Fundamental limits of Cognitive Networks and Cooperation in Millimeter Wave Networks

BY

DIANA MAAMARI
B.S., University of Balamand, 2009
M.S., University of Balamand, 2011

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Electrical and Computer Engineering
in the Graduate College of the
University of Illinois at Chicago, 2015

Chicago, Illinois

Defense Committee:

Daniela Tuninetti, Chair and Advisor
Natasha Devroye, Advisor
Rashid Ansari
Milos Zefran
Hulya Seferoglu
George Calcev, Huawei Corporation
Besma Smida, Purdue University

ACKNOWLEDGMENTS

I can honestly express that my academic journey would not have been possible without the constant technical guidance, support and tremendous belief received by both my Ph.D advisors Prof. Daniela Tuninetti and Prof. Natasha Devroye. I am deeply grateful for the continuous help and generous collaboration received which have helped me grow both on a personal and academic level. I would like to thank my committee members, Prof. Milos Zefran, Prof. Rashid Ansari, Prof. Besma Smida, Prof. Hulya Seferoglu and Dr. George Calcev for providing their valuable time and through their presence, a real role model.

Words will never be enough to describe the everlasting and unconditional sacrifice and support made by my parents to allow me to pursue my dreams and deepen my thirst for knowledge, I am forever indebted to them. Finally, I would like to thank my husband, my sister and brother for being supportive throughout those years.

It has been wonderful working with all my current and former lab mates at UIC. A special acknowledgement to my lab partner Alex Dytso who introduced me to both Prof. Tuninetti and Prof. Devroye and inadvertently launching my Ph.D journey as I first joined UIC.

DM

CONTRIBUTION OF AUTHORS

The contents of this thesis is the result of a joint effort between my PhD advisors Prof. Daniela Tuninetti, Prof. Natasha Devroye and I. Our main contributions and a detailed list of prior existing work are presented in Chapter 1. The main results of our collaboration are found in Chapters 2, 3 and 4.

Motivated by the problem of spectrum scarcity, we consider different multi-user cognitive interference networks and characterize their information theoretic limits of communication. More precisely, we characterize the transmitters' simultaneous rates and throughput at which reliable communication is guaranteed. We make progress by finding the optimal power control policy when a cognitive user is sharing the spectrum with a licensed transmitter.

Cooperation in millimeter wave networks is further considered and we make progress by characterizing the coverage probability for a network with base station cooperation. The coverage probabilities derived account for blockage, interference and different fading distributions.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

CHAPTER **PAGE**

TABLE OF CONTENTS (Continued)

CHAPTER **PAGE**

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES (Continued)

FIGURE **PAGE**

SUMMARY

One of the most promising technologies for next-generation spectrally efficient wireless communications is cognitive radio. Cognitive radios are artificially intelligent devices that are aware of their environment and can coexist with licensed users in the same frequency band, thus revolutionizing the way spectrum is allocated and alleviating the problem of underutilized frequency spectrum.

The cognitive radio theory is still in the early stages of development, it is thus important to understand the fundamental limits of communication possible when certain nodes use cognitive radio technology. In particular, we are interested characterizing the limits of communication when cognitive transmitters have side information (messages) of other transmitters and to understand how to optimally use the available resources (power) when a cognitive transmitter is in a fading environment. Our main interest is characterizing the information theoretic limits of communication for a multi-user cognitive network and the optimal resource allocation policy for maximizing the sum-capacity in a fading environment.

To further address the problem of spectrum scarcity and to fulfill the need for increased bandwidth, millimeter wave bands between 30 and 300 GHz have been considered for future wireless mobile networks. Until relatively recently, it was presumed that these ultra high frequency bands are unreliable for cellular communications since signals cannot penetrate as far as those at lower frequency bands, thus resulting in signal outages. With the advancement in CMOS and the ability to pack large arrays of antennas, this problem can be alleviated.

Moreover, with the anticipated vision that mmWave networks would have a dense deployment of base stations, interference from strong line-of-sight base stations increases too, thus further increasing the probability of outage. Motivated by this challenge, this thesis addresses the problem of base station cooperation in millimeter wave networks as a way to decrease signal outage with the increase in the number of signal paths. The main emphasis is characterizing the coverage probability for a millimeter wave downlink heterogenous network with coordinated multipoint joint transmission.

This thesis is organized into 5 chapters: Chapter 1 presents the main contributions of our work with a detailed list of prior related work. Chapter 2 is devoted to demonstrating the optimal resource allocation (power allocation policy) for maximizing the sum-capacity of an ergodic fading Gaussian cognitive interference channel consisting of a primary user and a cognitive user communicating over the same frequency band. Chapter 3 contains the information theoretic limits for a multi-user cognitive interference network consisting of one primary and an arbitrary number of cognitive users with a "hierarchal" message knowledge structure and its relationship with a network with only one cognitive user. Chapter 4 focuses on characterizing the coverage probability for a downlink millimeter wave heterogenous network with base station cooperation, while Chapter 5 includes a summary of the results.

CHAPTER 1

INTRODUCTION

Almost all of the prime frequency bands have already been assigned to various wireless services, leading to the scarcity of available frequency spectrum for emerging wireless applications and products. However reports and statistics show that the wireless spectrum is in fact underutilized. These two facts suggest that if one is able to exploit the underutilized spectrum using advanced software radio, then this would help solve the problem of spectrum gridlock.

Cognitive radios are proposed as a potential solution to alleviate the problem of spectrum scarcity currently faced by the increasing demand for wireless broadband. Having cognitive radios that can dynamically access the spectrum will allow new wireless devices to intelligently use the available licensed wireless channels and coexist with users that hold the priority access to the spectrum.

The cognitive radios are classified according to the type of side information they acquire. Cognitive radios can sense white spaces (time, space or frequency void) and adjust their transmissions to fill the sensed voids, this approach has been referred to as *interweave* and has the advantage of interference avoidance. Contrary to keeping its transmission orthogonal to the primary user's transmissions, as is the case in the interweave paradigm, the *underlay* paradigm

allows a cognitive radio to simultaneously transmit with primary user(s), if the interference caused at the primary receiver(s) is kept below a certain threshold commonly referred to as the interference temperature

A cognitive radio with even a higher level of cognition would have additional information of codebooks, messages and/or channel gains. When simultaneously transmitting with primary license holders it is referred to as the *overlay* cognitive interference channel [7]. We focus on the latter form of cognition where secondary users have a-priori non-causal message knowledge of primary license holders. Intuitively, this idealized assumption of message knowledge allows the cognitive radio to either cooperate in sending this message of the primary user and/or transmit its own message using some interference mitigating method.

Our main focus of interest is characterizing the information theoretic limits of communication of overlay cognitive networks with *arbitrary* number of secondary / cognitive users (having non-causal message knowledge) coexisting with a primary user in the same frequency band and thus interfering with one another; moreover additive white Gaussian noise is included in the channel model to account for the natural random processes that occur in nature.

The basic limit of communication for any channel is the *channel capacity.* The channel capacity is the maximum rate of communication for which arbitrary small probability of error can be achieved. In this work, we are interested in finding the highest sum of rates (*sum-capacity*) that can be simultaneously achieved at the receivers with small probability of decoding error. In many cases finding the sum-capacity of a network is not an easy problem; therefore it is helpful to study first the asymptotic approximations of the sum-capacity for example finding

the *generalized degrees of freedom*. The generalized degrees of freedom is a useful metric that approximates the interference-limited sum-capacity performance at high SNR and is a metric we seek to derive for the multi-user cognitive interference channel.

Moreover we are interested in understanding the behavior of a cognitive transmitter in a time-varying fading environment. The *ergodic channel capacity* is the capacity at which reliable communication is possible while averaging out the time variation of the channel and is the performance metric we seek to derive.

To further address the possibility of breaking the spectrum gridlock, millimeter wave communication has been proposed for next generation cellular networks. It was presumed that these ultra high frequency bands between 30 and 300 GHz are unreliable for cellular communication, since signals are susceptible to blockage, absorption, diffraction and penetration resulting in outages and thus unreliable communication. Therefore, it is expected that mmWave networks would likely be dense to attain reliable coverage. However, a dense mmWave network further increases the number of interfering base stations too thus counteracting its initial purpose. To address this issue, we explore the benefits of base station cooperation in the downlink of a mmWave heterogenous network. Using tools from stochastic geometry, we show that cooperation from randomly located base stations within different tiers can decrease the probability of outage by increasing the number of signal paths (and hence diversity) a receiver sees. For this problem, we are interested in characterizing the coverage probability, in other words, the probability of having the received power above a target minimum for which performance is acceptable.

1.1 Contributions

The single antenna overlay two-user cognitive radio channel, the 2-CIFC, first introduced in [8], models the communication in a network between a licensed user (has the exclusive right to transmit) and a secondary user that has a-priori the primary user's codebook as well as its message. The cognitive transmitter is able to transmit simultaneously over the same channel as apposed to limiting its transmissions to white spaces. With these idealized side information assumptions, the secondary transmitter can act selfishly and pre-cancel the effect of the primary's interference and transmit its own independent message (dirty paper coding) or it can act selflessly and behave as a relay (beam forming).

Ever since its introduction in [8], this two-user channel has been extensively studied. In particular [9] and [10] determine the sum-capacity in weak interference, while [11] characterizes the capacity under strong interference conditions. The most comprehensive results on the two-user cognitive interference channel (2-CIFC) can be found in [12,13], where [12] gives the largest achievable rate region and [13] gives capacity results for the Gaussian Noise channel.

We remove the assumption of constant channel gains and consider a more realistic fading (time varying) cognitive interference channel. To the best of our knowledge, there has been no work regarding a fading cognitive interference channel. For this time varying channel model, we are interested in characterizing the power allocation policy that maximizes its sum-capacity.

The literature on the fundamental performance of *multi-user* cognitive interference channels is limited, in part due to the fact that the two-user counterpart is not yet fully understood [12, 13]. Several 3-user extensions of the 2-CIFC model have been studied in the lit-

erature. For the 3-user case, the 3-CIFC, the work [14] proposed different overlay cognitive interference channel models that differ by the cognitive abilities at the transmitters. Achievable rate regions have been proposed in [15]. The work in [16, 17] considered the three user cognitive interference channel with one cognitive transmitter and provided the capacity region under strong interference conditions. To the best of our knowledge, capacity results for the fully connected multi-user cognitive networks have not yet been reported.

In this work, we are interested in characterizing the capacity for one particular message knowledge structure for a multi-user setting which consists of a primary user sharing the spectrum with arbitrary number of cognitive users with *cumulative message sharing*. Surprisingly, we show - by deriving fundamental information theoretic limits - that under certain channel gain conditions having "distributed cognition" or having a cumulative message knowledge structure at nodes may not be worth the overhead as (approximately) the same sum-capacity can be achieved by having only one "global cognitive" user whose role is to manage all the interference in the network.

In the following section we list the main contributions regarding the EGCIFC and the K-user cognitive interference channel with cumulative message sharing.

1.1.1 Contributions: Ergodic Fading Gaussian Cognitive Interference Channels

We study a time varying two user cognitive interference channel - *ergodic fading Gaussian cognitive interference channel* (EGCIFC). The main contributions regarding this time varying channel are:

1. Sum-capacity genie-aided outer bound for the EGCIFC. This genie-aided outer bound consists of giving side-information to the cognitive receiver. The side information consists of primary user's message and output. We then provide a matching achievability scheme, thus completely characterizing the sum-capacity under the assumption of perfect CSI at all transmitters and receivers.

2. When the primary receiver experiences strong interference, the sum-capacity achieving scheme is that of MISO channel with PerPC and perfect CSI at the transmitter and receiver. As a result of independent interest, we thus also characterize the sum-capacity of a MISO channel with full CSI at all terminals, and with an arbitrary number of transmit antennas.

3. The power allocation policy that maximizes the sum-capacity for the EGCIFC is derived which depends on the relative channel gains for a given fading state.

1.1.2 Contributions: Multi-user Cognitive Interference Channels

In the second part of this work, we are interested in the information theoretic limits of communication when arbitrary number of cognitive users - with different amount of side information (messages) of other transmitters - communicate over the same frequency band with a primary user, thus interfering with one another. We term this network the K *user cognitive interference channel with cumulative message sharing* and we are interested in understanding the minimum message knowledge needed at the transmitters that is *sufficient* for optimal sum-capacity performance. The assumption made for this channel is that cognitive transmitter $\mathsf{i} \in [2 : \mathsf{K}]$ has

non-causal knowledge of the messages of users with index less than i. The main contributions for the K-*user cognitive interference channel with cumulative message sharing* are:

1. A novel and general outer bound region. The bound is valid for any memoryless channel and any number of users. The bound does not contain auxiliary random variables and is therefore computable for many channels of interest, including the Gaussian channel.

2. The sum-capacity the 3-user Linear Deterministic Approximation of the Gaussian noise channel at high-SNR for any channel parameters is characterized. This optimal scheme inspires a scheme for the K-user symmetric channel. This latter scheme only requires cognition of all messages at one transmitter; all the others need only knowledge of their own message.

3. The sum-capacity for the symmetric Gaussian noise channel with K users to within a constant additive and multiplicative gap is derived. The additive gap is a function of the number of users and grows as $(\mathsf{K}-2)\log_2(\mathsf{K}-2)$. The proposed achievable scheme is based on Dirty Paper Coding (DPC) and may be thought of as a Gaussian MIMO-broadcast channel scheme where only one encoding order is possible due to the cumulative message sharing mechanism. As opposed to other multi-user interference channel models, a single scheme suffices for both the weak and strong interference regimes. Moreover, no interference alignment of structured coding seems to be needed. Numerical evaluations show that the actual gap is less than the analytical one; this is so because of necessary crude bounding steps needed to obtain analytically tractable sum-rate expressions. The multiplicative gap is K and is achieved by having all users beam form to the primary user.

