel have been discussed in chapter 5, section 5.2. The results have shown that if channel coefficients are drawn from continuously distributed random variable, and the number of users K is large, then:

1. Theorem 38 shows that under equal rate constraint, as $K \to \infty$, with probability one, the achievable sum slope of a channel realization must satisfy $\Delta\mathcal{S}_0 < 1/2 + \epsilon$ for all $\epsilon > 0$.
2. Theorem 40 shows that under equal power constraint, as $K \to \infty$, the achievable sum slope of a channel realization must satisfy $\Delta\mathcal{S}_0 < 1/2 + \epsilon + \delta$ all $\epsilon > 0$, where δ is a positive real number depending on the distribution function of the channel coefficients.

Comparing with the achievable wideband slope region discussed in the previous chapter, we can see that for the large bandwidth case, the interference alignment transmission scheme proposed in chapter 4 section 4.3, which achieves $\Delta\mathcal{S}_0 = 1/2$, is asymptotically optimal under the equal rate constraint, and may achieve fairly good performance under the equal power constraint.

K-user Channel: The Small Bandwidth Case

In the small bandwidth case, the low-SNR regime is approached by letting the input power P go to zero while let the bandwidth be fixed and finite. We have further assumed that B is so small as to render propagation delays negligible comparing with the symbol duration. The achievable wideband slope has been discussed in chapter 4, section 4.4. A phase interference alignment achievable scheme has been proposed. In this scheme, transmitter j only transmits in the direction $e^{\iota\theta_j}$, instead of using the complete 2-dimensional complex plane. Simulation results show that when the interference link gains are moderate, this interference alignment scheme performs better than existing transmission schemes such as treating interference as noise and TDMA with non-zero probability, under the assumption that the phase of the channel coefficient h_{ji} is drawn from a continuous random distribution over $[-\pi, \pi]$.

The wideband slope outer bound has been discussed in chapter 5, section 5.3. This outer bound is developed from a type of generalized Z-channel. Its performance implies that for a set of small bandwidth K-user interference channel realizations, the achievable sum slope of a channel realization must satisfy $\Delta\mathcal{S}_0 < 1/K + \delta$ for all $\delta > 0$, and the probability that a channel realization falls into this set is strictly greater than zero. Therefore, for this set of channels, TDMA is almost optimal.

Future directions

For the large bandwidth case, the interference alignment scheme requires full knowledge of the propagation delay at all transmitters and receivers. One of the future topics is to consider a channel with imperfect channel state information at transmitters and/or receivers.

For the small bandwidth case, our outer bound only works for a subset of channel realizations, whose statistic property is unclear. While if channels are not in this subset, the sum slope has outer bound $\Delta\mathcal{S}_0 < \frac{1}{2} + \epsilon$, for all $\epsilon > 0$, the gap between this outer bound and the achievable sum slope is still wide. Therefore, we hope tighter outer bound and better achievable schemes can be found for this case.

In a more general sense, variations of the interference channel also provide interesting research topics. Some of them are interference channel with user cooperation, additional relay nodes, or equipped with multiple antennas at the transmitters and receivers. For many of these cases, the minimum energy per bit is still unknown. Large open field is yet to be covered to obtain a better understanding on their low-SNR regime performance.

Appendix A

Proofs of Lemmas, Theorems, and Corollaries

A.1 Proof of Theorem 19

Following lemmas will be used in this proof.

Let $\underline{X}^n$=$\{X_1, X_2, \cdots, X_n\}$ be a sequence of random variables satisfying power constraint $\frac{1}{n}\sum_{i=1}^{n}\mathrm{cov}\left(X_i\right) \leq \mathrm{SNR}$. Let $\underline{X}_G^n$=$\{X_{1G}, X_{2G}, \cdots, X_{nG}\}$ be a sequence of i.i.d. Gaussian random variable, $X_G \sim \mathcal{N}\left(0,\, \mathrm{SNR}\right)$. Let $\underline{Z}_1^n$ and $\underline{Z}_2^n$ be two sequence of i.i.d. random variables with distributions $Z_1 \sim \mathcal{N}\left(0,\, \sigma_1^2\right)$ and $Z_2 \sim \mathcal{N}\left(0,\, \sigma_2^2\right)$. Then we have the following inequality

$$\begin{aligned} & h\left(\underline{X}^n + \underline{Z}_1^n\right) - h\left(\underline{X}^n + \underline{Z}_1^n + \underline{Z}_2^n\right) \\ \leq\; & h\left(X_G + Z_1\right) - h\left(X_G + Z_1 + Z_2\right) \end{aligned}$$

Let $\underline{X}^n = \{X_1, X_2 \cdots, X_n\}$ and $\underline{Y}^n = \{Y_1,\, Y_2 \cdots, Y_n\}$ be two sequence of random variables. Let $\hat{X}_G$, $\check{X}_G$ and $\hat{Y}_G$, $\check{Y}_G$ be random variables satisfying

$$\mathrm{cov}\begin{pmatrix}\hat{X}_G\\ \hat{Y}_G\end{pmatrix} \leq \frac{1}{n}\sum_{i=1}^{n}\mathrm{cov}\begin{pmatrix}X_i\\ Y_i\end{pmatrix} \leq \mathrm{cov}\begin{pmatrix}\check{X}_G\\ \check{Y}_G\end{pmatrix}. \tag{A.1}$$

Then

$$\begin{aligned} h\left(\underline{X}^n\right) &\leq nh\left(\hat{X}_G\right) \leq nh\left(\check{X}_G\right) \\ h\left(\underline{Y}^n \middle| \underline{X}^n\right) &\leq nh\left(\hat{Y}_G \middle| \hat{X}_G\right) \leq nh\left(\check{Y}_G \middle| \check{X}_G\right) \end{aligned}$$

This is a special case of [25, Lemma 2].

For the delay-free case where $\tilde{X}_i\left[n - n_{ji}\right] = X_i\left[n\right]$, this theorem is identical to the previous results in [17, 24, 25], and a later work [25]. Here we use similar technique as the proof of Theorem 6 in [25] to show that this results still hold for channel with non-zero delay.

Assume that the channel coefficients and input power constraints satisfy Equation 3.47. Provide side information S_1^n and S_2^n to receiver 1 and 2 respectively

$$\begin{aligned} S_1^n &= h_{21}X_1^n + W_1^n \\ S_2^n &= h_{12}X_2^n + W_2^n \end{aligned}$$

where W_j are zero mean i.i.d Gaussian noise. And the joint distribution of W_j and Z_j is

$$\begin{pmatrix} Z_j \\ W_j \end{pmatrix} \sim \mathcal{N}\left(0, \begin{pmatrix} 1 & \rho_j \\ \rho_j^* & \sigma_j^2 \end{pmatrix}\right)$$

ρ_j and σ_j^2 satisfy Equation 3.42 to Equation 3.45. From Fano's inequality, we have

$$\begin{aligned}
& n\left(R_1 + R_2\right) \\
\leq\ & I\left(X_1^n; Y_1^n\right) + I\left(X_2^n; Y_2^n\right) + o(n) \\
\leq\ & I\left(X_1^n; Y_1^n, S_1^n\right) + I\left(X_2^n; Y_2^n, S_2^n\right) + o(n) \\
\stackrel{(a)}{=}\ & h\left(S_1^n\right) - h\left(S_1^n \middle| X_1^n\right) + h\left(Y_1^n \middle| S_1^n\right) - h\left(Y_1^n \middle| S_1^n, X_1^n\right) \\
& + h\left(S_2^n\right) - h\left(S_2^n \middle| X_2^n\right) + h\left(Y_2^n \middle| S_2^n\right) - h\left(Y_2^n \middle| S_2^n, X_2^n\right) \qquad \text{(A.2)} \\
\stackrel{(b)}{=}\ & h\left(h_{21}X_1^n + W_1\right) - h\left(W_1^n\right) - h\left(h_{12}\tilde{X}_2^n + Z_1^n \middle| W_1^n\right) \\
& + h\left(h_{11}X_1^n + h_{12}\tilde{X}_2^n + Z_1^n \middle| h_{21}X_1^n + W_1^n\right) \\
& + h\left(h_{12}X_2^n + W_2^n\right) - h\left(W_2^n\right) - h\left(h_{21}\tilde{X}_1^n + Z_2^n \middle| W_2^n\right) \\
& + h\left(h_{21}\tilde{X}_1^n + h_{22}X_2^n + Z_2^n \middle| h_{12}X_2^n + W_2^n\right) + o(n) \qquad \text{(A.3)} \\
\stackrel{(c)}{\leq}\ & -nh\left(W_1\right) + h\left(h_{12}X_2^n + W_2^n\right) - h\left(h_{12}X_2^n + Z_1^n \middle| W_1^n\right) \\
& + h\left(h_{11}X_1^n + h_{12}\tilde{X}_2^n + Z_1^n \middle| h_{21}X_1^n + W_1^n\right) \\
& - nh\left(W_2\right) + h\left(h_{21}X_1^n + W_1\right) - h\left(h_{21}X_1^n + Z_2^n \middle| W_2^n\right) \\
& + h\left(h_{21}\tilde{X}_1^n + h_{22}X_2^n + Z_2^n \middle| h_{12}X_2^n + W_2^n\right) + o(n) \qquad \text{(A.4)}
\end{aligned}$$

$$\begin{aligned}
&\overset{(d)}{\leq} -nh\left(W_1\right)+nh\left(h_{12}X_{2G}+W_2\right)-nh\left(h_{12}X_{2G}+Z_1\middle|\,W_1\right) \\
&\quad +h\left(h_{11}X_1^n+h_{12}\tilde{X}_2^n+Z_1^n\middle|\,h_{21}X_1^n+W_1^n\right) \\
&\quad -nh\left(W_2\right)+nh\left(h_{21}X_{1G}+W_1\right)-nh\left(h_{21}X_{1G}+Z_2\middle|\,W_2\right) \\
&\quad +h\left(h_{21}\tilde{X}_1^n+h_{22}X_2^n+Z_2^n\middle|\,h_{12}X_2^n+W_2^n\right)+o(n) \qquad \text{(A.5)} \\
&\overset{(f)}{\leq} -nh\left(W_1\right)+nh\left(h_{12}X_{2G}+W_2\right)-nh\left(h_{12}X_{2G}+Z_1\middle|\,W_1\right) \\
&\quad +nh\left(h_{11}X_{1G}+h_{12}\tilde{X}_{2G}+Z_1\middle|\,h_{21}X_{1G}+W_1\right) \\
&\quad -nh\left(W_2\right)+nh\left(h_{21}X_{1G}+W_1\right)-nh\left(h_{21}X_{1G}+Z_2\middle|\,W_2\right) \\
&\quad +nh\left(h_{21}\tilde{X}_{1G}+h_{22}X_{2G}+Z_2\middle|\,h_{12}X_{2G}+W_2\right)+o(n) \qquad \text{(A.6)} \\
&\overset{(g)}{\leq} -nh\left(W_1\right)+nh\left(h_{12}X_{2G}+W_2\right)-nh\left(h_{12}X_{2G}+Z_1\middle|\,W_1\right) \\
&\quad +nh\left(h_{11}X_{1G}+h_{12}X_{2G}+Z_1\middle|\,h_{21}X_{1G}+W_1\right) \\
&\quad -nh\left(W_2\right)+nh\left(h_{21}X_{1G}+W_1\right)-nh\left(h_{21}X_{1G}+Z_2\middle|\,W_2\right) \\
&\quad +nh\left(h_{21}X_{1G}+h_{22}X_{2G}+Z_2\middle|\,h_{12}X_{2G}+W_2\right)+o(n) \qquad \text{(A.7)}
\end{aligned}$$

where $\lim_{n\to\infty} o(n)/n = 0$, X_{jG} with 'G' subscription means that input at transmitter j is i.i.d. Gaussian, with distribution $X_{jG} \sim \mathcal{N}(0, \mathrm{SNR}_j)$. (a) is from chain rule. (c) holds because both X_j and $\tilde{X}_j$ can be obtained from sampling the same continuous-time base band signal $X_j(t)$ at the Nyquist rate, while Z_j and W_j are sampled from white Gaussian noise, so that

$$\begin{aligned}
h\left(h_{21}\tilde{X}_1^n+Z_2^n\middle|\,W_2^n\right) &= h\left(h_{21}X_1^n+Z_2^n\middle|\,W_2^n\right)+o(n) \\
h\left(h_{12}\tilde{X}_2^n+Z_1^n\middle|\,W_1^n\right) &= h\left(h_{12}X_2^n+Z_1^n\middle|\,W_1^n\right)+o(n)
\end{aligned}$$

because Given Equation 3.42 and Equation 3.43, $\mathrm{cov}\left(W_1^n\right) \leq \mathrm{cov}\left(Z_2^n\middle|\,W_2^n\right)$ and $\mathrm{cov}\left(W_2^n\right) \leq \mathrm{cov}\left(Z_1^n\middle|\,W_1^n\right)$. Combining Lemma section A.1 and [23, Lemma 3], we have

$$\begin{aligned}
& h\left(h_{12}X_2^n+W_2^n\right)-h\left(h_{12}X_2^n+Z_1^n\middle|\,W_1^n\right) \\
\leq\ & nh\left(h_{12}X_{2G}+W_2\right)-nh\left(h_{12}X_{2G}+Z_1\middle|\,W_1\right) \qquad \text{(A.8)}
\end{aligned}$$

and

$$\begin{aligned}
& h\left(h_{21}X_1^n+W_1^n\right)-h\left(h_{21}X_1^n+Z_2^n\middle|\,W_2^n\right) \\
\leq\ & nh\left(h_{21}X_{1G}+W_1\right)-nh\left(h_{21}X_{1G}+Z_2\middle|\,W_2\right). \qquad \text{(A.9)}
\end{aligned}$$

Therefore (d) is true.

(f) is from Lemma section A.1, where $\tilde{X}_{jG}$ are i.i.d. Gaussian random variable satisfying $\operatorname{cov}\left(\tilde{X}_{jG}\right) = \frac{1}{n}\operatorname{Tr}\left(\tilde{X}_j^n\left(\tilde{X}_j^n\right)^\dagger\right)$. Denote $\mathrm{SNR}_j' \triangleq \frac{1}{n}\operatorname{Tr}\left(\tilde{X}_j^n\left(\tilde{X}_j^n\right)^\dagger\right)$ as the power of $\tilde{X}_j^n$. We could see that $\mathrm{SNR}_j' \leq \mathrm{SNR}_j$, because time-shifting of a signal sampled at the Nyquist rate does not change signal power. Therefore we have (g). This shows that the sum capacity of a channel with delay is outer bounded by that of a channel without delay.

