Randomized Resource Allocation in Decentralized Wireless Networks

by

Kamyar Moshksar

A thesis
presented to the University of Waterloo
in fulfilment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2011

AUTHOR'S DECLARATION

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Ad hoc networks and bluetooth systems operating over the unlicensed ISM[1] band are instances of decentralized wireless networks. By definition, a decentralized network is composed of separate transmitter-receiver pairs where there is no central controller to assign the resources to the users. As such, resource allocation must be performed locally at each node. Users are anonymous to each other, i.e., they are not aware of each other's code-books. This implies that multiuser detection is not possible and users treat each other as noise. Multiuser interference is known to be the main factor that limits the achievable rates in such networks particularly in the high Signal-to-Noise Ratio (SNR) regime. Therefore, all users must follow a distributed signaling scheme such that the destructive effect of interference on each user is minimized, while the resources are fairly shared.

In chapter 2 we consider a decentralized wireless communication network with a fixed number u of frequency sub-bands to be shared among N transmitter-receiver pairs. It is assumed that the number of active users is a realization of a random variable with a given probability mass function. Moreover, users are unaware of each other's codebooks and hence, no multiuser detection is possible. We propose a randomized Frequency Hopping (FH) scheme in which each transmitter randomly hops over a subset of u sub-bands from transmission slot to transmission slot. Assuming all users transmit Gaussian signals, the distribution of the noise plus interference is mixed Gaussian, which makes calculation of the mutual information between the transmitted and received signals of each user intractable. We derive lower and upper bounds on the mutual information of each user and demonstrate that, for large SNR values, the two bounds coincide. This observation enables us to compute the sum multiplexing gain of the system and obtain the optimum hopping strategy for maximizing this quantity. We compare the performance of the FH system with that of the Frequency Division (FD) system in terms of the following performance measures: average sum multiplexing gain ($\eta^{(1)}$) and average minimum multiplexing gain per user ($\eta^{(2)}$). We show that (depending on the probability mass function of the number of active users) the FH system can offer a significant improvement in terms of $\eta^{(1)}$ and $\eta^{(2)}$ (implying a more efficient usage of the spectrum). In the sequel, we consider a scenario where the transmitters are unaware of the number of active

[1] Industrial, Scientific and Medical.

users in the network as well as the channel gains. Developing a new upper bound on the differential entropy of a mixed Gaussian random vector and using entropy power inequality, we obtain lower bounds on the maximum transmission rate per user to ensure a specified outage probability at a given SNR level. We demonstrate that the so-called outage capacity can be considerably higher in the FH scheme than in the FD scenario for reasonable distributions on the number of active users. This guarantees a higher spectral efficiency in FH compared to FD.

Chapter 3 addresses spectral efficiency in decentralized wireless networks of separate transmitter-receiver pairs by generalizing the ideas developed in chapter 2. Motivated by random spreading in Code Division Multiple Access (CDMA), a signaling scheme is introduced where each user's code-book consists of two groups of codewords, referred to as signal codewords and signature codewords. Each signal codeword for the i^{th} user is a sequence of i.i.d. Gaussian random variables $\mathbf{x}_i[0], \mathbf{x}_i[1], \cdots$ and each signature codeword is a sequence of i.i.d. random vectors $\vec{\mathbf{s}}_i[0], \vec{\mathbf{s}}_i[1], \cdots$ constructed over a globally known alphabet. To represent a message, a unique codeword in each group is selected and the sequence of vectors $\mathbf{x}_i[0]\vec{\mathbf{s}}_i[0], \mathbf{x}_i[1]\vec{\mathbf{s}}_i[1], \cdots$ is transmitted. Using a conditional entropy power inequality and a key upper bound on the differential entropy of a mixed Gaussian random vector, we develop an inner bound on the capacity region of the decentralized network. To guarantee consistency and fairness, each user designs its signature codewords based on maximizing the average (with respect to a globally known distribution on the channel gains) of the achievable rate per user. It is demonstrated how the Sum Multiplexing Gain (SMG) in the network (regardless of the number of users) can be made arbitrarily close to the SMG of a centralized network with an orthogonal scheme such as Time Division (TD). An interesting observation is that in general the elements of the vectors in a signature codeword must not be equiprobable over the underlying alphabet in contrast to the use of binary Pseudo-random Noise (PN) signatures in randomly spread CDMA where the chip elements are $+1$ or -1 with equal probability. The main reason for this phenomenon is the interplay between two factors appearing in the expression of the achievable rate, i.e., multiplexing gain and the so-called interference entropy factor. In the sequel, invoking an information theoretic extremal inequality, we present an optimality result by showing that in randomized frequency hopping which is the main idea in the prevailing bluetooth devices in decentralized networks, transmission of i.i.d. signals in consecutive transmission slots is in general suboptimal regardless of the distribution of the signals.

Finally, chapter 4 addresses a decentralized Gaussian interference channel consisting of two block-asynchronous transmitter-receiver pairs. We consider a scenario where the rate of data arrival at the encoders is considerably low and codewords of each user are transmitted at random instants depending on the availability of enough data for transmission. This makes the transmitted signals by each user look like scattered bursts along the time axis. Users are block-asynchronous meaning there exists a delay between their transmitted signal bursts[2]. The proposed model for asynchrony assumes the starting point of an interference burst is uniformly distributed along the transmitted codeword of any user. There is also the possibility that each user does not experience interference on a transmitted codeword at all. Due to the randomness of delay, the channels are non-ergodic in the sense that the transmitters are unaware of the location of interference bursts along their transmitted codewords. In the proposed scheme, upon availability of enough data in its queue, each user follows a locally Randomized Masking (RM) strategy where the transmitter quits transmitting the Gaussian symbols in its codeword independently from symbol interval to symbol interval. An upper bound on the probability of outage per user is developed using entropy power inequality and a key upper bound on the differential entropy of a mixed Gaussian random variable. It is shown that by adopting the RM scheme, the probability of outage is considerably less than the case where both users transmit the Gaussian symbols in their codewords in consecutive symbol intervals, referred to as Continuous Transmission (CT).

[2] Users are synchronous at the symbol level.

Acknowledgements

First and foremost, I am deeply grateful to my supervisor, Professor Amir K. Khandani, for providing guidance and inspiration throughout the course of my studies at University of Waterloo. It was a great honor for me to work with such a brilliant, insightful, creative, and knowledgeable supervisor. His attitude and kindness has made him a compassionate and everlasting friend in my life.

I also wish to thank the members of my dissertation committee, Professors , Randall Berry, Patrick Mitran, Alexandru Nica and Liang-Liang Xie for having accepted to take the time to read my thesis and to provide me with their comments and suggestions.

I have been fortunate to work among members of the Coding and Signal Transmission (CST) laboratory. I would like to thank all CST members as they created a very pleasant, intellectually stimulating and friendly environment.

At the end, I express my gratitude to my teachers, Professors, Amir K. Khandani, Liang-Liang Xie, Oussama Damen, Andrew Heunis of the Electrical and Computer Enginnering Department and Brian Forrest, Alexandru Nica, Nico Spronk and Ruxandra Moraru of the Pure Mathematics Department at the University of Waterloo. I am truly glad that I had the experience of enjoying their lectures.

To my parents,
Zahra Mahmoodian and Rahim Moshksar

Table of Contents

List of Figures

List of Abbreviations

FH	Frequency Hopping
AFH	Adaptive Frequency Hopping
TH	Time Hopping
FD	Frequency Division
FBS	Full-Band Spreading
RM	Randomized Masking
CT	Continuous Transmission
MG	Multiplexing Gain
SMG	Sum Multiplexing Gain
IEF	Interference Entropy Factor
CSF	Channel-Signature Factor
CDMA	Code Division Multiple Access
MAC	Multiple Access Channel
NND	Nearest Neighbor Decoding
LFSR	Linear Feedback Shift Register
LDCT	Lebesgue Dominated Convergence Theorem
SNR	Signal to Noise Ratio
PDF	Probability Density Function
PMF	Probability Mass Function
MF	Matched Filter
i.i.d.	Independent and Identically Distributed

List of Notations

Random Quantities	Bolds Letters
$f(x) \overset{x}{\sim} g(x)$	If and only if $\lim_{x\to\infty} \frac{f(x)}{g(x)} = 1$.
$f \overset{x}{\lesssim} g$	if and only if $\lim_{x\to\infty} \frac{f(x)}{\log x} \leq \lim_{x\to\infty} \frac{g(x)}{\log x}$.
$\mathrm{Ber}(p)$	A Bernoulli random variable with parameter p
$\mathcal{CN}(\vec{m}, C)$	A complex Gaussian vector of mean $\vec{m}$ and covariance matrix C
$\mathcal{N}(\vec{m}, C)$	A real Gaussian vector of mean $\vec{m}$ and covariance matrix C
A^{t}	The transpose of a matrix A
$A^{\dagger}$	THe conjugate transpose of a matrix A
$\|A\|$	The Euclidean norm of a matrix A, i.e., $\|A\| = \sqrt{\mathrm{tr}(A^{\dagger}A)}$.
$[A]_i$	The i^{th} column of a matrix A
$\mathrm{col}(A)$	The set of columns of a matrix A
$\mathrm{csp}(A)$	The column space of a matrix A
$\mathrm{rank}(A)$	The rank of a matrix A
$[\vec{v}]_i$	The i^{th} element of a vector $\vec{v}$
$[\vec{v}]_i^j$	The sub-vector of $\vec{v}$ consisting of its i^{th} element to its j^{th} element
$\mathrm{span}(\mathcal{V})$	The subspace spanned by the vectors in the set $\mathcal{V}$
$\vec{x} \odot \vec{y}$	The element-wise product of $\vec{x}$ and $\vec{y}$
$\mathrm{h}(\boldsymbol{x})$	The differential entropy of $\boldsymbol{x}$
$\mathrm{H}(\boldsymbol{x})$	The discrete entropy of $\boldsymbol{x}$
$\mathscr{H}(x)$	The binary entropy function: $\mathscr{H}(x) = -x\log x - (1-x)\log(1-x)$
$(a_l)_{l=1}^n$	The sequence $(a_1, a_2, \cdots, a_n)$
$(a_l)_{l\in\mathcal{A}}$	For $\mathcal{A} \subset \{1, 2, \cdots, n\}$, $(a_l)_{l\in\mathcal{A}}$ is a subsequence of $(a_l)_{l=1}^n$
$\mathbb{I}\{.\}$	The indicator function
$\bar{x}$	$1 - x$
$1_{m\times n}$	The $m \times n$ matrix of all 1s
$0_{m\times n}$	The $m \times n$ matrix of all 0s

$\mathcal{A} \backslash \mathcal{B}$	The difference between two sets $\mathcal{A}$ and $\mathcal{B}$
$\mathcal{A} \times \mathcal{B}$	The cartesian product of two sets $\mathcal{A}$ and $\mathcal{B}$

Chapter 1

Introduction

Increasing demand for wireless applications on one hand, and the limited available resources on the other hand, provoke more efficient usage of such resources. Due to its significance, many researchers have addressed the problem of resource allocation in wireless networks. One major challenge in wireless networks is the destructive effect of multiuser interference, which degrades the performance when multiple users share the spectrum. As such, an efficient and low complexity resource allocation scheme that maximizes the quality of service while mitigating the impact of the multiuser interference is desirable. The existing resource allocation schemes are either *centralized*, i.e., a central controller manages the resources, or *decentralized*, where resource allocation is performed locally at each node. Due to the complexity of adapting the centralized schemes to the network structure (e.g. number of active users), these schemes are usually designed for a fixed network structure. This makes inefficient usage of resources because, in most cases, the number of active users may be considerably less than the value assumed in the design process. On the other hand, most of the decentralized resource allocation schemes suffer from the complexity, either in the algorithm (e.g. game-theoretic approaches involving iterative methods) or in the hardware (e.g. cognitive radio). Therefore, it is of interest to devise an efficient and low-complexity decentralized resource allocation scheme, which is the main goal of this thesis.

1.1 Review of Existing Centralized and Decentralized Resource Allocation Schemes

1.1.1 Centeralized Schemes

In recent years, many centralized power and spectrum allocation schemes have been studied in cellular and multihop wireless networks [1, 2, 3, 4, 5, 6, 7, 8]. Clearly, centralized schemes perform better than the decentralized (distributed) approaches, while requiring extensive knowledge of the network configuration. In particular, when the number of nodes is large, deploying such centralized schemes may not be practically feasible.

Traditional wireless systems aimed to avoid the interference among users by using orthogonal transmission schemes. The most common example is the Frequency Division (FD) system, in which different users transmit over disjoint frequency sub-bands. The assignment of frequency sub-bands is usually performed by a central controller. Despite its simplicity, FD is shown to achieve the highest throughput in certain scenarios. In particular, [9] proves that in a wireless network where interference is treated as noise (no multiuser detection is performed), if the crossover gains are sufficiently larger than the forward gains, FD is Pareto-rate-optimal. Due to practical considerations, such FD systems usually rely on a fixed number of frequency sub-bands. Hence, if the number of users changes, the system is not guaranteed to offer the best possible spectral efficiency because, most of the time, the majority of the potential users may be inactive.

1.1.2 Decentralized Schemes

In decentralized schemes, decisions concerning network resources are made by individual nodes based on their local information. Two important classes of decentralized schemes reported in the literature rely on either *game-theoretic* approaches or *cognitive radios*. Cognitive radios [10] have the ability to sense the unoccupied portion of the available spectrum and use this information in resource allocation. Fundamental limits of wireless networks with cognitive radios are studied in [11, 12, 13, 14]. Although cognitive radios avoid the use of a central controller, they require sophisticated detection techniques for sensing the spectrum holes and dynamic frequency assignment, which add to the overall system complexity [15]. Noting the above points, it is desirable to have a decentralized frequency sharing strategy without the need for cognitive radios, which allows the users to coexist while utilizing the spectrum efficiently and fairly.

Spread spectrum communications [16] offers techniques for sharing the same bandwidth by several users. Historically, it was used as a means of low interception for military use. This area has attracted tremendous attention during the past three decades in the context of centralized uplink/downlink multiuser systems. Appealing characteristics of spread spectrum systems have motivated researchers to utilize these schemes in networks without a fixed infrastructure, e.g., packet radio or ad-hoc networks [17]. In direct sequence spread spectrum systems, the signal of each user is spread by a PN signature. The challenging point is that if two users pick the same signature, they will not be capable of recovering the data at the receiver side due to the high amount of interference. A distributed code assignment technique is developed in [18] for a certain snapshot of an ad hoc network with an arbitrary communication topology. The main idea is to use signatures with variable spreading factors based on the local traffic in the network. Using a greedy approximation algorithm and invoking graph theory, a distributed code assignment protocol is suggested to maximize the throughput of the network.

Randomly spread direct sequence techniques proposed in [19, 20] represent practical CDMA systems in the context of Multiple Access Channels (MAC). To generate its signature in practice, each transmitter utilizes the output of a Linear Feedback Shift Register (LFSR). An LFSR generates a maximum length sequence (m-sequence) of relatively large period over the alphabet $\{-1, +1\}$. Thereafter, the transmitter breaks the m-sequence generated by its LFSR into a large number of pieces (signatures) and utilizes these signatures to spread the symbols in its codeword. This can be accurately modeled by random spreading where each transmitter generates a binary random signature whose elements are $+1$ or -1 equally likely. Authors in [19] assess the capacity loss of such a mechanism under various sub-optimum and low-complexity multiuser detection strategies such as conventional matched filter, decorrelator and minimum mean square error estimator. An important characteristic of this randomized signaling is that each transmitter runs its own LFSR locally and there is no need to design the optimum capacity-achieving signatures and assign them to the users. However, the common receiver is required to know the signatures of all users in order to decode their messages.

Being a standard technique in spread spectrum communications and due to its interference avoidance nature, hopping is the simplest spectrum sharing method to use in decentralized networks. As different users typically have no prior information about the codebooks of the other users, the most efficient method is avoiding interference by choosing unused channels. As mentioned earlier, searching the spectrum to find spectrum holes is not an easy task due to the dynamic spectrum usage. As such, Frequency Hopping (FH) is a realization of a transmission scheme without sensing, while avoiding the collisions as much as possible. Frequency Hopping is one of the standard signaling schemes [17] adopted in ad-hoc networks. In short range scenarios, bluetooth systems [21, 22, 23] are the most popular examples of a Wireless Personal Area Network (WPAN). Using FH over the unlicensed ISM band, a bluetooth system provides robust communication to unpredictable sources of interference. A modification of Frequency Hopping, called Dynamic Frequency Hopping (DFH), selects the hopping pattern based on interference measurements in order to avoid dominant interferers. The performance of a DFH scheme when applied to a cellular system is assessed in [24, 25, 26].

Although there has been a tremendous amount of work on the performance evaluation of hopping-based decentralized networks, there are only a few information-theoretic results reflecting the fundamental limits of such networks. We will offer a thorough analysis of a decentralized network where all users follow a randomized hopping strategy to share the resources, while the number of active users is a random variable with a given distribution. In one scenario, the channel gains and the number of active users are assumed to be static and known to the corresponding transmitter-receiver nodes. This assumption makes the concept of achievable rate in the Shannon sense meaningful. However, in case the channel gains and the number of active users are not known to the transmitters, the concept of achievable rate may no longer be valid. A common setup for such an assumption is a network where channel gains are quasi-static fading and unknown to the transmitters, which is also addressed in the thesis.

Rayleigh fading is an unavoidable phenomenon in wireless networks that can affect the performance of the system significantly. Traditionally, Rayleigh fading has been considered to be harmful due to reducing the transmission reliability in wireless networks. However, recently, researchers have been able to reduce this harmful effect by exploiting the so-called *multiuser diversity* [27, 28]. This can be considered a *scheduling gain* by allowing the users with favorable channels to be active. It is shown that multiuser diversity gain can be as large as $\log \log N$ in broadcast and multiple-access channels [29, 30, 31] and $\log N$ in single-hop ad hoc networks [32, 33], as N grows to infinity. However, achieving this scheduling gain requires the fading channels to vary over time such that all possible realizations of the fading process are covered. In the case that channel gains are selected randomly at the start of the transmission and remain constant during the whole transmission period (quasi-static fading), the channels do not have ergodic behavior. In this case, a suitable performance measure is the so-called *ε-outage capacity* [34], denoted by $R(\varepsilon)$, which is defined as the maximum transmission rate per user, ensuring an outage probability below ε, i.e.,

$$R(\varepsilon) = \sup\{R : \Pr\{\text{Outage}\} < \varepsilon\}.$$

The reality of wireless channels is more complicated to be simply represented by the Rayleigh fading model. A class of channel models considered in the literature is the one in which the signal power decays according to a distance-based attenuation law [35, 36, 37, 38, 39, 40, 41, 42]. Moreover, the presence of obstacles adds some randomness (known as shadowing) to the received signal. It is well known that the effects of such random phenomena can significantly affect the throughput of a spectrum sharing network in both multi-hop [43, 44, 45, 46] and single-hop scenarios [47] (Chapter 8), [48, 49, 50, 51, 52]. These features indeed increase the frequency reuse factor as they will attenuate the interference caused by a given transmitter on its neighboring receivers. In spite of the significance of the effects of distance-based attenuation and shadowing on the throughput of a spectrum sharing system, unfortunately, there is not a single commonly accepted model to represent these factors. It should be emphasized that the inclusion of signal attenuation due to distance and/or shadowing will indeed simplify the spectrum sharing as the multiuser interference will be attenuated and consequently its harmful effect will be reduced. On the other hand, these factors do not impact the performance of the orthogonal schemes in which the multiuser interference is altogether avoided. In spite of this fact, as the actual model used to represent distance-based attenuation and shadowing can have a profound impact on the system throughput (to the advantage of the randomized spectrum sharing schemes advocated in the current article), to avoid any confusion, we have relied on a simple Rayleigh fading model which in some sense captures the minimum advantage offered by the proposed scheme vs. those based on orthogonal separation of users.

Reference [53] studies a wireless network composed of a set of transmitter/receiver pairs in which a given link can be off or transmit with a constant power. Both the case of Rayleigh fading as well as a Rayleigh fading mixed with a proper distance-based attenuation are considered and a comparison between the scaling (with respect to the number of links) of the throughput in these two cases is provided.

In [54, 55], the authors study a decentralized wireless ad hoc network where different transmitters are connected to different receivers through channels with a similar path loss exponent. Assuming the transmitters are scattered over the two dimensional plane according to a Poisson point process, a fixed bandwidth is partitioned into a certain number of sub-bands, such that the so-called transmission intensity in the network is maximized, while the probability of outage per user is below a certain threshold[54]. The transmission strategy is based on choosing one sub-band randomly per transmission, which is a special case of Frequency Hopping. In [55], a non-iterative and distributed power control scheme is introduced, for which the constant power and the channel inversion schemes are extreme cases. It is observed that none of these cases are ideal in general. In fact, it is shown that regulating the transmission power proportionally to the inverse square root of the forward channel strength minimizes the outage probability.

Frequency hopping is also proposed in [14] in the context of cognitive radios, where each cognitive transmitter selects a frequency sub-band but quits transmitting if the sub-band is already occupied by a primary user.

The popularity of Frequency Hopping motivates us to consider this scheme operating over a fixed bandwidth where due to practical constraints, the spectrum is divided to a number u of frequency sub-bands.

It is worth mentioning that the long-known ALOHA system [56] is a potential candidate for data communication in a decentralized network. In an ALOHA-like system, each user randomly selects a number of frequency sub-bands upon activation and transmits the entire symbols in its codeword through the selected sub-bands. Therefore, if two users overlap (partially or completely) on their selected sub-bands, they lose (some or the entire) information in the transmitted packets. However, in the randomized FH scheme, each user transmits the different symbols in its codeword through randomly selected sub-bands which changes from transmission slot to transmission slot. Overall, in an ALOHA-like system there is a nonzero probability that any user loses its entire codeword regardless of the transmission rate, while in the randomized FH scenario, one can always transmit bellow a certain rate and recover its data completely for each transmission block.

Although, the average throughput of an ALOHA-like system can be shown to be the same as that of the randomized FH scheme in the high SNR regime, there are basic assumptions made in an ALOHA system that are not necessary in the FH scheme:

1- The need to retransmit the lost packets in ALOHA-like systems demands the receiver to notify its associated transmitter that a packet is lost. This leads to considerable delay and stability[1] issues regarding the backlogged packets [57, 58, 59]. Also, there must be a feedback link from any receiver to its affiliated transmitter in order to report the lost packets. This decreases the overall spectral efficiency.

2- In the ALOHA system, all users are assumed to be block and symbol synchronized. However, in the proposed FH scheme, the assumption on block synchronization is either unnecessary (blocks are transmitted without interruption) or the analysis can be easily generalized to include cases that there are periods of silence between blocks.

Recently, researchers have aimed to improve the average throughput of an ALOHA network in the context of random multiple access. For example, let there be at most two users in the network where each user is active with a probability p. Inspired by the fact that the capacity region of a two-user multiple access channel is a polytope with two corner points (r_1, r_2) and (r'_1, r'_2) where r_1, r_2, r'_1 and r'_2 are nonzero rates, the authors in [60] propose that user 1 (user 2) selects r_1 and r'_1 (r_2 and r'_2) with probabilities q and $1-q$, respectively. As far as the pair of selected rates lies in the capacity region of the corresponding multiple access channel, the messages sent by both users can be received with arbitrarily small probability of error. Finally, q is selected to maximize the average throughput. An elegant approach is proposed in [61] where users transmit over several data streams (using message splitting) and the receiver decodes only certain data streams depending on which users are active. This results in characterizing the capacity region of the Gaussian random multiple access channel within a constant gap in a scenario where the activation status of both users is only available at the receiver. We emphasize that these techniques do not apply in the setup of a decentralized network as the capacity region of a decentralized interference channel is unknown and multiuser detection is not possible.

1.1.3 Block-Asynchronous Decentralized Networks

Interference channels represent networks of separate transmitter-receiver pairs where encoding and decoding are performed without any cooperation among the users. The capacity region of interference channels has been unknown for more than thirty years. Even the two-user case is only partially solved [62, 63, 64, 65]. In fact, [62] proves that the classical random coding scheme developed by Han and Kobayashi [66] can achieve within one bit of the capacity region of the two-user

[1] An ALOHA system is called stable if the Markovian process of backlogged packets is ergodic.

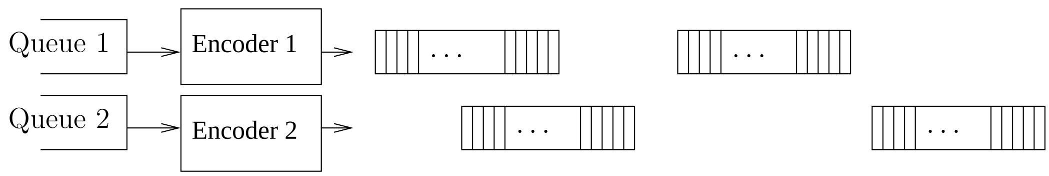

Figure 1.1: Asynchrony among the transmitted codes in a scenario where data traffic in each queue is relatively low.

interference channel for all ranges of channel coefficients and SNR values.

Two pivotal assumptions made in [62, 63, 64, 65, 66] and the references therein are:

1- The network is centralized, i.e., there is a central controller that assigns the resources (such as bandwidth) to the users. Moreover, users are aware of each other's code-books. One drawback of centralized scenarios is that the network is designed to service a certain number of users. Hence, if the number of active users is less than what is considered in the design phase, spectral efficiency will be very low, i.e, most of the bandwidth is not used.

2- Data is constantly available at the encoders for transmission. Also, all users are block-synchronous, i.e., they start to transmit their codewords in the same symbol interval in a slotted channel. This assumption is not necessarily valid in practice, because different users are not required to become active or inactive simultaneously. In fact, if the rate of data arrival at the encoders is relatively low, transmitters only transmit intermittently. In this case, codewords of different users overlap partially or do not overlap at all. This is illustrated in fig. 1.1.

Information theoretic studies on a network of block-asynchronous users is already investigated in [68, 69] in the context of centralized Multiple Access Channel (MAC) with two users. In case the amount of mutual delay between the codewords of the users in negligible compared to the length of codewords, it is shown in [68] that the capacity region of a MAC with block-asynchronous users coincides with the capacity region of a MAC with block-synchronous users. On the other hand, if the amount of delay is comparable with the block-length, the authors in [69] show that the capacity region is still similar to the capacity region of a MAC with block-synchronous users except for the fact that time-sharing is not possible anymore. In particular, it is concluded that the capacity region of a Gaussian MAC with block-asynchronous users is the same as the capacity region of block-syncronous Gaussian MAC regardless of the value of delay. The scheme consists of rate splitting at the encoders and embedding preamble intervals of data at the beginning of each codeword so that the receiver can estimate the delay between the two users' transmission blocks [70]. The collision MAC without feedback is introduced in [71] where it is shown how users can jointly design their transmitters in order to reliably communicate in the presence of unknown delays between any two users' transmission blocks whether the users are synchronous at the symbol level or not.

An approach is taken in [72] to study a *centralized* interference channel with block-asynchronous users. Using the general formula for capacity proposed in [73] and based on the concept of information density, the authors derive an expression for the capacity region of such networks. This expression is not a single-letter formulation of the capacity region and hence, it is not computable. By deriving a single-letter inner bound, the authors present an analysis of the block-asynchronous two-user interference channel where the main focus is to show that using Gaussian codewords is in general suboptimal.

The schemes proposed in [68, 69, 70, 71, 72] can not be applied to a *decentralized* interference channel with block-asynchronous users based on the model considered in the current article. This is mainly due to the fact that in general joint transmitter design is not possible in decentralized scenarios and users are unaware of each other's code-books. Moreover, [68, 69, 70, 71, 72] assume that the users have always enough data in their queues for transmission. This results in a similar delay pattern between consecutive transmission blocks of the users. However, if users do not have enough data for transmission, their transmission blocks look like scattered signal bursts along the time axis. This results in different values for the mutual delay between any two overlapping transmission blocks of the users. Therefore, it is of interest to devise an efficient and low-complexity resource allocation scheme in a two-user decentralized interference channel in which the users are block-asynchronous and the rate of data arrival at the encoders is relatively low, i.e., the transmitters do not have always enough data for transmission.

1.2 Contributions

1.2.1 Chapter 2

In chapter 2, we consider a decentralized wireless communication network with a fixed number u of frequency sub-bands to be shared among N transmitter-receiver pairs. It is assumed that the number of active users is a realization of a random variable with a given probability mass function. Moreover, users are unaware of each other's codebooks, and hence, no multiuser detection is possible. We propose a randomized Frequency Hopping scheme in which the i^{th} transmitter randomly hops over v_i (the hopping parameter) out of u sub-bands from transmission slot to transmission slot. Assuming i.i.d. Gaussian signals are transmitted over the chosen sub-bands, the distribution of the noise plus interference becomes mixed Gaussian, which makes the calculation of the achievable rate complicated. The main contributions of this chaper are presented in two different cases:

Case I: Transmitters are ***aware*** *of the number of active users and the channel gains.*

• We derive lower and upper bounds on the mutual information between the transmitted and received signals of each user and demonstrate that, for large SNR values, the two bounds coincide. Thereafter, we are able to show that the achievable rate of the i^{th} user scales like $\frac{v_i}{2}\prod_{\substack{j=1\\j\neq i}}^{N}\left(1-\frac{v_j}{u}\right)\log\mathrm{SNR}$.

• We show that each transmitter only needs the knowledge of the number of active users in the network, the forward channel gain and the maximum interference level at its associated receiver to regulate its transmission rate. Knowing these quantities, we demonstrate how the i^{th} user can achieve a multiplexing gain of $\frac{v_i}{2}\prod_{\substack{j=1\\j\neq i}}^{N}\left(1-\frac{v_j}{u}\right)$.

• We obtain the optimum design parameters $\{v_i\}_{i=1}^{N}$ in order to maximize various performance measures.

• We compare the performance of the FH with that of the Frequency Division in terms of the following performance measures: average sum multiplexing gain ($\eta^{(1)}$) and average minimum multiplexing gain per user ($\eta^{(2)}$). We show that (depending on the probability mass function of the number of active users) the FH system can offer a significant improvement in terms of $\eta^{(1)}$ and $\eta^{(2)}$ (implying a more efficient usage of the spectrum).

*Case II: Transmitters are **unaware** of the number of active users and the channel gains.*

• Developing a new upper bound on the differential entropy of a class of mixed Gaussian random vectors and using entropy power inequality, we offer three lower bounds on the ε-outage capacity of each user denoted by $R_{\mathrm{FH}}^{(1)}(\varepsilon)$, $R_{\mathrm{FH}}^{(2)}(\varepsilon)$, and $R_{\mathrm{FH}}^{(3)}(\varepsilon)$. To evaluate the system performance analytically, we use $R_{\mathrm{FH}}^{(3)}(\varepsilon)$, which can be computed easily. However, computation of $R_{\mathrm{FH}}^{(1)}(\varepsilon)$ and $R_{\mathrm{FH}}^{(2)}(\varepsilon)$ involves integrations that cannot be carried out in a closed form. In the simulation results, we use the lower bounds $R_{\mathrm{FH}}^{(1)}(\varepsilon)$ and $R_{\mathrm{FH}}^{(2)}(\varepsilon)$, which are tighter than $R_{\mathrm{FH}}^{(3)}(\varepsilon)$.

• We perform asymptotic analysis for the outage capacity in terms of ε and SNR. In the asymptotically small ε regime, we observe that the maximum of outage capacity is obtained when the hopping parameter for each user is either 1 or u. In the asymptotically small SNR regime, we demonstrate that the system achieves the optimal performance regardless of the value of the hopping parameter. For asymptotically large values of SNR, it is shown that the optimal hopping parameter is $\left\lceil \frac{u}{n_{\max}} \right\rceil$, where $n_{\max}$ is the maximum possible number of concurrently active users in the network.

• We compare the outage capacity of the underlying FH scenario with that of the FD scheme for various scenarios in terms of distributions on the number of active users, SNR and ε. It is shown that FH outperforms FD in terms of outage capacity in many cases. We observe that in the low SNR regime, FH and FD offer the same performance. In the low ε regime, FD is always better than FH, however, for many practical scenarios there exists a threshold, $\varepsilon_{\mathrm{th}}$, such that FH outperforms FD as far as $\varepsilon \geq \varepsilon_{\mathrm{th}}$. Also, we have shown that supremacy of FH over FD in the high

SNR regime occurs quite often.

1.2.2 Chapter 3

In chapter 3, we focus on a decentralized network of N transmitter-receiver pairs. It is assumed that different transmitters are linked to different receivers through channels with static and non-frequency selective gains that are realizations of i.i.d. complex Gaussian random variables representing Rayleigh fading. In a decentralized interference channel the knowledge of users over the channel gains is only partial. In fact, each user is only aware of the channel gains from all transmitters to its own receiver. Motivated by random spreading in CDMA, we assume that the i^{th} user generates two groups of code-books, referred to as the signal code-book and signature code-book with rates R_i and R'_i, respectively. The signal code-book is a collection of $2^{\lfloor TR_i \rfloor}$ codewords where each codeword consists of T i.i.d. Gaussian random variables $\boldsymbol{x}_i[0], \boldsymbol{x}_i[1], \cdots, \boldsymbol{x}_i[T-1]$. The signature code-book is a collection of $2^{\lfloor TR'_i \rfloor}$ codewords where each codeword consists of T i.i.d. random vectors $\vec{\boldsymbol{s}}_i[0], \vec{\boldsymbol{s}}_i[1], \cdots, \vec{\boldsymbol{s}}_i[T-1]$ with equal length K called signature vectors. The elements in each signature vector are independently generated based on a globally known Probability Mass Function (PMF) over a finite alphabet $\mathcal{A}$. Letting each user have $2^{\lfloor TR_i \rfloor + \lfloor TR'_i \rfloor}$ possible messages for transmission (message splitting), corresponding signal and signature codewords are picked and the sequence of vectors $\boldsymbol{x}_i[0]\vec{\boldsymbol{s}}_i[0], \boldsymbol{x}_i[1]\vec{\boldsymbol{s}}_i[1], \cdots, \boldsymbol{x}_i[T-1]\vec{\boldsymbol{s}}_i[T-1]$ is transmitted. As an example, let $K = 8$, the underlying alphabet for the signatures be $\mathcal{A} = \{-1, 0, +1\}$, a typical symbol in the signal codeword be $\boldsymbol{x}_i$ and the corresponding signature vector in the signature codeword be

$$\vec{\boldsymbol{s}}_i = \begin{pmatrix} 1 & -1 & 0 & 1 & 0 & -1 & 1 & 1 \end{pmatrix}^{\text{t}}.$$

Thereafter, the vector

$$\boldsymbol{x}_i\vec{\boldsymbol{s}}_i = \begin{pmatrix} \boldsymbol{x}_i & -\boldsymbol{x}_i & 0 & \boldsymbol{x}_i & 0 & -\boldsymbol{x}_i & \boldsymbol{x}_i & \boldsymbol{x}_i \end{pmatrix}^{\text{t}}$$

is transmitted in 8 consecutive transmission slots.

Let us denote the received vector in the K transmission slots during which $\boldsymbol{x}_i\vec{\boldsymbol{s}}_i$ is transmitted by $\vec{\boldsymbol{y}}_i$. Then, any transmission rate $R_i + R'_i < \frac{\mathrm{I}(\boldsymbol{x}_i, \vec{\boldsymbol{s}}_i; \vec{\boldsymbol{y}}_i)}{K}$ is achievable in the conventional sense by applying typical decoding [80]. To calculate the quantity $\frac{\mathrm{I}(\boldsymbol{x}_i, \vec{\boldsymbol{s}}_i; \vec{\boldsymbol{y}}_i)}{K}$, the i^{th} user needs to estimate the number of users and the gains of the channels connecting the transmitters to its receiver and send these data to its affiliated transmitter through a feedback link. The i^{th} user aims to design K and the underlying PMF used to generate the signature vectors such that $\mathrm{E}\left\{\frac{\mathrm{I}(\boldsymbol{x}_i, \vec{\boldsymbol{s}}_i; \vec{\boldsymbol{y}}_i)}{K}\right\}$ is as large as possible. Here, the expectation is with respect to the channel gains. To motivate the use of $\mathrm{E}\{.\}$, note that in decentralized networks the knowledge of the users about the channel gains is only partial due to the lack of a central controller or cooperation among users. For example, in a network of $N = 3$ users, it is reasonable to assume that user 1 is aware of the channel gains from transmitters 2 and 3 to receiver 1, however, user 1 is completely unaware of the channel gain from transmitter 2 to receiver 3. Since the quantity $\mathrm{I}(\boldsymbol{x}_i, \vec{\boldsymbol{s}}_i; \vec{\boldsymbol{y}}_i)$ depends on the gains of the channels connecting all transmitters to the receiver of the i^{th} user, maximizing $\frac{\mathrm{I}(\boldsymbol{x}_i, \vec{\boldsymbol{s}}_i; \vec{\boldsymbol{y}}_i)}{K}$ by this user results in possibly different solutions for K and the underlying PMF compared to the solutions obtained for these parameters by other users. Assuming all the channel gains are realizations of i.i.d. complex Gaussian random variables, the quantity $\mathrm{E}\left\{\frac{\mathrm{I}(\boldsymbol{x}_i, \vec{\boldsymbol{s}}_i; \vec{\boldsymbol{y}}_i)}{K}\right\}$ does not depend on the index i and it can be used as the utility function to be maximized by any user. In fact, optimizing $\mathrm{E}\left\{\frac{\mathrm{I}(\boldsymbol{x}_i, \vec{\boldsymbol{s}}_i; \vec{\boldsymbol{y}}_i)}{K}\right\}$ leads to a consistent result for all users, i.e., all users come up with the same set of values for the design parameters.

We are primarily interested in system design at a finite SNR level. Since different users are unaware of each other's signature vectors, the noise plus interference at each receiver is mixed Gaussian. Appropriate tools such as conditional entropy power inequality and a key upper bound on the differential entropy of a mixed Gaussian vector are applied to develop a lower bound on $\mathrm{I}(\boldsymbol{x}_i, \vec{\boldsymbol{s}}_i; \vec{\boldsymbol{y}}_i)$. This lower bound is tight enough to guarantee the same SNR scaling as that of $\mathrm{I}(\boldsymbol{x}_i, \vec{\boldsymbol{s}}_i; \vec{\boldsymbol{y}}_i)$ and a similar behavior in medium ranges of SNR as that of $\mathrm{I}(\boldsymbol{x}_i, \vec{\boldsymbol{s}}_i; \vec{\boldsymbol{y}}_i)$. The latter is quite essential as it states that the derived lower bound on $\mathrm{I}(\boldsymbol{x}_i, \vec{\boldsymbol{s}}_i; \vec{\boldsymbol{y}}_i)$ mimics the plot of $\mathrm{I}(\boldsymbol{x}_i, \vec{\boldsymbol{s}}_i; \vec{\boldsymbol{y}}_i)$ as a function of the design parameters. This enables us to design the signature code-books based on the proposed lower bound.

Comparison With Prior Art

Our scheme is an instance of random spreading [19] in CDMA systems with the difference that the spreading codes (signature vectors in our setup) also carry some information to the corresponding receivers. It is long known in the context of CDMA-based Gaussian MAC that if the signature codes of the users are $\vec{\boldsymbol{s}}_1, \cdots, \vec{\boldsymbol{s}}_N$, the sum capacity of the network drops unless the matrix $\left[\vec{\boldsymbol{s}}_1|\cdots|\vec{\boldsymbol{s}}_N\right]$ has orthogonal rows [74]. This condition is certainly not satisfied in a randomly spread CDMA-based Gaussian MAC. Hence, one expects that random spreading only achieves a fraction of the sum-capacity of a Gaussian MAC as shown in [19].

The analysis given in [19] considers the asymptotic case of $N, K \to \infty$, while the fraction $\frac{N}{K}$ is a constant. It is shown that the spectral efficiency of the network in the randomly spread CDMA depends on the empirical eigenvalue distribution of random matrices with particular structures. Invoking results from random matrix theory [81] in the mentioned asymptotic case, this empirical distribution tends to a deterministic distribution that makes the analysis tractable. In this work, we are interested in the behavior of a decentralized network with any finite number of users. As such, due to the finiteness of the number of active users, the Central Limit Theorem and its variants such as Lindeberg-type Theorems [89] that are used in [19] are not applicable in our setup. Although the emphasis is on finite N, we will also examine several asymptotic regimes.

The focus of [19] is on a MAC where the receiver is aware of the signatures of all users. The signature of any user is randomly generated to spread a symbol in its codeword, where the common receiver is assumed to be aware of the signatures of all users. However, in this work the focus is on a decentralized interference channel where any user is unaware of the signatures of other users. Moreover, the sequence of signatures used by any single user is already designed as part of its code-book and it carries some information to the corresponding receiver. Since each signature vector is a random vector with a finite number of possible realizations, the noise plus interference at any receiver is a mixed Gaussian vector. As mentioned before, this makes $\mathrm{I}(\boldsymbol{x}_i, \vec{\boldsymbol{s}}_i; \vec{\boldsymbol{y}}_i)$ have no closed expression.

Note that using the minimum distance decoder over a channel with additive non-Gaussian noise is generally suboptimal. In fact, it is shown in [76] that Nearest Neighbor Decoding (NND) over an additive white complex non-Gaussian noise channel, referred to as mismatched decoding, does not achieve any rate beyond $\mathscr{R}_{\mathrm{NND}} \triangleq \log\left(1 + \frac{P}{\sigma^2}\right)$ where P is the average transmission power and $2\sigma^2$ is the noise spectral density. Treating interference as noise in our setup, σ^2 is proportional to P, and hence, $\mathscr{R}_{\mathrm{NND}}$ saturates as the average transmission power tends to infinity. This in particular shows that using the minimum distance decoder achieves no multiplexing gain in a decentralized network where users apply the signaling proposed in this chapter. However, we will show that $\frac{\mathrm{I}(\boldsymbol{x}, \vec{\boldsymbol{s}}; \vec{\boldsymbol{y}})}{K}$ scales like $\left(\frac{1}{N} - \delta\right) \log \mathrm{SNR}$ in the high SNR regime (for any finite number of users N and arbitrary $\delta > 0$) by designing the signature vectors appropriately. This is guaranteed as far as a jointly typical decoder is adopted that is matched with the mixed Gaussian PDF of the noise plus interference.

Main Observations

In the sequel, we are able to make several observations as follows:

Observation 1- *There exists a tradeoff between the achievable rates in high and medium SNR:*

Investigation of the lower bound reveals a tradeoff between the value of the achievable rate in medium SNR and the multiplexing gain per user. In fact, the developed lower bound on the achievable rate of each user can be roughly expressed as $\mathrm{MG} \log \mathrm{SNR} - \mathrm{IEF}$ where MG is the multiplexing gain per user and IEF is what we call the interference entropy factor. This factor depends on the discrete entropy of a deterministic function of the signature vectors of the interferers. For any two different alphabets $\mathcal{A}_1$ and $\mathcal{A}_2$ that are used to generate the signatures of users, if $\mathrm{MG}_{\mathcal{A}_1} > \mathrm{MG}_{\mathcal{A}_2}$, then $\mathrm{IEF}_{\mathcal{A}_1} > \mathrm{IEF}_{\mathcal{A}_2}$ as well. This implies that choosing $\mathcal{A}_2$ is more appropriate for relatively smaller values of SNR, while choosing $\mathcal{A}_1$ results in larger values for the lower bound at relatively larger SNR values. This is illustrated in fig. 1.2. Simulation results confirm that this

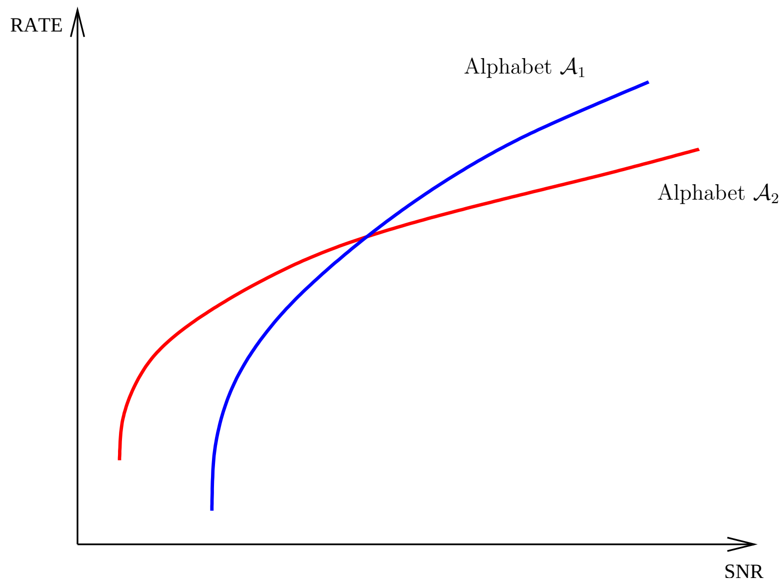

Figure 1.2: The tradeoff between the slope (high SNR) and the values (medium SNR) of the rate.

behavior is a fundamental property of the achievable rate, i.e., the observation based on the lower bound on $\frac{\mathrm{I}(\boldsymbol{x}_i;\vec{\boldsymbol{y}}_i|\vec{\boldsymbol{s}}_i)}{K}$ holds for $\frac{\mathrm{I}(\boldsymbol{x}_i;\vec{\boldsymbol{y}}_i|\vec{\boldsymbol{s}}_i)}{K}$ as well.

Observation 2- *The elements of the signature vectors are not necessarily equiprobable over the underlying alphabet:*

A surprising observation is that the designed PMF used to generate the signature vectors is not necessarily uniform. For example, let the underlying alphabet be $\{-1, +1\}$ and the number of users be $N = 4$. We show that the probabilities of $+1$ and -1 that maximize the achievable rate of each user are not equal to $\frac{1}{2}$, i.e., the elements of the signatures are not equiprobable over the underlying alphabet. This is in contrast to the customary use of binary PN signatures in conventional CDMA systems where the chip elements are $+1$ or -1 with equal probability. This observation is confirmed by calculating the exact value of the achievable rate for each user through simulations. Once again, the developed lower bound on the achievable rate of each user is shown to mimic the lower bound itself, i.e., it is maximized by letting $+1$ and -1 have unequal probabilities. The main reason for this phenomenon is the interplay between the factors MG and IEF described in the preceding part.

Observation 3- *It is possible to achieve the SMG of orthogonal schemes in decentralized networks:*

One may consider a generalized version of our approach where each signature vector is randomly generated to be one of the vectors in a globally known signature-book $\mathcal{C}$. For example, if the underlying alphabet is $\{-1, +1\}$, our signaling scheme reduces to selecting signatures in the set of all K-tuples whose elements are either 1 or -1. Assuming the signatures in $\mathcal{C}$ are selected with equal probability, we show that the SMG of the network is bounded from above by $\left(1 - \frac{1}{|\mathcal{C}|}\right)^{N-1}$ where $|\mathcal{C}|$ denotes the size of $\mathcal{C}$. Moreover, given any natural number L, we construct a signature-book $\mathcal{C}_L$ of size L such that an SMG of $\left(1 - \frac{1}{L}\right)^{N-1}$ is in fact achievable. Therefore, regardless of the number of users an SMG of 1 is approachable by increasing L and using the right signature-book $\mathcal{C}_L$. This is the SMG of an orthogonal signaling scheme such as time division multiplexing. However, the choice of $\mathcal{C}_L$ is not unique. This implies that users are not able to agree upon $\mathcal{C}_L$ without any coordination, i.e., they must somehow agree on a global choice of the optimum $\mathcal{C}_L$.

Observation 4- *The SMG of a decentralized network with matched filtering at the receivers and hence, the SMG of random spreading CDMA-based MAC with matched filtering, is nonzero as N tends to infinity:*

We offer an analysis of the multiplexing gain if the receivers are equipped with suboptimal detectors using matched filters. This detection method is widely used in conventional CDMA system due to its low complexity [19]. In this approach, each receiver first decodes the sequence of signature vectors $\vec{\boldsymbol{s}}_i[0], \vec{\boldsymbol{s}}_i[1], \cdots, \vec{\boldsymbol{s}}_i[T-1]$ and decides on the sequence of signals $\boldsymbol{x}_i[0], \boldsymbol{x}_i[1], \cdots, \boldsymbol{x}_i[T-1]$ based on the inner products $\vec{\boldsymbol{s}}_i^\dagger[0]\vec{\boldsymbol{y}}_i[0], \vec{\boldsymbol{s}}^\dagger[1]\vec{\boldsymbol{y}}_i[1], \cdots, \vec{\boldsymbol{s}}^\dagger[T-1]\vec{\boldsymbol{y}}_i[T-1]$. In the special case that the signature vectors are generated over $\{-1, +1\}$, we demonstrate that the elements of the signatures must not necessarily be $+1$ or -1 with equal probabilities. This particularly holds if the number of users is larger than or equal to 5. It is shown that if all users select the right values for the probabilities of $+1$ and -1, the SMG of the network settles on 0.182 as the number of users tends to infinity. However, if all users construct their signatures by selecting each element in the signature to be $+1$ or -1 with equal probabilities, the SMG in the network tends to 0 as the number of users increases. Moreover, the optimum value for the length K of signatures is $K = 2$ regardless of the value of N. The authors in [19] show that the sum rate in a randomly spread CDMA-based MAC with matched filters utilized at the receiver side saturates in the asymptotic case where N, K and SNR tend to infinity while the ratio $\frac{N}{K}$ is a constant. However, our results demonstrate that by appropriately selecting the probabilities of $+1$ and -1 (not equal to $\frac{1}{2}$) and setting $K = 2$, the sum rate in a randomly spread CDMA-based MAC exploiting matched filters scales with $\log \mathrm{SNR}$ instead of saturating.

1.2.3 Chapter 4

A network of two transmitter-receiver pairs is considered where the channel from each transmitter to each receiver is modeled by static and non-frequency selective Rayleigh fading. The network is decentralized, i.e., there is no central controller to assign the resources to the users and users do not explicitly cooperate. In particular, each user is unaware of the other user's code-book. Hence, multiuser detection is not possible at the receivers. We consider a scenario where the rate of data arrival at the encoders is low such that the codewords of each user are transmitted at random instants depending on the availability of data for transmission. As such, the transmitted signals by each user look like scattered bursts along the time axis. Moreover, the users are block-asynchronous meaning there exists a mutual delay between their transmitted codewords. There are two main reasons for asynchrony:

1- Users do not become active simultaneously, i.e., the two queues start to accept data at different instants. This entails a transient phase where one user is active and the other user becomes active at a later time. Although we are not interested in this transient period and we focus on a phase where both users are active, it is important to note that the length of the transient phase is in general a random variable which adds to the randomness of delay between the transmitted codewords of the two users.

2- The data streams entering the queues of the two users are independent. In fact, if one user has enough data in its queue for transmission, the other user may be still waiting to receive more data and will transmit at a later instant.

The proposed model for asynchrony assumes the starting point of an interference burst is uniformly distributed along the transmitted codeword by each user. Also, users may not experience any interference on a transmitted codeword at all.

Due to the randomness of delay, the communication channel for each user is non-ergodic in the sense that the transmitters are unaware of the starting point of interference bursts. Outage analysis is an appropriate computational tool to study the performance in this setup. Aside from the delay pattern, it is also assumed that the transmitters are unaware of the channel gains in the network. However, each receiver is able to measure (without error) the channel gains and the mutual delay by the end of the transmission of its corresponding codeword.

Following the randomized resource allocation scenario proposed in chapter 2, we adopt a strategy where each user transmits a Gaussian symbol in its codeword with a probability of θ and quits transmitting with a probability of $1-\theta$ independently from symbol interval to symbol interval in a slotted channel[2]. We call this scheme Randomized Masking (RM) with activity factor θ. For example, let us consider the scenario in fig. 1.1, and assume that both encoders randomly mask their outputs. This is demonstrated in fig. 1.3 where the masked symbol intervals are shown in black.

As no user is aware of the on-off pattern of the other user, the noise plus interference has a mixed Gaussian distribution. As a result, one can not find a closed expression for the mutual information between the channel input and channel output of any user. Hence, it is not possible to derive

[2]In fact, we assume the users are synchronous at the symbol level, i.e., they are symbol-synchronized. Throughout the chapter, asynchrony only exists at the block level.

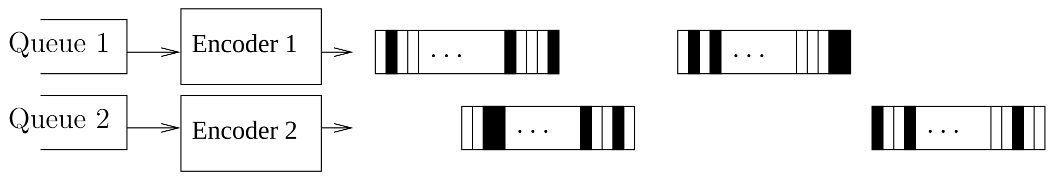

Figure 1.3: Each user masks its output independently from symbol interval to symbol interval.

a closed expression for the probability of outage. Instead, an upper bound on the probability of outage is developed using entropy power inequality together with a key upper bound on the differential entropy of a mixed Gaussian random variable. In the sequel, the activity factor θ is designed to minimize the upper bound on the probability of outage. This upper bound has the following properties:

• It is tight at $\theta = 1$, i.e., setting $\theta = 1$ in its expression yields the exact probability of outage for the case when masking is not applied.

• The activity factor θ that minimizes the upper bound is in general less than 1.

Therefore, it is shown that the proposed RM strategy reduces the probability of outage compared to the case where all transmitters keep transmitting their Gaussian signals in consecutive symbol intervals (no masking applied), referred to as Continuous Transmission (CT). We emphasize that even in the CT scheme, in case the users do not have enough data in their queues, they remain silent and wait for more data to arrive. Hence, CT must not be interpreted as a scenario where users have always data for transmisson.

Chapter 2

Randomized Frequency Hopping

2.1 System Model and Assumptions

We consider a wireless network with N users[1] operating over a spectrum consisting of u orthogonal sub-bands. The number of active users is assumed to be a realization of a random variable with a given distribution, however, it is fixed during the whole transmission once it is set first. The transmission blocks of each user comprise of an arbitrarily large number of transmission slots. We remark that the results of this chapter are valid regardless of having block synchronization among the users, however, we assume synchronization at the symbol level. The i^{th} user exploits $v_i(\leq u)$ out of the u sub-bands in each transmission slot and hops randomly to another set of v_i frequency sub-bands in the next transmission slot. This user transmits independent real Gaussian signals of variance $\frac{P}{v_i}$ over the chosen sub-bands, in which P denotes the total average power for each transmitter. Each receiver is assumed to know the hopping pattern of its affiliated transmitter. It is assumed that the users are not aware of each other's codebooks and hence, no multiuser detection or interference cancelation is possible at the receiver sides. The static and non-frequency selective channel gain of the link connecting the i^{th} transmitter to the j^{th} receiver is shown by $h_{i,j}$. As it will be shown in (2.50), the only information each transmitter needs in order to regulate its transmission rate (focusing on the achieved multiplexing gain) is its forward channel gain, the maximum interference level at its associated receiver and the number of active users in the network. This information can be obtained at the receiver side and provided to the corresponding transmitter via a feedback link[2].

As all users hop over different portions of the spectrum from transmission slot to transmission slot, no user is assumed to be capable of tracking the instantaneous interference level. This assump-

[1]Each user consists of a separate transmitter-receiver pair.

[2]This is addressed in section 2.4.

tion makes the interference plus noise PDF at the receiver side of each user be mixed Gaussian. In fact, depending on different choices the other users make to select the frequency sub-bands and values of the crossover gains, the interference on each frequency sub-band at any given receiver can have up to 2^{N-1} power levels. The vector consisting of the received signals on the frequency sub-bands at the i^{th} receiver in a typical transmission slot is

$$\vec{\boldsymbol{y}}_i = h_{i,i}\vec{\boldsymbol{x}}_i + \vec{\boldsymbol{z}}_i, \tag{2.1}$$

where $\vec{\boldsymbol{x}}_i$ is the $u \times 1$ transmitted vector and $\vec{\boldsymbol{z}}_i$ is the noise plus interference vector at the receiver side of the i^{th} user. One may write $p_{\vec{\boldsymbol{x}}_i}(.)$ as

$$p_{\vec{\boldsymbol{x}}_i}(\vec{x}) = \sum_{C \in \mathcal{C}} \frac{1}{\binom{u}{v_i}} g_u(\vec{x}, C), \tag{2.2}$$

which corresponds to the mixed Gaussian distribution. In the above equation, the set $\mathcal{C}$ includes all $u \times u$ diagonal matrices in which v_i out of the u diagonal elements are $\frac{P}{v_i}$ and the rest are zero. Denoting the noise plus interference on the j^{th} sub-band at the receiver side of the i^{th} user by $\boldsymbol{z}_{i,j}$ (the j^{th} component of $\vec{\boldsymbol{z}}_i$), it is clear that $p_{\boldsymbol{z}_{i,j}}(.)$ is not dependent on j. This is due to the fact that the crossover gains are not frequency selective and there is no particular interest in a specific frequency sub-band by any user. We assume there are $L_i + 1$ ($L_i \leq 2^{N-1} - 1$) possible non-zero power levels for $\boldsymbol{z}_{i,j}$, say $\{\sigma_{i,l}^2\}_{l=0}^{L_i}$. Denoting the occurrence probability of $\sigma_{i,l}^2$ by $a_{i,l}$, $p_{\boldsymbol{z}_{i,j}}(.)$ identifies a mixed Gaussian PDF as

$$p_{\boldsymbol{z}_{i,j}}(z) = \sum_{l=0}^{L_i} \frac{a_{i,l}}{\sqrt{2\pi}\sigma_{i,l}} \exp\left(-\frac{z^2}{2\sigma_{i,l}^2}\right), \tag{2.3}$$

where $\sigma^2 = \sigma_{i,0}^2 < \sigma_{i,1}^2 < \sigma_{i,2}^2 < ... < \sigma_{i,L_i}^2$ (σ^2 is the ambient noise power). We notice that for each $l \geq 0$, there exists a $\zeta_{i,l} \geq 0$ such that $\sigma_{i,l}^2 = \sigma^2 + \zeta_{i,l}P$ where $0 = \zeta_{i,0} < \zeta_{i,1} < \zeta_{i,2} < ... < \zeta_{i,L_i}$. One may write

$$\boldsymbol{z}_{i,j} = \sum_{\substack{k=1 \\ k \neq i}}^{N} h_{k,i}\boldsymbol{c}_{k,j}\boldsymbol{x}_{k,j} + \boldsymbol{\nu}_{i,j} \tag{2.4}$$

where $\boldsymbol{c}_{k,j}$ is a Bernoulli random variable showing if the k^{th} user has utilized the j^{th} sub-band, $\boldsymbol{x}_{k,j}$ is the signal of the k^{th} user sent on the j^{th} sub-band (assuming it has utilized the j^{th} sub-band), and $\boldsymbol{\nu}_{i,j}$ is the ambient noise which is a zero-mean Gaussian random variable with variance σ^2. The ratio $\frac{P}{\sigma^2}$ is taken as a measure of SNR and is denoted by γ throughout the chapter.

2.2 Analysis of the Achievable Rate

Let us denote the achievable rate of the i^{th} user by $\mathscr{R}_i$. It can be observed that the communication channel of this user is a channel with state $\boldsymbol{s}_i$, the hopping pattern of the i^{th} user, which is independently changing over different transmission slots, and is known to both the transmitter and the receiver. The achievable rate of such a channel is given by

$$\mathscr{R}_i = \mathrm{I}(\vec{\boldsymbol{x}}_i; \vec{\boldsymbol{y}}_i | \boldsymbol{s}_i) = \sum_{s_i \in \mathcal{S}_i} \Pr(\boldsymbol{s}_i = s_i) \mathrm{I}(\vec{\boldsymbol{x}}_i; \vec{\boldsymbol{y}}_i | \boldsymbol{s}_i = s_i), \tag{2.5}$$

where $\mathrm{I}(\vec{\boldsymbol{x}}_i; \vec{\boldsymbol{y}}_i | \boldsymbol{s}_i = s_i)$ is the mutual information between $\vec{\boldsymbol{x}}_i$ and $\vec{\boldsymbol{y}}_i$ for the specific sub-band selection corresponding to $\boldsymbol{s}_i = s_i$. The set $\mathcal{S}_i$ denotes all possible selections of v_i out of the u sub-bands. As $p_{\vec{\boldsymbol{z}}_i}(.)$ is a symmetric density function, meaning all its components have the same PDF given in (2.3), we deduce that $\mathrm{I}(\vec{\boldsymbol{x}}_i; \vec{\boldsymbol{y}}_i | \boldsymbol{s}_i = s_i)$ is independent of s_i. Therefore, to calculate $\mathscr{R}_i$, we may assume any specific sub-band selection for the i^{th} user in $\mathcal{S}_i$, say the first v_i sub-bands. Denoting this specific state by s_i^*, we get

$$\mathscr{R}_i = \mathrm{I}(\vec{\boldsymbol{x}}_i; \vec{\boldsymbol{y}}_i | \boldsymbol{s}_i = s_i^*). \tag{2.6}$$

In this case, we denote $\vec{\boldsymbol{y}}_i$ and $\vec{\boldsymbol{x}}_i$ by $\vec{\boldsymbol{y}}_i(s_i^*)$ and $\vec{\boldsymbol{x}}_i(s_i^*)$, respectively. Obviously, we have

$$\mathscr{R}_i = \mathrm{I}(\vec{\boldsymbol{x}}_i(s_i^*); \vec{\boldsymbol{y}}_i(s_i^*)) = \mathrm{h}(\vec{\boldsymbol{y}}_i(s_i^*)) - \mathrm{h}(\vec{\boldsymbol{z}}_i). \tag{2.7}$$

Because $\vec{\boldsymbol{y}}_i(s_i^*)$ and $\vec{\boldsymbol{z}}_i$ have mixed Gaussian distributions, there is no closed expression for the differential entropy of these vectors. As such, we provide an upper bound and a lower bound on the achievable rate of each user in the following subsections and show that these bounds coincide in the asymptotically high SNR regime.

2.2.1 Upper Bound on The Achievable Rates

In this section, we develop an upper bound $\mathscr{R}_i^{(\text{ub})}$ on the achievable rate of the i^{th} user that is tight enough to ensure that $\mathscr{R}_i^{(\text{ub})} - \mathscr{R}_i$ does not increase unboundedly as SNR increases. The idea behind this upper bound is the convexity of $\mathscr{R}_i$ in terms of $p_{\vec{\boldsymbol{y}}_i(s_i^*)|\vec{\boldsymbol{x}}_i(s_i^*)}(.|.)$.