4. The normalized generalized degrees of freedom (gDoF), defined as the pre-log of the sum-capacity as a function of SNR, is shown to be a function of K. This is in contrast with other channel models, like the non-cognitive case or the broadcast channel, where the gDoF are the same for any K. Interestingly, it is shown that as the number of users grows to infinity the gDoF of the K-user cognitive interference channel with CMS tends to the gDoF of a broadcast channel with a K-antenna transmitter and K single-antenna receivers.

   We then consider a K-user cognitive interference channel with two different message sharing structures in strong interference. In addition to the cumulative message sharing (CMS) network we consider the cognitive only message sharing (CoMS) cognitive interference channel - which consists of $\mathsf{K}-1$ primary users and only one cognitive transmitter that has all the messages of the primary users. The contributions regarding these two channels models experiencing strong interference are:

5. A sum-rate outer bound valid for any number of users under a certain *strong interference condition*, which amounts to having one receiver that can decode all transmitted signals without loss of optimality. The bound takes the same form for both CoMS and CMS models, but over different sets of input distributions. The bound does not contain auxiliary random variables and is therefore computable for many channels of interest, including the Gaussian channel. The bound is not the classical compound MAC result of similar strong interference capacity results.

6. We present coding schemes that achieve the sum-rate outer bound for both CoMS and CMS in strong interference.

7. For the Gaussian noise channel, we explicitly characterize the set of channel gains satisfying the strong interference condition and compare our results with [18] (for the 2-user case) and [16, 17] (for the 3-user case). Since we only focus on sum-capacity, our outer bound holds under more relaxed conditions than [16,17]. Moreover, our approach extends beyond the 3-user Gaussian case to any number of users and to any memoryless channel.

8. For the Gaussian noise channel, we show that under weak channel gain conditions, it is sufficient to have one global cognitive user in a K-user network, which knows all the message of the network, to achieve an upper bound of a network derived initially for the network with cumulative message sharing.

1.1.3 Contributions: Coverage for Base station Cooperation in Millimeter Wave Heterogenous Networks with Blockage

Characterizing coverage and capacity in mmWave cellular systems and in coordinated multipoint (CoMP) networks have been extensively studied. In [19] the authors compared the performance, in terms of coverage and capacity, of a stochastic geometry based mmWave network (without CoMP) to a microwave cellular network, at a single antenna receiver (typical user). In [19], directionality at the transmitters, intra-cell and inter-cell interference were accounted for but blockage was not included in the analysis. The authors show that coverage in mmWave systems increases with the decrease in the half-power beam width of the radiation pattern. In fact, having narrower beams decreases beam overlap, thus decreasing intra-cell and inter-cell interference and increasing coverage probability. We propose to study the problem of base station cooperation in the downlink of dense mmWave heterogenous network as a mean to

combat blockage and decrease signal outage. Our derivations of the coverage probability, similarly to [19], account for interference experienced at the typical user, but in addition blockage is incorporated in the analysis.

In [20] (see also journal version in [21]) the authors proposed a stochastic geometry framework to evaluate the performance of mmWave cellular networks (without CoMP) with blockage. The authors incorporate blockage by modeling the probability of a communication link - being either a line-of-sight (LOS) or non-LOS (NLOS) link - as function of the length of the communication link from the serving base station. Different pathloss laws were applied to the LOS and NLOS links. Interestingly, it was shown in [21] that the LOS region (a region where a user does not experience blockage) was approximated by an equivalent LOS ball whose radius is chosen such that the average number of base stations observed in the LOS ball is the same as that observed in the mmWave network. Numerical results in [21] suggest that higher data rates can be achieved when compared to microwave cellular networks. One of the interesting observations made in [20] is that mmWave networks should be dense but not too dense - since the number of LOS interfering base stations increases when the density of base stations increases. We willl leverage results from [21] to incorporate blockage and differentiate between having LOS links and NLOS links from the cooperating base stations in the analysis of the problem of joint transmission in mmWave networks.

In [22] (see also journal version in [23]) the authors used stochastic geometry for studying microwave joint transmission CoMP where single antenna base stations transmit the same data to single antenna users. Different performance metrics (including coverage probability) were

considered, to evaluate the performance at the typical user located at an arbitrary location (general user) and receiving data from base stations with the strongest average received power. The authors showed that in the interference-limited regime, the coverage probability for the typical user becomes independent of the number, density and respective power levels of the tiers present in a CoMP network since a variation in any of the listed parameters causes a change in both desired signal power as well as a change in the interference power in such a way that their ratio remains constant in the absence of noise.

The derivation of the coverage probability for a mmWave network with base station cooperation in this work is similar to that in [22] for the general user, except that key factors specific to the mmWave channel model have to incorporated, some of which are the high directional transmission at the base stations, blockage and improved fading distribution due to beamsteering and sparse scattering.

We propose to study the benefits of base station cooperation in the downlink of a heterogenous mmWave cellular system as a mean to address the problem of signal outage due to blockage. We anticipate that the probability of a user being in outage due to blockage and interference will decrease with the increase in diversity / number of signal paths from different cooperating base stations. Our extensive numerical examples show that this is in fact the case for the following scenarios, Case 1) for dense mmWave networks where the number of interfering LOS base stations increases and Case 2) when the mmWave network experiences no blockage but beam steering is still accounted for at the base stations. We also provide examples when cooperation does not provide substantial increase in coverage probability.

Our numerical results account for different number of transmit antennas, different tier intensities (cell radius) and different blockage parameters. We compare the derived coverage probabilities under different scenarios. The case with and without base station cooperation is considered in the numerical examples. Moreover, we consider the case when there is no cooperation but the base station has a transmit power equal to the sum of individual transmit power if $\mathfrak{n}$ base stations were to jointly transmit. We also compare a mmWave network with blockage when the typical user is subject to interference versus the case when there is no interference. The case when the fading distribution from the cooperating base stations is Nakagami fading is also considered and two upper bounds on the coverage probability are derived.

1.2 Motivation

Before proceeding with the chapters, we list some of the many network topologies which can be modeled as an overlay cognitive interference channel:

1. Channels with Retransmission: Consider a network of primary and cognitive users in which a fixed-rate, i.e., not a function of the channel gains, primary user's message was sent but was not decoded at the intended receiver. The primary receiver informs the primary transmitter by sending a NACK, and a retransmission takes place. If the cognitive transmitter overheard the primary's initial transmission and was able to successfully decode the message in the first round then, in the retransmission phase, the cognitive transmitter would have non-causal primary message knowledge and could transmit together with the primary in the ARQ round(s) [24].

2. Coordinated Multipoint Transmission: Consider a network in which multiple transmitters (may be thought of as base-stations) are connected by high capacity back-haul links, allowing them to perfectly exchange messages to be transmitted to the receivers (may be thought of as mobile users). This model may include the cognitive interference channel model with asymmetric non-causal message knowledge as a special case; studying such networks with high-capacity backhaul may reveal what type of message knowledge structure is useful [25].

3. Overlay Networks: This model is inspired by the idea of layered cognitive networks: the first layer consists of primary users and each additional layer consists of cognitive users that share the same spectrum. Each additional layer is given the codebook(s) of all previous layers. This hierarchical codebook knowledge enables them to causally learn the lower layers' messages and aid in their transmission. Thus studying the non-causal message knowledge setting provides an upper bound to the more realistic case where the messages are causally learned by the cognitive transmitter(s) [2].

1.3 Notations

Throughout Chapter 2 and Chapter 3 we adopt the following notation convention:

- $A^{\dagger}$ Transpose and complex-conjugate of the matrix A
- $A^{\star}$ Optimal solution for a given optimization problem
- A^{N} Vector of random variables A_i with $i \in [1 : N]$
- $\mathbb{E}[.]$ Expectation operator

- $\log(.)$ Logarithm in base 2
- $I(.;.)$ Mutual information
- $h(.)$ Differential entropy
- $\mathbb{P}(A)$ Probability of event A
- $\mathcal{N}(m, \sigma^2)$ Additive white Gaussian Noise with mean m and variance σ^2.
- $\mathbf{I}_j$ Identity matrix of dimension j
- $[n_1 : n_2]$ Set of integers from n_1 to n_2

Notations used in Chapter 4 are summarized in tables and are in given in Chapter 4.

CHAPTER 2

ERGODIC FADING COGNITIVE INTERFERENCE CHANNEL

There are two important aspects that make wireless communication challenging. The first is *fading*, that is, the time variation of the channel gains due to small scale effects, such as those due to multi-path, and large-scale effects, such as those due to path loss and shadowing. The second is *interference* between wireless users communicating over the same frequency band. In this work, we focus on the two-user fading interference channel where one of the transmitters is cognitive, or knows the message of the other independent transmitter. This channel model thus experiences both fading and interference, as well as a third phenomena seen in wireless communications: the ability of transmit nodes to *cooperate*. We aim to characterize the sum-capacity / throughput of this channel so as to highlight the impact of fading, interference and asymmetric cooperation between transmitters.

Cognitive networks are wireless networks in which certain nodes are cognitive radios, or artificially intelligent devices, and have been the subject of intensive investigation by the wireless communication community in the past decade. In a cognitive network, cognitive / secondary transmitters share the spectrum with primary / licensed users. The cognitive devices exploit *side information* about their environment to improve spectral management. Depending on the

nature of the side information [7], cognitive users either search for unused spectrum (interweave), or operate simultaneously with non-cognitive transmitters as long as the interference produced is within an acceptable threshold (underlay), or relay part of the primary user's message and cancel interference through advanced encoding schemes (overlay). We consider the overlay paradigm where a primary transmitter-receiver (Tx-Rx) pair share the same spectrum with a cognitive Tx-Rx pair. According to the overlay paradigm [7], the secondary Tx is assumed to have non-causal knowledge of the message of the primary Tx.

The channel between the Txs and the Rxs is assumed to experience ergodic fading (i.e., time average of every sufficiently long fading realization equals the statistical average). All nodes in the network are assumed to have perfect instantaneous knowledge of the channel fading coefficients, or full channel state information (CSI). Although the full CSI assumption might be impractical even in the presence of a dedicated feedback channel from the Rxs to the Txs, the resulting model serves as the customary first step towards understanding the performance of more realistic models. Under these assumptions, we seek to determine the sum-capacity of the network and the corresponding optimal power allocation policy under a long-term average transmit power constraint at the Txs. In doing so, we also seek to answer the question of whether coding separately across fading states is optimal [26].

2.1 Prior Work

In [27] the two-user Gaussian interference channel with ergodic fading was introduced and the optimal power allocation policy that maximizes the outer bound was investigated. In [26] the authors also considered the same model and showed that in general joint encoding and

decoding across fading states is necessary to achieve capacity when perfect CSI at all nodes is assumed (note that fading is an example of the more general result that parallel interference channels are not separable [28]). However, [26] showed that in some parameter regimes, such as very strong or very weak interference, separate encoding and decoding across fading states is optimal. In other words, in interference channels separability may hold for certain channel states but not for all channel states. While interference channels are not separable in general, it is known that multi-access [29] and broadcast [30] channels with the same CSI assumption are separable. Since the EGCIFC has elements of *both* an interference channel and a broadcast channel, it is not a priori obvious that a separable scheme is capacity achieving. A formal proof of the optimality of separability for the EGCIFC is shown in this work.

In [31] the authors also consider a fading cognitive network under the underlay paradigm. The primary user in this case is completely oblivious to the existence of the cognitive user. The relationship between the achievable capacity of the secondary channel and the interference caused at the primary receiver was quantified. Instantaneous and average interference power constraints were both considered and the optimal power allocation policy for the secondary user in each case was derived. In this work, we consider the overlay paradigm with long-term average power constraints at the primary and secondary transmitters. Moreover, our primary user is not oblivious to the existence of the secondary user.

In [32] the authors consider a cognitive radio network under the underlay paradigm where primary and secondary users are subject to block fading. The primary user is not capable of adapting its power allocation while the secondary user is able to do so. The authors derive the

optimal power allocation strategies for the cognitive user to maximize its ergodic and outage capacity. In our work *both* users are capable of adapting their transmit power over the different fading states based on the channel state information, and we consider the overlay paradigm.

In [33] the authors consider the fading cognitive single input single output (SISO) MAC channel for the underlay paradigm, where the secondary users are subject to both a transmit power constraint and interference power constraint to primary users. It is shown that the sum-capacity achieving power allocation policy is a water filling type of solution. The key difference between the EGCIFC and the model in [33] is the message structure; in [33] each sender only knows its own message and thus the transmit signals are independent; in the EGCIFC the transmit signals are correlated due to the non-causal primary message knowledge at the secondary transmitter. Because of this, the sum-capacity achieving power allocation policy for the EGCIFC will not be a water filling type.