We could see that the inequality (g) is independent of the propagation delay. It is identical to the first inequality in [25, (89)]. Therefore, from this point on, the proof will be the same as the delay-free case. We will not proceed here.

A.2 Proof of Lemma 27

We have

$$\begin{aligned}
\mathrm{E}\left[\tilde{x}_i^*[n_1,\delta_{ji}]\tilde{x}_i[n_2,\delta_{ji}]\right] &= \mathrm{E}\left[\left(\sum_{m=-\infty}^{\infty} x_i[2m]\operatorname{sinc}(n_1-2m+\delta_{ji})\right)^* \cdot \right.\\
&\qquad \left.\left(\sum_{m=-\infty}^{\infty} x_i[2m]\operatorname{sinc}(n_2-2m+\delta_{ji})\right)\right]\\
&= \sum_{m=-\infty}^{\infty} \mathrm{E}\left[|x_i[2m]|^2\right]\operatorname{sinc}(n_1-2m+\delta_{ji})\operatorname{sinc}(n_2-2m+\delta_{ji})\\
&= \sum_{m=0}^{\infty} \mathrm{E}\left[|x_i[2m]|^2\right]\operatorname{sinc}(n_1-2m+\delta_{ji})\operatorname{sinc}(n_2-2m+\delta_{ji})\\
&\quad + \sum_{m=1}^{\infty} \mathrm{E}\left[|x_i[-2m]|^2\right]\operatorname{sinc}(n_1+2m+\delta_{ji})\operatorname{sinc}(n_2+2m+\delta_{ji})\\
= 2P_i &\left(\sum_{m=0}^{\infty}\operatorname{sinc}(n_1-2m+\delta_{ji})\operatorname{sinc}(n_2-2m+\delta_{ji})\right.\\
&\quad \left. + \sum_{m=1}^{\infty}\operatorname{sinc}(n_1+2m+\delta_{ji})\operatorname{sinc}(n_2+2m+\delta_{ji})\right). \qquad \text{(A.10)}
\end{aligned}$$

Define $f_m(\delta_{ji})$ and $g_m(\delta_{ji})$ as

$$\begin{aligned}
f_m(\delta_{ji}) &\triangleq \operatorname{sinc}(n_1-2m+\delta_{ji})\operatorname{sinc}(n_2-2m+\delta_{ji})\\
g_m(\delta_{ji}) &\triangleq \operatorname{sinc}(n_1+2m+\delta_{ji})\operatorname{sinc}(n_2+2m+\delta_{ji})
\end{aligned}$$

and their partial sums $s_{f,M}(\delta_{ji}) = \sum_{m=0}^{M} f_m(\delta_{ji})$, $s_f(\delta_{ji}) \triangleq \lim_{M\to\infty} s_{f,M}(\delta_{ji})$; $s_{g,M}(\delta_{ji}) = \sum_{m=0}^{M} g_m(\delta_{ji})$, $s_g(\delta_{ji}) \triangleq \lim_{M\to\infty} s_{g,M}(\delta_{ji})$. Here

$$\begin{aligned}
|f_m(\delta_{ji})| &= \left|\frac{\sin(\pi(n_1-2m+\delta_{ji}))\sin(\pi(n_2-2m+\delta_{ji}))}{(n_1-2m+\delta_{ji})(n_2-2m+\delta_{ji})}\right| \\
&\le \frac{1}{(n_1-2m+\delta_{ji})(n_2-2m+\delta_{ji})}
\end{aligned}$$

and

$$|g_m(\delta_{ji})| \le \frac{1}{(n_1+2m+\delta_{ji})(n_2+2m+\delta_{ji})}.$$

Let $M_{f,k} \triangleq \frac{1}{(n_1-2m+\delta_{ji})(n_2-2m+\delta_{ji})}$, $M_{g,k} \triangleq \frac{1}{(n_1+2m+\delta_{ji})(n_2+2m+\delta_{ji})}$. Because $\sum_{m=1}^{\infty}\frac{1}{k^2}$ is convergent, $\sum_{k=0}^{\infty} M_{f,k}$ and $\sum_{k=1}^{\infty} M_{g,k}$ converge too. Due to Weierstrass's test for uniform convergence[42], $s_{f,M}(\delta_{ji})$ and $s_{g,M}(\delta_{ji})$ converge uniformly. And using Theorem 7.11 in [42], we have

$$\begin{aligned}
\lim_{\delta_{ji}\downarrow 0}\lim_{M\to\infty} s_{f,M}(\delta_{ji}) &= \lim_{M\to\infty}\lim_{\delta_{ji}\downarrow 0} s_{f,M}(\delta_{ji}), \\
\lim_{\delta_{ji}\downarrow 0}\lim_{M\to\infty} s_{g,M}(\delta_{ji}) &= \lim_{M\to\infty}\lim_{\delta_{ji}\downarrow 0} s_{g,M}(\delta_{ji}).
\end{aligned}$$

Thus, Equation A.10 becomes

$$\begin{aligned}
\lim_{\delta_{ji}\downarrow 0} \mathrm{E}\left[\tilde{x}_i^*[n_1,\delta_{ji}]\tilde{x}_i[n_2,\delta_{ji}]\right] &= 2P_i\Bigg(\lim_{M\to\infty}\lim_{\delta_{ji}\downarrow 0} s_{f,M}(\delta_{ji}) \\
&\quad + \lim_{M\to\infty}\lim_{\delta_{ji}\downarrow 0} s_{g,M}(\delta_{ji})\Bigg) \\
&= \begin{cases} 2P_i & \text{if } n_1 = n_2 = 2k, \\ & \text{for some integer } k \\ 0 & o.w. \end{cases}.
\end{aligned}$$

Given Theorem 7.12 in [42] and the continuity of sinc function, we can conclude that $\mathrm{E}\left[\tilde{x}_i^*[n_1,\delta_{ji}]\tilde{x}_i[n_2,\delta_{ji}]\right]$is a continuous function of δ_{ji}.

A.3 Proof of Lemma 35

Let $C_{i,\epsilon}$ be the event that user$i$ does not form an $(1-\epsilon)$-pair with any other user;. Then

$$\begin{aligned}
\Pr(B_{\epsilon,\hat{\epsilon}} \neq \emptyset) &= \Pr\left(\bigcup_{i=1}^{K} C_{i,\epsilon}\right) \\
&\leq \sum_{i=1}^{K} \Pr(C_{i,\epsilon}) \\
&= K \Pr(C_{1,\epsilon})
\end{aligned}$$

and

$$\begin{aligned}
\Pr(C_{1,\epsilon}) &= \Pr\left(\forall j > 1 : \frac{|h_{j1}|^2}{|h_{11}|^2} or \frac{|h_{1j}|^2}{|h_{jj}|^2} \notin (1-\epsilon, 1) \text{ and } |h_{11}|^2 \notin R_{\hat{\epsilon}}\right) \\
&= \Pr\left(\forall j > 1 : \frac{|h_{j1}|^2}{|h_{11}|^2} \notin (1-\epsilon, 1) \text{ and } |h_{11}|^2 \notin R_{\hat{\epsilon}}\right) \\
&\quad +\Pr\left(\forall j > 1 : \frac{|h_{1j}|^2}{|h_{jj}|^2} \notin (1-\epsilon, 1)\right) \\
&\quad -\Pr\left(\forall j > 1 : \frac{|h_{j1}|^2}{|h_{11}|^2} \text{ and } \frac{|h_{1j}|^2}{|h_{jj}|^2} \notin (1-\epsilon, 1) \text{ and } |h_{11}| \notin R_{\hat{\epsilon}}\right) \\
&= \left(1 - p_j^{K-1}\right) \Pr\left(\forall j > 1 : \frac{|h_{j1}|^2}{|h_{11}|^2} \notin (1-\epsilon, 1) \text{ and } |h_{11}|^2 \notin R_{\hat{\epsilon}}\right) + p_j^{K-1}
\end{aligned}$$

where $p_j = \Pr\left(\frac{|h_{1j}|^2}{|h_{jj}|^2} \notin (1-\epsilon, 1)\right) \in (0,1)$. Notice that the events $\frac{|h_{1j}|^2}{|h_{jj}|^2} \notin (1-\epsilon, 1)$ are independent for different j, and $p_j^{K-1} \to 0$. Thus,

$$\begin{aligned}
\Pr(C_{1,\epsilon}) \to \Pr\Bigg(\forall j > 1 : \frac{|h_{j1}|^2}{|h_{11}|^2} &\notin (1-\epsilon, 1) \\
\text{and } |h_{11}|^2 &\notin R_{\hat{\epsilon}}\Bigg).
\end{aligned}$$

As the $|h_{jj}|^2$ are independent,

$$\begin{aligned}
& \Pr\left(\forall j > 1 : \frac{|h_{j1}|^2}{|h_{11}|^2} \notin (1-\epsilon, 1) \text{ and } |h_{11}|^2 \notin R_{\hat{\epsilon}}\right) \\
= & \int_{x \notin R_{\hat{\epsilon}}} \prod_{j=2}^{K} \left(1 - \int_{(1-\epsilon)x}^{x} dF_{|h_{j1}|^2}(u)\right) dF_{|h_{11}|^2}(x) \\
= & \int_{x \notin R_{\hat{\epsilon}}} \left(1 - F_{|h_{ii}|^2}(x) + F_{|h_{ii}|^2}((1-\epsilon)x)\right)^{K-1} dF_{|h_{ii}|^2}(x) \\
\leq & (1-\hat{\epsilon})^{K-1} \int_{x \notin R_{\hat{\epsilon}}} dF_{|C_{ii}|^2}(x) \\
= & (1 - \mu_F(R_{\hat{\epsilon}}))(1-\hat{\epsilon})^{K-1}.
\end{aligned}$$

Thus, $P_\sigma \leq K(1 - \mu_F(R_{\hat{\epsilon}}))(1-\hat{\epsilon})^{K-1}$, and $\lim_{K\to\infty} \Pr(B_{\epsilon,\hat{\epsilon}} \neq \emptyset) = 0$.

A.4 Proof of Lemma 36

Given the definition $\forall x \in R_{\hat{\epsilon}} : F_{|h_{ii}|^2}(x) - F_{|h_{ii}|^2}((1-\epsilon)x) \leq \hat{\epsilon}$, and given the fact $F_{|h_{ii}|^2}(x) - F_{|h_{ii}|^2}((1-\epsilon)x) < \hat{\epsilon}_{n+1} < \hat{\epsilon}_n$, it clearly follows that $R_{\hat{\epsilon}_{n+1}} \subseteq R_{\hat{\epsilon}_n}$. Let $I_{\hat{\epsilon}_n}(x)$ be the indicator function of $R_{\hat{\epsilon}_n}$. Using Lebesgue dominated convergence we have

$$\begin{aligned}
& \lim_{n\to\infty} \int_0^\infty I_{\hat{\epsilon}_n}(x) dF_{|h_{ii}|^2}(x) \\
= & \int_0^\infty \lim_{n\to\infty} I_{\hat{\epsilon}_n}(x) dF_{|h_{ii}|^2}(x) \\
= & \int_0^\infty I_0(x) dF_{|h_{ii}|^2}(x)
\end{aligned}$$

where $I_0(x)$ is the indicator function of the set $R_0 = \{x \in \mathbb{R} : F_{|h_{ii}|^2}(x) - F_{|h_{ii}|^2}((1-\epsilon)x) = 0\}$. Since we have assumed that $E\left[|h_{ii}|^{-2}\right] < \infty$, also $\Pr(|h_{ii}|^2 = 0) = 0$ and clearly $\mu_F(R_0) = 0$, and $\mu(R_{\hat{\epsilon}_n}) \to \mu(R_0)$.

A.5 Proof of Lemma 37

To be explicit, let X be a random variable on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ [43]. Given the definition of X_i, we can conclude that

$$\lim_{i\to\infty} X_i \;=\; 0 \qquad \text{w.p. } 1.$$

Namely, if there is a set $B \in \mathcal{F}$ with $\mathbb{P}(B) > 0$ where $\lim_{i\to\infty} X_i \neq 0$ then $\mathbb{P}(B) \leq \mu_X\left(\bigcap_{i=1}^{\infty} G_i\right)$ which contradicts $\lim_{i\to\infty} \mu_X(G_i) = 0$.

Now $X_i \leq X$, and therefore by Lebesgue dominated convergence

$$\lim_{i\to\infty} E[X_i] \;=\; E[\lim_{i\to\infty} X_i] \;=\; 0.$$

A.6 Proof of Property 39

Model the interference channel as a graph G_K, with $K = 2M$ vertices $u_1, u_2, \cdots, u_{2n}$. Vertices u_i and u_j are connected by edge E_{ij} if they form a weak$(1-\epsilon)$-pair, i.e., $(1-\epsilon) \leq \frac{|h_{ij}|^2}{|h_{jj}|^2} \leq 1$ or $(1-\epsilon) \leq \frac{|h_{ji}|^2}{|h_{ii}|^2} \leq 1$. Divide vertices into two disjoint classes $V_1 = \{u_1, u_2, \cdots, u_M\}$ and $V_2 = \{u_{M+1}, u_{M+2}, \cdots, u_{2M}\}$. Now define event

$$\hat{A}_\epsilon \;=\; \{\text{there exists a perfect matching in the bipartite graph } G_{M,M}\}.$$

As $\hat{A}_\epsilon \subseteq A_\epsilon$, $P(A_\epsilon) \geq P\left(\hat{A}_\epsilon\right)$. Thus, if we can show that $P\left(\hat{A}_\epsilon\right) = 1 - o(1)$ as $K \to \infty$ then Property 39 holds.

For any bipartite graph, a perfect matching exists if Hall's condition is satisfied.