Theorem 1. *There exists an upper bound on the achievable rate of the i^{th} user given by*

$$\mathscr{R}_i^{(\text{ub})} = \frac{1}{2} v_i \prod_{\substack{k=1 \\ k \neq i}}^{N} \left(1 - \frac{v_k}{u}\right) \log\left(1 + \frac{|h_{i,i}|^2 \gamma}{v_i}\right) + \widetilde{\mathscr{R}}_i^{(\text{ub})} \tag{2.8}$$

where $\lim_{\gamma \to \infty} \widetilde{\mathscr{R}}_i^{(\text{ub})} < \infty$, *i.e.,*

$$\mathscr{R}_i^{(\text{ub})} \overset{\gamma}{\sim} \frac{1}{2} v_i \prod_{\substack{k=1 \\ k \neq i}}^{N} \left(1 - \frac{v_k}{u}\right) \log \gamma. \tag{2.9}$$

Proof. Let $\vec{\boldsymbol{w}}_i$ be the $u \times 1$ interference vector where its j^{th} component $\boldsymbol{w}_{i,j}$ is a random variable showing the interference term on the j^{th} frequency sub-band at the receiver of the i^{th} user. We have $\boldsymbol{w}_{i,j} = \sum_{\substack{k=1\\k\neq i}}^{N} h_{k,i}\boldsymbol{c}_{k,j}\boldsymbol{x}_{k,j}$. Clearly, $\vec{\boldsymbol{w}}_i$ is a mixed Gaussian random vector where the Gaussian components in its PDF represent different choices the other users make in selecting their sub-bands. In fact, we have $p_{\vec{\boldsymbol{w}}_i}(\vec{w}) = \frac{1}{M_i}\sum_{m=1}^{M_i} g_u(\vec{w}, D_{i,m})$, where $M_i = \prod_{k\neq i}\binom{u}{v_k}$ and $D_{i,m} = \operatorname{diag}(d_{i,m}^{(1)}, \cdots, d_{i,m}^{(u)})$, in which[3] $d_{i,m}^{(j)} = \sum_{\substack{k=1\\k\neq i}}^{N} |h_{k,i}|^2 c_{k,j,m}^2 \frac{P}{v_k}$ denotes the variance of $\boldsymbol{w}_{i,j}$ for the m^{th} realization of $\{\boldsymbol{c}_{k,j}\}_{k\neq i}$ out of M_i possible realizations that is denoted by $\{c_{k,j,m}\}_{k\neq i}$. If the probability density function of the interference vector consisted only of $g_u(., D_{i,m})$, the forward link of the i^{th} user would be converted into an additive Gaussian channel. The achievable rate of such a virtual channel is given by

$$\begin{aligned}
\mathscr{R}_{i,m} &= \frac{1}{2}\log\frac{\det\left(\operatorname{Cov}(\vec{\boldsymbol{x}}_i(s_i^*)) + D_{i,m} + \sigma^2 I_u\right)}{\det\left(D_{i,m} + \sigma^2 I_u\right)} \\
&= \frac{1}{2}\log\frac{\prod_{j=1}^{v_i}\left(\frac{|h_{i,i}|^2 P}{v_i} + d_{i,m}^{(j)} + \sigma^2\right)}{\prod_{j=1}^{v_i}(d_{i,m}^{(j)} + \sigma^2)} \\
&= \frac{1}{2}\sum_{j=1}^{v_i}\log\left(1 + \frac{|h_{i,i}|^2 P}{v_i(d_{i,m}^{(j)} + \sigma^2)}\right).
\end{aligned} \tag{2.10}$$

One may also state this as follows. Let

$$\mathcal{T}_{i,m} \triangleq \left\{j : 1 \leq j \leq v_i, d_{i,m}^{(j)} = 0\right\}. \tag{2.11}$$

Then,

$$\mathscr{R}_{i,m} = \frac{|\mathcal{T}_{i,m}|}{2}\log\left(1 + \frac{|h_{i,i}|^2\gamma}{v_i}\right) + \widetilde{\mathscr{R}}_{i,m}, \tag{2.12}$$

where

$$\widetilde{\mathscr{R}}_{i,m} = \frac{1}{2}\sum_{1\leq j\leq v_i : d_{i,m}^{(j)}\neq 0}\log\left(1 + \frac{|h_{i,i}|^2 P}{v_i(d_{i,m}^{(j)} + \sigma^2)}\right) \tag{2.13}$$

and $|\mathcal{T}_{i,m}|$ denotes the cardinality of the set $\mathcal{T}_{i,m}$. As each nonzero $d_{i,m}^{(j)}$ is proportional to P, it is clear that $\lim_{\gamma\to\infty}\widetilde{\mathscr{R}}_{i,m} < \infty$. We know that $\mathscr{R}_i$ is convex in terms of $p_{\vec{\boldsymbol{y}}_i(s_i^*)|\vec{\boldsymbol{x}}_i(s_i^*)}(\vec{y}|\vec{x}) = p_{\vec{\boldsymbol{z}}_i}(\vec{y} - h_{i,i}\vec{x})$

[3]Note that as each user transmits independent Gaussian signals over its chosen sub-bands, the matrices $\{D_{i,m}\}_{m=1}^{M_i}$ are diagonal.

[80]. Noting this and the fact that $p_{\vec{\boldsymbol{z}}_i}(\vec{z}) = \frac{1}{M_i}\sum_{m=1}^{M_i} g_u(\vec{z}, D_{i,m} + \sigma^2 I_u)$,

$$\begin{aligned} \mathscr{R}_i &\leq \frac{1}{M_i}\sum_{m=1}^{M_i} \mathscr{R}_{i,m} \\ &= \left(\frac{1}{M_i}\sum_{m=1}^{M_i} |\mathcal{T}_{i,m}|\right)\frac{1}{2}\log\left(1+\frac{|h_{i,i}|^2\gamma}{v_i}\right) + \widetilde{\mathscr{R}}_i^{(\mathrm{ub})}, \end{aligned} \tag{2.14}$$

where $\widetilde{\mathscr{R}}_i^{(\mathrm{ub})} = \frac{1}{M_i}\sum_{m=1}^{M_i} \widetilde{\mathscr{R}}_{i,m}$. As each $\widetilde{\mathscr{R}}_{i,m}$ saturates by increasing γ, one has $\lim_{\gamma\to\infty} \widetilde{\mathscr{R}}_i^{(\mathrm{ub})} < \infty$. Intuitively, $\frac{1}{M_i}\sum_{m=1}^{M_i} |\mathcal{T}_{i,m}|$ is the average size of the set of frequencies where the i^{th} user is transmitting and all other users are not transmitting. Since that the users choose the sub-bands independently, the probability of only the i^{th} user choosing a particular sub-band is given by $\frac{v_i}{u}\prod_{k\neq i}\left(1-\frac{v_k}{u}\right)$. Since there are a total of u sub-bands, the average size is $v_i\prod_{k\neq i}\left(1-\frac{v_k}{u}\right)$. This is proved formally in the following Lemma:

Lemma 1.

$$\frac{1}{M_i}\sum_{m=1}^{M_i} |\mathcal{T}_{i,m}| = v_i \prod_{\substack{k=1\\k\neq i}}^{N}\left(1-\frac{v_k}{u}\right). \tag{2.15}$$

Proof. Defining $\mathcal{A}_{i,j} \triangleq \{m : 1 \leq m \leq M_i, |\mathcal{T}_{i,m}| = j\}$ for each $1 \leq i \leq N$ and $1 \leq j \leq v_i$, one may express the left side of (2.15) as

$$\frac{1}{M_i}\sum_{m=1}^{M_i} |\mathcal{T}_{i,m}| = \frac{1}{M_i}\sum_{j=1}^{v_i} j|\mathcal{A}_{i,j}|. \tag{2.16}$$

Let $\boldsymbol{F}_i$ be a random variable showing the number of interference-free sub-bands among the v_i sub-bands selected by the i^{th} user. Using (2.16) and noting that $\Pr\{\boldsymbol{F}_i = j\} = \frac{|\mathcal{A}_{i,j}|}{M_i}$,

$$\frac{1}{M_i}\sum_{m=1}^{M_i} |\mathcal{T}_{i,m}| = \sum_{j=1}^{v_i} j\Pr\{\boldsymbol{F}_i = j\} = \mathrm{E}\{\boldsymbol{F}_i\}. \tag{2.17}$$

Let us define

$$\boldsymbol{F}_{i,j} \triangleq \begin{cases} 1 & \boldsymbol{W}_{i,j} = 0 \\ 0 & \boldsymbol{W}_{i,j} \neq 0 \end{cases} \tag{2.18}$$

for any $1 \leq i \leq N$ and $1 \leq j \leq v_i$. Obviously, $\boldsymbol{F}_i = \sum_{j=1}^{v_i} \boldsymbol{F}_{i,j}$. As such,

$$\mathrm{E}\{\boldsymbol{F}_i\} = \sum_{j=1}^{v_i} \mathrm{E}\{\boldsymbol{F}_{i,j}\} = \sum_{j=1}^{v_i} \Pr\{\boldsymbol{w}_{i,j} = 0\}. \tag{2.19}$$

Since $\Pr\{\boldsymbol{c}_{k,j} = 1\} = \frac{v_k}{u}$,

$$\Pr\{\boldsymbol{w}_{i,j} = 0\} = \Pr\{\boldsymbol{z}_{i,j} \text{ contains no interference}\} = \prod_{\substack{k=1 \\ k \neq i}}^{N} \Pr\{\boldsymbol{c}_{k,j} = 0\} = \prod_{\substack{k=1 \\ k \neq i}}^{N} \left(1 - \frac{v_k}{u}\right). \tag{2.20}$$

This yields

$$\mathrm{E}\{\boldsymbol{F}_i\} = v_i \prod_{\substack{k=1 \\ k \neq i}}^{N} \left(1 - \frac{v_k}{u}\right), \tag{2.21}$$

which completes the proof of Lemma 1. □

Using this in (2.14), the proof of Theorem 1 is complete. □

2.2.2 Lower Bound on the Achievable Rates

In this section, we derive a lower bound on the achievable rates of users. The idea behind deriving this lower bound is to invoke the classical Entropy Power Inequality (EPI). As we will see, this initial lower bound is not in a closed form as it depends on the differential entropy of a mixed Gaussian random variable. In appendix A, we obtain an appropriate upper bound on such an entropy that leads us to the final lower bound on $\mathscr{R}_i$.

Theorem 2. *There exists a lower bound $\mathscr{R}_i^{(\mathrm{lb})}$ on the achievable rate of the i^{th} user which can be written as*

$$\mathscr{R}_i^{(\mathrm{lb})} = \frac{1}{2} v_i \prod_{\substack{k=1 \\ k \neq i}}^{N} \left(1 - \frac{v_k}{u}\right) \log \gamma + \widetilde{\mathscr{R}}_i^{(\mathrm{lb})}, \tag{2.22}$$

such that $\lim_{\gamma \to \infty} \widetilde{\mathscr{R}}_i^{(\mathrm{lb})} < \infty$, *i.e.,*

$$\mathscr{R}_i^{(\mathrm{lb})} \overset{\gamma}{\sim} \frac{1}{2} v_i \prod_{\substack{k=1 \\ k \neq i}}^{N} \left(1 - \frac{v_k}{u}\right) \log \gamma. \tag{2.23}$$

Proof. We define $\vec{\boldsymbol{x}}_i'$ to be the $v_i \times 1$ signal vector corresponding to the first v_i elements of $\vec{\boldsymbol{x}}_i(s_i^*)$. Clearly, $\vec{\boldsymbol{x}}_i'$ is a Gaussian vector with covariance matrix $\frac{P}{v_i} I_{v_i}$. Let $\vec{\boldsymbol{y}}_i' = h_{i,i}\vec{\boldsymbol{x}}_i' + \vec{\boldsymbol{z}}_i'$ where $\vec{\boldsymbol{z}}_i'$ is the noise plus interference vector at the receiver side of the i^{th} user on the first v_i sub-bands. Using EPI,

$$2^{\frac{2}{v_i}\mathrm{h}(\vec{\boldsymbol{y}}_i')} \geq 2^{\frac{2}{v_i}\mathrm{h}(h_{i,i}\vec{\boldsymbol{x}}_i')} + 2^{\frac{2}{v_i}\mathrm{h}(\vec{\boldsymbol{z}}_i')}. \tag{2.24}$$

Dividing both sides by $2^{\frac{2}{v_i}\mathrm{h}(\vec{\boldsymbol{z}}'_i)}$, we get

$$\mathrm{h}(\vec{\boldsymbol{y}}'_i) - \mathrm{h}(\vec{\boldsymbol{z}}'_i) \geq \frac{v_i}{2}\log\left(2^{\frac{2}{v_i}\left(\mathrm{h}(h_{i,i}\vec{\boldsymbol{x}}'_i)-\mathrm{h}(\vec{\boldsymbol{z}}'_i)\right)} + 1\right). \tag{2.25}$$

On the other hand, since $\vec{\boldsymbol{y}}'_i$ is a subvector of $\vec{\boldsymbol{y}}_i(s_i^*)$, we have

$$\mathscr{R}_i = \mathrm{I}(\vec{\boldsymbol{x}}_i(s_i^*); \vec{\boldsymbol{y}}_i(s_i^*)) \geq \mathrm{I}(\vec{\boldsymbol{x}}'_i; \vec{\boldsymbol{y}}'_i) = \mathrm{h}(\vec{\boldsymbol{y}}'_i) - \mathrm{h}(\vec{\boldsymbol{z}}'_i). \tag{2.26}$$

Comparing (2.25) and (2.26) yields

$$\mathscr{R}_i \geq \frac{v_i}{2}\log\left(2^{\frac{2}{v_i}\left(\mathrm{h}(h_{i,i}\vec{\boldsymbol{x}}'_i)-\mathrm{h}(\vec{\boldsymbol{z}}'_i)\right)} + 1\right). \tag{2.27}$$

Clearly, $\mathrm{h}(h_{i,i}\vec{\boldsymbol{x}}'_i) = \frac{v_i}{2}\log\left(2\pi e\frac{|h_{i,i}|^2P}{v_i}\right)$. As $\vec{\boldsymbol{z}}'_i$ is a mixed Gaussian random vector, there is no closed formula for $\mathrm{h}(\vec{\boldsymbol{z}}'_i)$. Hence, we have to find an appropriate upper bound on $\mathrm{h}(\vec{\boldsymbol{z}}'_i)$ to further simplify (2.27). Using the chain rule for the differential entropy,

$$\mathrm{h}(\vec{\boldsymbol{z}}'_i) \leq \sum_{j=1}^{v_i}\mathrm{h}(\boldsymbol{z}_{i,j}). \tag{2.28}$$

Recalling the definitions of $\{a_{i,l}\}_{l=0}^{L_i}$ and $\{\zeta_{i,l}\}_{l=0}^{L_i}$ in the system model, the following Lemma yields an upper bound on $\mathrm{h}(\boldsymbol{z}_{i,j})$ for each $1 \leq j \leq v_i$.

Lemma 2. *For every $1 \leq j \leq v_i$ and for all values of γ, there exists an upper bound on $\mathrm{h}(\boldsymbol{z}_{i,j})$ given by*

$$\mathrm{h}(\boldsymbol{z}_{i,j}) \leq \frac{1-a_{i,0}}{2}\log(\zeta_{i,L_i}\gamma + 1) + \log(\sqrt{2\pi e}\sigma) - \sum_{l=0}^{L_i} a_{i,l}\log a_{i,l}. \tag{2.29}$$

Proof. In appendix A, we derive an upper bound on the differential entropy of a complex mixed Gaussian random variable in (5.11). Modifying the result in appendix A appropriately for the real mixed Gaussian random variable $\boldsymbol{z}_{i,j}$, we have

$$\mathrm{h}(\boldsymbol{z}_{i,j}) \leq \frac{1}{2}\sum_{l=0}^{L_i} a_{i,l}\log(2\pi e\sigma_{i,l}^2) - \sum_{l=0}^{L_i} a_{i,l}\log a_{i,l}. \tag{2.30}$$

Noting that $\sigma_{i,0} = \sigma$ and $\sigma_{i,l}^2 \leq \sigma^2(\zeta_{i,L_i}\gamma + 1)$ for any $1 \leq l \leq L_i$,

$$\mathrm{h}(\boldsymbol{z}_{i,j}) \leq \frac{1-a_{i,0}}{2}\log(\zeta_{i,L_i}\gamma + 1) + \log(\sqrt{2\pi e}\sigma) - \sum_{l=0}^{L_i} a_{i,l}\log a_{i,l}. \tag{2.31}$$

□

By (2.27), (2.28) and (2.29),

$$\begin{aligned}\mathscr{R}_i \geq \mathscr{R}_i^{(\mathrm{lb})} &\triangleq \frac{v_i}{2}\log\left(\frac{2^{2\sum_{l=0}^{L_i} a_{i,l}\log a_{i,l}}|h_{i,i}|^2\gamma}{v_i(\zeta_{i,L_i}\gamma+1)^{1-a_{i,0}}}+1\right)\\ &= \frac{v_i}{2}\log\left(\frac{2^{2\sum_{l=0}^{L_i} a_{i,l}\log a_{i,l}}|h_{i,i}|^2}{v_i(\zeta_{i,L_i}+\gamma^{-1})^{1-a_{i,0}}}+\gamma^{-a_{i,0}}\right)+\frac{1}{2}v_i a_{i,0}\log\gamma. \qquad (2.32)\end{aligned}$$

Defining $\widetilde{\mathscr{R}}_i^{(\mathrm{lb})} \triangleq \frac{v_i}{2}\log\left(\frac{2^{2\sum_{l=0}^{L_i} a_{i,l}\log a_{i,l}}|h_{i,i}|^2}{v_i(\zeta_{i,L_i}+\gamma^{-1})^{1-a_{i,0}}}+\gamma^{-a_{i,0}}\right)$, we note that $\lim_{\gamma\to\infty}\widetilde{\mathscr{R}}_i^{(\mathrm{lb})} < \infty$. Combining this with the fact that $a_{i,0} = \prod_{\substack{k=1\\k\neq i}}^{N}\left(1-\frac{v_k}{u}\right)$, the proof of Theorem 2 is complete[4]. $\square$

One may consider the following generalization of the FH scheme. Let us assume that the users are not restricted to choose a fixed number of frequency sub-bands in each transmission slot. In fact, in each transmission slot the number of selected sub-bands can be any integer between 0 and u, and the probability of choosing $v \in [0,u]\cap\mathbb{Z}$ sub-bands by the i^{th} user is denoted by $\mu_{i,v}$. Therefore, the i^{th} user has two random generators. The first random generator selects a number $v \in [0,u]\cap\mathbb{Z}$ according to the probability mass function $\{\mu_{i,v}\}_{v=0}^{u}$, while the other generator selects v sub-bands among the whole available u sub-bands. This repeats independently from transmission slot to transmission slot. Based on the arguments made at the beginning of this section, the achievable rate of the i^{th} user can be written as

$$\mathscr{R}_i = \sum_{v=0}^{u}\mu_{i,v}\mathrm{I}(\vec{\boldsymbol{x}}_i(s_{i,v}^*);\vec{\boldsymbol{y}}_i(s_{i,v}^*)), \qquad (2.33)$$

where $s_{i,v}^*$ denotes the state where the i^{th} user selects the first v sub-bands. Clearly, $\mathrm{I}(\vec{\boldsymbol{x}}_i(s_{i,0}^*);\vec{\boldsymbol{y}}_i(s_{i,0}^*)) =$

[4]In [78], we address another approach to propose a lower bound (looser than $\mathscr{R}_i^{(\mathrm{lb})}$) on the achievable rate of the i^{th} user with the same SNR scaling as $\mathscr{R}_i^{(\mathrm{lb})}$.

0 for any $1 \leq i \leq N$. Furthermore,

$$\begin{aligned}
a_{i,0} &= \Pr\{\boldsymbol{z}_{i,1} \text{ is interference-free}\} \\
&= \sum_{v_1=0}^{u} \cdots \sum_{v_{i-1}=0}^{u} \sum_{v_{i+1}=0}^{u} \cdots \sum_{v_N=0}^{u} \prod_{\substack{k=1\\k\neq i}}^{N} \mu_{k,v_k}\left(1-\frac{v_k}{u}\right) \\
&= \prod_{\substack{k=1\\k\neq i}}^{N} \sum_{v_k=0}^{u} \mu_{k,v_k}\left(1-\frac{v_k}{u}\right) \\
&= \prod_{\substack{k=1\\k\neq i}}^{N} \sum_{v=0}^{u} \mu_{k,v}\left(1-\frac{v}{u}\right) \\
&= \prod_{\substack{k=1\\k\neq i}}^{N} \left(1-\frac{\widetilde{v}_k}{u}\right) \qquad (2.34)
\end{aligned}$$

where $\widetilde{v}_k \triangleq \sum_{v=0}^{u} v\mu_{k,v}$. Applying the results of this section, we get

$$\mathrm{I}(\vec{\boldsymbol{x}}_i(s^*_{i,v}); \vec{\boldsymbol{y}}_i(s^*_{i,v})) \overset{\gamma}{\sim} \frac{1}{2} v a_{i,0} \log \gamma. \qquad (2.35)$$

By (2.33), (2.34) and (2.35),

$$\begin{aligned}
\mathscr{R}_i &\overset{\gamma}{\gtrsim} \sum_{v=0}^{u} \frac{1}{2}\mu_{i,v} v \prod_{\substack{k=1\\k\neq i}}^{N} \left(1-\frac{\widetilde{v}_k}{u}\right) \log \gamma \\
&= \frac{1}{2}\widetilde{v}_i \prod_{\substack{k=1\\k\neq i}}^{N} \left(1-\frac{\widetilde{v}_k}{u}\right) \log \gamma. \qquad (2.36)
\end{aligned}$$

In fact, (2.36) demonstrates that the generalized FH scheme is equivalent to the FH scheme through substituting $\{v_i\}_{i=1}^{N}$ by $\{\widetilde{v}_i\}_{i=1}^{N}$. However, it is remarkable that in contrast to the FH scheme in which $\{v_i\}_{i=1}^{N}$ are integer values, in the generalized FH scheme $\{\widetilde{v}_i\}_{i=1}^{N}$ are real values. This provides more flexibility in system design. The above observation motivates us to use this generalized scenario in the rest of the chapter and we simply refer to it as the FH scheme. In this scheme, the i^{th} user has a parameter $\widetilde{v}_i$, which can be chosen to be any real number in the interval $[0, u]$.

2.3 System Design

In this section, we find the optimum operation point for the FH scheme. This requires finding the optimum values of $\{\widetilde{v}_i\}_{i=1}^{N}$. Based on the results established in the previous section, there exist upper and lower bounds on the achievable rate of each user that coincide in the high SNR regime. As such, the achievable rate itself must be asymptotically equivalent to each of these bounds, i.e.,

$$\mathscr{R}_i \overset{\gamma}{\sim} \frac{1}{2}\widetilde{v}_i \prod_{\substack{k=1 \\ k \neq i}}^{N} \left(1 - \frac{\widetilde{v}_k}{u}\right) \log \gamma, \tag{2.37}$$

where, based on the conclusion made at the end of section 2.2, the parameters $\{\widetilde{v}_i\}_{i=1}^N$ can be adjusted to be any real number in the range $[0, u]$. By (2.37), the network sum rate can be asymptotically written as

$$\sum_{i=1}^{N} \mathscr{R}_i \overset{\gamma}{\sim} \mathrm{SMG}\left(\widetilde{v}_1, \cdots, \widetilde{v}_N\right) \log \gamma, \tag{2.38}$$

where

$$\mathrm{SMG}\left(\widetilde{v}_1, \cdots, \widetilde{v}_N\right) \triangleq \sum_{i=1}^{N} \frac{1}{2}\widetilde{v}_i \prod_{\substack{k=1 \\ k \neq i}}^{N} \left(1 - \frac{\widetilde{v}_k}{u}\right). \tag{2.39}$$

We call $\mathrm{SMG}\left(\widetilde{v}_1, \cdots, \widetilde{v}_N\right)$ the *sum multiplexing gain* of the system. $\mathrm{SMG}\left(\widetilde{v}_1, \cdots, \widetilde{v}_N\right)$ is a symmetric function of $\left(\widetilde{v}_1, \cdots, \widetilde{v}_N\right)$ and has a saddle point at $\widetilde{v}_i = \frac{u}{N}$ for $1 \leq i \leq N$. In a *fair* FH system, it is required that $\widetilde{v}_i = v$ for all $1 \leq i \leq N$ where v is any real number in the interval $[0, u]$. We call v the hopping parameter in the randomized FH scenario. Hence, we define

$$\begin{aligned} \mathrm{SMG}(v, N) &\triangleq \mathrm{SMG}\left(\widetilde{v}_1, \cdots, \widetilde{v}_N\right)\Big|_{\forall i: \widetilde{v}_i = v} \\ &= \frac{N}{2} v \left(1 - \frac{v}{u}\right)^{N-1}. \end{aligned} \tag{2.40}$$

Maximizing this in terms of v yields[5]

$$v_{\mathrm{opt}} = \frac{u}{N}. \tag{2.41}$$

Setting $v = v_{\mathrm{opt}}$, the highest sum multiplexing gain of the fair FH scheme is given by

$$\sup_{v} \ \mathrm{SMG}(v, N) = \frac{1}{2} u \left(1 - \frac{1}{N}\right)^{N-1}. \tag{2.42}$$

It is remarkable that $\frac{u}{N}$ may not be an integer. If we do not adopt the generalized FH scheme, then all users must hop randomly over sets of $\hat{v} = \max\left\{\lfloor \frac{u}{N} \rfloor, 1\right\}$ frequency sub-bands. This results in a sum multiplexing gain of $\frac{N}{2}\hat{v}\left(1 - \frac{\hat{v}}{u}\right)^{N-1}$. This is generally less than $\frac{1}{2}u\left(1 - \frac{1}{N}\right)^{N-1}$. By adopting the generalized FH scheme in case $\frac{u}{N} \notin \mathbb{Z}$, each user only needs to hop randomly over different sets of frequency sub-bands of cardinality $\lfloor \frac{u}{N} \rfloor$ or $\lceil \frac{u}{N} \rceil$. In fact, each user has two random generators.

[5]Computation of v_{opt} requires that all transmitters know the number of active users N in the network. This is more discussed in section 2.4.

The first random generator selects one of the numbers $\lfloor \frac{u}{N} \rfloor$ and $\lceil \frac{u}{N} \rceil$ with probabilities μ and $\overline{\mu}$, respectively, such that $\mu \lfloor \frac{u}{N} \rfloor + \overline{\mu} \lceil \frac{u}{N} \rceil = \frac{u}{N}$ or equivalently $\mu = \lceil \frac{u}{N} \rceil - \frac{u}{N}$. Let us assume the first random generator has selected a number $a \in \{\lfloor \frac{u}{N} \rfloor, \lceil \frac{u}{N} \rceil\}$. Then, the second random generator selects a subset of cardinality a among the u frequency sub-bands. Doing this independently from transmission slot to transmission slot, the sum multiplexing gain given in (2.42) is achieved.

Observation 1- One might suggest another well-known utility function that is popular in the game theory context, namely the *proportional fair function*, which is defined as $\sum_{i=1}^{N} \log \mathscr{R}_i$. We have

$$\begin{aligned} \sum_{i=1}^{n} \log \mathscr{R}_i \quad &\overset{\gamma}{\sim} \quad \sum_{i=1}^{N} \log \left(\frac{1}{2} \widetilde{v}_i \prod_{\substack{k=1 \\ k \neq i}}^{N} \left(1 - \frac{\widetilde{v}_k}{u}\right) \log \gamma \right) \\ &= \quad \sum_{i=1}^{N} \log \left(\frac{1}{2} \widetilde{v}_i \prod_{\substack{k=1 \\ k \neq i}}^{N} \left(1 - \frac{\widetilde{v}_k}{u}\right) \right) + N \log\log \gamma. \end{aligned} \tag{2.43}$$

It can be easily verified that $\sum_{i=1}^{N} \log \left(\frac{1}{2} \widetilde{v}_i \prod_{\substack{k=1 \\ k \neq i}}^{N} \left(1 - \frac{\widetilde{v}_k}{u}\right) \right)$ has an absolute maximum at $\widetilde{v}_i = \frac{u}{N}$ for $1 \leq i \leq N$.

Observation 2- As we will discuss in more detail in the next section, the number of active users in the system is in general a random variable $\boldsymbol{N}$ with realization N. Although users in the FH system can use their knowledge about the number of active users to adjust the hopping parameter (as explained earlier in this section), one may devise a sub-optimal rule to fix $v = v^*$ given by

$$v^* = \arg \max_{v \in [0,u]} \mathrm{E}\left\{\mathrm{SMG}(v, \boldsymbol{N})\right\}, \tag{2.44}$$

where the expectation is with respect to the number of active users in the network. This selection of v by all users makes the system robust against changes in the number of active users in the network[6]. We call this version of the FH system the robust randomized Frequency Hopping. We remark that the rule in (2.44) is a particular design approach for the robust FH system. In the next section, we consider another design rule based on maximizing the average of the minimum multiplexing gain per user in term of the number of active users in the network.

[6]In fact, the transmitters use their knowledge about the instantaneous number of active users only to regulate their transmission rates. This is explained in (2.50).

2.4 Comparison of the robust FH scenario with the FD scheme

In a centralized setup, under the condition that no user is aware of the other users' codebooks and the number of users is fixed and known to the central controller, it is shown in [9] that if the crossover channel gains are sufficiently larger than the forward channel gains, then every Pareto optimal rate vector is realized by Frequency Division for all ranges of SNR. However, in realistic scenarios, the number of active users is not fixed. This degrades the performance of the FD scheme as it is designed for a specified number of users. In particular, if the number of active users is less than the designed target of the FD scheme, a considerable portion of the spectrum may remain unused. This encourages us to compare the performance of the proposed robust FH scheme with that of the FD scheme in a setup where the number of active users is a random variable $\boldsymbol{N}$ with a given distribution[7]. The realization of $\boldsymbol{N}$ is denoted by N as before.

To perform the comparison, we introduce three different performance measures. In the following definitions, the sup operation is over possible adjustable parameters in the system, e.g., the hopping parameter in the robust randomized FH scenario. All expectations are taken with respect to $\boldsymbol{N}$. We define $q_n \triangleq \Pr\{\boldsymbol{N} = n\}$ for all $n \geq 0$. It is assumed that the maximum number of active users in the network is $n_{\max}$, i.e., $\Pr\{\boldsymbol{N} > n_{\max}\} = 0$. We usually take $q_0 = 0$ unless otherwise stated.

- **Average Sum Multiplexing Gain**, which is defined as

$$\eta^{(1)} \triangleq \sup \lim_{\gamma \to \infty} \frac{\mathrm{E}\left\{\sum_{i=1}^{\boldsymbol{N}} \mathscr{R}_i\right\}}{\log \gamma} = \sup \mathrm{E}\left\{\mathrm{SMG}\right\}, \tag{2.45}$$

  where $\mathrm{SMG} = \lim_{\gamma \to \infty} \frac{\sum_{i=1}^{\boldsymbol{N}} \mathscr{R}_i}{\log \gamma}$ is the sum multiplexing gain.

- **Average Minimum Multiplexing Gain per User**, which is defined as

$$\eta^{(2)} \triangleq \sup \lim_{\gamma \to \infty} \frac{\mathrm{E}\left\{\min_{1 \leq i \leq \boldsymbol{N}} \mathscr{R}_i\right\}}{\log \gamma}. \tag{2.46}$$

- **Minimum nonzero multiplexing gain per user**, which is defined as

$$\eta^{(3)} \triangleq \min_{n: q_n \neq 0} \quad \min_{\substack{\boldsymbol{N}_{\mathrm{serv}} = n \\ 1 \leq i \leq n}} \quad \lim_{\gamma \to \infty} \frac{\mathscr{R}_i}{\log \gamma} \tag{2.47}$$

  where $\boldsymbol{N}_{\mathrm{serv}}$ denotes the number of active users receiving service (i.e., their multiplexing gain is strictly positive).

[7] As explicitly mentioned in the system model, the number of users is assumed to be fixed for the whole transmission period of interest.

The FD system is designed to service a certain number of active users. We denote this design target in the FD scheme by n_{des}. Therefore, the spectrum is divided to n_{des} *bands* where each band contains $\frac{u}{n_{\mathrm{des}}}$ frequency sub-bands. This requires that u is divisible by n_{des}, which is assumed to be the case to guarantee fairness. Each user that becomes active occupies an empty band. If there is no empty band, no service is available. In case $n_{\max} \leq u$, the central controller in the FD system sets $n_{\mathrm{des}} = n_{\max}$ to ensure that all users can receive service upon activation[8]. In case $n_{\max} > u$, the central controller sets $n_{\mathrm{des}} = u$ to guarantee that as many users receive service as possible. Therefore, $n_{\mathrm{des}} = \min\{n_{\max}, u\}$. In fact, one can show that selecting $n_{\mathrm{des}} = \min\{n_{\max}, u\}$ maximizes the service capability in the FD system. Service capability is measured by $\mathrm{E}\left\{\frac{\boldsymbol{N}_{\mathrm{serv}}}{\boldsymbol{N}}\right\}$.

Due to the nature of the randomized FH scheme, as far as the hopping parameter v is strictly less than u, all users receive service, while if $v = u$ and $N > 1$, no user receives service, i.e., the multiplexing gain achieved by any active user is zero. As such, to get the largest service capability in the FH scenario, we require $v \in (0, u)$. As an example, if v^* in (2.44) is equal to u, the service capability will be less than 1. To avoid this, we set the hopping parameter $v = v^* - \delta = u - \delta$ for sufficiently small δ such that the performance of the robust FH is still above the performance of the FD scenario.

Note that the comparison between the robust FH and the FD schemes is a fair comparison based on the following facts:

1- Both schemes are required to achieve the largest possible service capability. This is enforced by setting $v \in (0, u)$ and $n_{\mathrm{des}} = \min\{n_{\max}, u\}$ in the robust FH and FD scenarios, respectively.

2- The hopping parameter v in robust FH and the number of bands n_{des} in FD are both designed based on the distribution of the number of active users rather than the realization of the number of active users itself.

- *Average Sum Multiplexing Gain*

This measure is a meaningful tool of comparison if $n_{\max} < \infty$. Hence, we assume $n_{\max}$ is a finite number and u is a multiple of $n_{\max}$ in this subsection. It is easily seen that the sum multiplexing gain in the FD scenario is

$$\mathrm{SMG}_{\mathrm{FD}}(n_{\mathrm{des}}, \boldsymbol{N}) = \begin{cases} \frac{\boldsymbol{N}}{2}\frac{u}{n_{\mathrm{des}}} & \boldsymbol{N} \leq n_{\mathrm{des}} \\ \frac{u}{2} & \boldsymbol{N} > n_{\mathrm{des}} \end{cases}. \tag{2.48}$$

By (2.40), $\mathrm{SMG}_{\mathrm{FH}}(v, \boldsymbol{N})$ is given by

$$\mathrm{SMG}_{\mathrm{FH}}(v, \boldsymbol{N}) = \frac{1}{2}\boldsymbol{N} v \left(1 - \frac{v}{u}\right)^{\boldsymbol{N}-1}. \tag{2.49}$$

As mentioned earlier, a candidate for robust hopping strategy against changes in the number of active users is the one given in (2.44). It is notable that although the value of v is fixed at v^*,

[8] We assume u is divisible by $n_{\max}$ in this case.

any user regulates its transmission rate based on the instantaneous number of active users to avoid transmission failure. Using the lower bound on the achievable rate of the i^{th} user given in (2.32), it is shown in appendix B that the i^{th} user may select its transmission rate R_i as

$$R_i = \mathrm{E}\left\{\frac{\boldsymbol{v}_i}{2}\log\left(\frac{2^{-2(N-1)\left(\vartheta_{v^*}+\frac{\mathscr{H}(\lceil v^*\rceil - v^*)}{\boldsymbol{v}_i}\right)}\gamma|h_{i,i}|^2}{\boldsymbol{v}_i\left(\frac{\gamma\sum_{j\neq i}|h_{j,i}|^2}{\lfloor v^*\rfloor}+1\right)^{1-\left(1-\frac{v^*}{u}\right)^{N-1}}}+1\right)\right\} \tag{2.50}$$

where $\boldsymbol{v}_i$ is a random variable taking the values $\lfloor v^*\rfloor$ and $\lceil v^*\rceil$ with probabilities $\lceil v^*\rceil - v^*$ and $v^* - \lfloor v^*\rfloor$, respectively. The quantity ϑ_{v^*} is defined as

$$\vartheta_{v^*} \triangleq \left\{\begin{array}{cc} \mathscr{H}\left(\frac{v^*}{u}\right) & v^* \in \mathbb{N} \\ (v^* - \lfloor v^*\rfloor)\,\mathscr{H}\left(\frac{\lceil v^*\rceil}{u}\right) + (\lceil v^*\rceil - v^*))\mathscr{H}\left(\frac{\lfloor v^*\rfloor}{u}\right) & v^* \notin \mathbb{N} \end{array}\right. . \tag{2.51}$$

It is seen that the quantities the i^{th} transmitter needs to evaluate R_i are $|h_{i,i}|$, $\sum_{j\neq i}|h_{j,i}|^2$ and N. The i^{th} receiver sends these required data to its transmitter via a feedback link. It is shown in appendix B in [78] how the i^{th} receiver can estimate N and $\{h_{j,i}\}_{j\neq i}$ using a combined technique based on the method of moments and maximum likelihood estimation [83]. In this work we assume the i^{th} user has perfect knowledge of N and the channel gains $\{h_{j,i}\}_{j=1}^{N}$. In section 2.6, we relax this assumption partially by not requiring any knowledge about the channel gains and the number of active users at the transmitters. Moreover, regulating the transmission rate at R_i by the i^{th} user for any realization N of $\boldsymbol{N}$, one may observe that an average sum multiplexing gain of $\mathrm{E}\left\{\boldsymbol{N}v^*\left(1-\frac{v^*}{u}\right)^{\boldsymbol{N}-1}\right\}$ is achievable.

We present an example to compare the performance of FH with that of FD in terms of $\eta^{(1)}$.

Example 1- Let us consider a network where $n_{\max} = 2$. The central controller in the FD system sets $n_{\text{des}} = 2$, and according to (2.48), $\eta^{(1)}_{\text{FD}} = \mathrm{E}\{\mathrm{SMG}_{\text{FD}}(2,\boldsymbol{N})\} = q_1\frac{u}{4} + q_2\frac{u}{2} = \frac{q_1+2q_2}{4}u$. Based on (2.49), $\mathrm{E}\{\mathrm{SMG}_{\text{FH}}(v,\boldsymbol{N})\} = \frac{1}{2}q_1 v + q_2 v\left(1-\frac{v}{u}\right)$. Using this in (2.44),

$$v^* = \arg\max_{v\in[0,u]} \mathrm{E}\{\mathrm{SMG}_{\text{FH}}(v,\boldsymbol{N})\} = \left\{\begin{array}{cc} \frac{q_1+2q_2}{4q_2}u & q_1 \leq 2q_2 \\ u & q_1 > 2q_2 \end{array}\right. . \tag{2.52}$$

Therefore,

$$\eta^{(1)}_{\text{FH}} = \sup_{v\in[0,u]} \mathrm{E}\{\mathrm{SMG}_{\text{FH}}(v,\boldsymbol{N})\} = \mathrm{E}\{\mathrm{SMG}_{\text{FH}}(v^*,\boldsymbol{N})\} = \left\{\begin{array}{cc} \frac{(q_1+2q_2)^2}{16q_2}u & q_1 \leq 2q_2 \\ \frac{q_1}{2}u & q_1 > 2q_2 \end{array}\right. . \tag{2.53}$$

It is easy to see that $\eta^{(1)}_{\text{FH}} > \eta^{(1)}_{\text{FD}}$ if and only if $q_1 > 2q_2$, or equivalently, $q_1 > \frac{2}{3}$. We note that in this case $v^* = u$, i.e., all users spread their power on the whole spectrum and no hopping is performed. This makes service capability be strictly less than %100 because, if both users are active, none of them receive service. As such, we take $v = u - \delta$. To ensure that the performance of the robust FH scenario is above that of the FD system, we require

$$\frac{1}{2}q_1(u-\delta) + q_2(u-\delta)\left(1 - \frac{u-\delta}{u}\right) > \frac{q_1 + 2q_2}{4}u. \tag{2.54}$$

As far as $\delta < \frac{u}{2}$, (2.54) is equivalent to $q_1 > 2q_2 \frac{1-\frac{2\delta}{u}\left(1-\frac{\delta}{u}\right)}{1-\frac{2\delta}{u}}$. This is a more restrictive condition than $q_1 > 2q_2$ which is the cost paid for having full service capability. □

- *Average Minimum Multiplexing Gain per User*

This measure can also be written as

$$\eta^{(2)} = \sup \mathrm{E}\left\{\frac{\mathrm{SMG}}{\boldsymbol{N}}\mathbb{I}\left\{\boldsymbol{N}_{\mathrm{serv}} = \boldsymbol{N}\right\}\right\}. \tag{2.55}$$

In fact, if $\boldsymbol{N}_{\mathrm{serv}} \neq \boldsymbol{N}$, there exists at least one user that achieves no multiplexing gain. Therefore, the minimum multiplexing gain per user is zero in this case. However, if $\boldsymbol{N}_{\mathrm{serv}} = \boldsymbol{N}$, all users achieve a nonzero multiplexing gain. This measure can be used whether $n_{\max}$ is finite or infinite.

In case of the FH scenario, the rule to choose the optimum value of the hopping parameter v, denoted by v_*, is given by

$$v_* = \arg \max_{v \in [0,u]} \mathrm{E}\left\{\frac{\mathrm{SMG}_{\mathrm{FH}}(v, \boldsymbol{N})}{\boldsymbol{N}}\mathbb{I}\left\{\boldsymbol{N}_{\mathrm{serv}} = \boldsymbol{N}\right\}\right\}. \tag{2.56}$$

In this case, the transmission rate of the i^{th} user is given by (2.50) where v^* is replaced by v_*.

In general, it is hard to determine v^* or v_* analytically for $n_{\max} \geq 3$. Therefore, one can not obtain a closed characterization for the regions $\left\{\{q_n\}_{n=1}^{n_{\max}} : \eta_{\mathrm{FH}}^{(1)} > \eta_{\mathrm{FD}}^{(1)}\right\}$ and $\left\{\{q_n\}_{n=1}^{n_{\max}} : \eta_{\mathrm{FH}}^{(2)} > \eta_{\mathrm{FD}}^{(2)}\right\}$. We aim to derive sufficient conditions on $\{q_n\}_{n=1}^{n_{\max}}$ (inner bounds to the mentioned regions) such that $\eta_{\mathrm{FH}}^{(1)} > \eta_{\mathrm{FD}}^{(1)}$ or $\eta_{\mathrm{FH}}^{(2)} > \eta_{\mathrm{FD}}^{(2)}$. Our conditions only involve $\mathrm{E}\{\boldsymbol{N}\}$ and $n_{\max}$.

Proposition 1. *If*

$$\mathrm{E}\{\boldsymbol{N}\} < \frac{1}{2}\ln\left((e^2 - 1)n_{\max}\right), \tag{2.57}$$

then, $\eta_{\mathrm{FD}}^{(1)} < \eta_{\mathrm{FH}}^{(1)}$. *Also, As far as*

$$\frac{1}{\mathrm{E}\{\boldsymbol{N}\}}\left(1 - \frac{1}{\mathrm{E}\{\boldsymbol{N}\}}\right)^{\mathrm{E}\{\boldsymbol{N}\}-1} > \frac{1}{n_{\max}}, \tag{2.58}$$

we have $\eta_{\mathrm{FD}}^{(2)} < \eta_{\mathrm{FH}}^{(2)}$.

Proof. See appendix C. □

For example, if $n_{\max} = 10$, $q_1 = 0.22$, $q_2 = q_3 = q_4 = 0.24$ and $q_5 = q_6 = \cdots = q_{10} = 0.01$, one has $\mathrm{E}\{\boldsymbol{N}\} = 2.78$, which satisfies (2.58). Therefore, we conclude $\eta_{\mathrm{FH}}^{(2)} > \eta_{\mathrm{FD}}^{(2)}$. Computing these quantities directly, we get $\eta_{\mathrm{FD}}^{(2)} = \frac{u}{16}$ and

$$\begin{aligned}
\eta_{\mathrm{FH}}^{(2)} &= \frac{1}{2} \max_{v \in [0,u]} \left\{ v \sum_{n=1}^{10} q_n \left(1 - \frac{v}{u}\right)^{n-1} \right\} \\
&\overset{(a)}{=} \frac{1}{2} u \max_{\omega_v \in [0,1]} (1 - \omega_v) \left(0.22 + 0.24 \sum_{n=1}^{3} \omega_v^n + 0.01 \sum_{n=4}^{9} \omega_v^n \right) \\
&\overset{(b)}{=} 0.1121u
\end{aligned} \tag{2.59}$$

where in (a), we define $\omega_v \triangleq 1 - \frac{v}{u}$ and (b) is obtained by setting $\omega_v = 0.28$, or equivalently $v = v_* = 0.72u$. This yields $\frac{\eta_{\mathrm{FH}}^{(2)}}{\eta_{\mathrm{FD}}^{(2)}} = 1.7936$.

We consider an example where $n_{\max}$ is not finite.

Example 2- We assume a Poisson distribution on the number of active users, i.e., $q_n = \frac{e^{-\lambda}\lambda^n}{n!}$ for $n \geq 0$ and $\lambda > 0$. We have

$$\begin{aligned}
& \mathrm{E}\left\{ \frac{\mathrm{SMG}_{\mathrm{FH}}(v, \boldsymbol{N})}{\boldsymbol{N}} \mathbb{I}\left\{ \boldsymbol{N}_{\mathrm{serv,FH}} = \boldsymbol{N} \right\} \right\} \\
\overset{(a)}{=}\; & \mathrm{E}\left\{ \frac{\mathrm{SMG}_{\mathrm{FH}}(v, \boldsymbol{N})}{\boldsymbol{N}} \right\} \\
=\; & \frac{1}{2} \sum_{n=1}^{\infty} \frac{e^{-\lambda}\lambda^n}{n!} \left(v \left(1 - \frac{v}{u}\right)^{n-1} \right) \\
=\; & \frac{1}{2} \frac{v}{1 - \frac{v}{u}} \sum_{n=1}^{\infty} \frac{e^{-\lambda}\lambda^n}{n!} \left(1 - \frac{v}{u}\right)^{n} \\
\overset{(b)}{=}\; & \frac{1}{2} \frac{v}{1 - \frac{v}{u}} \left(e^{\lambda\left(e^{\ln\left(1 - \frac{v}{u}\right)} - 1\right)} - e^{-\lambda} \right) \\
=\; & \frac{e^{-\lambda}(1 - \omega_v)\left(e^{\lambda\omega_v} - 1\right)}{2\omega_v} u.
\end{aligned} \tag{2.60}$$

In the above equation, (a) results from the fact that $\mathbb{I}\left\{ \boldsymbol{N}_{\mathrm{serv,FH}} = \boldsymbol{N} \right\} = 0$ whenever $v = u$ and $N > 1$, however, $\mathrm{SMG}_{\mathrm{FH}}(v, \boldsymbol{N}) = 0$ in this case. Also, (b) follows by the fact that $\mathrm{E}\{e^{t\boldsymbol{N}}\} = e^{\lambda(e^t - 1)}$ for any $t \in \mathbb{R}$ and we have defined $\omega_v \triangleq 1 - \frac{v}{u}$. It can be easily seen that the optimal $v = v_*$ satisfies the nonlinear equation $e^{-\lambda\omega_{v_*}} = 1 - \lambda\omega_{v_*}(1 - \omega_{v_*})$. Solving this for v_*, we find out that v_* is not equal to u for all $\lambda > 2$. The following table lists the values of ω_{v_*}, the values of v_* and also the corresponding average minimum multiplexing gain per user $\eta_{\mathrm{FH}}^{(2)}$ for $\lambda \in \{3, 4, 5, 6\}$.

$$\begin{array}{|c|c|c|c|c|}\hline \lambda & 3 & 4 & 5 & 6 \\ \hline \omega_{v_*} & 0.4536 & 0.6392 & 0.7347 & 0.7912 \\ \hline v_* & 0.5464u & 0.3608u & 0.2653u & 0.2088u \\ \hline \eta_{\mathrm{FH}}^{(2)} & 0.0869u & 0.0615u & 0.0467u & 0.0374u \\ \hline \end{array}. \tag{2.61}$$

The FD system aims to serve as many users as it can. Since it is not possible to serve more than u users, the number of bands is set at $n_{\mathrm{des}} = u$. Therefore, $\boldsymbol{N}_{\mathrm{serv,FD}} < \boldsymbol{N}$ if and only if $\boldsymbol{N} > u$. Using this and by (2.48),

$$\begin{aligned} \eta_{\mathrm{FD}}^{(2)} &= \mathrm{E}\left\{\frac{\mathrm{SMG}_{\mathrm{FD}}(n_{\mathrm{des}}, \boldsymbol{N})}{\boldsymbol{N}} \mathbb{I}\left\{\boldsymbol{N}_{\mathrm{serv,FD}} = \boldsymbol{N}\right\}\right\} \\ &= \mathrm{E}\left\{\frac{\mathrm{SMG}_{\mathrm{FD}}(u, \boldsymbol{N})}{\boldsymbol{N}} \middle| \boldsymbol{N} \le u\right\} \Pr\{\boldsymbol{N} \le u\} \\ &= \frac{1}{2}\sum_{n=1}^{u} \frac{e^{-\lambda}\lambda^n}{n!}. \end{aligned} \tag{2.62}$$

We have sketched $\eta_{\mathrm{FH}}^{(2)}$ and $\eta_{\mathrm{FD}}^{(2)}$ in terms of λ in fig. 2.1 for $u = 7$. It is noticeable that $\eta_{\mathrm{FH}}^{(2)}$ scales linearly with u. However, $\eta_{\mathrm{FD}}^{(2)}$ is always less than $\frac{1}{2}$ no matter how large u is. As u increases, the advantage of FH over FD becomes more apparent. $\square$

• *Minimum nonzero multiplexing gain per user*

The minimum nonzero multiplexing gain per user is the smallest nonzero multiplexing gain that a user in the network attains for different realizations in terms of the number of active users. Assuming $n_{\max} < \infty$, this happens when there are exactly $n_{\max}$ active users in the system. As the FD system is already designed to handle the case where $n_{\max}$ users are present in the network, the minimum multiplexing gain per user is automatically higher in FD as compared to FH. Setting $n_{\mathrm{des}} = n_{\max}$, we have $\eta_{\mathrm{FD}}^{(3)} = \frac{\mathrm{SMG}_{\mathrm{FD}}(u, n_{\max})}{n_{\max}} = \frac{u}{2n_{\max}}$. In the case of FH, we assume that all users select $v = \frac{u}{n_{\max}}$. Hence, $\eta_{\mathrm{FH}}^{(3)} = \frac{\mathrm{SMG}_{\mathrm{FH}}\left(\frac{u}{n_{\max}}, n_{\max}\right)}{n_{\max}} = \frac{u}{2n_{\max}}\left(1 - \frac{1}{n_{\max}}\right)^{n_{\max}-1}$. Clearly, $\frac{1}{e} \le \frac{\eta_{\mathrm{FH}}^{(3)}}{\eta_{\mathrm{FD}}^{(3)}} \le 1$ as $\left(1 - \frac{1}{n_{\max}}\right)^{n_{\max}-1}$ approaches $\frac{1}{e}$ from above by increasing $n_{\max}$. Therefore, the loss incurred in terms of $\eta^{(3)}$ for the FH system is always less than $\frac{1}{e}$.

2.5 Adaptive Frequency Hopping

The results of the previous section are obtained based on the assumption that the hopping parameter v is fixed and is not adaptively changed based on the number of active users. The performance of

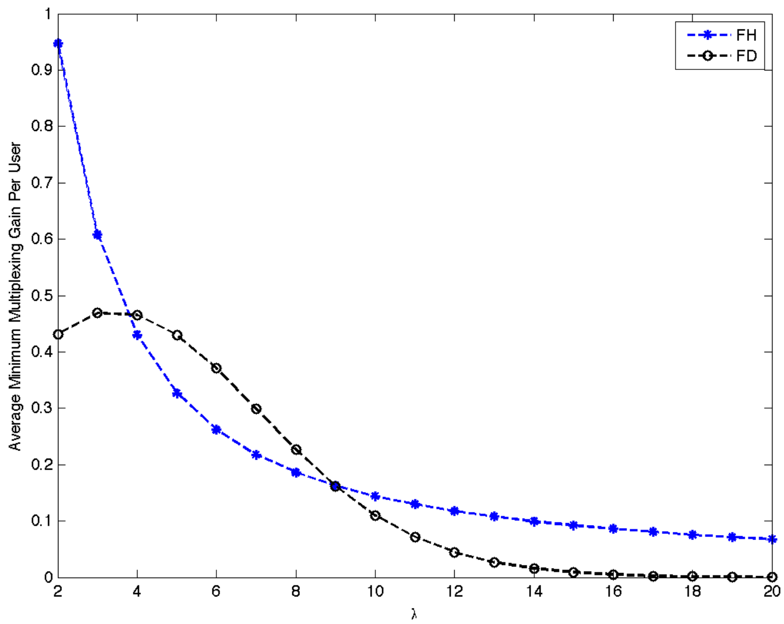

Figure 2.1: Curves of $\eta^{(2)}_{\mathrm{FH}}$ and $\eta^{(2)}_{\mathrm{FD}}$ in terms of λ in a network with $u = 7$ sub-bands.

the FH system can be improved by letting the transmitters adapt their hopping parameter based on the number of active users using (2.41). We refer to this scenario as Adaptive Frequency Hopping (AFH). In the following example, we study the performance improvement offered by AFH over FH in terms of $\eta^{(1)}$ and $\eta^{(2)}$.

Example 3- Let us assume that the number of active users is a Poisson random variable with parameter $\lambda > 1$. To compute v^* in the FH scenario, we have

$$\begin{aligned}
\mathrm{E}\{\mathrm{SMG}_{\mathrm{FH}}(v, \boldsymbol{N})\} &= \frac{1}{2}\sum_{n=1}^{\infty}\frac{e^{-\lambda}\lambda^n}{n!}\left(nv\left(1-\frac{v}{u}\right)^{n-1}\right) \\
&= \frac{1}{2}v\sum_{n=1}^{\infty}\frac{e^{-\lambda}\lambda^n}{(n-1)!}\left(1-\frac{v}{u}\right)^{n-1} \\
&= \frac{1}{2}\lambda v\sum_{n=0}^{\infty}\frac{e^{-\lambda}\lambda^n}{n!}\left(1-\frac{v}{u}\right)^{n} \qquad (2.63) \\
&= \frac{1}{2}\lambda v e^{-\frac{\lambda v}{u}}. \qquad (2.64)
\end{aligned}$$

Therefore,

$$v^* = \arg\max_v \mathrm{E}\{\mathrm{SMG}_{\mathrm{FH}}(v, \boldsymbol{N})\} = \frac{u}{\lambda}. \qquad (2.65)$$

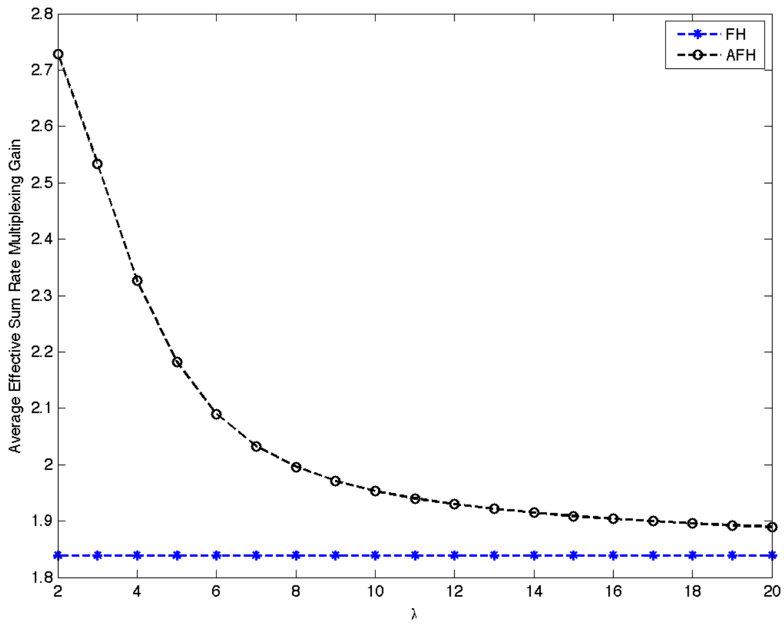

Figure 2.2: Plots of $\eta^{(1)}_{\text{AFH}}$ and $\eta^{(1)}_{\text{FH}}$ in terms of λ for $u = 10$.

Since $\lambda \neq 1$, we get $v^* \neq u$. Thus, choosing $v = v^*$ guarantees a %100 service capability. Finally,

$$\eta^{(1)}_{\text{FH}} = \frac{u}{2e}. \tag{2.66}$$

By (2.42),

$$\eta^{(1)}_{\text{AFH}} = \frac{u}{2}\sum_{n=1}^{\infty}\frac{e^{-\lambda}\lambda^n}{n!}\Big(1-\frac{1}{n}\Big)^{n-1}. \tag{2.67}$$

Figure 2.2 shows the plots of $\eta^{(1)}_{\text{FH}}$ and $\eta^{(1)}_{\text{AFH}}$ versus λ for $u = 10$. It is observed that $\eta^{(1)}_{\text{FH}}$ does not change with λ, while $\eta^{(1)}_{\text{AFH}}$ decreases by increasing λ. This indicates that in a crowded network (large λ), AFH does not provide any significant advantage over FH in terms of $\eta^{(1)}$.

We have already calculated $\eta^{(2)}_{\text{FH}}$ in example 2 in a system where $3 \leq \lambda \leq 10$. Also,

$$\eta^{(2)}_{\text{AFH}} = \frac{u}{2}\sum_{n=1}^{\infty}\frac{e^{-\lambda}\lambda^n}{n!}\frac{1}{n}\Big(1-\frac{1}{n}\Big)^{n-1}. \tag{2.68}$$

Figure 2.3 presents the plots of $\eta^{(2)}_{\text{FH}}$ and $\eta^{(2)}_{\text{AFH}}$ versus λ for $u = 10$. Both $\eta^{(2)}_{\text{FH}}$ and $\eta^{(2)}_{\text{AFH}}$ decrease by increasing λ. However, the ratio $\frac{\eta^{(2)}_{\text{AFH}}}{\eta^{(2)}_{\text{FH}}}$ decreases as λ increases. This indicates that for large values of λ, AFH does also not provide any significant advantage over FH in terms of $\eta^{(2)}$. $\square$

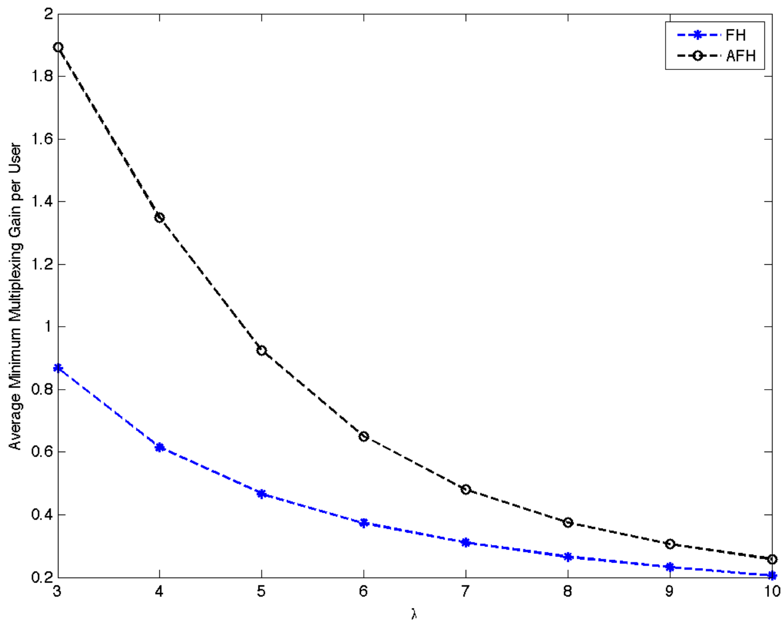

Figure 2.3: Plots of $\eta^{(2)}_{\text{AFH}}$ and $\eta^{(2)}_{\text{FH}}$ in terms of λ for $u = 10$.

2.6 Analysis of the Outage Capacity

As shown in section 2.4, the transmitter of the i^{th} user requires N, $|h_{i,i}|$ and $\sum_{j\neq i} |h_{j,i}|^2$ to regulate its transmission rate. Assuming the transmitters are not aware of the channel gains and the number of active users in the network, the Shannon capacity is not meaningful anymore. In this case, a suitable performance measure is the ε-*outage capacity*, denoted by $R(\varepsilon)$, which is defined as the maximum transmission rate per user ensuring an outage probability below ε. Throughout the rest of the chapter, N is considered to be a realaization of the random variable $\boldsymbol{N}$. Also, we assume the channel gains $\{h_{i,j}\}_{i,j=1}^{N}$ are realizations of independent and circularly symmetric zero-mean complex Gaussian random variables with unit variances denoted by $\{\boldsymbol{h}_{i,j}\}_{i,j=1}^{\boldsymbol{N}}$. Accordingly, $\zeta_{i,l}$ and $a_{i,l}$ are realizations of the random variables $\boldsymbol{\zeta}_{i,l}$ and $\boldsymbol{a}_{i,l}$ for $1 \leq i \leq \boldsymbol{N}$ and $0 \leq l \leq 2^{\boldsymbol{N}-1} - 1$. Since the channel gains are complex, we consider the randomized FH scheme in the complex setup where users transmit independent and circularly symmetric complex Gaussian signals. This causes minor changes in the formulation of the bounds developed earlier. For tractability reasons, we only deal with the primary version of randomized FH, i.e., all users only hop over a *fixed number* $v_1 = v_2 = \cdots = v_{\boldsymbol{N}} = v$ of frequency sub-bands from transmission slot to transmission slot where $0 \leq v \leq u$ is an integer. Moreover, $L_i = 2^{N-1} - 1$ is a realization of $\boldsymbol{L}_i = 2^{\boldsymbol{N}-1} - 1$.

Let $\vec{\boldsymbol{h}}_i$ contain the channel coefficients concerning the i^{th} user, i.e., $\vec{\boldsymbol{h}}_i = \begin{pmatrix} \boldsymbol{h}_{1,i} & \cdots & \boldsymbol{h}_{\boldsymbol{N},i} \end{pmatrix}^{\text{t}}$. In this case, we denote the achievable rate of the i^{th} user by $\mathscr{R}_i(\vec{\boldsymbol{h}}_i)$. The outage event for this user is

$$\mathcal{O}_i(R) \triangleq \{\vec{\boldsymbol{h}}_i : \mathscr{R}_i(\vec{\boldsymbol{h}}_i) < R\}, \tag{2.69}$$

where R is the transmission rate. Hence,

$$R(\varepsilon) = \sup\Big\{R : \Pr\{\mathcal{O}_i(R)\} < \varepsilon\Big\}. \tag{2.70}$$

We emphasize that the randomness of the number of active users is involved in the outage event, as $\boldsymbol{N}$ represents the size of $\vec{\boldsymbol{h}}_i$. Moreover, due to symmetry, the ε-outage capacity is the same for all users.

Since there is no closed formula for $\mathscr{R}_i(\vec{\boldsymbol{h}}_i)$, we need to provide a lower bound $\mathscr{R}_i^{(\text{lb})}(\vec{\boldsymbol{h}}_i)$ on $\mathscr{R}_i(\vec{\boldsymbol{h}}_i)$. Subsequently, using $\mathscr{R}_i^{(\text{lb})}(\vec{\boldsymbol{h}}_i)$, we derive a lower bound on the outage capacity of the i^{th} user as $\sup\Big\{R : \Pr\Big\{\vec{\boldsymbol{h}}_i : \mathscr{R}_i^{(\text{lb})}(\vec{\boldsymbol{h}}_i) < R\Big\} < \varepsilon\Big\}$, and show that this lower bound is higher than the actual outage capacity in the FD scheme in many scenarios.