As a byproduct of our analysis and as a result of independent interest, we shall show that the optimal power policy for the EGCIFC in some regimes requires both transmitters to beam-form to the primary receiver; in this case the model reduces to a point-to-point multiple input single output (MISO) channel with per-antenna power constraints (PerPC). In [34] the author finds the capacity for the point-to-point MISO channel with PerPC with two different assumptions of CSI. The capacity for a constant channel with CSI at both terminals and with the assumption that Rayleigh is the fading distribution channel with CSI at the receiver only were derived. The author compares the result with the capacity of a MISO channel with sum power constraints (SumPC) through numerical examples. In both cases the capacity with PerPC is, as expected,

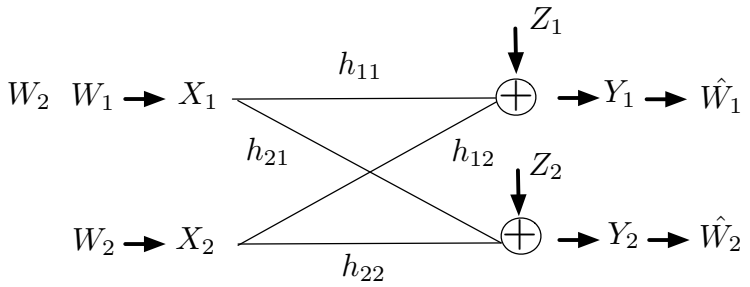


Figure 1. Ergodic Fading Gaussian Cognitive Interference Channel (EGCIFC).

less than that with SumPC. Here we characterize the sum-capacity of the EGCIFC with CSI at the transmitters and receivers and find the optimal signaling scheme, which is thus different from the solutions found in [34].

In [8] the authors first introduced the information theoretic study of the cognitive radio channel (same as the cognitive interference channel) which falls into the overlay paradigm. In that work, the channel gains were constant and achievable rate regions and outer bounds were derived. In [9] the authors found the capacity of the cognitive interference channel in the weak interference regime. In particular, the power split which ensures that the primary receiver rate continues to be the same as that without interference from the cognitive user. For the state-of-the-art on the two-user cognitive interference channel with constant channel gains (known to all nodes) we refer the reader to [12, 13]. We remove the assumption of constant channel gains and consider the fading (time varying) cognitive interference channel.

2.2 Channel Model

The cognitive interference channel consists of two transmit-receive pairs Tx_1 to Rx_1 and Tx_2 to Rx_2 representing the cognitive and primary users, respectively, as shown in Figure 1. Each transmitter Tx_k wishes to convey to its destination Rx_k an independent message W_k, which is uniformly distributed over the set $[1 : 2^{NR_k}]$, where R_k is the rate in bits per channel use, and N represents the codeword length, for $k \in [1 : 2]$. Tx_1 is *cognitive* in the sense that it has non-causal message knowledge of the primary user's (Tx_2) message W_2. A rate vector (R_1, R_2) is said to be achievable if there exists a family of codes indexed by N such that the probability of decoding error can be made arbitrarily small [35]. The sum-capacity is defined as the maximum achievable $R_1 + R_2$.

In Gaussian noise and with ergodic fading, the ECIFC input-output relationship at every time instant (time index is omitted for easier notation) is given by

$$Y_1 = h_{11}X_1 + h_{12}X_2 + Z_1, \quad Z_1 \sim \mathcal{N}(0, 1), \tag{2.1}$$

$$Y_2 = h_{22}X_2 + h_{21}X_1 + Z_2, \quad Z_2 \sim \mathcal{N}(0, 1), \tag{2.2}$$

where $\mathbf{H} := \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix}$ denotes the random channel gain matrix (with complex entries generated randomly at each time instant / channel use according to a known, stationary and ergodic random process), with $[\mathbf{H}]_{i,j} = h_{ij} \in \mathbb{C},\ i, j \in [1 : 2]$ representing the fading channel gain between Tx_j and Rx_i. A realization of $\mathbf{H}$ is indicated as $\mathbf{h}$. Tx_j, with channel input X_j is subject to the long-term average power constraint $\mathbb{E}[|X_j|^2] \leq \overline{P}_j,\ j \in [1 : 2]$. With CSI

at all terminals, the transmitters can perform dynamic power allocation, and transmit with power $P_j(\mathbf{h}) \geq 0,\ j \in [1:2]$, at a channel use with fading state $\mathbf{h}$. We seek to determine the sum-capacity optimal power allocation for each user such that $\mathbb{E}[P_j(\mathbf{h})] \leq \overline{P}_j,\ j \in [1:2]$.

Thanks to cognition, the channel inputs can be correlated; the input covariance matrix in fading state $\mathbf{h}$ is denoted by

$$\Sigma_x(\mathbf{h}) := \begin{bmatrix} P_1(\mathbf{h}) & \rho^\dagger(\mathbf{h})\sqrt{P_1(\mathbf{h})P_2(\mathbf{h})} \\ \rho(\mathbf{h})\sqrt{P_1(\mathbf{h})P_2(\mathbf{h})} & P_2(\mathbf{h}) \end{bmatrix}, \tag{2.3}$$

where the correlation coefficient must satisfy $|\rho(h)| \leq 1$.

2.3 Main Results

This section includes the main results for the EGCIFC. A sum-capacity outer bound is presented as a maximization problem over three constraints: the two long-term average power constraints at the transmitters and the constraint on the correlation coefficient. We then prove that the outer bound with the optimal power allocation policy is achievable through a variation of the achievability scheme of [9] proposed for the constant channel gain cognitive interference channel in weak interference. The achievability scheme, unlike that for certain parameter regimes of the interference channel, is separable, that is, encoding need not be done across fading states. For a certain parameter regime (relative channel gains), the sum-capacity achieving power allocation corresponds to the optimal power allocation scheme for a point-to-point MISO channel with a PerPC. Thus as a topic of independent interest we characterize

the power allocation policy and the capacity of the fading point-to-point MISO channel with PerPC and an arbitrary number of antennas.

The sum-capacity of the EGCIFC is presented next and is shown to be the solution of a maximization problem in the variables $(P_1(\mathbf{h}), P_2(\mathbf{h}), \rho(\mathbf{h}))$ in (Equation 2.3).

Theorem 2.3.1. *The ergodic sum-capacity of the EGCIFC is*

$$C_{\text{EGCIFC,sum}} = \max \mathbb{E}\Bigg[\log\left(1 + |h_{21}|^2 P_1(\mathbf{h}) + |h_{22}|^2 P_2(\mathbf{h}) + 2|\rho(\mathbf{h})|\sqrt{|h_{21}|^2 P_1(\mathbf{h})\ |h_{22}|^2 P_2(\mathbf{h})}\right) + \log\left(\frac{1 + (1-|\rho(\mathbf{h})|^2)\max\{|h_{11}|^2, |h_{21}|^2\}P_1(\mathbf{h})}{1 + (1-|\rho(\mathbf{h})|^2)|h_{21}|^2 P_1(\mathbf{h})}\right)\Bigg], \tag{2.4}$$

where the maximization is over $P_i(\mathbf{h}) \geq 0 : \mathbb{E}[P_i(\mathbf{h})] \leq \overline{P}_i, i \in [1:2],$ *and* $|\rho(\mathbf{h})| \leq 1$.

Proof. As a generalization of the outer bound technique of [12, 13] to the fading case, the sum-capacity is upper-bounded by

$$\begin{aligned}
N(R_1 + R_2 - 2\epsilon_N) &\overset{(a)}{\leq} I(W_1; Y_1^N|\mathbf{h}^N) + I(W_2; Y_2^N|\mathbf{h}^N) \overset{(b)}{\leq} I(W_1; Y_1^N, Y_2^N|\mathbf{h}^N, W_2) + I(W_2; Y_2^N|\mathbf{h}^N) \\
&\overset{(c)}{=} I(W_1, Y_1^N|\mathbf{h}^N, W_2, Y_2^N) + I(W_1; Y_2^N|W_2, \mathbf{h}^N) + I(W_2; Y_2^N|\mathbf{h}^N) \\
&\overset{(d)}{=} I(W_1; Y_1^N|\mathbf{h}^N, W_2, Y_2^N) + I(W_1, W_2; Y_2^N|\mathbf{h}^N) \\
&\overset{(e)}{\leq} I(X_1^N; Y_1^N|\mathbf{h}^N, X_2^N, Y_2^N) + I(X_1^N, X_2^N; Y_2^N|\mathbf{h}^N) \\
&\overset{(f)}{=} h(Y_1^N|\mathbf{h}^N, X_2^N, Y_2^N) - h(Y_1^N|\mathbf{h}^N, X_2^N, Y_2^N, X_1^N) + h(Y_2^N|\mathbf{h}^N) - h(Y_2^N|\mathbf{h}^N, X_1^N, X_2^N) \\
&\overset{(g)}{=} \sum_{i=1}^{N} h(Y_{1i}|\mathbf{h}^N, X_2^N, Y_2^N, (Y_1)_1^{i-1}) - h(Y_{1i}|\mathbf{h}^N, X_2^N, Y_2^N, X_1^N, (Y_1)_1^{i-1})
\end{aligned}$$

$$
\begin{aligned}
&+h(Y_{2i}|\mathbf{h}^N,(Y_2)_1^{i-1}) - h(Y_{2i}|\mathbf{h}^N,X_1^N,X_2^N,(Y_2)_1^{i-1}) \\
&\overset{(h)}{\leq} \sum_{i=1}^{N} h(Y_{1i}|\mathbf{h}_i,X_{2i},Y_{2i}) - h(Y_{1i}|\mathbf{h}_i,X_{2i},Y_{2i},X_{1i}) + h(Y_{2i}|\mathbf{h}_i) - h(Y_{2i}|\mathbf{h}_i,X_{1i},X_{2i}) \\
&\overset{(i)}{=} \sum_{i=1}^{N} I(X_{1i};Y_{1i}|X_{2i},Y_{2i},\mathbf{h}_i) + I(X_{1i},X_{2i};Y_{2i}|\mathbf{h}_i) \\
&\overset{(j)}{\leq} \sum_{i=1}^{N} I(X_{1Gi};Y_{1i}|Y_{2i},X_{2Gi},\mathbf{h}_i) + I(X_{1Gi},X_{2Gi};Y_{2i}|\mathbf{h}_i),
\end{aligned}
$$

where the different inequalities follow from: (a) Fano's inequality (here $\epsilon_N \to 0$ as $N \to \infty$), (b) a genie provides side information (Y_2^N, W_2) to Rx_1 and independence of messages, (c) chain rule for mutual information, (d) recombining mutual information terms, (e) data processing inequality and definition of encoding functions, (f) definition of mutual information, (g) chain rule for entropy, (h) conditioning reduces entropy and memoryless channel, (i) definition of mutual information, and (j) Gaussian maximizes entropy, where X_{1Gi}, X_{2Gi} are jointly Gaussian with the same covariance matrix as X_{1i}, X_{2i}. The dependence on the time index i can be eliminated by taking the appropriate limit over N as done in [26]. We note that in (g) we can choose the correlation coefficient among Z_1 and Z_2 since the Rxs do not cooperate and hence the capacity region only depends on the noise marginal distributions (i.e., we can choose the worst noise correlation as long as the marginal distributions are preserved). From the results on the static channel [13] we know that the worst noise correlation is $\min\left\{\frac{|h_{11}|}{|h_{21}|}, \frac{|h_{21}|}{|h_{11}|}\right\}$. With this worst noise correlation and with the input covariance as in (Equation 2.3), the sum-capacity outer bound in (g) for the EGCIFC can be expressed as in (Equation 2.4).

We now demonstrate that the derived outer bound is achievable. A variable rate coding scheme is used: in this case at each channel use (each coordinate of the codewords X_1^N and X_2^N), the optimal powers for that particular fading state are used $P_1^\star(\mathbf{h})$ and $P_2^\star(\mathbf{h})$ by the secondary and primary nodes, respectively with an optimal $\rho^\star(\mathbf{h})$. The cognitive transmitter assigns part of its power to relay W_2 and uses the remaining power to send its own message by Dirty Paper Coding (DPC) [36] against W_2, which it knows non-causally. This is similar to the scheme for the static channel [9] with the difference that at each channel use, the optimal parameters $P_1^\star(\mathbf{h})$, $P_2^\star(\mathbf{h})$ and $\rho^\star(\mathbf{h})$ for that particular fading state that maximize (Equation 2.4) are used. In particular, let U_1 and U_2 be independent Gaussian random variables with zero mean and unit variance, and

- Primary user sends

$$X_2 = \sqrt{P_2^\star(\mathbf{h})}U_2, \tag{2.5a}$$

- Cognitive user sends $X_1 = X_{2R} + X_{1DPC}$ where

$$X_{2R} = \sqrt{|\rho^\star(\mathbf{h})|^2 P_1^\star(\mathbf{h})}e^{j(-\angle h_{21}+\angle h_{22})}U_2, \tag{2.5b}$$

$$X_{1DPC} = \sqrt{(1-|\rho^\star(\mathbf{h})|^2)P_1^\star(\mathbf{h})}U_1, \tag{2.5c}$$

and where $X_{1\mathrm{DPC}}$ is DPC against the 'non-causally known state'

$$S = \left(h_{12}\sqrt{P_2^\star(\mathbf{h})} + h_{11}\sqrt{|\rho^\star(\mathbf{h})|^2 P_1^\star(\mathbf{h})}e^{j(-\angle h_{21}+\angle h_{22})} \right) U_2. \tag{2.5d}$$

- The received signals are

$$Y_1 = h_{11}\sqrt{(1-|\rho^\star(\mathbf{h})|^2)P_1^\star(\mathbf{h})}U_1 + S + Z_1, \tag{2.5e}$$

$$Y_2 = e^{j\angle h_{22}}\left(\sqrt{|\rho^\star(\mathbf{h})|^2 P_1^\star(\mathbf{h})}|h_{21}| + \sqrt{P_2^*(\mathbf{h})}|h_{22}| \right) U_2 + h_{21}\sqrt{(1-|\rho^\star(\mathbf{h})|^2)P_1^\star(\mathbf{h})}U_1 + Z_2. \tag{2.5f}$$

- Since Tx_1 used DPC we have

$$R_1 = \log\left(1 + |h_{11}|^2(1-|\rho^\star(\mathbf{h})|^2)P_1^\star(\mathbf{h})\right), \tag{2.5g}$$

and if Rx_2 treats U_1 as noise we have

$$R_2 = \log\left(1 + \frac{(\sqrt{|h_{21}|^2|\rho^\star(\mathbf{h})|^2 P_1^\star(\mathbf{h})} + \sqrt{|h_{22}|^2 P_2^\star(\mathbf{h})})^2}{1+|h_{21}|^2(1-|\rho^\star(\mathbf{h})|^2)P_1^\star(\mathbf{h})}\right). \tag{2.5h}$$

By summing (Equation 2.5g) and (Equation 2.5h), re-arranging and taking the expectation yields the sum-rate in (Equation 2.4) when $|h_{11}|^2 \geq |h_{21}|^2$. When $|h_{11}|^2 < |h_{21}|^2$ the sum-rate in (Equation 2.4) in maximized by $|\rho^\star(\mathbf{h})| = 1$ and is again achievable by (Equation 2.5g) and (Equation 2.5h). □

Remark 1: One can think of parallel Gaussian cognitive interference channels (PGCIFC) as an EGCIFC in which each sub-channel occurs with equal probability and so the sum-capacity result described in this work gives also the sum-capacity for PGCIFC.