Given a bipartite graph $G_{M,M}$ with disjoint vertices class V_1 and V_2, $V_1 \bigcup V_2 = V$, $|V_i| = M$, whose set of edges is $E(G_{M,M})$, a perfect matching exists if and only if for every $S \subseteq V_i, i = 1or2$, $|N(S)| \geq |S|$, where

$$N(S) = \{y : \; xy \in E(G_{M,M}) \;\; for\ some\ x \in S\}.$$

Any bipartite graph that does not have a perfect matching has following properties

Suppose $G_{M,M}$ has no isolated vertices and it does not have a perfect matching. Then Hall's condition must be violated by some set $A \subset V_i$, $i = 1\, or\, 2$. And such set with minimal cardinality satisfies the following necessary conditions

(i) $|N(A)| = |A| - 1$;

(ii) $2 \leq |A| \leq \left\lceil \frac{M}{2} \right\rceil$

(iii) the subgraph of G spanned by $A \bigcup N(A)$ is connected, and it has at least $2a-2$ edges;

(iv) every vertex in $N(A)$ is adjacent to at least two vertices in A;

(v) any subsets of $N(A)$ can find a perfect matching in $|A|$;

(i), (ii), (iii), and (iv) are proved by Lemma 7.12 in [44], and p.82 of [45]. And (iv) is true because if there exists a subset B of $N(A)$ that can not find a perfect match, we could just let B be $\hat{A}$, and its neighbors in A be $N\left(\hat{A}\right)$. Then $\hat{A}$ violates Hall's condition, while $\left|\hat{A}\right| < a$. This contradicts the assumption that A is the minimal set violating Hall's condition.

Define the event F_a: there is a set $A \subset V_i$, $i = 1\, or\, 2$, $|A| = a$. satisfying (i), (ii) and (iii) in Lemma section A.6. [44] shows that for a graph with no isolated vertex, $P(A_\epsilon) = 1 - o(1)$ is equivalent to $P\left(\bigcup_{a=2}^{\lceil \frac{M}{2} \rceil} F_a\right) = o(1)$. Define F_1 as the event that there exists at least one isolated vertex in $G_{M,M}$. In our case, we want to show that

$$P\left(\bigcup_{a=2}^{\lceil \frac{M}{2} \rceil} F_a\right) + P(F_1) \quad = \quad o(1).$$

Using the union bound, we have

$$\begin{aligned} P(F_1) &\le \sum_{i=1}^{2M} P(u_i\ isolated) \\ &\le 2M \cdot P(u_1\ isolated) \\ &\overset{(a)}{\le} 2M \cdot (1 - p_{1j})^M \\ &\overset{(b)}{=} o(1) \end{aligned}$$

where $p_{1j} \triangleq P\left((1-\epsilon) \le \frac{|h_{1j}|^2}{|h_{jj}|^2} \le 1\right)$, $j = M+1, \cdots, 2M$. We also define $p_0 \triangleq P\left((1-\epsilon) \le \frac{|h_{ij}|^2}{|h_{jj}|^2} \le 1, or\,(1-\epsilon) \le \frac{|h_{ji}|^2}{|h_{ii}|^2} \le 1\right)$ for later use. (a) holds because the event $\frac{|h_{1j}|^2}{|h_{jj}|^2} \notin (1-\epsilon, 1)$ is independent of j. And it is a necessary condition for V_1 to be is isolated.

Now, let us look into F_a for $2 \le a \le \left\lceil \frac{M}{2} \right\rceil$. Let $A_1 \subset V_1$, $A_2 \subset V_2$, and $|A_1| = |A_2| + 1 = a$. Denote $P(\mathcal{A}_a)$ as the probability that the subgraph of $G_{M,M}$ spanned by $A_1 \bigcup A_2$ satisfies (i), (ii), and (iii) in Lemma section A.6. We have

$$\begin{aligned} P\left(\bigcup_{a=2}^{\lceil\frac{M}{2}\rceil} F_a\right) &\overset{(d)}{\leq} \sum_{a=2}^{\lceil\frac{M}{2}\rceil} P\left(F_a\right) \\ &\overset{(e)}{\leq} 2\sum_{a=2}^{\lceil\frac{M}{2}\rceil} \binom{M}{a}\binom{M}{a-1} P\left(\mathcal{A}_a\right) \end{aligned} \tag{A.11}$$

(d) is from the union bound; (e) is from the union bound, and from that fact that there are $2\binom{M}{a}$ choices for A with $|A| = a$, and $\binom{M}{a-1}$ more choices for $N(A)$. In [44, 45], the case where edge probabilities are i.i.d, whose value is p, is considered. In [44], $P(\mathcal{A}_a)$ is bounded using condition (i), (ii) and (iii), which gives $P(\mathcal{A}_a) \leq \binom{a(a-1)}{2a-2} p^{2a-2}p^{a(n-a+1)}$. The term $p^{a(n-a+1)}$ is the probability that the vertices in A_1 do not connect to vertices in $V_2 - A_2$. And in [45], condition (iv) instead of (iii) are used, which gives $P(\mathcal{A}_a) \leq \binom{a}{2}^{a-1} p^{2a-2}p^{a(n-a+1)}$. However, in our case, any two edges having adjacent vertices are dependent. So we use condition (v). Since for $N(A)$, a perfect match exists, then the subgraph spanned by $A \bigcup N(A)$ has $a-1$ edges that are not adjacent with each other. Thus, $P(\mathcal{A}_a)$ can be bounded by

$$\begin{aligned} P\left(\mathcal{A}_a\right) \leq\; & Pr\,(\text{condition (i), (ii) and (iv) are satisfied,} && \text{(A.12)} \\ & \text{vertices in A}_1 \text{ do not connect to vertices in } V_2 - A_2) && \text{(A.13)} \\ & \left(p_0^{a-1}\prod_{k=2}^{a}\binom{k}{1}\right) p_{A_1\bar{A}_2} && \text{(A.14)} \end{aligned}$$

where

$$\begin{aligned} p_{A_1\bar{A}_2} &= P\left(\frac{|h_{ij}|^2}{|h_{jj}|^2} \notin \left[(1-\epsilon),1\right], \text{ for all } u_i \in A_1 \,and\, u_j \in (V_2 - A_2)\right) \\ &\leq \prod_{u_j\in(V_2-A_2)} P\left(\frac{|h_{ij}|^2}{|h_{jj}|^2} \notin \left[(1-\epsilon),1\right], \text{ for all } u_i \in A_1\right) \\ &= \left(P\left(\frac{|h_{ji}|^2}{|h_{jj}|^2} \notin \left[(1-\epsilon),1\right], i = 1,\cdots,a\; j = M+a\right)\right)^{M-a+1} \end{aligned} \tag{A.15}$$

notice that the event $\frac{|h_{ij}|^2}{|h_{jj}|^2} \notin \left[(1-\epsilon),1\right]$, for all $u_i \in A_1 \,and\, u_j \in (V_2 - A_2)$ is an necessary Substitute $|h_{j1}|^2$ by x_j, and $|h_{11}|^2$ by x_1, denote their CDF by $F_x(x)$, and their joint CDF $F_{\underline{\mathrm{X}}}(\underline{\mathrm{x}})$. Notice that $|h_{ij}|^2$ are i.i.d. distributed. Then

$$\begin{aligned}
& P\left(\frac{|h_{ji}|^2}{|h_{jj}|^2} \notin [(1-\epsilon),1], i=1,\cdots,a\ j=M+a\right) \\
= & \int_{A_{\underline{\mathrm{X}}}} dF_{\underline{\mathrm{X}}}(\underline{\mathrm{x}}) \\
= & \int_0^\infty f_{x_{M+a}}(x_{M+a}) \prod_{i=1}^{a} \left(\int_{\frac{x_i}{x_{M+a}} \notin [(1-\epsilon),1]} f_{x_i}(x_i)\, dx_i\right) dx_{M+a} \\
= & \int_0^\infty f_{x_1}(x_1) \left(\int_{\frac{x_1}{x_{M+a}} \notin [(1-\epsilon),1]} f_{x_{M+1}}(x_{M+1})\, dx_{M+1}\right)^a dx_{M+a} \\
\overset{(f)}{\leq} & \left(\int_0^\infty f_{x_1}^2(x_1)\, dx_1\right)^{1/2} \left(\int_0^\infty g_{x_1}^2(x_1)\, dx_1\right)^{a/2}
\end{aligned} \tag{A.16}$$

where $g(x_1) \triangleq \int_{\frac{x_1}{x_{M+a}} \notin [(1-\epsilon),1]} f_{x_{M+1}}(x_{M+1})\, dx_{M+1}$. (f) is from Cauchy-Schwartz inequality. Denote

$$q_1 \triangleq \left(\int_0^\infty f_{x_1}^2(x_1)\, dx_1\right)^{\frac{1}{2(M-a+1)}} \left(\int_0^\infty g_{x_1}^2(x_1)\, dx_1\right)^{1/2}$$

$q_1 < 1$, and $\lim_{M\to\infty} q_1 = \left(\int_0^\infty g_{x_1}^2(x_1)\, dx_1\right)^{1/2}$. Notice that this limit value do not depend on the value of M. Now combining Equation A.15 and Equation A.16, we have

$$P(\mathcal{A}_a) \leq \left(p_0^{a-1} \prod_{k=2}^{a} \begin{pmatrix} k \\ 1 \end{pmatrix}\right) q_1^{a(M-a+1)} \tag{A.17}$$

. Given Equation A.11 and Equation A.17,

$$
\begin{aligned}
P\left(\bigcup_{a=2}^{\lceil\frac{M}{2}\rceil} F_a\right) &\leq 2\sum_{a=2}^{\lceil\frac{M}{2}\rceil}\begin{pmatrix} M \\ a \end{pmatrix}\begin{pmatrix} M \\ a-1 \end{pmatrix}\left(p_0^{a-1}\prod_{k=2}^{a}\begin{pmatrix} k \\ 1 \end{pmatrix}\right) q_1^{a(M-a+1)} \\
&\leq 2\sum_{a=2}^{\lceil\frac{M}{2}\rceil}\left(\frac{eM}{a}\right)^{a}\left(\frac{eM}{a-1}\right)^{a-1} a^{a-1}p_0^{a-1}q_1^{a(M-a+1)} \\
&\leq 2q_1\sum_{a=2}^{\lceil\frac{M}{2}\rceil}\left(\frac{e^2M^2}{(a-1)^2}ap_0q_1^{\frac{M}{2}}\right)^{a-1} \\
&\leq 2q_1\sum_{a=2}^{\lceil\frac{M}{2}\rceil}\left(2e^2p_0M^2q_1^{\frac{M}{2}}\right)^{a-1} \\
&\leq 2q_1\frac{M}{2}\left(2e^2p_0M^2q_1^{\frac{M}{2}}\right) \\
&= o(1)
\end{aligned}
$$

. This means we can find a perfect matching with high probability, i.e., $1-o(1)$.

A.7 Proof of Lemma 41

Given Equation 5.18, we have

$$
\begin{aligned}
\underline{S}^n_{(j-1)p} &= \left|h_{(j-1)j}\right| \mathbf{U}_{p(j-1)}\underline{X}^n_{(j-1)} + \sum_{i=j}^{K}\left|h_{pi}\right| \mathbf{U}_{pi}\underline{X}^n_i + \underline{W}^n_{(j-1)p} \\
\underline{S}^n_{(j-1)(p-1)} &= \left|h_{(j-1)j}\right| \mathbf{U}_{(p-1)(j-1)}\underline{X}^n_{(j-1)} + \sum_{i=j}^{K}\left|h_{(p-1)i}\right| \mathbf{U}_{(p-1)i}\underline{X}^n_i + \underline{W}^n_{(j-1)(p-1)} \\
&\vdots \\
\underline{S}^n_{(j-1)1} &= \left|h_{(j-1)j}\right| \mathbf{U}_{1(j-1)}\underline{X}^n_{(j-1)} + \sum_{i=j}^{K}\left|h_{1i}\right| \mathbf{U}_{1i}\underline{X}^n_i + \underline{W}^n_{(j-1)1}
\end{aligned}
$$

When $\underline{X}^n_{(j-1)}$ is given, it can be subtracted from $\underline{S}^n_{(j-1)p}, \underline{S}^n_{(j-1)(p-1)}, \cdots, \underline{S}^n_{(j-1)1}$ to give

$$\begin{aligned}
\underline{\hat{S}}^n_{(j-1)p} &= \underline{S}^n_{(j-1)p} - \left|h_{(j-1)j}\right| \mathbf{U}_{p(j-1)}\underline{X}^n_{(j-1)} \\
&= \sum_{i=j}^{K} |h_{pi}| \mathbf{U}_{pi}\underline{X}^n_i + \underline{W}^n_{(j-1)p} \qquad \text{(A.18)} \\
\underline{\hat{S}}^n_{(j-1)(p-1)} &= \underline{S}^n_{(j-1)(p-1)} - \left|h_{(j-1)j}\right| \mathbf{U}_{(p-1)(j-1)}\underline{X}^n_{(j-1)} \\
&= \sum_{i=j}^{K} \left|h_{(p-1)i}\right| \mathbf{U}_{(p-1)i}\underline{X}^n_i + \underline{W}^n_{(j-1)(p-1)} \qquad \text{(A.19)} \\
&\vdots \\
\underline{\hat{S}}^n_{(j-1)1} &= \underline{S}^n_{(j-1)1} - \left|h_{(j-1)j}\right| \mathbf{U}_{1(j-1)}\underline{X}^n_{(j-1)} \\
&= \sum_{i=j}^{K} |h_{1i}| \mathbf{U}_{1i}\underline{X}^n_i + \underline{W}^n_{(j-1)1} \qquad \text{(A.20)}
\end{aligned}$$

while

$$\begin{aligned}
\underline{S}^n_{jp} &= \sum_{i=j}^{K} |h_{pi}| \mathbf{U}_{pi}\underline{X}^n_i + \underline{W}^n_{jp} \qquad \text{(A.21)} \\
\underline{S}^n_{j(p-1)} &= \sum_{i=j}^{K} \left|h_{(p-1)i}\right| \mathbf{U}_{(p-1)i}\underline{X}^n_i + \underline{W}^n_{j(p-1)} \qquad \text{(A.22)} \\
&\vdots \\
\underline{S}^n_{j1} &= \sum_{i=j}^{K} |h_{1i}| \mathbf{U}_{1i}\underline{X}^n_i + \underline{W}^n_{j1}. \qquad \text{(A.23)}
\end{aligned}$$

We know that $\left(\underline{Z}_j, \underline{W}_{j(j-1)}, \cdots, \underline{W}_{j1}\right)$ are jointly Gaussian random variables, with zero mean and covariance matrix $\mathbf{K}_{S_j}$ equal to:

$$\begin{pmatrix}
\mathbf{I} & \mathbf{A}_{j(j-1)} & \cdots & \mathbf{A}_{j2} & \mathbf{A}_{j1} \\
\mathbf{A}^T_{j(j-1)} & \mathbf{I} & \mathbf{A}_{(j-1)(j-2)} & \cdots & \mathbf{A}_{(j-1)1} \\
\vdots & & \ddots & \ddots & \vdots \\
\mathbf{A}^T_{j1} & & & \mathbf{I} & \mathbf{A}_{21} \\
\mathbf{A}^T_{j1} & \mathbf{A}^T_{(j-1)1} & \cdots & \mathbf{A}^T_{21} & \mathbf{I}
\end{pmatrix}$$

which is defined in Equation 5.21. It is clear that the covariance matrices of the jointly Gaussian random variables $\left(\underline{W}_{(j-1)p}, \underline{W}_{(j-1)(p-1)}, \cdots, \underline{W}_{(j-1)1}\right)$ and