We have already developed a lower bound on $\mathscr{R}_i(\vec{\boldsymbol{h}}_i)$ in Theorem 2. In the following, we improve this lower bound to obtain better results in terms of outage capacity at a certain finite SNR level.

2.6.1 Lower Bounds on $\mathscr{R}_i(\vec{\boldsymbol{h}}_i)$

Following the same lines as in the proof of Theorem 2,

$$\mathscr{R}_i(\vec{\boldsymbol{h}}_i) \geq v \log\left(\frac{\pi e|\boldsymbol{h}_{i,i}|^2 P}{v} 2^{-\mathrm{h}(\boldsymbol{z}_{i,1})} + 1\right). \tag{2.71}$$

We start with the following Lemma.

Lemma 3. *There exists an upper bound on* $\mathrm{h}(\boldsymbol{z}_{i,1})$ *given by*

$$\mathrm{h}(\boldsymbol{z}_{i,1}) \leq \sum_{l=0}^{\boldsymbol{L}_i} \boldsymbol{a}_{i,l} \log\left(\pi e(\boldsymbol{\zeta}_{i,l} P + \sigma^2)\right) + \boldsymbol{\mathfrak{H}}_i - \boldsymbol{\mathfrak{G}}_i, \tag{2.72}$$

where

$$\boldsymbol{\mathfrak{H}}_i \triangleq -\sum_{l=0}^{\boldsymbol{L}_i} \boldsymbol{a}_{i,l} \log \boldsymbol{a}_{i,l} \tag{2.73}$$

and

$$\boldsymbol{\mathfrak{G}}_i \triangleq \frac{1}{\boldsymbol{\zeta}_{i,\boldsymbol{L}_i}\gamma + 1} \sum_{l=1}^{\boldsymbol{L}_i} \boldsymbol{a}_{i,l} \log\left(1 + \frac{(\boldsymbol{\zeta}_{i,\boldsymbol{L}_i}\gamma + 1)\sum_{m=0}^{l-1} \boldsymbol{a}_{i,m}}{\boldsymbol{a}_{i,l}}\right). \tag{2.74}$$

Proof. See appendix A. $\square$

Applying Lemma 2 to (2.71), we obtain a lower bound on $\mathscr{R}_i(\vec{\boldsymbol{h}}_i)$ as

$$\mathscr{R}_i^{(1)}(\vec{\boldsymbol{h}}_i) \triangleq v \log \left(\frac{2^{-\mathfrak{H}_i} 2^{\mathfrak{G}_i} |\boldsymbol{h}_{i,i}|^2 \gamma}{v \prod_{l=1}^{\boldsymbol{L}_i} (\boldsymbol{\zeta}_{i,l}\gamma + 1)^{\boldsymbol{a}_{i,l}}} + 1 \right). \tag{2.75}$$

We make the following observations:

Observation 3- It can be immediately verified that $\mathfrak{H}_i$ does not depend on the crossover gains. However, $\mathfrak{G}_i$ is implicitly a function of all crossover gains as the partial sums $\sum_{m=1}^{l-1} \boldsymbol{a}_{i,m}$ for $2 \leq l \leq \boldsymbol{L}_i$ depend on the ordering of the crossover gains. This will be investigated more in Lemma 3.

Observation 4- Since $\prod_{l=1}^{\boldsymbol{L}_i}(\boldsymbol{\zeta}_{i,l}\gamma+1)^{\boldsymbol{a}_{i,l}} \leq \prod_{l=1}^{\boldsymbol{L}_i}(\boldsymbol{\zeta}_{i,\boldsymbol{L}_i}\gamma+1)^{\boldsymbol{a}_{i,l}} = (\boldsymbol{\zeta}_{i,\boldsymbol{L}_i}\gamma+1)^{(1-\boldsymbol{a}_{i,0})}$, one obtains a looser version of $\mathscr{R}_i^{(1)}(\vec{\boldsymbol{h}}_i)$ given by

$$\mathscr{R}_i^{(2)}(\vec{\boldsymbol{h}}_i) \triangleq v \log \left(\frac{2^{-\mathfrak{H}_i} 2^{-\mathfrak{G}_i} |\boldsymbol{h}_{i,i}|^2 \gamma}{v (\boldsymbol{\zeta}_{i,\boldsymbol{L}_i}\gamma + 1)^{1-\boldsymbol{a}_{i,0}}} + 1 \right). \tag{2.76}$$

We note that $\mathscr{R}_i^{(2)}(\vec{\boldsymbol{h}}_i)$ still has the same asymptotic expression as that of $\mathscr{R}_i^{(1)}(\vec{\boldsymbol{h}}_i)$ in the high SNR regime. As we will see later, the computational complexity of the lower bound on the outage capacity obtained by using $\mathscr{R}_i^{(2)}(\vec{\boldsymbol{h}}_i)$ is much lower than that of $\mathscr{R}_i^{(1)}(\vec{\boldsymbol{h}}_i)$.

As explained before, $\mathfrak{G}_i$ depends on the ordering of $\{\boldsymbol{\zeta}_{i,l}\}_{l=0}^{\boldsymbol{L}_i}$, which requires analyzing the order statistics of the channel gains. To avoid this, the following Lemma introduces a lower bound on $\mathfrak{G}_i$ that only depends on $\boldsymbol{\zeta}_{i,\boldsymbol{L}_i}$.

Lemma 4. *Let $\boldsymbol{B}$ be a Binomial random variable with parameters $(\boldsymbol{N}-1, p)$ where $p = \frac{v}{u}$. Then,*

$$\mathfrak{H}_i = (\boldsymbol{N}-1)\mathscr{H}(p) \tag{2.77}$$

and $\mathfrak{G}_i \geq \mathfrak{G}_i^{(\mathrm{lb})}$ where

$$\mathfrak{G}_i^{(\mathrm{lb})} \triangleq \frac{1}{\boldsymbol{\zeta}_{i,\boldsymbol{L}_i}\gamma + 1} \Big(\mathrm{E}_{\boldsymbol{B}} \left\{ \log \left(1 + \left(1 - p^{\boldsymbol{B}}\right) \boldsymbol{\zeta}_{i,\boldsymbol{L}_i}\gamma \right) \right\} - (\boldsymbol{N}-1) p \log p \Big). \tag{2.78}$$

Proof. As $\boldsymbol{L}_i = 2^{\boldsymbol{N}-1} - 1$ with probability 1 and each user selects a certain frequency sub-band with probability p, the collection $\{\boldsymbol{a}_{i,l}\}_{l=0}^{\boldsymbol{L}_i}$ consists of the numbers $p^j(1-p)^{\boldsymbol{N}-1-j}$ repeated $\binom{\boldsymbol{N}-1}{j}$ times for $0 \leq j \leq \boldsymbol{N}-1$. Hence,

$$\begin{aligned} \mathfrak{H}_i &= -\sum_{j=0}^{\boldsymbol{N}-1} \Pr\{\boldsymbol{B} = j\} \log\left(p^j (1-p)^{\boldsymbol{N}-1-j}\right) \\ &= -(\boldsymbol{N}-1)p\log p - (\boldsymbol{N}-1-(\boldsymbol{N}-1)p)\log(1-p) \\ &= -(\boldsymbol{N}-1)\left(p\log p + (1-p)\log(1-p)\right). \end{aligned} \tag{2.79}$$

As for $\mathfrak{G}_i$, computation of $\sum_{m=0}^{l-1}\boldsymbol{a}_{i,m}$ is not an easy task. In fact, it depends on the ordering of the crossover gains. For example, if $\boldsymbol{N}=4$, $l=4$ and $i=1$,

$$\sum_{m=0}^{3}\boldsymbol{a}_{1,m}=\begin{cases}(1-p)^3+2p(1-p)^2+p^2(1-p) & |\boldsymbol{h}_{2,1}|^2<|\boldsymbol{h}_{3,1}|^2<|\boldsymbol{h}_{2,1}|^2+|\boldsymbol{h}_{3,1}|^2<|\boldsymbol{h}_{4,1}|^2\\ (1-p)^3+3p(1-p)^2 & |\boldsymbol{h}_{2,1}|^2<|\boldsymbol{h}_{3,1}|^2<|\boldsymbol{h}_{4,1}|^2<|\boldsymbol{h}_{2,1}|^2+|\boldsymbol{h}_{3,1}|^2\end{cases}. \tag{2.80}$$

To avoid this difficulty in describing $\mathfrak{G}_i$, we derive a lower bound on this quantity, which is not sensitive to the ordering of crossover gains. Taking each $\boldsymbol{a}_{i,l}$, there exists a $0\le n\le \boldsymbol{N}-1$ such that $\boldsymbol{a}_{i,l}=p^n(1-p)^{\boldsymbol{N}-1-n}$. This implies that $\boldsymbol{a}_{i,l}$ corresponds to the interference plus noise power level $\frac{\sum_{k=1}^{n}|\boldsymbol{h}_{j_k,i}|^2}{v}P+\sigma^2$ for some $1\le j_1<\cdots<j_n\le\boldsymbol{N}$ where $j_k\ne i$ for $1\le k\le n$. Since $\frac{\sum_{k=1}^{n}|\boldsymbol{h}_{j_k,i}|^2}{v}P+\sigma^2>\frac{\sum_{t\in\mathcal{A}\subsetneq\{1,2,\cdots,n\}}|\boldsymbol{h}_{j_t,i}|^2}{v}P+\sigma^2$ for any set $\mathcal{A}\subsetneq\{1,2,\cdots,n\}$, and $\frac{\sum_{t\in\mathcal{A}\subsetneq\{1,2,\cdots,n\}}|\boldsymbol{h}_{j_t,i}|^2}{v}P+\sigma^2$ is itself a power level in the PDF of the noise plus interference on each frequency sub-band, we conclude that its associated probability $p^{|\mathcal{A}|}(1-p)^{\boldsymbol{N}-1-|\mathcal{A}|}$ is an element in the sequence $(\boldsymbol{a}_{i,0},\boldsymbol{a}_{i,1},\cdots,\boldsymbol{a}_{i,l-1})$. Therefore, we come up with the following lower bound,

$$\begin{aligned}\sum_{m=0}^{l-1}\boldsymbol{a}_{i,m} &\ge \sum_{\mathcal{A}\subsetneq\{1,2,\cdots,n\}}p^{|\mathcal{A}|}(1-p)^{\boldsymbol{N}-1-|\mathcal{A}|}\\ &= \sum_{n'=0}^{n-1}\binom{n}{n'}p^{n'}(1-p)^{\boldsymbol{N}-1-n'}.\end{aligned} \tag{2.81}$$

Using (2.81) in (2.74), one can develop a lower bound on $\mathfrak{G}_i$ as in (2.82).

$$\begin{aligned}\mathfrak{G}_i &\ge \frac{1}{\boldsymbol{\zeta}_{i,\boldsymbol{L}_i}\gamma+1}\sum_{n=1}^{\boldsymbol{N}-1}\binom{\boldsymbol{N}-1}{n}p^n(1-p)^{\boldsymbol{N}-1-n}\log\left(1+\frac{(\boldsymbol{\zeta}_{i,\boldsymbol{L}_i}\gamma+1)\sum_{n'=0}^{n-1}\binom{n}{n'}p^{n'}(1-p)^{\boldsymbol{N}-1-n'}}{p^n(1-p)^{\boldsymbol{N}-1-n}}\right)\\ &= \frac{1}{\boldsymbol{\zeta}_{i,\boldsymbol{L}_i}\gamma+1}\sum_{n=1}^{\boldsymbol{N}-1}\binom{\boldsymbol{N}-1}{n}p^n(1-p)^{\boldsymbol{N}-1-n}\log\left(1+\frac{(\boldsymbol{\zeta}_{i,\boldsymbol{L}_i}\gamma+1)\sum_{n'=0}^{n-1}\binom{n}{n'}p^{n'}(1-p)^{n-n'}}{p^n}\right)\\ &= \frac{1}{\boldsymbol{\zeta}_{i,\boldsymbol{L}_i}\gamma+1}\sum_{n=1}^{\boldsymbol{N}-1}\binom{\boldsymbol{N}-1}{n}p^n(1-p)^{\boldsymbol{N}-1-n}\log\left(1+\frac{(1-p^n)(\boldsymbol{\zeta}_{i,\boldsymbol{L}_i}\gamma+1)}{p^n}\right)\\ &= -\frac{1}{\boldsymbol{\zeta}_{i,\boldsymbol{L}_i}\gamma+1}\sum_{n=0}^{\boldsymbol{N}-1}\binom{\boldsymbol{N}-1}{n}np^n(1-p)^{\boldsymbol{N}-1-n}\log p\\ &\quad+\frac{1}{\boldsymbol{\zeta}_{i,\boldsymbol{L}_i}\gamma+1}\sum_{n=0}^{\boldsymbol{N}-1}\binom{\boldsymbol{N}-1}{n}p^n(1-p)^{\boldsymbol{N}-1-n}\log\left(1+(1-p^n)\boldsymbol{\zeta}_{i,\boldsymbol{L}_i}\gamma\right)\\ &= \frac{\mathrm{E}_{\boldsymbol{B}}\left\{\log\left(1+(1-p^{\boldsymbol{B}})\boldsymbol{\zeta}_{i,\boldsymbol{L}_i}\gamma\right)\right\}-(\boldsymbol{N}-1)p\log p}{\boldsymbol{\zeta}_{i,\boldsymbol{L}_i}\gamma+1}.\end{aligned} \tag{2.82}$$

This concludes the Lemma. □

From now on, we replace $\mathfrak{G}_i$ by $\mathfrak{G}_i^{(\mathrm{lb})}$ defined in (2.78) in all expressions offered for the lower bounds on $\mathscr{R}_i(\vec{\boldsymbol{h}}_i)$. We denote $\boldsymbol{a}_{i,0}$, $\mathfrak{H}_i$ and $\mathfrak{G}_i^{(\mathrm{lb})}$ by $a(v,\boldsymbol{N}) = \left(1-\frac{v}{u}\right)^{\boldsymbol{N}-1}$, $\mathfrak{H}(v,\boldsymbol{N})$ and $\mathfrak{G}_i^{(\mathrm{lb})}(v,\boldsymbol{N})$ respectively[9], to emphasize their dependence on $v, \boldsymbol{N}$.

As a special case, let us assume $v = u$, i.e., all users spread their power on the whole spectrum. This scheme is called Full-Band Spreading (FBS). In this case, it can be observed that $\boldsymbol{a}_{i,l} = 0$ for $l \leq \boldsymbol{L}_i - 1$ and $\boldsymbol{a}_{i,\boldsymbol{L}_i} = 1$. This yields $a(u,\boldsymbol{N}) = \mathfrak{H}(u,\boldsymbol{N}) = \mathfrak{G}_i^{(\mathrm{lb})}(u,\boldsymbol{N}) = 0$. In fact, $\mathscr{R}_i^{(2)}(\vec{\boldsymbol{h}}_i)$ is tight for $v = u$, i.e., $\mathscr{R}_i^{(2)}(\vec{\boldsymbol{h}}_i)$ is exactly the achievable rate of the i^{th} user when all users transmit over the whole spectrum. We denote this rate by $\mathscr{R}_{i,\mathrm{FBS}}(\vec{\boldsymbol{h}}_i)$, which is given by

$$\mathscr{R}_{i,\mathrm{FBS}}(\vec{\boldsymbol{h}}_i) = u \log\left(\frac{|\boldsymbol{h}_{i,i}|^2\gamma}{u\left(\frac{\boldsymbol{\zeta}_{i,\boldsymbol{L}_i}\gamma}{u}+1\right)}+1\right). \tag{2.83}$$

Observation 5- A straightforward method to develop a lower bound on the achievable rate in an additive white non-Gaussian noise channel is to replace the noise with a Gaussian noise of the same covariance matrix [82]. Following this approach, it is easy to derive the following lower bound on $\mathscr{R}_i(\vec{\boldsymbol{h}}_i)$,

$$\mathscr{R}_{i,\mathrm{g}}(\vec{\boldsymbol{h}}_i) \triangleq v \log\left(\frac{|\boldsymbol{h}_{i,i}|^2\gamma}{v\left(\frac{\boldsymbol{\zeta}_{i,\boldsymbol{L}_i}\gamma}{u}+1\right)}+1\right) \tag{2.84}$$

where the index "g" stands for Gaussian. There are two facts that are worth mentioning about $\mathscr{R}_{i,\mathrm{g}}(\vec{\boldsymbol{h}}_i)$. First, it is seen that $\lim_{\gamma\to\infty}\mathscr{R}_{i,\mathrm{g}}(\vec{\boldsymbol{h}}_i) < \infty$. Another point is that $\mathscr{R}_{i,\mathrm{g}}(\vec{\boldsymbol{h}}_i)$ is an increasing function of v. However, setting $v = u$ in the expression of $\mathscr{R}_{i,\mathrm{g}}(\vec{\boldsymbol{h}}_i)$ yields the expression of $\mathscr{R}_{i,\mathrm{FBS}}(\vec{\boldsymbol{h}}_i)$. Therefore, for all realizations of the channel gains and all ranges of γ,

$$\mathscr{R}_{i,\mathrm{g}}(\vec{\boldsymbol{h}}_i) \leq \mathscr{R}_{i,\mathrm{FBS}}(\vec{\boldsymbol{h}}_i). \tag{2.85}$$

This indicates that using $\mathscr{R}_{i,\mathrm{g}}(\vec{\boldsymbol{h}}_i)$ as a lower bound on the achievable rate of users in the FH scheme provides no proof of advantage for FH over FBS.

2.6.2 Lower Bounds on $R_{\mathrm{FH}}(\varepsilon)$ and System Design

Based on the preceding discussion, we can derive lower bounds on $R_{\mathrm{FH}}(\varepsilon)$, namely $R_{\mathrm{FH}}^{(1)}(\varepsilon)$ and $R_{\mathrm{FH}}^{(2)}(\varepsilon)$ associated with the lower bounds $\mathscr{R}_i^{(1)}(\vec{\boldsymbol{h}}_i)$ and $\mathscr{R}_i^{(2)}(\vec{\boldsymbol{h}}_i)$ respectively. Consequently, we can obtain estimates of the optimum hopping parameter[10] by maximizing $R_{\mathrm{FH}}^{(1)}(\varepsilon)$ or $R_{\mathrm{FH}}^{(2)}(\varepsilon)$ over v. In the following subsections, we separately compute $R_{\mathrm{FH}}^{(1)}(\varepsilon)$ and $R_{\mathrm{FH}}^{(2)}(\varepsilon)$.

1- Computation of $R_{\mathrm{FH}}^{(1)}(\varepsilon)$

[9]We note that $\boldsymbol{a}_{i,0}$ and $\mathfrak{H}_i$ are similar for different i, however, $\mathfrak{G}_i^{(\mathrm{lb})}$ depends on i through $\sum_{j\neq i}|\boldsymbol{h}_{j,i}|^2$.

[10]The optimum hopping parameter is the value of v that maximizes $R_{\mathrm{FH}}(\varepsilon)$.

We start with the following definitions.

Definition 1- Let $n \in \mathbb{N}$. For $p \in [0,1]$ and $b > 0$, the function $\alpha_n(.;b,p) : \mathbb{R}^+ \to \mathbb{R}^+$ is defined by

$$\alpha_n(\theta; b, p) \triangleq \frac{\mathrm{E}_{\boldsymbol{B}}\left\{\log\left(1 + (1 - p^{\boldsymbol{B}})b\theta\right)\right\} - (n-1)p\log p}{b\theta + 1}, \tag{2.86}$$

where $\boldsymbol{B}$ is a Binomial random variable with parameters $(n-1, p)$. $\square$

Definition 2- Let $n \geq 2$ be an integer. For $b_1 < 0$, $b_2 > 0$ and $p \in [0,1]$, we define $\psi_n(b_1, b_2, p)$ as

$$\psi_n(b_1, b_2, p) \triangleq \int_{[0,\infty)^{n-1}} \exp\left(b_1 2^{-\alpha_n(\theta_{n-1,1}; b_2, p)} \prod_{m=1}^{n-1} \prod_{m'=1}^{\binom{n-1}{m}} (b_2\theta_{m,m'} + 1)^{\beta_{m,n}(p)} - \theta_{n-1,1}\right) d\theta_1 \cdots d\theta_{n-1} \tag{2.87}$$

where for each m, $\{\theta_{m,m'}\}_{m'=1}^{\binom{n-1}{m}}$ consists of all possible summations of m elements in the set of dummies $\{\theta_i\}_{i=1}^{n-1}$ and $\beta_{m,n}(p) \triangleq p^m(1-p)^{n-1-m}$. $\square$

For example,

$$\psi_2(b_1, b_2, p) = \int_0^\infty \exp\left(b_1 2^{-\alpha_2(\theta; b_2, p)}(b_2\theta + 1)^p - \theta\right) d\theta, \tag{2.88}$$

and

$$\psi_3(b_1, b_2, p) = \int_{\theta_1, \theta_2 > 0} \exp\left(b_1 2^{-\alpha_3(\theta_{2,1}; b_2, p)} \left((b_2\theta_1 + 1)(b_2\theta_2 + 1)\right)^{p(1-p)} (b_2\theta_{2,1} + 1)^{p^2} - \theta_{2,1}\right) d\theta_1 d\theta_2, \tag{2.89}$$

where $\theta_{2,1} = \theta_1 + \theta_2$ by definition.

The following Proposition offers an expression to compute $R_{\mathrm{FH}}^{(1)}(\varepsilon)$.

Proposition 2.

$$R_{\mathrm{FH}}^{(1)}(\varepsilon) = \sup\left\{R : \Xi_1(R) > \overline{\varepsilon}\right\} \tag{2.90}$$

where

$$\Xi_1(R) = q_1 e^{\frac{v}{\gamma}\left(1 - 2^{\frac{R}{v}}\right)} + \sum_{n=2}^{n_{\max}} q_n \psi_n\left(b_{1,n}, b_2, p\right), \tag{2.91}$$

$b_{1,n} = \frac{2^{\mathfrak{H}(v,n)}\left(1 - 2^{\frac{R}{v}}\right)v}{\gamma}$, $b_2 = \frac{\gamma}{v}$ *and* $p = \frac{v}{u}$.

Proof. See appendix D. $\square$

The expression given in (2.90) is quite complicated. On one hand, the multiple integrals do not have a closed form. On the other hand, the maximization $\max_v R_{\mathrm{FH}}^{(1)}(\varepsilon)$ must be computed numerically. However, $R_{\mathrm{FH}}^{(1)}(\varepsilon)$ is the best lower bound on $R_{\mathrm{FH}}(\varepsilon)$ as $\mathscr{R}_i^{(1)}(\vec{\boldsymbol{h}}_i)$ is the best lower bound we have found on the achievable rate of the i^{th} user in the FH scenario.

2- Computation of $R_{\mathrm{FH}}^{(2)}(\varepsilon)$

We start with the following definition.

Definition 3- Let $n \geq 2$ be an integer. For $b_1 < 0$, $b_2 > 0$ and $p_1, p_2 \in [0,1]$, we define $\phi_n(b_1, b_2, p_1, p_2)$ as

$$\phi_n(b_1, b_2, p_1, p_2) \triangleq \frac{1}{(n-2)!} \int_0^{\infty} \theta^{n-2} \exp\left(b_1(b_2\theta + 1)^{p_1} 2^{-\alpha_n(\theta; b_2, p_2)} - \theta\right) d\theta. \tag{2.92}$$

Using this class of functions, the following Proposition yields $R_{\mathrm{FH}}^{(2)}(\varepsilon)$.

Proposition 3.

$$R_{\mathrm{FH}}^{(2)}(\varepsilon) = \sup\left\{R : \Xi_2(R) > \overline{\varepsilon}\right\}, \tag{2.93}$$

where

$$\Xi_2(R) = q_1 e^{\frac{v}{\gamma}\left(1 - 2^{\frac{R}{v}}\right)} + \sum_{n=2}^{n_{\max}} q_n \phi_n\left(b_{1,n}, b_2, p_{1,n}, p_2\right), \tag{2.94}$$

$p_{1,n} = 1 - a(v,n)$, $p_2 = \frac{v}{u}$ *and* $b_{1,n}$ *and* b_2 *are given in Proposition 2.*

Proof. The proof follows from similar steps along the lines of the proof for Proposition 2. □

Comparing the expressions for $R_{\mathrm{FH}}^{(1)}(\varepsilon)$ and $R_{\mathrm{FH}}^{(2)}(\varepsilon)$, it can be observed that computation of $R_{\mathrm{FH}}^{(2)}(\varepsilon)$ involves only single integrations, while the computation of $R_{\mathrm{FH}}^{(1)}(\varepsilon)$ involves iterated integrations that are not tractable for many cases. To further reduce the complexity of computation, the following Corollary yields another lower bound on $R_{\mathrm{FH}}(\varepsilon)$ that involves no numerical integrations. We denote this lower bound by $R_{\mathrm{FH}}^{(3)}(\varepsilon)$.

Corollary 1. *Let*

$$R_{\mathrm{FH}}^{(3)}(\varepsilon) \triangleq \sup\left\{R : \Xi_3(R) > \overline{\varepsilon}\right\} \tag{2.95}$$

where

$$\Xi_3(R) = \sum_{n=1}^{n_{\max}} q_n \exp\left(b_{1,n}\left((n-1)b_2 + 1\right)^{1-a(v,n)}\right) \tag{2.96}$$

and $b_{1,n}$ *and* b_2 *are given in Proposition 2. Then,*

$$R_{\mathrm{FH}}(\varepsilon) \geq R_{\mathrm{FH}}^{(1)}(\varepsilon) \geq R_{\mathrm{FH}}^{(2)}(\varepsilon) \geq R_{\mathrm{FH}}^{(3)}(\varepsilon). \tag{2.97}$$

Proof. See appendix E. □

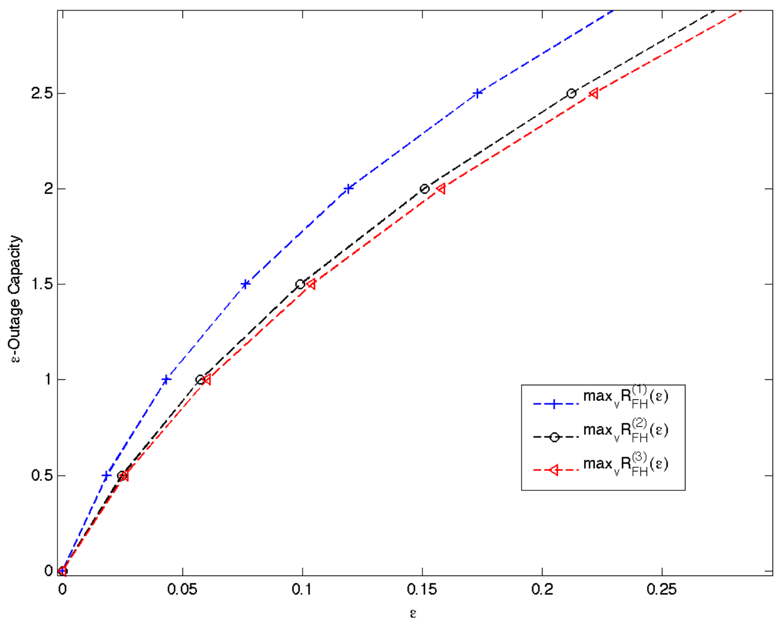

Figure 2.4: Sketch of $\max_v R_{\mathrm{FH}}^{(k)}(\varepsilon)$ (bits/transmission slot) for $k = 1, 2, 3$ in a setup where $(q_1, q_2, q_3, q_4) = (0.4, 0.2, 0.2, 0.2)$, $u = 8$ and $\gamma = 20$dB.

Figure 2.4 shows the three lower bounds $\max_v R_{\mathrm{FH}}^{(k)}(\varepsilon)$ on $\max_v R_{\mathrm{FH}}(\varepsilon)$ for $k = 1, 2, 3$ in a system where, at most, four users become active simultaneously with $(q_1, q_2, q_3, q_4) = (0.4, 0.2, 0.2, 0.2)$, $u = 8$ and $\gamma = 20$dB. As can be observed from this figure, $\max_v R_{\mathrm{FH}}^{(2)}(\varepsilon)$ and $\max_v R_{\mathrm{FH}}^{(3)}(\varepsilon)$ are pretty close to each other while they are within a considerable gap to $\max_v R_{\mathrm{FH}}^{(1)}(\varepsilon)$, especially for larger values of ε. It is notable that the maximization over v is performed separately for each ε.

2.6.3 Asymptotic Analysis

Having the expression for the ε-outage capacity, we can find the best operational point of the system in terms of v, the number of selected sub-bands. For this purpose, we consider some asymptotic cases in terms of ε and γ and discuss the optimum value of v in these regimes. As the expressions of $R_{\mathrm{FH}}^{(1)}(\varepsilon)$ and $R_{\mathrm{FH}}^{(2)}(\varepsilon)$ are not analytically tractable, we use $R_{\mathrm{FH}}^{(3)}(\varepsilon)$ for our analysis. For simulation purposes, we use $R_{\mathrm{FH}}^{(1)}(\varepsilon)$ or $R_{\mathrm{FH}}^{(2)}(\varepsilon)$.

Asymptotically Small ε

In this case, one can easily show that $\lim_{\varepsilon \to 0} R_{\mathrm{FH}}^{(3)}(\varepsilon) = 0$. Therefore, in (2.95), we can approximate $\frac{\left(1-2^{\frac{R}{v}}\right)v}{\gamma}$ by $-\frac{R \ln 2}{\gamma}$ and hence,

$$e^{\frac{v}{\gamma}\left(1-2^{\frac{R}{v}}\right)} \approx 1 - \frac{R\ln(2)}{\gamma}. \tag{2.98}$$

By the same token, the term on the right side of (2.95) can be approximated as

$$\begin{aligned} & \sum_{n=1}^{n_{\max}} q_n \exp\left(b_{1,n}\big((n-1)b_2+1\big)^{1-a(v,n)}\right) \\ \approx & \sum_{n=1}^{n_{\max}} q_n \exp\left(-\frac{R\ln 2}{\gamma} 2^{\mathfrak{H}(v,n)} \left(\frac{n-1}{v}\gamma + 1\right)^{1-a(v,n)}\right) \\ \overset{(a)}{\approx} & \sum_{n=1}^{n_{\max}} q_n \left(1 - \frac{R\ln 2}{\gamma} f(v,n,\gamma)\right), \end{aligned} \tag{2.99}$$

where (a) follows from the fact that the term

$$f(v,n,\gamma) \triangleq 2^{\mathfrak{H}(v,n)} \left(\frac{n-1}{v}\gamma + 1\right)^{1-a(v,n)} \tag{2.100}$$

does not depend on ε. Using (2.99) in (2.95) yields

$$R_{\mathrm{FH}}^{(3)}(\varepsilon) \approx \frac{\varepsilon\gamma}{\left(\sum_{n=1}^{n_{\max}} q_n f(v,n,\gamma)\right)\ln 2}. \tag{2.101}$$

One can observe that the function $f(v,n,\gamma)$ is concave in terms of v for all $n \geq 1$. Hence, the function $g(v) \triangleq \sum_{n=1}^{n_{\max}} q_n f(v,n,\gamma)$ is concave as well. As such, the minimum of $g(v)$ occurs either at $v = 1$ or $v = u$. In partiular, for $v = 1$ to be the optimum value, we must have $g(1) < g(u)$ or equivalently,

$$\begin{aligned} \sum_{n=1}^{n_{\max}} \frac{q_n u^{n-1}\big((n-1)\gamma+1\big)^{1-\left(1-\frac{1}{u}\right)^{n-1}}}{(u-1)^{\left(1-\frac{1}{u}\right)(n-1)}} &< \sum_{n=1}^{n_{\max}} q_n \left(1 + \frac{(n-1)\gamma}{u}\right) \\ &= 1 + \frac{\mathrm{E}\{N\}-1}{u}\gamma, \end{aligned} \tag{2.102}$$

where we have used the fact that $2^{\mathfrak{H}(1,n)} = \left(\frac{u}{(u-1)^{1-\frac{1}{u}}}\right)^{n-1}$.

Figure 2.5 offers the curves of $R_{\mathrm{FH}}^{(1)}(\varepsilon)$ in terms of v for $\varepsilon = 0.01, 0.05$ and 0.15. The underlying network is characterized with $q_1 = q_2 = 0.5$, $u = 10$ and $\gamma = 20$dB. It is seen that for $\varepsilon = 0.01$,

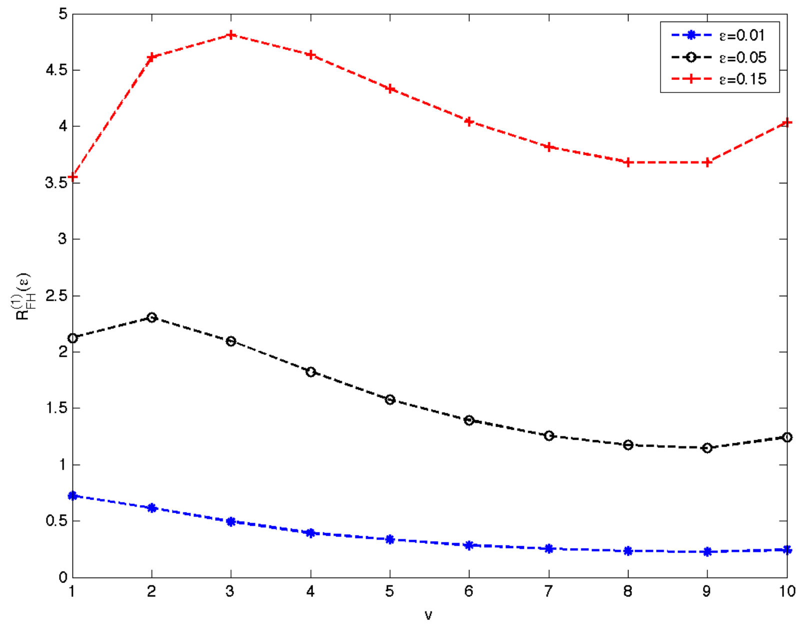

Figure 2.5: Sketch of $R^{(1)}_{\text{FH}}(\varepsilon)$ (bits/transmission slot) for different values of ε in a network with $(q_1, q_2) = (0.5, 0.5)$, $u = 10$ and $\gamma = 20$dB.

taking $v = 1$ yields the best performance. As far as $q_1 = q_2 = 0.5$ and for sufficiently small ε, the equation (2.102) simplifies to $\frac{u^u}{(u-1)^{u-1}} < \frac{1}{1+\gamma}\left(1 + \frac{\gamma}{u}\right)^u$. Setting $u = 10$ yields 7.052dB as the minimum SNR value that guarantees $v = 1$ is the best choice. Since $\gamma = 20\text{dB} > 7.052\text{dB}$, we expect $v = 1$ is the best choice, which is confirmed in the plot. However, as ε increases, we move away from the *asymptotically small* ε region and $v = 1$ is no longer an optimal choice.

Asymptotically Small γ

In this case, one can easily show that $\lim_{\gamma \to 0} R^{(3)}_{\text{FH}}(\varepsilon) = 0$. Therefore, similar to the previous case, we can use the approximation

$$\frac{\left(1 - 2^{\frac{R}{v}}\right) v}{\gamma} \approx -\frac{R \ln 2}{\gamma}. \tag{2.103}$$

Defining $\tau \triangleq \exp\left(-\frac{R \ln 2}{\gamma}\right)$, one can rewrite (2.95) as

$$R^{(3)}_{\text{FH}}(\varepsilon) \approx \sup\left\{ R : \sum_{n=1}^{n_{\max}} q_n \tau^{f(v,n,\gamma)} > \bar{\varepsilon} \right\}, \tag{2.104}$$

where $f(v,n,\gamma)$ is given in (2.100). As τ is a decreasing function in terms of R, we can write

$$R_{\mathrm{FH}}^{(3)}(\varepsilon) \;=\; -\frac{\gamma \ln(\tau^*)}{\ln 2}, \tag{2.105}$$

where

$$\tau^* = \inf\left\{\tau : \sum_{n=1}^{n_{\max}} q_n \tau^{f(v,n,\gamma)} > \overline{\varepsilon}\right\}. \tag{2.106}$$

Therefore,

$$\max_v R_{\mathrm{FH}}^{(3)}(\varepsilon) \approx -\frac{\gamma}{\ln 2} \ln\left(\min_v \tau^*\right). \tag{2.107}$$

It can be shown that $\min_v \tau^*$ occurs when $f(v,n,\gamma)$ takes its minimum value over v for each n. But, as $\gamma \ll 1$, the term $\left(\frac{n-1}{v}\gamma+1\right)^{1-a(v,n)} \approx 1$ and hence, $f(v,n,\gamma) \approx 2^{\mathfrak{H}(v,n)}$, which is uniformly minimized for all values of n by taking $v=u$. This gives $\tau^* = \overline{\varepsilon}$, and hence,

$$R_{\mathrm{FH}}^{(3)}(\varepsilon) \approx -\gamma \log \overline{\varepsilon}. \tag{2.108}$$

This is exactly the outage capacity of a point-to-point system without interference. Therefore, in the low SNR regime, interference has no destructive effect on the outage capacity justifying the optimality of $v=u$.

Figure 2.6 presents the plot of $R_{\mathrm{FH}}^{(1)}(0.1)$ versus v for $\gamma = -10$dB in a system with $q_1 = q_2 = 0.5$ and $u = 10$. It is seen that $v = u = 10$ is the best choice, which is expected by our analysis. Also, we observe that the outage capacity is not very sensitive to the value of v.

Asymptotically High γ

Recalling the expression of $R_{\mathrm{FH}}^{(3)}(\varepsilon)$ given in (2.95), we have

$$R_{\mathrm{FH}}^{(3)}(\varepsilon) = \sup\left\{R : \sum_{n=1}^{n_{\max}} q_n e^{-\frac{v}{\gamma} f(v,n,\gamma)\left(2^{R/v}-1\right)} > \overline{\varepsilon}\right\}. \tag{2.109}$$

As $\gamma \to \infty$, the term $f(v,n,\gamma)$ grows polynomially with γ for any $n \geq 1$. The outage event is determined by the term with the maximum $f(v,n,\gamma)$. In fact, since $f(v,n,\gamma)$ is an increasing function in terms of n, it follows that $\sum_{n=1}^{n_{\max}} q_n e^{-\frac{v}{\gamma} f(v,n,\gamma)\left(2^{R/v}-1\right)} \approx \sum_{n=1}^{n_{\max}-1} q_n + q_{n_{\max}} e^{-\frac{v}{\gamma} f(v,n_{\max},\gamma)\left(2^{R/v}-1\right)}$. Therefore, (2.109) simplifies to

$$R_{\mathrm{FH}}^{(3)}(\varepsilon) \approx \sup\left\{R : e^{-\frac{v}{\gamma} f(v,n_{\max},\gamma)\left(2^{R/v}-1\right)} > 1 - \frac{\varepsilon}{q_{n_{\max}}}\right\}, \tag{2.110}$$

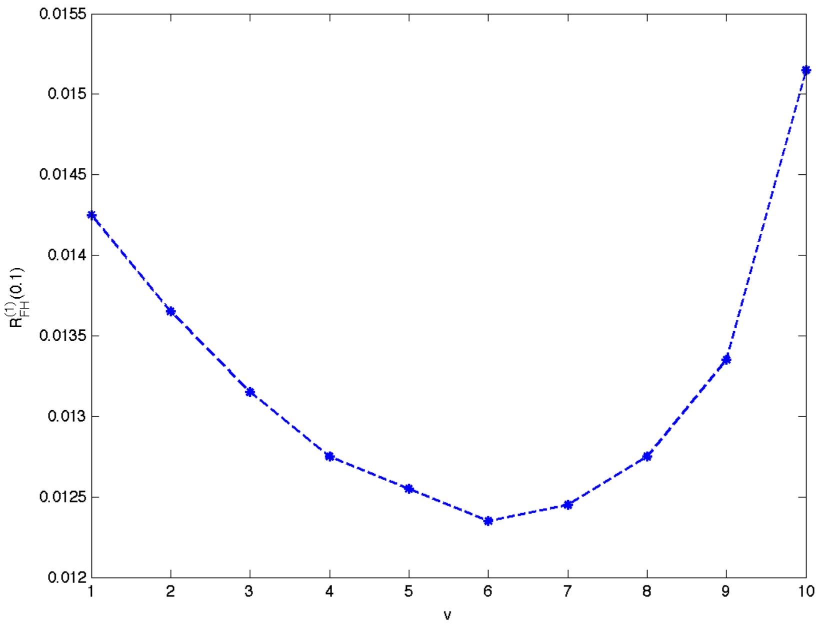

Figure 2.6: Sketch of $R^{(1)}_{\mathrm{FH}}(0.1)$ (bits/transmission slot) in a network with $(q_1, q_2) = (0.5, 0.5)$, $\gamma = -10$dB and $u = 10$.

which gives

$$\begin{aligned} R^{(3)}_{\mathrm{FH}}(\varepsilon) &\approx v \log\left(1 - \frac{\gamma \ln\left(1 - \frac{\varepsilon}{q_{n_{\max}}}\right)}{v f(v, n_{\max}, \gamma)}\right) \\ &= v \log\left(1 - \frac{\gamma 2^{-\mathfrak{H}(v, n_{\max})} \ln\left(1 - \frac{\varepsilon}{q_{n_{\max}}}\right)}{v\left(1 + \frac{n_{\max}-1}{v}\gamma\right)^{1-\left(1-\frac{v}{u}\right)^{n_{\max}-1}}}\right) \\ &\overset{(a)}{\approx} v \log\left(1 - \kappa_v \gamma^{\left(1-\frac{v}{u}\right)^{n_{\max}-1}} \ln\left(1 - \frac{\varepsilon}{q_{n_{\max}}}\right)\right), \end{aligned} \tag{2.111}$$

where

$$\kappa_v \triangleq \frac{2^{-\mathfrak{H}(v, n_{\max})}}{v}\left(\frac{n_{\max}-1}{v}\right)^{\left(1-\frac{v}{u}\right)^{n_{\max}-1}-1} \tag{2.112}$$

and (a) follows from the assumption that γ lies in the high SNR range. It is observed that the maximization of the above expression with respect to v is equivalent to the maximization of the

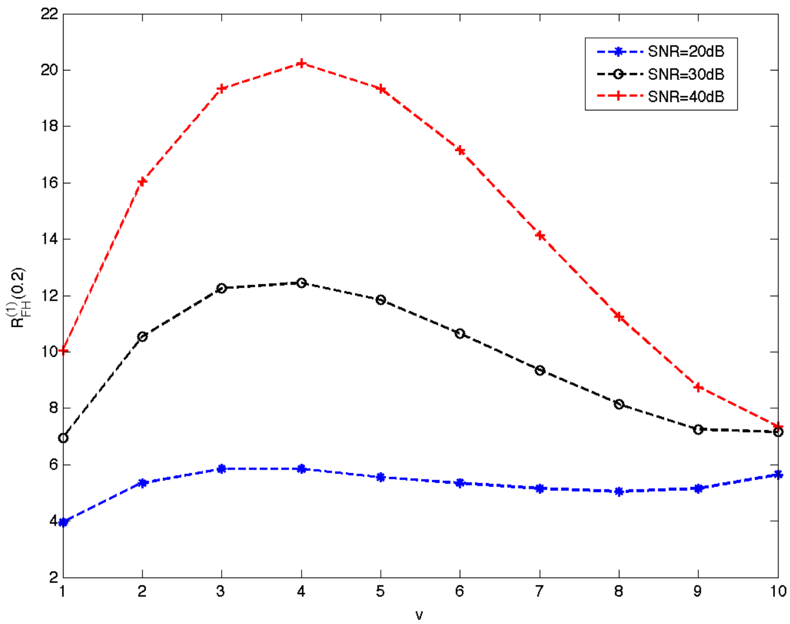

Figure 2.7: Sketch of $R^{(1)}_{\text{FH}}(0.2)$ (bits/transmission slot) at different SNR levels in a network with $n_{\max} = 2$, $(q_1, q_2) = (0.5, 0.5)$ and $u = 10$.

term $v\left(1 - \frac{v}{u}\right)^{n_{\max}-1}$ with respect to v, as SNR tends to infinity. This yields

$$v = \left\lceil \frac{u}{n_{\max}} \right\rceil. \tag{2.113}$$

Figure 2.7 shows the curves of $R^{(1)}_{\text{FH}}(0.2)$ versus v for different values of SNR in a system with parameters $n_{\max} = 2$, $(q_1, q_2) = (0.5, 0.5)$ and $u = 10$. In general, for a sufficiently large, however finite, value of SNR, one can obtain the optimum value for v as

$$v = \arg \max_{1 \leq \mathsf{v} \leq u} \left\{ \mathsf{v} \log \left(1 - \kappa_{\mathsf{v}} \gamma^{\left(1-\frac{\mathsf{v}}{u}\right)^{n_{\max}-1}} \ln \left(1 - \frac{\varepsilon}{q_{n_{\max}}} \right) \right) \right\}. \tag{2.114}$$

For example, it is easy to verify that in a system with the above parameters at $\gamma = 40$dB, one gets $v = 4$, while $\left\lceil \frac{u}{n_{\max}} \right\rceil = \left\lceil \frac{10}{2} \right\rceil = 5$. This is in agreement with the plot of $R^{(1)}_{\text{FH}}(0.2)$ given in fig. 2.7 for $\gamma = 40$dB.

2.6.4 Comparison with other Schemes

In this section, we compare the performance of the proposed FH scenario with that of the FD scheme in terms of the ε-outage capacity. In the FD scheme, the spectrum is primarily divided into $n_{\text{des}} = \min\{n_{\max}, u\}$ bands and each active user only occupies one band. If $n_{\max} \le u$ and all the $n_{\max}$ users are active all the time, i.e., $q_{n_{\max}} = 1$, this scheme results in the most efficient usage of the bandwidth. This makes the FD scenario superior to other schemes proposed in the literature, especially in the high SNR regime. However, in a practical situation, the number of concurrently active users is smaller than n_{des} with a nonzero probability. This makes FD highly inefficient on the heels that a considerable portion of the sub-bands is unused. In addition to FD, we also study the ε-outage capacity of the FBS scenario, which is a special case of FH. In fact, FD and FBS can be considered as two extreme spectrum management schemes where the former avoids any interference among the users, while the latter makes all users share the same spectrum all the time. In the sequel, we compute $R_{\text{FD}}(\varepsilon)$ and $R_{\text{FBS}}(\varepsilon)$.

Computation of $R_{\text{FD}}(\varepsilon)$

In the FD scenario, the spectrum is already divided into n_{des} non-overlaping bands each containing $\frac{u}{n_{\text{des}}}$ sub-bands. Each user that becomes active occupies one of the units. As there is no interference among users, the outage event for the i^{th} user can be written as

$$\mathcal{O}_{i,\text{FD}}(R) = \left\{ \boldsymbol{h}_{i,i} : \frac{u}{n_{\text{des}}} \log\left(1 + \frac{n_{\text{des}}|\boldsymbol{h}_{i,i}|^2\gamma}{u}\right) < R \right\}. \tag{2.115}$$

As $\boldsymbol{h}_{i,i}$ is a complex Gaussian random variable with variance $\frac{1}{2}$ per dimension, $|\boldsymbol{h}_{i,i}|^2$ is an exponential random variable with parameter 1. Thus,

$$\Pr\{\mathcal{O}_{i,\text{FD}}(R)\} = 1 - \exp\left(\frac{u(1-2^{\frac{n_{\text{des}}R}{u}})}{n_{\text{des}}}\frac{1}{\gamma}\right), \tag{2.116}$$

and

$$\begin{aligned} R_{\text{FD}}(\varepsilon) &= \sup\left\{R : \Pr\left\{\mathcal{O}_{i,\text{FD}}(R)\right\} < \varepsilon\right\} \\ &= \sup\left\{R : \exp\left(\frac{u(1-2^{\frac{n_{\text{des}}R}{u}})}{n_{\text{des}}}\frac{1}{\gamma}\right) > \overline{\varepsilon}\right\} \\ &= \frac{u}{n_{\text{des}}}\log\left(1 - \frac{n_{\text{des}}\gamma\ln\overline{\varepsilon}}{u}\right). \end{aligned} \tag{2.117}$$

Computation of $R_{\text{FBS}}(\varepsilon)$

Using (2.83), the following Proposition yields $R_{\text{FBS}}(\varepsilon)$.

Proposition 4.

$$R_{\text{FBS}}(\varepsilon) = \sup\left\{R : e^{\frac{u}{\gamma}\left(1-2^{\frac{R}{u}}\right)}\sum_{n=1}^{n_{\max}} q_n 2^{-\frac{(n-1)R}{u}} > \overline{\varepsilon}\right\}. \tag{2.118}$$

Proof. The proof follows from similar steps along the lines of the proof for Proposition 2. $\square$

Comparison of $R_{\text{FH}}(\varepsilon)$, $R_{\text{FD}}(\varepsilon)$, and $R_{\text{FBS}}(\varepsilon)$ in The Asymptotically Small ε Regime

Noting that $\ln \overline{\varepsilon} \approx -\varepsilon$, we get $\log\left(1 - \frac{n_{\text{des}}\gamma \ln \overline{\varepsilon}}{u}\right) \approx \frac{n_{\text{des}}\gamma\varepsilon}{u \ln 2}$. Therefore,

$$R_{\text{FD}}(\varepsilon) \approx \frac{\gamma\varepsilon}{\ln 2}, \tag{2.119}$$

which is the maximum achievable ε-outage capacity in the underlying network for small values of ε. As for FBS, using (2.118) and noting that $\frac{R}{u} \ll 1$, we get

$$R_{\text{FBS}}(\varepsilon) \approx \frac{\varepsilon\gamma}{\left(1 + \frac{\mathrm{E}\{\boldsymbol{N}\}-1}{u}\gamma\right)\ln 2}. \tag{2.120}$$

Comparing (2.101) and (2.120) reveals that $R^{(3)}_{\text{FH}}(\varepsilon)$ is larger than $R_{\text{FBS}}(\varepsilon)$ in the low ε regime as far as (2.102) is satisfied. Furthermore, in the case that $\mathrm{E}\{\boldsymbol{N}\} \ll \frac{u}{\gamma}$, the above equation implies that the FBS scheme, and consequently the FH scenario, achieves the optimal performance of the FD scheme.

Comparison of $R_{\text{FH}}(\varepsilon)$, $R_{\text{FD}}(\varepsilon)$, and $R_{\text{FBS}}(\varepsilon)$ in The Asymptotically Small γ Regime

Using (2.117) and (2.118), it can be realized that for the asymptotically small γ,

$$R_{\text{FD}}(\varepsilon) = R_{\text{FBS}}(\varepsilon) = -\gamma \log \overline{\varepsilon}, \tag{2.121}$$

which is the same value obtained in the previous section for $R_{\text{FH}}(\varepsilon)$ (by setting $v = u$) and is the maximum achievable outage capacity in the network. Therefore, in this regime, spectrum division and spectrum sharing both achieve the optimal performance. Furthermore, the Gaussian lower bound given in (2.84) implies that the FH scheme is optimal in the low SNR regime, regardless of v. However, if we use the proposed lower bounds (e.g., $R^{(3)}_{\text{FH}}(\varepsilon)$), due to the presence of the term $2^{\mathfrak{H}(v,n)}$ in the expression of such lower bounds, taking $v < u$ does not yield this conclusion.

Comparison of $R_{\text{FH}}(\varepsilon)$, $R_{\text{FD}}(\varepsilon)$, and $R_{\text{FBS}}(\varepsilon)$ in The Asymptotically High γ Regime

From (2.117),

$$\begin{aligned} R_{\text{FD}}(\varepsilon) &\approx \frac{u}{n_{\text{des}}} \log \gamma + \frac{u}{n_{\text{des}}} \log\left(-\frac{n_{\text{des}}}{u} \ln \overline{\varepsilon}\right) \\ &\approx \frac{u}{n_{\text{des}}} \log \gamma, \end{aligned} \tag{2.122}$$

as $\gamma \to \infty$. Also, (2.118) indicates that $R_{\text{FBS}}(\varepsilon)$ saturates as γ tends to infinity and as such, FBS is highly inefficient in this regime. In fact,

$$\lim_{\gamma \to \infty} R_{\text{FBS}}(\varepsilon) = \sup\left\{R : \sum_{n=1}^{n_{\max}} q_n 2^{-\frac{(n-1)R}{u}} > \overline{\varepsilon}\right\}. \tag{2.123}$$

To show the supremacy of FH over FD, we consider a practical scenario with $u = n_{\max}$ sub-bands. Each user can be active with some probability p independently of other users. It is assumed that $n_{\max} \gg 1$ and $p \ll 1$ such that $\lambda \triangleq pn_{\max}$ is a constant. Let us assume that the i^{th} user is active. The number of users, other than the i^{th} user, which are simultaneously active together with the i^{th} user is a Binomial random variable with parameters $(n_{\max} - 1, p)$. This random variable, denoted by $\widetilde{\boldsymbol{N}}$, can be well approximated[11] by a Poisson random variable with parameter $(n_{\max} - 1)p \approx \lambda$.

For some $\hat{n} \geq 1$ where u is divisible by $\hat{n}$, the outage probability in the FH system can be written as

$$\begin{aligned}
\Pr\{\mathcal{O}_{i,\mathrm{FH}}(R)\} &= \Pr\{\mathcal{O}_{i,\mathrm{FH}}(R)|\boldsymbol{N} \leq \hat{n}\}\Pr\{\boldsymbol{N} \leq \hat{n}\} \\
&\quad + \Pr\{\mathcal{O}_{i,\mathrm{FH}}(R)|\boldsymbol{N} > \hat{n}\}\Pr\{\boldsymbol{N} > \hat{n}\} \\
&\leq \Pr\{\mathcal{O}_{i,\mathrm{FH}}(R)|\boldsymbol{N} \leq \hat{n}\} + \Pr\{\boldsymbol{N} > \hat{n}\} \\
&\leq \Pr\{\mathcal{O}_{i,\mathrm{FH}}(R)|\boldsymbol{N} = \hat{n}\} + \Pr\{\boldsymbol{N} > \hat{n}\}.
\end{aligned} \tag{2.124}$$

The last line follows from the fact that for a fixed R the outage probability is an increasing function of the number of active users. Choosing $v = \frac{u}{\hat{n}}$ in (2.111) and selecting the transmission rate as

$$\hat{R} = \frac{u}{\hat{n}} \log\left(1 - \kappa_{\frac{u}{\hat{n}}} \gamma^{\left(1-\frac{1}{\hat{n}}\right)^{\hat{n}-1}} \ln\left(1 - \frac{\varepsilon}{2}\right)\right) \tag{2.125}$$

results in $\Pr\{\mathcal{O}_{i,\mathrm{FH}}(\hat{R})|\boldsymbol{N} = \hat{n}\} \leq \frac{\varepsilon}{2}$ where $\kappa_{\frac{u}{\hat{n}}} = \frac{2^{-\boldsymbol{h}\left(\frac{u}{\hat{n}},\hat{n}\right)}\hat{n}}{u}\left(\frac{\hat{n}(\hat{n}-1)}{u}\right)^{\left(1-\frac{1}{\hat{n}}\right)^{\hat{n}-1}-1}$. Furthermore, we select $\hat{n}$ such that

$$\Pr\{\boldsymbol{N} > \hat{n}\} \leq \frac{\varepsilon}{2}. \tag{2.126}$$

Noting that

$$\begin{aligned}
\Pr\{\boldsymbol{N} > \hat{n}\} &= \Pr\{\widetilde{\boldsymbol{N}} \geq \hat{n}\} \\
&\approx \sum_{n=\hat{n}}^{\infty} \frac{e^{-\lambda}\lambda^n}{n!},
\end{aligned} \tag{2.127}$$

it is shown in appendix F that for small enough ε, selecting $\hat{n} = -\lambda \ln \varepsilon$ guarantees $\Pr\{\boldsymbol{N} > \hat{n}\} \leq \frac{\varepsilon}{2}$.

[11] We note that $\widetilde{\boldsymbol{N}} = \boldsymbol{N} - 1$.

Hence, using (2.125),

$$\begin{aligned}
R_{\mathrm{FH}}(\varepsilon) &\overset{(a)}{\geq} \hat{R} \\
&\geq \frac{u}{\hat{n}} \log\left(-\kappa_{\frac{u}{\hat{n}}} \gamma^{\left(1-\frac{1}{\hat{n}}\right)^{\hat{n}-1}} \ln\left(1-\frac{\varepsilon}{2}\right)\right) \\
&= \frac{u}{\hat{n}}\left(1-\frac{1}{\hat{n}}\right)^{\hat{n}-1} \log\gamma \\
&\quad +\frac{u}{\hat{n}} \log\left(-\kappa_{\frac{u}{\hat{n}}} \ln\left(1-\frac{\varepsilon}{2}\right)\right) \\
&\approx \frac{u}{\hat{n}}\left(1-\frac{1}{\hat{n}}\right)^{\hat{n}-1} \log\gamma \\
&\overset{(b)}{\geq} -\frac{u}{e\lambda\ln\varepsilon} \log\gamma,
\end{aligned} \tag{2.128}$$

as $\gamma \to \infty$. In (2.128), (a) follows by the fact that setting the transmission rate at $\hat{R}$, we get $\Pr\{\mathcal{O}_{i,\mathrm{FH}}(\hat{R})\} \leq \varepsilon$. This can be easily seen by (2.124) for the particular choice of $\hat{n} = -\lambda \ln\varepsilon$. Also, (b) holds since $\left(1-\frac{1}{\hat{n}}\right)^{\hat{n}-1}$ is greater or equal to $\frac{1}{e}$.

By (2.122) and setting $n_{\mathrm{des}} = u$,

$$R_{\mathrm{FD}}(\varepsilon) \approx \log\gamma. \tag{2.129}$$

The above equations imply that as long as $u > -e\lambda\ln\varepsilon$, FH is superior to FD. This condition can be alternatively written as $p < -\frac{1}{e\ln\varepsilon}$.

2.6.5 Numerical Results

We consider a practical scenario where each user operates in two different environments with distinct distributions on the number of active users. To describe the distribution of $\boldsymbol{N}$, let $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$ be two independent Binomial random variables with parameters (n_1, p_1) and (n_2, p_2), respectively. For some $p_0 \in (0,1)$, we set

$$\begin{aligned}
q_n &= p_0 \Pr\{\boldsymbol{B}_1 = n | \boldsymbol{B}_1 \geq 1\} + \overline{p}_0 \Pr\{\boldsymbol{B}_2 = n | \boldsymbol{B}_2 \geq 1\} \\
&= p_0 \frac{\binom{n_1}{n} p_1^n (1-p_1)^{n_1-n}}{1-(1-p_1)^{n_1}} + \overline{p}_0 \frac{\binom{n_2}{n} p_2^n (1-p_2)^{n_2-n}}{1-(1-p_2)^{n_2}}
\end{aligned} \tag{2.130}$$

where $1 \leq n \leq \max\{n_1, n_2\}$. In fact, the number of active users is a Binomial random variable with parameters (n_1, p_1) for a fraction p_0 of the time and a Binomial random variable with parameters (n_2, p_2) for the rest of the time. We have made the assumption that there is always at least one active user in the network.

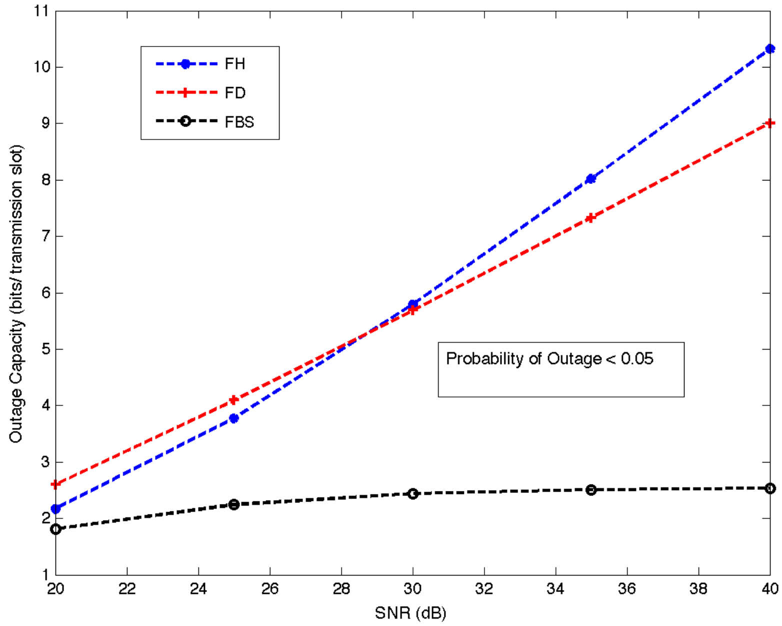

Figure 2.8: Comparison of FH, FD and FBS in terms of outage capacity (bits/transmission slot) for $u = n_{\max} = 10$ and $\{q_n\}_{n=1}^{n_{\max}}$ given in (2.130). The probability of outage is not allowed to exceed 0.05.

Let us set $n_1 = 3$, $n_2 = 10$, $p_1 = 0.1$, $p_2 = 0.2$ and $p_0 = 0.8$. Therefore, $n_{\max} = \max\{n_1, n_2\} = 10$. We assume there are $u = 10$ frequency sub-bands. Figure 2.8 depicts $\max_v R_{\mathrm{FH}}^{(2)}(\varepsilon)$, $R_{\mathrm{FD}}(\varepsilon)$ and $R_{\mathrm{FBS}}(\varepsilon)$ in terms of SNR for $\varepsilon = 0.05$. It is seen that for $\gamma > 29$dB the FH scenario offers a better outage capacity compared to the FD scheme. It is also observed that the FBS scheme offers a poor performance compared to both FD and FH.

Chapter 3

Randomized Signature-Based Transmission

3.1 System Model and The Signaling Scheme

We consider a decentralized wireless communication network of N separate transmitter-receiver pairs. The static and non frequency-selective gain of the channel from the i^{th} transmitter to the j^{th} receiver is shown by $h_{i,j} \in \mathbb{C}$. There is no central controller to assign the resources to the users. Moreover, users do not explicitly cooperate and they are unaware of each other's code-books. Hence, multiuser detection is not possible, i.e., users treat interference as noise. Due to the fact that the network has no fixed infrastructure, resource allocation and rate assignment must be performed locally at every transmitter-receiver pair. To realize this, we propose a distributed signaling scheme motivated by the randomly spread CDMA systems.

Fixing $1 \leq i \leq N$ and for $T \in \mathbb{N}$ and $R_i, R'_i > 0$, a code-book of rate $R_i + R'_i$ for the i^{th} user consists of two groups of codewords, i.e., a collection of $2^{\lfloor TR_i \rfloor}$ signal codewords and a collection of $2^{\lfloor TR'_i \rfloor}$ signature codewords. Each signal codeword is a sequence $(\boldsymbol{x}_i[t])_{t=0}^{T-1}$ of i.i.d. Gaussian signals with zero mean and variance P. Each signature codeword is a sequence $(\vec{\boldsymbol{s}}_i[t])_{t=0}^{T-1}$ of i.i.d. vectors with length K called the signature vectors. The elements of each signature vector are generated independently over an underlying alphabet $\mathcal{A}$ according to a globally known PMF $(q_a)_{a\in\mathcal{A}}$ where q_a is the probability that an element in a signature vector is equal to $a \in \mathcal{A}$. For technical reasons, we let

$$q_0 = \overline{\varepsilon} \tag{3.1}$$

and

$$q_a = \varepsilon p_a \ ; a \in \mathcal{A}\backslash\{0\} \tag{3.2}$$

where $\varepsilon \in (0,1]$ and $p_a \geq 0$ for any $a \in \mathcal{A}\backslash\{0\}$ such that $\sum_{a\in\mathcal{A}\backslash\{0\}} p_a = 1$.