While Theorem 2.3.1 expresses the ergodic sum-capacity as an optimization problem, we now proceed to determine the optimal power allocation policy. To do so, we first investigate a topic of independent interest: the ergodic capacity of the point-to-point MISO channel with PerPC, which gives the sum-capacity of the EGCIFC when $|h_{11}|^2 < |h_{21}|^2$.

2.3.1 The ergodic capacity of the point-to-point MISO channel with per-antenna power constraints and perfect CSI at all terminals

The MISO channel with n transmit antennas with PerPC has output

$$Y = [H_1 \; H_2 \cdots H_n] \; \mathbf{X} + Z \in \mathbb{C}, \quad Z \sim \mathcal{N}(0, 1),$$

where each entry of the input vector $\mathbf{X} := [X_1, \cdots, X_n]^T$ has a separate long-term average transmit power constraint $\mathbb{E}[|X_i|^2] \leq \overline{P}_i$, for $i \in [1 : n]$. The channel vector $[H_1 \; H_2 \cdots H_n]$ has complex-valued entries representing the channel gain coefficient from each transmit antenna to the receive antenna and is generated from an ergodic process whose instantaneous realization is known to the transmitter and the receiver. We aim to characterize the ergodic capacity of this channel, where capacity is defined as usual [35]. In the following, we denote the instantaneous realization of the channel vector as $\mathbf{h} := [h_1 \; \cdots \; h_n] \in \mathbb{C}^n$ and the power allocated on antenna i in fading realization $\mathbf{h}$ as $P_i(\mathbf{h}), i \in [1 : n]$.

Theorem 2.3.2. *The ergodic capacity of the Gaussian fading MISO channel with PerPC is*

$$C_{\text{MISOPerPC}} = \max\ \mathbb{E}\left[\log\left(1+\Big(\sum_{i\in[1:n]}|h_i|\sqrt{P_i(\mathbf{h})}\Big)^2\right)\right] \tag{2.6}$$

where the maximization in (Equation 2.6) *is over* $P_i(\mathbf{h}) \geq 0 : \mathbb{E}[P_i(\mathbf{h})] \leq \overline{P}_i,\ i \in [1:n]$. *The optimal power allocation policy is given by*

$$P_j^\star(\mathbf{h}) = \frac{\left[\sum_{i\in[1:n]}\frac{|h_i|^2}{\lambda_i} - 1\right]^+}{\left(\sum_{i\in[1:n]}\frac{|h_i|^2}{\lambda_i}\right)^2}\ \frac{|h_j|^2}{\lambda_j^2}, \tag{2.7}$$

where the Lagrange multipliers $\{\lambda_i,\ i \in [1:n]\}$ *solve the non-linear system of equations* $\mathbb{E}[P_j^\star(\mathbf{h})] = \overline{P}_j,\ j \in [1:n]$, *and attains*

$$C_{\text{MISOPerPC}} = \mathbb{E}\left[\log^+\left(\sum_{i\in[1:n]}\frac{|h_i|^2}{\lambda_i}\right)\right]. \tag{2.8}$$

Proof. The proof is based on solving the dual problem to (Equation 2.6) and is provided in Appendix B.

The capacity in (Equation 2.8) can be obtained by *beamforming*: each antenna transmits

$$X_i = \exp\{-j\angle h_i\}\sqrt{P_i^\star(\mathbf{h})}\ U,\ U \sim \mathcal{N}(0,1),\ i \in [1:n]$$

where $P_i^\star(\mathbf{h})$ is the optimal power allocation given by (Equation 2.7). By taking the average over all fading states, the capacity can be expressed as (Equation 2.8). □

Remark 2: If Lagrange multipliers in (Equation 2.7) are all equal to λ, then the power allocation becomes

$$P_i^\star(\mathbf{h}) = \left[\frac{1}{\lambda} - \frac{1}{\|\mathbf{h}\|^2}\right]^+ \frac{|h_i|^2}{\|\mathbf{h}\|^2}, \tag{2.9}$$

with $\|\mathbf{h}\|^2 =: \sum_{\in[1:n]} |h_i|^2$. The expression in (Equation 2.9) corresponds to the water-filling power allocation optimal under SumPC, in which case the Lagrange multiplier would satisfy $\mathbb{E}\Big[\sum_{i\in[1:n]} P_i^\star(\mathbf{h})\Big] = \mathbb{E}\left[\left[\frac{1}{\lambda} - \frac{1}{\|\mathbf{h}\|^2}\right]^+\right] = \sum_{i\in[1:n]} \overline{P}_j$. This can happen if the power constraint on each antenna is the same and the distribution of the fading vector does not change by permuting its components, such as with identical and independent distributed fading.

Remark 3: In [34] the capacity of the fading MISO channel with PerPC was derived analytically under the assumption of CSI at the receiver only — in Theorem 2.3.2 we consider the case of CSI at both the transmitter and receiver, and obtain the capacity with PerPC in closed-form.

Remark 4: The capacity of the fading MISO point-to-point channel with PerPC can not be deduced from capacity results for the fading SISO MAC [37]. Although one could think of a user in the fading SISO MAC as an antenna in the fading MISO point-to-point channel, the analogy stops there; the reason is that in the fading SISO MAC the users send independent inputs while in the fading MISO point-to-point channel the signals sent by the antennas can

be correlated. As a result, the sum-capacity achieving power policy for the fading SISO MAC is the following water-filling solution

$$P_i^\star(\mathbf{h}) = \left[\frac{1}{\lambda_i} - \frac{1}{|h_i|^2}\right]^+ \quad \text{if} \quad \frac{|h_i|^2}{\lambda_i} = \max_{k=1,\ldots,K}\left\{\frac{|h_k|^2}{\lambda_k}\right\}. \tag{2.10}$$

which does not correspond to (Equation 2.7) — for example, under (Equation 2.7) either all the antennas send with a strictly positive power or all stay silent, while under (Equation 2.10) at most one user/antenna sends at any give time.

2.3.2 Fading Cognitive Interference Channel

The sum-capacity of the EGCIFC in (Equation 2.4) involves a maximization over the power allocation policy of both transmitters and a correlation coefficient between the inputs over different fading states. We now seek to solve this optimization problem, which in turn depends on the relative strengths of the channel gains between the transmitters and receivers in the channel. We have the following theorems that describe the optimal solution.

Theorem 2.3.3 (Strong interference at the cognitive receiver / Rx_2)**.** *When* $|h_{21}|^2 \geq |h_{11}|^2$*, the optimal power allocation policy for the EGCIFC in Theorem 2.3.1 corresponds to that of point-to-point MISO with PerPC in Theorem 2.3.2.*

Proof. Given that channel gain relationship $|h_{21}|^2 \geq |h_{11}|^2$ is satisfied, then it is clear that $\rho^\star(\mathbf{h}) = 1$ is optimal in (Equation 2.4). We are then left with solving for the optimal power allocation. Setting $\rho^\star(\mathbf{h}) = 1$ reduces the optimization problem in (Equation 2.4) to that in (Equation 2.6) and hence the optimal power allocation strategy is given by Theorem 2.3.2.

TABLE I. OPTIMAL POWER POLICY WHEN $|h_{21}|^2 < |h_{11}|^2$.

Channel Gain Conditions	$P_1^\star(\mathbf{h})$	$P_2^\star(\mathbf{h})$	$\rho^\star(\mathbf{h})$
$\mathcal{R}_1 := \left\{ \frac{\lvert h_{11}\rvert^2}{\lambda_1} \le 1, \frac{\lvert h_{22}\rvert^2}{\lambda_2} \le 1 \right\}$	0	0	0
$\mathcal{R}_2 := \left\{ \frac{\lvert h_{11}\rvert^2}{\lambda_1} > 1, \frac{\frac{\lvert h_{22}\rvert^2}{\lambda_2}}{1+\frac{\lvert h_{21}\rvert^2}{\lambda_1}-\frac{\lvert h_{21}\rvert^2}{\lvert h_{11}\rvert^2}} \le 1 \right\}$	$\left[\frac{1}{\lambda_1} - \frac{1}{\lvert h_{11}\rvert^2}\right]^+$	0	0
$\mathcal{R}_3 := \left\{ \frac{\lvert h_{22}\rvert^2}{\lambda_2} > 1, \frac{\frac{\lvert h_{21}\rvert^2}{\lambda_1}}{\frac{\lvert h_{22}\rvert^2}{\lambda_2}} + \frac{\lvert h_{11}\rvert^2}{\lambda_1} - \frac{\lvert h_{21}\rvert^2}{\lambda_1} \le 1 \right\}$	0	$\left[\frac{1}{\lambda_2} - \frac{1}{\lvert h_{22}\rvert^2}\right]^+$	0
$\mathcal{R}_4 := \left\{ \frac{\lvert h_{11}\rvert^2}{\lambda_1} \le \frac{\lvert h_{21}\rvert^2}{\lambda_1} + \frac{\frac{\lvert h_{22}\rvert^2}{\lambda_2}}{\frac{\lvert h_{21}\rvert^2}{\lambda_1}+\frac{\lvert h_{22}\rvert^2}{\lambda_2}}, \frac{\lvert h_{21}\rvert^2}{\lambda_1} + \frac{\lvert h_{22}\rvert^2}{\lambda_2} > 1 \right\}$	$\frac{\frac{\lvert h_{21}\rvert^2}{\lambda_1}+\frac{\lvert h_{22}\rvert^2}{\lambda_2}-1}{\left(\frac{\lvert h_{21}\rvert^2}{\lambda_1}+\frac{\lvert h_{22}\rvert^2}{\lambda_2}\right)^2} \frac{\lvert h_{21}\rvert^2}{\lambda_1^2}$	$\frac{\frac{\lvert h_{21}\rvert^2}{\lambda_1}+\frac{\lvert h_{22}\rvert^2}{\lambda_2}-1}{\left(\frac{\lvert h_{21}\rvert^2}{\lambda_1}+\frac{\lvert h_{22}\rvert^2}{\lambda_2}\right)^2} \frac{\lvert h_{22}\rvert^2}{\lambda_1^2}$	1
$\mathcal{R}_5 := \left\{ \left(\mathcal{R}_1 \cup \mathcal{R}_2 \cup \mathcal{R}_3 \cup \mathcal{R}_4\right)^c \right\}$	numerically	numerically	numerically

Therefore, setting $R_1 = 0$ turns out to be optimal for the EGCIFC in this regime, i.e., the best use of cognitive user's ability is to broadcast the primary's message. □

Theorem 2.3.4 (Weak interference at the cognitive receiver / Rx_2)**.** *When $|h_{21}|^2 < |h_{11}|^2$, the optimal power allocation policy for the EGCIFC is summarized in Table I and is one of either of the following policies: (1) both users refrain from transmitting, (2) cognitive transmitter water-fills over $|h_{11}|$, (3) primary transmitter water-fills over $|h_{22}|$, (4) MISO with SumPC type of power allocation, or (5) both user transmits to their intended receivers with non-zero powers (in this case the optimal policy must be determined numerically).*

Proof. The proof, based on solving the Lagrangian dual problem of (Equation 2.4), is provided in Appendix C.