$\left(\underline{W}_{jp}, \underline{W}_{j(p-1)}, \cdots, \underline{W}_{j1}\right)$ are the same:

$$\begin{aligned}
& \operatorname{cov}\left(\underline{W}_{(j-1)p}, \underline{W}_{(j-1)(p-1)}, \cdots, \underline{W}_{(j-1)1}\right) \\
= & \operatorname{cov}\left(\underline{W}_{jp}, \underline{W}_{j(p-1)}, \cdots, \underline{W}_{j1}\right) \\
= & \begin{pmatrix}
\mathbf{I} & \mathbf{A}_{p(p-1)} & \cdots & \mathbf{A}_{p2} & \mathbf{A}_{p1} \\
\mathbf{A}^T_{p(p-1)} & \mathbf{I} & \mathbf{A}_{(p-1)(p-2)} & \cdots & \mathbf{A}_{(p-1)1} \\
\vdots & & \ddots & \ddots & \vdots \\
\mathbf{A}^T_{p2} & & & \mathbf{I} & \mathbf{A}_{21} \\
\mathbf{A}^T_{p1} & \mathbf{A}^T_{(p-1)1} & \cdots & \mathbf{A}^T_{21} & \mathbf{I}
\end{pmatrix}
\end{aligned} \tag{A.24}$$

Comparing Equation A.18~Equation A.20 and Equation A.21~Equation A.23, we can see that distribution of $\underline{S}^n_{(j-1)p}\Big|\underline{S}^n_{(j-1)(p-1)}, \cdots, \underline{S}^n_{(j-1)1}, \underline{X}^n_{(j-1)}$ and $\underline{S}^n_{jp}\Big|\underline{S}^n_{j(p-1)}, \cdots, \underline{S}^n_{j1}$ are equal as long as $\underline{W}^n_{(j-1)p}\Big|\underline{W}^n_{(j-1)(p-1)}, \cdots, \underline{W}^n_{(j-1)1}$ and $\underline{W}^n_{jp}\Big|\underline{W}^n_{j(p-1)}, \cdots, \underline{W}^n_{j1}$ have the same distribution. Recall that $\underline{W}_{ji}$ is i.i.d. Gaussian random variables which is independent from the input signals $\underline{X}^n$. Therefore given Equation A.24, Lemma 41 is proved.

A.8 Proof of Lemma 42

Given Equation 5.17 and Equation 5.18, we have

$$\begin{aligned}
\underline{\hat{Y}}_{(j-1)} &= \left|h_{(j-1)(j-1)}\right| \mathbf{U}_{(j-1)(j-1)}\underline{X}_{(j-1)} + \sum_{i=j}^{K} \left|h_{(j-1)i}\right| \mathbf{U}_{(j-1)i}\underline{X}_i + \underline{N}_{(j-1)}. \\
\underline{S}^n_{(j-1)(j-2)} &= \left|h_{(j-1)j}\right| \mathbf{U}_{(j-2)(j-1)}\underline{X}^n_{(j-1)} + \sum_{i=j}^{K} \left|h_{(j-2)i}\right| \mathbf{U}_{(j-2)i}\underline{X}^n_i + \underline{W}^n_{(j-1)(j-2)} \\
&\vdots \\
\underline{S}^n_{(j-1)1} &= \left|h_{(j-1)j}\right| \mathbf{U}_{1(j-1)}\underline{X}^n_{(j-1)} + \sum_{i=j}^{K} \left|h_{1i}\right| \mathbf{U}_{1i}\underline{X}^n_i + \underline{W}^n_{(j-1)1}
\end{aligned}$$

When $\underline{X}^n_{(j-1)}$ is given, it can be subtracted from $\underline{S}^n_{(j-1)(j-2)}, \cdots, \underline{S}^n_{(j-1)1}$ to give

$$\begin{aligned}
\hat{\underline{S}}^n_{(j-1)(j-2)} &= \underline{S}^n_{(j-1)(j-2)} - \left|h_{(j-1)j}\right| \mathbf{U}_{(j-2)(j-1)}\underline{X}^n_{(j-1)} \\
&= \sum_{i=j}^{K} \left|h_{(j-2)i}\right| \mathbf{U}_{(j-2)i}\underline{X}^n_i + \underline{W}^n_{(j-1)(j-2)} \qquad \text{(A.25)} \\
&\vdots \\
\hat{\underline{S}}^n_{(j-1)1} &= \underline{S}^n_{(j-1)1} - \left|h_{(j-1)j}\right| \mathbf{U}_{1(j-1)}\underline{X}^n_{(j-1)} \\
&= \sum_{i=j}^{K} |h_{1i}| \mathbf{U}_{1i}\underline{X}^n_i + \underline{W}^n_{(j-1)1} \qquad \text{(A.26)}
\end{aligned}$$

while

$$\begin{aligned}
\underline{S}^n_{j(j-1)} &= \sum_{i=j}^{K} \left|h_{(j-1)i}\right| \mathbf{U}_{(j-1)i}\underline{X}^n_i + \underline{W}^n_{j(j-1)} \qquad \text{(A.27)} \\
&\vdots \\
\underline{S}^n_{j1} &= \sum_{i=j}^{K} |h_{1i}| \mathbf{U}_{1i}\underline{X}^n_i + \underline{W}^n_{j1}. \qquad \text{(A.28)}
\end{aligned}$$

We know that $\left(\underline{Z}_j, \underline{W}_{j(j-1)}, \cdots, \underline{W}_{j1}\right)$ are jointly Gaussian random variables, with zero mean and covariance matrix $\mathbf{K}_{S_j}$ equal to:

$$\begin{pmatrix}
\mathbf{I} & \mathbf{A}_{j(j-1)} & \cdots & \mathbf{A}_{j2} & \mathbf{A}_{j1} \\
\mathbf{A}^T_{j(j-1)} & \mathbf{I} & \mathbf{A}_{(j-1)(j-2)} & \cdots & \mathbf{A}_{(j-1)1} \\
\vdots & & \ddots & \ddots & \vdots \\
\mathbf{A}^T_{j1} & & & \mathbf{I} & \mathbf{A}_{21} \\
\mathbf{A}^T_{j1} & \mathbf{A}^T_{(j-1)1} & \cdots & \mathbf{A}^T_{21} & \mathbf{I}
\end{pmatrix}$$

which is defined in Equation 5.21. It is clear that the covariance matrices of jointly Gaussian random variables

$$\left(\underline{Z}_{(j-1)}, \underline{W}_{(j-1)(j-2)}, \cdots, \underline{W}_{(j-1)1}\right)$$

and

$$\left(\underline{W}_{j(j-1)}, \underline{W}_{j(j-2)}, \cdots, \underline{W}_{j1}\right)$$

are the same:

$$\begin{aligned}
& \operatorname{cov}\left(\underline{Z}_{(j-1)}, \underline{W}_{(j-1)(j-2)}, \cdots, \underline{W}_{(j-1)1}\right) \\
= & \operatorname{cov}\left(\underline{W}_{j(j-1)}, \underline{W}_{j(j-2)}, \cdots, \underline{W}_{j1}\right) \\
= & \begin{pmatrix}
\mathbf{I} & \mathbf{A}_{(j-1)(j-2)} & \cdots & \mathbf{A}_{(j-1)2} & \mathbf{A}_{(j-1)1} \\
\mathbf{A}^T_{(j-1)(j-2)} & \mathbf{I} & \mathbf{A}_{(p-1)(p-2)} & \cdots & \mathbf{A}_{(j-2)1} \\
\vdots & & \ddots & \ddots & \vdots \\
\mathbf{A}^T_{(j-1)2} & & & \mathbf{I} & \mathbf{A}_{21} \\
\mathbf{A}^T_{(j-1)1} & \mathbf{A}^T_{(j-2)1} & \cdots & \mathbf{A}^T_{21} & \mathbf{I}
\end{pmatrix}
\end{aligned} \tag{A.29}$$

Comparing Equation A.25~Equation A.26 and Equation A.27~Equation A.28, we can see that random variables

$$\underline{\hat{Y}}^n_{j-1}\Big|\, \underline{S}^n_{(j-1)(j-2)}, \cdots, \underline{S}^n_{(j-1)1}, \underline{X}^n_{(j-1)}$$

and

$$\underline{S}^n_{j(j-1)}\Big|\, \underline{S}^n_{j(j-2)}, \cdots, \underline{S}^n_{j1}$$

have the same marginal distribution as long as

$$\underline{Z}^n_{(j-1)}\Big|\, \underline{W}^n_{(j-1)(j-2)}, \cdots, \underline{W}^n_{(j-1)1}$$

and

$$\underline{W}^n_{j(j-1)}\Big|\, \underline{W}^n_{j(j-2)}, \cdots, \underline{W}^n_{j1}$$

have the same marginal distribution. Remember that $\underline{W}_{ji}$ is i.i.d. Gaussian random variables which is independent from the input signals $\underline{X}^n$. Therefore given Equation A.29, Lemma 42 is proved.

A.9 Proof of Lemma 43

Lemma 43 is proved using the following lemma from [46].

Let $\underline{X}$, $\underline{Y}$ and $\underline{Z}$ be jointly Gaussian vectors. If $\operatorname{cov}(\underline{Y})$ is invertible, then $\underline{X} \to \underline{Y} \to \underline{Z}$ forms a Markov chain if and only if

$$\operatorname{cov}(\underline{X}, \underline{Z}) = \operatorname{cov}(\underline{X}, \underline{Y}) \operatorname{cov}(\underline{Y})^{-1} \operatorname{cov}(\underline{Y}, \underline{Z})$$

Given Lemma section A.9 and the fact that $\operatorname{cov}\left(\underline{\hat{Y}}_{jG}\right)$ is invertible, $\underline{X}_{jG} \to \underline{\hat{Y}}_{jG} \to$

$\underline{\mathbf{S}}_j$ forms a Markov chain if and only if

$$\begin{aligned}
& \operatorname{cov}\left(\underline{X}_{jG}, \underline{\mathbf{S}}_j\right) \\
= \; & \operatorname{cov}\left(\underline{X}_{jG}, \underline{\hat{Y}}_{jG}\right) \operatorname{cov}\left(\underline{\hat{Y}}_{jG}\right)^{-1} \operatorname{cov}\left(\underline{\hat{Y}}_{jG}, \underline{\mathbf{S}}_j\right) \qquad \text{(A.30)}
\end{aligned}$$

Given Equation 5.17, Equation 5.18 and the independence of $\underline{W}_{jp}$ and $\underline{X}_i$, the left hand side of Equation A.30 is

$$LHS = \begin{pmatrix} |h_{1j}|^2 \mathbf{V}_j \mathbf{U}\left(-\phi_{1j}\right) \\ |h_{2j}|^2 \mathbf{V}_j \mathbf{U}\left(-\phi_{2j}\right) \\ \vdots \\ \left|h_{(j-1)j}\right|^2 \mathbf{V}_j \mathbf{U}\left(-\phi_{(j-1)j}\right) \end{pmatrix}^T$$

and the right hand side is

$$\begin{aligned}
RHS = \; & |h_{jj}|^2 \mathbf{V}_j \mathbf{U}\left(-\phi_{jj}\right) \left(\sum_{i=j}^{K} |h_{ji}|^2 \mathbf{U}\left(\phi_{ji}\right) \mathbf{V}_i \mathbf{U}\left(-\phi_{ji}\right) + \mathbf{I} \right)^{-1} \\
& \begin{pmatrix} \sum_{i=j}^{K} |h_{ji}|^2 |h_{1i}|^2 \mathbf{U}\left(\phi_{ji}\right) \mathbf{V}_i \mathbf{U}\left(-\phi_{1i}\right) + \mathbf{A}_{j1} \\ \sum_{i=j}^{K} |h_{ji}|^2 |h_{2i}|^2 \mathbf{U}\left(\phi_{ji}\right) \mathbf{V}_i \mathbf{U}\left(-\phi_{2i}\right) + \mathbf{A}_{j2} \\ \vdots \\ \sum_{i=j+1}^{K} |h_{ji}|^2 \left|h_{(j-1)i}\right|^2 \mathbf{U}\left(\phi_{ji}\right) \mathbf{V}_i \mathbf{U}\left(-\phi_{(j-1)i}\right) + \mathbf{A}_{j(j-1)} \end{pmatrix}^T
\end{aligned}$$

In order for $LHS = RHS$, we must have

$$\begin{aligned}
& |h_{pj}|^2 \mathbf{V}_j \mathbf{U}\left(-\phi_{pj}\right) \\
= \; & |h_{jj}|^2 \mathbf{V}_j \mathbf{U}\left(-\phi_{jj}\right) \left(\sum_{i=j}^{K} |h_{ji}|^2 \mathbf{U}\left(\phi_{ji}\right) \mathbf{V}_i \mathbf{U}\left(-\phi_{ji}\right) + \mathbf{I} \right)^{-1} \\
& \left(\sum_{i=j}^{K} |h_{ji}|^2 |h_{pi}|^2 \mathbf{U}\left(\phi_{ji}\right) \mathbf{V}_i \mathbf{U}\left(-\phi_{pi}\right) + \mathbf{A}_{jp} \right)
\end{aligned}$$

Solving the equation above, we have

$$\begin{aligned}\mathbf{A}_{jp} &= \frac{|h_{pj}|^2}{|h_{jj}|^2}\mathbf{U}(\phi_{jj}-\phi_{pj}) \\ &+\frac{|h_{pj}|^2}{|h_{jj}|^2}\sum_{i=j+1}^{K}|h_{ji}|^2\mathbf{U}(\phi_{ji})\mathbf{V}_i\mathbf{U}(\phi_{jj}-\phi_{pj}-\phi_{ji}) \\ &-\sum_{i=j+1}^{K}|h_{ji}|^2|h_{pi}|^2\mathbf{U}(\phi_{ji})\mathbf{V}_i\mathbf{U}(-\phi_{pi})\end{aligned}$$

A.10 Proof of Lemma 44

First, consider the simple case where $\frac{|h_{pj}|^2}{|h_{jj}|^2}=\alpha$, $\phi_{ji}=0$ and $P_j=0$, that is, $\mathbf{K}_{x_j}=\mathbf{0}$. For this case, given Equation 5.21 we have $\mathbf{A}_{ji}=\mathbf{B}=\begin{pmatrix}\alpha & 0\\ 0 & \alpha\end{pmatrix}$ for all i, j. It is easy to check that the eigenvalues of $\mathbf{K}_{S_j}=\begin{pmatrix}\mathbf{I} & \mathbf{B} & \cdots & \mathbf{B}\\ \mathbf{B}^T & \mathbf{I} & \cdots & \mathbf{B}\\ \vdots & & \ddots & \vdots\\ \mathbf{B}^T & & \cdots & \mathbf{I}\end{pmatrix}$ are $\lambda_1=1-\alpha$ and $\lambda_2=1+(j-1)\alpha$, with multiplicity $2(j-1)$ and 2 respectively. Therefore, $\mathbf{K}_{S_j}$ is positive definite if $0<\alpha<1$.