The i^{th} user has $2^{\lfloor TR_i \rfloor + \lfloor TR'_i \rfloor}$ messages where each message is uniquely represented by a pair of integers (message splitting) in the set $\left\{1, 2, \cdots, 2^{\lfloor TR_i \rfloor}\right\} \times \left\{1, 2, \cdots, 2^{\lfloor TR'_i \rfloor}\right\}$. For any message $(\mathrm{W}_i, \mathrm{W}'_i)$, the corresponding signal and signature codewords are selected and the sequence of vectors $(\boldsymbol{x}_i[t]\vec{\boldsymbol{s}}_i[t])_{t=0}^{T-1}$ is transmitted. We assume the channel from each transmitter to each receiver is a slotted channel. Moreover, the slots on different channels perfectly overlap, i.e., they start and end at the same time instants. This implies that all users are slot-synchronized. For each $0 \leq t \leq T-1$, the i^{th} user transmits $\boldsymbol{x}_i[t]\vec{\boldsymbol{s}}_i[t]$ in K consecutive transmission slots, referred to as a signature interval. Denoting the received vector at the receiver side of the i^{th} user during a typical signature interval by $\vec{\boldsymbol{y}}_i$ and dropping the index t for simplicity of notation, we have

$$\vec{\boldsymbol{y}}_i = \alpha h_{i,i}\boldsymbol{x}_i\vec{\boldsymbol{s}}_i + \sum_{j\neq i} \alpha h_{j,i}\boldsymbol{x}_j\vec{\boldsymbol{s}}_j + \vec{\boldsymbol{z}}_i \tag{3.3}$$

where $\vec{\boldsymbol{z}}_i$ is a $\mathcal{CN}(\vec{0}_K, I_K)$ random vector representing the ambient noise. The expression for $\vec{\boldsymbol{y}}_i$ in (3.3) implies that in addition to being slot-synchronized, we require the users to be signature-synchronized, i.e., the signatur intervals of all users coincide. The parameter α is a normalization factor that ensures the average transmitted power per transmission slot of the i^{th} user is γ, i.e.,

$$\alpha^2 = \frac{K}{\mathrm{E}\left\{\|\vec{\boldsymbol{s}}_i\|^2\right\}}\frac{\gamma}{P}. \tag{3.4}$$

Since the elements of $\vec{\boldsymbol{z}}_i$ have unit variance, the parameter γ is considered as the measure of SNR.

The interference vector in a signature interval at the receiver side of the i^{th} user is denoted by $\vec{\boldsymbol{w}}_i$, i.e., $\vec{\boldsymbol{w}}_i = \sum_{j\neq i} \alpha h_{j,i}\boldsymbol{x}_j\vec{\boldsymbol{s}}_j$. One can state $\vec{\boldsymbol{w}}_i$ as

$$\vec{\boldsymbol{w}}_i = \alpha \boldsymbol{S}_i \Xi_i \begin{pmatrix} \boldsymbol{x}_1 & \cdots & \boldsymbol{x}_{i-1} & \boldsymbol{x}_{i+1} & \cdots & \boldsymbol{x}_N \end{pmatrix}^{\mathrm{t}} \tag{3.5}$$

where

$$\boldsymbol{S}_i \triangleq \begin{pmatrix} \vec{\boldsymbol{s}}_1 & \cdots & \vec{\boldsymbol{s}}_{i-1} & \vec{\boldsymbol{s}}_{i+1} & \cdots & \vec{\boldsymbol{s}}_N \end{pmatrix} \tag{3.6}$$

and

$$\Xi_i \triangleq \mathrm{diag}(h_{1,i}, \cdots, h_{i-1,i}, h_{i+1,i}, \cdots, h_{N,i}). \tag{3.7}$$

Using joint typicality decoding at the receiver side of the i^{th} user, any data rate $R_i + R'_i < \mathscr{R}_i(\vec{h}_i)$ is achievable in the conventional sense where[1]

$$\mathscr{R}_i(\vec{h}_i) \triangleq \frac{\mathrm{I}(\boldsymbol{x}_i, \vec{\boldsymbol{s}}_i; \vec{\boldsymbol{y}}_i)}{K} \tag{3.8}$$

[1]Note that, in general, the pair (R_i, R'_i) must lie in a multiple access channel rate region defined by $R_i \leq \frac{\mathrm{I}(\boldsymbol{x}_i;\vec{\boldsymbol{y}}_i|\vec{\boldsymbol{s}}_i)}{K}$, $R'_i \leq \frac{\mathrm{I}(\vec{\boldsymbol{s}}_i;\vec{\boldsymbol{y}}_i|\boldsymbol{x}_i)}{K}$ and $R_i + R'_i \leq \frac{\mathrm{I}(\boldsymbol{x}_i,\vec{\boldsymbol{s}}_i;\vec{\boldsymbol{y}}_i)}{K}$.

and $\vec{h}_i \triangleq \begin{pmatrix} h_{1,i} & \cdots & h_{N,i} \end{pmatrix}^{\mathrm{t}}$ is a vector that contains the channel gains related to the i^{th} user.

We conclude this section by mentioning that the i^{th} receiver is assumed to be able to estimate (without error) the channel gains $\{h_{j,i}\}_{j=1}^{N}$ and the number of active users N and transmit these data to its corresponding transmitter through a feedback link. In [78], using a combined technique based on the method of moments and maximum likelihood estimation [83, 84], we demonstrate how each user can perform this task in practice.

3.2 A Lower Bound on $\mathrm{I}(\boldsymbol{x}_i, \vec{\boldsymbol{s}}_i; \vec{\boldsymbol{y}}_i)$

One can write $\mathrm{I}(\boldsymbol{x}_i, \vec{\boldsymbol{s}}_i; \vec{\boldsymbol{y}}_i)$ as

$$\mathrm{I}(\boldsymbol{x}_i, \vec{\boldsymbol{s}}_i; \vec{\boldsymbol{y}}_i) = \mathrm{I}(\vec{\boldsymbol{s}}_i; \vec{\boldsymbol{y}}_i) + \mathrm{I}(\boldsymbol{x}_i; \vec{\boldsymbol{y}}_i | \vec{\boldsymbol{s}}_i). \tag{3.9}$$

The i^{th} user can successfully decode $\vec{\boldsymbol{s}}_i$ as far as $R'_i < \frac{\mathrm{I}(\vec{\boldsymbol{s}}_i; \vec{\boldsymbol{y}}_i)}{K}$. Thereafter, using its knowledge of $\vec{\boldsymbol{s}}_i$, this user decodes $\boldsymbol{x}_i$ successfully if $R_i < \frac{\mathrm{I}(\boldsymbol{x}_i; \vec{\boldsymbol{y}}_i | \vec{\boldsymbol{s}}_i)}{K}$.

Due to the fact that the signature codeword of any user is not known to the other users, the noise plus interference at any receiver has a mixed Gaussian PDF. As a result, the terms $\mathrm{I}(\vec{\boldsymbol{s}}_i; \vec{\boldsymbol{y}}_i)$ and $\mathrm{I}(\boldsymbol{x}_i; \vec{\boldsymbol{y}}_i | \vec{\boldsymbol{s}}_i)$ do not have closed expressions. The term $\mathrm{I}(\vec{\boldsymbol{s}}_i; \vec{\boldsymbol{y}}_i)$ is bounded from above by $\mathrm{H}(\vec{\boldsymbol{s}}_i)$ which is not a function of SNR. Therefore, $\mathrm{I}(\boldsymbol{x}_i; \vec{\boldsymbol{y}}_i | \vec{\boldsymbol{s}}_i)$ and $\mathrm{I}(\boldsymbol{x}_i, \vec{\boldsymbol{s}}_i; \vec{\boldsymbol{y}}_i)$ have similar SNR scalings, i.e.,

$$\mathrm{I}(\boldsymbol{x}_i, \vec{\boldsymbol{s}}_i; \vec{\boldsymbol{y}}_i) \overset{\gamma}{\sim} \mathrm{I}(\boldsymbol{x}_i; \vec{\boldsymbol{y}}_i | \vec{\boldsymbol{s}}_i). \tag{3.10}$$

Motivated by (3.10), we ignore the term $\mathrm{I}(\vec{\boldsymbol{s}}_i; \vec{\boldsymbol{y}}_i)$ in this work and focus our attention on $\mathrm{I}(\boldsymbol{x}_i; \vec{\boldsymbol{y}}_i | \vec{\boldsymbol{s}}_i)$. We are able to develop a lower bound on $\mathrm{I}(\boldsymbol{x}_i; \vec{\boldsymbol{y}}_i | \vec{\boldsymbol{s}}_i)$ that is tight in the sense that:

1- It achieves the same SNR scaling as that of $\mathrm{I}(\boldsymbol{x}_i; \vec{\boldsymbol{y}}_i | \vec{\boldsymbol{s}}_i)$.

2- Its behavior in terms of the design parameters mimics that of $\mathrm{I}(\boldsymbol{x}_i; \vec{\boldsymbol{y}}_i | \vec{\boldsymbol{s}}_i)$.

As such, system design is performed only based on the term $\mathrm{I}(\boldsymbol{x}_i; \vec{\boldsymbol{y}}_i | \vec{\boldsymbol{s}}_i)$.

To develop a lower bound on $\mathrm{I}(\boldsymbol{x}_i; \vec{\boldsymbol{y}}_i | \vec{\boldsymbol{s}}_i)$, our major tools are linear processing of the channel output based on writing the SVD for the signature $\vec{\boldsymbol{s}}_i$, a conditional entropy power inequality and a key upper bound on the differential entropy of a mixed Gaussian random vector.

We have

$$\mathrm{I}(\boldsymbol{x}_i; \vec{\boldsymbol{y}}_i | \vec{\boldsymbol{s}}_i) = \sum_{\vec{s} \in \mathrm{supp}(\vec{\boldsymbol{s}}_i) \backslash \{\vec{0}_K\}} \Pr\{\vec{\boldsymbol{s}}_i = \vec{s}\} \mathrm{I}(\boldsymbol{x}_i; \vec{\boldsymbol{y}}_i | \vec{\boldsymbol{s}}_i = \vec{s}) \tag{3.11}$$

In the following, we find a lower bound on $\mathrm{I}(\boldsymbol{x}_i; \vec{\boldsymbol{y}}_i | \vec{\boldsymbol{s}}_i = \vec{s})$ for any $\vec{s} \in \mathrm{supp}(\vec{\boldsymbol{s}}_i) \backslash \{\vec{0}_K\}$.

Step 1- The matrix $\vec{s}\vec{s}^{\dagger}$ has two eigenvalues, namely, 0 and $\|\vec{s}\|^2$. The eigenvector corresponding to $\|\vec{s}\|^2$ is $\vec{s}$ and the eigenvectors corresponding to 0 are $K-1$ orthonormal vectors denoted by

$\vec{g}_1, \cdots, \vec{g}_{K-1}$ which together with $\frac{\vec{s}}{\|\vec{s}\|}$ make an orthonormal basis for $\mathbb{R}^K$. Let us define

$$G_i(\vec{s}) \triangleq \begin{pmatrix} \vec{g}_{i,1} & \cdots & \vec{g}_{i,K-1} \end{pmatrix}, \tag{3.12}$$

$$U_i(\vec{s}) \triangleq \begin{pmatrix} \frac{\vec{s}}{\|\vec{s}\|} & G_i(\vec{s}) \end{pmatrix} \tag{3.13}$$

and

$$\vec{d}(\vec{s}) \triangleq \begin{pmatrix} \|\vec{s}\| & \vec{0}^{\mathrm{t}}_{K-1} \end{pmatrix}^{\mathrm{t}}. \tag{3.14}$$

Writing the SVD for $\vec{s}$,

$$\vec{s} = U_i(\vec{s})\vec{d}(\vec{s}). \tag{3.15}$$

The i^{th} receiver constructs the vector $U_i^{\dagger}(\vec{s})\vec{\boldsymbol{y}}_i(\vec{s})$ where $\vec{\boldsymbol{y}}_i(\vec{s})$ has the same expression as $\vec{\boldsymbol{y}}_i$ in which $\vec{\boldsymbol{s}}_i$ is replaced by $\vec{s}$, i.e.,

$$\vec{\boldsymbol{y}}_i(\vec{s}) \triangleq \alpha h_{i,i}\boldsymbol{x}_i\vec{s} + \sum_{j\neq i}\alpha h_{j,i}\boldsymbol{x}_j\vec{\boldsymbol{s}}_j + \vec{\boldsymbol{z}}_i. \tag{3.16}$$

We have

$$U_i^{\dagger}(\vec{s})\vec{\boldsymbol{y}}_i(\vec{s}) = \alpha h_{i,i}\boldsymbol{x}_i\vec{d}(\vec{s}) + U_i^{\dagger}(\vec{s})\left(\vec{\boldsymbol{w}}_i + \vec{\boldsymbol{z}}_i\right). \tag{3.17}$$

Define

$$\boldsymbol{\varphi}_i \triangleq \left[U_i^{\dagger}(\vec{s})\left(\vec{\boldsymbol{w}}_i + \vec{\boldsymbol{z}}_i\right)\right]_1 = \frac{\alpha\vec{s}^{\dagger}\left(\vec{\boldsymbol{w}}_i + \vec{\boldsymbol{z}}_i\right)}{\|\vec{s}\|}, \tag{3.18}$$

$$\boldsymbol{\omega}_i \triangleq \left[U_i^{\dagger}(\vec{s})\vec{\boldsymbol{y}}_i(\vec{s})\right]_1 = \alpha h_{i,i}\|\vec{s}\|\boldsymbol{x}_i + \boldsymbol{\varphi}_i \tag{3.19}$$

and

$$\vec{\boldsymbol{\vartheta}}_i \triangleq \left[U_i^{\dagger}(\vec{s})\vec{\boldsymbol{y}}_i(\vec{s})\right]_2^K = G_i^{\dagger}(\vec{s})\left(\vec{\boldsymbol{w}}_i + \vec{\boldsymbol{z}}_i\right). \tag{3.20}$$

We have the following thread of equalities,

$$\begin{aligned} \mathrm{I}(\boldsymbol{x}_i;\vec{\boldsymbol{y}}_i|\vec{\boldsymbol{s}}_i = \vec{s}) &\overset{(a)}{=} \mathrm{I}\left(\boldsymbol{x}_i;U_i^{\dagger}(\vec{s})\vec{\boldsymbol{y}}_i(\vec{s})\right) \\ &= \mathrm{I}(\boldsymbol{x}_i;\boldsymbol{\omega}_i,\vec{\boldsymbol{\vartheta}}_i) \\ &= \mathrm{I}(\boldsymbol{x}_i;\vec{\boldsymbol{\vartheta}}_i) + \mathrm{I}(\boldsymbol{x}_i;\boldsymbol{\omega}_i|\vec{\boldsymbol{\vartheta}}_i) \\ &\overset{(b)}{=} \mathrm{I}(\boldsymbol{x}_i;\boldsymbol{\omega}_i|\vec{\boldsymbol{\vartheta}}_i) \end{aligned} \tag{3.21}$$

where (a) holds since $U_i(\vec{s})$ is a unitary matrix and (b) is by the fact that $\boldsymbol{x}_i$ and $\vec{\boldsymbol{\vartheta}}_i$ are independent, i.e., $\mathrm{I}(\boldsymbol{x}_i;\vec{\boldsymbol{\vartheta}}_i) = 0$.

Step 2- We need the following Lemma which is a conditional entropy power inequality [70] in the complex setup.

Lemma 5. *Let* z_1 *and* z_2 *be complex random variables and* z_3 *be any random quantity (scalar or vector) with densities. Also, assume that the conditional densities* $p_{\mathsf{z}_1|\mathsf{z}_3}(.|.)$ *and* $p_{\mathsf{z}_2|\mathsf{z}_3}(.|.)$ *exist. If* z_1 *and* z_2 *are conditionally independent given* z_3*, then*

$$2^{\mathrm{h}(\mathsf{z}_1+\mathsf{z}_2|\mathsf{z}_3)} \geq 2^{\mathrm{h}(\mathsf{z}_1|\mathsf{z}_3)} + 2^{\mathrm{h}(\mathsf{z}_2|\mathsf{z}_3)}. \tag{3.22}$$

Proof. A straightforward application of the non-conditional entropy power inequality leads to (3.22). The proof is omitted here. □

We have

$$\begin{aligned}
\mathrm{I}(\boldsymbol{x}_i;\boldsymbol{\omega}_i|\vec{\boldsymbol{\vartheta}}_i) &= \mathrm{h}(\boldsymbol{\omega}_i|\vec{\boldsymbol{\vartheta}}_i) - \mathrm{h}(\boldsymbol{\omega}_i|\boldsymbol{x}_i,\vec{\boldsymbol{\vartheta}}_i) \\
&= \mathrm{h}(\boldsymbol{\omega}_i|\vec{\boldsymbol{\vartheta}}_i) - \mathrm{h}\left(\alpha h_{i,i}\|\vec{s}\|\boldsymbol{x}_i + \boldsymbol{\varphi}_i \middle| \boldsymbol{x}_i,\vec{\boldsymbol{\vartheta}}_i\right) \\
&\overset{(a)}{=} \mathrm{h}(\boldsymbol{\omega}_i|\vec{\boldsymbol{\vartheta}}_i) - \mathrm{h}(\boldsymbol{\varphi}_i|\vec{\boldsymbol{\vartheta}}_i)
\end{aligned} \tag{3.23}$$

where (a) is by independence of $\boldsymbol{x}_i$ and $(\boldsymbol{\varphi}_i,\vec{\boldsymbol{\vartheta}}_i)$. On the other hand, we know that $\boldsymbol{\omega}_i = \alpha h_{i,i}\|\vec{s}\|\boldsymbol{x}_i + \boldsymbol{\varphi}_i$. Defining $\mathsf{z}_1 \triangleq \alpha h_{i,i}\|\vec{s}\|\boldsymbol{x}_i$ and $\mathsf{z}_2 \triangleq \boldsymbol{\varphi}_i$, it is clear that z_1 and z_2 are conditionally independent given $\mathsf{z}_3 \triangleq \vec{\boldsymbol{\vartheta}}_i$. As the conditional densities $p_{\mathsf{z}_1|\mathsf{z}_3}(.|.)$ and $p_{\mathsf{z}_2|\mathsf{z}_3}(.|.)$ exist, by Lemma 5,

$$\begin{aligned}
2^{\mathrm{h}\left(\boldsymbol{\omega}_i|\vec{\boldsymbol{\vartheta}}_i\right)} &\geq 2^{\mathrm{h}\left(\alpha h_{i,i}\|\vec{s}\|\boldsymbol{x}_i|\vec{\boldsymbol{\vartheta}}_i\right)} + 2^{\mathrm{h}\left(\boldsymbol{\varphi}_i|\vec{\boldsymbol{\vartheta}}_i\right)} \\
&\overset{(a)}{=} 2^{\mathrm{h}\left(\alpha h_{i,i}\|\vec{s}\|\boldsymbol{x}_i\right)} + 2^{\mathrm{h}\left(\boldsymbol{\varphi}_i|\vec{\boldsymbol{\vartheta}}_i\right)}
\end{aligned} \tag{3.24}$$

where (a) is by the fact that the $\boldsymbol{x}_i$ is independent of $\vec{\boldsymbol{\vartheta}}_i$. Dividing both sides of (3.24) by $2^{\mathrm{h}(\boldsymbol{\varphi}_i|\vec{\boldsymbol{\vartheta}}_i)}$,

$$\mathrm{h}\left(\boldsymbol{\omega}_i|\vec{\boldsymbol{\vartheta}}_i\right) - \mathrm{h}\left(\boldsymbol{\varphi}_i|\vec{\boldsymbol{\vartheta}}_i\right) \geq \log\left(2^{\left(\mathrm{h}\left(\alpha h_{i,i}\|\vec{s}\|\boldsymbol{x}_i\right) - \mathrm{h}\left(\boldsymbol{\varphi}_i|\vec{\boldsymbol{\vartheta}}_i\right)\right)} + 1\right). \tag{3.25}$$

By (3.21), (3.23) and (3.25),

$$\mathrm{I}(\boldsymbol{x}_i;\vec{\boldsymbol{y}}_i|\vec{\boldsymbol{s}}_i = \vec{s}) \geq \log\left(2^{\left(\mathrm{h}\left(\alpha h_{i,i}\|\vec{s}\|\boldsymbol{x}_i\right) - \mathrm{h}\left(\boldsymbol{\varphi}_i|\vec{\boldsymbol{\vartheta}}_i\right)\right)} + 1\right). \tag{3.26}$$

Step 3- We need the following Lemma to proceed.

Lemma 6. *Let* $\vec{\mathsf{z}}$ *be a* $t\times 1$ *complex and circularly symmetric mixed Gaussian random vector with* PDF

$$p_{\vec{\mathsf{z}}}(\vec{\mathsf{z}}) = \sum_{l=1}^{L} \frac{a_l}{\pi^t \det \Omega_l} \exp\left(-\vec{\mathsf{z}}^\dagger \Omega_l^{-1} \vec{\mathsf{z}}\right) \tag{3.27}$$

where $a_l \geq 0$ *for* $1 \leq l \leq L$ *and* $\sum_{l=1}^{L} a_l = 1$*. Then,*

$$\sum_{l=1}^{L} a_l \log\left((\pi e)^t \det \Omega_l\right) \leq \mathrm{h}(\vec{\mathsf{z}}) \leq \sum_{l=1}^{L} a_l \log\left((\pi e)^t \det \Omega_l\right) + \mathrm{H}((a_l)_{l=1}^{L}) \tag{3.28}$$

Proof. Let us define the random matrix $\mathbf{\Omega} \in \{\Omega_l : 1 \leq l \leq L\}$ such that $\Pr\{\mathbf{\Omega} = \Omega_l\} = a_l$ and let $\vec{\mathbf{x}}$ be a $\mathcal{CN}(\vec{0}_t, I_t)$ random vector. Then, one can easily see that $\vec{\mathbf{z}} = \sqrt{\mathbf{\Omega}}\vec{\mathbf{x}}$ in which $\sqrt{\mathbf{\Omega}}$ is the conventional square root of a positive semi-definite matrix. Using the inequalities

$$\begin{aligned} \mathrm{h}(\vec{\mathbf{z}}|\mathbf{\Omega}) \leq \mathrm{h}(\vec{\mathbf{z}}) &\leq \mathrm{h}(\vec{\mathbf{z}}, \mathbf{\Omega}) \\ &= \mathrm{h}(\vec{\mathbf{z}}|\mathbf{\Omega}) + \mathrm{H}(\mathbf{\Omega}) \\ &= \mathrm{h}(\vec{\mathbf{z}}|\mathbf{\Omega}) + \mathrm{H}((a_l)_{l=1}^{L}) \end{aligned} \tag{3.29}$$

and noting that $\mathrm{h}(\vec{\mathbf{z}}|\mathbf{\Omega}) = \sum_{l=1}^{L} a_l \log\left((\pi e)^t \det \Omega_l\right)$, the result is immediate. □

By (3.5), the interference vector $\vec{\boldsymbol{w}}_i$ has a mixed Gaussian distribution where the covariance matrices of its Gaussian components correspond to different realizations of the matrix $\alpha^2 P \boldsymbol{S}_i \Xi_i \Xi_i^{\dagger} \boldsymbol{S}_i^{\dagger}$. This together with Lemma 6 yields

$$\mathrm{h}(\vec{\boldsymbol{w}}_i + \vec{\boldsymbol{z}}_i) \leq \mathrm{h}(\vec{\boldsymbol{w}}_i + \vec{\boldsymbol{z}}_i|\boldsymbol{S}_i) + \mathrm{H}\left(\boldsymbol{S}_i \Xi_i \Xi_i^{\dagger} \boldsymbol{S}_i^{\dagger}\right) \tag{3.30}$$

where we have used the fact that $\mathrm{H}\left(\alpha^2 P \boldsymbol{S}_i \Xi_i \Xi_i^{\dagger} \boldsymbol{S}_i^{\dagger}\right) = \mathrm{H}\left(\boldsymbol{S}_i \Xi_i \Xi_i^{\dagger} \boldsymbol{S}_i^{\dagger}\right)$. One has

$$\begin{aligned} \mathrm{h}(\boldsymbol{\varphi}_i|\vec{\boldsymbol{\vartheta}}_i) &= \mathrm{h}(\boldsymbol{\varphi}_i, \vec{\boldsymbol{\vartheta}}_i) - \mathrm{h}(\vec{\boldsymbol{\vartheta}}_i) \\ &= \mathrm{h}\left(U_i^{\dagger}(\vec{s})\left(\vec{\boldsymbol{w}}_i + \vec{\boldsymbol{z}}_i\right)\right) - \mathrm{h}\left(G_i^{\dagger}(\vec{s})\left(\vec{\boldsymbol{w}}_i + \vec{\boldsymbol{z}}_i\right)\right) \\ &\stackrel{(a)}{=} \mathrm{h}\left(\vec{\boldsymbol{w}}_i + \vec{\boldsymbol{z}}_i\right) - \mathrm{h}\left(G_i^{\dagger}(\vec{s})\left(\vec{\boldsymbol{w}}_i + \vec{\boldsymbol{z}}_i\right)\right) \\ &\stackrel{(b)}{\leq} \mathrm{h}\left(\vec{\boldsymbol{w}}_i + \vec{\boldsymbol{z}}_i|\boldsymbol{S}_i\right) + \mathrm{H}\left(\boldsymbol{S}_i \Xi_i \Xi_i^{\dagger} \boldsymbol{S}_i^{\dagger}\right) \\ &\quad - \mathrm{h}\left(G_i^{\dagger}(\vec{s})\left(\vec{\boldsymbol{w}}_i + \vec{\boldsymbol{z}}_i\right)\right) \\ &\stackrel{(c)}{\leq} \mathrm{h}\left(\vec{\boldsymbol{w}}_i + \vec{\boldsymbol{z}}_i|\boldsymbol{S}_i\right) + \mathrm{H}\left(\boldsymbol{S}_i \Xi_i \Xi_i^{\dagger} \boldsymbol{S}_i^{\dagger}\right) \\ &\quad - \mathrm{h}\left(G_i^{\dagger}(\vec{s})\left(\vec{\boldsymbol{w}}_i + \vec{\boldsymbol{z}}_i\right)|\boldsymbol{S}_i\right) \end{aligned} \tag{3.31}$$

where (a) follows by noting that the matrix $U_i(\vec{s})$ is unitary, i.e., $\log|\det(U_i(\vec{s}))| = 0$, (b) follows from (3.30) and (c) is a direct consequence of Lemma 6. On the other hand, for any realization of $\boldsymbol{S}_i$, the vector $\vec{\boldsymbol{w}}_i + \vec{\boldsymbol{z}}_i$ turns into a complex Gaussian vector. Hence,

$$\mathrm{h}\left(\vec{\boldsymbol{w}}_i + \vec{\boldsymbol{z}}_i|\boldsymbol{S}_i\right) = K\log(\pi e) + \sum_{S \in \mathrm{supp}(\boldsymbol{S}_i)} \Pr\{\boldsymbol{S}_i = S\} \log\det\left(I_K + \alpha^2 P S \Xi_i \Xi_i^{\dagger} S^{\dagger}\right). \tag{3.32}$$

By the same token,

$$\begin{aligned} &\mathrm{h}\left(G_i^{\dagger}(\vec{s})\left(\vec{\boldsymbol{w}}_i + \vec{\boldsymbol{z}}_i\right)|\boldsymbol{S}_i\right) = (K-1)\log(\pi e) \\ &+ \sum_{S \in \mathrm{supp}(\boldsymbol{S}_i)} \Pr\{\boldsymbol{S}_i = S\} \log\det\left(I_{K-1} + \alpha^2 P G_i^{\dagger}(\vec{s}) S \Xi_i \Xi_i^{\dagger} S^{\dagger} G_i^{\dagger}(\vec{s})\right). \end{aligned} \tag{3.33}$$

Using (3.32) and (3.33) in (3.31),

$$\mathrm{h}(\boldsymbol{\varphi}_i|\vec{\boldsymbol{\vartheta}}_i) \leq \log(\pi e) + \mathrm{H}\left(\boldsymbol{S}_i\Xi_i\Xi_i^\dagger\boldsymbol{S}_i^\dagger\right)$$
$$+ \sum_{S\in\mathrm{supp}(\boldsymbol{S}_i)} \Pr\{\boldsymbol{S}_i = S\} \log\det\left(I_K + \alpha^2 PS\Xi_i\Xi_i^\dagger S^\dagger\right)$$
$$- \sum_{S\in\mathrm{supp}(\boldsymbol{S}_i)} \Pr\{\boldsymbol{S}_i = S\} \log\det\left(I_{K-1} + \alpha^2 PG_i^\dagger(\vec{s})S\Xi_i\Xi_i^\dagger S^\dagger G_i^\dagger(\vec{s})\right). \quad (3.34)$$

Moreover, $\mathrm{h}(\alpha h_{i,i}\|\vec{s}\|\boldsymbol{x}_i) = \log\left(\pi e\alpha^2|h_{i,i}|^2\|\vec{s}\|^2P\right)$. Hence, $\mathrm{h}(\alpha h_{i,i}\|\vec{s}\|\boldsymbol{x}_i) - \mathrm{h}\left(\boldsymbol{\varphi}_i|\vec{\boldsymbol{\vartheta}}_i\right)$ appearing in (3.26) can be bounded from below as

$$\mathrm{h}(\alpha h_{i,i}\|\vec{s}\|\boldsymbol{x}_i) - \mathrm{h}\left(\boldsymbol{\varphi}_i|\vec{\boldsymbol{\vartheta}}_i\right) \geq \log\left(\alpha^2|h_{i,i}|^2\|\vec{s}\|^2P\right) - \mathrm{H}\left(\boldsymbol{S}_i\Xi_i\Xi_i^\dagger\boldsymbol{S}_i^\dagger\right)$$
$$- \sum_{S\in\mathrm{supp}(\boldsymbol{S}_i)} \Pr\{\boldsymbol{S}_i = S\} \log\det\left(I_K + \alpha^2 PS\Xi_i\Xi_i^\dagger S^\dagger\right)$$
$$+ \sum_{S\in\mathrm{supp}(\boldsymbol{S}_i)} \Pr\{\boldsymbol{S}_i = S\} \log\det\left(I_{K-1} + \alpha^2 PG_i^\dagger(\vec{s})S\Xi_i\Xi_i^\dagger S^\dagger G_i^\dagger(\vec{s})\right). \quad (3.35)$$

Substituting (3.35) in (3.26),

$$\mathrm{I}(\boldsymbol{x}_i;\vec{\boldsymbol{y}}_i|\vec{\boldsymbol{s}}_i = \vec{s}) \geq \log\left(2^{-\mathrm{H}\left(\boldsymbol{S}_i\Xi_i\Xi_i^\dagger\boldsymbol{S}_i^\dagger\right)}\varrho_i(P;\vec{s},\Xi_i) + 1\right) \quad (3.36)$$

where

$$\varrho_i(P;\vec{s},\vec{h}_i) \triangleq \alpha^2|h_{i,i}|^2\|\vec{s}\|^2P \prod_{S\in\mathrm{supp}(\boldsymbol{S}_i)} \left(\frac{\det\left(I_{K-1} + \alpha^2 PG_i^\dagger(\vec{s})S\Xi_i\Xi_i^\dagger S^\dagger G_i^\dagger(\vec{s})\right)}{\det\left(I_K + \alpha^2 PS\Xi_i\Xi_i^\dagger S^\dagger\right)}\right)^{\Pr\{\boldsymbol{S}_i=S\}}. \quad (3.37)$$

Finally, we get a lower bound on $\mathscr{R}_i(\vec{h}_i)$ given by

$$\mathscr{R}_i^{(\mathrm{lb})}(\vec{h}_i) \triangleq \frac{1}{K} \sum_{\vec{s}\in\mathrm{supp}(\vec{\boldsymbol{s}}_i)\backslash\{\vec{0}_K\}} \Pr\{\vec{\boldsymbol{s}}_i = \vec{s}\} \log\left(2^{-\mathrm{H}\left(\boldsymbol{S}_i\Xi_i\Xi_i^\dagger\boldsymbol{S}_i^\dagger\right)}\varrho_i(P;\vec{s},\vec{h}_i) + 1\right). \quad (3.38)$$

3.3 Analysis of Sum Multiplexing Gain

An important observation is the SNR scaling of $\mathscr{R}_i^{(\mathrm{lb})}(\vec{h}_i)$ that is presented in the following Proposition:

Proposition 5. *Regulating its transmission rate at $\mathscr{R}_i^{(\mathrm{lb})}(\vec{h}_i)$, the i^{th} user achieves an* SNR *scaling of*

$$\lim_{\gamma\to\infty} \frac{\mathscr{R}_i^{(\mathrm{lb})}(\vec{h}_i)}{\log\gamma} = \frac{\Pr\{\vec{\boldsymbol{s}}_i \notin \mathrm{csp}(\boldsymbol{S}_i)\}}{K}. \quad (3.39)$$

Proof. See appendix G. $\square$

Finally, the following Corollary proves that $\mathscr{R}_i^{(\text{lb})}(\vec{h}_i)$ achieves the SNR scaling of $\mathscr{R}_i(\vec{h}_i)$.

Corollary 2. $\mathscr{R}_i^{(\text{lb})}(\vec{h}_i) \overset{\mathcal{L}}{\sim} \mathscr{R}_i(\vec{h}_i)$.

Proof. See Appendix H. $\square$

Since all users utilize the same PMF to construct their signature vectors, the achievable SMG can be written as

$$\text{SMG}_N = \frac{N \Pr\left\{\vec{\boldsymbol{s}}_1 \notin \text{csp}\left([\vec{\boldsymbol{s}}_2|\vec{\boldsymbol{s}}_3|\cdots|\vec{\boldsymbol{s}}_{N-1}|\vec{\boldsymbol{s}}_N]\right)\right\}}{K}. \tag{3.40}$$

Computing $\Pr\left\{\vec{\boldsymbol{s}}_1 \notin \text{csp}\left([\vec{\boldsymbol{s}}_2|\vec{\boldsymbol{s}}_3|\cdots|\vec{\boldsymbol{s}}_{N-1}|\vec{\boldsymbol{s}}_N]\right)\right\}$ is a tedious task for $N \geq 3$. Using the so-called Marčenko-Pastur law[2] in random matrix theory [81], it is easily seen that $\vec{\boldsymbol{s}}_1 \notin \text{csp}\left([\vec{\boldsymbol{s}}_2|\vec{\boldsymbol{s}}_3|\cdots|\vec{\boldsymbol{s}}_{N-1}|\vec{\boldsymbol{s}}_N]\right)$ holds almost surely in the asymptotic case where $K = N \to \infty$. As such, a sum multiplexing gain of 1 is achievable if the number of active users is large. In this work, our focus is on a finite number of users. To get a rough idea about the behavior of SMG_N, we present some examples.

Example 1- Effects of The Underlying Alphabet $\mathcal{A}$ on SMG_N*.*

In this example, we compare four different types of signature vectors by computing the achievable SMG via Monte-Carlo simulations. The following cases are considered:

1- $\mathcal{A} = \{-1, 1\}$, $q_{-1} = q_1 = \frac{1}{2}$, $K = N$.

2- $\mathcal{A} = \{0, 1\}$, $q_0 = q_1 = \frac{1}{2}$, $K = N$.

3- $\mathcal{A} = \{-1, 0, 1\}$, $q_{-1} = q_0 = q_1 = \frac{1}{3}$, $K = N$

4- $\mathcal{A} = \{0, 1\}$, $q_0 = q_1 = \frac{1}{2}$, $K = 1$.

The elements of the signature vectors are generated uniformly over the underlying alphabet in all the schemes. This is not the best one can do as we will see in example 2. The achieved SMG for each scheme is sketched in fig. 3.1. We are able to make the following observations:

• SMG_N is an increasing function of N as far as $K = N$. Moreover, it approaches 1 as N increases. This is in agreement with the Marčenko-Pastur law.

• Increasing the size of the underlying alphabet $\mathcal{A}$ leads to a larger SMG.

• The alphabet $\{0, 1\}$ yields a slightly better SMG compared to the alphabet of the same size $\{-1, 1\}$.

Let us investigate the case $\mathcal{A} = \{0, 1\}$ and $K = 1$ in more detail. For each $1 \leq i \leq N$,

[2] A central result in random matrix theory states that when the entries of a random matrix $\boldsymbol{A}$ of size $n \times n$ are i.i.d., then $\frac{\text{rank}(\boldsymbol{A})}{n} \to 1$ as $n \to \infty$. A more general statement of this result is the so-called Marčenko-Pastur law.

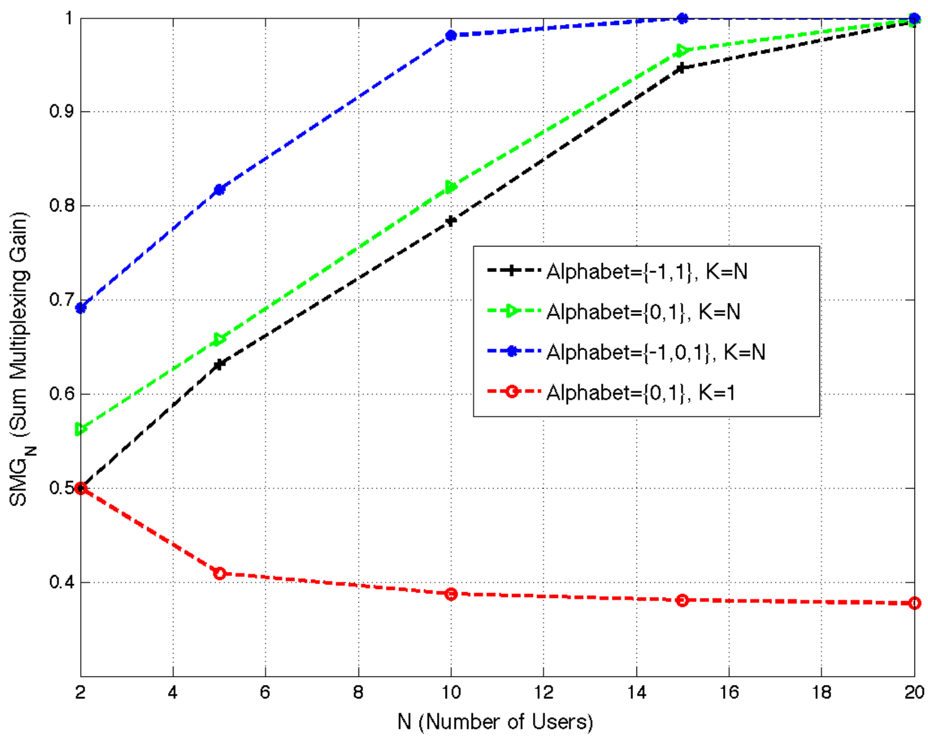

Figure 3.1: Effects of the choice of the underlying alphabet $\mathcal{A}$ on SMG.

$\vec{s}_i = s_i \in \{0, 1\}$ is simply a Ber(ε) random variable. Hence,

$$\begin{aligned}
\Pr\{\vec{s}_1 \notin \mathrm{csp}\,([\vec{s}_2|\vec{s}_3|\cdots|\vec{s}_{N-1}|\vec{s}_N])\} &= \Pr\left\{s_1 \notin \mathrm{span}\big(\{s_2, s_3, \cdots, s_{N-1}, s_N\}\big)\right\} \\
&= \varepsilon \Pr\left\{0 \notin \mathrm{span}\big(\{s_2, s_3, \cdots, s_{N-1}, s_N\}\big)\right\} \\
&\quad + \bar{\varepsilon} \Pr\left\{1 \notin \mathrm{span}\big(\{s_2, s_3, \cdots, s_{N-1}, s_N\}\big)\right\} \\
&\overset{(a)}{=} \bar{\varepsilon} \Pr\left\{1 \notin \mathrm{span}\big(\{s_2, s_3, \cdots, s_{N-1}, s_N\}\big)\right\} \\
&\overset{(b)}{=} \bar{\varepsilon} \Pr\{s_2 = s_3 = \cdots = s_{N-1} = s_N = 0\} \\
&= \bar{\varepsilon}\varepsilon^{N-1} \qquad (3.41)
\end{aligned}$$

where (a) is by the fact that $\Pr\left\{0 \notin \mathrm{span}\big(\{s_2, s_3, \cdots, s_{N-1}, s_N\}\big)\right\} = 0$ and (b) is by the fact that $1 \notin \mathrm{span}\big(\{s_2, s_3, \cdots, s_{N-1}, s_N\}\big)$ if and only if $s_i = 0$ for $2 \le i \le N$. Maximizing $\bar{\varepsilon}\varepsilon^{N-1}$ over ε, a sum multiplexing gain of $\left(1 - \frac{1}{N}\right)^{N-1}$ is achieved for $\varepsilon = 1 - \frac{1}{N}$. This is sketched in fig. 3.1 as the only decreasing curve. As N increases, SMG_N settles on $\frac{1}{e}$. □

Example 2- The Elements of The Signatures Are Not Necessarily Equiprobable Over $\mathcal{A}$.

Let us fix $\mathcal{A} = \{-1, 0, 1\}$ and $N = 2$. For $i = 1, 2$, the elements of $\vec{\boldsymbol{s}}_i$ are i.i.d. random variables taking the values 0, 1 and -1 with probabilities ε, $\overline{\varepsilon}p_1$ and $\overline{\varepsilon}p_{-1}$ respectively. We have

$$\begin{aligned}
\Pr\{\vec{\boldsymbol{s}}_1 \notin \mathrm{csp}([\vec{\boldsymbol{s}}_2])\} &= 1 - \Pr\{\vec{\boldsymbol{s}}_1 \in \mathrm{csp}([\vec{\boldsymbol{s}}_2])\} \\
&= 1 - \Pr\{\vec{\boldsymbol{s}}_1 = \vec{0}_K\} - \Pr\{\vec{\boldsymbol{s}}_1 \neq \vec{0}_K, \vec{\boldsymbol{s}}_1 = \pm\vec{\boldsymbol{s}}_2\} \\
&= 1 - \varepsilon^K - \Pr\{\vec{\boldsymbol{s}}_1 \neq \vec{0}_K, \vec{\boldsymbol{s}}_1 = \vec{\boldsymbol{s}}_2\} - \Pr\{\vec{\boldsymbol{s}}_1 \neq \vec{0}_K, \vec{\boldsymbol{s}}_1 = -\vec{\boldsymbol{s}}_2\}. \qquad (3.42)
\end{aligned}$$

However,

$$\begin{aligned}
\Pr\{\vec{\boldsymbol{s}}_1 \neq \vec{0}_K, \vec{\boldsymbol{s}}_1 = \vec{\boldsymbol{s}}_2\} &= \sum_{\vec{s} \in \mathrm{supp}(\vec{\boldsymbol{s}}_1) \setminus \{\vec{0}_K\}} \Big(\Pr\{\vec{\boldsymbol{s}}_1 = \vec{s}\}\Big)^2 \\
&= \sum_{k=0}^{K-1} \sum_{l=0}^{K-k} \binom{K}{k} \binom{K-k}{l} \varepsilon^{2k} (\overline{\varepsilon}p_1)^{2l} (\overline{\varepsilon}p_{-1})^{2(K-k-l)} \\
&= \left(\varepsilon^2 + \overline{\varepsilon}^2(p_1^2 + p_{-1}^2)\right)^K - \varepsilon^{2K}. \qquad (3.43)
\end{aligned}$$

Similarly,

$$\Pr\{\vec{\boldsymbol{s}}_1 \neq \vec{0}_K, \vec{\boldsymbol{s}}_1 = \vec{\boldsymbol{s}}_2\} = \left(\varepsilon^2 + 2\overline{\varepsilon}^2 p_1 p_{-1}\right)^K - \varepsilon^{2K}. \qquad (3.44)$$

Therefore,

$$\begin{aligned}
\mathrm{SMG}_2 &= \frac{2\Pr\{\vec{\boldsymbol{s}}_1 \notin \mathrm{csp}([\vec{\boldsymbol{s}}_2])\}}{K} \\
&= \frac{2}{K}\left(1 - \varepsilon^K + 2\varepsilon^{2K} - \left(\varepsilon^2 + \overline{\varepsilon}^2(p_1^2 + p_{-1}^2)\right)^K - \left(\varepsilon^2 + 2\overline{\varepsilon}^2 p_1 p_{-1}\right)^K\right). \qquad (3.45)
\end{aligned}$$

This expression is maximized at $p_1 = p_{-1} = \frac{1}{2}$ for any ε and K. Thus,

$$\sup_{p_1, p_{-1}} \mathrm{SMG}_2 = \frac{2\left(1 - \varepsilon^K + 2\varepsilon^{2K} - 2\left(\varepsilon^2 + \frac{\overline{\varepsilon}^2}{2}\right)^K\right)}{K}. \qquad (3.46)$$

Finally, the expression in the right side of (3.46) is maximized for $K = 2$ and $\varepsilon = 0.756$ where a sum multiplexing gain of 0.7091 is achieved. Although one expects that $\varepsilon = \frac{2}{3}$ is the best choice, it is not the case. The SMG is 0.6914 for $\varepsilon = \frac{2}{3}$ which is slightly less than 0.7091. □

Remark 1- For any matrix A, let $(\lambda_A^{(l)})_{l=1}^{\mathrm{rank}(A)}$ be the nonzero eigenvalues of the matrix $AA^\dagger$. Writing the eigenvalue decomposition [85] for $I_{K-1} + \alpha^2 P G_i^\dagger(\vec{s}) S \Xi_i \Xi_i^\dagger S^\dagger G_i^\dagger(\vec{s})$ and $I_K + \alpha^2 P S \Xi_i \Xi_i^\dagger S^\dagger$

appearing in the expression of $\varrho_i(P; \vec{s}, \vec{h}_i)$ in (3.37), we get

$$\begin{aligned}\varrho_i(P; \vec{s}, \vec{h}_i) &= \alpha^2 |h_{i,i}|^2 \|\vec{s}\|^2 P \prod_{S \in \mathrm{supp}(\boldsymbol{S}_i)} \frac{\prod_{l=1}^{\mathrm{rank}(G_i^\dagger(\vec{s})S)} \left(1 + \alpha^2 P \lambda^{(l)}_{G_i^\dagger(\vec{s})S\Xi_i}\right)^{\Pr\{\boldsymbol{S}_i = S\}}}{\prod_{l=1}^{\mathrm{rank(S)}} \left(1 + \alpha^2 P \lambda^{(l)}_{S\Xi_i}\right)^{\Pr\{\boldsymbol{S}_i = S\}}} \\ &= \alpha^2 |h_{i,i}|^2 \|\vec{s}\|^2 P^{\mu_i(\vec{s})} \prod_{S \in \mathrm{supp}(\boldsymbol{S}_i)} \frac{\prod_{l=1}^{\mathrm{rank}(G_i^\dagger(\vec{s})S)} \left(\frac{1}{P} + \alpha^2 \lambda^{(l)}_{G_i^\dagger(\vec{s})S\Xi_i}\right)^{\Pr\{\boldsymbol{S}_i = S\}}}{\prod_{l=1}^{\mathrm{rank}(S)} \left(\frac{1}{P} + \alpha^2 \lambda^{(l)}_{S\Xi_i}\right)^{\Pr\{\boldsymbol{S}_i = S\}}}\end{aligned} \tag{3.47}$$

where

$$\begin{aligned}\mu_i(\vec{s}) &= 1 - \sum_{S \in \mathrm{supp}(\boldsymbol{S}_i)} \Pr\{\boldsymbol{S}_i = S\} \mathrm{rank}(G_i^\dagger(\vec{s})S) - \sum_{S \in \mathrm{supp}(\boldsymbol{S}_i)} \Pr\{\boldsymbol{S}_i = S\} \mathrm{rank}(S) \\ &= 1 - \mathrm{E}\left\{\mathrm{rank}(G_i^\dagger(\vec{s})\boldsymbol{S}_i)\right\} - \mathrm{E}\{\mathrm{rank}(\boldsymbol{S}_i)\}.\end{aligned} \tag{3.48}$$

On the other hand, using (3.38),

$$\begin{aligned}\mathscr{R}_i^{(\mathrm{lb})}(\vec{h}_i) &\geq \frac{1}{K} \sum_{\vec{s} \in \mathrm{supp}(\vec{\boldsymbol{s}}_i) \setminus \{\vec{0}_K\}} \Pr\{\vec{\boldsymbol{s}}_i = \vec{s}\} \log\left(2^{-\mathrm{H}\left(\boldsymbol{S}_i \Xi_i \Xi_i^\dagger \boldsymbol{S}_i^\dagger\right)} \varrho_i(P; \vec{s}, \vec{h}_i)\right) \\ &= \frac{1}{K} \sum_{\vec{s} \in \mathrm{supp}(\vec{\boldsymbol{s}}_i) \setminus \{\vec{0}_K\}} \Pr\{\vec{\boldsymbol{s}}_i = \vec{s}\} \log \varrho_i(P; \vec{s}, \vec{h}_i) - \frac{\Pr\{\vec{\boldsymbol{s}}_i \neq \vec{0}_K\} \mathrm{H}\left(\boldsymbol{S}_i \Xi_i \Xi_i^\dagger \boldsymbol{S}_i^\dagger\right)}{K}.\end{aligned} \tag{3.49}$$

Combining (3.47) and (3.49),

$$\mathscr{R}_i^{(\mathrm{lb})}(\vec{h}_i) \geq \mathrm{MG} \log \gamma - \mathrm{IEF} + \mathrm{CSF}_i \tag{3.50}$$

where the three terms appearing in the right side of (3.50) are the *Multiplexing Gain* defined by

$$\mathrm{MG} \triangleq \frac{1}{K} \sum_{\vec{s} \in \mathrm{supp}(\vec{\boldsymbol{s}}_i) \setminus \{\vec{0}_K\}} \Pr\{\vec{\boldsymbol{s}}_i = \vec{s}\} \mu_i(\vec{s}), \tag{3.51}$$

the *Interference Entropy Factor* defined by

$$\mathrm{IEF} \triangleq \frac{\Pr\{\vec{\boldsymbol{s}}_i \neq \vec{0}_K\} \mathrm{H}\left(\boldsymbol{S}_i \Xi_i \Xi_i^\dagger \boldsymbol{S}_i^\dagger\right)}{K} \tag{3.52}$$

and the *Channel-Signature Factor* defined by

$$\mathrm{CSF}_i \triangleq \frac{1}{K}\sum_{\vec{s}\in \mathrm{supp}(\vec{\boldsymbol{s}}_i)\backslash\{\vec{0}_K\}} \Pr\{\vec{\boldsymbol{s}}_i=\vec{s}\}\log\left(\alpha^2|h_{i,i}|^2\|\vec{s}\|^2\right)$$

$$+\frac{1}{K}\sum_{\vec{s}\in \mathrm{supp}(\vec{\boldsymbol{s}}_i)\backslash\{\vec{0}_K\}} \Pr\{\vec{\boldsymbol{s}}_i=\vec{s}\}\sum_{S\in\mathrm{supp}(\boldsymbol{S}_i)}\log\frac{\prod_{l=1}^{\mathrm{rank}(G_i^\dagger(\vec{s})S)}\left(\frac{1}{P}+\alpha^2\lambda^{(l)}_{G_i^\dagger(\vec{s})S\Xi_i}\right)^{\Pr\{\boldsymbol{S}_i=S\}}}{\prod_{l=1}^{\mathrm{rank}(S)}\left(\frac{1}{P}+\alpha^2\lambda^{(l)}_{S\Xi_i}\right)^{\Pr\{\boldsymbol{S}_i=S\}}}. \tag{3.53}$$

By (3.48), (3.51) and the identity (5.58) appearing in the proof of Proposition 5 in appendix G, it is seen that

$$\mathrm{MG}=\frac{\Pr\{\vec{\boldsymbol{s}}_i\notin \mathrm{csp}(\boldsymbol{S}_i)\}}{K} \tag{3.54}$$

which is in agreement with the result of Proposition 5.

In general, MG does not depend on the user index. Moreover, assuming the channel gains are realizations of i.i.d. and continuous random variables, the discrete entropy $\mathrm{H}\left(\boldsymbol{S}_i\Xi_i\Xi_i^\dagger\boldsymbol{S}_i^\dagger\right)$ is not a function of the user index. In fact, a simple argument shows that

$$\mathrm{H}\left(\boldsymbol{S}_i\Xi_i\Xi_i^\dagger\boldsymbol{S}_i^\dagger\right)=(N-1)\mathrm{H}\left(\vec{\boldsymbol{s}}_1\vec{\boldsymbol{s}}_1^\dagger\right). \tag{3.55}$$

The interplay between MG, IEF and CSF_i determines the behavior of the achievable rate. As we will see in the next section, a larger MG is achieved at the cost of a larger IEF. This can be in particular understood if we study the effect of the underlying alphabet $\mathcal{A}$ on these two factors. If $\mathcal{A}$ increases, IEF becomes larger as the randomness of the signature vectors rises. At the same time, the chance for the signature vector of a user to be spanned by signature vectors of other users drops implying a larger MG as shown in fig. 3.1. By (3.50), a larger IEF reduces the rate and a larger MG increases the rate. Due to the fact that MG identifies the scaling of rate with $\log P$, these opposing effects identify a tradeoff between the achievable rates in moderate and high SNR regimes. Note that $\mathscr{R}_i^{(\mathrm{lb})}(\vec{h}_i)$ is only a lower bound to $\mathscr{R}_i(\vec{h}_i)$. Monte-Carlo simulations (to be offered in section 3.4) verify that the same tradeoff holds for $\mathscr{R}_i(\vec{h}_i)$, i.e., $\mathscr{R}_i^{(\mathrm{lb})}(\vec{h}_i)$ is tight enough to mimic the behavior of $\mathscr{R}_i(\vec{h}_i)$. □

In the next two subsection we address the following two questions, respectively:

1- Let all users have access to a globally known signature-book $\mathcal{C}$ and each user selects a signature in $\mathcal{C}$ uniformly and independently from signature interval to signature interval. What is the ultimate SMG in this scenario and how can one design $\mathcal{C}$ to achieve this SMG?

2- Let all users utilize matched filters at their receiver side as a low-complexity detection technique. What is the effect of matched filtering on SMG?

3.3.1 A Modified Approach To Generate The Signatures

The expression for SMG_N given in (3.40) is valid for other classes of randomized algorithms in generating the signature vectors. One can consider a general scheme where the i^{th} user generates its signature vector $\vec{\boldsymbol{s}}_i$ by randomly selecting a vector in a globally known set of L signatures $\mathcal{C}_L = \{\vec{\mathsf{c}}_1, \cdots, \vec{\mathsf{c}}_L\} \subset \mathbb{R}^K$. For example, generating the signature vectors over some alphabet $\mathcal{A} = \{-1, +1\}$ is equivalent to setting $L = 2^K$ and $\mathcal{C}_{2^K} = A^K$.

We determine the largest achievable SMG in a network of N users under the following assumptions:

1- All users have access to a globally known signature-book $\mathcal{C}_L \subset \mathbb{R}^K$ containing L *nonzero* and *non-parallel* vectors.

2- The signature vectors are selected with equal probability in $\mathcal{C}_L$.

For each $1 \leq l \leq L$ and $1 \leq m \leq L-1$, let us denote by $r_{l,m}$ the number of distinct subsets $\mathcal{B}$ of $\mathcal{C}_L$ such that $|\mathcal{B}| = m$ and $\vec{\mathsf{c}}_l \notin \operatorname{span}(\mathcal{B})$. We denote these subsets explicitly by $\mathcal{B}_{l,m}^{(1)}, \cdots, \mathcal{B}_{l,m}^{(r_{l,m})}$.Then,

$$\Pr\{\vec{\boldsymbol{s}}_1 \notin \operatorname{csp}([\vec{\boldsymbol{s}}_2|\vec{\boldsymbol{s}}_3|\cdots|\vec{\boldsymbol{s}}_{N-1}|\vec{\boldsymbol{s}}_N])\} = \frac{1}{L}\sum_{l=1}^{L}\Pr\{\vec{\mathsf{c}}_l \notin \operatorname{csp}([\vec{\boldsymbol{s}}_2|\vec{\boldsymbol{s}}_3|\cdots|\vec{\boldsymbol{s}}_{N-1}|\vec{\boldsymbol{s}}_N])\} \tag{3.56}$$

where for each $1 \leq l \leq L$,

$$\Pr\{\vec{\mathsf{c}}_l \notin \operatorname{csp}([\vec{\boldsymbol{s}}_2|\vec{\boldsymbol{s}}_3|\cdots|\vec{\boldsymbol{s}}_{N-1}|\vec{\boldsymbol{s}}_N])\} = \sum_{m=1}^{L-1}\sum_{r=1}^{r_{l,m}}\Pr\left\{\{\vec{\boldsymbol{s}}_2, \cdots, \vec{\boldsymbol{s}}_{N-1}, \vec{\boldsymbol{s}}_N\} = \mathcal{B}_{l,m}^{(r)}\right\}. \tag{3.57}$$

It is easy to see that[3]

$$\Pr\left\{\{\vec{\boldsymbol{s}}_2, \cdots, \vec{\boldsymbol{s}}_{N-1}, \vec{\boldsymbol{s}}_N\} = \mathcal{B}_{l,m}^{(r)}\right\} = \frac{\phi_{N,m}}{L^{N-1}} \tag{3.58}$$

where

$$\phi_{N,m} \triangleq \sum_{\substack{n_1+\cdots+n_m=N-1 \\ n_1,\cdots,n_m\geq 1}} \frac{(N-1)!}{n_1!\cdots n_m!} \tag{3.59}$$

for any $1 \leq l \leq L$, $1 \leq m \leq L-1$ and $1 \leq r \leq r_{l,m}$. By (3.56), (3.57) and (3.58),

$$\Pr\{\vec{\boldsymbol{s}}_1 \notin \operatorname{csp}([\vec{\boldsymbol{s}}_2|\vec{\boldsymbol{s}}_3|\cdots|\vec{\boldsymbol{s}}_{N-1}|\vec{\boldsymbol{s}}_N])\} = \frac{1}{L^N}\sum_{l=1}^{L}\sum_{m=1}^{L-1} r_{l,m}\phi_{N,m} \tag{3.60}$$

Therefore,

$$\mathrm{SMG}_N = \frac{N\sum_{l=1}^{L}\sum_{m=1}^{L-1} r_{l,m}\phi_{N,m}}{KL^N}. \tag{3.61}$$

[3] Assuming the $1^{\text{st}}, \cdots, (m-1)^{\text{th}}$ and m^{th} elements of $\mathcal{B}_{l,m}^{(r)}$ are chosen by $n_1, \cdots, n_{m-1}$ and n_m users respectively (each user selects one signature), this can happen in $\sum_{\substack{n_1+\cdots+n_m=N-1 \\ n_1,\cdots,n_m\geq 1}} \frac{(N-1)!}{n_1!\cdots n_m!}$ different ways.

There is no closed expression for $\phi_{N,m}$. However, one can use the recursion

$$m^{N-1} = \phi_{N,m} + \sum_{v=1}^{m-1} \binom{m}{v} \phi_{N,m-v} \tag{3.62}$$

to compute this quantity[4]. The following Proposition shows that SMG_N given in (3.61) can be made arbitrarily close to 1 for any finite N.

Proposition 6. *For any number of users N and any integer $L \geq N$, there exists a signature-book $\mathcal{C}_L \subset \mathbb{R}^N$ of size L such that*

$$\mathrm{SMG}_N = \left(1 - \frac{1}{L}\right)^{N-1}. \tag{3.63}$$

Moreover, this is the highest sum multiplexing gain one can achieve in a decentralized network with N users where each user selects its signatures uniformly in a globally known signature-book of size L.

Proof. Noting that $\phi_{N,m} = 0$ for $m \geq N$, we can write

$$\mathrm{SMG}_N = \frac{N \sum_{l=1}^{L} \sum_{m=1}^{N-1} r_{l,m} \phi_{N,m}}{K L^N}. \tag{3.64}$$

It is seen that SMG_N is maximized if $r_{l,m}$ is as large as possible for each $1 \leq l \leq L$ and $1 \leq m \leq N-1$. The quantity $r_{l,m}$ can not be larger than $\binom{L-1}{m}$. For the moment, let us assume that there exists a signature-book $\mathcal{C}_L \subset \mathbb{R}^K$ such that any N elements of $\mathcal{C}_L$ are linearly independent. We will soon provide the constructive proof for the existence of this signature-book. Obviously, one requires $K \geq N$. In this case, $r_{l,m} = \binom{L-1}{m}$ for any $1 \leq l \leq L$ and $1 \leq m \leq N-1$. Hence,

$$\mathrm{SMG}_N = \frac{N \sum_{m=1}^{N-1} \binom{L-1}{m} \phi_{N,m}}{K L^{N-1}}. \tag{3.65}$$

To proceed, we need the following Lemma.

Lemma 7.

$$\sum_{m=1}^{N-1} \binom{L-1}{m} \phi_{N,m} = (L-1)^{N-1}. \tag{3.66}$$

Proof. We offer a combinatorial proof. There are $(L-1)^{N-1}$ different ways to distribute $N-1$ distinct balls in $L-1$ distinct boxes. If the balls were similar and we put v_i balls in the i^{th} box, the number of different ways is the number of non-negative solutions in $\{v_i\}_{i=1}^{L-1}$ for the equation

[4] We omit the proof of (3.62) for brevity.

$v_1 + \cdots + v_{L-1} = N-1$. Fixing $\{v_i\}_{i=1}^{L-1}$ and recalling that the balls are distinct, one gets $\frac{(N-1)!}{v_1! \cdots v_{L-1}!}$ different distributions. Hence,

$$(L-1)^{N-1} = \sum_{\substack{v_1+\cdots+v_{L-1}=N-1 \\ v_1,\cdots,v_{L-1}\geq 0}} \frac{(N-1)!}{v_1! \cdots v_{L-1}!}. \tag{3.67}$$

A simple counting argument shows that

$$\sum_{\substack{v_1+\cdots+v_{L-1}=N-1 \\ v_1,\cdots,v_{L-1}\geq 0}} \frac{(N-1)!}{v_1! \cdots v_{L-1}!} = \sum_{m'=L-N}^{L-2} \binom{L-1}{m'} \sum_{\substack{v_1+\cdots+v_{L-m'-1}=N-1 \\ v_1,\cdots,v_{L-m'-1}\geq 1}} \frac{(N-1)!}{v_1! \cdots v_{L-m'-1}!}. \tag{3.68}$$

By (3.59), (3.67) and (3.68),

$$\sum_{m'=L-N}^{L-2} \binom{L-1}{m'} \phi_{N,L-m'-1} = (L-1)^{N-1}. \tag{3.69}$$

Replacing the dummy variable m' by $L-m-1$ yields

$$\sum_{m=1}^{N-1} \binom{L-1}{m} \phi_{N,m} = (L-1)^{N-1}. \tag{3.70}$$

This is the desired result. $\square$

Using (3.66) in (3.65),

$$\mathrm{SMG}_N = \frac{N}{K}\left(1-\frac{1}{L}\right)^{N-1}. \tag{3.71}$$

To get the largest SMG, one may set $K=N$. Then,

$$\mathrm{SMG}_N = \left(1-\frac{1}{L}\right)^{N-1}. \tag{3.72}$$

This can be made arbitrarily close to one by increasing L.