One may interpret the policies in Table I as the following. For $\mathcal{R}_1$ both direct link channel gains are "weak" (smaller than the corresponding optimal Lagrange multiplier) and so the

optimal scheme for both transmitters is to refrain from allocating power, saving the power for better channel states. In $\mathcal{R}_2$ the cognitive transmitter water-fills over its direct link $|h_{11}|$ while the primary user refrains from allocating any power because in this regime $|h_{11}|$ is "not weak" while $|h_{22}|$ is "weak". One can interpret $\mathcal{R}_3$ similarly to $\mathcal{R}_2$ but with the roles of the users swapped. The channel gain condition in $\mathcal{R}_4$ implies that $\frac{|h_{11}|^2}{\lambda_1} \leq \frac{|h_{21}|^2}{\lambda_1} + \frac{|h_{22}|^2}{\lambda_2}$ (given that $\frac{|h_{21}|^2}{\lambda_1} + \frac{|h_{22}|^2}{\lambda_2} > 1$); in this case the sum of the channel gains to the primary receiver is "stronger" than that of the direct gain to the cognitive receiver and performing a point-to-point MISO-type power allocation is optimal. In $\mathcal{R}_5$ both transmitters send with non-zero power; in this case a closed-form solution for the optimal power allocation policy is not available. □

Remark 5: The above analysis showed that a separable achievable scheme is sum-capacity optimal. In order to characterize the whole capacity region one needs bounds on R_1 and R_2 too; the cut-set approach gives such bounds. Therefore the capacity region of the EGCIFC is outer bounded by

$$R_1 - \epsilon_N \leq \frac{1}{N}\sum_{i=1}^{N} I(X_{1Gi}; Y_{1i}|X_{2Gi}, \mathbf{h}_i)$$

$$R_2 - \epsilon_N \leq \frac{1}{N}\sum_{i=1}^{N} I(Y_{2Gi}; X_{1i}, X_{2Gi}|\mathbf{h}_i)$$

$$R_1 + R_2 - 2\epsilon_N \leq \frac{1}{N}\sum_{i=1}^{N} I(Y_{2i}; X_{1Gi}, X_{2Gi}|\mathbf{h}_i) + I(X_{1Gi}; Y_{1i}|X_{2Gi}, Y'_{2i}, \mathbf{h}_i),$$

where the last bound is from Theorem 2.3.1 and the single rate bounds are cut-set bounds, similarly to [13, eq.(8)]. As for the sum-capacity, the whole region is exhausted by considering jointly Gaussian inputs and where the region can be tighten by choosing any $Y'_2 \sim Y_2$. By further

taking the appropriate limit over N as done in [26] and by considering the input covariance as in (Equation 2.3), the upper bound region can be expressed as

$$R_1 \leq \mathbb{E}\left[\log\left(1+|h_{11}|^2 P_1(\mathbf{h})(1-|\rho(\mathbf{h})|^2)\right)\right], \tag{2.11a}$$

$$R_2 \leq \mathbb{E}\left[\log\left(1+|h_{21}|^2 P_1(\mathbf{h})+|h_{22}|^2 P_2(\mathbf{h})+2|\rho(\mathbf{h})|\sqrt{|h_{21}|^2 P_1(\mathbf{h})\ |h_{22}|^2 P_2(\mathbf{h})}\right)\right], \tag{2.11b}$$

$$\begin{aligned} R_1+R_2 \leq\ & \mathbb{E}\left[\log\left(1+|h_{21}|^2 P_1(\mathbf{h})+|h_{22}|^2 P_2(\mathbf{h})+2|\rho(\mathbf{h})|\sqrt{|h_{21}|^2 P_1(\mathbf{h})\ |h_{22}|^2 P_2(\mathbf{h})}\right)\right] \\ & + \mathbb{E}\left[\left[\log\left(1+|h_{11}|^2 P_1(\mathbf{h})(1-|\rho(\mathbf{h})|^2)\right)-\log\left(1+|h_{21}|^2 P_1(\mathbf{h})(1-|\rho(\mathbf{h})|^2)\right)\right]^+\right]. \end{aligned} \tag{2.11c}$$

By considering the achievable scheme in [13, eqs.(22)-(23)], which was shown to be at most to within 1 bit per channel use per user of the capacity region outer bound for the static / non-fading case, the outer bound in (Equation 2.11) can be shown to be achievable to within 1 bit per channel use per user as well as in the EGCIFC. We note that this achievable scheme involves two DPC steps, one per user (while the scheme in (Equation 2.5) that only has one DPC step) and it is only approximately optimal to within a constant gap (while the scheme in (Equation 2.5) is exactly sum-rate optimal).

We therefore conclude that, in order to characterize the whole capacity region of the EGCIFC to within a gap, it suffices to characterize the closure of the outer bound in (Equation 2.11),

which can be done by solving the following family of convex optimization problems: for each $\lambda \in [0, 1]$

$$C_{\text{EGCIFC,region}}(\lambda) := \max\{\lambda R_1 + (1-\lambda)R_2\}, \tag{2.12}$$

where the maximization is over the rate pairs (R_1, R_2) in (Equation 2.11). Solving the optimization problem in (Equation 2.12) is not a trivial extension of the sum-capacity results presented in Theorems 2.3.3 and 2.3.4 and is beyond the scope of this work. This is so because one needs to consider which bounds are active in (Equation 2.11) in order to determine the optimal 'corner point' as a function of $\lambda \in [0, 1]$; the coordinate of such a point must be plugged in the optimization problem in (Equation 2.12) and the corresponding KKT conditions must be worked out similarly to Appendix C. We expect that there will be parameters regimes in which the KKT conditions must be solved numerically as for the sum-capacity.

2.4 Numerical Results

In this section we numerically evaluate Theorems 2.3.1 and 2.3.2 for the case where the channel gains are independent Rayleigh random variables, not necessarily with the same mean parameter.

2.4.1 The point-to-point MISO channel with PerPC

We first consider a 2×1 point-to-point MISO channel. The channel vector $[h_1, h_2]$ at each channel use has independent and exponentially distributed components with means $\gamma_1 = \mathbb{E}[|h_1|^2] = 5$ and $\gamma_2 = \mathbb{E}[|h_2|^2] = 2$. The transmit antennas are subject to the average power

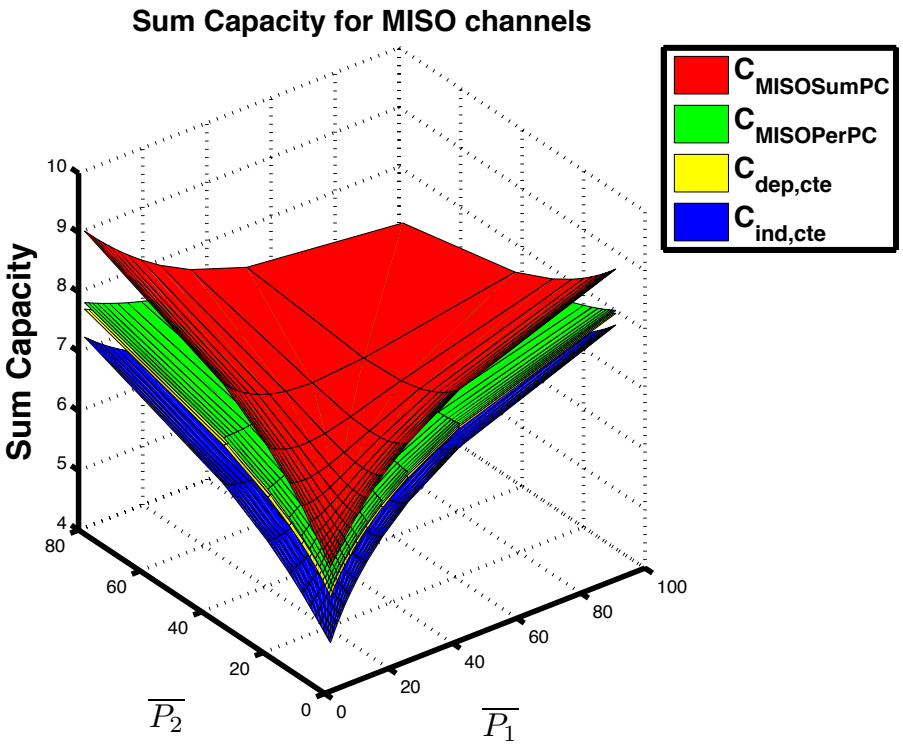


Figure 2. The capacity of the point-to-point MISO channel with PerPC (surface in green, from (Equation 2.13)) is upper bounded by that of a MISO channel with SumPC (surface in red, from (Equation 2.16)) and and lower by that of a MISO channel with constant power allocation and dependent inputs (surface in yellow, from (Equation 2.14)) and with constant power allocation and independent inputs (surface in blue, from (Equation 2.15)).

constraints $\overline{P}_1$ and $\overline{P}_2$. In Figure 2 four surfaces representing the capacities of different MISO channels are plotted as functions of the average transmit antenna power constraints $\overline{P}_1$ and $\overline{P}_2$. The surfaces correspond to:

1) the MISO channel with PerPC

$$C_{\text{MISOPerPC}} = \mathbb{E}\left[\log\left(1+\left(\sqrt{|h_1|^2 P_1^\star(\mathbf{h})}+\sqrt{|h_2|^2 P_2^\star(\mathbf{h})}\right)^2\right)\right] \tag{2.13}$$

where $P_1^\star(\mathbf{h})$ and $P_2^\star(\mathbf{h})$ are given in (Equation 2.7) for $j \in [1:2]$;

2) the MISO channel with constant power allocation and beam-forming (where the instantaneous phases are known to the transmitters to allow for coherent beam-forming)

$$C_{\text{dep,cte}} = \mathbb{E}\left[\log\left(1+\left(\sqrt{|h_1|^2\overline{P}_1}+\sqrt{|h_2|^2\overline{P}_2)}\right)^2\right)\right]; \tag{2.14}$$

3) the MISO channel with constant power allocation and independent signaling (instantaneous phases are not known at the transmitter and they are independent and uniformly distributed in $[0, 2\pi]$ as in [34])

$$C_{\text{indep,cte}} = \mathbb{E}\left[\log\left(1+|h_1|^2\overline{P}_1+|h_2|^2\overline{P}_2\right)\right]; \tag{2.15}$$

and 4) the MISO channel with a SumPC

$$C_{\text{MISOSumPC}} = \mathbb{E}\left[\log\left(1+\left(\sqrt{|h_1|^2 P_1^\star(\mathbf{h})}+\sqrt{|h_2|^2 P_2^\star(\mathbf{h})}\right)^2\right)\right] \tag{2.16}$$

where $P_1^\star(\mathbf{h})$ and $P_2^\star(\mathbf{h})$ are given in (Equation 2.9) for $i \in [1:2]$.

Numerical evaluations show that the capacity of the point-to-point MISO with PerPC is upper and lower bounded by that of the MISO with SumPC and that of constant power allocation respectively (as expected). Also as expected, dependent constant inputs in (Equation 2.14) outperform independent constant inputs in (Equation 2.15).

Remark 6: In [34] the point-to-point MISO channel with PerPC with Rayleigh fading with no CSI at the transmitters was compared to that with SumPC and with independent signaling; the author noted that $C_{\mathrm{MISOPerPC}} = C_{\mathrm{indep,cte}}$ because of the phase being independent and uniformly distributed in $[0, 2\pi]$. Since here we assume the transmitter has CSI we have $C_{\mathrm{MISOPerPC}} \geq C_{\mathrm{indep,cte}}$.

Remark 7: In the case of dependent inputs and beam-forming, the channel is assumed to have CSI at the transmitter to account for the channel gain phases and the ability to coherently beam-form. This explains why $C_{\mathrm{MISOPerPC}}$ is almost the same as $C_{\mathrm{dep,cte}}$, which may have practical implications.

2.4.2 The EGCIFC Sum-Capacity

We now consider two different scenarios for the EGCIFC corresponding again to Rayleigh fading channels with different means: Case 1) the means are chosen such that the channel experiences strong interference with high probability, and Case 2) the means are chosen to experience weak interference with high probability. Monte Carlo simulations were used to evaluate the capacities.

In Figure 3 the sum-capacity for the EGCIFC having $|h_{11}|^2, |h_{21}|^2$ and $|h_{22}|^2$ exponentially distributed with $\gamma_0 = \mathbb{E}[|h_{11}|^2] = 1$, $\gamma_1 = \mathbb{E}[|h_{21}|^2] = 5$ and $\gamma_2 = \mathbb{E}[|h_{22}|^2] = 2$ (skewed with high probability to be in strong interference since $\mathbb{P}\left[|h_{21}|^2 \geq |h_{11}|^2\right] = \frac{\gamma_1}{\gamma_1+\gamma_0} = \frac{5}{6}$) is plotted along with the sum-capacity of the system using the MISO with PerPC transmit strategy. The latter is not optimal in general when the channel experiences weak interference states. The surface representing the sum-capacity $C_{\text{MISOPerPC}}$ approaches that of C_{EGCIFC} (as expected since we are skewed to be in strong interference where the scheme corresponding to a MISO channel with PerPC is optimal).

In Figure 4 the sum-capacity for the EGCIFC having mean parameters $\gamma_0 = 5$, $\gamma_1 = 1$ and $\gamma_2 = 2$ is plotted (skewed with high probability to be in weak interference $\mathbb{P}\left[|h_{21}|^2 < |h_{11}|^2\right] = \frac{5}{6}$) along with the sum-capacity achieved by using the power allocation corresponding to that of a MISO channel with PerPC. As expected, the MISO scheme with PerPC (not optimal for weak interference) does not perform as well as in the regime where the channel is skewed to be in strong interference with high probability.