Now let us consider the case where ϕ_{ji} and P_j are small but non-zero, and $\frac{|h_{pj}|^2}{|h_{jj}|^2}$ are not necessarily equal to α. Denote the $(p,q)\,th$ element of $\mathbf{B}$ by b_{pq}. It is well known that the eigenvalues of symmetric matrix are locally (Lipschitz) continuous[47] with respect to its elements. Therefore, corresponding to every $\alpha\in(0,\,1)$, for any $\hat{\epsilon}>0$, there exist some strictly positive real numbers ϵ_α, ϵ'_α and $\epsilon''_\alpha(\underline{h})$ such that if $\left|\frac{|h_{pj}|^2}{|h_{jj}|^2}-\alpha\right|<\epsilon_\alpha$, $|\phi_{ji}|<\epsilon'_\alpha$, and $P_j<\epsilon''_\alpha(\underline{h})$ then every eigenvalues λ_s of $\mathbf{K}_{S_j}$ satisfies $|\lambda_s-\lambda_1|<\hat{\epsilon}$ or $|\lambda_s-\lambda_2|<\hat{\epsilon}$. The bound on P_j may depend $\underline{h}$ to ensure that the two last terms in Equation 5.21 are of bounded variation. For any $0<\alpha<1$ we can always find some $\hat{\epsilon}>0$ that guarantees $\lambda_s>0$, and $\mathbf{K}_{S_j}$ is positive definite as a result.

A.11 Proof of Theorem 45

First we state a useful result from [46].

([46, Lemma 2]) Let $\underline{X}^n=(\underline{X}_1,\cdots,\underline{X}_n)$ and $\underline{Y}^n=(\underline{Y}_1,\cdots,\underline{Y}_n)$ be two sequences of random vectors, and let $\underline{X}'_G$, $\underline{X}_G$, $\underline{Y}'_G$, and $\underline{Y}_G$ be Gaussian vectors with covariance matrices satisfying

$$\operatorname{cov}\begin{pmatrix} \underline{X}'_G \\ \underline{Y}'_G \end{pmatrix} = \frac{1}{n}\sum_{i=1}^{n} \operatorname{cov}\begin{pmatrix} \underline{X}_i \\ \underline{Y}_i \end{pmatrix} \preceq \operatorname{cov}\begin{pmatrix} \underline{X}_G \\ \underline{Y}_G \end{pmatrix}$$

then we have

$$\begin{aligned} h\left(\underline{X}^n\right) &\leq nh\left(\underline{X}'_G\right) &&\leq nh\left(\underline{X}_G\right) \\ h\left(\underline{Y}^n \middle| \underline{X}^n\right) &\leq nh\left(\underline{Y}'_G \middle| \underline{X}'_G\right) &&\leq nh\left(\underline{Y}_G \middle| \underline{X}_G\right) \end{aligned}$$

By Fano's inequality, the sum capacity of the generalized Z-channel Equation 5.17

must satisfy

$$
\begin{aligned}
& n\sum_{j=1}^{K} R_j - n\epsilon \\
& \overset{(a)}{\leq} \;\; I\left(\underline{X}_1^n;\, \underline{\hat{Y}}_1^n\right) + \sum_{j=2}^{K} I\left(\underline{X}_j^n;\, \underline{\hat{Y}}_j^n,\, \underline{\mathbf{S}}_j\right) \\
& \overset{(b)}{=} \;\; h\left(\underline{\hat{Y}}_1^n\right) - h\left(\underline{\hat{Y}}_1^n\middle|\,\underline{X}_1^n\right) \\
& \qquad + \sum_{j=2}^{K} I\left(\underline{X}_j^n;\, \underline{\mathbf{S}}_j\right) + \sum_{j=2}^{K} I\left(\underline{X}_j^n;\, \underline{\hat{Y}}_j^n\middle|\,\underline{\mathbf{S}}_j\right) \\
& \overset{(c)}{=} \;\; h\left(\underline{\hat{Y}}_1^n\right) - h\left(\underline{\hat{Y}}_1^n\middle|\,\underline{X}_1^n\right) \\
& \qquad + \sum_{j=2}^{K}\sum_{p=1}^{j-1} I\left(\underline{X}_j^n;\, \underline{S}_{jp}\middle|\,\underline{S}_{j(p-1)}^n, \cdots, \underline{S}_{j1}^n\right) \\
& \qquad + \sum_{j=2}^{K} I\left(\underline{X}_j^n;\, \underline{\hat{Y}}_j^n\middle|\,\underline{\mathbf{S}}_j\right) \\
& \overset{(d)}{=} \;\; h\left(\underline{\hat{Y}}_1^n\right) - h\left(\underline{\hat{Y}}_1^n\middle|\,\underline{X}_1^n\right) \\
& \qquad + \sum_{j=2}^{K}\sum_{p=1}^{j-1} \Bigg(h\left(\underline{S}_{jp}^n\middle|\,\underline{S}_{j(p-1)}^n, \cdots, \underline{S}_{j1}^n\right) \\
& \qquad - h\left(\underline{S}_{jp}^n\middle|\,\underline{S}_{j(p-1)}^n, \cdots, \underline{S}_{j1}^n, \underline{X}_j^n\right)\Bigg) \\
& \qquad + \sum_{j=2}^{K} \Bigg(h\left(\underline{\hat{Y}}_j^n\middle|\,\underline{S}_{j(j-1)}^n, \underline{S}_{j(j-2)}^n, \cdots, \underline{S}_{j1}^n\right) \\
& \qquad - h\left(\underline{\hat{Y}}_j^n\middle|\,\underline{S}_{j(j-1)}^n, \underline{S}_{j(j-2)}^n, \cdots, \underline{S}_{j1}^n, \underline{X}_j^n\right)\Bigg) \\
& \overset{(e)}{\leq} \;\; nh\left(\underline{\hat{Y}}_{1G}\right) - h\left(\underline{\hat{Y}}_1^n\middle|\,\underline{X}_1^n\right) \\
& \qquad + h\left(\underline{S}_{21}^n\right) - h\left(\underline{S}_{21}^n\middle|\,\underline{X}_2^n\right) \\
& \qquad + \sum_{j=3}^{K}\sum_{p=j-1}^{j-1} h\left(\underline{S}_{jp}^n\middle|\,\underline{S}_{j(p-1)}^n, \cdots, \underline{S}_{j1}^n\right) \\
& \qquad + \sum_{j=3}^{K}\sum_{p=1}^{j-2} h\left(\underline{S}_{jp}^n\middle|\,\underline{S}_{j(p-1)}^n, \cdots, \underline{S}_{j1}^n\right) \\
& \qquad - \sum_{j=3}^{K-1}\sum_{p=1}^{j-1} h\left(\underline{S}_{jp}^n\middle|\,\underline{S}_{j(p-1)}^n, \cdots, \underline{S}_{j1}^n, \underline{X}_j^n\right) \\
& \qquad - \sum_{j=K}^{K}\sum_{p=1}^{j-1} h\left(\underline{S}_{jp}^n\middle|\,\underline{S}_{j(p-1)}^n, \cdots, \underline{S}_{j1}^n, \underline{X}_j^n\right) \\
& \qquad + \sum_{j=2}^{K} h\left(\underline{\hat{Y}}_j^n\middle|\,\underline{S}_{j(j-1)}^n, \underline{S}_{j(j-2)}^n, \cdots, \underline{S}_{j1}^n\right) \\
& \qquad - \sum_{j=2}^{K} h\left(\underline{\hat{Y}}_j^n\middle|\,\underline{S}_{j(j-1)}^n, \underline{S}_{j(j-2)}^n, \cdots, \underline{S}_{j1}^n, \underline{X}_j^n\right)
\end{aligned}
$$

$$
\begin{aligned}
\overset{(f)}{=}\; & nh\left(\underline{\hat{Y}}_{1G}\right) - h\left(\underline{\hat{Y}}_1^n \middle| \underline{X}_1^n\right) \\
& + h\left(\underline{S}_{21}^n\right) + \sum_{j=3}^{K} h\left(\underline{S}_{j(j-1)}^n \middle| \underline{S}_{j(j-2)}^n, \cdots, \underline{S}_{j1}^n\right) \\
& + \sum_{j=3}^{K}\sum_{p=1}^{j-2}\Bigg(h\left(\underline{S}_{jp}^n \middle| \underline{S}_{j(p-1)}^n, \cdots, \underline{S}_{j1}^n\right) \\
& - h\left(\underline{S}_{(j-1)p}^n \middle| \underline{S}_{(j-1)(p-1)}^n, \cdots, \underline{S}_{(j-1)1}^n, \underline{X}_{(j-1)}^n\right)\Bigg) \\
& - \sum_{p=1}^{K-1} h\left(\underline{S}_{Kp}^n \middle| \underline{S}_{K(p-1)}^n, \cdots, \underline{S}_{K1}^n, \underline{X}_K^n\right) \\
& + \sum_{j=2}^{K} h\left(\underline{\hat{Y}}_j^n \middle| \underline{S}_{j(j-1)}^n, \underline{S}_{j(j-2)}^n, \cdots, \underline{S}_{j1}^n\right) \\
& - \sum_{j=2}^{K} h\left(\underline{\hat{Y}}_j^n \middle| \underline{S}_{j(j-1)}^n, \underline{S}_{j(j-2)}^n, \cdots, \underline{S}_{j1}^n, \underline{X}_j^n\right) \\
\overset{(g)}{=}\; & nh\left(\underline{\hat{Y}}_{1G}\right) - h\left(\underline{\hat{Y}}_1^n \middle| \underline{X}_1^n\right) \\
& + h\left(\underline{S}_{21}^n\right) + \sum_{j=3}^{K} h\left(\underline{S}_{j(j-1)}^n \middle| \underline{S}_{j(j-2)}^n, \cdots, \underline{S}_{j1}^n\right) \\
& - \sum_{p=1}^{K-1} h\left(\underline{S}_{Kp}^n \middle| \underline{S}_{K(p-1)}^n, \cdots, \underline{S}_{K1}^n, \underline{X}_K^n\right) \\
& + \sum_{j=2}^{K} h\left(\underline{\hat{Y}}_j^n \middle| \underline{S}_{j(j-1)}^n, \underline{S}_{j(j-2)}^n, \cdots, \underline{S}_{j1}^n\right) \\
& - \sum_{j=2}^{K} h\left(\underline{\hat{Y}}_j^n \middle| \underline{S}_{j(j-1)}^n, \underline{S}_{j(j-2)}^n, \cdots, \underline{S}_{j1}^n, \underline{X}_j^n\right) \\
\overset{(h)}{=}\; & nh\left(\underline{\hat{Y}}_{1G}\right) + \sum_{j=3}^{K} h\left(\underline{S}_{j(j-1)}^n \middle| \underline{S}_{j(j-2)}^n, \cdots, \underline{S}_{j1}^n\right) \\
& - nh\left(\underline{W}_{K(K-1)}, \cdots, \underline{W}_{K1}\right) \\
& + \sum_{j=2}^{K} h\left(\underline{\hat{Y}}_j^n \middle| \underline{S}_{j(j-1)}^n, \underline{S}_{j(j-2)}^n, \cdots, \underline{S}_{j1}^n\right) \\
& - \sum_{j=2}^{K} h\left(\underline{\hat{Y}}_j^n \middle| \underline{S}_{j(j-1)}^n, \underline{S}_{j(j-2)}^n, \cdots, \underline{S}_{j1}^n, \underline{X}_j^n\right)
\end{aligned}
$$

$$
\begin{aligned}
&\overset{(i)}{=} nh\left(\underline{\hat{Y}}_{1G}\right) - nh\left(\underline{W}_{K(K-1)}, \cdots, \underline{W}_{K1}\right) \\
&\quad + \sum_{j=2}^{K} h\left(\underline{\hat{Y}}_j^n \middle| \underline{S}_{j(j-1)}^n, \underline{S}_{j(j-2)}^n, \cdots, \underline{S}_{j1}^n\right) \\
&\quad - nh\left(\underline{N}_K \middle| \underline{W}_{K(K-1)}, \cdots, \underline{W}_{K1}\right) \\
&\overset{(j)}{=} nh\left(\underline{\hat{Y}}_{1G}\right) - nh\left(\underline{N}_K, \underline{W}_{K(K-1)}, \cdots, \underline{W}_{K1}\right) \\
&\quad + \sum_{j=2}^{K} h\left(\underline{\hat{Y}}_j^n \middle| \underline{S}_{j(j-1)}^n, \underline{S}_{j(j-2)}^n, \cdots, \underline{S}_{j1}^n\right) \\
&\overset{(k)}{\leq} nh\left(\underline{\hat{Y}}_{1G}\right) - nh\left(\underline{N}_K, \underline{W}_{K(K-1)}, \cdots, \underline{W}_{K1}\right) \\
&\quad + n\sum_{j=2}^{K} h\left(\underline{\hat{Y}}_{jG} \middle| \underline{S}_{j(j-1)G}, \underline{S}_{j(j-2)G}, \cdots, \underline{S}_{j1G}\right) \\
&\overset{(l)}{=} nh\left(\underline{\hat{Y}}_{1G}\right) - nh\left(\underline{N}_K, \underline{W}_{K(K-1)}, \cdots, \underline{W}_{K1}\right) \\
&\quad + n\sum_{j=2}^{K} h\left(\underline{\hat{Y}}_{jG}\right) \\
&\quad - n\sum_{j=2}^{K} h\left(\underline{S}_{j(j-1)G}, \underline{S}_{j(j-2)G}, \cdots, \underline{S}_{j1G}\right) \\
&\quad + n\sum_{j=2}^{K} h\left(\underline{S}_{j(j-1)G}, \underline{S}_{j(j-2)G}, \cdots, \underline{S}_{j1G} \middle| \underline{\hat{Y}}_{jG}\right)
\end{aligned}
$$