To complete the proof, we must show the existence of the signature-book $\mathcal{C}_L$ with the given properties. Let $K=N$ and $b_1, \cdots, b_L$ be $L \geq N$ distinct and nonzero real numbers. We define

$$\mathcal{C}_L = \left\{ \vec{\mathsf{c}}_l = \begin{pmatrix} 1 & b_l & b_l^2 & \cdots & b_l^{N-1} \end{pmatrix}^{\mathrm{t}} : 1 \leq l \leq L \right\}. \tag{3.73}$$

Let us take any N elements in $\mathcal{C}_L$, say $\vec{\mathsf{c}}_{l_1}, \cdots, \vec{\mathsf{c}}_{l_{N-1}}$ and $\vec{\mathsf{c}}_{l_N}$. Assume there are real numbers $a_1, \cdots, a_{N-1}$ and a_N such that

$$a_1\vec{\mathsf{c}}_{l_1} + \cdots + a_{N-1}\vec{\mathsf{c}}_{l_{N-1}} + a_N\vec{\mathsf{c}}_{l_N} = \vec{0}_N. \tag{3.74}$$

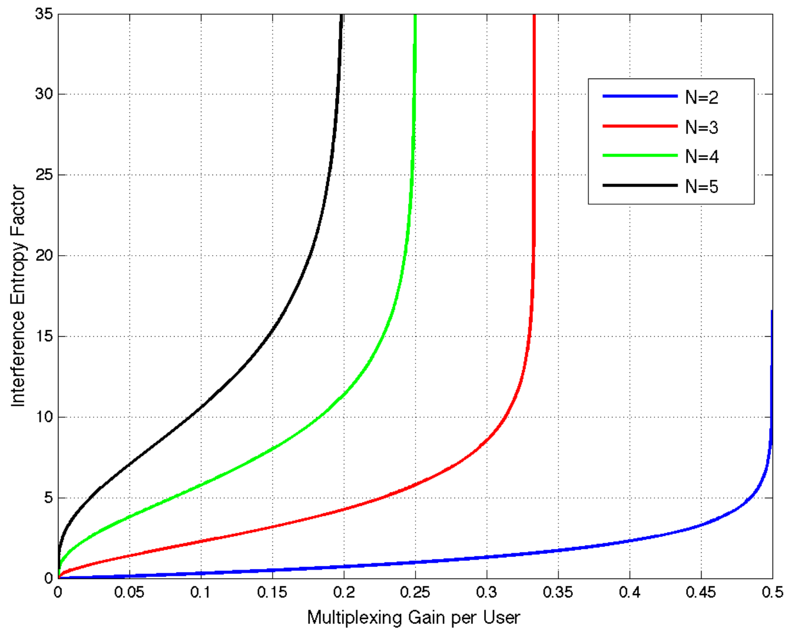

Figure 3.2: The tradeoff between multiplexing gain per user and interference entropy factor achieved by the construction in the proof of Proposition 6.

This yields,

$$\begin{pmatrix} 1 & 1 & \cdots & 1 & 1 \\ b_{l_1} & b_{l_2} & \cdots & b_{l_{N-1}} & b_{l_N} \\ b_{l_1}^2 & b_{l_2}^2 & \cdots & b_{l_{N-1}}^2 & b_{l_N}^2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ b_{l_1}^{N-1} & b_{l_2}^{N-1} & \cdots & b_{l_{N-1}}^{N-1} & b_{l_N}^{N-1} \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_{N-1} \\ a_N \end{pmatrix} = \vec{0}_N. \tag{3.75}$$

The matrix appearing in (3.75) is a Vandermonde matrix. The determinant of this matrix is $\prod_{1\le i<j\le N}\left(b_{l_i}-b_{l_j}\right)$ which is a nonzero number. This proves that any N elements in $\mathcal{C}_L$ are linearly independent. The proof of Proposition 6 is complete. $\square$

By Proposition 6, $\mathrm{SMG}_N = 1-\delta$ can be achieved in a network of N users for any $\delta > 0$. This is at the expense of an interference entropy factor equal to $-(N-1)\log\left(1-(1-\delta)^{\frac{1}{N-1}}\right)$. The corresponding tradeoff between multiplexing gain per user and the interference entropy factor is sketched in fig. 3.2 for several values of N.

Note that this scheme requires all users to select their codes in an appropriately designed and globally known set of signatures. This might be a restriction in a decentralized network where there is no explicit cooperation among users.

3.3.2 Effect of Matched Filtering on SMG

In order to reduce the computational complexity in the decoder, a suboptimal detector such as single-user Matched Filter (MF) is frequently used in conventional CDMA systems. For a detailed analysis, we refer the reader to [19]. Next, we investigate how the presence of such a detector affects the SMG in a decentralized network of users.

The output of the MF at the receiver side of the i^{th} user is the inner product of $\vec{\boldsymbol{y}}_i$ and $\vec{\boldsymbol{s}}_i$. We have

$$\vec{\boldsymbol{s}}_i^{\dagger}\vec{\boldsymbol{y}}_i = \alpha h_{i,i}\boldsymbol{\eta}_{i,i}\boldsymbol{x}_i + \alpha \sum_{j\neq i} h_{i,j}\boldsymbol{\eta}_{i,j}\boldsymbol{x}_j + \vec{\boldsymbol{s}}_i^{\dagger}\vec{\boldsymbol{z}}_i \tag{3.76}$$

where

$$\boldsymbol{\eta}_{i,j} \triangleq \vec{\boldsymbol{s}}_i^{\dagger}\vec{\boldsymbol{s}}_j \tag{3.77}$$

for $1 \leq j \leq N$. Following the same analysis that led us to (3.40), the achievable SMG denoted by $\mathrm{SMG}_N^{(\mathrm{MF})}$ is given by

$$\begin{aligned}
\mathrm{SMG}_N^{(\mathrm{MF})} &= \frac{N}{K}\Pr\left\{\boldsymbol{\eta}_{1,1} \notin \mathrm{span}(\{\boldsymbol{\eta}_{1,2},\cdots,\boldsymbol{\eta}_{1,N}\})\right\} \\
&\stackrel{(a)}{=} \frac{N}{K}\Pr\left\{\boldsymbol{\eta}_{1,1} \neq 0, \boldsymbol{\eta}_{1,2} = \cdots = \boldsymbol{\eta}_{1,N} = 0\right\} \\
&= \frac{N}{K}\sum_{\vec{s}\in\mathrm{supp}(\vec{\boldsymbol{s}}_i)\backslash\{\vec{0}_K\}} \Pr\left\{\vec{\boldsymbol{s}}_i = \vec{s}, \boldsymbol{\eta}_{1,2} = \cdots = \boldsymbol{\eta}_{1,N} = 0\right\} \\
&\stackrel{(b)}{=} \frac{N}{K}\sum_{\vec{s}\in\mathrm{supp}(\vec{\boldsymbol{s}}_i)\backslash\{\vec{0}_K\}} \Pr\left\{\vec{s}^{\,\dagger}\vec{\boldsymbol{s}}_2 = \cdots = \vec{s}^{\,\dagger}\vec{\boldsymbol{s}}_N = 0\right\}\Pr\{\vec{\boldsymbol{s}}_i = \vec{s}\} \\
&\stackrel{(c)}{=} \frac{N}{K}\sum_{\vec{s}\in\mathrm{supp}(\vec{\boldsymbol{s}}_i)\backslash\{\vec{0}_K\}} \left(\Pr\left\{\vec{s}^{\,\dagger}\vec{\boldsymbol{s}}_2 = 0\right\}\right)^{N-1}\Pr\{\vec{\boldsymbol{s}}_i = \vec{s}\}.
\end{aligned} \tag{3.78}$$

In 3.78, (a) is due to the fact that $\boldsymbol{\eta}_{1,j}$ is a scalar for $1 \leq j \leq N$, (b) holds by independence of $\vec{\boldsymbol{s}}_i$ and the collection $(\vec{\boldsymbol{s}}_j)_{j\neq i}$, and (c) is due to the fact that $(\vec{\boldsymbol{s}}_j)_{j\neq i}$ is a collection of i.i.d. random variables.

As a working example, let $\mathcal{A} = \{-1, 1\}$, $p_1 = \nu$ and $p_{-1} = \overline{\nu}$. Then,

$$\begin{aligned}
\mathrm{SMG}_N^{(\mathrm{MF})} &= \frac{N}{K}\sum_{k=0}^{K}\sum_{\vec{s}:k \text{ elements of } \vec{s} \text{ are } +1} \left(\Pr\left\{\vec{s}^{\,\dagger}\vec{\boldsymbol{s}}_2 = 0\right\}\right)^{N-1}\Pr\{\vec{\boldsymbol{s}}_i = \vec{s}\} \\
&= \frac{N}{K}\sum_{k=0}^{K}\binom{K}{k}\nu^k\overline{\nu}^{K-k}\left(\Pr\left\{\vec{s}^{\,\dagger}\vec{\boldsymbol{s}}_2 = 0, k \text{ elements of } \vec{s} \text{ are } 1\right\}\right)^{N-1}.
\end{aligned} \tag{3.79}$$

In appendix I, it is shown that

$$\Pr\left\{\vec{s}^{\dagger}\vec{\boldsymbol{s}}_2 = 0, k \text{ elements of } \vec{s} \text{ are } 1\right\} = \binom{K}{\frac{K}{2}} \nu^{\frac{K}{2}-k} \overline{\nu}^{\frac{K}{2}+k} \operatorname{E}\left\{\left(\frac{\nu}{\overline{\nu}}\right)^{2\boldsymbol{\kappa}}\right\} \tag{3.80}$$

where $\boldsymbol{\kappa}$ is a Hypergeometric random variable with PMF

$$\Pr\{\boldsymbol{\kappa} = l\} = \frac{\binom{k}{l}\binom{K-k}{\frac{K}{2}-l}}{\binom{K}{\frac{K}{2}}}. \tag{3.81}$$

It can be verified that $\mathrm{SMG}_N^{(\mathrm{MF})}$ is maximized for $K = 2$ for all values of $\nu \in [0, 1]$. In this case,

$$\Pr\left\{\vec{s}^{\dagger}\vec{\boldsymbol{s}}_2 = 0, \text{no element of } \vec{s} \text{ is } -1\right\} = 2\nu\overline{\nu}, \tag{3.82}$$

$$\Pr\left\{\vec{s}^{\dagger}\vec{\boldsymbol{s}}_2 = 0, \text{both elements of } \vec{s} \text{ are } -1\right\} = 2\nu\overline{\nu}, \tag{3.83}$$

and

$$\Pr\left\{\vec{s}^{\dagger}\vec{\boldsymbol{s}}_2 = 0, \text{only one element of } \vec{s} \text{ is } -1\right\} = \nu^2 + \overline{\nu}^2. \tag{3.84}$$

Therefore,

$$\mathrm{SMG}_N^{(\mathrm{MF})} = \frac{N}{2}\left((\nu^2+\overline{\nu}^2)(2\nu\overline{\nu})^{N-1} + 2\nu\overline{\nu}(\nu^2+\overline{\nu}^2)^{N-1}\right). \tag{3.85}$$

Fig. 3.3 sketches $\mathrm{SMG}_N^{(\mathrm{MF})}$ in terms of ν for different values of N. It is seen that:

• $\sup_{\nu\in(0,1]} \mathrm{SMG}_2^{(\mathrm{MF})} = \frac{1}{2}$ which is equal to the SMG of the optimal detector (no matched filtering). This holds in spite of the fact that $\vec{\boldsymbol{y}}_i$ and $\vec{\boldsymbol{s}}_i^{\dagger}\vec{\boldsymbol{y}}_i$ are not the same from an information theoretic point of view, i.e., $\mathrm{I}(\boldsymbol{x}_i; \vec{\boldsymbol{y}}_i|\vec{\boldsymbol{s}}_i) > \mathrm{I}(\boldsymbol{x}_i; \vec{\boldsymbol{s}}_i^{\dagger}\vec{\boldsymbol{y}}_i|\vec{\boldsymbol{s}}_i)$ or equivalently, $\vec{\boldsymbol{s}}_i^{\dagger}\vec{\boldsymbol{y}}_i$ is not a sufficient statistic. However, for $N \geq 3$, the SMG of MF is considerably lower than that of the optimal scenario[5].

• The optimum value of ν is not equal to 0.5 in general. This happens for any $N \geq 5$, while for $N \leq 4$, the optimum value of ν is 0.5.

• As the number of users increases, $\mathrm{SMG}_N^{(\mathrm{MF})}$ decreases and settles on 0.182. Selecting $\nu = 0.5$ in (3.85) yields an SMG of $\frac{N}{2^N}$ that tends to 0 as N becomes large.

3.4 System Design in Finite SNR

An important issue in a decentralized network is to propose a globally known utility function to be optimized by all users without any cooperation. As mentioned earlier, there is no closed expression for $\mathscr{R}_i(\vec{h}_i)$. However, we have developed a lower bound $\mathscr{R}_i^{(\mathrm{lb})}(\vec{h}_i)$ on $\mathscr{R}_i(\vec{h}_i)$ which is tight enough to guarantee $\mathscr{R}_i^{(\mathrm{lb})}(\vec{h}_i) \overset{\gamma}{\sim} \mathscr{R}_i$.

[5]Refer to fig. 3.1 for the SMG of the optimal scheme.

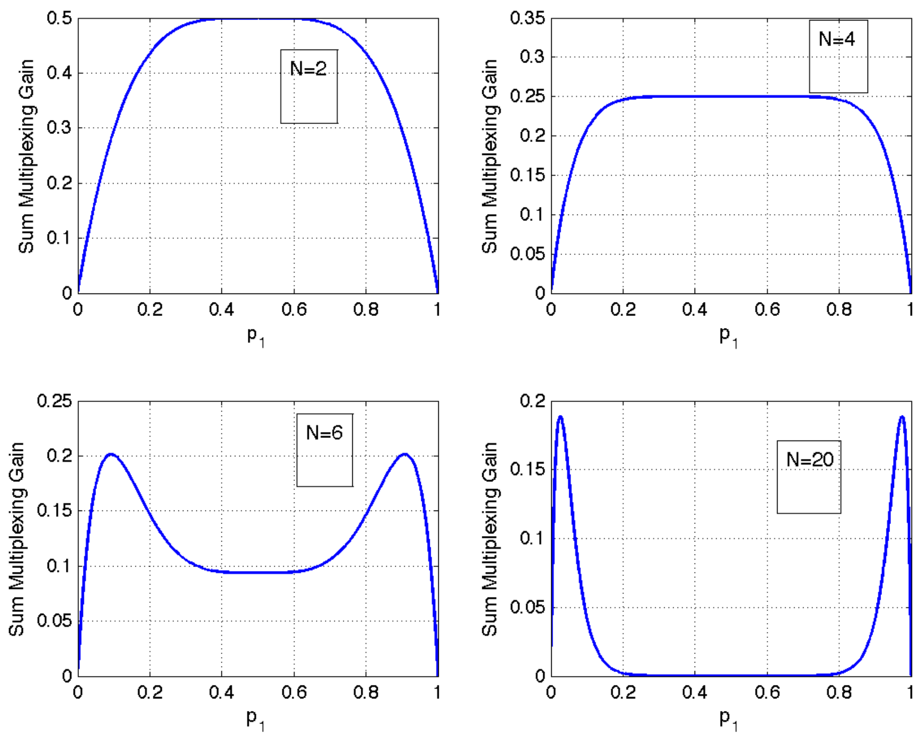

Figure 3.3: Plots of $\mathrm{SMG}_N^{(\mathrm{MF})}$ for different number of users. The signatures are only based on randomized spreading over $\{-1,+1\}$. It is seen that the elements of the signatures are not in general equiprobable over the underlying alphabet.

Let the channel gains $h_{i,j}$ be realizations of independent $\mathcal{CN}(0,1)$ random variables $\boldsymbol{h}_{i,j}$ representing Rayleigh fading. We propose that the i^{th} user selects K, ε and $(p_a)_{a\in\mathcal{A}\backslash\{0\}}$ (or a subset of these parameters) based on

$$(\hat{K},\hat{\varepsilon},(\hat{p}_a)_{a\in\mathcal{A}\backslash\{0\}}) = \arg \sup_{(K,\varepsilon,(p_a)_{a\in\mathcal{A}\backslash\{0\}})} \mathrm{E}\left\{\mathscr{R}_i^{(\mathrm{lb})}(\vec{\boldsymbol{h}}_i)\right\}. \tag{3.86}$$

Thereafter, the i^{th} user regulates the actual transmission rate of its signal code-book at $R_i = \mathscr{R}_i^{(\mathrm{lb})}(\vec{h}_i)$ using the realization $\vec{h}_i$ of $\vec{\boldsymbol{h}}_i$.

Example 3- Spreading vs. No Spreading: A Win-Win Situation.

Let us consider a network with $N = 2$ users. We use (3.37) to compute $\varrho_1(P; \vec{s}, \vec{\boldsymbol{h}}_1)$. It is seen that

1- $\boldsymbol{\Xi}_1 = \boldsymbol{h}_{2,1}$.

2- Since $\boldsymbol{S}_1 = \vec{\boldsymbol{s}}_2$, for each $\vec{t} \in \mathrm{supp}(\boldsymbol{S}_1)\backslash\{\vec{0}_K\}$, we have $\mathrm{rank}(\vec{t}) = 1$ and $\lambda^{(1)}_{\vec{t}\boldsymbol{\Xi}_1} = |\boldsymbol{h}_{2,1}|^2\|\vec{t}\|^2$.

3- For each $\vec{s} \in \mathrm{supp}(\vec{\boldsymbol{s}}_1)$ and $\vec{t} \in \mathrm{supp}(\boldsymbol{S}_1)$, we have $\mathrm{rank}(G_1^\dagger(\vec{s})\vec{t}\,) \leq 1$. Indeed, $G_1^\dagger(\vec{s})\vec{t} \in \mathbb{R}$ and if $G_1^\dagger(\vec{s})\vec{t} \neq 0$, then $\lambda^{(1)}_{G_1^\dagger(\vec{s})\vec{t}\boldsymbol{\Xi}_1} = |\boldsymbol{h}_{2,1}|^2|G_1^\dagger(\vec{s})\vec{t}|^2$.

Therefore,

$$\varrho_1(P; \vec{s}, \vec{\boldsymbol{h}}_1) = \alpha^2|\boldsymbol{h}_{1,1}|^2\|\vec{s}\|^2 P \prod_{\vec{t}\in\mathrm{supp}(\vec{\boldsymbol{s}}_2)} \left(\frac{1 + \alpha^2 P|\boldsymbol{h}_{2,1}|^2|G_1^\dagger(\vec{s})\vec{t}|^2}{1 + \alpha^2 P|\boldsymbol{h}_{2,1}|^2\|\vec{t}\|^2}\right)^{\Pr\{\vec{\boldsymbol{s}}_2=\vec{t}\}}. \tag{3.87}$$

In this example, we only consider the optimization in (3.86) with respect to ε and $(p_a)_{a\in\mathcal{A}\backslash\{0\}}$. In fact, we assume K is globally known. We consider two different schemes corresponding to two different values of K, i.e., $K = 2$ and $K = 1$, respectively.

<u>Scheme A</u>- Let $K = 2$, $\mathcal{A} = \{-1, 0, 1\}$, $p_1 = \nu$ and $p_{-1} = \overline{\nu}$ for some $\nu \in (0, 1]$. To simplify the expression for $\varrho_1(P; \vec{s}, \vec{\boldsymbol{h}}_1)$ in (3.87), we make the following observations:

1- If $\vec{s}$ has only one nonzero element,

$$|G_1^\dagger(\vec{s})\vec{t}|^2 = \begin{cases} 0 & \vec{t} = \vec{0}_2 \text{ or } \vec{t} = \pm\vec{s} \\ 1 & \text{oth.} \end{cases}. \tag{3.88}$$

2- If every element of $\vec{s}$ is nonzero,

$$|G_1^\dagger(\vec{s})\vec{t}|^2 = \begin{cases} 0 & \vec{t} = \vec{0}_2 \text{ or } \vec{t} = \pm\vec{s} \\ 2 & \vec{t}^{\mathrm{t}}\vec{s} = 0, \vec{t} \neq \vec{0}_2 \\ \frac{1}{2} & \vec{t}^{\mathrm{t}}\vec{s} \neq 0 \end{cases}. \tag{3.89}$$

Noting that $\alpha^2 = \frac{\gamma}{\overline{\varepsilon}P}$, we get

1- If $\vec{s}$ has only one nonzero element,

$$\varrho_1(P; \vec{s}, \vec{\boldsymbol{h}}_1) = \frac{|\boldsymbol{h}_{1,1}|^2\gamma}{\overline{\varepsilon}\left(1 + \frac{|\boldsymbol{h}_{2,1}|^2\gamma}{\overline{\varepsilon}}\right)^{\overline{\varepsilon}\varepsilon-\overline{\varepsilon}^2}\left(1 + \frac{2|\boldsymbol{h}_{2,1}|^2\gamma}{\overline{\varepsilon}}\right)^{\overline{\varepsilon}^2}}. \tag{3.90}$$

2- If $\vec{s} \in \{(1 \quad 1)^{\mathrm{t}}, (-1 \quad -1)^{\mathrm{t}}\}$,

$$\varrho_1(P; \vec{s}, \vec{\boldsymbol{h}}_1) = \frac{2|\boldsymbol{h}_{1,1}|^2\gamma\left(1 + \frac{|\boldsymbol{h}_{2,1}|^2\gamma}{2\overline{\varepsilon}}\right)^{2\overline{\varepsilon}\varepsilon}}{\overline{\varepsilon}\left(1 + \frac{|\boldsymbol{h}_{2,1}|^2\gamma}{\overline{\varepsilon}}\right)^{2\overline{\varepsilon}\varepsilon}\left(1 + \frac{2|\boldsymbol{h}_{2,1}|^2\gamma}{\overline{\varepsilon}}\right)^{\overline{\varepsilon}^2(\nu^2+\overline{\nu}^2)}}. \tag{3.91}$$

3- If $\vec{s} \in \{(1 \quad -1)^{\mathrm{t}}, (-1 \quad 1)^{\mathrm{t}}\}$,

$$\varrho_1(P; \vec{s}, \vec{\boldsymbol{h}}_1) = \frac{2|\boldsymbol{h}_{1,1}|^2\gamma\left(1+\frac{|\boldsymbol{h}_{2,1}|^2\gamma}{2\overline{\varepsilon}}\right)^{2\overline{\varepsilon}\varepsilon}}{\overline{\varepsilon}\left(1+\frac{|\boldsymbol{h}_{2,1}|^2\gamma}{\overline{\varepsilon}}\right)^{2\overline{\varepsilon}\varepsilon}\left(1+\frac{2|\boldsymbol{h}_{2,1}|^2\gamma}{\overline{\varepsilon}}\right)^{2\overline{\varepsilon}^2\nu\overline{\nu}}}. \tag{3.92}$$

Moreover, it is shown in appendix J that

$$\mathrm{H}\left(\vec{\boldsymbol{s}}_1\vec{\boldsymbol{s}}_1^{\dagger}\right) = 2\mathscr{H}(\varepsilon) + \overline{\varepsilon}^2\mathscr{H}(\nu^2+\overline{\nu}^2). \tag{3.93}$$

It is not hard to see that setting $\nu = 0.5$ maximizes $\mathscr{R}_i^{(\mathrm{lb})}(\vec{\boldsymbol{h}}_i)$. Hence, we come up with

$$\begin{aligned}\sup_{\nu\in(0,1]}\mathscr{R}_1^{(\mathrm{lb})}(\vec{\boldsymbol{h}}_1) &= \overline{\varepsilon}\varepsilon\log\left(1+\frac{2^{-2\mathscr{H}(\varepsilon)-\overline{\varepsilon}^2}|\boldsymbol{h}_{1,1}|^2\gamma}{\overline{\varepsilon}\left(1+\frac{|\boldsymbol{h}_{2,1}|^2\gamma}{\overline{\varepsilon}}\right)^{\overline{\varepsilon}\varepsilon-\overline{\varepsilon}^2}\left(1+\frac{2|\boldsymbol{h}_{2,1}|^2\gamma}{\overline{\varepsilon}}\right)^{\overline{\varepsilon}^2}}\right)\\ &\quad+\frac{\overline{\varepsilon}^2}{2}\log\left(1+\frac{2^{1-2\mathscr{H}(\varepsilon)-\overline{\varepsilon}^2}|\boldsymbol{h}_{1,1}|^2\gamma\left(1+\frac{|\boldsymbol{h}_{2,1}|^2\gamma}{2\overline{\varepsilon}}\right)^{2\overline{\varepsilon}\varepsilon}}{\overline{\varepsilon}\left(1+\frac{|\boldsymbol{h}_{2,1}|^2\gamma}{\overline{\varepsilon}}\right)^{2\overline{\varepsilon}\varepsilon}\left(1+\frac{2|\boldsymbol{h}_{2,1}|^2\gamma}{\overline{\varepsilon}}\right)^{\frac{\overline{\varepsilon}^2}{2}}}\right).\end{aligned} \tag{3.94}$$

It is seen that

$$\mathrm{MG}_{\mathrm{A}} = \frac{1}{2} - \frac{\varepsilon^2}{2} - (\overline{\varepsilon}\varepsilon)^2 - \frac{\overline{\varepsilon}^4}{4} \tag{3.95}$$

and

$$\mathrm{IEF}_{\mathrm{A}} = \overline{\varepsilon}\left(\mathscr{H}(\varepsilon) + \frac{\overline{\varepsilon}^2}{2}\right). \tag{3.96}$$

<u>Scheme B</u>- Let $K = 1$ and $\mathcal{A} = \{0, 1\}$. Noting that $\mathrm{supp}(\vec{\boldsymbol{s}}_1) = \mathrm{supp}(\vec{\boldsymbol{s}}_2) = \{0, 1\}$ and $G_1(1) = 0$,

$$\varrho_1(P; 1, \vec{\boldsymbol{h}}_1) = \frac{|\boldsymbol{h}_{1,1}|\gamma}{\overline{\varepsilon}\left(1+\frac{|\boldsymbol{h}_{2,1}|^2\gamma}{\overline{\varepsilon}}\right)^{\overline{\varepsilon}}}. \tag{3.97}$$

Evidently, $\mathrm{H}(\vec{\boldsymbol{s}}_1\vec{\boldsymbol{s}}_1^{\dagger}) = \mathscr{H}(\varepsilon)$. Therefore,

$$\mathscr{R}_1^{(\mathrm{lb})}(\vec{\boldsymbol{h}}_1) = \overline{\varepsilon}\log\left(1+\frac{2^{-\mathscr{H}(\varepsilon)}|\boldsymbol{h}_{1,1}|^2\gamma}{\overline{\varepsilon}\left(1+\frac{|\boldsymbol{h}_{2,1}|^2\gamma}{\overline{\varepsilon}}\right)^{\overline{\varepsilon}}}\right). \tag{3.98}$$

Also,

$$\mathrm{MG}_{\mathrm{B}} = \overline{\varepsilon}\varepsilon \tag{3.99}$$

and

$$\mathrm{IEF}_{\mathrm{B}} = \overline{\varepsilon}\mathscr{H}(\varepsilon). \tag{3.100}$$

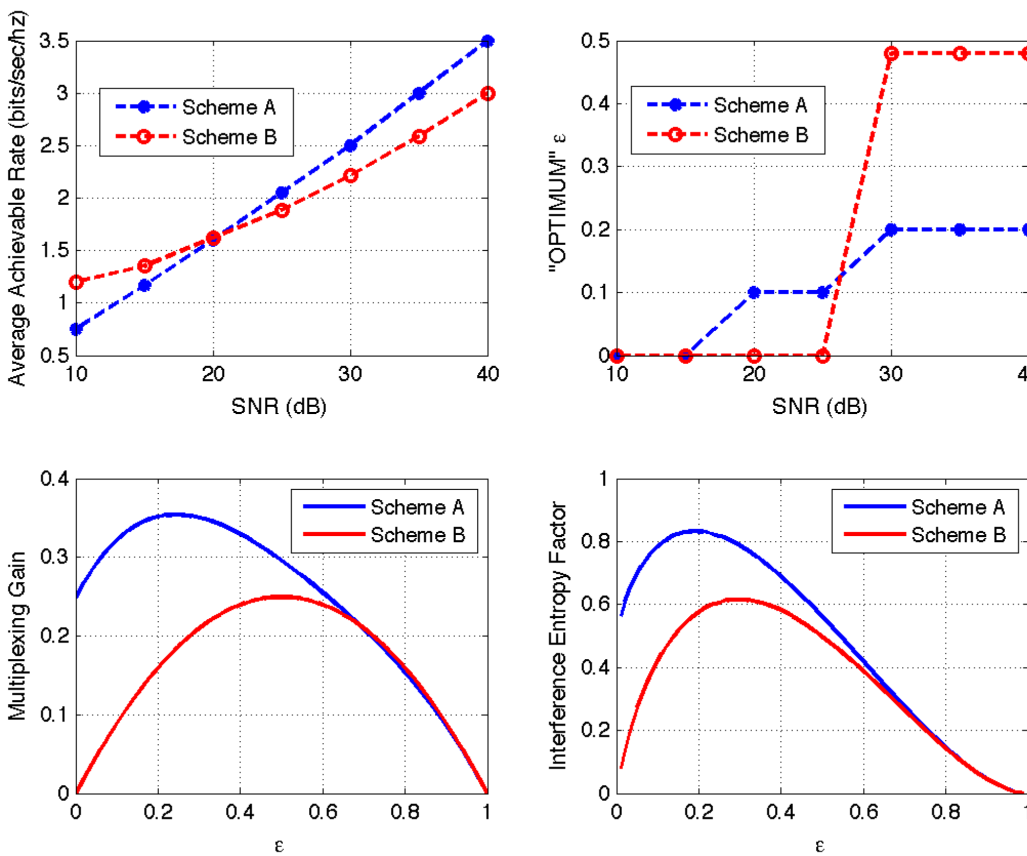

Figure 3.4: Comparison between $\sup_{\varepsilon,\nu\in(0,1]} \mathrm{E}\left\{\mathscr{R}_1^{(\mathrm{lb})}(\vec{\boldsymbol{h}}_1)\right\}$ in the schemes A and B for different SNR values.

Fig. 3.4 sketches $\sup_{\varepsilon,\nu\in(0,1]} \mathrm{E}\left\{\mathscr{R}_1^{(\mathrm{lb})}(\vec{\boldsymbol{h}}_1)\right\}$ for both schemes at different SNR values. This figure also demonstrates the best ε chosen by the users. It is seen that there is a tradeoff between the rates at medium and high SNR values. As discussed in remark 1, this happens as a result of the interplay between MG and IEF. This observation is based on the formulation for $\mathscr{R}_1^{(\mathrm{lb})}(\vec{h}_1)$ which is only a lower bound on $\mathscr{R}_1(\vec{h}_1)$. To show that this tradeoff is an intrinsic behavior of $\mathscr{R}_1(\vec{h}_1)$ itself, let the channel gains related to user 1 be given by $h_{1,1} = -0.3059 - 1.1777\sqrt{-1}$ and $h_{2,1} = 0.0886 + 0.2034\sqrt{-1}$. Assuming that user 1 utilizes the plots given in fig. 3.4 to regulate ε, fig. 3.5 sketches $\mathscr{R}_1(\vec{h}_1)$ (using Monte-Carlo simulation) for both schemes. It is seen that the tradeoff between the rate at medium and high SNR regimes is a fundamental property of decentralized networks.

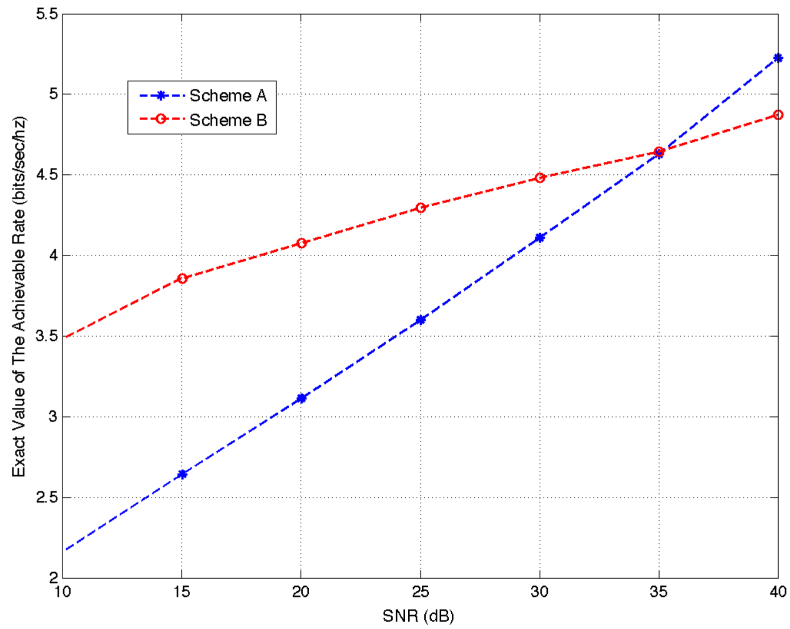

Figure 3.5: Plots of $\mathscr{R}_1(\vec{h}_1)$ (using Monte-Carlo simulation) for schemes A and B. The channel gains are $h_{1,1} = -0.3059 - 1.1777\sqrt{-1}$ and $h_{2,1} = 0.0886 + 0.2034\sqrt{-1}$.

Example 4- A Non-Uniform PMF on The Underlying Alphabet Leads To A Rise in Rate.

In this example, we let $\mathcal{A} = \{-1, 1\}$, $p_1 = \nu$ and $p_{-1} = \overline{\nu}$ for some $\nu \in (0, 1]$. This coincides with the random spreading technique used in CDMA systems [19, 20]. The purpose of this example is to show that in contrast to example 3 (where $N = 2$), the optimum value of ν is surprisingly not $\frac{1}{2}$ in a network with $N = 4$ users.

Before proceeding, let us consider a different and common approach to obtain a lower bound on $\mathscr{R}_1(\vec{h}_1)$. It is well-known that in an additive noise channel with a stationary and ergodic noise process, as far as the correlation function[6] of the noise process is fixed and SNR is sufficiently large, a Gaussian noise process yields the smallest mutual information between the input and output [77]. Using this fact, one can obtain a lower bound on $\mathrm{I}(\boldsymbol{x}_1; \vec{\boldsymbol{y}}_1 | \vec{\boldsymbol{s}}_1)$ as

$$\mathrm{I}(\boldsymbol{x}_1; \vec{\boldsymbol{y}}_1 | \vec{\boldsymbol{s}}_1) \geq \log \frac{\det \mathrm{Cov}(\vec{\boldsymbol{y}}_1)}{\det \mathrm{Cov}(\vec{\boldsymbol{w}}_1 + \vec{\boldsymbol{z}}_1)}. \tag{3.101}$$

[6] The correlation function of a zero-mean process $\mathbf{x}[t]$ is the function $\mathrm{E}\{\mathbf{x}[t_1]\mathbf{x}^\dagger[t_2]\}$ for $t_1, t_2 \in \mathbb{R}$.

It is easy to see that

$$\mathrm{Cov}(\vec{\boldsymbol{y}}_1) = I_K + \gamma \sum_{i=1}^{N} |\boldsymbol{h}_{i,1}|^2 \left(1-(2\nu-1)^2\right) I_K \tag{3.102}$$

and

$$\mathrm{Cov}(\vec{\boldsymbol{w}}_1 + \vec{\boldsymbol{z}}_1) = I_K + \gamma \sum_{i=2}^{N} |\boldsymbol{h}_{i,1}|^2 \left(1-(2\nu-1)^2\right) I_K. \tag{3.103}$$

Therefore,

$$\frac{\mathrm{I}(\boldsymbol{x}_1;\vec{\boldsymbol{y}}_1|\vec{\boldsymbol{s}}_1)}{K} \geq \log\left(1+\frac{\gamma|\boldsymbol{h}_{1,1}|^2(1-(2\nu-1)^2)}{1+\gamma\sum_{i=2}^{N}|\boldsymbol{h}_{i,1}|^2(1-(2\nu-1)^2)}\right) \tag{3.104}$$

One can see that this lower bound is maximized for $\nu = \frac{1}{2}$. Also, it is not sensitive to the value of K. Hence,

$$\sup_{\nu,K} \frac{\mathrm{I}(\boldsymbol{x}_1;\vec{\boldsymbol{y}}_1|\vec{\boldsymbol{s}}_1)}{K} \geq \log\left(1+\frac{\gamma|\boldsymbol{h}_{1,1}|^2}{1+\gamma\sum_{i=2}^{N}|\boldsymbol{h}_{i,1}|^2}\right). \tag{3.105}$$

Let us define

$$\begin{aligned}\tau_N &\triangleq \sup_{\gamma} \mathrm{E}\left\{\log\left(1+\frac{\gamma|\boldsymbol{h}_{1,1}|^2}{1+\gamma\sum_{i=2}^{N}|\boldsymbol{h}_{i,2}|^2}\right)\right\} \\ &= \frac{1}{(N-2)!}\int_{h\in\mathbb{R}}\int_{h'\in\mathbb{R}} h'^{N-2}\log\left(1+\frac{h}{h'}\right)e^{-h-h'}dhdh' \end{aligned}\tag{3.106}$$

where we have used the fact that $|\boldsymbol{h}_{1,1}|^2$ is an exponential random variable with parameter 1 and $2\sum_{i=2}^{N}|\boldsymbol{h}_{i,1}|^2$ is a $\chi^2_{2(N-1)}$ random variable. We call τ_N the Gaussian bounding threshold. It is the average with respect to channel gains of the bound given in (3.105) as SNR goes to infinity.

We demonstrate that if $N=4$, taking $K>1$ and $\nu \neq \frac{1}{2}$ makes $\mathrm{E}\left\{\mathscr{R}_1^{(\mathrm{lb})}(\vec{\boldsymbol{h}}_1)\right\}$ be considerably larger than τ_N at finite SNR values. This implies that there exists a set of nonzero measure for the realizations of $\vec{\boldsymbol{h}}_1$ such that $\mathscr{R}_1^{(\mathrm{lb})}(\vec{\boldsymbol{h}}_1)$ is larger than $\log\left(1+\frac{\gamma|\boldsymbol{h}_{1,1}|^2}{1+\gamma\sum_{i=2}^{N}|\boldsymbol{h}_{i,1}|^2}\right)$.

In order to calculate $\mathscr{R}_1^{(\mathrm{lb})}(\vec{\boldsymbol{h}}_1)$, one needs to find $\mathrm{H}\left(\boldsymbol{S}_i\Xi_i\Xi_i^{\dagger}\boldsymbol{S}_i^{\dagger}\right)$. In appendix K, it is shown that

$$\mathrm{H}\left(\boldsymbol{S}_i\Xi_i\Xi_i^{\dagger}\boldsymbol{S}_i^{\dagger}\right) = -3\sum_{k=0}^{K}\binom{K}{k}\left(\nu^{k+1}\overline{\nu}^{K-k}+\nu^{K-k}\overline{\nu}^{k+1}\right)\log\left(\nu^{k+1}\overline{\nu}^{K-k}+\nu^{K-k}\overline{\nu}^{k+1}\right). \tag{3.107}$$

In contrast to example 3, writing the expressions for the eigenvalues involved in the formula for $\mathscr{R}_1^{(\mathrm{lb})}(\vec{\boldsymbol{h}}_1)$ is a tedious task. As such, we finish the calculation of $\mathrm{E}\left\{\mathscr{R}_1^{(\mathrm{lb})}(\vec{\boldsymbol{h}}_1)\right\}$ by generating the complete list of these eigenvalues via computer simulation. Setting the SNR at $\gamma = 60$dB, fig. 3.6 sketches $\mathrm{E}\left\{\mathscr{R}_1^{(\mathrm{lb})}(\vec{\boldsymbol{h}}_1)\right\}$ in terms of $p_1 = \nu$ for different values of K. It is seen that the average

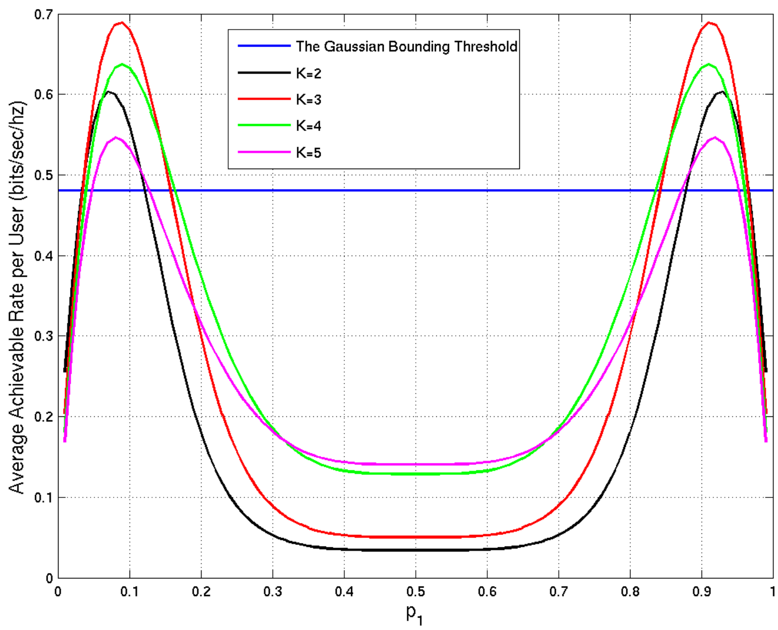

Figure 3.6: Plots of $\mathrm{E}\left\{\mathscr{R}_1^{(\mathrm{lb})}(\vec{\boldsymbol{h}}_1)\right\}$ in terms of $p_1 = \nu$ for different values of K. The value of SNR is 60dB.

achievable rate per user has two humps and is not maximized at $\nu = \frac{1}{2}$. The best performance is obtained for $K = 3$ and $\nu = 0.09$ or $\nu = 0.91$.

To show that $\mathscr{R}_1(\vec{h}_1)$ has the same property, i.e., it is not maximized for $\nu = \frac{1}{2}$, let us consider a special case where the channel gains related to user 1 are given by $h_{1,1} = 0.8409 - 0.0266\sqrt{-1}$, $h_{2,1} = -0.3059 - 1.1777\sqrt{-1}$, $h_{3,1} = -0.8107 + 0.8421\sqrt{-1}$ and $h_{4,1} = 0.0886 + 0.2034\sqrt{-1}$. Setting $K = 3$, fig. 3.7 presents the plot of $\mathscr{R}_1(\vec{h}_1)$ in terms of ν for SNR=60dB. The maximum of $\mathscr{R}_1(\vec{h}_1)$ is 1.577 bits/sec/hz that occurs for $\nu = 0.2$ or 0.8. We have also sketched the Gaussian lower bound $\log\left(1 + \frac{\gamma|h_{1,1}|^2}{1+\gamma(|h_{2,1}|^2+|h_{3,1}|^2+|h_{4,1}|^2)}\right)$ that is equal to 0.3155 bits/sec/hz.

The explanation for the existence of the humps is that for any K the MG and IEF factors[7] are both maximized at $\nu = \frac{1}{2}$. Since a larger MG yields a larger rate and a larger IEF leads to smaller values for the rate, the competition between these factors is settled at some ν that is not equal to $\frac{1}{2}$. $\square$

[7] See remark 1.

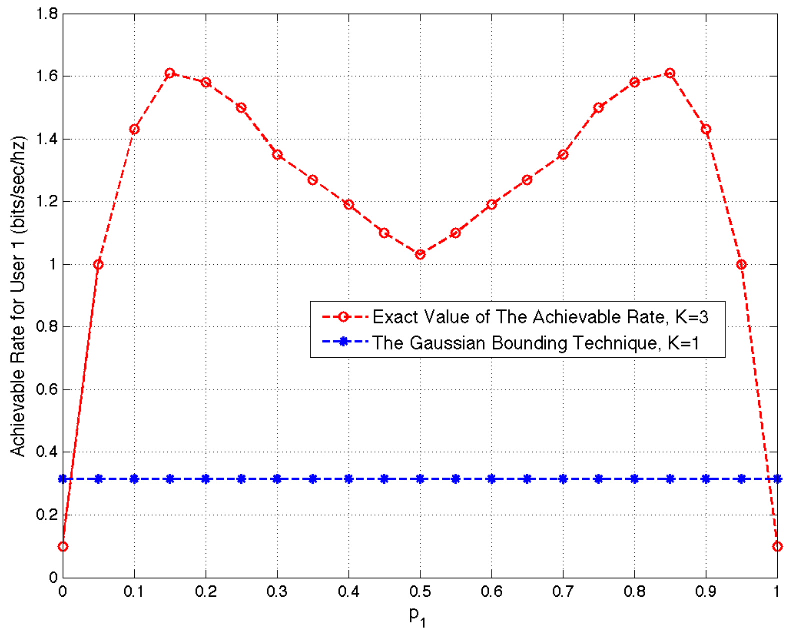

Figure 3.7: Plot of $\mathscr{R}_1(\vec{h}_1)$ in terms of $p_1 = \nu$ for $h_{1,1} = 0.8409 - 0.0266\sqrt{-1}$, $h_{2,1} = -0.3059 - 1.1777\sqrt{-1}$, $h_{3,1} = -0.8107 + 0.8421\sqrt{-1}$ and $h_{4,1} = 0.0886 + 0.2034\sqrt{-1}$. The value of SNR is 60dB and $K = 3$.

3.5 Optimality Results

In chapter 2, we studied a decentralized network of users sharing a number of frequency sub-bands where each user transmits independent Gaussian signals over a randomly selected frequency sub-bands and repeats this process independently from transmission slot to transmission slot. Using the time-frequency duality, this is equivalent to the randomized Time Hopping (TH) scenario where each user remains silent randomly from transmission slot to transmission slot so that other users have the opportunity of transmission without interference. Let the i^{th} user transmit the signals $\boldsymbol{u}_{i,1}, \boldsymbol{u}_{i,2}, \cdots$ consecutively during the slots that it is active. This is shown in fig. 3.8 where the transmission slots during which a user remains silent are shown in black.

Let each user quit transmitting its signal from transmission slot to transmission slot with a probability of $1 - \theta$. Dropping the index t in $\boldsymbol{u}_{i,t}$ for the sake of notational simplicity, the signal received at the receiver side of user 1 in a typical transmission slot can be written as

$$\boldsymbol{y}_1 = \theta^{-\frac{1}{2}} h_{1,1} \boldsymbol{c}_1 \boldsymbol{u}_1 + \theta^{-\frac{1}{2}} \sum_{i=2}^{N} h_{i,2} \boldsymbol{c}_i \boldsymbol{u}_i + \boldsymbol{z}_1 \tag{3.108}$$

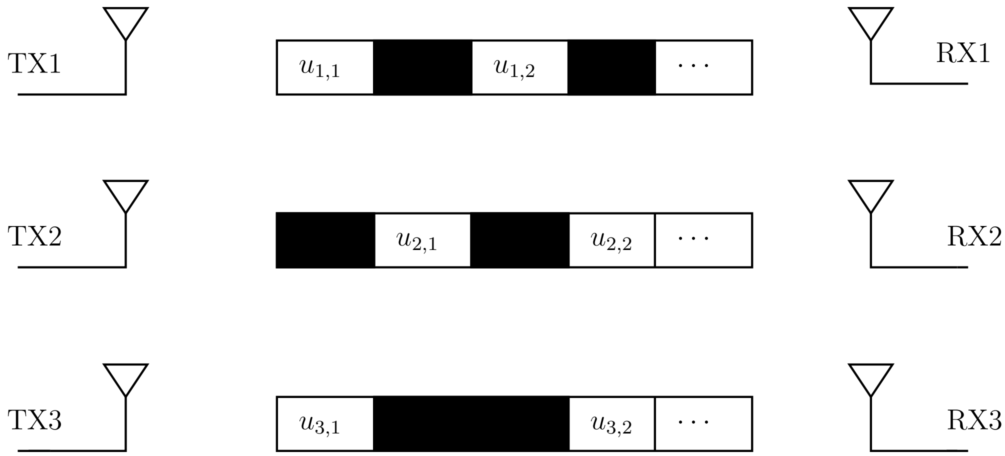

Figure 3.8: A typical decentralized network with $N = 3$ users based on randomized TH. Each user is independently and randomly active and silent with probabilities θ and $\overline{\theta}$ respectively.

where $\boldsymbol{c}_i$ is a Ber(θ) random variable that determines if the i^{th} user is active, i.e.,

$$\boldsymbol{c}_i = \begin{cases} 1 & \text{The } i^{\text{th}} \text{ user is active.} \\ 0 & \text{Oth.} \end{cases} . \tag{3.109}$$

Each $\boldsymbol{u}_i$ satisfies the power constraint $\mathrm{E}\{|\boldsymbol{u}_i|^2\} \leq \gamma$, and the factors $\theta^{-\frac{1}{2}}$ in (3.108) are incorporated to guarantee an average transmission power of less than or equal to γ per user.

For any $1 \leq i \leq N$, if the i^{th} user transmits i.i.d. signals $(\boldsymbol{u}_{i,t})_{t \geq 0}$ with PDF $p_{\boldsymbol{u}}(.)$, the highest achievable rate for user 1 is given by

$$\mathscr{C}_1(\vec{h}_1; \theta) \triangleq \sup_{\substack{\boldsymbol{u}_1, \cdots, \boldsymbol{u}_N \text{ are i.i.d with PDF } p_{\boldsymbol{u}}(.) \\ \mathrm{E}\{|\boldsymbol{u}_1|^2\} \leq \gamma}} \mathrm{I}(\boldsymbol{u}_1, \boldsymbol{c}_1; \boldsymbol{y}_1). \tag{3.110}$$

In case $N = 2$, one might think that selecting $\theta = \frac{1}{2}$, the highest achievable rate one can guarantee for user 1 is $\mathscr{C}_1(\vec{h}_1; \frac{1}{2})$. It turns out that this is not the case.

The main result of this section is the following Theorem.

Theorem 3. *In a decentralized network of $N = 2$ users operating based on randomized TH, for any $h_{1,1}, h_{2,1} \in \mathbb{C}$ and $\theta \in (0, 1)$, user 1 can achieve rates that are larger than $\mathscr{C}_1(\vec{h}_1; \theta)$ for sufficiently large values of γ. Equivalently, transmission of* i.i.d. *signals in consecutive transmission slots that the users are active is suboptimal in the high SNR regime.*

To prove Theorem 3, we need the following Lemma.

Lemma 8. *Let $\mathbf{z}_1$ and $\mathbf{z}_2$ be circularly symmetric complex Gaussian random variables with variances σ_1^2 and σ_2^2 respectively and $\mathbf{x}$ be independent of $(\mathbf{z}_1, \mathbf{z}_2)$. Then, the answer to the optimization problem*

$$\sup_{\mathbf{x}:\mathrm{E}\{|\mathbf{x}|^2\} \leq \sigma^2} \mathrm{h}(\mathbf{x} + \mathbf{z}_1) - \xi \mathrm{h}(\mathbf{x} + \mathbf{z}_2)$$

is a circularly symmetric complex Gaussian $\mathbf{x}$ for any σ^2 and any $\xi \geq 1$. Also, if $\sigma_1^2 \leq \sigma_2^2$, the same conclusion holds for any $\xi \in \mathbb{R}$.

Proof. This is a direct consequence of Theorem 1 in [75]. □

Our strategy is to find an upper bound on $\mathscr{C}_1(\vec{h}_1; \theta)$ for arbitrary $\theta \in (0, 1]$ and propose an achievable rate that is larger than this upper bound. Our main tool is the signaling proposed in this chapter based on random spreading. Throughout the rest of this section, we assume $N = 2$.

3.5.1 An Upper Bound on $\mathscr{C}_1(\vec{h}_1;\theta)$

One can write

$$
\begin{aligned}
\mathrm{I}(\boldsymbol{u}_1,\boldsymbol{c}_1;\boldsymbol{y}_1) &= \mathrm{I}(\boldsymbol{u}_1;\boldsymbol{y}_1|\boldsymbol{c}_1)+\mathrm{I}(\boldsymbol{c}_1;\boldsymbol{y}_1)\\
&\le \mathrm{I}(\boldsymbol{u}_1;\boldsymbol{y}_1|\boldsymbol{c}_1)+\mathrm{H}(\boldsymbol{c}_1)\\
&= \mathrm{I}(\boldsymbol{u}_1;\theta^{-\frac{1}{2}}h_{1,1}\boldsymbol{c}_1\boldsymbol{u}_1+h_{2,1}\boldsymbol{c}_2\boldsymbol{u}_2+\boldsymbol{z}_1|\boldsymbol{c}_1=1)\Pr\{\boldsymbol{c}_1=1\}\\
&\quad+\mathrm{I}(\boldsymbol{u}_1;\theta^{-\frac{1}{2}}h_{1,1}\boldsymbol{c}_1\boldsymbol{u}_1+\theta^{-\frac{1}{2}}h_{2,1}\boldsymbol{c}_2\boldsymbol{u}_2+\boldsymbol{z}_1|\boldsymbol{c}_1=0)\Pr\{\boldsymbol{c}_1=0\}+\mathscr{H}(\theta)\\
&\overset{(a)}{=} \theta\mathrm{I}(\boldsymbol{u}_1;\theta^{-\frac{1}{2}}h_{1,1}\boldsymbol{u}_1+\theta^{-\frac{1}{2}}h_{2,1}\boldsymbol{c}_2\boldsymbol{u}_2+\boldsymbol{z}_1)+\mathscr{H}(\theta)\\
&\overset{(b)}{\le} \theta\mathrm{I}(\boldsymbol{u}_1;\theta^{-\frac{1}{2}}h_{1,1}\boldsymbol{u}_1+\theta^{-\frac{1}{2}}h_{2,1}\boldsymbol{c}_2\boldsymbol{u}_2+\boldsymbol{z}_1|\boldsymbol{c}_2)+\mathscr{H}(\theta)\\
&= \theta\mathrm{I}(\boldsymbol{u}_1;\theta^{-\frac{1}{2}}h_{1,1}\boldsymbol{u}_1+\theta^{-\frac{1}{2}}h_{2,1}\boldsymbol{c}_2\boldsymbol{u}_2+\boldsymbol{z}_1|\boldsymbol{c}_2=0)\Pr\{\boldsymbol{c}_2=0\}\\
&\quad+\theta\mathrm{I}(\boldsymbol{u}_1;\theta^{-\frac{1}{2}}h_{1,1}\boldsymbol{u}_1+\theta^{-\frac{1}{2}}h_{2,1}\boldsymbol{c}_2\boldsymbol{u}_2+\boldsymbol{z}_1|\boldsymbol{c}_2=1)\Pr\{\boldsymbol{c}_2=1\}+\mathscr{H}(\theta)\\
&= \theta\overline{\theta}\,\mathrm{I}(\boldsymbol{u}_1;\theta^{-\frac{1}{2}}h_{1,1}\boldsymbol{u}_1+\boldsymbol{z}_1)+\theta^2\mathrm{I}(\boldsymbol{u}_1;\theta^{-\frac{1}{2}}h_{1,1}\boldsymbol{u}_1+\theta^{-\frac{1}{2}}h_{2,1}\boldsymbol{u}_2+\boldsymbol{z}_1)+\mathscr{H}(\theta)\\
&= \theta\overline{\theta}\Big(\mathrm{h}(\theta^{-\frac{1}{2}}h_{1,1}\boldsymbol{u}_1+\boldsymbol{z}_1)-\mathrm{h}(\boldsymbol{z}_1)\Big)\\
&\quad+\theta^2\Big(\mathrm{h}(\theta^{-\frac{1}{2}}h_{1,1}\boldsymbol{u}_1+\theta^{-\frac{1}{2}}h_{2,1}\boldsymbol{u}_2+\boldsymbol{z}_1)-\mathrm{h}(\theta^{-\frac{1}{2}}h_{2,1}\boldsymbol{u}_2+\boldsymbol{z}_1)\Big)+\mathscr{H}(\theta)\\
&\overset{(c)}{=} \theta\overline{\theta}\left(\mathrm{h}(\theta^{-\frac{1}{2}}h_{1,1}\boldsymbol{u}_1+\boldsymbol{z}_1)-\frac{\theta}{\overline{\theta}}\ \mathrm{h}(\theta^{-\frac{1}{2}}h_{2,1}\boldsymbol{u}_1+\boldsymbol{z}_1)\right)\\
&\quad+\theta^2\mathrm{h}(\theta^{-\frac{1}{2}}h_{1,1}\boldsymbol{u}_1+\theta^{-\frac{1}{2}}h_{2,1}\boldsymbol{u}_2+\boldsymbol{z}_1)-\theta\overline{\theta}\log(\pi e)+\mathscr{H}(\theta)\\
&\overset{(d)}{=} \theta\overline{\theta}\left(\mathrm{h}(\boldsymbol{u}_1+\boldsymbol{z}_1')-\frac{\theta}{\overline{\theta}}\ \mathrm{h}(\boldsymbol{u}_1+\boldsymbol{z}_1'')\right)+\theta^2\mathrm{h}(\theta^{-\frac{1}{2}}h_{1,1}\boldsymbol{u}_1+\theta^{-\frac{1}{2}}h_{2,1}\boldsymbol{u}_2+\boldsymbol{z}_1)\\
&\quad+\theta\overline{\theta}\log(\theta^{-1}|h_{1,1}|^2)-\theta^2\log(\theta^{-1}|h_{2,1}|^2)-\theta\overline{\theta}\log(\pi e)+\mathscr{H}(\theta)
\end{aligned}
\tag{3.111}
$$

where (a) follows due to $\mathrm{I}(\boldsymbol{u}_1;h_{1,1}\boldsymbol{c}_1\boldsymbol{u}_1+h_{2,1}\boldsymbol{c}_2\boldsymbol{u}_2+\boldsymbol{z}_1|\boldsymbol{c}_1=0)=0$, (b) is by the fact that the mutual information between the input and output of the channel increases if a "genie" provides the receiver side of user 1 with $\boldsymbol{c}_2$, (c) follows by the fact that $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ are identically distributed and $\mathrm{h}(\boldsymbol{z}_1)=\log(\pi e)$ and finally (d) follows by the fact that for any complex random variable $\mathbf{x}$ and $a\in\mathbb{C}$, we have $\mathrm{h}(a\mathbf{x})=\mathrm{h}(\mathbf{x})+\log|a|^2$. Also, $\boldsymbol{z}_1'$ and $\boldsymbol{z}_1''$ are $\mathcal{CN}\left(0,\frac{\theta}{|h_{1,1}|^2}\right)$ and $\mathcal{CN}\left(0,\frac{\theta}{|h_{2,1}|^2}\right)$ random variables in the last equality in (3.111), respectively. By (3.110) and (3.111), we conclude that

$$
\begin{aligned}
\mathscr{C}_1(\vec{h}_1;\theta) \le\ & \theta\overline{\theta}\sup_{\substack{\boldsymbol{u}_1,\boldsymbol{u}_2\ \text{are i.i.d}\\ \mathrm{E}\{|\boldsymbol{u}_1|^2\}\le\gamma}}\left(\mathrm{h}(\boldsymbol{u}_1+\boldsymbol{z}_1')-\frac{\theta}{\overline{\theta}}\ \mathrm{h}(\boldsymbol{u}_1+\boldsymbol{z}_1'')\right)\\
&+\theta^2\sup_{\substack{\boldsymbol{u}_1,\boldsymbol{u}_2\ \text{are i.i.d}\\ \mathrm{E}\{|\boldsymbol{u}_1|^2\}\le\gamma}}\mathrm{h}(\theta^{-\frac{1}{2}}h_{1,1}\boldsymbol{u}_1+\theta^{-\frac{1}{2}}h_{2,1}\boldsymbol{u}_2+\boldsymbol{z}_1)\\
&+\theta\overline{\theta}\log(\theta^{-1}|h_{1,1}|^2)-\theta^2\log(\theta^{-1}|h_{2,1}|^2)-\theta\overline{\theta}\log(\pi e).
\end{aligned}
\tag{3.112}
$$

By the Maximum Entropy Lemma[80], it is trivial that

$$\sup_{\substack{\boldsymbol{u}_1,\boldsymbol{u}_2 \text{ are i.i.d} \\ \mathrm{E}\{|\boldsymbol{u}_1|^2\}\leq\gamma}} \mathrm{h}(\theta^{-\frac{1}{2}}h_{1,1}\boldsymbol{u}_1+\theta^{-\frac{1}{2}}h_{2,1}\boldsymbol{u}_2+\boldsymbol{z}_1)=\log\left(\pi e\theta^{-1}\left(|h_{1,1}|^2+|h_{2,1}|^2\right)\gamma+1\right) \tag{3.113}$$

Invoking Lemma 8, if $\frac{\theta}{\overline{\theta}}\geq 1$ or $|h_{1,1}|>|h_{2,1}|$, or equivalently, $\theta\geq\frac{1}{2}$ or $|h_{1,1}|>|h_{2,1}|$, the answer to the optimization $\max_{\substack{\boldsymbol{u}_1,\boldsymbol{u}_2 \text{ are i.i.d} \\ \mathrm{E}\{|\boldsymbol{u}_1|^2\}\leq\gamma}}\left(\mathrm{h}(\boldsymbol{u}_1+\boldsymbol{z}'_1)-\frac{\theta}{\overline{\theta}}\,\mathrm{h}(\boldsymbol{u}_1+\boldsymbol{z}''_1)\right)$ is a complex Gaussian $\boldsymbol{u}_1$ whose variance is possibly less than γ. Let the optimum $\boldsymbol{u}_1$ for this optimization problem be a $\mathcal{CN}(0,\sigma^2)$ random variable. We distinguish the following cases.

Case 1- If $\theta\geq\frac{1}{2}$ and $\frac{|h_{1,1}|}{|h_{2,1}|}<\left(\frac{\theta}{\overline{\theta}}\right)^{\frac{1}{2}}$, then $\sigma^2=0$.

Case 2- If $\theta>\frac{1}{2}$, $\frac{|h_{1,1}|}{|h_{2,1}|}>\left(\frac{\theta}{\overline{\theta}}\right)^{\frac{1}{2}}$ and $\gamma>\frac{\theta^2\overline{\theta}}{2\theta-1}\left(\frac{1}{|h_{2,1}|^2}-\frac{\theta}{\overline{\theta}|h_{1,1}|^2}\right)$, then $\sigma^2=\frac{\theta\overline{\theta}}{2\theta-1}\left(\frac{1}{|h_{2,1}|^2}-\frac{\theta}{\overline{\theta}|h_{1,1}|^2}\right)$.

Case 3- If $\theta\leq\frac{1}{2}$ and $\frac{|h_{1,1}|}{|h_{2,1}|}>1$, then $\sigma^2=\frac{\gamma}{\theta}$.

Verification of these cases is a straightforward task which is omitted for the sake of brevity. Therefore, as far as $\theta\geq\frac{1}{2}$, the term $\sup_{\substack{\boldsymbol{u}_1,\boldsymbol{u}_2 \text{ are i.i.d} \\ \mathrm{E}\{|\boldsymbol{u}_1|^2\}\leq\gamma}}\left(\mathrm{h}(\boldsymbol{u}_1+\boldsymbol{z}'_1)-\frac{\theta}{\overline{\theta}}\,\mathrm{h}(\boldsymbol{u}_1+\boldsymbol{z}''_1)\right)$ saturates by increasing γ. Using this fact together with (3.112) and (3.113),

$$\mathscr{C}_1(\vec{h}_1;\theta)\overset{\gamma}{\lesssim}\theta^2\log\gamma. \tag{3.114}$$

as far as $\theta\geq\frac{1}{2}$ and for any $h_{1,1},h_{2,1}\in\mathbb{C}$.