In Figure 5 we choose the channel gains to be identically distributed with mean parameter $\gamma_0 = \gamma_1 = \gamma_2 = 1$ and plot the sum-capacity of the EGCIFC and an achievability scheme corresponding to constant power allocation, but with the optimal correlation coefficient which changes with each fading state, i.e. the solution of

$$C_{\text{sum}} = \mathbb{E}\Bigg[\max_{|\rho(\mathbf{h})|\leq 1} \log\left(1 + |h_{21}|^2\overline{P_1} + |h_{22}|^2\overline{P_2} + 2|\rho(\mathbf{h})|\sqrt{|h_{21}|^2\overline{P_1}\ |h_{22}|^2\overline{P_2}}\right) + \\ + \log\left(\frac{1 + (1-|\rho(\mathbf{h})|^2)\max\{|h_{11}|^2, |h_{21}|^2\}\overline{P_1}}{1 + (1-|\rho(\mathbf{h})|^2)|h_{21}|^2\overline{P_1}}\right)\Bigg]. \tag{2.17}$$

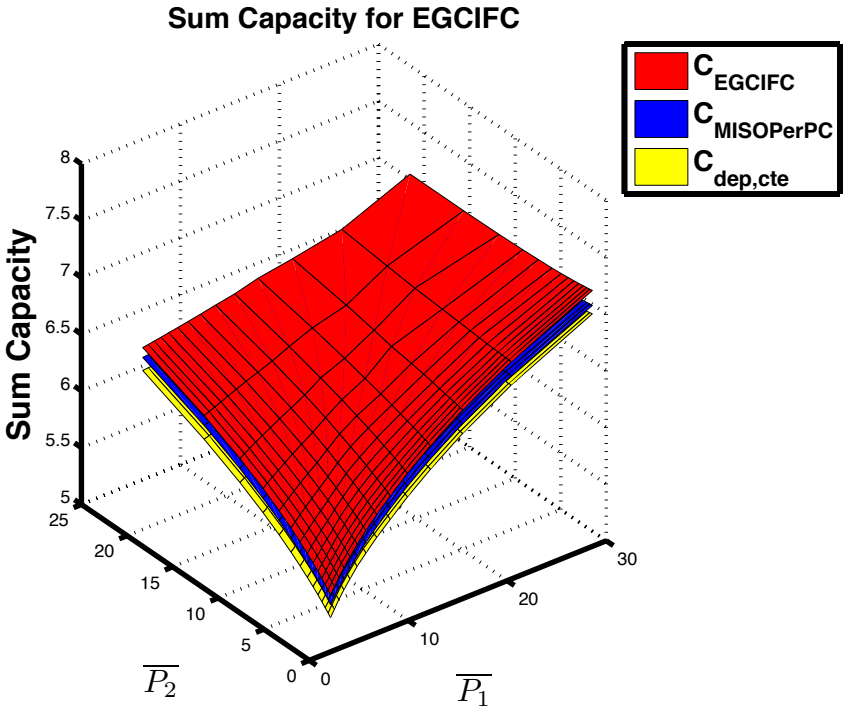


Figure 3. The sum-capacity of the EGCIFC which is skewed to be in strong interference with high probability (surface in red) is plotted with the sum-capacity achieved when considering a MISO achievability with PerPC (surface in blue). The MISO power allocation is almost sum-capacity achieving as the two surfaces almost overlap. The constant power allocation scheme with dependent inputs (surface in yellow) is again a lower bound on the sum-capacity of EGCIFC. Note that perfect CSI at both transmitters is needed for coherent beam-forming.

Note that (Equation 2.17) is solved for the optimal correlation coefficient numerically. It is interesting to note that by optimizing the correlation coefficient only (and not the power allocation, i.e., keeping the power constant) one can approach (a difference of around 0.23 bits/channel use), at least for these channel conditions, the sum-capacity of the EGCIFC. This may have

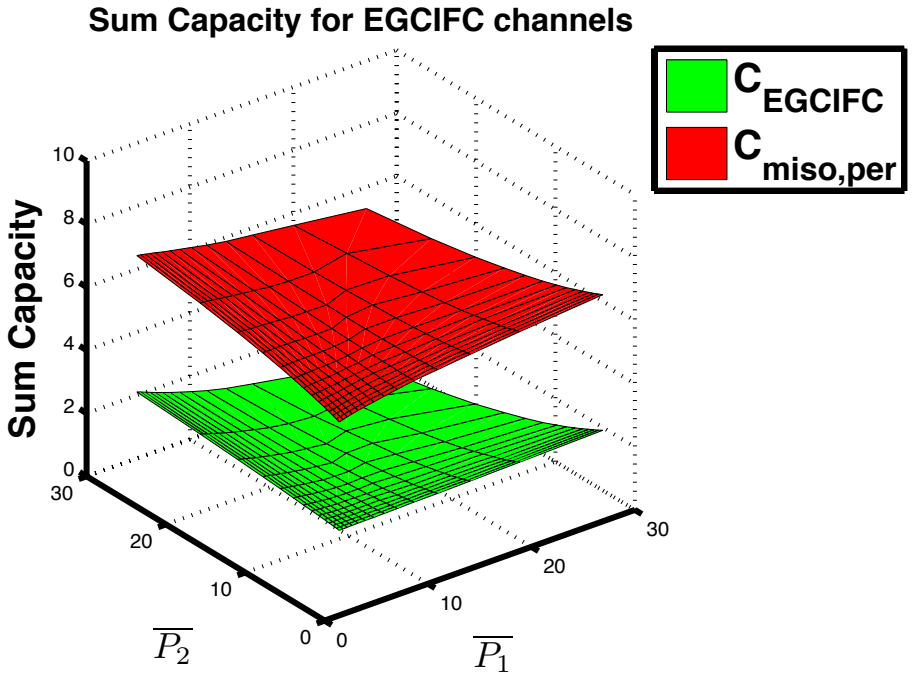


Figure 4. The sum-capacity of EGCIFC which is skewed to be in weak interference with high probability (surface in green) is plotted with the sum-capacity achieved when considering a MISO achievability scheme with PerPC. The power allocation as in the MISO with PerPC (surface in green) is not optimal as in the case when the channel is skewed to be in strong interference.

implications in practice – i.e, constant power may be good enough if one optimally finds the correlation coefficient.

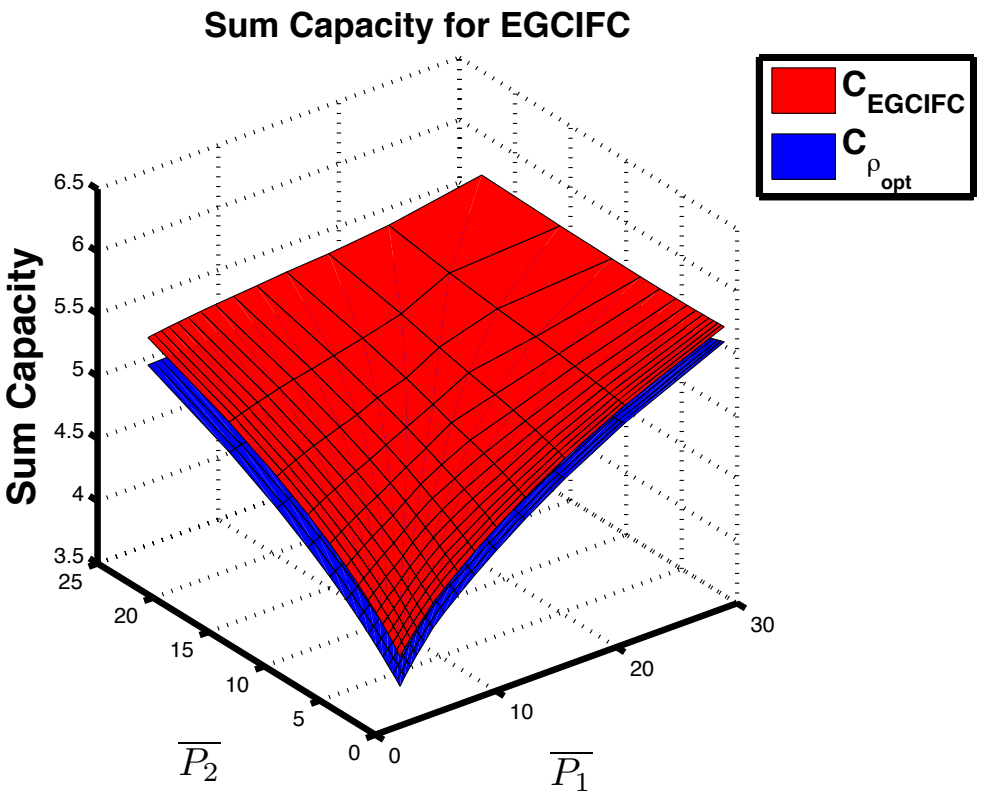


Figure 5. Two surfaces representing the sum-capacity of EGCIFC with the optimal power allocation (red) and the sum-capacity while considering constant power allocation and an optimized correlation coefficient at each fading state (blue). Although the optimal power allocation was not utilized, the blue surface is almost capacity achieving.

CHAPTER 3

K USER COGNITIVE INTERFERENCE CHANNEL WITH CUMULATIVE MESSAGE SHARING

All this chapter has been previously published in [2]; while parts of it has been published in [4], [5], [6].

The K-user cognitive interference channel with cumulative message sharing is a wireless channel consisting of one primary and $\mathsf{K}-1$ secondary/cognitive transmitters with cognitive transmitter $\mathsf{i} \in [2:\mathsf{K}]$ having non-causal knowledge of the messages of users with index less than i. For this channel we are interested in understanding the information theoretic limits of communication and amount of side information that is sufficient for optimal communication. In order to accomplish this, a computable outer bound valid for any memoryless channel is proposed. The sum-rate outer bound is evaluated first for the high-SNR linear deterministic approximation of the Gaussian noise channel. This is shown to be both the sum-capacity for the 3-user channel with arbitrary channel gains, and the sum-capacity for the symmetric K-user channel. Interestingly, for the K user channel, we observe that cognition at transmitters 2 to $\mathsf{K}-1$ is not needed, and knowledge of all messages at the K-th transmitter only is sufficient to achieve the sum-capacity. Next, the sum-capacity of the symmetric Gaussian noise channel is characterized to within a constant additive and multiplicative gap, both of which are functions

of K. As opposed to other multi-user interference channel models, a single scheme (in this case based on dirty-paper coding) suffices for both the weak and strong interference regimes. The generalized degrees of freedom (gDoF) are then derived and are shown, unlike interference and broadcast channels, to be a function of K. Interestingly, it is shown that as the number of users grows to infinity the gDoF of the K-user cognitive interference channel with cumulative message sharing tends to the gDoF of a broadcast channel with a K-antenna transmitter and K single-antenna receivers. Numerical evaluations show that the actual gaps between the presented inner and outer bounds are significantly smaller than the analytically derived gaps.

3.1 Prior Work

While not much work on $\mathsf{K} > 3$ channels exists, in [14, 15, 17, 38, 39] different three-user cognitive channels are considered; we note that the models differ from the one considered here either in the number of transmitter/receivers, or in the message sharing/cognition structure in all but [14, 15]. In the more comprehensive [14], several types of 3-user cognitive interference channels are proposed: that with "cumulative message sharing" (CMS) as considered here, that with "primary message sharing" where the message of the single primary user is known at both cognitive transmitters (who do not know each others' messages), and finally "cognitive only message sharing" (CoMS) where there are two primary users who do not know each others' message and a single cognitive user which knows both primary messages. Achievable rate regions are obtained which are evaluated in Gaussian noise. The CoMS mechanism yields almost the same message structure as in the interference channel with a cognitive relay – identical if the relay were to further have a message of its own (see [40, 41] and references therein for the

interference channel with a cognitive relay). In [38] the CoMS was first introduced where an achievable rate region was obtained which employs a combination of superposition coding and Gel'fand-Pinsker's binning, which was numerically evaluated for the Gaussian noise channel. In [39] the CoMS structure is assumed and the cognitive user is furthermore assumed not to interfere with the primary users; an inner and an outer bound are obtained. In [16,17] capacity under "strong interference" for the CoMS is obtained. We thus emphasize that the channel considered here is more general than others studied as we consider K users, a fully connected interference channel, and consider the less studied CMS sharing structure. In [25], the authors characterized the Degrees of Freedom (DoF) of a K-user interference channel in which each transmitter, in addition to its own message, has access to a subset of the other users' messages, was obtained; in particular, it was shown that the DoF=K (maximum possible) if the sum of the number of jointly cooperating transmitters and the number of jointly decoding receivers is greater than or equal to $\mathsf{K}+1$. An outer bound on the DOF region and the sum DOF is presented.

3.2 Channel Model

The general memoryless K-user cognitive interference channel with cumulative message sharing (K-CIFC-CMS) consists of K source-destination pairs sharing the same physical channel, where some transmitters have non-causal knowledge of the messages of other transmitters. Here transmitter 1 is referred to as the primary user and is assumed to have no cognitive abilities. Transmitter i, $\mathsf{i} \in [2:\mathsf{K}]$, is non-causally cognizant of the messages of the users with index smaller than i. More formally, the K-CIFC-CMS channel consists of

- Channel inputs $X_i \in \mathcal{X}_i,\ i \in [1:K]$,
- Channel outputs $Y_i \in \mathcal{Y}_i,\ i \in [1:K]$,
- A memoryless channel with joint transition probability $P(Y_1, \ldots, Y_K | X_1, \ldots, X_K)$,
- Messages W_i known to users $1, 2, \ldots, i$, $i \in [1:K]$.

A code with non-negative rate vector $(R_1, \ldots, R_K)$ and blocklength N is defined by

- Messages $W_i,\ i \in [1:K]$, uniformly distributed over $[1:2^{NR_i}]$ and independent of everything else,
- Encoding functions $f_i^{(N)} : [1:2^{NR_1}] \times \ldots \times [1:2^{NR_i}] \to \mathcal{X}_i^N$ such that $X_i^N := f_i^{(N)}(W_1, \ldots, W_i)$, $i \in [1:K]$,
- Decoding functions $g_i^{(N)} : \mathcal{Y}_i^N \to [1:2^{NR_i}]$ such that $\widehat{W}_i = g_i^{(N)}(Y_i^N)$, $i \in [1:K]$,
- Probability of error $P_e^{(N)} := \max_{i \in [1:K]} \mathbb{P}[\widehat{W}_i \neq W_i]$.

The capacity of the K-CIFC-CMS channel consists of all non-negative rate tuples $(R_1, \ldots, R_K)$ for which there exist a sequence of codes indexed by the block length N such that $P_e^{(N)} \to 0$ as $N \to \infty$. Since the decoders cannot cooperate and the channel is used without feedback, the capacity may be shown to depend only on the marginal noise distributions rather than the joint noise distribution by an argument similar to that used for the broadcast channel (BC) [42]. In this work we focus on two channel models, the Gaussian noise channel and its linear deterministic approximation at high SNR.