$$
\begin{aligned}
&\stackrel{(m)}{=} nh\left(\underline{\hat{Y}}_{1G}\right) - nh\left(\underline{N}_K,\, \underline{W}_{K(K-1)}, \cdots, \underline{W}_{K1}\right) \\
&\quad +n\sum_{j=2}^{K} h\left(\underline{\hat{Y}}_{jG}\right) \\
&\quad -n\sum_{j=2}^{K} h\left(\underline{S}_{j(j-1)G}, \underline{S}_{j(j-2)G}, \cdots, \underline{S}_{j1G}\right) \\
&\quad +n\sum_{j=2}^{K-1} h\left(\underline{S}_{j(j-1)G}, \underline{S}_{j(j-2)G}, \cdots, \underline{S}_{j1G}\middle| \underline{\hat{Y}}_{jG}\right) \\
&\quad +nh\left(\underline{S}_{K(K-1)G}, \underline{S}_{K(K-2)G}, \cdots, \underline{S}_{K1G}\middle| \underline{\hat{Y}}_{KG}\right) \\
&\stackrel{(n)}{=} nh\left(\underline{\hat{Y}}_{1G}\right) - nh\left(\underline{N}_K,\, \underline{W}_{K(K-1)}, \cdots, \underline{W}_{K1}\right) \\
&\quad +n\sum_{j=2}^{K} h\left(\underline{\hat{Y}}_{jG}\right) \\
&\quad -n\sum_{j=2}^{K} h\left(\underline{S}_{j(j-1)G}, \underline{S}_{j(j-2)G}, \cdots, \underline{S}_{j1G}\right) \\
&\quad +n\sum_{j=2}^{K-1} h\left(\underline{S}_{(j+1)(j-1)G}, \underline{S}_{(j+1)(j-2)G}, \cdots, \underline{S}_{(j+1)1G}\middle| \underline{S}_{(j+1)jG}\right) \\
&\quad +nh\left(\underline{W}_{K(K-1)}, \underline{W}_{K(K-2)}, \cdots, \underline{W}_{K1}\middle| \underline{N}_K\right) \\
&= nh\left(\underline{\hat{Y}}_{1G}\right) - nh\left(\underline{N}_K,\, \underline{W}_{K(K-1)}, \cdots, \underline{W}_{K1}\right) \\
&\quad +n\sum_{j=2}^{K} h\left(\underline{\hat{Y}}_{jG}\right) \\
&\quad -n\sum_{j=2}^{K} h\left(\underline{S}_{j(j-1)G}, \underline{S}_{j(j-2)G}, \cdots, \underline{S}_{j1G}\right) \\
&\quad +n\sum_{j=3}^{K} h\left(\underline{S}_{j(j-2)G}, \underline{S}_{j(j-3)G}, \cdots, \underline{S}_{j1G}\middle| \underline{S}_{j(j-1)G}\right) \\
&\quad +nh\left(\underline{W}_{K(K-1)}, \underline{W}_{K(K-2)}, \cdots, \underline{W}_{K1}\middle| \underline{N}_K\right) \\
&= nh\left(\underline{\hat{Y}}_{1G}\right) - nh\left(\underline{N}_K,\, \underline{W}_{K(K-1)}, \cdots, \underline{W}_{K1}\right) \\
&\quad +n\sum_{j=2}^{K} h\left(\underline{\hat{Y}}_{jG}\right) - h\left(\underline{S}_{21G}\right) \\
&\quad -n\sum_{j=3}^{K} \Big(h\left(\underline{S}_{j(j-1)G}, \underline{S}_{j(j-2)G}, \cdots, \underline{S}_{j1G}\right) \\
&\quad - h\left(\underline{S}_{j(j-2)G}, \underline{S}_{j(j-3)G}, \cdots, \underline{S}_{j1G}\middle| \underline{S}_{j(j-1)G}\right)\Big) \\
&\quad +nh\left(\underline{W}_{K(K-1)}, \underline{W}_{K(K-2)}, \cdots, \underline{W}_{K1}\middle| \underline{N}_K\right)
\end{aligned}
$$

$$
\begin{aligned}
= \; & nh\left(\underline{\hat{Y}}_{1G}\right) - nh\left(\underline{N}_K, \underline{W}_{K(K-1)}, \cdots, \underline{W}_{K1}\right) \\
& +n\sum_{j=2}^{K} h\left(\underline{\hat{Y}}_{jG}\right) - h\left(\underline{S}_{21G}\right) \\
& -n\sum_{j=3}^{K} h\left(\underline{S}_{j(j-1)G}\right) \\
& +nh\left(\underline{W}_{K(K-1)}, \underline{W}_{K(K-2)}, \cdots, \underline{W}_{K1}\middle| \underline{N}_K\right) \\
= \; & nh\left(\underline{\hat{Y}}_{1G}\right) - nh\left(\underline{N}_K\right) \\
& +n\sum_{j=2}^{K} h\left(\underline{\hat{Y}}_{jG}\right) - h\left(\underline{S}_{21G}\right) \\
& -n\sum_{j=3}^{K} h\left(\underline{S}_{j(j-1)G}\right) \\
= \; & n\sum_{j=1}^{K-1}\left(h\left(\underline{\hat{Y}}_{jG}\right) - h\left(\underline{S}_{(j+1)jG}\right)\right) \\
& +nh\left(\underline{\hat{Y}}_{KG}\right) - nh\left(\underline{N}_K\right) \\
= \; & n\sum_{j=1}^{K} I\left(\underline{X}_{jG}; \underline{\hat{Y}}_{jG}\right)
\end{aligned}
$$

(a) is from Fano's inequality.

(b) is from the expansion of mutual information: $I\left(\underline{X}_1^n; \underline{\hat{Y}}_1^n\right) = h\left(\underline{\hat{Y}}_1^n\right) - h\left(\underline{\hat{Y}}_1^n \middle| \underline{X}_1^n\right)$, and the chain rule which gives $I\left(\underline{X}_j^n; \underline{\hat{Y}}_j^n, \mathbf{S}_j\right) = I\left(\underline{X}_j^n; \mathbf{S}_j\right) + I\left(\underline{X}_j^n; \underline{\hat{Y}}_j^n \middle| \mathbf{S}_j\right)$..

(c) is from the chain rule, which gives $I\left(\underline{X}_j^n; \mathbf{S}_j\right) = \sum_{p=1}^{j-1} I\left(\underline{X}_j^n; \underline{S}_{jp}\middle| \underline{S}_{j(p-1)}^n, \cdots, \underline{S}_{j1}^n\right)$.

(d) is from the expansion of mutual information.

(e) is from the inequality $h\left(\underline{\hat{Y}}_1^n\right) \le nh\left(\underline{\hat{Y}}_{1G}\right)$. It holds because Gaussian random variable maximize entropy under given power constraint, and line 2 to line 6 in (e) is equivalent to line 2 and line 3 in (d).

(f) is from the following equation:

$$
\begin{aligned}
& -h\left(\underline{S}_{21}^n \middle| \underline{X}_2^n\right) - \sum_{j=3}^{K-1}\sum_{p=1}^{j-1} h\left(\underline{S}_{jp}^n \middle| \underline{S}_{j(p-1)}^n, \cdots, \underline{S}_{j1}^n, \underline{X}_j^n\right) \\
= \; & -\sum_{j=3}^{K}\sum_{p=1}^{j-1} h\left(\underline{S}_{(j-1)p}^n \middle| \underline{S}_{(j-1)(p-1)}^n, \cdots, \underline{S}_{(j-1)1}^n, \underline{X}_{(j-1)}^n\right). \qquad \text{(A.31)}
\end{aligned}
$$

(g) is from Lemma 41. Because random variables

$$
\underline{S}_{(j-1)p}^n \middle| \underline{S}_{(j-1)(p-1)}^n, \cdots, \underline{S}_{(j-1)1}^n, \underline{X}_{(j-1)}^n
$$

and

$$\underline{S}_{jp}^{n} \Big| \underline{S}_{j(p-1)}^{n}, \cdots, \underline{S}_{j1}^{n}$$

have the same marginal distribution,

$$h\left(\underline{S}_{(j-1)p}^{n} \Big| \underline{S}_{(j-1)(p-1)}^{n}, \cdots, \underline{S}_{(j-1)1}^{n}, \underline{X}_{(j-1)}^{n}\right)$$

and

$$h\left(\underline{S}_{jp}^{n} \Big| \underline{S}_{j(p-1)}^{n}, \cdots, \underline{S}_{j1}^{n}\right)$$

are equal, which gives

$$\begin{aligned} &\textstyle\sum_{j=3}^{K} \sum_{p=1}^{j-2} \Big(h\left(\underline{S}_{jp}^{n} \Big| \underline{S}_{j(p-1)}^{n}, \cdots, \underline{S}_{j1}^{n}\right) \\ &-h\left(\underline{S}_{(j-1)p}^{n} \Big| \underline{S}_{(j-1)(p-1)}^{n}, \cdots, \underline{S}_{(j-1)1}^{n}, \underline{X}_{(j-1)}^{n}\right)\Big) \quad = 0. \end{aligned}$$

(h) Given $\underline{S}_{Kp} = |C_{pK}| \mathbf{U}_{pK} \underline{X}_K + \underline{W}_{Kp}$, the summation in the third line after (g) gives

$$\begin{aligned} & \sum_{p=1}^{K-1} h\left(\underline{S}_{Kp}^{n} \Big| \underline{S}_{K(p-1)}^{n}, \cdots, \underline{S}_{K1}^{n}, \underline{X}_{K}^{n}\right) \\ = & \sum_{p=1}^{K-1} h\left(\underline{W}_{Kp}^{n} \Big| \underline{W}_{K(p-1)}^{n}, \cdots, \underline{W}_{K1}^{n}\right) && \text{(A.32)} \\ = & \; h\left(\underline{W}_{K(K-1)}^{n}, \cdots, \underline{W}_{K1}^{n}\right) && \text{(A.33)} \\ = & \; nh\left(\underline{W}_{K(K-1)}, \cdots, \underline{W}_{K1}\right) && \text{(A.34)} \end{aligned}$$

It is also easy to see that $\underline{S}_{21}^{n}$ and $\underline{\hat{Y}}_{1}^{n} \Big| \underline{X}_{1}^{n}$ have same marginal distribution, therefore

$$h\left(\underline{S}_{21}^{n}\right) - h\left(\underline{\hat{Y}}_{1}^{n} \Big| \underline{X}_{1}^{n}\right) \quad = \quad 0 \qquad \text{(A.35)}$$

(i) Now combine the second and the last terms after (h):

$$\begin{aligned}
& \sum_{j=3}^{K} h\left(\underline{S}_{j(j-1)}^{n}\Big|\,\underline{S}_{j(j-2)}^{n},\cdots,\underline{S}_{j1}^{n}\right) \\
& -\sum_{j=2}^{K} h\left(\underline{\hat{Y}}_{j}^{n}\Big|\,\underline{S}_{j(j-1)}^{n},\underline{S}_{j(j-2)}^{n},\cdots,\underline{S}_{j1}^{n},\underline{X}_{j}^{n}\right) && \text{(A.36)}\\
= \; & \sum_{j=3}^{K} h\left(\underline{S}_{j(j-1)}^{n}\Big|\,\underline{S}_{j(j-2)}^{n},\cdots,\underline{S}_{j1}^{n}\right) \\
& -\sum_{j=3}^{K} h\left(\underline{\hat{Y}}_{(j-1)}^{n}\Big|\,\underline{S}_{(j-1)(j-2)}^{n},\cdots,\underline{S}_{(j-1)1}^{n},\underline{X}_{(j-1)}^{n}\right) && \text{(A.37)}\\
& -h\left(\underline{\hat{Y}}_{K}^{n}\Big|\,\underline{S}_{K(K-1)}^{n},\cdots,\underline{S}_{K1}^{n},\underline{X}_{K}^{n}\right) && \text{(A.38)}\\
\overset{(h-1)}{=} \; & -h\left(\underline{\hat{Y}}_{K}^{n}\Big|\,\underline{S}_{K(K-1)}^{n},\cdots,\underline{S}_{K1}^{n},\underline{X}_{K}^{n}\right) && \text{(A.39)}\\
= \; & nh\left(\underline{N}_{K}|\,\underline{W}_{K(K-1)},\cdots,\underline{W}_{K1}\right) && \text{(A.40)}
\end{aligned}$$

(h-1) is from Lemma 42. Given that random variables

$$\underline{\hat{Y}}_{j-1}^{n}\Big|\,\underline{S}_{(j-1)(j-2)}^{n},\cdots,\underline{S}_{(j-1)1}^{n},\underline{X}_{(j-1)}^{n}$$

and

$$\underline{S}_{j(j-1)}^{n}\Big|\,\underline{S}_{j(j-2)}^{n},\cdots,\underline{S}_{j1}^{n}$$

have the same marginal distribution, we have

$$h\left(\underline{S}_{j(j-1)}^{n}\Big|\,\underline{S}_{j(j-2)}^{n},\cdots,\underline{S}_{j1}^{n}\right)=h\left(\underline{\hat{Y}}_{(j-1)}^{n}\Big|\,\underline{S}_{(j-1)(j-2)}^{n},\cdots,\underline{S}_{(j-1)1}^{n},\underline{X}_{(j-1)}^{n}\right),$$

(j) From chain rule of entropy, we know that

$$\begin{aligned}
& h\left(\underline{W}_{K(K-1)},\cdots,\underline{W}_{K1}\right)+h\left(\underline{N}_{K}|\,\underline{W}_{K(K-1)},\cdots,\underline{W}_{K1}\right) \\
= \; & h\left(\underline{N}_{K},\underline{W}_{K(K-1)},\cdots,\underline{W}_{K1}\right). && \text{(A.41)}
\end{aligned}$$

(k) is from Lemma section A.11.

(l) is from the formula $h\left(X|\,Y\right)=h\left(X\right)+h\left(Y|\,X\right)-h\left(Y\right)$.