On the other hand, if $\theta<\frac{1}{2}$ and $\frac{|h_{1,1}|}{|h_{2,1}|}>1$,

$$\sup_{\substack{\boldsymbol{u}_1,\boldsymbol{u}_2 \text{ are i.i.d} \\ \mathrm{E}\{|\boldsymbol{u}_1|^2\}\leq\gamma}}\left(\mathrm{h}(\boldsymbol{u}_1+\boldsymbol{z}'_1)-\frac{\theta}{\overline{\theta}}\,\mathrm{h}(\boldsymbol{u}_1+\boldsymbol{z}''_1)\right)\overset{\gamma}{\sim}\frac{\overline{\theta}-\theta}{\overline{\theta}}\log\gamma. \tag{3.115}$$

Using this together with (3.112) and (3.113),

$$\mathscr{C}_1(\vec{h}_1;\theta)\overset{\gamma}{\lesssim}\theta\overline{\theta}\log\gamma \tag{3.116}$$

as far as $\theta<\frac{1}{2}$ and $\frac{|h_{1,1}|}{|h_{2,1}|}<1$. However, we can remove the condition $\frac{|h_{1,1}|}{|h_{2,1}|}>1$ by noting that $\mathscr{C}_1(\vec{h}_1;\theta)$ is an increasing function of $|h_{1,1}|$. Hence, (3.116) holds for all $\theta<\frac{1}{2}$ regardless of the values of $h_{1,1}$ and $h_{2,1}$.

To recap,

$$\mathscr{C}_1(\vec{h}_1;\theta)\overset{\gamma}{\lesssim}\max\{\theta^2,\theta\overline{\theta}\}\log\gamma. \tag{3.117}$$

3.5.2 Achieving Rates Larger Than $\mathscr{C}_1(\vec{h}_1;\theta)$

A variant approach for signaling by any user is to not transmit i.i.d. signals in consecutive activity intervals (transmission slots where the user is not silent). Let each user apply the signaling scheme presented in this chapter where the signature vectors $\vec{\boldsymbol{s}}_1$ and $\vec{\boldsymbol{s}}_2$ have length $K=2$ and are generated according to the PMF $(q_a)_{a\in\mathcal{A}}$ over an alphabet $\mathcal{A}$ such that $0\notin\mathcal{A}$. Since the output of the transmitter for each user is automatically nulled independently from transmission slot to transmission slot, one can think of the signature vectors at the output of the switch for user 1 and user 2 as vectors $\widetilde{\vec{\boldsymbol{s}}}_1$ and $\widetilde{\vec{\boldsymbol{s}}}_2$ being constructed over the augmented alphabet $\widetilde{\mathcal{A}}=\mathcal{A}\cup\{0\}$ where based on our previous notation $q_0=\varepsilon=\overline{\theta}$. By Proposition 5, the achievable rate of user 1 scales like

$$\mathscr{R}_1(\vec{h}_1)\overset{\gamma}{\sim}\frac{1}{2}\Pr\left\{\widetilde{\vec{\boldsymbol{s}}}_1\notin\operatorname{span}\{\widetilde{\vec{\boldsymbol{s}}}_2\}\right\}. \tag{3.118}$$

We are interested in values of θ so that $\mathscr{C}_1(\vec{h}_1;\theta)\overset{\gamma}{\lesssim}\mathscr{R}_1(\vec{h}_1)$ is satisfied strictly. By (3.117) and (3.118), it is sufficient to explore the values of θ in the set $\mathcal{I}_A$ defined by

$$\mathcal{I}_A\triangleq\left\{\theta\in(0,1]:\frac{1}{2}\Pr\left\{\widetilde{\vec{\boldsymbol{s}}}_1\notin\operatorname{span}\{\widetilde{\vec{\boldsymbol{s}}}_2\}\right\}>\max\{\theta^2,\theta\overline{\theta}\}\right\}. \tag{3.119}$$

Let $\mathcal{A}=\{-1,1\}$ and $q_1=q_{-1}=\frac{1}{2}$. By (3.46) in example 2, one can see that

$$\Pr\left\{\widetilde{\vec{\boldsymbol{s}}}_1\notin\operatorname{span}\{\widetilde{\vec{\boldsymbol{s}}}_2\}\right\}=1-\overline{\theta}^2-2(\theta\overline{\theta})^2-\frac{\theta^4}{2}. \tag{3.120}$$

By (3.119), we require that

$$1-\overline{\theta}^2-2(\theta\overline{\theta})^2-\frac{\theta^4}{2}>2\max\{\theta^2,\theta\overline{\theta}\}. \tag{3.121}$$

Solving this inequality yields

$$\mathcal{I}_{\{-1,1\}}=\in(0.31,0.566). \tag{3.122}$$

Next, we show through an example that for any two alphabets $\mathcal{A}_1$ and $\mathcal{A}_2$ with the property $\mathcal{A}_1\subsetneq\mathcal{A}_2$, we have $\mathcal{I}_{\mathcal{A}_1}\subsetneq\mathcal{I}_{\mathcal{A}_2}$.

Let $\mathcal{A}=\{-2,-1,1,2\}$ and $q_{-2}=q_{-1}=q_1=q_2=\frac{1}{4}$. Then, the elements of $\widetilde{\vec{\boldsymbol{s}}}_1$ and $\widetilde{\vec{\boldsymbol{s}}}_2$ are i.i.d. random variables taking the values 0, -2, -1, 1 and 2 with probabilities $\overline{\theta}$, $\frac{\theta}{4}$, $\frac{\theta}{4}$, $\frac{\theta}{4}$ and $\frac{\theta}{4}$ respectively. The event $\widetilde{\vec{\boldsymbol{s}}}_1\in\operatorname{span}\{\widetilde{\vec{\boldsymbol{s}}}_2\}$ occurs if and only if $\widetilde{\vec{\boldsymbol{s}}}_1=\vec{0}_2$ or $\widetilde{\vec{\boldsymbol{s}}}_1\neq\vec{0}_2$ and $\widetilde{\vec{\boldsymbol{s}}}_1=\pm\widetilde{\vec{\boldsymbol{s}}}_2$ or $\widetilde{\vec{\boldsymbol{s}}}_1=\pm2\widetilde{\vec{\boldsymbol{s}}}_2$ or $\widetilde{\vec{\boldsymbol{s}}}_1=\pm\frac{1}{2}\widetilde{\vec{\boldsymbol{s}}}_2$. It is easy to see that

$$\Pr\left\{\widetilde{\vec{\boldsymbol{s}}}_1\notin\operatorname{span}\{\widetilde{\vec{\boldsymbol{s}}}_2\}\right\}=1-\overline{\theta}^2-2(\theta\overline{\theta})^2-\frac{3\theta^4}{16}. \tag{3.123}$$

By (3.119), characterization of $\mathcal{I}_{\{-2,-1,1,2\}}$ requires

$$1 - \bar{\theta}^2 - 2(\theta\bar{\theta})^2 - \frac{3\theta^4}{16} > 2\max\{\theta^2, \theta\bar{\theta}\}. \tag{3.124}$$

Solving this inequality yields

$$\mathcal{I}_{\{-2,-1,1,2\}} = (0.297, 0.588). \tag{3.125}$$

Therefore, $\mathcal{I}_{\{-1,1\}} \subsetneq \mathcal{I}_{\{-2,-1,1,2\}}$ as desired. This completes the proof of Theorem 3, because by expanding the alphabet, one can cover the whole interval $(0,1)$ for θ.

Chapter 4

Block-Asynchronous Decentralized Networks

4.1 System Model

We consider an interference channel with two users of separate transmitter-receiver pairs. The network is assumed to be decentralized, i.e., there is no explicit cooperation among users. In particular, users are not aware of each other's code-books. This implies that interference cancellation is not possible and users treat each other's signal as noise. The channel from transmitter i to receiver j is modeled by a static and non-frequency selective gain $h_{i,j}$. For simplicity of analysis, we assume the channel from each transmitter to each receiver is slotted and the symbol intervals on any of the four channels from different transmitters to different receivers coincide. Therefore, the two users are symbol-synchronized, i.e., they are synchronous at the symbol level. Both users utilize codewords of length T, i.e., any codeword has T symbols and each symbol is transmitted in one symbol interval. User 1 (similar notation is applied for user 2) chooses its message in a set of size $2^{\lfloor TR \rfloor}$ where R is the transmission rate[1] in bits/symbol. User 1 assigns a codeword $(\boldsymbol{x}_{1,t})_{t=0}^{T-1} = (\boldsymbol{x}_{1,0}, \cdots, \boldsymbol{x}_{1,T-1})$ consisting of zero-mean random variables[2] to each of its messages. Thereafter, $(\boldsymbol{x}_{1,t})_{t=0}^{T-1}$ is transmitted in T consecutive symbol intervals called a transmission block. We assume

$$\frac{1}{T}\sum_{t=0}^{T-1} \mathrm{E}\left\{|\boldsymbol{x}_{1,t}|^2\right\} \leq \gamma, \tag{4.1}$$

i.e., the average transmission power per code symbol is less than or equal to γ.

[1] For fairness, we assume the transmission rates of both users are the same.

[2] We define $\boldsymbol{x}_{1,t} = 0$ for $t \geq T$ and $t < 0$. Also, the PDF of $\boldsymbol{x}_{1,t}$ for any $0 \leq t \leq T-1$ is not Gaussian in general.

We assume the rate of data arrival for both users is such that they may not have enough data in their buffers and hence, the transmitted codewords by any user look like intermittent bursts along the time axis. Whether a user is transmitting a codeword or not, a number of k bits arrive in its buffer with a probability of q or no new data is received with a probability of $1-q$ during each symbol interval[3]. As soon as the number of bits in the buffer exceeds or is equal to $\lfloor TR \rfloor$, a codeword representing $\lfloor TR \rfloor$ bits is transmitted. Afterwards, the user remains silent until, once again, enough data is accumulated in its queue. It is shown in [70] that the probability of data loss with a finite buffer size can be made arbitrarily small if and only if $R > kq$. Users are block-asynchronous in the sense that there exists a delay between their transmitted codewords. This is due to the fact that the data streams of the two users are independent, i.e., if one user has enough data for transmission, the other user may still need to wait to receive more data in its buffer. Another reason is the mismatch between the activation instants of the users. In practice, one user may become active before the other user joins the network.

Noting that the codewords are of equal length, the transmitted codeword by any user can at most overlap with two transmitted codewords of the other user. For simplicity of analysis and in order to get insight into the behavior of the system, we assume that the rate of data arrival is low enough to ensure each transmitted codeword by any user overlaps with at most one transmitted codeword by the other user. This can be guaranteed if the gap between any two consecutive transmitted codewords by each user is larger than or equal to T symbol intervals[4]. This certainly holds if $R \geq 2k$. For $R < 2k$ and as far as $\frac{R}{k}$ is a rational number, we are able to find a range for q such that for sufficiently large T, the gap between any two consecutive codewords is greater than or equal to T almost surely. The following Proposition states our result.

Proposition 7. *Let $\frac{R}{k} = \frac{a}{b} < 2$ where $a, b \in \mathbb{N}$ and the greatest common divisor of a and b is 1. For sufficiently large codeword length T, the gap between any two consecutive codewords is greater than or equal to T if*

$$\sum_{m=1}^{\infty} \left(1 - I_{1-q}\left((2b-a)m - 1, am\right)\right) < \infty \tag{4.2}$$

where

$$I_c(c', c'') \triangleq \frac{\int_0^c \tau^{c'-1}(1-\tau)^{c''-1} d\tau}{\int_0^1 \tau^{c'-1}(1-\tau)^{c''-1} d\tau} \tag{4.3}$$

is the regularized incomplete beta function [86] for $c', c'' \in \mathbb{N}$ and $c \in (0, 1)$.

[3]The same stochastic model for the input data streams is considered in [70] and references therein.

[4]It is straightforward to extend our analysis to include the general case. To fulfill this purpose, one must compute the distribution of the gap between any two consecutive transmitted codewords by each user. This distribution depends on the parameters k, q, T and R.

Proof. Let $\lfloor TR \rfloor$ be divisible by k, i.e., $\lfloor TR \rfloor = nk$ for some $n \in \mathbb{N}$. As soon as the number of bits in the queue of user 1 is $\lfloor TR \rfloor$, a codeword is sent in T symbol intervals. Let A_T be the event that the number of bits received in the queue of user 1 during $2T-2$ consecutive symbol intervals is greater than or equal to $\lfloor TR \rfloor$. In this case, the number of symbol intervals that user 1 remains silent and waits for enough bits in its queue in order to transmit the next codeword is less than or equal to $T-1$. Therefore, a codeword transmitted by user 2 may overlap with these two consecutively transmitted codewords by user 1. We have

$$\begin{aligned}\Pr\{A_T\} &= \sum_{l=n}^{2T-2} \binom{2T-2}{l} q^l (1-q)^{2T-2-l} \\ &= 1 - \sum_{l=0}^{n-1} \binom{2T-2}{l} q^l (1-q)^{2T-2-l} \\ &= 1 - I_{1-q}(2T-n-1, n) \end{aligned} \tag{4.4}$$

where the last step is by the properties of the regularized incomplete beta function [86] defined in (4.3).

For fixed R and k, let $T_1, T_2, T_3, \cdots$ be an increasing sequence of values for T such that $\lfloor T_m R \rfloor$ is divisible by k for $m \geq 1$. If R, k and q are such that

$$\sum_{m=1}^{\infty} \Pr\{A_{T_m}\} = \sum_{m=1}^{\infty} \left(1 - I_{1-q}\left(2T_m - \frac{\lfloor T_m R \rfloor}{k} - 1, \frac{\lfloor T_m R \rfloor}{k}\right)\right) < \infty, \tag{4.5}$$

we conclude by the Borel-Cantelli Lemma [89] that there is $m_0 \in \mathbb{N}$ such that $\Pr\{A_{T_m}\} = 0$ for $m \geq m_0$. Assuming $\frac{R}{k} = \frac{a}{b}$, we let $T_m = bm$ in (4.5). The result of the Proposition is immediate. $\square$

As an example, let $\frac{R}{k} = \frac{4}{5}$. By (4.2), we are interested in values of q such that

$$\sum_{m=1}^{\infty} (1 - I_{1-q}(6m-1, 4m)) < \infty. \tag{4.6}$$

Simulation results show that (4.6) holds for any $q < 0.4$ and does not hold for any $q \geq 0.4$. Therefore, if $\frac{R}{k} = \frac{4}{5}$ and $q < 0.4$, our assumption is valid for sufficiently large T.

Let $\boldsymbol{x}_{1,0}$ be transmitted at time instant $t = 0$. The signal received at the receiver side of user 1 at time instant $t \geq 0$ is

$$\boldsymbol{y}_1[t] = h_{1,1}\boldsymbol{x}_{1,t} + h_{2,1}\boldsymbol{x}_{2,t-\boldsymbol{d}_T} + \boldsymbol{z}_1[t] \tag{4.7}$$

where $\boldsymbol{d}_T \in \mathbb{Z}$ is a random integer representing the asynchrony between the users and $\boldsymbol{z}_1[t] \sim \mathcal{CN}(0,1)$ is the ambient noise[5] at the receiver side of user 1 at time instant t. Since the ambient

[5] Similarly, $\boldsymbol{z}_2[t] \sim \mathcal{CN}(0,1)$.

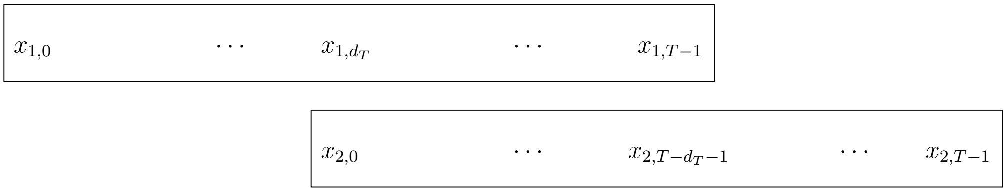

Figure 4.1: There is a mutual delay d_T between the transmitted codewords of the two users. Users are assumed to be synchronous at the symbol level.

noise has unit variance, the parameter γ is a measure of SNR throughout the chapter. It is seen that if $|\boldsymbol{d}_T| \geq T$, there is no interference between the two users. On the other hand, if $|\boldsymbol{d}_T| \leq T-1$, the codewords of the two users overlap.

In order to model the asynchrony, we define an auxiliary random variable $\boldsymbol{\alpha}$ with support $\text{supp}(\boldsymbol{\alpha}) = \left[0, \frac{1}{2}\right]$ called the asynchrony random variable whose PDF, $p_{\boldsymbol{\alpha}}(.)$, is globally known to all users. This random variable is used to describe the distribution of $\boldsymbol{d}_T$. Given a realization α of $\boldsymbol{\alpha}$, we assume $\boldsymbol{d}_T$ is uniformly distributed over the set of integers in the interval $[1-T, T-1]$. Let us define

$$\lambda_T(\alpha) \triangleq \Pr\left\{\boldsymbol{d}_T = t \middle| \boldsymbol{\alpha} = \alpha\right\} \tag{4.8}$$

for all $t \in [1-T, T-1] \cap \mathbb{Z}$. Since $\Pr\{|\boldsymbol{d}_T| \geq T\} = 1 - (2T-1)\lambda_T(\alpha) \geq 0$, we have $\lambda_T(\alpha) \leq \frac{1}{2T-1}$. This yields

$$\lim_{T\to\infty} \lambda_T(\alpha) = 0. \tag{4.9}$$

Moreover, we assume $\lim_{T\to\infty} T\lambda_T(\alpha)$ exists and is equal to α, i.e.,

$$\lim_{T\to\infty} T\lambda_T(\alpha) = \alpha. \tag{4.10}$$

Since $\lim_{T\to\infty} T\lambda_T(\alpha) \leq \lim_{T\to\infty} \frac{T}{2T-1} = \frac{1}{2}$, the assumption made in (4.10) is consistent with the definition of the support of $\boldsymbol{\alpha}$. If $\alpha = 0$, the probability that there is no overlap between the codewords of the users tends to 1 by increasing T. On the other hand, if $\alpha = \frac{1}{2}$, the codewords overlap (partially or completely) with a probability that approaches 1 as T tends to infinity.

Let us denote the interference at the receiver side of user 1 caused by user 2 at time instant t by

$$\boldsymbol{w}_1[t] \triangleq h_{2,1}\boldsymbol{x}_{2,t-\boldsymbol{d}_T}. \tag{4.11}$$

From the viewpoint of receiver 1, if the knowledge about the realization of $\boldsymbol{d}_T$ is not available, the sequence $(\boldsymbol{w}_1[t])_{t=0}^{T-1}$ is not an ergodic process in the sense that the Entropy Ergodic Theorem [87] does not hold for $(\boldsymbol{w}_1[t])_{t=0}^{T-1}$, i.e., $-\frac{\log p_{\boldsymbol{w}_1[0],\cdots,\boldsymbol{w}_1[T-1]}(\boldsymbol{w}_1[0],\cdots,\boldsymbol{w}_1[T-1])}{T}$ does not converge (with probability 1) to the entropy rate of $(\boldsymbol{w}_1[t])_{t=0}^{T-1}$ as T tends to infinity. This is true even if $(\boldsymbol{x}_{2,t})_{t=0}^{T-1}$

is entropy ergodic. We assume the receivers are able to measure (without error) the realization of $\boldsymbol{d}_T$ by the end of the transmission block (consisting of T symbol intervals) and perform the decoding procedure afterwards. On the other hand, even if the receivers are aware of $\boldsymbol{d}_T$, due to the fact that the transmitters are unaware of the realizations of $\boldsymbol{\alpha}$ and $\boldsymbol{d}_T$, the channel is non-ergodic from the transmitters' standpoint[6]. In this setup, a commonly used computational tool to assess the system performance is the probability of outage which is the subject of the next section.

4.2 Outage Analysis

In this section we assume the channel gains $h_{i,j}$ are realizations of i.i.d. random variables $\boldsymbol{h}_{i,j}$ representing Rayleigh fading, i.e., $\boldsymbol{h}_{i,j} \sim \mathcal{CN}(0,1)$. As mentioned before, transmitters are unaware of the realizations of $\boldsymbol{\alpha}$ and $\boldsymbol{d}_T$. We assume the transmitters are unaware of the realizations of the channel gains as well. The receivers are assumed to be capable of estimating $\boldsymbol{d}_T$ and the channel gains. In fact, the decoders need to know the true amount of delay and the values of channel gains for successful maximum likelihood or joint typicality decoding.

Let $\vec{\boldsymbol{h}}_1 \triangleq \begin{pmatrix} \boldsymbol{h}_{1,1} & \boldsymbol{h}_{2,1} \end{pmatrix}^{\mathrm{t}}$ be the vector containing the channel gains related to user 1. The outage event for user 1 is defined by

$$\mathcal{O}_{1,T} \triangleq \left\{ \vec{\boldsymbol{h}}_1, \boldsymbol{d}_T, \boldsymbol{\alpha} : \mathscr{R}_{1,T} < R \right\} \tag{4.12}$$

where

$$\mathscr{R}_{1,T} \triangleq \sup \frac{\mathrm{I}\left((\boldsymbol{x}_{1,t})_{t=0}^{T-1}; (\boldsymbol{y}_{1,t})_{t=0}^{T-1} \right)}{T} \tag{4.13}$$

is the highest achievable rate of user 1 as if this user was aware of $\vec{\boldsymbol{h}}_1, \boldsymbol{d}_T$ and $\boldsymbol{\alpha}$. The sup is taken over the set of all joint PDFs for $(\boldsymbol{x}_{1,t})_{t=0}^{T-1}$ that satisfy the Entropy Ergodic Theorem. We note that the mutual information $\mathrm{I}\left((\boldsymbol{x}_{1,t})_{t=0}^{T-1}; (\boldsymbol{y}_{1,t})_{t=0}^{T-1} \right)$ is a function of the random quantities $\vec{\boldsymbol{h}}_1$, $\boldsymbol{d}_T$ and $\boldsymbol{\alpha}$, i.e., these quantities are treated as parameters in order to calculate $\mathrm{I}\left((\boldsymbol{x}_{1,t})_{t=0}^{T-1}; (\boldsymbol{y}_{1,t})_{t=0}^{T-1} \right)$. Hence, $\mathscr{R}_{1,T}$ is a function of $\vec{\boldsymbol{h}}_1$, $\boldsymbol{d}_T$ and $\boldsymbol{\alpha}$ and therefore, it is a random variable. In general, finding $\mathscr{R}_{1,T}$ is an open problem. In fact, the optimum joint distribution of $(\boldsymbol{x}_{1,t})_{t=0}^{T-1}$ is unknown. Our strategy is to assume a particular choice for the joint PDF of $(\boldsymbol{x}_{1,t})_{t=0}^{T-1}$ that yields a lower bound $\mathscr{R}_{1,T}^{(\mathrm{lb})}$ on $\mathscr{R}_{1,T}$. These constitute the materials to be offered in this section.

Let us define the event $\mathcal{O}_{1,T}^{(\mathrm{lb})}$ by

$$\mathcal{O}_{1,T}^{(\mathrm{lb})} \triangleq \left\{ \vec{\boldsymbol{h}}_1, \boldsymbol{d}_T, \boldsymbol{\alpha} : \mathscr{R}_{1,T}^{(\mathrm{lb})} < R \right\}. \tag{4.14}$$

[6]In the context of compound channels[88], the achievable rate of each user corresponds to the worst case where both users always overlap completely.

Noting that $\Pr\{\mathcal{O}_{1,T}\} \leq \Pr\left\{\mathcal{O}_{1,T}^{(\mathrm{lb})}\right\}$, we develop an upper bound on the probability of outage by computing $\lim_{T\to\infty}\Pr\left\{\mathcal{O}_{1,T}^{(\mathrm{lb})}\right\}$ in the limit as T tends to infinity. In part 3 of the current section, invoking the celebrated Lebesgue Dominated Convergence Theorem (LDCT) [89] in Proposition 8, we find an expression for $\lim_{T\to\infty}\Pr\left\{\mathcal{O}_{1,T}^{(\mathrm{lb})}\right\}$. The main problem that rises at this point is that computing $\Pr\left\{\mathcal{O}_{1,T}^{(\mathrm{lb})}\right\}$ is a difficult task. Therefore, one can not compute $\Pr\left\{\mathcal{O}_{1,T}^{(\mathrm{lb})}\right\}$ in closed form and find its limit directly as T tends to infinity. A way around this difficulty is to write $\Pr\left\{\mathcal{O}_{1,T}^{(\mathrm{lb})}\right\}$ as $\mathrm{E}\left\{\mathbb{I}\left\{\mathscr{R}_{1,T}^{(\mathrm{lb})} \leq R\right\}\right\}$. Hence, if one can show that $\lim_{T\to\infty}\mathbb{I}\left\{\mathscr{R}_{1,T}^{(\mathrm{lb})} \leq R\right\}$ exists almost surely, then LDCT can be applied to yield $\lim_{T\to\infty}\Pr\left\{\mathcal{O}_{1,T}^{(\mathrm{lb})}\right\} = \mathrm{E}\left\{\lim_{T\to\infty}\mathbb{I}\left\{\mathscr{R}_{1,T}^{(\mathrm{lb})} \leq R\right\}\right\}$.

Next, we introduce our signaling scheme called Randomized Masking (RM).

4.2.1 Randomized Masking

Following the randomized resource allocation strategies introduced in the previous chapters, we assume each transmitter transmits a Gaussian symbol in each symbol interval with a probability of $\theta \in (0,1]$ and remains silent with a probability of $\overline{\theta}$. This process repeats independently from symbol interval to symbol interval. This scheme is called Randomized Masking with activity factor θ. In fact, $(\boldsymbol{x}_{1,t})_{t=0}^{T-1}$ is an i.i.d. sequence where the PDF of $\boldsymbol{x}_{1,t}$ for any $0 \leq t \leq T-1$ is given by $\overline{\theta}\delta(.) + \theta \mathrm{g}_1^{(\mathrm{c})}\left(., \frac{\gamma}{\theta}\right)$. The codeword of user 1 can be written as

$$(\boldsymbol{x}_{1,0}, \cdots, \boldsymbol{x}_{1,T-1}) = (\beta\boldsymbol{c}_{1,0}\boldsymbol{s}_{1,0}, \cdots, \beta\boldsymbol{c}_{1,T-1}\boldsymbol{s}_{1,T-1}) \tag{4.15}$$

where $(\boldsymbol{c}_{1,t})_{t=0}^{T-1}$ are independent $\mathrm{Ber}(\theta)$ random variables, $(\boldsymbol{s}_{1,t})_{t=0}^{T-1}$ are independent Gaussian random variables with variance γ, and $\beta = \frac{1}{\sqrt{\theta}}$ is a scaling factor that is included in order to guarantee (4.1). Note that taking $\theta = 1$ is equivalent to continuous transmission of Gaussian signals.

By (4.7), the signal received at the receiver side of user 1 at time instant $t \geq 0$ can be written as

$$\boldsymbol{y}_1[t] = \beta h_{1,1}\boldsymbol{c}_{1,t}\boldsymbol{s}_{1,t} + \beta h_{2,1}\boldsymbol{c}_{2,t-\boldsymbol{d}_T}\boldsymbol{s}_{2,t-\boldsymbol{d}_T} + \boldsymbol{z}_1[t]. \tag{4.16}$$

Since user 1 is unaware of the on-off pattern of user 2, the noise plus interference at receiver 1 has a mixed Gaussian PDF. As such, the mutual information (assuming user 1 is aware of the realizations of delay and channel gains) has no closed formulation. Invoking entropy power inequality [80] and developing a new upper bound on the differential entropy of a complex and circularly symmetric mixed Gaussian random variable, we obtain an achievable rate for user 1 as if this user knew the realizations $\vec{h}_1 = \begin{pmatrix} h_{1,1} & h_{2,1} \end{pmatrix}^{\mathrm{t}}$, d_T and α.

4.2.2 Developing $\mathscr{R}_{1,T}^{(\mathrm{lb})}$

Let us define

$$\boldsymbol{\mathcal{T}} \triangleq \{0 \le t \le T-1 : \boldsymbol{c}_{1,t} = 1\} \tag{4.17}$$

and

$$\mathcal{S} \triangleq \{d_T, d_T+1, \cdots, d_T+T-1\}. \tag{4.18}$$

The random set $\boldsymbol{\mathcal{T}}$ contains the time instants that masking is not performed by user 1, while $\mathcal{S}$ is the set of time instants that the codeword of user 2 may overlap[7] with the codeword of user 1. We can write

$$\begin{aligned}
\mathrm{I}\left((\boldsymbol{x}_{1,t})_{t=0}^{T-1}; (\boldsymbol{y}_1[t])_{t=0}^{T-1}\right) &= \mathrm{I}\left((\boldsymbol{c}_{1,t}\boldsymbol{s}_{1,t})_{t=0}^{T-1}; (\boldsymbol{y}_1[t])_{t=0}^{T-1}\right) \\
&\overset{(a)}{=} \mathrm{I}\left((\boldsymbol{s}_{1,t})_{t\in\boldsymbol{\mathcal{T}}}, \boldsymbol{\mathcal{T}}; (\boldsymbol{y}_1[t])_{t=0}^{T-1}\right) \\
&= \mathrm{I}\left((\boldsymbol{s}_{1,t})_{t\in\boldsymbol{\mathcal{T}}}, \boldsymbol{\mathcal{T}}; (\boldsymbol{y}_1[t])_{t\in\boldsymbol{\mathcal{T}}}\right) \\
&\quad + \mathrm{I}\left((\boldsymbol{s}_{1,t})_{t\in\boldsymbol{\mathcal{T}}}, \boldsymbol{\mathcal{T}}; (\boldsymbol{y}_1[t])_{t\notin\boldsymbol{\mathcal{T}}} \middle| (\boldsymbol{y}_1[t])_{t\in\boldsymbol{\mathcal{T}}}\right) \\
&\overset{(b)}{\ge} \mathrm{I}\left((\boldsymbol{s}_{1,t})_{t\in\boldsymbol{\mathcal{T}}}, \boldsymbol{\mathcal{T}}; (\boldsymbol{y}_1[t])_{t\in\boldsymbol{\mathcal{T}}}\right) \\
&= \mathrm{I}\left(\boldsymbol{\mathcal{T}}; (\boldsymbol{y}_1[t])_{t\in\boldsymbol{\mathcal{T}}}\right) + \mathrm{I}\left((\boldsymbol{s}_{1,t})_{t\in\boldsymbol{\mathcal{T}}}; (\boldsymbol{y}_1[t])_{t\in\boldsymbol{\mathcal{T}}} \middle| \boldsymbol{\mathcal{T}}\right) \\
&\overset{(c)}{\ge} \mathrm{I}\left((\boldsymbol{s}_{1,t})_{t\in\boldsymbol{\mathcal{T}}}; (\boldsymbol{y}_1[t])_{t\in\boldsymbol{\mathcal{T}}} \middle| \boldsymbol{\mathcal{T}}\right) \\
&\overset{(d)}{=} \mathrm{I}\left((\boldsymbol{s}_{1,t})_{t\in\boldsymbol{\mathcal{T}}\setminus\mathcal{S}}; (\boldsymbol{y}_1[t])_{t\in\boldsymbol{\mathcal{T}}\setminus\mathcal{S}} \middle| \boldsymbol{\mathcal{T}}\right) \\
&\quad + \mathrm{I}\left((\boldsymbol{s}_{1,t})_{t\in\boldsymbol{\mathcal{T}}\cap\mathcal{S}}; (\boldsymbol{y}_1[t])_{t\in\boldsymbol{\mathcal{T}}\cap\mathcal{S}} \middle| \boldsymbol{\mathcal{T}}\right)
\end{aligned} \tag{4.19}$$

where (a) holds since the two collections $(\boldsymbol{c}_{1,t}\boldsymbol{s}_{1,t})_{t=0}^{T-1}$ and $((\boldsymbol{s}_{1,t})_{t\in\boldsymbol{\mathcal{T}}}, \boldsymbol{\mathcal{T}})$ are equivalent, i.e., one yields the other, (b) and (c) are due to $\mathrm{I}\left((\boldsymbol{s}_{1,t})_{t\in\boldsymbol{\mathcal{T}}}, \boldsymbol{\mathcal{T}}; (\boldsymbol{y}_1[t])_{t\notin\boldsymbol{\mathcal{T}}} \middle| (\boldsymbol{y}_1[t])_{t\in\boldsymbol{\mathcal{T}}}\right) \ge 0$ and $\mathrm{I}\left(\boldsymbol{\mathcal{T}}; (\boldsymbol{y}_1[t])_{t\in\boldsymbol{\mathcal{T}}}\right) \ge 0$, respectively, and finally, (d) follows from the fact that for each realization $\mathcal{T}$ of $\boldsymbol{\mathcal{T}}$, the collection of random variables $\left((\boldsymbol{s}_{1,t})_{t\in\mathcal{T}\setminus\mathcal{S}}, (\boldsymbol{y}_1[t])_{t\in\mathcal{T}\setminus\mathcal{S}}\right)$ is independent of the collection $((\boldsymbol{s}_{1,t})_{t\in\boldsymbol{\mathcal{T}}\cap\mathcal{S}}, (\boldsymbol{y}_1[t])_{t\in\boldsymbol{\mathcal{T}}\cap\mathcal{S}})$.

We compute the two terms on the right side of (4.19) separately.

[7] Overlap happens if and only if $|d_T| \le T-1$.

Computation of $\mathrm{I}\left((\boldsymbol{s}_{1,t})_{t\in\boldsymbol{\mathcal{T}}\setminus\mathcal{S}};(\boldsymbol{y}_1[t])_{t\in\boldsymbol{\mathcal{T}}\setminus\mathcal{S}}\middle|\boldsymbol{\mathcal{T}}\right)$

We have

$$\begin{aligned}
\mathrm{I}\left((\boldsymbol{s}_{1,t})_{t\in\boldsymbol{\mathcal{T}}\setminus\mathcal{S}};(\boldsymbol{y}_1[t])_{t\in\boldsymbol{\mathcal{T}}\setminus\mathcal{S}}\middle|\boldsymbol{\mathcal{T}}\right) &= \sum_{\mathcal{T}\in\mathrm{supp}(\boldsymbol{\mathcal{T}})}\Pr\{\boldsymbol{\mathcal{T}}=\mathcal{T}\}\,\mathrm{I}\left((\boldsymbol{s}_{1,t})_{t\in\boldsymbol{\mathcal{T}}\setminus\mathcal{S}};(\boldsymbol{y}_1[t])_{t\in\boldsymbol{\mathcal{T}}\setminus\mathcal{S}}\middle|\boldsymbol{\mathcal{T}}=\mathcal{T}\right)\\
&\overset{(a)}{=} \sum_{\mathcal{T}\in\mathrm{supp}(\boldsymbol{\mathcal{T}})}\Pr\{\boldsymbol{\mathcal{T}}=\mathcal{T}\}\,\mathrm{I}\left((\boldsymbol{s}_{1,t})_{t\in\mathcal{T}\setminus\mathcal{S}};(\boldsymbol{y}_1[t])_{t\in\mathcal{T}\setminus\mathcal{S}}\right)\\
&\overset{(b)}{=} \sum_{\mathcal{T}\in\mathrm{supp}(\boldsymbol{\mathcal{T}})}\Pr\{\boldsymbol{\mathcal{T}}=\mathcal{T}\}\,|\mathcal{T}\setminus\mathcal{S}|\log\left(1+\frac{|h_{1,1}|^2\gamma}{\theta}\right)\\
&= \mathrm{E}\{|\boldsymbol{\mathcal{T}}\setminus\mathcal{S}|\}\,u(\vec{h}_1) \qquad (4.20)
\end{aligned}$$

where

$$u(\vec{h}_1)\triangleq\log\left(1+\frac{|h_{1,1}|^2\gamma}{\theta}\right). \qquad (4.21)$$

In (4.20), (a) holds by independence of $\boldsymbol{\mathcal{T}}$ and the collection $\left((\boldsymbol{s}_{1,t})_{t\in\mathcal{T}\setminus\mathcal{S}},(\boldsymbol{y}_1[t])_{t\in\mathcal{T}\setminus\mathcal{S}}\right)$ for any $\mathcal{T}\in\mathrm{supp}(\boldsymbol{\mathcal{T}})$. Also, (b) is due to the fact that for any $\mathcal{T}\in\mathrm{supp}(\boldsymbol{\mathcal{T}})$,

$$(\boldsymbol{y}_1[t])_{t\in\mathcal{T}\setminus\mathcal{S}}=(\beta h_{1,1}\boldsymbol{s}_{1,t})_{t\in\mathcal{T}\setminus\mathcal{S}}+(\boldsymbol{z}_1[t])_{t\in\mathcal{T}\setminus\mathcal{S}}. \qquad (4.22)$$

Finally, it is easy to verify that

$$\mathrm{E}\{|\boldsymbol{\mathcal{T}}\setminus\mathcal{S}|\}=\begin{cases}\theta T & |d_T|\geq T\\ \theta|d_T| & |d_T|\leq T-1\end{cases}. \qquad (4.23)$$

Computation of $\mathrm{I}\left((\boldsymbol{s}_{1,t})_{t\in\boldsymbol{\mathcal{T}}\cap\mathcal{S}};(\boldsymbol{y}_1[t])_{t\in\boldsymbol{\mathcal{T}}\cap\mathcal{S}}\middle|\boldsymbol{\mathcal{T}}\right)$

We have

$$\begin{aligned}
\mathrm{I}\left((\boldsymbol{s}_{1,t})_{t\in\boldsymbol{\mathcal{T}}\cap\mathcal{S}};(\boldsymbol{y}_1[t])_{t\in\boldsymbol{\mathcal{T}}\cap\mathcal{S}}\middle|\boldsymbol{\mathcal{T}}\right) &= \sum_{\mathcal{T}\in\mathrm{supp}(\boldsymbol{\mathcal{T}})}\Pr\{\boldsymbol{\mathcal{T}}=\mathcal{T}\}\,\mathrm{I}\left((\boldsymbol{s}_{1,t})_{t\in\boldsymbol{\mathcal{T}}\cap\mathcal{S}};(\boldsymbol{y}_1[t])_{t\in\boldsymbol{\mathcal{T}}\cap\mathcal{S}}\middle|\boldsymbol{\mathcal{T}}=\mathcal{T}\right)\\
&= \sum_{\mathcal{T}\in\mathrm{supp}(\boldsymbol{\mathcal{T}})}\Pr\{\boldsymbol{\mathcal{T}}=\mathcal{T}\}\,\mathrm{I}\left((\boldsymbol{s}_{1,t})_{t\in\mathcal{T}\cap\mathcal{S}};(\boldsymbol{y}_1[t])_{t\in\mathcal{T}\cap\mathcal{S}}\right).
\end{aligned} \qquad (4.24)$$

For any $\mathcal{T}\in\mathrm{supp}(\boldsymbol{\mathcal{T}})$,

$$(\boldsymbol{y}_1[t])_{t\in\mathcal{T}\cap\mathcal{S}}=(\beta h_{1,1}\boldsymbol{s}_{1,t})_{t\in\mathcal{T}\cap\mathcal{S}}+(\boldsymbol{w}_1[t]+\boldsymbol{z}_1[t])_{t\in\mathcal{T}\cap\mathcal{S}}. \qquad (4.25)$$

This yields

$$\mathrm{I}\left((\boldsymbol{s}_{1,t})_{t\in\mathcal{T}\cap\mathcal{S}};(\boldsymbol{y}_1[t])_{t\in\mathcal{T}\cap\mathcal{S}}\right)=\mathrm{h}\left((\boldsymbol{y}_1[t])_{t\in\mathcal{T}\cap\mathcal{S}}\right)-\mathrm{h}\left((\boldsymbol{w}_1[t]+\boldsymbol{z}_1[t])_{t\in\mathcal{T}\cap\mathcal{S}}\right). \qquad (4.26)$$

Applying the entropy power inequality[8] to (4.25),

$$2^{\frac{1}{|\mathcal{T}\cap\mathcal{S}|}\mathrm{h}((\boldsymbol{y}_1[t])_{t\in\mathcal{T}\cap\mathcal{S}})} \geq 2^{\frac{1}{|\mathcal{T}\cap\mathcal{S}|}\mathrm{h}((\beta h_{1,1}\boldsymbol{s}_{1,t})_{t\in\mathcal{T}\cap\mathcal{S}})} + 2^{\frac{1}{|\mathcal{T}\cap\mathcal{S}|}\mathrm{h}((\boldsymbol{w}_1[t]+\boldsymbol{z}_1[t])_{t\in\mathcal{T}\cap\mathcal{S}})}. \tag{4.27}$$

Dividing both sides by $2^{\frac{1}{|\mathcal{T}\cap\mathcal{S}|}\mathrm{h}((\boldsymbol{w}_1[t]+\boldsymbol{z}_1[t])_{t\in\mathcal{T}\cap\mathcal{S}})}$,

$$\begin{aligned}
& \mathrm{h}\left((\boldsymbol{y}_1[t])_{t\in\mathcal{T}\cap\mathcal{S}}\right) - \mathrm{h}\left((\boldsymbol{w}_1[t]+\boldsymbol{z}_1[t])_{t\in\mathcal{T}\cap\mathcal{S}}\right) \\
\geq\ & |\mathcal{T}\cap\mathcal{S}|\log\left(2^{\frac{1}{|\mathcal{T}\cap\mathcal{S}|}(\mathrm{h}((\beta h_{1,1}\boldsymbol{s}_{1,t})_{t\in\mathcal{T}\cap\mathcal{S}})-\mathrm{h}((\boldsymbol{w}_1[t]+\boldsymbol{z}_1[t])_{t\in\mathcal{T}\cap\mathcal{S}}))}+1\right) \\
=\ & |\mathcal{T}\cap\mathcal{S}|\log\left(2^{\frac{1}{|\mathcal{T}\cap\mathcal{S}|}\left(|\mathcal{T}\cap\mathcal{S}|\log\left(\frac{\pi e|h_{1,1}|^2\gamma}{\theta}\right)-\mathrm{h}((\boldsymbol{w}_1[t]+\boldsymbol{z}_1[t])_{t\in\mathcal{T}\cap\mathcal{S}})\right)}+1\right)
\end{aligned} \tag{4.28}$$

where the last step follows by $\mathrm{h}\left((\beta h_{1,1}\boldsymbol{s}_{1,t})_{t\in\mathcal{T}\cap\mathcal{S}}\right) = |\mathcal{T}\cap\mathcal{S}|\log\frac{\pi e|h_{1,1}|^2\gamma}{\theta}$. However, $(\boldsymbol{w}_1[t]+\boldsymbol{z}_1[t])_{t\in\mathcal{T}\cap\mathcal{S}}$ is a mixed Gaussian sequence. As a result, its differential entropy does not have a closed expression. Using the following Lemma, we obtain an upper bound on this quantity.

Lemma 9. *Let* x *be a circularly symmetric and complex mixed Gaussian random variable with* PDF

$$p_{\mathsf{x}}(\mathrm{x}) = p_1\mathrm{g}_1^{(\mathrm{c})}(\mathrm{x};\sigma_1^2) + p_2\mathrm{g}_1^{(\mathrm{c})}(\mathrm{x},\sigma_2^2) \tag{4.29}$$

where $\sigma_1^2 < \sigma_2^2$ *and* p_1 *and* p_2 *are positive numbers such that* $p_1+p_2=1$. *Then,*

$$\begin{aligned}
\mathrm{h}(\mathsf{x}) \leq\ & -p_1\log p_1 - p_2\log p_2 + p_1\log\left(\pi e\sigma_1^2\right) + p_2\log\left(\pi e\sigma_2^2\right) \\
& -p_1\log\left(1+\frac{p_2}{p_1}\frac{\sigma_1^2}{\sigma_2^2}\right) - \frac{p_2\sigma_1^2}{\sigma_2^2}\log\left(1+\frac{p_1}{p_2}\frac{\sigma_2^2}{\sigma_1^2}\right).
\end{aligned} \tag{4.30}$$

Proof. This is a special case of the bound developed in appendix A for a mixed Gaussian random variable with two variance levels. □

Since the elements of the sequence $(\boldsymbol{w}_1[t]+\boldsymbol{z}_1[t])_{t\in\mathcal{T}\cap\mathcal{S}}$ are i.i.d.,

$$\mathrm{h}\left((\boldsymbol{w}_1[t]+\boldsymbol{z}_1[t])_{t\in\mathcal{T}\cap\mathcal{S}}\right) = |\mathcal{T}\cap\mathcal{S}|\mathrm{h}\left(\boldsymbol{w}_1[t_0]+\boldsymbol{z}_1[t_0]\right) \tag{4.31}$$

where t_0 is an arbitrary element of $\mathcal{T}\cap\mathcal{S}$. The PDF of $\boldsymbol{w}_1[t_0]+\boldsymbol{z}_1[t_0]$ is a mixed Gaussian PDF consisting of two Gaussian PDFs with variances 1 and $1+\frac{|h_{2,1}|^2\gamma}{\theta}$ and corresponding probabilities $\overline{\theta}$ and θ, respectively. Applying Lemma 9,

$$\begin{aligned}
\mathrm{h}\left(\boldsymbol{w}_1[t_0]+\boldsymbol{z}_1[t_0]\right) \leq\ & \log(\pi e) + \mathscr{H}(\theta) + \theta\log\left(1+\frac{|h_{2,1}|^2\gamma}{\theta}\right) \\
& -\overline{\theta}\log\left(1+\frac{\theta}{\overline{\theta}}\frac{1}{1+\frac{|h_{2,1}|^2\gamma}{\theta}}\right) - \frac{\theta}{1+\frac{|h_{2,1}|^2\gamma}{\theta}}\log\left(1+\frac{\overline{\theta}}{\theta}\left(1+\frac{|h_{2,1}|^2\gamma}{\theta}\right)\right).
\end{aligned} \tag{4.32}$$

[8] We note that all sequences are complex, i.e., they have $2|\mathcal{T}\cap\mathcal{S}|$ real dimensions.

As such,

$$\begin{aligned}\mathrm{h}\left((\boldsymbol{w}_1[t]+\boldsymbol{z}_1[t])_{t\in\mathcal{T}\cap\mathcal{S}}\right) \quad \le \quad & |\mathcal{T}\cap\mathcal{S}|\left(\log(\pi e)+\mathscr{H}(\theta)+\theta\log\left(1+\frac{|h_{2,1}|^2\gamma}{\theta}\right)\right)\\ & -|\mathcal{T}\cap\mathcal{S}|\overline{\theta}\log\left(1+\frac{\theta}{\overline{\theta}}\frac{1}{1+\frac{|h_{2,1}|^2\gamma}{\theta}}\right)\\ & -\frac{|\mathcal{T}\cap\mathcal{S}|\theta}{1+\frac{|h_{2,1}|^2\gamma}{\theta}}\log\left(1+\frac{\overline{\theta}}{\theta}\left(1+\frac{|h_{2,1}|^2\gamma}{\theta}\right)\right). \end{aligned} \tag{4.33}$$

Substituting this in (4.28),

$$\mathrm{h}\left((\boldsymbol{y}_1[t])_{t\in\mathcal{T}\cap\mathcal{S}}\right)-\mathrm{h}\left((\boldsymbol{w}_1[t]+\boldsymbol{z}_1[t])_{t\in\mathcal{T}\cap\mathcal{S}}\right)\ge|\mathcal{T}\cap\mathcal{S}|v(\vec{h}_1) \tag{4.34}$$

where

$$v(\vec{h}_1)\triangleq\log\left(\frac{2^{-\mathscr{H}(\theta)}|h_{1,1}|^2\gamma\left(1+\frac{\theta}{\overline{\theta}}\frac{1}{1+\frac{|h_{2,1}|^2\gamma}{\theta}}\right)^{\overline{\theta}}\left(1+\frac{\overline{\theta}}{\theta}\left(1+\frac{|h_{2,1}|^2\gamma}{\theta}\right)\right)^{\frac{\theta}{1+\frac{|h_{2,1}|^2\gamma}{\theta}}}}{\theta\left(1+\frac{|h_{2,1}|^2\gamma}{\theta}\right)^{\theta}}+1\right). \tag{4.35}$$

By (4.24) and (4.34),

$$\mathrm{I}\left((\boldsymbol{s}_{1,t})_{t\in\boldsymbol{\mathcal{T}}\cap\mathcal{S}};(\boldsymbol{y}_1[t])_{t\in\boldsymbol{\mathcal{T}}\cap\mathcal{S}}\middle|\boldsymbol{\mathcal{T}}\right)\ge\mathrm{E}\left\{|\boldsymbol{\mathcal{T}}\cap\mathcal{S}|\right\}v(\vec{h}_1). \tag{4.36}$$

Finally, $\mathrm{E}\left\{|\boldsymbol{\mathcal{T}}\cap\mathcal{S}|\right\}$ can be calculated as

$$\mathrm{E}\left\{|\boldsymbol{\mathcal{T}}\cap\mathcal{S}|\right\}=\begin{cases}0 & |d_T|\ge T\\ \theta(T-|d_T|) & |d_T|\le T-1\end{cases}. \tag{4.37}$$

Note that $\mathrm{I}\left((\boldsymbol{s}_{1,t})_{t\in\boldsymbol{\mathcal{T}}\backslash\mathcal{S}};(\boldsymbol{y}_1[t])_{t\in\boldsymbol{\mathcal{T}}\backslash\mathcal{S}}\middle|\boldsymbol{\mathcal{T}}\right)$ is computed in closed form, however, we have only relied on a lower bound on $\mathrm{I}\left((\boldsymbol{s}_{1,t})_{t\in\boldsymbol{\mathcal{T}}\cap\mathcal{S}};(\boldsymbol{y}_1[t])_{t\in\boldsymbol{\mathcal{T}}\cap\mathcal{S}}\middle|\boldsymbol{\mathcal{T}}\right)$.

According to (4.19), (4.20) and (4.36),

$$\mathrm{I}\left((\boldsymbol{x}_{1,t})_{t=0}^{T-1};(\boldsymbol{y}_1[t])_{t=0}^{T-1}\right)\ge\mathrm{E}\left\{|\boldsymbol{\mathcal{T}}\backslash\mathcal{S}|\right\}u(\vec{h}_1)+\mathrm{E}\left\{|\boldsymbol{\mathcal{T}}\cap\mathcal{S}|\right\}v(\vec{h}_1). \tag{4.38}$$

Recalling the definition of $\mathscr{R}_{1,T}$ in (4.13), the expression in (4.38) motivates us to define $\mathscr{R}_{1,T}^{(\mathrm{lb})}$ as

$$\mathscr{R}_{1,T}^{(\mathrm{lb})}\triangleq u(\vec{\boldsymbol{h}}_1)\boldsymbol{\varphi}_T+v(\vec{\boldsymbol{h}}_1)\boldsymbol{\psi}_T \tag{4.39}$$

where the random variables $\boldsymbol{\varphi}_T$ and $\boldsymbol{\psi}_T$ are given by

$$\boldsymbol{\varphi}_T\triangleq\frac{\mathrm{E}\left\{|\boldsymbol{\mathcal{T}}\backslash\mathcal{S}|\right\}}{T}=\begin{cases}\theta & |\boldsymbol{d}_T|\ge T\\ \frac{\theta|\boldsymbol{d}_T|}{T} & |\boldsymbol{d}_T|\le T-1\end{cases} \tag{4.40}$$

and

$$\boldsymbol{\psi}_T\triangleq\frac{\mathrm{E}\left\{|\boldsymbol{\mathcal{T}}\cap\mathcal{S}|\right\}}{T}=\begin{cases}0 & |\boldsymbol{d}_T|\ge T\\ \theta(1-\frac{|\boldsymbol{d}_T|}{T}) & |\boldsymbol{d}_T|\le T-1\end{cases}. \tag{4.41}$$

4.2.3 Computation of $\lim_{T\to\infty} \Pr\left\{\mathcal{O}^{(\mathrm{lb})}_{1,T}\right\}$

For any $t \in [1-T, T-1] \cap \mathbb{Z}$, let

$$
\begin{aligned}
\nu_T &\triangleq \Pr\{\boldsymbol{d}_T = t\} \\
&= \int_0^{\frac{1}{2}} \Pr\left\{\boldsymbol{d}_T = t \middle| \boldsymbol{\alpha} = \alpha\right\} p_{\boldsymbol{\alpha}}(\alpha) d\alpha \\
&= \int_0^{\frac{1}{2}} \lambda_T(\alpha) p_{\boldsymbol{\alpha}}(\alpha) d\alpha. \qquad (4.42)
\end{aligned}
$$

For simplicity of notation, we show the event $\left\{\vec{\boldsymbol{h}}_1, \boldsymbol{d}_T, \boldsymbol{\alpha} : \mathscr{R}^{(\mathrm{lb})}_{1,T} < R\right\}$ by $\left\{\mathscr{R}^{(\mathrm{lb})}_{1,T} < R\right\}$. Using the tower property for conditional expectations [89],

$$
\begin{aligned}
\Pr\left\{\mathcal{O}^{(\mathrm{lb})}_{1,T}\right\} &= \mathrm{E}\left\{\mathbb{I}\left\{\mathscr{R}^{(\mathrm{lb})}_{1,T} < R\right\}\right\} \\
&= \mathrm{E}\left\{\mathrm{E}\left\{\mathbb{I}\left\{\mathscr{R}^{(\mathrm{lb})}_{1,T} < R\right\} \middle| \vec{\boldsymbol{h}}_1\right\}\right\}. \qquad (4.43)
\end{aligned}
$$

As such, we need to compute $\mathrm{E}\left\{\mathbb{I}\left\{\mathscr{R}^{(\mathrm{lb})}_{1,T} < R\right\} \middle| \vec{\boldsymbol{h}}_1\right\}$. We have

$$
\begin{aligned}
\mathrm{E}\left\{\mathbb{I}\left\{\mathscr{R}^{(\mathrm{lb})}_{1,T} < R\right\} \middle| \vec{\boldsymbol{h}}_1\right\} &= \mathrm{E}\left\{\mathbb{I}\left\{u(\vec{\boldsymbol{h}}_1)\boldsymbol{\varphi}_T + v(\vec{\boldsymbol{h}}_1)\boldsymbol{\psi}_T < R\right\} \middle| \vec{\boldsymbol{h}}_1\right\} \\
&= \mathrm{E}\left\{\mathbb{I}\left\{u(\vec{\boldsymbol{h}}_1)\boldsymbol{\varphi}_T + v(\vec{\boldsymbol{h}}_1)\boldsymbol{\psi}_T < R\right\}\mathbb{I}\big\{|\boldsymbol{d}_T| \geq T\big\} \middle| \vec{\boldsymbol{h}}_1\right\} \\
&\quad + \sum_{d=-(T-1)}^{T-1} \mathrm{E}\left\{\mathbb{I}\left\{u(\vec{\boldsymbol{h}}_1)\boldsymbol{\varphi}_T + v(\vec{\boldsymbol{h}}_1)\boldsymbol{\psi}_T < R\right\}\mathbb{I}\big\{\boldsymbol{d}_T = d\big\} \middle| \vec{\boldsymbol{h}}_1\right\}
\end{aligned}
\qquad (4.44)
$$

where the last step is due to the fact that $\mathbb{I}\big\{|\boldsymbol{d}_T| \geq T\big\} + \sum_{d=-(T-1)}^{T-1} \mathbb{I}\big\{\boldsymbol{d}_T = d\big\} = 1$. To compute the first term on the right side of (4.44), we note that as long as $|\boldsymbol{d}_T| \geq T$, then $\boldsymbol{\varphi}_T = \theta$ and $\boldsymbol{\psi}_T = 0$. Therefore,

$$
\begin{aligned}
&\mathrm{E}\left\{\mathbb{I}\left\{u(\vec{\boldsymbol{h}}_1)\boldsymbol{\varphi}_T + v(\vec{\boldsymbol{h}}_1)\boldsymbol{\psi}_T < R\right\}\mathbb{I}\big\{|\boldsymbol{d}_T| \geq T\big\} \middle| \vec{\boldsymbol{h}}_1\right\} \\
&= \mathrm{E}\left\{\mathbb{I}\left\{\theta u(\vec{\boldsymbol{h}}_1) < R\right\}\mathbb{I}\big\{|\boldsymbol{d}_T| \geq T\big\} \middle| \vec{\boldsymbol{h}}_1\right\} \\
&\overset{(a)}{=} \mathrm{E}\left\{\mathbb{I}\left\{\theta u(\vec{\boldsymbol{h}}_1) < R\right\} \middle| \vec{\boldsymbol{h}}_1\right\} \mathrm{E}\big\{\mathbb{I}\left\{|\boldsymbol{d}_T| \geq T\right\}\big\} \\
&\overset{(b)}{=} \mathbb{I}\left\{\theta u(\vec{\boldsymbol{h}}_1) < R\right\} \Pr\left\{|\boldsymbol{d}_T| \geq T\right\} \\
&\overset{(c)}{=} \mathbb{I}\left\{\theta u(\vec{\boldsymbol{h}}_1) < R\right\} (1 - (2T-1)\nu_T) \qquad (4.45)
\end{aligned}
$$

where (a) follows by independence of $\boldsymbol{d}_T$ and $\vec{\boldsymbol{h}}_1$, (b) is due to the fact that $\mathbb{I}\left\{\theta u(\vec{\boldsymbol{h}}_1) < R\right\}$ is a deterministic function of $\vec{\boldsymbol{h}}_1$, and (c) is by definition (4.42).

To compute the second term on the right side of (4.44), we need the following Lemma.

Lemma 10. *The inequality $u(\vec{\boldsymbol{h}}_1) > v(\vec{\boldsymbol{h}}_1)$ holds almost surely.*

Proof. See appendix L. $\square$

We can write

$$\begin{aligned}
& \sum_{d=-(T-1)}^{T-1} \mathrm{E}\left\{\mathbb{I}\left\{u(\vec{\boldsymbol{h}}_1)\boldsymbol{\varphi}_T + v(\vec{\boldsymbol{h}}_1)\boldsymbol{\psi}_T < R\right\}\mathbb{I}\{\boldsymbol{d}_T = d\}\middle|\vec{\boldsymbol{h}}_1\right\} \\
= \; & \sum_{d=-(T-1)}^{T-1} \mathrm{E}\left\{\mathbb{I}\left\{\tfrac{\theta|d|}{T}u(\vec{\boldsymbol{h}}_1) + \theta\left(1-\tfrac{|d|}{T}\right)v(\vec{\boldsymbol{h}}_1) < R\right\}\mathbb{I}\{\boldsymbol{d}_T = d\}\middle|\vec{\boldsymbol{h}}_1\right\} \\
= \; & \sum_{d=-(T-1)}^{T-1} \mathrm{E}\left\{\mathbb{I}\left\{\tfrac{\theta|d|}{T}u(\vec{\boldsymbol{h}}_1) + \theta\left(1-\tfrac{|d|}{T}\right)v(\vec{\boldsymbol{h}}_1) < R\right\}\middle|\vec{\boldsymbol{h}}_1\right\}\mathrm{E}\{\mathbb{I}\{\boldsymbol{d}_T = d\}\} \\
= \; & \sum_{d=-(T-1)}^{T-1} \mathbb{I}\left\{\tfrac{\theta|d|}{T}u(\vec{\boldsymbol{h}}_1) + \theta\left(1-\tfrac{|d|}{T}\right)v(\vec{\boldsymbol{h}}_1) < R\right\}\Pr\left\{\boldsymbol{d}_T = d\right\} \\
\overset{(a)}{=} \; & \nu_T \sum_{d=-(T-1)}^{T-1} \mathbb{I}\left\{|d| < \tfrac{R-\theta v(\vec{\boldsymbol{h}}_1)}{\theta(u(\vec{\boldsymbol{h}}_1)-v(\vec{\boldsymbol{h}}_1))}T\right\} \\
\overset{(b)}{=} \; & \mathbb{I}\left\{\theta v(\vec{\boldsymbol{h}}_1) < R\right\}\nu_T \sum_{d=-(T-1)}^{T-1} \mathbb{I}\left\{|d| < \tfrac{R-\theta v(\vec{\boldsymbol{h}}_1)}{\theta(u(\vec{\boldsymbol{h}}_1)-v(\vec{\boldsymbol{h}}_1))}T\right\} \\
= \; & \mathbb{I}\left\{\theta v(\vec{\boldsymbol{h}}_1) < R\right\}\nu_T \sum_{d=-(T-1)}^{T-1} \mathbb{I}\left\{|d| \le \left\lfloor\tfrac{R-\theta v(\vec{\boldsymbol{h}}_1)}{\theta(u(\vec{\boldsymbol{h}}_1)-v(\vec{\boldsymbol{h}}_1))}T\right\rfloor\right\} \\
\overset{(c)}{=} \; & \mathbb{I}\left\{\theta v(\vec{\boldsymbol{h}}_1) < R\right\}\nu_T \left(1 + 2\min\left\{T-1, \left\lfloor\tfrac{R-\theta v(\vec{\boldsymbol{h}}_1)}{\theta(u(\vec{\boldsymbol{h}}_1)-v(\vec{\boldsymbol{h}}_1))}T\right\rfloor\right\}\right) \\
\overset{(d)}{=} \; & \mathbb{I}\left\{\theta v(\vec{\boldsymbol{h}}_1) < R\right\}\mathbb{I}\left\{\tfrac{R-\theta v(\vec{\boldsymbol{h}}_1)}{\theta(u(\vec{\boldsymbol{h}}_1)-v(\vec{\boldsymbol{h}}_1))} \ge 1 - \tfrac{1}{T}\right\}(2T-1)\nu_T \\
& + \mathbb{I}\left\{\theta v(\vec{\boldsymbol{h}}_1) < R\right\}\mathbb{I}\left\{\tfrac{R-\theta v(\vec{\boldsymbol{h}}_1)}{\theta(u(\vec{\boldsymbol{h}}_1)-v(\vec{\boldsymbol{h}}_1))} < 1 - \tfrac{1}{T}\right\}\left(1 + 2\left\lfloor\tfrac{R-\theta v(\vec{\boldsymbol{h}}_1)}{\theta(u(\vec{\boldsymbol{h}}_1)-v(\vec{\boldsymbol{h}}_1))}T\right\rfloor\right)\nu_T
\end{aligned} \tag{4.46}$$

where (a) follows by Lemma 10, (b) is due to the fact that if $R \le \theta v(\vec{\boldsymbol{h}}_1)$, then we have $\mathbb{I}\left\{|d| < \frac{R-\theta v(\vec{\boldsymbol{h}}_1)}{\theta(u(\vec{\boldsymbol{h}}_1)-v(\vec{\boldsymbol{h}}_1))}T\right\} = 0$, (c) is a consequence of $\sum_{d=-(T-1)}^{T-1} \mathbb{I}\left\{|d| \le \left\lfloor\frac{R-\theta v(\vec{\boldsymbol{h}}_1)}{\theta(u(\vec{\boldsymbol{h}}_1)-v(\vec{\boldsymbol{h}}_1))}T\right\rfloor\right\}$ being equal to the size of the set $\left\{d \in \mathbb{Z} : |d| \le \min\left\{T-1, \left\lfloor\frac{R-\theta v(\vec{\boldsymbol{h}}_1)}{\theta(u(\vec{\boldsymbol{h}}_1)-v(\vec{\boldsymbol{h}}_1))}T\right\rfloor\right\}\right\}$, and (d) holds by noting that $T-1 \le \left\lfloor\frac{R-\theta v(\vec{\boldsymbol{h}}_1)}{\theta(u(\vec{\boldsymbol{h}}_1)-v(\vec{\boldsymbol{h}}_1))}T\right\rfloor$ if and only if $\frac{R-\theta v(\vec{\boldsymbol{h}}_1)}{\theta(u(\vec{\boldsymbol{h}}_1)-v(\vec{\boldsymbol{h}}_1))} \ge 1 - \frac{1}{T}$.