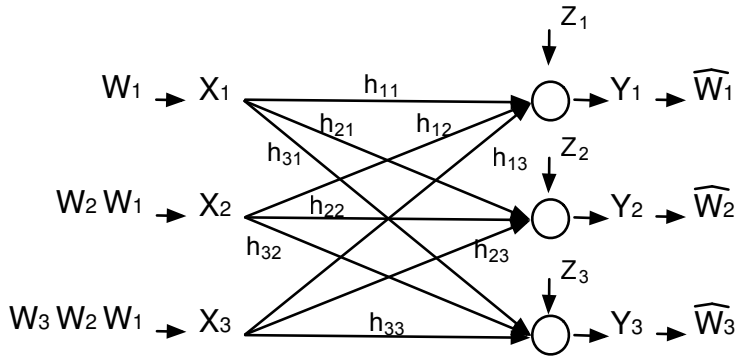


Figure 6. The Gaussian 3-CIFC-CMS.

3.2.1 The Gaussian noise channel

The single-antenna complex-valued K-CIFC-CMS with Additive White Gaussian Noise (AWGN), shown in Figure 6 for $K = 3$, has input-output relationship

$$Y_\ell = \sum_{i\in[1:K]} h_{\ell i} X_i + Z_\ell,\ \ell \in [1:K], \tag{3.1a}$$

where, without loss of generality, the inputs are subject to the power constraint

$$\mathbb{E}[|X_i|^2] \le 1,\ i \in [1:K], \tag{3.1b}$$

and the noises are marginally proper-complex Gaussian random variables with parameters

$$Z_\ell \sim \mathcal{N}(0,1),\ \ell \in [1:K]. \tag{3.1c}$$

The channel gains h_{ij}, $(i,j) \in [1:K]^2$, are constant. Without loss of generality we may assume the direct links h_{ii}, $i \in [1:K]$ to be real-valued and non-negative. The Generalized Degrees-of-Freedom (gDoF) of the *symmetric* Gaussian channel is a performance metric that characterizes the high-SNR behavior of the sum-capacity and is defined as follows: Let SNR be a non-negative number and parameterize

$$|h_{ii}|^2 := SNR, \ i \in [1:K], \tag{3.2a}$$

$$|h_{\ell i}|^2 := SNR^{\alpha}, \ (\ell, i) \in [1:K]^2, \ell \neq i, \tag{3.2b}$$

for some non-negative α. The gDoF is

$$d(\alpha) := \lim_{SNR \to +\infty} \frac{C_\Sigma}{\log(1+SNR)}, \tag{3.3}$$

where $C_\Sigma := \max\{R_1 + \ldots + R_K\}$ and where the maximization is over all achievable rates. The sum-capacity is said to be known to within a constant gap of b bits if one can show rates $R_\Sigma^{(in)}$ and $R_\Sigma^{(out)}$ such that

$$R_\Sigma^{(in)} \leq C_\Sigma \leq R_\Sigma^{(out)} \leq R_\Sigma^{(in)} + b\log(2). \tag{3.4}$$

The gDoF and constant gap characterization of the symmetric sum capacity imply that

$$C_\Sigma = d(\alpha)\log(1+SNR) + o(1),$$

where $o(1)$ indicates a quantity that is finite at all SNR.

3.2.2 Linear deterministic approximation of the Gaussian noise channel at high SNR

The Linear Deterministic approximation of the Gaussian Noise Channel at high SNR (LDC) was first introduced in [43] to allow focusing on the effect of signal interactions between users rather than on the effect of additive noise. The proposed framework has been powerful in revealing key issues for the problem of communicating over interfering networks. The insights gained from the LDC have often translated into Gaussian capacity results to within a constant gap for any finite SNR [13,44,45]. In light of these success stories we also start our investigation from the LDC. The LDC has input-output relationship

$$Y_\ell = \sum_{i\in[1:K]} \mathbf{S}^{m-n_{\ell i}} X_i,\ \ell \in [1:K], \tag{3.5}$$

where $m := \max\{n_{ij}\}$, $\mathbf{S}$ is the binary shift matrix of dimension m, all inputs and outputs are binary column vectors of dimension m, the summation is bit-wise over of the binary field, and the channel gains $n_{\ell i}$ for $(\ell, i) \in [1:K]^2$, are positive integers. In a *symmetric* LDC all direct links have the same strength $n_{ii} = n_d \geq 0, i \in [1:K]$, and all the interfering links have the same strength $n_{\ell i} = n_i = \alpha\, n_d \geq 0, (\ell, i) \in [1:K]^2,\ \ell \neq i$. Note that the subscript i (roman font) of n_i stands for "interference" and is not an index; as such it should not be confused with index i (italic font).

The channel in (Equation 3.5) can be thought of as the high SNR approximation of the channel in (Equation 3.1) with their parameters related as $n_{ij} = \lfloor \log(1+|h_{ij}|^2) \rfloor$, $(i,j) \in [1:K]^2$.

3.3 Outer Bound

In this section we derive an outer-bound region for the general memoryless K-CIFC-CMS. We start with the case of $K = 3$ users to highlight the main proof techniques and ease the reader into the extension to any number of users.

Theorem 3.3.1. *The capacity region of the general memoryless 3-CIFC-CMS is contained in the region defined by*

$$R_1 \leq I(Y_1; X_1, X_2, X_3), \tag{3.6a}$$

$$R_2 \leq I(Y_2; X_2, X_3 | X_1), \tag{3.6b}$$

$$R_3 \leq I(Y_3; X_3 | X_1, X_2), \tag{3.6c}$$

$$R_2 + R_3 \leq I(Y_2; X_2, X_3 | X_1) + I(Y_3; X_3 | X_1, X_2, Y_2), \tag{3.6d}$$

$$R_1 + R_2 + R_3 \leq I(Y_1; X_1, X_2, X_3) + I(Y_2; X_2, X_3 | X_1, Y_1) + I(Y_3; X_3 | X_1, Y_1, X_2, Y_2), \tag{3.6e}$$

for some input distribution P_{X_1,X_2,X_3}. *The joint conditional distribution* $P_{Y_1,Y_2,Y_3|X_1,X_2,X_3}$ *can be chosen so as to tighten the different bounds as long as the conditional marginal distributions* $P_{Y_i|X_1,X_2,X_3}$, $i \in [1:3]$, *are preserved.*

Proof. The proof is found in Appendix D. □

Remarks:

1. The region in Th. 3.3.1 reduces to the outer bound in [12, Th. 6] by setting $X_3 = Y_3 = \emptyset$.

2. The outer bound region in (Equation 3.6) does not contain auxiliary random variables. Moreover, every mutual information term contains all the inputs. These two facts imply that the outer bound region in Th. 3.3.1 can be easily evaluated for many channels of interest. For example, for the Gaussian noise channel in Section 3.2.1, the "Gaussian maximizes entropy" principle suffices to show that jointly Gaussian inputs exhaust the outer bound.

3. The sum-capacity bound in (Equation 3.6e) is obtained by giving S_i as side information to receiver i, $i \in [1 : K]$, where $S_i = [S_{i-1}, W_{i-1}, Y_{i-1}^N]$ starting with $S_1 = \emptyset$. With this "nested" side information, the mutual information terms can be expressed in terms of entropies which may be recombined in ways that can be easily single-letterized. This form of side information allows us to extend the result from the 3-user case to any number of users.

4. The mutual information terms in (Equation 3.6e) have the form

$$I(Y_i; X_i, \ldots, X_K | X_1, Y_1, \ldots, X_{i-1}, Y_{i-1}), 1 \leq i \leq K$$

which can be given the following interpretation: Since message W_i is available at transmitters i through K, inputs $(X_i, \ldots, X_K)$ are "informative" for receiver i, while inputs $(X_1, \ldots, X_{i-1})$ are independent of W_i; receiver i decodes from Y_i the information carried

in $(X_i, \ldots, X_K)$ that could not be recovered by users with lesser index as represented by $(X_1, Y_1, \ldots, X_{i-1}, Y_{i-1})$.

Th. 3.3.1 can be extended to any K as follows.

Theorem 3.3.2. *The capacity region of the general memoryless K-CIFC-CMS is contained in the region defined by*

$$R_i \leq I(Y_i; X_{[i:K]} | X_{[1:i-1]}), \tag{3.7a}$$

$$\sum_{j=i}^{K} R_j \leq \sum_{j=i}^{K} I(Y_j; X_{[j:K]} | X_{[1:j-1]}, Y_{[1:j-1]}), \tag{3.7b}$$

for some input distribution $P_{X_1,\ldots,X_K}$. Moreover, each rate bound in may be tightened with respect to the channel conditional distribution as long as the channel conditional marginal distributions are preserved.

Proof. The proof is found in Appendix E. □

In the following section we shall derive achievable schemes matching the sum-capacity outer bound in Th. 3.3.2 for the LDC in (Equation 3.5) and schemes that achieve the sum-capacity outer bound to within a constant bounded gap regardless of the channel parameters for the Gaussian channel in (Equation 3.1).

3.4 Sum-capacity for the Linear Deterministic K-CIFC-CMS

In Sections 3.4.1 and 3.4.2 we determine the sum-capacity of the LDC with $K = 3$ users and any value of the channel gains. In Sections 3.4.4 and 3.4.5 we derive the sum-capacity for any K but for symmetric channel gains only. The main results of this section are

Theorem 3.4.1. *The sum-capacity bound in* (Equation 3.6e) *is achievable for the LDC* 3*-CIFC-CMS with generic channel gains.*

Theorem 3.4.2. *The sum-capacity bound in* (Equation 3.7b) *is achievable for the LDC* K*-CIFC-CMS with symmetric channel gains. The capacity achieving scheme only requires cognition of all messages at one single transmitter.*

The rest of the section is devoted to their proofs.

3.4.1 Sum-capacity outer bound for the 3-user case and generic channel gains

The sum-capacity outer bound in Th. 3.3.1 specialized to a deterministic 3-CIFC-CMS (i.e., $H(Y_i|X_1, X_2, X_3) = 0, i \in [1:3]$) gives the following sum-capacity upper-bound

$$R_1 + R_2 + R_3 \leq \max \Big\{ H(Y_1) + H(Y_2|X_1, Y_1) + H(Y_3|X_1, Y_1, X_2, Y_2) \Big\},$$

where the maximization is over all possible joint distributions P_{X_1,X_2,X_3}. For the LDC in (Equation 3.5) with $K = 3$ we obtain

$$R_1 + R_2 + R_3 \leq \max\{n_{11}, n_{12}, n_{13}\} + f(n_{22}, n_{23}|n_{12}, n_{13}) + [n_{33} - \max\{n_{13}, n_{23}\}]^+, \qquad (3.8a)$$

where $f(c, d|a, b)$ follows from [46, eq.(5)] and is defined as

$$f(c, d|a, b) := \begin{cases} \max\{c + b, a + d\} - \max\{a, b\} & \text{if } c - d \neq a - b, \\ \max\{a, b, c, d\} - \max\{a, b\} & \text{if } c - d = a - b. \end{cases}$$

The bound in (Equation 3.8) follows by maximizing each mutual information term individually and is as follows

$$\begin{aligned}
H(Y_1) &= H(\mathbf{S}^{m-n_{11}}X_1 + \mathbf{S}^{m-n_{12}}X_2 + \mathbf{S}^{m-n_{13}}X_3) \leq \max\{n_{11}, n_{12}, n_{13}\},\\
H(Y_2|X_1, Y_1) &= H(\mathbf{S}^{m-n_{22}}X_2 + \mathbf{S}^{m-n_{23}}X_3|X_1, \mathbf{S}^{m-n_{12}}X_2 + \mathbf{S}^{m-n_{13}}X_3)\\
&\leq H(\mathbf{S}^{m-n_{22}}X_2 + \mathbf{S}^{m-n_{23}}X_3|\mathbf{S}^{m-n_{12}}X_2 + \mathbf{S}^{m-n_{13}}X_3) \leq f(n_{22}, n_{23}|n_{12}, n_{13}),\\
H(Y_3|X_1, Y_1, X_2, Y_2) &= H(\mathbf{S}^{m-n_{33}}X_3|X_1, X_2, \mathbf{S}^{m-n_{13}}X_3, \mathbf{S}^{m-n_{23}}X_3) \leq H(\mathbf{S}^{m-n_{33}}X_3|\mathbf{S}^{m-\max\{n_{13},n_{23}\}}X_3)\\
&\leq [n_{33} - \max\{n_{13}, n_{23}\}]^+,
\end{aligned}$$

where $[x]^+ := \max\{0, x\}$. Notice that i.i.d. Bernoulli(1/2) input bits simultaneously maximize each of the above entropy terms.

3.4.2 Achievability of the sum-capacity outer bound for the 3-user case and generic channel gains

In the following, depending on whether $[n_{33} - \max\{n_{13}, n_{23}\}]^+$ in (Equation 3.8a) is zero or positive, different interference scenarios are identified and transmission schemes that are capable of achieving the sum-capacity outer bound in (Equation 3.8) are proposed. In particular:

Case 1: If the signal sent by the most cognitive transmitter is received the weakest at the intended destination, that is, if

$$n_{33} \leq \max\{n_{13}, n_{23}\}, \tag{3.9}$$

the sum-capacity in (Equation 3.8) becomes

$$R_1 + R_2 + R_3 \leq \max\{n_{11}, n_{12}, n_{13}\} + f(n_{22}, n_{23}|n_{12}, n_{13}).$$

The condition in (Equation 3.9) corresponds to $H(Y_3|X_1, Y_1, X_2, Y_2) = 0$, i.e., conditioned on (X_1, X_2) the signal received at the most cognitive receiver is a degraded version of the signal received at the other two receivers. Recall that user 3 may send information to all receivers as it knows all messages. The condition in (Equation 3.9) implies that the signal X_3 may convey more information to receivers 1 and 2 than it can to receiver 3. In this case, one might thus suspect that $R_3 = 0$ is optimal and that the best use of the cognitive capabilities of user 3 is to broadcast to the receivers. We will next show that this is indeed the case.