(m) is from

$$\begin{aligned}
& \sum_{j=2}^{K} h\left(\underline{S}_{j(j-1)G}, \underline{S}_{j(j-2)G}, \cdots, \underline{S}_{j1G} \middle| \underline{\hat{Y}}_{jG}\right) \\
= \; & + \sum_{j=2}^{K-1} h\left(\underline{S}_{j(j-1)G}, \underline{S}_{j(j-2)G}, \cdots, \underline{S}_{j1G} \middle| \underline{\hat{Y}}_{jG}\right) \\
& + h\left(\underline{S}_{K(K-1)G}, \underline{S}_{K(K-2)G}, \cdots, \underline{S}_{K1G} \middle| \underline{\hat{Y}}_{KG}\right)
\end{aligned}$$

(n) Combining Lemma 43 and Lemma 44, we know that for channels in $\mathcal{C}_\alpha$, if the power constraint P_j satisfies $P_j \leq \epsilon''_\alpha$, then

$$\underline{X}_{jG} \rightarrow \underline{\hat{Y}}_{jG} \rightarrow \underline{\mathbf{S}}_j \tag{A.42}$$

form a Markov chain, and the following equality holds:

$$\begin{aligned}
& h\left(\underline{S}_{j(j-1)G}, \underline{S}_{j(j-2)G}, \cdots, \underline{S}_{j1G} \middle| \underline{\hat{Y}}_{jG}\right) \\
= \; & h\left(\underline{S}_{j(j-1)G}, \underline{S}_{j(j-2)G}, \cdots, \underline{S}_{j1G} \middle| \underline{\hat{Y}}_{jG}, \underline{X}_{jG}\right) \\
= \; & h\left(\underline{S}_{(j+1)(j-1)G}, \underline{S}_{(j+1)(j-2)G}, \cdots, \underline{S}_{(j+1)1G} \middle| \underline{S}_{(j+1)jG}\right)
\end{aligned} \tag{A.43}$$

We can conclude that the achievable sum capacity of the generalized Z-channel must satisfy

$$\sum_{j=1}^{K} R_j \tag{A.44}$$

$$\leq \max_{\substack{\mathrm{Tr}\left(\mathbf{V}_j\right) \leq \mathrm{SNR}_j \\ \mathbf{V}_j \succeq \mathbf{0},\; j=1,\cdots,K}} \sum_{j=1}^{K} I\left(\underline{X}_{jG}; \underline{\hat{Y}}_{jG}\right) \tag{A.45}$$

$$= \max_{\substack{\mathrm{Tr}\left(\mathbf{V}_j\right) \leq \mathrm{SNR}_j \\ \mathbf{V}_j \succeq \mathbf{0},\; j=1,\cdots,K}} \sum_{j=1}^{K} \log \left| \left(\mathbf{I} + \sum_{i=j}^{K} \left|C_{ji}\right|^2 \mathbf{V}_i\right) \left(\mathbf{I} + \sum_{i=j+1}^{K} \left|C_{ji}\right|^2 \mathbf{V}_i\right)^{-1} \right| \tag{A.46}$$

Notice that for Z-channel, this sum capacity outer bound is achievable because the expression above is identical to the sum capacity achieved by treating interference as noise. Since the generalized Z-channel is obtained by eliminating some of the interference links from the interference channel, Equation A.44 is an outer bound for the sum capacity of the interference channel. Theorem 45 is proved.

A.12 Proof of Theorem 46

In Theorem 45, we have proved that the sum capacity Equation 5.25 of the generalized Z-channel is achieved by i.i.d. Gaussian input,

$$\begin{aligned}
\sum_{j=1}^{K} R_j &\leq \max_{\substack{\mathrm{Tr}(\mathbf{V}_j) \leq P_j \\ \mathbf{V}_j \succeq \mathbf{0},\, j=1,\cdots,K}} \sum_{j=1}^{K} I\left(\underline{X}_{jG}; \underline{\hat{Y}}_{jG}\right) && \text{(A.47)} \\
&= \max_{\substack{\mathrm{Tr}(\mathbf{V}_j) \leq P_j \\ \mathbf{V}_j \succeq \mathbf{0},\, j=1,\cdots,K}} \sum_{j=1}^{K} \log \left| \left(\mathbf{I} + \sum_{i=j}^{K} |h_{ji}|^2 \mathbf{V}_i\right) \left(\mathbf{I} + \sum_{i=j+1}^{K} |h_{ji}|^2 \mathbf{V}_i\right)^{-1} \right| && \text{(A.48)}
\end{aligned}$$

Define the normalized covariance matrix $\hat{\mathbf{V}}_j = \frac{\mathbf{V}_j}{P_j}$, $\mathrm{Tr}\left(\hat{\mathbf{V}}_j\right) = 1$. Consider the equal power constraint where $P_j = {}^{P_{sum}}/_{K}$ for all users.

For an expression of the form $\log |\mathbf{I} + x\mathbf{A}|$, let the eigenvalue of matrix $\mathbf{A}$ be $0 \leq \lambda_i (\mathbf{A}) < \infty$. Then

$$\begin{aligned}
\log |\mathbf{I} + x\mathbf{A}| &= \sum_{i=1}^{n} \log (1 + x\lambda_i (\mathbf{A})) \\
&= \sum_{i=1}^{n} \left(x\lambda_i (\mathbf{A}) - \frac{1}{2} x^2 \lambda_i^2 (\mathbf{A}) + o\left(x^2\right) \right) \\
&= x\mathrm{Tr} (\mathbf{A}) - \frac{1}{2} x^2 \mathrm{Tr} \left(\mathbf{A}^2\right) + o\left(x^2\right) && \text{(A.49)}
\end{aligned}$$

The second equation uses Taylor's theorem for several variables at $\hat{\lambda}_i (\mathbf{A}) = x\lambda_i (\mathbf{A})$, since when $x \to 0$, $x\lambda_i (\mathbf{A}) \to 0$ as well.

Combining Equation A.49, Equation 1.11, Equation 1.16 and Equation 5.25, we find Equation 5.29 and Equation 5.30.

A.13 Proof of Theorem 47

To maximize the right hand side of Equation 5.30, we need to solve the following optimization problem

$$\begin{aligned}
\min_{\hat{\mathbf{V}}_1,\cdots,\hat{\mathbf{V}}_K} \quad & \sum_{j=1}^{K} |h_{jj}|^4 \operatorname{Tr}\left(\hat{\mathbf{V}}_j^2\right) \\
& +2\sum_{j=1}^{K-1}\sum_{i=j+1}^{K} |h_{jj}|^2 |h_{ji}|^2 \operatorname{Tr}\left(\hat{\mathbf{V}}_j \mathbf{U}_{ji} \hat{\mathbf{V}}_i \mathbf{U}_{ji}^{\dagger}\right) && \text{(A.50)} \\
s.t. \quad & \operatorname{Tr}\left(\hat{\mathbf{V}}_j\right) = 1 \\
& \hat{\mathbf{V}}_j \succeq \mathbf{0}. && \text{(A.51)}
\end{aligned}$$

First, consider a simple case where the channel is strictly symmetric: $\phi_{ji} = 0$, $|h_{jj}|^2 = 1$ and $|h_{ji}|^2 = \alpha < 1$ for all i, j. Equation A.50 becomes

$$\begin{aligned}
\min_{\hat{\mathbf{V}}_1,\cdots,\hat{\mathbf{V}}_K} \quad & \sum_{j=1}^{K} \operatorname{Tr}\left(\hat{\mathbf{V}}_j^2\right) + 2\alpha \sum_{j=1}^{K-1}\sum_{i=j+1}^{K} \operatorname{Tr}\left(\hat{\mathbf{V}}_j \hat{\mathbf{V}}_i\right) && \text{(A.52)} \\
s.t. \quad & \operatorname{Tr}\left(\hat{\mathbf{V}}_j\right) = 1 \\
& \hat{\mathbf{V}}_j \succeq \mathbf{0}. && \text{(A.53)}
\end{aligned}$$

Let the 2×2 real positive definite matrix $\hat{\mathbf{V}}_j$ be

$$\hat{\mathbf{V}}_j = \begin{pmatrix} k_{j1} & k_{j3} \\ k_{j3} & k_{j2} \end{pmatrix}. \tag{A.54}$$

Substituting Equation A.54 into Equation A.52, we construct a non-linear optimization problem from Equation A.52 on standard form:

$$\begin{aligned}
\min_{k_{11},k_{12},k_{13},\cdots,k_{K1},k_{K2},k_{K3}} \quad & \sum_{j=1}^{K} \left(k_{j1}^2 + k_{j2}^2 + 2k_{j3}^2\right) && \text{(A.55)} \\
& +2\alpha \sum_{j=1}^{K-1}\sum_{i=j+1}^{K} \left(k_{j1}k_{i1} + k_{j2}k_{i2} + 2k_{j3}k_{i3}\right) \\
s.t. \quad & -k_{j1} \le 0 && \text{(A.56)} \\
& -k_{j2} \le 0 && \text{(A.57)} \\
& k_{j3}^2 - k_{j1}k_{j2} \le 0 && \text{(A.58)} \\
& k_{j1} + k_{j2} = 1 && \text{(A.59)} \\
& for\, all\, j = 1, \cdots, K
\end{aligned}$$

The optimal solution of the problem defined by Equation A.55~Equation A.59 is also the optimal solution of the problem defined by Equation A.52. Denote the optimization problem defined by Equation A.55~Equation A.59 as $\left(P_{\underline{k}}\right)$, where $\underline{k} = (k_{11}, k_{12}, k_{13}, \cdots, k_{K1}, k_{K2}, k_{K3})$ represents the set of feasible solutions. Notice that while any positive k_{j1}, k_{j2} with $k_{j1} + k_{j2} \leq 1$ satisfies the power constraint, we require constraint Equation A.59 to be an equality. Because only when it is satisfied with equality, the system can achieve correct $\left.\frac{E_b}{N_0}\right|_{\min_0}$.

Denote the objective function in Equation A.55 by $f(\underline{k})$. Construct the Lagrangian function for problem Equation A.55 as

$$\begin{aligned} F\left(\underline{k}, \underline{u}_1, \underline{u}_2, \underline{u}_3, \underline{v}\right) &= f\left(\underline{k}\right) - \sum_{j=1}^{K} u_{j1} k_{j1} - \sum_{j=1}^{K} u_{j2} k_{j2} \\ &\quad + \sum_{j=1}^{K} u_{j3} \left(k_{j3}^2 - k_{j1} k_{j2}\right) + \sum_{j=1}^{K} v_j \left(k_{j1} + k_{j2} - 1\right) \end{aligned} \tag{A.60}$$

To find a optimal solution for this problem, we use Karush-Kuhn-Tucker (KKT) sufficient condition. It is stated as followed.

(KKT Sufficient Condition[**?**]) Consider an optimization problem (P) defined as

$$\begin{aligned} \min_{\underline{x}} \quad & f\left(\underline{x}\right) \\ \text{subject to} \quad & g_k\left(\underline{x}\right) \leq 0,\ k = 1, \cdots, m \\ & h_l\left(\underline{x}\right) = 0,\ l = 1, \cdots, n, \end{aligned}$$

with Lagrangian function

$$L\left(\underline{x}, \underline{u}, \underline{v}\right) = f\left(\underline{x}\right) + g\left(\underline{x}\right)^T \underline{u} + h\left(\underline{x}\right)^T \underline{v}$$

Let $\underline{x}$ be a feasible solution of (P), and suppose $(\underline{x}, \underline{u}, \underline{v})$ satisfy

$$\begin{aligned} \nabla_{\underline{x}} L\left(\underline{x}, \underline{u}, \underline{v}\right) &= 0 \\ \underline{u} &\geq 0 \\ u_k g_k\left(\underline{x}\right) &= 0 \end{aligned}$$

Then if $f(\underline{x})$ is a pseudoconvex function, $g_k(\underline{x})$, $k = 1, \cdots, m$ are quasiconvex functions, and $h_l(\underline{x})$, $l = 1, \cdots, n$ are linear functions, then $\underline{x}$ is a global optimal solution.

Given $\left(P_{\underline{k}}\right)$, it is clear that the objective function $f(\underline{k})$ is a convex function, the equality constraints Equation A.59 are linear, and the sets of inequality constraints Equation A.56, Equation A.57, and Equation A.58 are convex. Notice that a convex function is a special case of pseudoconvex and quasiconvex. Comparing the standard problem (P) in Theorem Equation A.13 with our optimization problem $(P_{\mathbf{K}})$, we can conclude that any feasible $\underline{k}$ satisfying

$$\begin{aligned}
\nabla_{\underline{k}} F\left(\underline{k},\, \underline{u}_1, \underline{u}_2,\, \underline{u}_3, \underline{v}\right) &= 0\\
\underline{u}_1, \underline{u}_2 \,\text{and}\, \underline{u}_3 &\geq 0\\
u_{j1}k_{j1} &= 0\\
u_{j2}k_{j2} &= 0\\
u_{j3}\left(k_{j3}^2 - k_{j1}k_{j2}\right) &= 0
\end{aligned}$$

is a global optimal for $\left(P_{\underline{k}}\right)$. Solving $\nabla_{\underline{k}} F\left(\underline{k},\, \underline{u}_1, \underline{u}_2,\, \underline{u}_3, \underline{v}\right)$ we have

$$\begin{aligned}
\frac{\nabla F}{\nabla k_{j1}} &= 2k_{j1} + 2\alpha \sum_{i=1, i\neq j}^{K} k_{i1} - u_{j1} - u_{j3}k_{j2} + v_j = 0\\
\frac{\nabla F}{\nabla k_{j2}} &= 2k_{j2} + 2\alpha \sum_{i=1, i\neq j}^{K} k_{i2} - u_{j2} - u_{j3}k_{j1} + v_j = 0\\
\frac{\nabla F}{\nabla k_{j3}} &= 4k_{j3} + 4\alpha \sum_{i=1, i\neq j}^{K} k_{i3} + 2u_{j3}k_{j3} = 0.
\end{aligned}$$

It is easy to check that $k_{j1} = k_{j2} = \frac{1}{2}$, $k_{j3} = 0$ while the Lagrange multipliers $u_{j1} = u_{j2} = u_{j3} = 0$, and $v_j = -1 - \alpha\,(K-1)$ satisfy KKT condition.