By (4.44), (4.45) and (4.46), one can write $\mathrm{E}\left\{\mathbb{I}\left\{\mathscr{R}_{1,T}^{(\mathrm{lb})} < R\right\}\middle|\vec{\boldsymbol{h}}_1\right\}$ as

$$\begin{aligned}
& \mathrm{E}\left\{\mathbb{I}\left\{\mathscr{R}_{1,T}^{(\mathrm{lb})} < R\right\}\middle|\vec{\boldsymbol{h}}_1\right\} = (1-(2T-1)\nu_T)\mathbb{I}\left\{\theta u(\vec{\boldsymbol{h}}_1) < R\right\} \\
& + \mathbb{I}\left\{\theta v(\vec{\boldsymbol{h}}_1) < R\right\}\mathbb{I}\left\{\frac{R-\theta v(\vec{\boldsymbol{h}}_1)}{\theta(u(\vec{\boldsymbol{h}}_1)-v(\vec{\boldsymbol{h}}_1))} \ge 1 - \frac{1}{T}\right\}(2T-1)\nu_T \\
& + \mathbb{I}\left\{\theta v(\vec{\boldsymbol{h}}_1) < R\right\}\mathbb{I}\left\{\frac{R-\theta v(\vec{\boldsymbol{h}}_1)}{\theta(u(\vec{\boldsymbol{h}}_1)-v(\vec{\boldsymbol{h}}_1))} < 1 - \frac{1}{T}\right\}\left(1 + 2\left\lfloor\frac{R-\theta v(\vec{\boldsymbol{h}}_1)}{\theta(u(\vec{\boldsymbol{h}}_1)-v(\vec{\boldsymbol{h}}_1))}T\right\rfloor\right)\nu_T.
\end{aligned} \tag{4.47}$$

The following Proposition yields $\lim_{T\to\infty}\Pr\left\{\mathcal{O}_{1,T}^{(\mathrm{lb})}\right\}$. It is shown that $\lim_{T\to\infty}\Pr\left\{\mathcal{O}_{1,T}^{(\mathrm{lb})}\right\}$ depends on the PDF of $\boldsymbol{\alpha}$ only through $\mathrm{E}\{\boldsymbol{\alpha}\}$.

Proposition 8. *Let*

$$\chi(R,\gamma,\theta) \triangleq \Pr\left\{\theta u(\vec{\boldsymbol{h}}_1) < R\right\} \tag{4.48}$$

and

$$\omega(R,\gamma,\theta) \triangleq 2\mathrm{E}\left\{\frac{(R-\theta v(\vec{\boldsymbol{h}}_1))\mathbb{I}\left\{\theta v(\vec{\boldsymbol{h}}_1) < R < \theta u(\vec{\boldsymbol{h}}_1)\right\}}{\theta(u(\vec{\boldsymbol{h}}_1)-v(\vec{\boldsymbol{h}}_1))}\right\}. \tag{4.49}$$

Then,

$$\lim_{T\to\infty}\Pr\left\{\mathcal{O}_{1,T}^{(\mathrm{lb})}\right\} = \chi(R,\gamma,\theta)+\omega(R,\gamma,\theta)\mathrm{E}\{\boldsymbol{\alpha}\}. \tag{4.50}$$

Proof. To prove the Proposition, we need to verify

$$\lim_{T\to\infty}\nu_T = 0 \tag{4.51}$$

and

$$\lim_{T\to\infty}T\nu_T = \mathrm{E}\left\{\boldsymbol{\alpha}\right\}. \tag{4.52}$$

For now, let us take these facts for granted. We will get back to the verification of (4.51) and (4.52) after proving the result of the Proposition.

The key idea is to use the Lebesgue Dominated Convergence Theorem (LDCT) [89]. By (4.43),

$$\begin{aligned}\lim_{T\to\infty}\Pr\left\{\mathcal{O}_{1,T}^{(\mathrm{lb})}\right\} &= \lim_{T\to\infty}\mathrm{E}\left\{\mathrm{E}\left\{\mathbb{I}\left\{\mathscr{R}_{1,T}^{(\mathrm{lb})}<R\right\}\Big|\vec{\boldsymbol{h}}_1\right\}\right\}\\ &\overset{(a)}{=} \mathrm{E}\left\{\lim_{T\to\infty}\mathrm{E}\left\{\mathbb{I}\left\{\mathscr{R}_{1,T}^{(\mathrm{lb})}<R\right\}\Big|\vec{\boldsymbol{h}}_1\right\}\right\}\end{aligned} \tag{4.53}$$

where (a) is due to LDCT. To justify the use of LDCT, we need to check the following. Firstly, we need to show that $\lim_{T\to\infty}\mathrm{E}\left\{\mathbb{I}\left\{\mathscr{R}_{1,T}^{(\mathrm{lb})}<R\right\}\Big|\vec{\boldsymbol{h}}_1\right\}$ exists almost surely. To verify this existence, we use (4.47), (4.51) and (4.52). It is clear that

$$\lim_{T\to\infty}\Pr\left\{|\boldsymbol{d}_T|\geq T\right\}\mathbb{I}\left\{\theta u(\vec{\boldsymbol{h}}_1)<R\right\} = (1-2\mathrm{E}\left\{\boldsymbol{\alpha}\right\})\mathbb{I}\left\{\theta u(\vec{\boldsymbol{h}}_1)<R\right\} \tag{4.54}$$

and

$$\begin{aligned}&\lim_{T\to\infty}\mathbb{I}\left\{\theta v(\vec{\boldsymbol{h}}_1)<R\right\}\mathbb{I}\left\{\frac{R-\theta v(\vec{\boldsymbol{h}}_1)}{\theta(u(\vec{\boldsymbol{h}}_1)-v(\vec{\boldsymbol{h}}_1))}\geq 1-\frac{1}{T}\right\}(2T-1)\nu_T\\ \overset{(a)}{=}\;& 2\mathrm{E}\left\{\boldsymbol{\alpha}\right\}\mathbb{I}\left\{\theta v(\vec{\boldsymbol{h}}_1)<R\right\}\mathbb{I}\left\{\theta u(\vec{\boldsymbol{h}}_1)<R\right\}\\ =\;& 2\mathrm{E}\left\{\boldsymbol{\alpha}\right\}\mathbb{I}\left\{\theta\max\left\{u(\vec{\boldsymbol{h}}_1),v(\vec{\boldsymbol{h}}_1)\right\}<R\right\}\\ \overset{(b)}{=}\;& 2\mathrm{E}\left\{\boldsymbol{\alpha}\right\}\mathbb{I}\left\{\theta u(\vec{\boldsymbol{h}}_1)<R\right\}.\end{aligned} \tag{4.55}$$

where (a) is due to the fact that $\lim_{T\to\infty} \mathbb{I}\left\{\frac{R-\theta v(\vec{\boldsymbol{h}}_1)}{\theta(u(\vec{\boldsymbol{h}}_1)-v(\vec{\boldsymbol{h}}_1))} \geq 1-\frac{1}{T}\right\} = \mathbb{I}\left\{\theta u(\vec{\boldsymbol{h}}_1) < R\right\}$ and (b) is a consequence of Lemma 10. Also, using $\mathrm{a}-1 \leq \lfloor \mathrm{a} \rfloor \leq \mathrm{a}$ for any $\mathrm{a} \in \mathbb{R}$, it is easily seen that

$$\begin{aligned}
&\lim_{T\to\infty} \mathbb{I}\left\{\theta v(\vec{\boldsymbol{h}}_1) < R\right\} \mathbb{I}\left\{\frac{R-\theta v(\vec{\boldsymbol{h}}_1)}{\theta(u(\vec{\boldsymbol{h}}_1)-v(\vec{\boldsymbol{h}}_1))} < 1-\frac{1}{T}\right\}\left(1+2\left\lfloor \frac{R-\theta v(\vec{\boldsymbol{h}}_1)}{\theta(u(\vec{\boldsymbol{h}}_1)-v(\vec{\boldsymbol{h}}_1))}T\right\rfloor\right)\nu_T \\
= \;& \frac{2\mathrm{E}\left\{\boldsymbol{\alpha}\right\}(R-\theta v(\vec{\boldsymbol{h}}_1))}{\theta(u(\vec{\boldsymbol{h}}_1)-v(\vec{\boldsymbol{h}}_1))} \mathbb{I}\left\{\theta v(\vec{\boldsymbol{h}}_1) < R < \theta u(\vec{\boldsymbol{h}}_1)\right\}. \qquad (4.56)
\end{aligned}$$

By (4.47), (4.54), (4.55) and (4.56),

$$\begin{aligned}
\lim_{T\to\infty} \mathrm{E}\left\{\mathbb{I}\left\{\mathscr{R}_{1,T}^{(\mathrm{lb})} < R\right\} \middle| \vec{\boldsymbol{h}}_1\right\} &= \mathbb{I}\left\{\theta u(\vec{\boldsymbol{h}}_1) < R\right\} \\
&\quad + \frac{2\mathrm{E}\left\{\boldsymbol{\alpha}\right\}(R-\theta v(\vec{\boldsymbol{h}}_1))}{\theta(u(\vec{\boldsymbol{h}}_1)-v(\vec{\boldsymbol{h}}_1))} \mathbb{I}\left\{\theta v(\vec{\boldsymbol{h}}_1) < R < \theta u(\vec{\boldsymbol{h}}_1)\right\}. \qquad (4.57)
\end{aligned}$$

Secondly, there must exist a random variable (or a constant) $\boldsymbol{\Upsilon}$ such that

$$\mathrm{E}\left\{\mathbb{I}\left\{\mathscr{R}_{1,T}^{(\mathrm{lb})} < R\right\} \middle| \vec{\boldsymbol{h}}_1\right\} \leq \boldsymbol{\Upsilon} \qquad (4.58)$$

for any $T \geq 1$ and $\mathrm{E}\left\{\boldsymbol{\Upsilon}\right\} < \infty$. Clearly, one can take $\boldsymbol{\Upsilon} = 1$.

This concludes the proof of (4.50).

Finally, we need to verify (4.51) and (4.52). By (4.42), $\nu_T = \mathrm{E}\left\{\lambda_T(\boldsymbol{\alpha})\right\}$. However, $\lambda_T(\boldsymbol{\alpha}) \leq 1$ for any $T \geq 1$ and $\lim_{T\to\infty} \lambda_T(\boldsymbol{\alpha}) = 0$ almost surely. Hence, by LDCT,

$$\begin{aligned}
\lim_{T\to\infty} \nu_T &= \lim_{T\to\infty} \mathrm{E}\left\{\lambda_T(\boldsymbol{\alpha})\right\} \\
&= \mathrm{E}\left\{\lim_{T\to\infty} \lambda_T(\boldsymbol{\alpha})\right\} \\
&= 0. \qquad (4.59)
\end{aligned}$$

The expression in (4.52) can be proved similarly. This concludes the Proposition. □

The following Corollary yields the exact (not an upper bound) probability of outage for continuous transmission ($\theta = 1$) of Gaussian codewords known as the CT scheme.

Corollary 3. *In the* CT *scheme, the exact value of the probability of outage is given by* $\chi(R,1) + \omega(R,1)\mathrm{E}\{\boldsymbol{\alpha}\}$.

Proof. By Proposition 8, $\chi(R,1)+\omega(R,1)\mathrm{E}\{\boldsymbol{\alpha}\}$ is an upper bound on the probability of outage for the CT scheme. To prove that this upper bound is tight, it suffices to show that $\mathscr{R}_{1,T}^{(\mathrm{lb})}$ is in fact the exact achievable rate attained by transmitting Gaussian codewords without adopting the masking scenario. We already know that $\mathrm{I}\left((\boldsymbol{s}_{1,t})_{t\in\boldsymbol{\mathcal{T}}\backslash\mathcal{S}};(\boldsymbol{y}_1[t])_{t\in\boldsymbol{\mathcal{T}}\backslash\mathcal{S}}\big|\boldsymbol{\mathcal{T}}\right)$ is exactly equal to $\mathrm{E}\left\{|\boldsymbol{\mathcal{T}}\backslash\mathcal{S}|\right\}u(\vec{\boldsymbol{h}}_1)$. However, $\mathrm{I}\left((\boldsymbol{s}_{1,t})_{t\in\boldsymbol{\mathcal{T}}\cap\mathcal{S}};(\boldsymbol{y}_1[t])_{t\in\boldsymbol{\mathcal{T}}\cap\mathcal{S}}\big|\boldsymbol{\mathcal{T}}\right)$ is only an upper bound to $\mathrm{E}\left\{|\boldsymbol{\mathcal{T}}\cap\mathcal{S}|\right\}v(\vec{\boldsymbol{h}}_1)$. By (4.39), it is sufficient to show that setting $\theta=1$ implies that $\mathrm{I}\left((\boldsymbol{s}_{1,t})_{t\in\boldsymbol{\mathcal{T}}\cap\mathcal{S}};(\boldsymbol{y}_1[t])_{t\in\boldsymbol{\mathcal{T}}\cap\mathcal{S}}\big|\boldsymbol{\mathcal{T}}\right)$ is equal to $\mathrm{E}\left\{|\boldsymbol{\mathcal{T}}\cap\mathcal{S}|\right\}v(\vec{\boldsymbol{h}}_1)$. The only steps that make $\mathrm{I}\left((\boldsymbol{s}_{1,t})_{t\in\boldsymbol{\mathcal{T}}\cap\mathcal{S}};(\boldsymbol{y}_1[t])_{t\in\boldsymbol{\mathcal{T}}\cap\mathcal{S}}\big|\boldsymbol{\mathcal{T}}\right)$ be an upper bound to $\mathrm{E}\left\{|\boldsymbol{\mathcal{T}}\cap\mathcal{S}|\right\}v(\vec{\boldsymbol{h}}_1)$ are in (4.27) (use of entropy power inequality) and (4.32) (the upper bound on the differential entropy of a complex mixed Gaussian random variable). Setting $\theta=1$ makes the sequences $(\boldsymbol{x}_{1,t})_{t=0}^{T-1}$ and $(\boldsymbol{x}_{2,t})_{t=0}^{T-1}$ be Gaussian. Since the entropy power inequality is tight for Gaussian vectors and the upper bound given in Lemma 9 is tight for Gaussian random variables, the proof of the Corollary is complete. □

The upper bound on the probability of outage given in Proposition 8 is not tight for $\theta<1$. In the next section, we demonstrate that the minimum of $\chi(R,\gamma,\theta)+\omega(R,\gamma,\theta)\mathrm{E}\{\boldsymbol{\alpha}\}$ occurs for some θ that is less than 1. Since setting $\theta=1$ results in the exact probability of outage, we conclude that the RM scheme in fact decreases the probability of outage per user. Equivalently, through using the RM scheme, it is possible to transmit at higher transmission rates, while the probability of outage is kept below a certain threshold.

4.3 System Design and Simulation Results

In the previous section we have developed an upper bound on the probability of outage given by

$$\lim_{T\to\infty}\Pr\left\{\mathcal{O}_{1,T}\right\}\leq\chi(R,\gamma,\theta)+\omega(R,\gamma,\theta)\mathrm{E}\{\boldsymbol{\alpha}\}\tag{4.60}$$

Our goal is to find θ that minimizes this upper bound on the probability of outage. As such, we set $\theta=\hat{\theta}(R,\gamma)$ where

$$\hat{\theta}(R,\gamma)\triangleq\arg\min_{\theta\in(0,1]}\Big(\chi(R,\gamma,\theta)+\omega(R,\gamma,\theta)\mathrm{E}\{\boldsymbol{\alpha}\}\Big).\tag{4.61}$$

The design criteria in (4.61) only depends on $\mathrm{E}\{\boldsymbol{\alpha}\}$. In the following, we refer to $\mathrm{E}\{\boldsymbol{\alpha}\}$ as α_0.

The performance evaluation is based on fixing a threshold $\varepsilon\in(0,1)$ and finding the largest transmission rate $R(\varepsilon)$ such that the probability of outage does not exceed ε. However, since we only have an upper bound on the probability of outage, we can only obtain a lower bound on $R(\varepsilon)$ as far as $\theta<1$. Also, by Corollary 3 we have the exact value of $R(\varepsilon)$ for $\theta=1$.

In practice, the users may not have any knowledge about α_0. In this case, the design rule in (4.61) is applied for the extreme case where $\alpha_0=0.5$. Figure 4.2 (the leftmost plot) demonstrates the performance of the RM strategy designed for $\alpha_0=0.5$ compared to the performance achieved by CT that is designed for various actual values of α_0. The curves show the highest transmission rates in terms of SNR to guarantee a probability of outage less than $\varepsilon=0.01$. Note that the RM scheme would offer a better performance if it was designed for the actual value of α_0. It is seen that even though the RM scheme is designed to handle the worst case, its performance is still superior to that of the CT scenario for a wide range of α_0. For example, it is seen that if $\alpha_0 = 0.1$, the RM scheme (designed for $\alpha_0 = 0.5$) offers a better performance than CT (designed for the actual value $\alpha_0 = 0.1$) for at least all values of SNR larger than 36dB. In fig. 4.2 (the middle plot), the corresponding activity factors used in the RM scheme designed for $\alpha_0 = 0.5$ at different SNR values are shown. In case the actual value of α_0 is less than 0.5, one expects the probability of outage be less than 0.01 by regulating the transmission rate and the activity factor in the RM scheme at the values given in fig. 4.2 (the leftmost and middle plots, respectively). This is confirmed in fig. 4.2 (the rightmost plot).

In another scenario, we let both RM and CT be designed for the actual values of α_0. Figure 4.3 compares the performance of the RM and CT schemes in terms of the transmission rate vs. SNR for different values of α_0. The probability of outage is not allowed to exceed $\varepsilon = 0.01$. The supremacy of RM over CT becomes more apparent for larger values of α_0. This is because the probability that the codewords of the two users overlap increases as α_0 becomes larger.

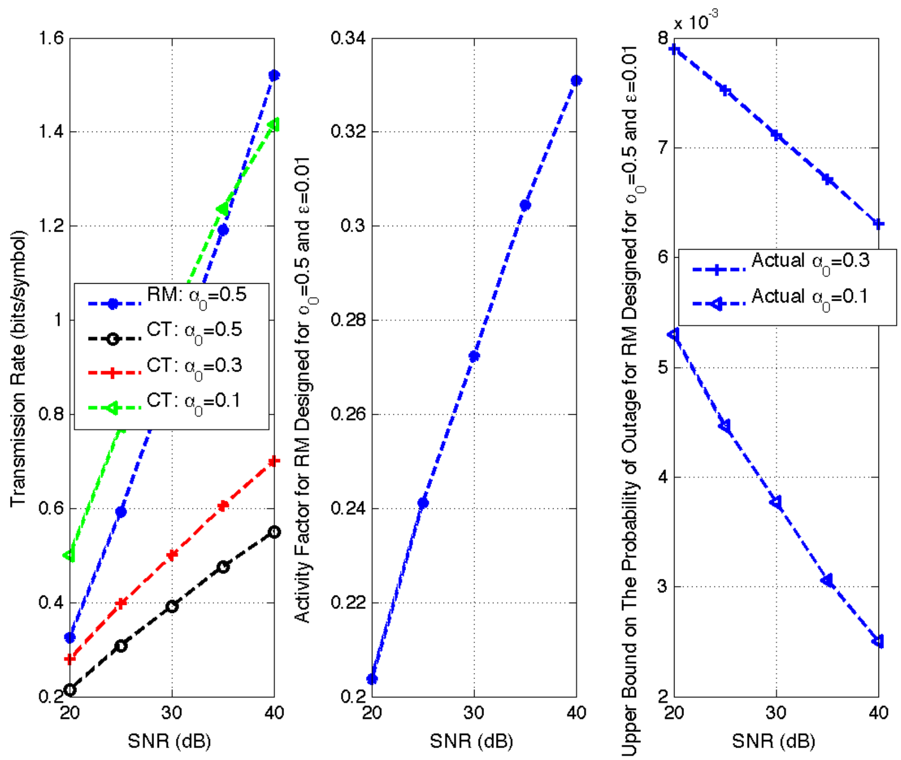

Figure 4.2: The leftmost plot presents sketch of transmission rate vs. SNR for RM and CT. The RM scheme is designed for $\alpha_0 = 0.5$. However, the curves related to the CT scheme are sketched for actual values of $\alpha_0 = 0.1, 0.3$ or 0.5. The probability of outage is not allowed to exceed 0.01, i.e., $\varepsilon = 0.01$. The middle plot demonstrates the activity factor (θ) related to the RM scheme designed for $\alpha_0 = 0.5$ and $\varepsilon = 0.01$. The rightmost plot shows the upper bound on the probability of outage for the RM scheme (designed in terms of transmission rate and activity factor for $\alpha_0 = 0.5$ and $\varepsilon = 0.01$) in case the actual value of α_0 is 0.1 or 0.3.

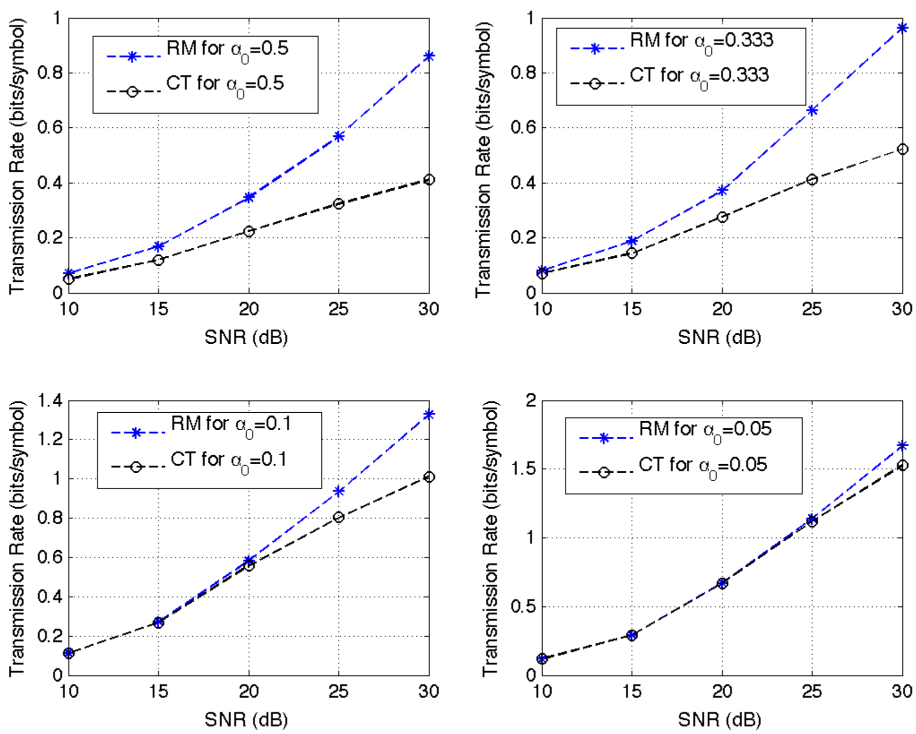

Figure 4.3: Plots of transmission rate vs. SNR for the RM and CT schemes for different values of the delay parameter. The probability of outage is not allowed to exceed 0.01.

Chapter 5

Conclusion and Future Research

5.1 Conclusion

In chapter 2, we have addressed a decentralized wireless communication network with a fixed number u of frequency sub-bands to be shared among N transmitter-receiver pairs. It is assumed that the number of active users is a realization of a random variable with a given distribution. Moreover, users are assumed to be unaware of each other's codebook and hence, no multiuser detection is possible. We proposed a randomized Frequency Hopping (FH) scheme in which each transmitter randomly hops over subsets of the u sub-bands from transmission slot to transmission slot. Assuming all users transmit Gaussian signals, the distribution of noise plus interference is mixed Gaussian, which makes the calculation of the mutual information between the input and output of each user intractable. We derived lower and upper bounds on this mutual information and demonstrated that for large SNR values, the two bounds coincide. This observation enabled us to compute the sum multiplexing gain of the system and obtain the optimum hopping strategy for maximizing this value. We compared the performance of the FH with that of the FD in terms of the following performance measures: average sum multiplexing gain ($\eta^{(1)}$) and average minimum multiplexing gain per user ($\eta^{(2)}$). We showed that (depending on the probability mass function of the number of active users) the FH system can offer a significant improvement in terms of $\eta^{(1)}$ and $\eta^{(2)}$ (implying a more efficient usage of the spectrum). In the sequel, we considered a scenario where the channel gains are quasi-static Rayleigh fading. The transmitters are assumed to be unaware of the number of active users in the network as well as the channel gains. Developing a new upper bound on the differential entropy of a mixed Gaussian random vector and via entropy power inequality, we offered three lower bounds on the ε-outage capacity for each user in the proposed scheme. Asymptotic analysis was presented in terms of SNR and outage threshold. In the asymptotically small ε regime, we observed that the maximum outage capacity is obtained for either $v = 1$ or $v = u$; in the asymptotically small SNR regime, we demonstrated that for all values of v the system achieves the optimal performance; for asymptotically large SNR, it is shown that $v_{\text{opt}} = \left\lceil \frac{u}{n_{\max}} \right\rceil$, where $n_{\max}$ is the maximum number of concurrently active users in the network. We compared the outage capacity of the underlying FH scheme with that of the FD scenario for various setups in terms of distributions on the number of active users, SNR and ε and showed that FH outperforms FD in many cases.

In chapter 3, spectral efficiency in decentralized wireless networks of separate transmitter-receiver pairs was studied. A signaling scheme was introduced where the code-book for each user consisted of two groups of codewords, referred to as signal codewords and signature codewords. Utilizing a conditional version of entropy power inequality and a key Lemma on the differential entropy of a random vector with mixed PDF an inner bound in the capacity region of the network was developed. For consistency reasons, each user designs its signature codewords based on maximizing the average (with respect to a globally known PDF for the channel gains) of the achievable rate per user. It was demonstrated how the SMG in the network (regardless of the number of users) can be made arbitrarily close to the SMG of a TD system. A pivotal observation was that the elements of the signature vectors are not equiprobable over the underlying alphabet in contrast to the customary use PN signatures in randomly spread CDMA where the chip elements are $+1$ or -1 with equal probability. Based on this observation, we showed how to achieve a nonzero SMG in a randomly spread CDMA-based multiple access channel equipped with matched filters at the common receiver. In the sequel, invoking an extremal inequality, we presented an optimality result by showing that in bluetooth systems, transmission of i.i.d. signals in consecutive transmission slots is suboptimal regardless of the PDF of the signals.

Finally, in chapter 4 a decentralized Gaussian interference channel consisting of two block-asynchronous users is studied. The network is decentralized, i.e., there is no central controller to assign the resources to the users and users do not explicitly cooperate. In particular, no user is aware of the other user's code-book. As such, multiuser detection is not possible at the receivers. We considered a scenario where the rate of data arrival at the encoders is considerably low and codewords of each user are transmitted at random instants depending on the availability of data for transmission. Users are block-asynchronous meaning there exists a mutual delay between their transmitted signal bursts. Due to the randomness of delay, no user is aware of the location of interference bursts on its transmitted data. A model for delay was suggested in which the starting point of an interference burst is uniformly distributed along the transmitted codeword. We also included the possibility that with a certain probability each user might not experience interference on a transmitted codeword at all. Since this model is non-ergodic, probability of outage was used as a computational tool to study the performance. It was proposed that each user follows a locally Randomized Masking (RM) scheme where the transmitter quits transmitting the Gaussian symbols in its codeword independently from transmission to transmission. An upper bound on the probability of outage per user was developed using entropy power inequality and a key upper bound on the differential entropy of a mixed Gaussian random variable. It was shown that by adopting the RM scheme, the probability of outage is strictly lower than the case where both users continuously transmit the Gaussian symbols in their codewords.

5.2 Future Research Directions

5.2.1 Sum Multiplexing Gain

In chapter 3, we have shown how an SMG of 1 is achievable in a decentralized interference channel with N users for any $N \geq 2$. The signature-book that achieves this SMG does not depend on the channel gains, i.e., our result holds for realization of the channel gains. As it is already pointed out, the knowledge of users about the channel gains is only partial and each user is only expected to know the gains of the channels from all transmitters to its own receiver. This raises the question if an SMG larger than 1 can be attained by incorporating the partial knowledge of users about the channel gains. Showing that an SMG larger than 1 is not achievable in a decentralized network of arbitrary number of users and for almost any realization of the channel gains would also be an interesting direction to investigate.

5.2.2 Coexistence Through Cognition

An alternative is to consider a network of one primary and $\boldsymbol{N}$ secondary users modeled by an interference channel with $\boldsymbol{N}+1$ users. We assume $\Pr\{\boldsymbol{N}=0\} > 0$, i.e., it is likely that the primary user is the only user in the network. The primary user is "dumb" in the sense that:

1- The primary transmitter is unaware of the channel gains.

2- The primary transmitter is unaware of the realization of $\boldsymbol{N}$.

3- The primary user is unaware of the code-books of the secondary users.

The secondary users are "smart" in the sense that:

1- Each secondary user is aware of its forward channel gain and the gains of the channels connecting all transmitters to its receiver.

2- Each secondary user is aware of the realization of $\boldsymbol{N}$.

3- Each secondary user is aware of the code-book of the primary user. However, the secondary users are unaware of each other's code-book, i.e., they are anonymous to each other.

Note that we only assume the secondary users are aware of the primary user's *code-book* and not its *message*. This is in contrast to the assumption made in the literature on cognitive radios that the secondary transmitter is aware of the primary user's message. Therefore, the secondary transmitter in our setup is unable to perform dirty paper coding, however, the secondary receiver is capable of multiuser detection (decoding its message and message of the primary user simultaneously, while treating other secondary users as noise).

This setup for the cognitive network is a decentralized setup where the secondary users only sneak into the network and have to make sure the primary user's performance is always maintained at a satisfactory level, while the quality of service for each secondary user is convincing as well. The strategy for the primary user is to continuously transmit Gaussian codewords as if the secondary users were not present at all. On the other hand, the secondary users perform the RM scenario proposed in chapter 4 in order to provide the primary user and each other with partially interference-free reception.

As the primary transmitter is unaware of the channel gains and the presence of the secondary users, outage probability is an appropriate measure to assess the quality of service for this user. We propose that the transmission rate R_{p} of the primary user and the activity factor θ of the secondary users must be such that the probability of outage for the primary user is kept less than a threshold ε. We define the ε-admissible region $\mathcal{A}_\varepsilon$ as

$$\mathcal{A}_\varepsilon \triangleq \{(\theta, R_{\mathrm{p}}) : \Pr\{\text{Outage for The Primary User}\} \leq \varepsilon\}.$$

The primary receiver treat the interference as noise. However, the secondary receiver has two options, i.e., treating interference as noise and multiuser detection. Let us denote the rate region associated with the i^{th} secondary user by $\mathcal{R}_i(\theta)$ and the transmission rate of this secondary user by $R_{\mathrm{s},i}$. Fixing R_{p} and θ, we define $\omega_i(\theta, R_1)$ by

$$\omega_i(\theta, R_{\mathrm{p}}) \triangleq \sup\{R_{\mathrm{s},i} : (R_{\mathrm{p}}, R_{\mathrm{s},i}) \in \mathcal{R}_i(\theta)\}.$$

Thereafter, assuming $\mathrm{U}(R_{\mathrm{p}}, R_{\mathrm{s},1}, \cdots, R_{\mathrm{s},\boldsymbol{N}})$ is a globally known utility function of the rates of the users, the rate R_{p} and the activity factor θ are selected based on the rule

$$(\hat{\theta}, \hat{R}_{\mathrm{p}}) = \arg \max_{(\theta, R_{\mathrm{p}}) \in \mathcal{A}_\varepsilon} \mathrm{E}\Big\{\mathrm{U}\Big(R_{\mathrm{p}}, \omega_1(\theta, R_{\mathrm{p}}), \cdots, \omega_{\boldsymbol{N}}(\theta, R_{\mathrm{p}})\Big)\Big\}$$

where the expectation is with respect to the channel gains and the number of secondary users. Hence, the primary user regulates its transmission rate at $\hat{R}_{\mathrm{p}}$ and each secondary transmitter regulates its activity factor at $\hat{\theta}$. Moreover, the i^{th} secondary user sets its transmission rate at $\omega_i(\hat{\theta}, \hat{R}_{\mathrm{p}})$ using its knowledge of $\boldsymbol{N}$ and the channel gains from all transmitters to its receiver.

In the sequel, one may consider a different strategy where each secondary transmitter listens to transmission of the primary transmitter and forwards this data towards the primary receiver in the next transmission slot. It is clear that the secondary receiver also experiences more interference as a result of this scheme. We can repeat the previous design technique again with the slight difference that the power splitting performed at each secondary transmitter in order to forward the signal of the primary transmitter and its own signal is also part of the optimization rule.

Appendix A; Proof of Lemma 1

Let us consider a $t \times 1$ complex vector $\vec{\mathsf{z}}$ with mixed Gaussian distribution $p_{\vec{\mathsf{z}}}(\vec{\mathsf{z}})$ with different covariance matrices $\{C_l\}_{l=1}^{L}$ and associated probabilities $\{p_l\}_{l=1}^{L}$ given by

$$p_{\vec{\mathsf{z}}}(\vec{\mathsf{z}}) = \sum_{l=1}^{L} p_l g_t^{(\mathrm{c})}(\vec{\mathsf{z}}, C_l), \tag{5.1}$$

where $g_t^{(\mathrm{c})}(\vec{\mathsf{z}}, C_l) = \frac{1}{\pi^t \det C_l} \exp\left(-\vec{\mathsf{z}}^\dagger C_l^{-1} \vec{\mathsf{z}}\right)$ and it is assumed that $C_l = \varrho_l^2 I_t$ and $\varrho_1^2 < \varrho_2^2 < \cdots < \varrho_L^2$. One can write

$$\int p_{\vec{\mathsf{z}}}(\vec{\mathsf{z}}) \log p_{\vec{\mathsf{z}}}(\vec{\mathsf{z}}) d\vec{\mathsf{z}} = \sum_{l=1}^{L} J_l \tag{5.2}$$

where $J_l = p_l \int g_t^{(\mathrm{c})}(\vec{\mathsf{z}}, C_l) \log p_{\vec{\mathsf{z}}}(\vec{\mathsf{z}}) d\vec{\mathsf{z}}$ for $1 \le l \le L$. To find a proper lower bound on each J_l, we proceed as follows. We have

$$J_l = p_l \int g_t^{(\mathrm{c})}(\vec{\mathsf{z}}, C_l) \log \left(\sum_{m=1}^{L} p_m g_t^{(\mathrm{c})}(\vec{\mathsf{z}}, C_m) \right) d\vec{\mathsf{z}}. \tag{5.3}$$

On the other hand, one can write $\log\left(\sum_{m=1}^{L} p_m g_t^{(\mathrm{c})}(\vec{\mathsf{z}}, C_m)\right)$ as

$$\begin{aligned} \log\left(\sum_{m=1}^{L} p_m g_t^{(\mathrm{c})}(\vec{\mathsf{z}}, C_m)\right) &= \log\left(p_l g_t^{(\mathrm{c})}(\vec{\mathsf{z}}, C_l)\right) \\ &\quad + \log\left(1 + \sum_{m=1}^{l-1} \frac{p_m}{p_l} \frac{g_t^{(\mathrm{c})}(\vec{\mathsf{z}}, C_m)}{g_t^{(\mathrm{c})}(\vec{\mathsf{z}}, C_l)} + \sum_{m=l+1}^{L} \frac{p_m}{p_l} \frac{g_t^{(\mathrm{c})}(\vec{\mathsf{z}}, C_m)}{g_t^{(\mathrm{c})}(\vec{\mathsf{z}}, C_l)}\right). \end{aligned} \tag{5.4}$$

However, the term $\sum_{m=l+1}^{L} \frac{p_m}{p_l} \frac{g_t^{(c)}(\vec{z},C_m)}{g_t^{(c)}(\vec{z},C_l)} = \sum_{m=l+1}^{L} \frac{p_m}{p_l} \frac{\varrho_l^{2t}}{\varrho_m^{2t}} \exp -\big((\frac{1}{\varrho_m^2} - \frac{1}{\varrho_l^2}) \vec{z}^\dagger \vec{z} \big)$ is always greater than $\sum_{m=l+1}^{L} \frac{p_m}{p_l} \frac{\varrho_l^{2t}}{\varrho_m^{2t}}$. Hence, we arrive at

$$\begin{aligned} \log\left(\sum_{m=1}^{L} p_m g_t^{(c)}(\vec{z}, C_m)\right) &\geq \log\left(p_l g_t^{(c)}(\vec{z}, C_l)\right) \\ &\quad + \log\left(1 + \sum_{m=l+1}^{L} \frac{p_m}{p_l} \frac{\varrho_l^{2t}}{\varrho_m^{2t}} + \sum_{m=1}^{l-1} \frac{p_m}{p_l} \frac{g_t^{(c)}(\vec{z}, C_m)}{g_t^{(c)}(\vec{z}, C_l)}\right). \end{aligned} \tag{5.5}$$

On the other hand, the term $\sum_{m=1}^{l-1} \frac{p_m}{p_l} \frac{g_t^{(c)}(\vec{z},C_m)}{g_t^{(c)}(\vec{z},C_l)} = \sum_{m=1}^{l-1} \frac{p_m}{p_l} \frac{\varrho_l^{2t}}{\varrho_m^{2t}} \exp -\big((\frac{1}{\varrho_m^2} - \frac{1}{\varrho_l^2}) \vec{z}^\dagger \vec{z} \big)$ is always less than $\sum_{m=1}^{l-1} \frac{p_m}{p_l} \frac{\varrho_l^{2t}}{\varrho_m^{2t}}$. Now, we use the following inequality[1], which is valid for any $b > 0$ and $0 \leq x \leq a$,

$$\log(1 + b + x) \geq \left(1 - \frac{x}{a}\right) \log(1 + b) + \frac{x}{a} \log(1 + a + b). \tag{5.6}$$

Utilizing this in the expression on the right side of (5.5), we get

$$\begin{aligned} \log\left(\sum_{m=1}^{L} p_m g_t^{(c)}(\vec{z}, C_m)\right) &\geq \left(1 - \frac{1}{\nu_l} \sum_{m=1}^{l-1} \frac{p_m}{p_l} \frac{g_t^{(c)}(\vec{z}, C_m)}{g_t^{(c)}(\vec{z}, C_l)}\right) \log(1 + \mu_l) \\ &\quad + \frac{1}{\nu_l} \sum_{m=1}^{l-1} \frac{p_m}{p_l} \frac{g_t^{(c)}(\vec{z}, C_m)}{g_t^{(c)}(\vec{z}, C_l)} \log(1 + \nu_l + \mu_l). \end{aligned} \tag{5.7}$$

where $\mu_l = \sum_{m=l+1}^{L} \frac{p_m}{p_l} \frac{\varrho_l^{2t}}{\varrho_m^{2t}}$ and $\nu_l = \sum_{m=1}^{l-1} \frac{p_m}{p_l} \frac{\varrho_l^{2t}}{\varrho_m^{2t}}$. Using this in (5.3) yields

$$\begin{aligned} J_l &\geq p_l \int g_t^{(c)}(\vec{z}, C_l) \log(p_l g_t^{(c)}(\vec{z}, C_l)) d\vec{z} + \left(p_l - \frac{\sum_{m=1}^{l-1} p_m}{\nu_l}\right) \log(1 + \mu_l) \\ &\quad + \frac{\sum_{m=1}^{l-1} p_m}{\nu_l} \log(1 + \nu_l + \mu_l). \end{aligned} \tag{5.8}$$

The first term on the right side can be calculated as

$$\begin{aligned} p_l \int g_t^{(c)}(\vec{z}, C_l) \log\left(p_l g_t^{(c)}(\vec{z}, C_l)\right) d\vec{z} &= p_l \log(p_l) \int g_t^{(c)}(\vec{z}, C_l) d\vec{z} \\ &\quad + p_l \int g_t^{(c)}(\vec{z}, C_l) \log g_t^{(c)}(\vec{z}, C_l) d\vec{z} \\ &= p_l \log p_l + p_l \int g_t^{(c)}(\vec{z}, C_l) \log g_t^{(c)}(\vec{z}, C_l) d\vec{z} \\ &\overset{(a)}{=} p_l \log p_l - p_l \log\left((\pi e)^t \det C_l\right) \\ &= p_l \log p_l - t p_l \log(\pi e \varrho_l^2) \end{aligned} \tag{5.9}$$

[1]One may verify this using Jensen's inequality and concavity of the log(.) function.

where in (a) we have used the fact that the differential entropy of a $t \times 1$ complex Gaussian vector with covariance matrix C_l is $\log\left((\pi e)^t \det C_l\right)$. Thus,

$$\begin{aligned}
\mathrm{h}(\vec{\mathbf{z}}) &= -\int p_{\vec{\mathbf{z}}}(\vec{\mathbf{z}}) \log p_{\vec{\mathbf{z}}}(\vec{\mathbf{z}}) d\vec{\mathbf{z}} = -\sum_{l=1}^{L} J_l \\
&\leq -\sum_{l=1}^{L} p_l \log p_l + t \sum_{l=1}^{L} p_l \log\left(\pi e \varrho_l^2\right) \\
&\quad -\sum_{l=1}^{L} \left(p_l - \frac{\sum_{m=1}^{l-1} p_m}{\nu_l}\right) \log(1+\mu_l) \\
&\quad -\sum_{l=1}^{L} \frac{\sum_{m=1}^{l-1} p_m}{\nu_l} \log(1+\nu_l+\mu_l).
\end{aligned} \tag{5.10}$$

Briefly,

$$\mathrm{h}(\vec{\mathbf{z}}) \leq t \sum_{l=1}^{L} p_l \log(\pi e \varrho_l^2) + \mathfrak{H} - \mathfrak{G}'' \tag{5.11}$$

where

$$\mathfrak{H} = -\sum_{l=1}^{L} p_l \log p_l \tag{5.12}$$

and

$$\mathfrak{G}'' = \sum_{l=1}^{L} \left(p_l - \frac{\sum_{m=1}^{l-1} p_m}{\nu_l}\right) \log(1+\mu_l) + \sum_{l=1}^{L} \frac{\sum_{m=1}^{l-1} p_m}{\nu_l} \log(1+\nu_l+\mu_l). \tag{5.13}$$

$\mathfrak{G}''$ is a complicated function of $\{\varrho_l\}_{l=1}^{L}$. To simplify it, one may notice that $\mathfrak{G}''$ is an increasing function of μ_l. Hence, using $\mu_l \geq 0$, we get a lower bound on $\mathfrak{G}''$, namely $\mathfrak{G}'$ given by

$$\mathfrak{G}' = \sum_{l=2}^{L} \frac{\log(1+\nu_l)}{\nu_l} \sum_{m=1}^{l-1} p_m. \tag{5.14}$$

On the other hand, using the fact that $\frac{\log(1+x)}{x}$ is a decreasing function of x, one may obtain a lower bound on $\mathfrak{G}'$ by finding an upper bound on ν_l for each l. One option is $\nu_l \leq \frac{\varrho_L^{2t}}{\varrho_1^{2t}} \frac{\sum_{m=1}^{l-1} p_m}{p_l}$. Thus, we come up with the following lower bound on $\mathfrak{G}'$

$$\mathfrak{G}' \geq \mathfrak{G} \triangleq \frac{\varrho_1^{2t}}{\varrho_L^{2t}} \sum_{l=2}^{L} p_l \log\left(1 + \frac{\varrho_L^{2t}}{\varrho_1^{2t}} \frac{\sum_{m=1}^{l-1} p_m}{p_l}\right). \tag{5.15}$$

Appendix B; Proof of (2.50)

Let us consider the generalized FH scheme with hopping parameter $v = v^*$. The number of frequency sub-bands chosen by the i^{th} user in a typical transmission slot is denoted by the random variable $\boldsymbol{v}_i$. This number changes independently from transmission slot to transmission slot. In fact, $\Pr\{\boldsymbol{v}_i = \lfloor v^* \rfloor\} = \lceil v^* \rceil - v^*$ and $\Pr\{\boldsymbol{v}_i = \lceil v^* \rceil\} = v^* - \lfloor v^* \rfloor$. We mention that $\{\boldsymbol{v}_j\}_{j=1}^N$ are independent random variables. As before, we denote the set of frequency sub-bands selected by the i^{th} user by $\boldsymbol{s}_i$ where $|\boldsymbol{s}_i| = \boldsymbol{v}_i$. To be more clear, we use the notation $\boldsymbol{s}_i(\boldsymbol{v}_i)$ instead of $\boldsymbol{s}_i$. The achievable rate of the i^{th} user is given by

$$\mathscr{R}_i = \mathrm{I}(\vec{\boldsymbol{x}}_i; \vec{\boldsymbol{y}}_i | \boldsymbol{v}_i, \boldsymbol{s}_i(\boldsymbol{v}_i)) = \sum_{v_i \in \{\lceil v^* \rceil, \lfloor v^* \rfloor\}} \Pr\{\boldsymbol{v}_i = v_i\} \mathrm{I}(\vec{\boldsymbol{x}}_i; \vec{\boldsymbol{y}}_i | \boldsymbol{v}_i = v_i, \boldsymbol{s}_i(\boldsymbol{v}_i) = s_i^*(v_i)) \tag{5.16}$$

where $s_i^*(v_i)$ denotes the selection of the first v_i sub-bands by the i^{th} user. Following the same lines that led to the lower bound in Theorem 2,

$$\mathrm{I}(\vec{\boldsymbol{x}}_i; \vec{\boldsymbol{y}}_i | \boldsymbol{v}_i = v_i, \boldsymbol{s}_i(\boldsymbol{v}_i) = s_i^*(v_i)) \geq \frac{v_i}{2} \log \left(2^{\frac{2}{v_i}\left(\mathrm{h}(\vec{\boldsymbol{x}}_i(s_i^*(v_i))) - \mathrm{h}(\vec{\boldsymbol{z}}_i')\right)} \right) \tag{5.17}$$

where $\vec{\boldsymbol{z}}_i'$ consists of the first v_i elements of $\vec{\boldsymbol{z}}_i$. On the other hand,

$$\begin{aligned} \mathrm{h}(\vec{\boldsymbol{z}}_i') &\leq \mathrm{h}(\vec{\boldsymbol{z}}_i', \{\boldsymbol{v}_j\}_{j \neq i}) \\ &= \mathrm{h}(\vec{\boldsymbol{z}}_i' | \{\boldsymbol{v}_j\}_{j \neq i}) + \mathrm{H}(\{\boldsymbol{v}_j\}_{j \neq i}) \\ &= \mathrm{h}(\vec{\boldsymbol{z}}_i' | \{\boldsymbol{v}_j\}_{j \neq i}) + \sum_{j \neq i} \mathrm{H}(\boldsymbol{v}_j) \\ &= \mathrm{h}(\vec{\boldsymbol{z}}_i' | \{\boldsymbol{v}_j\}_{j \neq i}) + (N-1)\mathscr{H}(\lceil v^* \rceil - v^*). \end{aligned} \tag{5.18}$$

Let us define

$$\boldsymbol{v}_j' = \begin{cases} \boldsymbol{v}_j & 1 \leq j \leq i-1 \\ \boldsymbol{v}_{j+1} & i \leq j \leq N-1 \end{cases} . \tag{5.19}$$

Then, we have the thread in (5.20).

$$\begin{aligned}
\mathrm{h}(\vec{\boldsymbol{z}}'_i|\{\boldsymbol{v}_j\}_{j\neq i}) &= \mathrm{h}(\vec{\boldsymbol{z}}'_i|\{\boldsymbol{v}'_j\}_{j=1}^{N-1}) \\
&= \sum_{v'_1,\cdots,v'_{N-1}\in\{\lceil v^*\rceil,\lfloor v^*\rfloor\}} \prod_{j=1}^{N-1} \Pr\{\boldsymbol{v}'_j = v'_j\}\mathrm{h}(\vec{\boldsymbol{z}}'_i|\boldsymbol{v}'_1 = v'_1,\cdots,\boldsymbol{v}'_{N-1} = v'_{N-1}) \\
&\leq v_i \sum_{v'_1,\cdots,v'_{N-1}\in\{\lceil v^*\rceil,\lfloor v^*\rfloor\}} \prod_{j=1}^{N-1} \Pr\{\boldsymbol{v}'_j = v'_j\}\mathrm{h}(\boldsymbol{z}_{i,1}|\boldsymbol{v}'_1 = v'_1,\cdots,\boldsymbol{v}'_{N-1} = v'_{N-1}).
\end{aligned} \tag{5.20}$$

Using Lemma 1,

$$\begin{aligned}
\mathrm{h}(\boldsymbol{z}_{i,1}|\boldsymbol{v}'_1 = v'_1,\cdots,\boldsymbol{v}'_{N-1} = v'_{N-1}) \leq\ & \frac{1-a_{i,0}(v'_1,\cdots,v'_{N-1})}{2}\log(\zeta_{i,L_i}(v'_1,\cdots,v'_{N-1})\gamma+1) \\
& -\sum_{l=0}^{L_i} a_{i,l}(v'_1,\cdots,v'_{N-1})\log a_{i,l}(v'_1,\cdots,v'_{N-1}) \\
& +\log(\sqrt{2\pi e}\sigma).
\end{aligned} \tag{5.21}$$

Assuming the channel gains are realizations of independent and continuous random variables, then $L_i = 2^{N-1}-1$. The set $\{a_{i,l}(v'_1,\cdots,v'_{N-1})\}_{l=0}^{2^{N-1}-1}$ consists of the numbers $\prod_{k=1}^{n}\frac{v'_{m_k}}{u}\prod_{\substack{1\leq m\leq N-1\\ m\notin\{m_k\}_{k=1}^{n}}}\left(1-\frac{v'_m}{u}\right)$ for $1\leq n\leq N-1$ and $1\leq m_1<\cdots<m_n\leq N-1$. Also,

$$a_{i,0}(v'_1,\cdots,v'_{N-1}) = \prod_{j=1}^{N-1}\left(1-\frac{v'_j}{u}\right) \tag{5.22}$$

and

$$\zeta_{i,L_i}(v'_1,\cdots,v'_{N-1}) = \sum_{j\neq i}\frac{|h_{j,i}|^2}{v'_j}. \tag{5.23}$$

Clearly,

$$\zeta_{i,L_i}(v'_1,\cdots,v'_{N-1}) \leq \frac{\sum_{j\neq i}|h_{j,i}|^2}{\lfloor v^*\rfloor}. \tag{5.24}$$

The equations (5.20) to (5.24) yield (5.25).

$$
\begin{aligned}
\mathrm{h}(\vec{\boldsymbol{z}}_i|\{\boldsymbol{v}_j\}_{j\neq i}) \le{}& \frac{v_i}{2}\left(1-\prod_{j=1}^{N-1}\left(1-\frac{\mathrm{E}\left\{\boldsymbol{v}'_j\right\}}{u}\right)\right)\log\left(\frac{\gamma\sum_{j\neq i}|h_{j,i}|^2}{\lfloor v^*\rfloor}+1\right)+v_i\log(\sqrt{2\pi e}\sigma) \\
&-v_i\sum_{n=0}^{N-1}\sum_{\{m_k\}_{k=1}^n}\mathrm{E}\left\{\prod_{k=1}^{n}\frac{\boldsymbol{v}'_{m_k}}{u}\prod_{\substack{1\le m\le N-1\\ m\notin\{m_k\}_{k=1}^n}}\left(1-\frac{\boldsymbol{v}'_m}{u}\right)\log\left(\prod_{k=1}^{n}\frac{\boldsymbol{v}'_{m_k}}{u}\prod_{\substack{1\le m\le N-1\\ m\notin\{m_k\}_{k=1}^n}}\left(1-\frac{\boldsymbol{v}'_m}{u}\right)\right)\right\}.
\end{aligned}
\tag{5.25}
$$

It is straightforward to show that the last term in (5.25) is equal to $(N-1)v_i\vartheta_{v^*}$ where ϑ_{v^*} is given in (2.51). Finally, by (5.16), (5.17), (5.18) and (5.25),

$$
\mathscr{R}_i \ge \mathrm{E}\left\{\frac{\boldsymbol{v}_i}{2}\log\left(\frac{2^{-2(N-1)\left(\vartheta_{v^*}+\frac{\mathscr{H}(\lceil v^*\rceil-v^*)}{\boldsymbol{v}_i}\right)}\gamma|h_{i,i}|^2}{\boldsymbol{v}_i\left(\frac{\gamma\sum_{j\neq i}|h_{j,i}|^2}{\lfloor v^*\rfloor}+1\right)^{1-\left(1-\frac{v^*}{u}\right)^{N-1}}}+1\right)\right\}. \tag{5.26}
$$

Appendix C; Proof of Proposition 1

We have that $\eta_{\mathrm{FD}}^{(1)} = \frac{\mathrm{E}\{\boldsymbol{N}\}u}{2n_{\max}}$ and $\eta_{\mathrm{FH}}^{(1)} = \frac{1}{2}\max_v \left\{v\mathrm{E}\left\{\boldsymbol{N}\left(1-\frac{v}{u}\right)^{\boldsymbol{N}-1}\right\}\right\}$. Let

$$\Omega(v,\boldsymbol{N}) \triangleq \boldsymbol{N}\omega_v^{\boldsymbol{N}-1} \tag{5.27}$$

where $\omega_v = 1-\frac{v}{u}$. Thinking of $\boldsymbol{N}$ as a real parameter for the moment, we have $\frac{\partial^2}{\partial \boldsymbol{N}^2}\Omega(v,\boldsymbol{N}) = \omega_v^{\boldsymbol{N}-1}\left(\boldsymbol{N}\left(\ln\omega_v\right)^2 + 2\ln\omega_v\right)$. As $\boldsymbol{N} \geq 1$, then $\frac{\partial^2}{\partial \boldsymbol{N}^2}\Omega(v,\boldsymbol{N}) \geq \omega_v^{\boldsymbol{N}-1}\left(\left(\ln\omega_v\right)^2 + 2\ln\omega_v\right)$. But, $\left(\ln\omega_v\right)^2 + 2\ln\omega_v \geq 0$ if and only if $\omega_v \leq \frac{1}{e^2}$ or $\omega_v \geq 1$. Since $\omega_v \leq 1$, we get $\omega_v \leq \frac{1}{e^2}$. This implies that the function $\Omega(v,\boldsymbol{N})$ is a convex function of $\boldsymbol{N}$ as far as $\omega_v \leq \frac{1}{e^2}$. Therefore, by Jensen's inequality,

$$\begin{aligned}
\mathrm{E}\left\{\boldsymbol{N}\left(1-\frac{v}{u}\right)^{\boldsymbol{N}-1}\right\} &= \mathrm{E}\{\Omega(v,\boldsymbol{N})\} \\
&\geq \Omega\left(v,\mathrm{E}\{\boldsymbol{N}\}\right) \\
&= \mathrm{E}\{\boldsymbol{N}\}\left(1-\frac{v}{u}\right)^{\mathrm{E}\{\boldsymbol{N}\}-1}
\end{aligned} \tag{5.28}$$

which is valid as far as $v \geq \left(1-\frac{1}{e^2}\right)u$. Hence,

$$\begin{aligned}
\eta_{\mathrm{FH}}^{(1)} &= \frac{1}{2}\max_v\left\{v\mathrm{E}\left\{\boldsymbol{N}\left(1-\frac{v}{u}\right)^{\boldsymbol{N}-1}\right\}\right\} \\
&\geq \frac{1}{2}\max_{v\in\left[\left(1-\frac{1}{e^2}\right)u,u\right]}\left\{v\mathrm{E}\left\{\boldsymbol{N}\left(1-\frac{v}{u}\right)^{\boldsymbol{N}-1}\right\}\right\} \\
&\geq \frac{1}{2}\mathrm{E}\{\boldsymbol{N}\}\max_{v\in\left[\left(1-\frac{1}{e^2}\right)u,u\right]}\left\{v\left(1-\frac{v}{u}\right)^{\mathrm{E}\{\boldsymbol{N}\}-1}\right\}.
\end{aligned} \tag{5.29}$$

The function $v\left(1-\frac{v}{u}\right)^{\mathrm{E}\{\boldsymbol{N}\}-1}$ is a concave function in terms of v that achieves its absolute maximum at $\frac{u}{\mathrm{E}\{\boldsymbol{N}\}}$. Therefore,

$$\max_{v\in\left[\left(1-\frac{1}{e^2}\right)u,u\right]}\left\{v\left(1-\frac{v}{u}\right)^{\mathrm{E}\{\boldsymbol{N}\}-1}\right\} = \mathcal{Q}\left(1-\mathcal{Q}\right)^{\mathrm{E}\{\boldsymbol{N}\}-1}u \tag{5.30}$$

where $\mathcal{Q} \triangleq \max\left\{1-\frac{1}{e^2}, \frac{1}{\mathrm{E}\{\boldsymbol{N}\}}\right\}$. Using (5.29) and (5.30),

$$\eta_{\mathrm{FH}}^{(1)} \geq \frac{1}{2}\mathcal{Q}\left(1-\mathcal{Q}\right)^{\mathrm{E}\{\boldsymbol{N}\}-1}\mathrm{E}\{\boldsymbol{N}\}u. \tag{5.31}$$

Hence, a sufficient condition for $\eta_{\mathrm{FH}}^{(1)} > \eta_{\mathrm{FD}}^{(1)}$ to hold is that

$$\mathcal{Q}\left(1-\mathcal{Q}\right)^{\mathrm{E}\{\boldsymbol{N}\}-1} > \frac{1}{n_{\max}}. \tag{5.32}$$

If $\mathrm{E}\{\boldsymbol{N}\} \geq \frac{e^2}{e^2-1}$, we have $\mathcal{Q} = 1-\frac{1}{e^2}$. Hence, (5.32) reduces to the inequality $\mathrm{E}\{\boldsymbol{N}\} < \frac{1}{2}\ln\left((e^2-1)n_{\max}\right)$. Therefore, if $\frac{e^2}{e^2-1} \leq \mathrm{E}\{\boldsymbol{N}\} < \frac{1}{2}\ln\left((e^2-1)n_{\max}\right)$, then (5.32) is satisfied. On the other hand, if $\mathrm{E}\{\boldsymbol{N}\} \leq \frac{e^2}{e^2-1} = 1.1565$, we get $\mathcal{Q} = \frac{1}{\mathrm{E}\{\boldsymbol{N}\}}$. Thus, (5.32) reduces to the inequality $\frac{1}{\mathrm{E}\{\boldsymbol{N}\}}\left(1-\frac{1}{\mathrm{E}\{\boldsymbol{N}\}}\right)^{\mathrm{E}\{\boldsymbol{N}\}-1} > \frac{1}{n_{\max}}$. For each $n_{\max} \geq 2$, this yields an upper bound on $\mathrm{E}\{\boldsymbol{N}\}$. Since $\frac{1}{\mathrm{E}\{\boldsymbol{N}\}}\left(1-\frac{1}{\mathrm{E}\{\boldsymbol{N}\}}\right)^{\mathrm{E}\{\boldsymbol{N}\}-1}$ is a decreasing function of $\mathrm{E}\{\boldsymbol{N}\}$, the smallest of these upper bounds is obtained for $n_{\max} = 2$ and is equal to 1.2938. This means that for $\mathrm{E}\{\boldsymbol{N}\} \leq 1.1565$, (5.32) is automatically satisfied. Thus, (5.32) is equivalent to

$$\mathrm{E}\{\boldsymbol{N}\} < \frac{1}{2}\ln\left((e^2-1)n_{\max}\right). \tag{5.33}$$

To prove the second part, we note that $\eta_{\mathrm{FD}}^{(2)} = \frac{u}{2n_{\max}}$ and $\eta_{\mathrm{FH}}^{(2)} = \frac{1}{2}\max_v\left\{v\mathrm{E}\left\{\left(1-\frac{v}{u}\right)^{N-1}\right\}\right\}$. The function $\left(1-\frac{v}{u}\right)^{N-1}$ is convex in terms of N. Using Jenson's inequality,

$$\begin{aligned}\eta_{\mathrm{FH}}^{(2)} &\geq \frac{1}{2}\max_v\left\{v\left(1-\frac{v}{u}\right)^{\mathrm{E}\{N\}-1}\right\}\\ &= \frac{u}{2\mathrm{E}\{N\}}\left(1-\frac{1}{\mathrm{E}\{N\}}\right)^{\mathrm{E}\{N\}-1}.\end{aligned} \tag{5.34}$$

Hence, a sufficient condition for $\eta_{\mathrm{FH}}^{(2)} > \eta_{\mathrm{FD}}^{(2)}$ to hold is

$$\frac{1}{\mathrm{E}\{N\}}\left(1-\frac{1}{\mathrm{E}\{N\}}\right)^{\mathrm{E}\{N\}-1} > \frac{1}{n_{\max}}. \tag{5.35}$$

Appendix D; Computation of $R_{\mathrm{FH}}^{(1)}(\varepsilon)$

By (2.75),

$$\Pr\left\{\mathscr{R}_i^{(1)}(\vec{\boldsymbol{h}}_i) < R\right\} = \Pr\left\{v\log\left(\frac{2^{-\mathfrak{H}(v,\boldsymbol{N})}2^{\alpha_{\boldsymbol{N}}(\boldsymbol{\mathcal{I}}_{\boldsymbol{N}-1,1};\frac{\gamma}{v},\frac{v}{u})}|\boldsymbol{h}_{i,i}|^2\gamma}{v\prod_{m=1}^{\boldsymbol{N}-1}\prod_{m'=1}^{\binom{\boldsymbol{N}-1}{m}}\left(\frac{\gamma}{v}\boldsymbol{\mathcal{I}}_{m,m'}+1\right)^{\beta_{m,\boldsymbol{N}}\left(\frac{v}{u}\right)}}+1\right) < R\right\} \tag{5.36}$$

where $\beta_{m,\boldsymbol{N}}\left(\frac{v}{u}\right) = \left(\frac{v}{u}\right)^m\left(1-\frac{v}{u}\right)^{\boldsymbol{N}-1-m}$ and for each m, $\{\boldsymbol{\mathcal{I}}_{m,m'}\}_{m'=1}^{\binom{\boldsymbol{N}-1}{m}}$ consists of all possible summations of m elements in the set $\{|\boldsymbol{h}_{j,i}|^2\}_{j=1,j\neq i}^{\boldsymbol{N}}$, such that $\prod_{m=1}^{\boldsymbol{N}-1}\prod_{m'=1}^{\binom{\boldsymbol{N}-1}{m}}(\frac{\gamma}{v}\boldsymbol{\mathcal{I}}_{m,m'}+1)^{\beta_{m,\boldsymbol{N}}\left(\frac{v}{u}\right)} = \prod_{l=1}^{\boldsymbol{L}_i}(\boldsymbol{\zeta}_{i,l}\gamma+1)^{\boldsymbol{a}_{i,l}}$. We have also substituted $\boldsymbol{\mathfrak{G}}_i$ by $\boldsymbol{\mathfrak{G}}_i^{(\mathrm{lb})}(v,\boldsymbol{N}) = \alpha_{\boldsymbol{N}}(\boldsymbol{\mathcal{I}}_{\boldsymbol{N}-1,1};\frac{\gamma}{v},\frac{v}{u})$. Since $\boldsymbol{N}$ itself is a random variable, the outage probability can be written as

$$\Pr\{\mathscr{R}_i^{(1)}(\vec{\boldsymbol{h}}_i) < R\} = q_1\xi_1 + \sum_{n=2}^{n_{\max}} q_n\xi_n, \tag{5.37}$$

where

$$\begin{aligned}
\xi_1 &= \Pr\{\mathscr{R}_i^{(1)}(\vec{\boldsymbol{h}}_i) < R|\boldsymbol{N}=1\}\\
&= \Pr\left\{v\log\left(1+\frac{|\boldsymbol{h}_{i,i}|^2\gamma}{v}\right) < R\right\}\\
&= 1-\exp\left(\frac{\left(1-2^{\frac{R}{v}}\right)v}{\gamma}\right).
\end{aligned} \tag{5.38}$$

Denoting the collection of random variables $\{\boldsymbol{\mathcal{I}}_{m,m'}\}_{\substack{1\leq m\leq n-1\\ 1\leq m'\leq\binom{n-1}{m}}}$ by $\mathscr{I}_n$, the quantity ξ_n is calculated as

$$\begin{aligned}
\xi_n &= \Pr\{\mathscr{R}_i^{(1)}(\vec{\boldsymbol{h}}_i) < R|\boldsymbol{N}=n\} \\
&= \Pr\left\{v\log\left(\frac{2^{-\mathfrak{H}(v,n)}2^{\alpha_n(\boldsymbol{\mathcal{I}}_{n-1,1};\frac{\gamma}{v},\frac{v}{u})}|\boldsymbol{h}_{i,i}|^2\gamma}{v\prod_{m=1}^{n-1}\prod_{m'=1}^{\binom{n-1}{m}}(\frac{\gamma}{v}\boldsymbol{\mathcal{I}}_{m,m'}+1)^{\beta_{m,n}\left(\frac{v}{u}\right)}}+1\right)<R\right\} \\
&= \mathrm{E}\left\{\Pr\left\{v\log\left(\frac{2^{-\mathfrak{H}(v,n)}2^{\alpha_n(\boldsymbol{\mathcal{I}}_{n-1,1};\frac{\gamma}{v},\frac{v}{u})}|\boldsymbol{h}_{i,i}|^2\gamma}{v\prod_{m=1}^{n-1}\prod_{m'=1}^{\binom{n-1}{m}}(\frac{\gamma}{v}\boldsymbol{\mathcal{I}}_{m,m'}+1)^{\beta_{m,n}\left(\frac{v}{u}\right)}}+1\right)<R\middle|\mathscr{I}_n\right\}\right\} \\
&\overset{(a)}{=} \mathrm{E}\left\{1-\exp\left(\frac{2^{\mathfrak{H}(v,n)}\left(1-2^{\frac{R}{v}}\right)v}{\gamma}2^{-\alpha_n(\boldsymbol{\mathcal{I}}_{n-1,1};\frac{\gamma}{v},\frac{v}{u})}\prod_{m=1}^{n-1}\prod_{m'=1}^{\binom{n-1}{m}}\left(\frac{\gamma}{v}\boldsymbol{\mathcal{I}}_{m,m'}+1\right)^{\beta_{m,n}\left(\frac{v}{u}\right)}\right)\right\} \\
&= 1-\mathrm{E}\left\{\exp\left(\frac{2^{\mathfrak{H}(v,n)}\left(1-2^{\frac{R}{v}}\right)v}{\gamma}2^{-\alpha_n(\boldsymbol{\mathcal{I}}_{n-1,1};\frac{\gamma}{v},\frac{v}{u})}\prod_{m=1}^{n-1}\prod_{m'=1}^{\binom{n-1}{m}}\left(\frac{\gamma}{v}\boldsymbol{\mathcal{I}}_{m,m'}+1\right)^{\beta_{m,n}\left(\frac{v}{u}\right)}\right)\right\} \\
&\overset{(b)}{=} 1-\psi_n\left(\frac{2^{\mathfrak{H}(v,n)}\left(1-2^{\frac{R}{v}}\right)v}{\gamma},\frac{\gamma}{v},\frac{v}{u}\right) \qquad (5.39)
\end{aligned}$$

where (a) follows from the fact that after conditioning on $\mathscr{I}_n$, the only random quantity is $|\boldsymbol{h}_{i,i}|^2$, which is exponentially distributed. Also, (b) holds by the definition of ψ_n. This yields the desired result.

Appendix E; Proof of Corollary 1

As $b_1 < 0$ and $0 < 2^{-\alpha_n(\mathsf{z};b_2,p_2)} < 1$ for all $\mathsf{z} > 0$, we have

$$\phi_n(b_{1,n}, b_2, p_{1,n}, p_2) \geq \frac{1}{(n-2)!} \int_0^\infty \mathsf{z}^{n-2} \exp\big(b_{1,n}(b_2\mathsf{z}+1)^{p_{1,n}} - \mathsf{z}\big) d\mathsf{z}. \tag{5.40}$$

One may easily check that $\frac{\mathsf{z}^{n-2}}{(n-2)!}\exp(-\mathsf{z})\mathbb{I}\{\mathsf{z} > 0\}$ is a PDF for some nonnegative random variable $\mathbf{z}$. Thus, (5.40) can be written as

$$\phi_n(b_{1,n}, b_2, p_{1,n}, p_2) \geq \mathrm{E}\left\{\exp\left(b_{1,n}(b_2\mathbf{z}+1)^{p_{1,n}}\right)\right\}. \tag{5.41}$$

However, as $b_{1,n} < 0$ and $0 < p_{1,n} < 1$, the function $\exp\big(b_{1,n}(b_2\mathbf{z}+1)^{p_{1,n}}\big)$ is a convex function of $\mathbf{z}$. Hence, applying Jensen's inequality yields

$$\begin{aligned} \phi_n(b_{1,n}, b_2, p_{1,n}, p_2) &\geq \exp\big(b_{1,n}(b_2\mathrm{E}\{\mathbf{z}\}+1)^{p_{1,n}}\big) \\ &= \exp\Big(b_{1,n}\big((n-1)b_2+1\big)^{p_{1,n}}\Big), \end{aligned} \tag{5.42}$$

where we have used $\mathrm{E}\{\mathbf{z}\} = n-1$. Using (5.42) in (2.93) and noting that $b_{1,n} = \frac{\left(1-2^{\frac{R}{v}}\right)v}{\gamma}2^{\mathfrak{H}(v,n)}$, $b_2 = \frac{\gamma}{v}$, and $p_{1,n} = 1 - a(v,n)$, we get the desired lower bound.

Appendix F

Setting $\hat{n} = -\lambda \ln \varepsilon$, we have

$$
\begin{aligned}
\Pr\{\widetilde{\boldsymbol{N}} \geq \hat{n}\} &\approx \sum_{n=\hat{n}}^{\infty} \frac{e^{-\lambda}\lambda^{n}}{n!} \\
&= e^{-\lambda}\frac{\lambda^{\hat{n}}}{\hat{n}!}\left(1 + \sum_{j=1}^{\infty}\prod_{k=1}^{j}\frac{\lambda}{\hat{n}+k}\right) \\
&\leq e^{-\lambda}\frac{\lambda^{\hat{n}}}{\hat{n}!}\left[\sum_{j=0}^{\infty}\left(\frac{\lambda}{\hat{n}}\right)^{j}\right] \\
&= e^{-\lambda}\frac{\lambda^{\hat{n}}}{\hat{n}!}\frac{1}{1-\frac{\lambda}{\hat{n}}} \\
&\overset{(a)}{\approx} \frac{e^{-\lambda}}{\sqrt{2\pi\hat{n}}}\left(\frac{\lambda e}{\hat{n}}\right)^{\hat{n}}\frac{1}{1-\frac{\lambda}{\hat{n}}} \\
&= \frac{e^{-\lambda}}{\sqrt{2\pi\hat{n}}}\left(\frac{e}{-\ln\varepsilon}\right)^{\hat{n}}\frac{1}{1+\frac{1}{\ln\varepsilon}} \qquad (5.43)
\end{aligned}
$$

where (a) follows by stirling approximation for $\hat{n}!$. The term $\left(\frac{e}{-\ln\varepsilon}\right)^{\hat{n}}$ can be written as

$$
\begin{aligned}
\left(\frac{e}{-\ln\varepsilon}\right)^{\hat{n}} &= e^{-\hat{n}(\ln(-\ln\varepsilon)-1)} \\
&= e^{\lambda(\ln(-\ln\varepsilon)-1)\ln\varepsilon} \\
&\overset{(a)}{\leq} e^{\ln\varepsilon} \\
&= \varepsilon. \qquad (5.44)
\end{aligned}
$$

where (a) is valid if $\lambda(\ln(-\ln\varepsilon)-1) \geq 1$, which is the case for sufficiently small ε. Combining this with the fact that $\frac{e^{-\lambda}}{\sqrt{2\pi\hat{n}}}\frac{1}{1+\frac{1}{\ln\varepsilon}} < \frac{1}{2}$ gives the desired result.

Appendix G; Proof of Proposition 5

To prove the desired result, we need some preliminary facts in linear analysis [85]. Let $\mathcal{E}$ be an Euclidean space over $\mathbb{C}$ with inner product function $\langle .,. \rangle : \mathcal{E} \times \mathcal{E} \to \mathbb{C}$.

Definition 1- Let $\mathcal{U}$ be a subspace of $\mathcal{E}$. The orthogonal complement of $\mathcal{U}$ is defined by

$$\mathcal{U}^{\perp} \triangleq \{v \in \mathcal{E} : \langle u, v \rangle = 0, \forall u \in \mathcal{U}\}. \tag{5.45}$$

Fact 1- If $\mathcal{U}$ is a subspace of $\mathcal{E}$, then for each $v \in \mathcal{E}$, there are unique elements $v_1 \in \mathcal{U}$ and $v_2 \in \mathcal{U}^{\perp}$ such that $v = v_1 + v_2$.

Definition 2- In the setup of Fact 1, v_1 is called the projection of v in $\mathcal{U}$ and is denoted by $\mathrm{proj}(v;\mathcal{U})$. By the same token, $v_2 = \mathrm{proj}(v;\mathcal{U}^{\perp})$.

Definition 3- Let $\mathcal{U}_1$ and $\mathcal{U}_2$ be subspaces of $\mathcal{E}$. We define

$$\mathrm{proj}(\mathcal{U}_1;\mathcal{U}_2) \triangleq \mathrm{span}\{\mathrm{proj}(v;\mathcal{U}_2) : v \in \mathcal{U}_1\}. \tag{5.46}$$

Fact 2- Let $\mathcal{U}_1$ and $\mathcal{U}_2$ be subspaces of $\mathcal{E}$. Then,

$$\dim(\mathcal{U}_1 \cup \mathcal{U}_2) = \dim(\mathcal{U}_1) + \dim(\mathrm{proj}(\mathcal{U}_2;\mathcal{U}_1^{\perp})). \tag{5.47}$$

Fact 3- Let X be a $p \times q$ matrix such that $\mathrm{rank}(X) = q$. Then, for any $q \times r$ matrix Y, we have $\mathrm{rank}(XY) = \mathrm{rank}(Y)$.

Using the fact that for any matrix X, $\mathrm{rank}(XX^{\dagger}) = \mathrm{rank}(X)$, it is easy to see that for any $\vec{s} \in \mathrm{supp}(\vec{\boldsymbol{s}}_i)$ and $S \in \mathrm{supp}(\boldsymbol{S}_i)$, we have $\log\det\left(I_{K-1} + \alpha^2 P G_i^{\dagger}(\vec{s}) S \Xi_i \Xi_i^{\dagger} S^{\dagger} G_i^{\dagger}(\vec{s})\right)$ scales like $\mathrm{rank}(G_i^{\dagger} S) \log P$ and $\log\det\left(I_K + \alpha^2 P S \Xi_i \Xi_i^{\dagger} S^{\dagger}\right)$ scales like $\mathrm{rank}(S) \log P$. Applying these observations to (3.38) and noting that P and γ are proportional, we get

$$\begin{aligned}
K \lim_{\gamma\to\infty} \frac{\mathscr{R}_i^{(\mathrm{lb})}(\vec{h}_i)}{\log\gamma} &= \sum_{\vec{s}\in\mathrm{supp}(\vec{\boldsymbol{s}}_i)\backslash\{\vec{0}_K\}} \Pr\{\vec{\boldsymbol{s}}_i = \vec{s}\} \\
&\quad + \sum_{\substack{S\in\mathrm{supp}(\boldsymbol{S}_i)\\ \vec{s}\in\mathrm{supp}(\vec{\boldsymbol{s}}_i)\backslash\{\vec{0}_K\}}} \Pr\{\boldsymbol{S}_i = S\}\Pr\{\vec{\boldsymbol{s}}_i = \vec{s}\}\mathrm{rank}(G_i^\dagger(\vec{s})S) \\
&\quad - \sum_{\substack{S\in\mathrm{supp}(\boldsymbol{S}_i)\\ \vec{s}\in\mathrm{supp}(\vec{\boldsymbol{s}}_i)\backslash\{\vec{0}_K\}}} \Pr\{\boldsymbol{S}_i = S\}\Pr\{\vec{\boldsymbol{s}}_i = \vec{s}\}\mathrm{rank}(S) \\
&= \Pr\{\vec{\boldsymbol{s}}_i \neq \vec{0}_K\} + \sum_{\substack{S\in\mathrm{supp}(\boldsymbol{S}_i)\\ \vec{s}\in\mathrm{supp}(\vec{\boldsymbol{s}}_i)}} \Pr\{\boldsymbol{S}_i = S\}\Pr\{\vec{\boldsymbol{s}}_i = \vec{s}\}\mathrm{rank}(G_i^\dagger(\vec{s})S) \\
&\quad - \sum_{S\in\mathrm{supp}(\boldsymbol{S}_i)} \Pr\{\boldsymbol{S}_i = S\}\Pr\{\vec{\boldsymbol{s}}_i = \vec{0}_K\}\mathrm{rank}(G_i^\dagger(\vec{0}_K)S) \\
&\quad - \mathrm{E}\{\mathrm{rank}(\boldsymbol{S}_i)\}\Pr\{\vec{\boldsymbol{s}}_i \neq \vec{0}_K\} \\
&\overset{(a)}{=} \Pr\{\vec{\boldsymbol{s}}_i \neq \vec{0}_K\} + \mathrm{E}\left\{\mathrm{rank}(\boldsymbol{G}_i^\dagger(\vec{\boldsymbol{s}}_i)\boldsymbol{S}_i)\right\} - \mathrm{E}\{\mathrm{rank}(\boldsymbol{S}_i)\}\Pr\{\vec{\boldsymbol{s}}_i = \vec{0}_K\} \\
&\quad - \mathrm{E}\{\mathrm{rank}(\boldsymbol{S}_i)\}\Pr\{\vec{\boldsymbol{s}}_i \neq \vec{0}_K\} \\
&= \Pr\{\vec{\boldsymbol{s}}_i \neq \vec{0}_K\} + \mathrm{E}\left\{\mathrm{rank}(\boldsymbol{G}_i^\dagger(\vec{\boldsymbol{s}}_i)\boldsymbol{S}_i) - \mathrm{rank}(\boldsymbol{S}_i)\right\} \qquad (5.48)
\end{aligned}$$

where (a) is by the fact that $G_i(\vec{0}_K) = I_K$. To complete the proof, we show that

$$\mathbb{I}\{\vec{\boldsymbol{s}}_i \neq \vec{0}_K\} + \mathrm{rank}(\boldsymbol{G}_i^\dagger(\vec{\boldsymbol{s}}_i)\boldsymbol{S}_i) = \mathrm{rank}\left([\boldsymbol{S}_i|\vec{\boldsymbol{s}}_i]\right) \qquad (5.49)$$

holds almost surely.

Let us write

$$\mathrm{rank}([\boldsymbol{S}_i|\vec{\boldsymbol{s}}_i]) = \dim(\mathrm{span}(\vec{\boldsymbol{s}}_i) \cup \mathrm{csp}(\boldsymbol{S}_i)). \qquad (5.50)$$

By Fact 2,

$$\begin{aligned}
\mathrm{rank}([\boldsymbol{S}_i|\vec{\boldsymbol{s}}_i]) &= \dim(\mathrm{span}(\vec{\boldsymbol{s}}_i)) + \dim(\mathrm{proj}(\mathrm{csp}(\boldsymbol{S}_i); (\mathrm{span}(\vec{\boldsymbol{s}}_i))^\perp)) \\
&= \mathbb{I}\{\vec{\boldsymbol{s}}_i \neq \vec{0}_K\} + \dim(\mathrm{proj}(\mathrm{csp}(\boldsymbol{S}_i); (\mathrm{span}(\vec{\boldsymbol{s}}_i))^\perp)). \qquad (5.51)
\end{aligned}$$

On the other hand, by the definition of $\boldsymbol{G}_i(\vec{\boldsymbol{s}}_i)$,

$$(\mathrm{span}(\vec{\boldsymbol{s}}_i))^\perp = \mathrm{csp}(\boldsymbol{G}_i(\vec{\boldsymbol{s}}_i)). \qquad (5.52)$$

It is easily seen that for any $1 \leq k \leq K-1$, the k^{th} column of the matrix $\boldsymbol{G}_i^\dagger(\vec{\boldsymbol{s}}_i)\boldsymbol{S}_i$ yields the proper linear combination of the columns of $\boldsymbol{G}_i(\vec{\boldsymbol{s}}_i)$ that constructs the projection of the k^{th} column of $\boldsymbol{S}_i$ into the space $\mathrm{csp}(\boldsymbol{G}_i(\vec{\boldsymbol{s}}_i))$, i.e.,

$$\boldsymbol{G}_i(\vec{\boldsymbol{s}}_i)[\boldsymbol{G}_i^\dagger(\vec{\boldsymbol{s}}_i)\boldsymbol{S}_i]_k = \mathrm{proj}\left([\boldsymbol{S}_i]_k; \mathrm{csp}(\boldsymbol{G}_i(\vec{\boldsymbol{s}}_i))\right). \qquad (5.53)$$

Therefore,

$$\begin{aligned}\operatorname{span}\left(\left\{\boldsymbol{G}_i(\vec{\boldsymbol{s}}_i)[\boldsymbol{G}_i^{\dagger}(\vec{\boldsymbol{s}}_i)\boldsymbol{S}_i]_k\right\}_{k=1}^{K-1}\right) &= \operatorname{proj}\left(\operatorname{span}\left(\left\{[\boldsymbol{S}_i]_k\right\}_{k=1}^{K-1}\right);\operatorname{csp}(\boldsymbol{G}_i(\vec{\boldsymbol{s}}_i))\right) \\ &= \operatorname{proj}(\operatorname{csp}(\boldsymbol{S}_i);\operatorname{csp}(\boldsymbol{G}_i(\vec{\boldsymbol{s}}_i))). \end{aligned} \tag{5.54}$$

However,

$$\operatorname{span}\left(\left\{\boldsymbol{G}_i(\vec{\boldsymbol{s}}_i)[\boldsymbol{G}_i^{\dagger}(\vec{\boldsymbol{s}}_i)\boldsymbol{S}_i]_k\right\}_{k=1}^{K}\right) = \operatorname{csp}(\boldsymbol{G}_i(\vec{\boldsymbol{s}}_i)\boldsymbol{G}_i^{\dagger}(\vec{\boldsymbol{s}}_i)\boldsymbol{S}_i). \tag{5.55}$$

By (5.54) and (5.55),

$$\operatorname{proj}(\operatorname{csp}(\boldsymbol{S}_i);\operatorname{csp}(\boldsymbol{G}_i(\vec{\boldsymbol{s}}_i))) = \operatorname{csp}(\boldsymbol{G}_i(\vec{\boldsymbol{s}}_i)\boldsymbol{G}_i^{\dagger}(\vec{\boldsymbol{s}}_i)\boldsymbol{S}_i). \tag{5.56}$$

Using (5.56) and (5.52) in (5.51),

$$\begin{aligned}\operatorname{rank}([\boldsymbol{S}_i|\vec{\boldsymbol{s}}_i]) &= \mathbb{I}\{\vec{\boldsymbol{s}}_i \neq \vec{0}_K\} + \dim(\operatorname{csp}(\boldsymbol{G}_i(\vec{\boldsymbol{s}}_i)\boldsymbol{G}_i^{\dagger}(\vec{\boldsymbol{s}}_i)\boldsymbol{S}_i)) \\ &= \mathbb{I}\{\vec{\boldsymbol{s}}_i \neq \vec{0}_K\} + \operatorname{rank}(\boldsymbol{G}_i(\vec{\boldsymbol{s}}_i)\boldsymbol{G}_i^{\dagger}(\vec{\boldsymbol{s}}_i)\boldsymbol{S}_i) \\ &\overset{(a)}{=} \mathbb{I}\{\vec{\boldsymbol{s}}_i \neq \vec{0}_K\} + \operatorname{rank}(\boldsymbol{G}_i^{\dagger}(\vec{\boldsymbol{s}}_i)\boldsymbol{S}_i) \end{aligned} \tag{5.57}$$

where (a) follows by Fact 3 as $\boldsymbol{G}_i(\vec{\boldsymbol{s}}_i)$ has independent columns. Taking expectation from both sides,

$$\mathrm{E}\left\{\operatorname{rank}([\boldsymbol{S}_i|\vec{\boldsymbol{s}}_i])\right\} = \Pr\{\vec{\boldsymbol{s}}_i \neq \vec{0}_K\} + \mathrm{E}\left\{\operatorname{rank}(\boldsymbol{G}_i^{\dagger}(\vec{\boldsymbol{s}}_i)\boldsymbol{S}_i)\right\}. \tag{5.58}$$

Using this in (5.48),

$$\begin{aligned} K\lim_{\gamma\to\infty}\frac{\mathscr{R}_i^{(\mathrm{lb})}(\vec{h}_i)}{\log\gamma} &= \mathrm{E}\left\{\operatorname{rank}([\boldsymbol{S}_i|\vec{\boldsymbol{s}}_i]) - \operatorname{rank}(\boldsymbol{S}_i)\right\} \\ &= \Pr\left\{\vec{\boldsymbol{s}}_i \notin \operatorname{csp}(\boldsymbol{S}_i)\right\}. \end{aligned} \tag{5.59}$$

This completes the proof.

Appendix H; Proof of Corollary 2

By Proposition 5,

$$\lim_{\gamma\to\infty} \frac{\mathscr{R}_i(\vec{h}_i)}{\log\gamma} \geq \frac{\Pr\left\{\vec{\boldsymbol{s}}_i \notin \operatorname{csp}(\boldsymbol{S}_i)\right\}}{K}. \tag{5.60}$$

In this appendix, we prove that

$$\lim_{\gamma\to\infty} \frac{\mathscr{R}_i(\vec{h}_i)}{\log\gamma} \leq \frac{\Pr\left\{\vec{\boldsymbol{s}}_i \notin \operatorname{csp}(\boldsymbol{S}_i)\right\}}{K}. \tag{5.61}$$

By (3.10), it suffices to show that $\lim_{\gamma\to\infty} \frac{\mathrm{I}(\boldsymbol{x}_i;\vec{\boldsymbol{y}}_i|\vec{\boldsymbol{s}}_i)}{\log\gamma} \leq \Pr\left\{\vec{\boldsymbol{s}}_i \notin \operatorname{csp}(\boldsymbol{S}_i)\right\}$. Let us consider the *informed* i^{th} user where the receiver is aware of $\vec{\boldsymbol{s}}_i$ and $\boldsymbol{S}_i$. The achievable rate of this virtual user is $\frac{\mathrm{I}(\boldsymbol{x}_i;\vec{\boldsymbol{y}}_i|\boldsymbol{s}_i,\boldsymbol{S}_i)}{K}$. It is clear that $\mathrm{I}(\boldsymbol{x}_i;\vec{\boldsymbol{y}}_i|\vec{\boldsymbol{s}}_i) \leq \mathrm{I}(\boldsymbol{x}_i;\vec{\boldsymbol{y}}_i|\vec{\boldsymbol{s}}_i,\boldsymbol{S}_i)$. However,

$$\begin{aligned}
\mathrm{I}(\boldsymbol{x}_i;\vec{\boldsymbol{y}}_i|\vec{\boldsymbol{s}}_i,\boldsymbol{S}_i) &= \sum_{\substack{\vec{s}\in\operatorname{supp}(\vec{\boldsymbol{s}}_i)\\ S\in\operatorname{range}(\boldsymbol{S}_i)}} \Pr\{\vec{\boldsymbol{s}}_i=\vec{s}\}\Pr\{\boldsymbol{S}_i=S\}\mathrm{I}(\boldsymbol{x}_i;\vec{\boldsymbol{y}}_i|\vec{\boldsymbol{s}}_i=\vec{s},\boldsymbol{S}_i=S)\\
&\overset{(a)}{=} \sum_{\substack{\vec{s}\in\operatorname{supp}(\vec{\boldsymbol{s}}_i)\\ S\in\operatorname{range}(\boldsymbol{S}_i)}} \Pr\{\vec{\boldsymbol{s}}_i=\vec{s}\}\Pr\{\boldsymbol{S}_i=S\}\log\frac{\det\left(\operatorname{cov}\left(\vec{\boldsymbol{y}}_i|\vec{\boldsymbol{s}}_i=\vec{s},\boldsymbol{S}_i=S\right)\right)}{\det\left(\operatorname{cov}\left(\vec{\boldsymbol{w}}_i+\vec{\boldsymbol{z}}_i|\boldsymbol{S}_i=S\right)\right)}\\
&= \sum_{\substack{\vec{s}\in\operatorname{supp}(\vec{\boldsymbol{s}}_i)\\ S\in\operatorname{range}(\boldsymbol{S}_i)}} \Pr\{\vec{\boldsymbol{s}}_i=\vec{s}\}\Pr\{\boldsymbol{S}_i=S\}\log\det\left(\operatorname{cov}\left(\vec{\boldsymbol{y}}_i|\vec{\boldsymbol{s}}_i=\vec{s},\boldsymbol{S}_i=S\right)\right)\\
&\quad - \sum_{S\in\operatorname{range}(\boldsymbol{S}_i)} \Pr\{\boldsymbol{S}_i=S\}\log\det\left(\operatorname{cov}\left(\vec{\boldsymbol{w}}_i+\vec{\boldsymbol{z}}_i|\boldsymbol{S}_i=S\right)\right)
\end{aligned} \tag{5.62}$$

where (a) follows by the fact that fixing $\boldsymbol{S}_i = S$ converts the channel of the i^{th} informed user to an additive Gaussian channel. On the other hand,

$$\begin{aligned}
&\sum_{\substack{\vec{s}\in\operatorname{supp}(\vec{\boldsymbol{s}}_i)\\ S\in\operatorname{range}(\boldsymbol{S}_i)}} \Pr\{\vec{\boldsymbol{s}}_i=\vec{s}\}\Pr\{\boldsymbol{S}_i=S\}\log\det\left(\operatorname{cov}\left(\vec{\boldsymbol{y}}_i|\vec{\boldsymbol{s}}_i=\vec{s},\boldsymbol{S}_i=S\right)\right)\\
=&\sum_{\substack{\vec{s}\in\operatorname{supp}(\vec{\boldsymbol{s}}_i)\\ S\in\operatorname{range}(\boldsymbol{S}_i)}} \Pr\{\vec{\boldsymbol{s}}_i=\vec{s}\}\Pr\{\boldsymbol{S}_i=S\}\log\det\left(I_K+\alpha^2P|h_{i,i}|^2\vec{s}\vec{s}^\dagger+\alpha^2PS\Xi_i\Xi_i^\dagger S^\dagger\right).
\end{aligned} \tag{5.63}$$

Noting that $\log\det\left(I_K + \alpha^2 P|h_{i,i}|^2\vec{s}\vec{s}^\dagger + \alpha^2 PS\Xi_i\Xi_i^\dagger S^\dagger\right)$ scales like $\mathrm{rank}\left([\vec{s}\,|S]\right)\log P$, we conclude that the first term on the right side of (5.62) scales like $\mathrm{E}\left\{\mathrm{rank}\left([\vec{\boldsymbol{s}}_i\,|\boldsymbol{S}_i]\right)\right\}\log P$. By the same token, the second term on the right side of (5.62) scales like $\mathrm{E}\left\{\mathrm{rank}(\boldsymbol{S}_i)\right\}\log P$. Therefore, $\mathrm{I}(\boldsymbol{x}_i;\vec{\boldsymbol{y}}_i|\vec{\boldsymbol{s}}_i)$ is upper bounded by a quantity which scales like $\left(\mathrm{E}\{\mathrm{rank}\left([\vec{\boldsymbol{s}}_i\,|\boldsymbol{S}_i]\right)\} - \mathrm{E}\{\mathrm{rank}(\boldsymbol{S}_i)\}\right)\log P$. Noting that P and γ are proportional, the result of the Corollary is immediate.

Appendix I; Proof of (3.80)

Let $\vec{s}$ be $\Big(\underbrace{1\ \ 1\ \ \cdots 1}_{k\,\text{times}}\ \ \underbrace{-1\ \ -1\ \ \cdots -1}_{K-k\,\text{times}}\Big)^{\mathrm{t}}$ and $\vec{\boldsymbol{s}}_2 = \Big(\boldsymbol{s}_{2,1}\quad \boldsymbol{s}_{2,2}\quad \cdots\quad \boldsymbol{s}_{2,K}\Big)^{\mathrm{t}}$. Then,

$$\Pr\left\{\vec{s}^{\dagger}\vec{\boldsymbol{s}}_2 = 0\right\} = \Pr\left\{\boldsymbol{s}_{2,1} + \cdots + \boldsymbol{s}_{2,k} = \boldsymbol{s}_{2,k+1} + \cdots + \boldsymbol{s}_{2,K}\right\}. \tag{5.64}$$

If l of $\boldsymbol{s}_{2,1}, \cdots, \boldsymbol{s}_{2,k}$ are $+1$ and l' of $\boldsymbol{s}_{2,k+1}, \cdots, \boldsymbol{s}_{2,K}$ are also $+1$, then for $\boldsymbol{s}_{2,1} + \cdots + \boldsymbol{s}_{2,k} = \boldsymbol{s}_{2,k+1} + \cdots + \boldsymbol{s}_{2,K}$ to happen, we must have $l' = l - k + \frac{K}{2}$ and therefore, K must be an even integer. Hence,

$$\begin{aligned} &\Pr\left\{\boldsymbol{s}_{2,1} + \cdots + \boldsymbol{s}_{2,k} = \boldsymbol{s}_{2,k+1} + \cdots + \boldsymbol{s}_{2,K}\right\} \\ =\ & \sum_{l=0}^{k}\sum_{l'=0}^{K-k}\binom{k}{l}\binom{K-k}{l'}\nu^{l+l'}\overline{\nu}^{K-l-l'}\,\mathbb{I}\left\{l' = l - k + \frac{K}{2}\right\}. \end{aligned} \tag{5.65}$$

The constraints $0 \le l' \le K-k$ and $0 \le l \le k$ yield $\max\left\{k - \frac{K}{2}, 0\right\} \le l \le \min\left\{k, \frac{K}{2}\right\}$. Therefore,

$$\begin{aligned} &\Pr\left\{\boldsymbol{s}_{2,1} + \cdots + \boldsymbol{s}_{2,k} = \boldsymbol{s}_{2,k+1} + \cdots + \boldsymbol{s}_{2,K}\right\} \\ =\ & \nu^{\frac{K}{2}-k}\overline{\nu}^{\frac{K}{2}+k}\sum_{l=\max\left\{k-\frac{K}{2},0\right\}}^{\min\left\{k,\frac{K}{2}\right\}}\binom{k}{l}\binom{K-k}{l-k+\frac{K}{2}}\left(\frac{\nu}{\overline{\nu}}\right)^{2l} \\ =\ & \nu^{\frac{K}{2}-k}\overline{\nu}^{\frac{K}{2}+k}\sum_{l=\max\left\{k-\frac{K}{2},0\right\}}^{\min\left\{k,\frac{K}{2}\right\}}\binom{k}{l}\binom{K-k}{\frac{K}{2}-l}\left(\frac{\nu}{\overline{\nu}}\right)^{2l} \\ =\ & \binom{K}{\frac{K}{2}}\nu^{\frac{K}{2}-k}\overline{\nu}^{\frac{K}{2}+k}\,\mathrm{E}\left\{\left(\frac{\nu}{\overline{\nu}}\right)^{2\boldsymbol{\kappa}}\right\} \end{aligned} \tag{5.66}$$

where $\boldsymbol{\kappa}$ is a Hypergeometric random variable with PMF

$$\Pr\{\boldsymbol{\kappa} = l\} = \frac{\binom{k}{l}\binom{K-k}{\frac{K}{2}-l}}{\binom{K}{\frac{K}{2}}}. \tag{5.67}$$

Appendix J; Proof of (3.93)

Let $\vec{\boldsymbol{s}}_1 = \begin{pmatrix} \boldsymbol{s}_{1,1} & \boldsymbol{s}_{1,2} \end{pmatrix}^{\mathrm{t}}$. Therefore,

$$\begin{aligned}
\mathrm{H}\left(\vec{\boldsymbol{s}}_1\vec{\boldsymbol{s}}_1^{\dagger}\right) &= \mathrm{H}\left(|\boldsymbol{s}_{1,1}|^2, |\boldsymbol{s}_{1,2}|^2, \boldsymbol{s}_{1,1}\boldsymbol{s}_{1,2}\right) \\
&= \mathrm{H}\left(|\boldsymbol{s}_{1,1}|^2, |\boldsymbol{s}_{1,2}|^2\right) + \mathrm{H}\left(\boldsymbol{s}_{1,1}\boldsymbol{s}_{1,2}\big||\boldsymbol{s}_{1,1}|^2, |\boldsymbol{s}_{1,2}|^2\right) \\
&= \mathrm{H}\left(|\boldsymbol{s}_{1,1}|^2\right) + \mathrm{H}\left(|\boldsymbol{s}_{1,2}|^2\right) + \mathrm{H}\left(\boldsymbol{s}_{1,1}\boldsymbol{s}_{1,2}\big||\boldsymbol{s}_{1,1}|^2, |\boldsymbol{s}_{1,2}|^2\right) \\
&= \mathrm{H}\left(|\boldsymbol{s}_{1,1}|\right) + \mathrm{H}\left(|\boldsymbol{s}_{1,2}|\right) + \mathrm{H}\left(\boldsymbol{s}_{1,1}\boldsymbol{s}_{1,2}\big||\boldsymbol{s}_{1,1}|, |\boldsymbol{s}_{1,2}|\right) \\
&= 2\mathscr{H}(\varepsilon) + \mathrm{H}\left(\boldsymbol{s}_{1,1}\boldsymbol{s}_{1,2}\big||\boldsymbol{s}_{1,1}|, |\boldsymbol{s}_{1,2}|\right). \qquad (5.68)
\end{aligned}$$

To compute $\mathrm{H}\left(\boldsymbol{s}_{1,1}\boldsymbol{s}_{1,2}\big||\boldsymbol{s}_{1,1}|, |\boldsymbol{s}_{1,2}|\right)$, we have

$$\begin{aligned}
\mathrm{H}\left(\boldsymbol{s}_{1,1}\boldsymbol{s}_{1,2}\big||\boldsymbol{s}_{1,1}|, |\boldsymbol{s}_{1,2}|\right) &= \mathrm{H}\left(\boldsymbol{s}_{1,1}\boldsymbol{s}_{1,2}\big||\boldsymbol{s}_{1,1}| = 1, |\boldsymbol{s}_{1,2}| = 1\right)\Pr\{|\boldsymbol{s}_{1,1}| = 1, |\boldsymbol{s}_{1,2}| = 1\} \\
&\quad + \mathrm{H}\left(\boldsymbol{s}_{1,1}\boldsymbol{s}_{1,2}\big||\boldsymbol{s}_{1,1}| = 0, |\boldsymbol{s}_{1,2}| = 1\right)\Pr\{|\boldsymbol{s}_{1,1}| = 0, |\boldsymbol{s}_{1,2}| = 1\} \\
&\quad + \mathrm{H}\left(\boldsymbol{s}_{1,1}\boldsymbol{s}_{1,2}\big||\boldsymbol{s}_{1,1}| = 1, |\boldsymbol{s}_{1,2}| = 0\right)\Pr\{|\boldsymbol{s}_{1,1}| = 1, |\boldsymbol{s}_{1,2}| = 0\} \\
&\quad + \mathrm{H}\left(\boldsymbol{s}_{1,1}\boldsymbol{s}_{1,2}\big||\boldsymbol{s}_{1,1}| = 0, |\boldsymbol{s}_{1,2}| = 0\right)\Pr\{|\boldsymbol{s}_{1,1}| = 0, |\boldsymbol{s}_{1,2}| = 0\} \\
&\overset{(a)}{=} \mathrm{H}\left(\boldsymbol{s}_{1,1}\boldsymbol{s}_{1,2}\big||\boldsymbol{s}_{1,1}| = 1, |\boldsymbol{s}_{1,2}| = 1\right)\Pr\{|\boldsymbol{s}_{1,1}| = 1, |\boldsymbol{s}_{1,2}| = 1\}
\end{aligned} \qquad (5.69)$$

where (a) is due to $\mathrm{H}\left(\boldsymbol{s}_{1,1}\boldsymbol{s}_{1,2}\big||\boldsymbol{s}_{1,1}| = 0, |\boldsymbol{s}_{1,2}| = 1\right) = \mathrm{H}\left(\boldsymbol{s}_{1,1}\boldsymbol{s}_{1,2}\big||\boldsymbol{s}_{1,1}| = 1, |\boldsymbol{s}_{1,2}| = 0\right) = \mathrm{H}\left(\boldsymbol{s}_{1,1}\boldsymbol{s}_{1,2}\big||\boldsymbol{s}_{1,1}| = 0, |\boldsymbol{s}_{1,2}|\right.$ 0. On the other hand, it is easy to see that $\Pr\{\boldsymbol{s}_{1,1}\boldsymbol{s}_{1,2} = 1\big||\boldsymbol{s}_{1,1}| = 1, |\boldsymbol{s}_{1,2}| = 1\} = \nu^2 + \overline{\nu}^2$. This implies $\mathrm{H}\left(\boldsymbol{s}_{1,1}\boldsymbol{s}_{1,2}\big||\boldsymbol{s}_{1,1}| = 1, |\boldsymbol{s}_{1,2}| = 1\right) = \mathscr{H}(\nu^2 + \overline{\nu}^2)$. Therefore,

$$\mathrm{H}\left(\boldsymbol{s}_{1,1}\boldsymbol{s}_{1,2}\big||\boldsymbol{s}_{1,1}|, |\boldsymbol{s}_{1,2}|\right) = \mathscr{H}(\nu^2 + \overline{\nu}^2)\Pr\{|\boldsymbol{s}_{1,1}| = 1, |\boldsymbol{s}_{1,2}| = 1\} = \overline{\varepsilon}^2\mathscr{H}(\nu^2 + \overline{\nu}^2). \qquad (5.70)$$

Using (5.68) and (5.70),

$$\mathrm{H}\left(\vec{\boldsymbol{s}}_1\vec{\boldsymbol{s}}_1^{\dagger}\right) = 2\mathscr{H}(\varepsilon) + \overline{\varepsilon}^2\mathscr{H}(\nu^2 + \overline{\nu}^2). \qquad (5.71)$$

Appendix K; Proof of (3.107)

Let $\vec{\boldsymbol{s}}_1 = \begin{pmatrix} \boldsymbol{s}_{1,1} & \cdots & \boldsymbol{s}_{1,K} \end{pmatrix}^{\mathrm{t}}$. We have

$$\begin{aligned} \mathrm{H}(\vec{\boldsymbol{s}}_1 \vec{\boldsymbol{s}}_1^{\dagger}) &= \mathrm{H}\left((\boldsymbol{s}_{1,k}\boldsymbol{s}_{1,l})_{1\le k,l\le K}\right) \\ &\overset{(a)}{=} \mathrm{H}\left(\left(\boldsymbol{s}_{1,k}\boldsymbol{s}_{1,l}\right)_{1\le k,l\le K, k\neq l}\right) \\ &\overset{(b)}{=} \mathrm{H}\left(\boldsymbol{s}_{1,1}\boldsymbol{s}_{1,2}, \boldsymbol{s}_{1,1}\boldsymbol{s}_{1,3}, \cdots, \boldsymbol{s}_{1,1}\boldsymbol{s}_{1,K}\right) \end{aligned} \tag{5.72}$$

where (a) is by the fact that $\boldsymbol{s}_{1,k}^2 = 1$ for any $1 \le k \le K$ and (b) is by the fact that for any two distinct numbers $k, l \in \{2, \cdots, K\}$, the knowledge about $\boldsymbol{s}_{1,k}\boldsymbol{s}_{1,l}$ can be obtained by knowing $\boldsymbol{s}_{1,1}\boldsymbol{s}_{1,k}$ and $\boldsymbol{s}_{1,1}\boldsymbol{s}_{1,l}$ as $\boldsymbol{s}_{1,1}\boldsymbol{s}_{1,k}\boldsymbol{s}_{1,1}\boldsymbol{s}_{1,l} = \boldsymbol{s}_{1,k}\boldsymbol{s}_{1,l}\boldsymbol{s}_{1,1}^2 = \boldsymbol{s}_{1,k}\boldsymbol{s}_{1,l}$. Let us define

$$\vec{\mathbf{s}} = \begin{pmatrix} \boldsymbol{s}_{1,2} & \cdots & \boldsymbol{s}_{1,K} \end{pmatrix}^{\mathrm{t}}. \tag{5.73}$$

By (5.72), $\mathrm{H}(\vec{\boldsymbol{s}}_1 \vec{\boldsymbol{s}}_1^{\dagger}) = \mathrm{H}(\boldsymbol{s}_{1,1}\vec{\mathbf{s}})$. Let $\mathcal{F}$ be the event where $\boldsymbol{s}_{1,1} = 1$, while k of the elements of $\vec{\mathbf{s}}$, namely, $\boldsymbol{s}_{1,l_1}, \cdots, \boldsymbol{s}_{1,l_{k-1}}$ and $\boldsymbol{s}_{1,l_k}$ are 1 and the rest are -1 for some $0 \le k \le K$ and $2 \le l_1 < l_2 < \cdots < l_k \le K$. Also, let $\mathcal{G}$ be the event where $\boldsymbol{s}_{1,1} = -1$, $\boldsymbol{s}_{1,l} = -1$ for $l \in \{l_1, l_2, \cdots, l_k\}$ and $\boldsymbol{s}_{1,l} = 1$ for $l \notin \{l_1, l_2, \cdots, l_k\}$. It is clear that

$$\boldsymbol{s}_{1,1}\vec{\mathbf{s}}\,\mathbb{I}\{\mathcal{F}\} = \boldsymbol{s}_{1,1}\vec{\mathbf{s}}\,\mathbb{I}\{\mathcal{G}\}. \tag{5.74}$$

We know that $\Pr\{\mathcal{F}\} = \nu^{k+1}\overline{\nu}^{K-k}$ and $\Pr\{\mathcal{G}\} = \nu^{K-k}\overline{\nu}^{k+1}$. Hence, using (5.74),

$$\mathrm{H}(\vec{\boldsymbol{s}}_1 \vec{\boldsymbol{s}}_1^{\dagger}) = -\sum_{k=0}^{K} \binom{K}{k} \left(\nu^{k+1}\overline{\nu}^{K-k} + \nu^{K-k}\overline{\nu}^{k+1}\right) \log\left(\nu^{k+1}\overline{\nu}^{K-k} + \nu^{K-k}\overline{\nu}^{k+1}\right). \tag{5.75}$$

Appendix L; Proof of Lemma 10

Intuitively speaking, $u(\vec{h}_1)$ is the *capacity* of a channel with forward gain $h_{1,1}$ and additive ambient noise $\mathcal{CN}(0,1)$, while $v(\vec{h}_1)$ is an achievable rate in the same channel where in addition to the ambient noise, a mixed Gaussian interference is present. Therefore, one expects $u(\vec{h}_1)$ to be larger than $v(\vec{h}_1)$.

To give a precise proof, we note that $v(\vec{h}_1) < u(\vec{h}_1)$ holds if and only if

$$2^{-\mathscr{H}(\theta)}\left(1+\frac{\theta}{\overline{\theta}}\frac{1}{1+\frac{|h_{2,1}|^2\gamma}{\theta}}\right)^{\overline{\theta}}\left(1+\frac{\overline{\theta}}{\theta}\left(1+\frac{|h_{2,1}|^2\gamma}{\theta}\right)\right)^{\frac{\theta}{1+\frac{|h_{2,1}|^2\gamma}{\theta}}} \leq 1. \tag{5.76}$$

This is equivalent to

$$\overline{\theta}\log\left(1+\frac{\theta}{\overline{\theta}}\frac{1}{1+\frac{|h_{2,1}|^2\gamma}{\theta}}\right)+\frac{\theta}{1+\frac{|h_{2,1}|^2\gamma}{\theta}}\log\left(1+\frac{\overline{\theta}}{\theta}\left(1+\frac{|h_{2,1}|^2\gamma}{\theta}\right)\right)\leq \mathscr{H}(\theta). \tag{5.77}$$

To prove (5.77), let x be a complex mixed Gaussian random variable with PDF given in (4.29). By the concavity property of differential entropy [80],

$$\mathrm{h}(\mathsf{x}) \geq p_1\log\left(\pi e\sigma_1^2\right)+p_2\log\left(\pi e\sigma_2^2\right). \tag{5.78}$$

Combining (5.78) with the result of Lemma 9 given in (4.30),

$$-p_1\log p_1 - p_2\log p_2 - p_1\log\left(1+\frac{p_2}{p_1}\frac{\sigma_1^2}{\sigma_2^2}\right)-\frac{p_2\sigma_1^2}{\sigma_2^2}\log\left(1+\frac{p_1}{p_2}\frac{\sigma_2^2}{\sigma_1^2}\right)\geq 0. \tag{5.79}$$

Setting $p_1=\overline{\theta}$, $p_2=\theta$, $\sigma_1^2=1$ and $\sigma_2^2=1+\frac{|h_{2,1}|^2}{\theta}$ in (5.79) yields (5.77) which is the desired result.

Bibliography

[1] Y. Liang, V. V. Veeravalli, and H. V. Poor, "Resource allocation for wireless fading relay channels: Max-min solution," *IEEE Trans. on Inf. Theory*, vol. 53, no. 10, pp. 3432–3453, October 2007.

[2] K. Kumaran and H. Viswanathan, "Joint power and bandwidth allocation in downlink transmission," *IEEE Trans. on Wireless Commun*, vol. 4, no. 3, pp. 1008–1016, May 2005.

[3] T. Holliday, A. Goldsmith, N. Bambos and P. Glynn, "Distributed power and admission control for time-varying wireless networks," in *Proc. IEEE International Symposium on Information Theory (ISIT'04)*, July 2004, pp. 352-352.

[4] T. ElBatt and A. Ephremides, "Joint scheduling and power control for wireless ad hoc networks," *IEEE Trans. on Wireless Comm.*, vol. 3, no. 1, pp. 74–85, Jan 2004.

[5] O. Seong-Jun, D. Zhang, and K. M. Wasserman, "Optimal resource allocation in multiservice CDMA networks," *IEEE Trans. on Wireless Comm.*, vol. 2, no. 4, pp. 811–821, July 2003.

[6] Z. Han, Z. Ji, and K. J. R. Liu, "Fair multiuser channel allocation for OFDMA networks using Nash bargaining solutions and coalitions," *IEEE Trans. on Comm.*, vol. 53, no. 8, pp. 1366–1376, August 2005.

[7] I. Katzela and M. Naghshineh, "Channel assignment schemes for cellular mobile telecommunication systems: a comprehensive survey," *IEEE Personal Communications*, vol. 3, no. 3, pp. 10–31, June 1996.

[8] S. G. Kiani and D. Gesbert, "Maximizing the capacity of large wireless networks: optimal and distributed solutions," in *Proc. IEEE International Symposium on Information Theory (ISIT'06)*, Seattle, USA, July 2006, pp. 2501–2505.

[9] R. Etkin, A.P. Parekh and D. Tse, "Spectrum sharing for unlicensed bands", *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 3, pp. 517-528, April 2007.

[10] J. Mitola, “Cognitive Radio: An integrated agent architecture for software defined radio”, *PhD Diss.*, KTH, Stockholm, Sweden, Dec. 2000.

[11] N. Devroye, P. Mitran and V. Tarokh, “Achievable rates in cognitive radio channels”, *IEEE Trans. on Inf. Theory*, vol. 52, no. 5, pp. 1813-1827, May 2006.

[12] A. Jovicic and P. Viswanath, “Cognitive radio: an information theoretic perspective”, *Submitted to IEEE Trans. on Inf. Theory*, April 2006.

[13] I. Maric, A. Goldsmith, G. Kramer and S. Shamai(Shitz), ”An achievable rate region for interference channels with a cognitive transmitter”, *The 41st Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, Vov. 2007.

[14] S. A. Jafar and S. Srinivasa, “Capacity limits of Cognitive radio with distributd and dynamic spectral activity”, *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 3, pp. 529-537, April 2007.

[15] A. Sahai, N. Hoven, S. M. Mishra and R. Tandra, “Fundamental tradeoffs in robust spectrum sensing for opportunistic frequency reuse”, *Technical Reports*, March 2006, Available online at http:// www.eecs,berkeley.edu/ sahai/papers/cognitivetechreport06.pdf.

[16] R. L. Peterson, R. E. Ziemer and D. E. Borth, “Introduction to spread spectrum communications”, *Prentice-Hall Inc.*, 1995.

[17] C. M. Cordeiro and D. P. Agrawal, “Ad-hoc and sensor networks- theory and applications”, *World Sci. Pub. Company*, 2006.

[18] S. Srivastava, S. Tripathi, D. Sanghi and A. K. Chaturvedi, “A code allocation protocol for maximizing throughput in CDMA based ad-hoc networks”, *IEEE Wireless Commun. and Networking Conf.*, vol. 2, pp. 1385-1390, March 2003.

[19] S. Verdú and S. Shamai (Shitz), ”Spectral efficiency of CDMA with random spreading”, *IEEE Trans. Inf. Theory*, vol. 45, no.2, pp. 622-640, March 1999.

[20] S. Shamai (Shitz) and S. Verdú, “The impact of frequency-flat fading on the spectral efficiency of CDMA”, *IEEE Trans. Inf. Theory*, vol. 47, no. 4, pp. 1302-1327, May 2001.

[21] J. Haartsen, “Bluetooth-The universal radio interface for ad-hoc wireless connectivity”, *Erricson Review*, no. 3. 1998.

[22] T. Salonidis, P. Bhagwat, L. Tassiulas and R. Lamaire, "Distributed topology construction of bluetooth wireless personal area networks", *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 3, pp. 633-643, March 2005.

[23] J. Haartsen and S. Mattisson, "Bluetooth- A new low power radio interface providing short range connectivity", *Proceedings of the IEEE*, vol. 88, no. 10, pp. 1651-1661, Oct. 2000.

[24] Z. Kostic and N. Sollenberger, "Performance and implementation of dynamic frequency hopping in limited-bandwidth cellular systems", *IEEE Trans. on Wireless Comm.*, vol. 1, no. 1, pp. 28-36, Jan. 2002.

[25] O. Mubarek, H. Yanikomeroglu and S. Periyalwar, "Dynamic frequency hopping in cellular fixed relay networks", *IEEE Veh. Tech. Conf.*, vol. 5, pp. 3112-3116, June 2005.

[26] Z. Kostic, I. Maric and X. Wang, "Fundamentals of dynamic frequency hopping in cellular systems", *IEEE Journal of Selected Areas in Comm.*, vol. 19, no. 11, pp. 2254-2266, Nov. 2001.

[27] R. Knopp and P. Humblet, "Information capacity and power control in single-cell multiuser communications," in *IEEE Int. Conf. Commun.*, June 1995, vol. 1, pp. 331-335.

[28] P. Viswanath, D. Tse, and R. Laroia, "Opportunistic beamforming using dumb antennas," *IEEE Trans. on Inf. theory*, vol. 48, no. 6, pp. 1277-1294, June 2002.

[29] M. Sharif and B. Hassibi, "On the capacity of MIMO broadcast channel with partial side information", *IEEE Trans. on Inf. Theory*, vol. 51, no. 2, pp. 506-522, Feb. 2005.

[30] A. Bayesteh and A. K. Khandani, "On the User Selection for MIMO Broadcast Channels", *IEEE Trans. on Inf. Theory*, vol. 54, no. 3, pp. 1086-1107, March 2008.

[31] T. Yoo and A. Goldsmith, "On the optimality of multi-antenna broadcast scheduling using zero-forcing Beamforming," *IEEE JSAC Special Issue on 4G Wireless Systems*, vol. 24, no. 3, pp. 528-541, March 2006.

[32] M. Ebrahimi, M.A. Maddah-Ali, A.K. Khandani, "Throughput Scaling Laws for Wireless Networks With Fading Channels", *IEEE Trans. on Inf. Theory*, vol. 53, no. 11, pp. 4250-4254, Nov. 2007.

[33] J. Abouei, A. Bayesteh, M. Ebrahimi, and A. K. Khandani, "On the throughput maximization in decentralized wireless networks", *Submitted to IEEE Trans. on Inf. Theory*, Oct. 2008.

[34] L.H. Ozarow, S. Shamai, and A. D. Wyner, "Information theoretic considerations for cellular mobile radio", *IEEE Trans. on Inf. Theory*, vol. 43, no. 2, pp. 359-378, May 1994.

[35] P. Gupta and P. R. Kumar, "The capacity of wireless networks", *IEEE Trans. on Inf. Theory*, vol. 46, no. 2, pp. 388-404, March 2000.

[36] M. Franceschetti, O. Dousse, D. Tse, and P. Thiran, "Closing the gap in the capacity of wireless networks via percolation theory", *IEEE Trans. on Inf. Theory*, vol. 53, no. 3, pp. 1009-1018, March 2007.

[37] L.-L. Xie and P. R. Kumar, "A network information theory for wireless communication: scaling laws and optimal operation", *IEEE Trans. on Inf. Theory*, vol. 50, no. 5, pp. 748-767, May 2004.

[38] M. Gastpar and M. Vetterli, "On the capacity of large gaussian relay networks", *IEEE Trans. on Inf. Theory*, vol. 51, no. 3, pp. 765-779, March 2005.

[39] O. Leveque and E. Telatar, "Information theoretic upper bounds on the capacity of large extended ad hoc wireless networks", *IEEE Trans. on Inf. Theory*, vol. 51, no. 3, pp. 858-865, March 2005.

[40] O. Dousse, M. Franceschetti, and P. Thiran, "On the throughput scaling of wireless relay networks", *IEEE Trans. on Inf. Theory*, vol. 52, no. 6, pp. 2756-2761, June 2006.

[41] M. Grossglauser and D. N. C. Tse, "Mobility increases the capacity of ad hoc wireless networks", *IEEE/ACM Trans. Networking*, vol. 10, no. 4, pp. 477-486, August 2002.

[42] S. R. Kulkarni and P. Viswanath, "A deterministic approach to throughput scaling in wireless networks", *IEEE Trans. Inf. Theory,* vol. 50, no. 6, pp. 1041-1049, June 2004.

[43] R. Gowaikar, B. Hochwald, and B. Hassibi, "Communication over a wireless network with random connections", *IEEE Trans. Inf. Theory*, vol. 52, no. 7, pp. 2857-2871, July 2006.

[44] R. Gowaikar and B. Hassibi, "On the achievable throughput in two-scale wireless networks", in Proc. *IEEE ISIT*, 2006.

[45] S. Toumpis and A. J. Goldsmith, "Large wireless networks under fading, mobility, and delay constraints", in Proc. *IEEE Infocom*, vol. 1, Hong Kong, 2004, pp. 609-619.

[46] F. Xue, L.-L. Xie, and P. R. Kumar, "The transport capacity of wireless networks over fading channels", *IEEE Trans. Inf. Theory*, vol. 51, no. 3, pp. 834-847, March 2005.

[47] R. Etkin, “Spectrum sharing: Fundamental limits, scaling laws, and self-enforcing protocols”, Ph.D. dissertation, EECS Department, University of California, Berkeley, 2006.

[48] S. Weber, J. G. Andrews, and N. Jindal, “Throughput and transmission capacity of ad hoc networks with channel state information”, in *Allerton Conference on Communication, Control, and Computing*, Monticello, IL, September 2006.

[49] —- “The effect of fading, channel inversion, and threshold scheduling on ad hoc networks”, to appear in *IEEE Trans. Inf. Theory*, November 2007.

[50] D. Gesbert and M. Kountouris, “Resource allocation in multicell wireless networks: Some capacity scaling laws”, in *Proc. Workshop on Resource Allocation in Wireless NETworks* (RAWNET 07), 2007.

[51] — “Joint power control and user scheduling in multicell wireless networks: Capacity scaling laws”, submitted to *IEEE Trans. on Inf. Theory*, September 2007.

[52] N. Jindal, J. G. Andrews, and S. Weber, “Bandwidth-SINR tradeoffs in spatial networks”, in *Proc. IEEE International Symposium on Inf. Theory*, Nice, France, June 2007.

[53] Masoud Ebrahimi and Amir K. Khandani, “Rate-Constrained Wireless Networks with Fading Channels: Interference-Limited and Noise-Limited Regimes”, submitted to *IEEE Trans. on Inf. Theory.*

[54] N. Jindal, J. G. Andrews and S. Weber, “Bandwidth partitioning in decentralized wireless networks”, *IEEE Trans. on Wireless Comm.*, vol. 7, no. 12, pp. 5408- 5419, Dec. 2008.

[55] N. Jindal, J. G. Andrews and S. Weber, “Fractional power control for decentralized wireless networks”, *IEEE Trans. on Wireless Comm.*, vol. 7, no. 12, pp. 5482- 5492, Dec. 2008.

[56] N. Abramson, “The throughput of packet broadcasting channels”, *IEEE Trans. on Comm.*, vol. 25, no. 1, pp. 117-128, Jan. 1977.

[57] S. Ghez, S. Verdú and S. C. Schwartz, “Stability properties of slotted Aloha with multipacket reception capability”, *IEEE Trans. on Automatic Control*, vol. 33, no. 7, pp. 640-649, July 1988.

[58] S. Ghez, S. Verdú and S. C. Schwartz, “Optimal decentralized control in the random access multipacket channel”, *IEEE Trans. on Automatic Control*, vol. 34, no. 11, pp. 1153-1163, Nov. 1989.

[59] B. Hajek and T. Loon, "Decentralized dynamic control of a multiaccess broadcast channel", *IEEE Trans. on Automatic Control*, vol. 27, no. 3, pp. 559-569, June 1982.

[60] T. Cui and T. Ho, "On multiple access random medium access control", *International Symposium on Inf. Theory*, ISIT09, Seoul, Korea, June 2009.

[61] P. Minero, M. Franceschetti and D.N.C. Tse, Random access: an information-theoretic perspective, submitted to *IEEE Trans. on Inf. Theory*, December 2009.

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

[63] A. Motahari and A.K. Khandani, "Capacity bounds for the Gaussian interference channel", *IEEE Trans. Inf. Theory*, vol. 55, no. 2, pp. 620-643, Feb. 2009.

[64] X. Shang, G. Kramer, and B. Chen, "A new outer bound and the noisy-interference sum-rate capacity for Gaussian interference channels", *IEEE Trans. Inf. Theory*, vol. 55, no. 2, pp. 689-699, Feb. 2009.

[65] V.S. Annapureddy and V.V. Veeravalli, "Gaussian Interference Networks: Sum Capacity in the Low Interference Regime and New Outer Bounds on the Capacity Region", *Submitted to the IEEE Trans. Inf. Theory*, Feb. 2008.

[66] T. S. Han and K. Kobayashi, "A new achievable rate region for the interference channel, *IEEE Trans. Inf. Theory*, vol. 27, no. 1, pp. 4960, Jan. 1981.

[67] M. Costa, "Writing on a dirty paper", *IEEE Trans. Inf. Theory*, vol. 29, no. 3, pp. 439-441, May 1983.

[68] T. M. Cover, R. J. Mceliece and E.C. Posner, "Asynchronous multiple-access channel capacity", *IEEE Trans. Inf. Theory*, vol. 27, no. 4, pp. 409-413, July 1981.

[69] J. Y. N. Hui and P. A. Humblet, "The capacity region of the totally asynchronous multiple access channel", *IEEE Trans. Inf. Theory*, vol 31, no. 2, pp. 207-216, March 1985.

[70] A. El Gamal and Y. H. Kim, "Lecture notes on network information theory", Available online at http://arxiv.org/pdf/1001.3404.

[71] J. L. Massey and P. Mathys, "The collision channel without feedback", *IEEE Trans. Inf. Theory*, vol. 31, no. 2, pp. 192-204, March 1985.

[72] E. Calvo, J. R. Fonollosa and J. Vidal, "On the totally asynchronous interference channel with single-user receivers", *International Symp. Inf. Theory*, ISIT 2009, Seoul, Korea, June 2009.

[73] S. Verdú and T.S. Han, "A general formula for channel capacity", *IEEE Trans. Inf. Theory*, vol. 40, no. 7, pp. 1147-1157, July 1994.

[74] M. Rupf and J. L. Massey, "Optimum sequence multisets for synchronous code-division multiple-access channels", *IEEE Trans. Inf. Theory*, vol. 40, no. 4, pp. 1261-1266, July 1994.

[75] T. Liu and P. Viswanath, "An extremal inequality motivated by multi-terminal information-theoretic problems", *IEEE Trans. Inf. Theory*, vol. 53, no. 5, May 2007.

[76] A. Lapidoth, "Nearest neighbor decoding for additive non-Gaussian noise channels", *IEEE Trans. Inf. Theory*, vol. 42, no. 5, pp. 1520-1529, May 1996.

[77] S. N. Diggavi and T. M. Cover, "The worst additive noise under a covariance constraint", *IEEE Trans. Inf. Theory*, vol. 47, no. 7, pp. 3072-3081, Nov. 2001.

[78] K. Moshksar, A. Bayesteh and A. K. Khandani, "Coexistence in wireless decentralized networks; Part I: Multiplexing Gain", *Tech. Report* No. UW-ECE 2009-16, University of Waterloo, *Available online at http://shannon2.uwaterloo.ca/∼kmoshksa/Paper%20I_TR7.pdf.*

[79] K. Moshksar, A. Bayesteh and A. K. Khandani, "A model for randomized recourse allocation in decentralized wireless networks", *Accepted for publication in IEEE Trans. Inf. Theory*, 2010.

[80] T. M. Cover and J. A. Thomas, "Elements of information theory", *John Wiley and Sons, Inc.*, 1991.

[81] A. M. Tulino and S. Verdú, "Random matrix theory and wireless communications", *Foundations and Trends in Commun. and Inf. Theory*, vol. 1, issue 1, 2004.

[82] R.G. Gallager, "Information theory and reliable communication", John Wiley and Sons, Inc., 1968

[83] S. M. Kay, "Fundamentals of statistical signal processing, Vol. I: Estimation Theory", *Prentice-Hall*, 1993.

[84] S. M. Kay, "Fundamentals of statistical signal processing, Vol. II: Detection Theory", *Prentice-Hall*, 1998.

[85] B. Bollabás, "Linear analysis: An introductory course", *Cambridge University Press*, 1999.

[86] M. Abramowitz and I. A., Stegun, “Handbook of Mathematical Functions”, *Dover Publications, Inc.*, 1965.

[87] R. M. Gray, “Entropy and information theory”, *Springer-Verlag*, 2009.

[88] J. Wolfowitz, “Coding theorems of information theory”, *Springer-Verlag*, 1978.

[89] R. M. Dudley, “Real analysis and probability”, *Cambridge University Press*, 2002.

[90] A. W. van der Vaart, “Asymptotic statistics”, *Cambridge University Press*, 1998.