Therefore, $k_{j1} = k_{j2} = \frac{1}{2}$, $k_{j3} = 0$, i.e. $\hat{\mathbf{V}}_{x_j} = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$ is a global optimal solution. Substitute this optimal solution into the formula of sum slope Equation 5.30, the sum slope has upper bound

$$\mathcal{S}_0 \;\leq\; \frac{2K}{\alpha K + (1-\alpha)}$$

A.14 Proof of Corollary 48

Before proving this result, we state existing results for general parametric optimization problems. A general parametric optimization problem $P\left(\underline{t}\right)$ depending on parameters $\underline{t} \in \mathbb{R}^r$ is defined by

$$\begin{aligned} \min \quad & f\left(\underline{x}, \underline{t}\right) \\ \text{subject to} \quad & \underline{x} \in \mathbb{R}^n \\ & g_i\left(\underline{x}, \underline{t}\right) \leq 0,\ i = 1, \cdots, s \\ & g_i\left(\underline{x}, \underline{t}\right) = 0,\ i = s+1, \cdots, m \end{aligned}$$

where f and g_i are real functions. Denote the parametric feasible region by

$$\begin{aligned} A\left(\underline{t}\right) \triangleq & \left\{\underline{x} \middle| \underline{x} \in \mathbb{R}^n;\ g_i\left(\underline{x}, \underline{t}\right) \leq 0\, if\ i = 1, \cdots, s; \right. \\ & \left. g_i\left(\underline{x}, \underline{t}\right) = 0\, if\ i = s+1, \cdots, m\right\}. \end{aligned}$$

And denote the parametric optimal value function by $\nu\left(\underline{t}\right) \triangleq \inf_{\underline{x} \in A(\underline{t})} f\left(\underline{x}, \underline{t}\right)$. The following theorem gives the sufficient condition under which $\nu\left(\underline{t}\right)$ is a continuous function of $\underline{t}$.

Suppose that

1. the function f is continuous on $\underline{x} \times \underline{t}$;

   a) the correspondence A is continuous on $\underline{t}$;

   b) the subsets $A\left(\underline{t}\right)$ are non empty and compact

   Then the optimal value function $\nu\left(\underline{t}\right)$ is continuous and the correspondence optimal solution set is upper semi-continuous.

Let $\underline{h}$ correspond to $\underline{t}$, and let the $\underline{k}$ as that defined in Appendix section A.13 correspond to $\underline{x}$ of Theorem section A.14. It is easy to see that the objective function of Equation A.50 is continuous on $\underline{k} \times \underline{h}$, while the feasible region $A\left(\underline{h}\right)$ is non empty, compact, and independent of $\underline{h}$. Therefore, all three conditions in Theorem section A.14 are satisfied and the optimal value function $f\left(\underline{k},\, \underline{h}\right)$ is continuous on $\underline{h}$.

Further, in Theorem 47 we have shown that when $\underline{h}_o = \left\{\underline{h} :\ \phi_{ji} = 0,\ \left|h_{jj}\right|^2 = 1,\ \left|h_{ji}\right|^2 = \alpha\right\}$, the optimal value of the objective function of the optimization problem $P_{\underline{k}}\left(\underline{h}_o\right)$ is

$$f\left(\underline{k},\, \underline{h}_o\right) = \frac{2K}{\alpha K + (1-\alpha)}.$$

Given the continuity of $f\left(\underline{k},\, \underline{h}_o\right)$ provided by Theorem section A.14 , for any σ, there exist σ_1, σ_2, σ_3 such that for the channels $\underline{h} \in \tilde{\mathcal{H}}_\sigma$, where the set $\tilde{\mathcal{H}}_\sigma$ is defined as

$$\begin{aligned}
\tilde{\mathcal{H}}_{\sigma} &= \{\underline{h} : |\phi_{ji}| \leq \sigma_1 \\
& \quad \left||h_{jj}|^2 - 1\right| \leq \sigma_2 \\
& \quad \left|\sqrt{|h_{ji}|^2} - \alpha\right| \leq \sigma_3 \\
& \quad {}^{|h_{ij}|^2}\!/\!_{|h_{jj}|^2} < 1 \\
& \quad \underline{h} \in \mathcal{C}_{\alpha}\},
\end{aligned}$$

the optimal value of the objective function of the optimization problem $P_{\underline{k}}(\underline{h})$ satisfies

$$|f(\underline{k}, \underline{h}) - f(\underline{k}, \underline{h}_o)| < \sigma.$$

Notice that $\mathcal{H}_{\alpha}$ is defined in Theorem 45.

Because $\mathbf{1} \in \mathrm{cl}\left(\tilde{\mathcal{H}}_{\sigma}\right)$, as $\alpha \rightarrow 1$, for any positive σ, there exists $\tilde{\mathcal{H}}_{\sigma}$, such that for $\underline{h} \in \tilde{\mathcal{H}}_{\sigma}$ its sum slope satisfies

$$\mathcal{S}_0 \leq 2 + \sigma, \tag{A.61}$$

If the magnitude and phase of the channel coefficients are drawn from continuous random distribution, $Pr(\mathcal{H}_{\sigma}) > 0$.

And as $\sigma \rightarrow 0$,

$$\lim_{\sigma \rightarrow 0} \Delta\mathcal{S}_0 = \frac{1}{K}$$

Bibliography

[1] S. Verdú, "Spectral efficiency in the wideband regime," *IEEE Transactions on Information Theory*, vol. 48, no. 6, pp. 1319–1343, 2002.

[2] G. Caire, D. Tuninetti, and S. Verdú, "Suboptimality of tdma in the low-power regime," *IEEE Transactions on Information Theory*, vol. 50, no. 4, pp. 608–620, 2004.

[3] C. E. Shannon, "A mathematical theory of communication," *The Bell Systems Technical Journal*, vol. 27, pp. 379–423, 623–656, 1948.

[4] A. D. Wyner, "The capacity of the band-limited gaussian channel," *Bell Sys. Tech. J.*, vol. 45, pp. 359–395, 1966.

[5] I. Jacobs, "The asymptotic behavior of incoherent m-ary communication systems," *Proceedings of the IEEE*, vol. 51, no. 1, pp. 251 – 252, jan. 1963.

[6] R. Kennedy, *Fading dispersive communication channels.* Wiley-Interscience, 1969. [Online]. Available: http://books.google.com/books?id=tt5SAAAAMAAJ

[7] L. Zheng, D. N. C. Tse, and M. Medard, "Channel coherence in the low-snr regime," *Information Theory, IEEE Transactions on*, vol. 53, no. 3, pp. 976 –997, march 2007.

[8] S. Ray, M. Medard, and L. Zheng, "On noncoherent mimo channels in the wideband regime: Capacity and reliability," *Information Theory, IEEE Transactions on*, vol. 53, no. 6, pp. 1983 –2009, june 2007.

[9] M. J. E. Golay, "Note on the theoretical efficiency of information reception with ppm," *Proceedings of the IRE*, vol. 37, no. 9, pp. 1029 – 1031, sept. 1949.

[10] H. Sato, "The capacity of the gaussian interference channel under strong interference (corresp.)," *Information Theory, IEEE Transactions on*, vol. 27, no. 6, pp. 786–788, 1981. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1056416

[11] M. Costa, "On the gaussian interference channel," *Information Theory, IEEE Transactions on*, vol. 31, no. 5, pp. 607–615, 1985. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1057085

[12] T. M. Cover and J. A. Thomas, *Elements of Information Theory 2nd Edition*, ser. Wiley Series in Telecommunications and Signal Processing. Wiley-Interscience, July 2006. [Online]. Available: http://www.amazon.ca/exec/obidos/redirect?tag=citeulike09-20&path=ASIN/0471241954

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

[14] H.-F. Chong, M. Motani, H. Garg, and H. El Gamal, "On the han kobayashi region for theinterference channel," *Information Theory, IEEE Transactions on*, vol. 54, no. 7, pp. 3188 –3195, july 2008.

[15] G. Kramer, "Outer bounds on the capacity of gaussian interference channels," *Information Theory, IEEE Transactions on*, vol. 50, no. 3, pp. 581–586, 2004. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1273673

[16] R. Etkin, D. Tse, and H. Wang, "Gaussian interference channel capacity to within one bit," Feb 2007. [Online]. Available: http://arxiv.org/abs/cs/0702045v1

[17] V. Sreekanth Annapureddy and V. Veeravalli, "Sum capacity of the gaussian interference channel in the low interference regime," in *Information Theory and Applications Workshop, 2008*, 27 2008-Feb. 1 2008, pp. 422–427.

[18] A. Host-Madsen and A. Nosratinia, "The multiplexing gain of wireless networks," in *Information Theory, 2005. ISIT 2005. Proceedings. International Symposium on*, sept. 2005, pp. 2065 –2069.

[19] V. Cadambe and S. Jafar, "Interference alignment and degrees of freedom of the -user interference channel," *Information Theory, IEEE Transactions on*, vol. 54, no. 8, pp. 3425 –3441, aug. 2008.

[20] R. H. Etkin and E. Ordentlich, "The degrees-of-freedom of the k-user gaussian interference channel is discontinuous at rational channel coefficients," *IEEE Trans. Inf. Theor.*, vol. 55, no. 11, pp. 4932–4946, 2009.

[21] V. R. Cadambe, S. A. Jafar, and C. Wang, "Interference alignment with asymmetric complex signaling - settling the host-madsen-nosratinia conjecture," *CoRR*, vol. abs/0904.0274, 2009.

[22] P. Bergmans, "A simple converse for broadcast channels with additive white Gaussian noise," *IEEE Transactions on Information Theory*, vol. IT-20, no. 2, pp. 279–280, March 1974.

[23] X. Shang, G. Kramer, and B. Chen, “"A New Outer Bound and the Noisy-Interference Sum-Rate Capacity for Gaussian Interference Channels",” *ArXiv e-prints*, vol. 712, dec 2007.

[24] A. Motahari and A. Khandani, “Capacity bounds for the gaussian interference channel,” *Information Theory, IEEE Transactions on*, vol. 55, no. 2, pp. 620–643, Feb. 2009.

[25] X. Shang, G. Kramer, and B. Chen, “A new outer bound and the noisy-interference sum rate capacity for gaussian interference channels,” *Information Theory, IEEE Transactions on*, vol. 55, no. 2, pp. 689 –699, feb. 2009.

[26] R. Etkin and E. Ordentlich, “On the degrees-of-freedom of the k-user gaussian interference channel,” in *Information Theory, 2009. ISIT 2009. IEEE International Symposium on*, 28 2009-july 3 2009, pp. 1919 –1923.

[27] K. Gomadam, V. Cadambe, and S. Jafar, “Approaching the capacity of wireless networks through distributed interference alignment,” in *Global Telecommunications Conference, 2008. IEEE GLOBECOM 2008. IEEE*, 30 2008-dec. 4 2008, pp. 1 –6.

[28] D. Tuninetti, “K-user interference channels: General outer bound and sum-capacity for certain gaussian channels,” in *Information Theory Proceedings (ISIT), 2011 IEEE International Symposium on*, 31 2011-aug. 5 2011, pp. 1166 –1170.

[29] V. Cadambe, S. Jafar, and S. Shamai, “Interference alignment on the deterministic channel and application to fully connected awgn interference networks,” in *Information Theory Workshop, 2008. ITW ’08. IEEE*, may 2008, pp. 41 –45.

[30] M. Maddah-Ali, A. Motahari, and A. Khandani, “Communication over mimo x channels: Interference alignment, decomposition, and performance analysis,” *Information Theory, IEEE Transactions on*, vol. 54, no. 8, pp. 3457 –3470, aug. 2008.

[31] H. Weingarten, T. Liu, S. Shamai, Y. Steinberg, and P. Viswanath, “The capacity region of the degraded multiple-input multiple-output compound broadcast channel,” *Information Theory, IEEE Transactions on*, vol. 55, no. 11, pp. 5011 –5023, nov. 2009.

[32] R. Etkin and E. Ordentlich, “The degrees-of-freedom of the -user gaussian interference channel is discontinuous at rational channel coefficients,” *Infor-*

mation Theory, IEEE Transactions on, vol. 55, no. 11, pp. 4932 –4946, nov. 2009.

[33] A. Jafarian, J. Jose, and S. Vishwanath, "Algebraic lattice alignment for k-user interference channels," in *Communication, Control, and Computing, 2009. Allerton 2009. 47th Annual Allerton Conference on*, 30 2009-oct. 2 2009, pp. 88 –93.

[34] G. Bresler, A. Parekh, and D. Tse, "The approximate capacity of the many-to-one and one-to-many gaussian interference channels," *Information Theory, IEEE Transactions on*, vol. 56, no. 9, pp. 4566 –4592, sept. 2010.

[35] M. Shen and A. Høst-Madsen, "The wideband slope of interference channels: The large bandwidth case," *Information Theory, IEEE Transactions on*, Accepted for publication, available at http://arxiv.org/abs/1010.5661.

[36] V. Cadambe and S. Jafar, "Degrees of freedom of wireless networks - what a difference delay makes," nov. 2007, pp. 133 –137.

[37] L. Grokop, D. N. C. Tse, and R. D. Yates, "Interference alignment for line-of-sight channels," *CoRR*, vol. abs/0809.3035, 2008.

[38] T. M. Apostol, *Modular Functions and Dirichlet Series in Number Theory*, 2nd ed. New York: Springer-Verlag Inc., 1990, graduate Texts in Mathematics;41.

[39] M. Shen and A. Hø andst Madsen, "Wideband slope of interference channel: Finite bandwidth case," in *Communication, Control, and Computing (Allerton), 2011 49th Annual Allerton Conference on*, 2011, allerton 2011.

[40] M. Shen and A. Høst-Madsen, "The wideband slope of interference channels: The small bandwidth case," *Information Theory, IEEE Transactions on*, Submitted, 2012.

[41] G. Bresler, A. Parekh, and D. Tse, "The approximate capacity of the many-to-one and one-to-many gaussian interference channels," *Information Theory, IEEE Transactions on*, vol. 56, no. 9, pp. 4566 –4592, sept. 2010.

[42] W. Rudin, *Principles of mathematical analysis*, 3rd ed. New York: McGraw-Hill Book Co., 1976, international Series in Pure and Applied Mathematics.

[43] G. Grimmett and D. Stirzaker, *Probability and random processes.* Oxford University Press, Jul. 2001.

[44] B. Bollobas, *Random Graphs*, W. Fulton, A. Katok, F. Kirwan, P. Sarnak, B. Simon, and B. Totaro, Eds. Cambridge University Press, 2001.

[45] S. Janson, "Random graphs." Wiley, 2000.

[46] X. Shang, B. Chen, G. Kramer, and H. Poor, "Capacity regions and sum-rate capacities of vector gaussian interference channels," *Information Theory, IEEE Transactions on*, vol. 56, no. 10, pp. 5030 –5044, oct. 2010.

[47] R. Bhatia, *Matrix analysis*, ser. Graduate texts in mathematics. Springer, 1997. [Online]. Available: http://books.google.com/books?id=f0ioPwAACAAJ