We set $R_3 = 0$ and we therefore convert the LCD 3-CIFC-CMS into a deterministic 2-CIFC-CMS where user 1 is the primary user (with input X_1 and output Y_1) and the cognitive user has vector input $[X_2, X_3]$ and output Y_2. The capacity of a general deterministic 2-user cognitive interference channel is [12, Th. 12]

$$R_1 \leq H(Y_1), \; R_2 \leq H(Y_2|X_1),$$

$$R_1 + R_2 \leq H(Y_1) + H(Y_2|X_1, Y_1),$$

for some input distribution $P_{X_1,[X_2,X_3]}$. Hence the sum-capacity is

$$\begin{aligned} R_1 + R_2 &= \max_{P_{X_1,[X_2,X_3]}} \Big\{ H(Y_1) + H(Y_2|X_1, Y_1) \Big\} \\ &= \max\{n_{11}, n_{12}, n_{13}\} + f(n_{22}, n_{23}|n_{12}, n_{13}), \end{aligned}$$

Case 2: In the regime not covered by the condition in (Equation 3.9), that is, for

$$n_{33} > \max\{n_{13}, n_{23}\}, \tag{3.10}$$

the sum-capacity in (Equation 3.8) becomes

$$R_1 + R_2 + R_3 \leq \max\{n_{11}, n_{12}, n_{13}\} + f(n_{22}, n_{23}|n_{12}, n_{13}) + n_{33} - \max\{n_{13}, n_{23}\}.$$

In this case, the condition in (Equation 3.10) suggests that the intended signal at receiver 3 is sufficiently strong to be able to support a non-zero rate. The form of the sum-capacity also suggests that a plausible strategy is to use the optimal strategy for Case 1 and "sneak in" extra bits for user 3 in such a way that they do not appear at the other receivers. We next show that this is optimal. The signal of transmitter 3 in two parts

$$X_3 := X_{3a} + X_{3b},$$

where X_{3a} is intended to mimic the scheme for Case 1 (i.e., as if user 2 had input $[X_2, X_{3a}]$) and X_{3b} carries the information to Y_3, possibly "pre-coded" against the interference of (X_1, X_2, X_{3a}), and such that X_{3b} is not received at receivers 1 and 2. We define

$$X_{3b} := \mathbf{S}^{\max\{n_{13}, n_{23}\}} V_3,$$

for some vector V_3 defined in the following. Note that the shift caused by $\mathbf{S}^{\max\{n_{13}, n_{23}\}}$ is such that V_3 is not received at Y_1 and at Y_2. We note that V_3 is "private information" for receiver 3 that is dirty paper coded against the interference caused by $[X_1, X_2, X_{3a}]$ at receiver 3; with this receiver 3 is virtually interference-free. We then implement the optimal strategy for Case 1 with $[X_1, X_2, X_{3a}]$ and with the remaining bits in X_{3b} we transmit to receiver 3 thereby achieving the sum-capacity in (Equation 3.8).

3.4.3 Example of sum-capacity optimal schemes for the 3-user case and symmetric channel gains

We now present several concrete examples of the achievability scheme presented in Section 3.4.2. We consider $n_d > 0,\ n_i = n_d\ \alpha,\ \alpha \geq 0$. Define the normalized sum-capacity as

$$d_\Sigma(\alpha; 3) := \frac{\max\{R_1 + R_2 + R_3\}}{n_d}.$$

Note that when $n_d = 0$ the channel reduces to a broadcast channel from transmitter $[X_2, X_3]$ to receivers Y_1 and Y_2 (receiver 3 cannot be reached by its transmitter and hence $R_3 = 0$ is

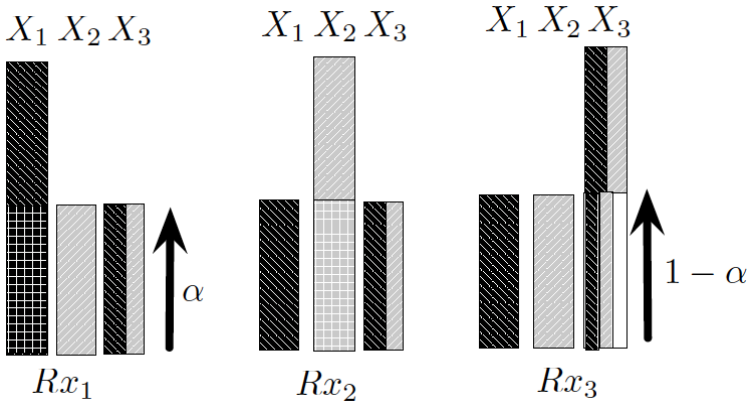


Figure 7. LDC 3-CIFC-CMS in weak interference with $\alpha = 1/2$. The achievable rates are $R_1/n_d = R_2/n_d = 1, R_3/n_d = 1-\alpha$ thereby achieving the sum-capacity outer bound in (Equation 3.8) under the condition in (Equation 3.10). Dark black bits are intended to Rx_1, gray bits are intended to Rx_2 and white bits are intended to Rx_3.

optimal; similarly the primary user cannot reach its intended destination and cannot deliver any information to the other destinations, hence $X_1 = 0$ is optimal); the capacity region of a deterministic broadcast channel is known [47] and for the symmetric LDC with $n_d = 0$ it reduces to $R_1 + R_2 = 2n_i$.

When $n_d > 0$ the sum-capacity can be expressed as

$$d_\Sigma(\alpha;3) = \max\{1,\alpha\} + \frac{f(n_d, n_d\ \alpha; n_d\ \alpha, n_d\ \alpha)}{n_d} + [1-\alpha]^+$$
$$= \begin{cases} 3\max\{1,\alpha\} - \alpha & \text{for } \alpha \neq 1, \\ 1 & \text{for } \alpha = 1. \end{cases}$$

Figure 7 shows an example of the achievable strategy for *weak interference* defined as $\alpha < 1$ (corresponding to Case 2 in Section 3.4.1). The case $\alpha = 1$ corresponds to a channel where all received signals are statistically equivalent and therefore its capacity region is that of a 3-user Multiple Access Channel. The *strong interference* regime defined as $\alpha > 1$ (corresponding to Case 1 in Section 3.4.1) is not explicitly considered as the achievable strategy is the same as for the weak interference regime except for the fact that the most cognitive user sends at zero rate, as its bits would create interference at the non-intended receivers. Notice the important role of cognition in Figure 7: the third transmitter (cognitive of all 3 messages) sends a linear combination of the messages of users 1 and 2 in such a way that the effect of the aggregate interference is neutralized at all receivers, i.e., leaving the receivers of 1 and 2 interference-free. The third transmitters also sends some "private" information bits in such a way that these bits do not appear at the other receivers. It is important also to observe that user 2, who is cognizant of the message of user 1, does not use this message knowledge. In other words, user 2 need not be cognizant in order to achieve the sum-capacity in the symmetric case.

3.4.4 Sum-capacity outer bound for the K-user case and symmetric channel gains

For the K-user symmetric LDC the sum-capacity is upper bounded by

$$d_\Sigma(\alpha, K) =\leq \begin{cases} K\max\{1,\alpha\} - \alpha & \text{for } \alpha \neq 1, \\ 1 & \text{for } \alpha = 1. \end{cases} \tag{3.11}$$

The proof that the sum-capacity upper bound in Th. 3.3.2 evaluates to the expression in (Equation 3.11) is provided next. For the K-user symmetric LDC with $m = n_d \max\{1, \alpha\}$ the sum-capacity is upper bounded by

$$\begin{aligned}
\sum_{k=1}^{K} R_k &\leq \sum_{k=1}^{K} H\left(Y_k | X_1, \ldots, X_{k-1}, Y_1, \ldots, Y_{k-1}\right) \\
&= \sum_{k=1}^{K-1} H\left(\mathbf{S}^{m-n_d} X_k + \mathbf{S}^{m-n_i}\left(\sum_{i=k+1}^{K} X_i\right) \Big| X_1, \ldots, X_{K-1}, \mathbf{S}^{m-n_i}\left(\sum_{i=k}^{K} X_i\right)\right) \\
&+ H\Big(\mathbf{S}^{m-n_d} X_K | X_1, \ldots, X_{K-1}, \mathbf{S}^{m-n_i} X_K\Big) \\
&\leq \sum_{k=1}^{K-1} H\Big((\mathbf{S}^{m-n_d} + \mathbf{S}^{m-n_i}) X_k\Big) + H\Big(\mathbf{S}^{m-n_d} X_K | \mathbf{S}^{m-n_i} X_K\Big) \\
&\leq (K-1)\max\{n_d, n_i\} + [n_d - n_i]^+ \\
&= n_d\Big(K \max\{1, \alpha\} - \alpha\Big).
\end{aligned}$$

The discontinuity at $\alpha = 1$ in (Equation 3.11) is because when $n_d = n_i$ all received signal are equivalent, i.e., $Y_1 = \ldots = Y_K = \sum_{i=1}^{K} X_i$, and the channel reduces to a K-user MAC with sum-capacity $\max H(Y_1) = n_d$.

3.4.5 Achievability of the sum-capacity outer bound for the K-user case and symmetric channel gains

The schemes which were shown to be optimal for the LCD 3-CIFC-CMS in Section 3.4.3 may be extended to any arbitrary number of users. Let U_j, $j \in [1 : K]$, be the signal intended

for receiver j, that is, U_j is only a function of message W_j, and composed of i.i.d. Bernoulli(1/2) bits. Let the transmit signals be

$$X_j = U_j,\ j \in [1 : K-1],$$

$$X_K = \begin{bmatrix} I_{n_i} & 0_{n_i \times [n_d - n_i]^+} \\ 0_{[n_d - n_i]^+ \times n_i} & 0_{[n_d - n_i]^+ \times [n_d - n_i]^+} \end{bmatrix} \left(\sum_{j=1}^{K-1} U_j \right) + \begin{bmatrix} 0_{n_i \times n_i} & 0_{n_i \times [n_d - n_i]^+} \\ 0_{[n_d - n_i]^+ \times n_i} & I_{[n_d - n_i]^+} \end{bmatrix} U_K,$$

$$\sum_{j=1}^{K} X_j = \begin{bmatrix} 0_{n_i \times n_i} & 0_{n_i \times [n_d - n_i]^+} \\ 0_{[n_d - n_i]^+ \times n_i} & I_{[n_d - n_i]^+} \end{bmatrix} \left(\sum_{j=1}^{K} U_j \right),$$

where $0_{n \times m}$ indicates the all zero matrix of dimension $n \times m$ and I_n the identity matrix of dimension n. With these choices, the signal at receiver ℓ, $\ell \in [1 : K]$, is

$$\begin{aligned} Y_\ell &= (\mathbf{S}^{m-n_d} + \mathbf{S}^{m-n_i}) X_\ell + \mathbf{S}^{m-n_i} \left(\sum_{j=1}^{K} X_j \right) \\ &= (\mathbf{S}^{m-n_d} + \mathbf{S}^{m-n_i}) X_\ell, \quad m = \max\{n_d, n_i\}. \end{aligned}$$

Since the matrix $\mathbf{S}^{m-n_d} + \mathbf{S}^{m-n_i}$ is full rank for $n_d \neq n_i$, receiver ℓ, $\ell \in [1 : K]$, decodes U_ℓ from $(\mathbf{S}^{m-n_d} + \mathbf{S}^{m-n_i})^{-1} Y_\ell = X_\ell$. Hence receiver ℓ, $\ell \in [1 : K-1]$, decodes $m = \max\{n_d, n_i\}$ bits since $X_\ell = U_\ell$, while receiver K can decodes the lower $[n_d - n_i]^+$ bits of U_K from X_K.

Interestingly, receivers from 1 to $K-1$ are interference free, while receiver K decodes n_i bits of the "interference function" $\sum_{j=1}^{K-1} U_j$. Notice that cognition is only needed at one transmitter

in all interference regimes. This implies that this sum-capacity result holds for all cognitive channels where user i is cognizant of any subset (including the empty set) of the messages of users with index less than i. We suspect that the fact that only the last user need cognition of all the other messages is a consequence of: 1) the extreme symmetry in the channel model (which is needed for analytical tractability), which naturally aligns the interfering signals at all users. Thus, if the most cognitive user cancels interference at one receiver, it essentially cancels it at all receivers by symmetry. 2) the LDA channel model in which "coherent" gains often seen in Gaussian channels, when two users have the same message may beam form that message to a particular receiver at higher rates, is not possible. That is, the modulo 2 addition at a bit-wise level prohibits such coherent gains and as such it may not be useful to share the messages with other transmitters since the last fully cognitive user is already eliminating interference and additional gains are not possible. We note that these are heuristic rather than rigorous statements, and we do not expect this to hold for Gaussian channels where coherent gains are possible.

3.4.6 Sum-capacity comparison between different channel models

We compare the symmetric sum-capacity of channels with different levels of cognition. Our base line for comparison is the K-user interference channel without any cognition, whose sum-capacity is [48]

$$d_{\Sigma}^{(IFC)}(\alpha; K) = \frac{K}{2} d_{\Sigma}^{(IFC)}(\alpha; 2) \tag{3.12}$$

and where $d_\Sigma^{(\mathrm{IFC})}(\alpha;2)$ is the so-called W-curve of [44] except for a discontinuity at $\alpha = 1$ where $d_\Sigma^{(\mathrm{IFC})}(\alpha;K) = 1$ for all K [48]. Note that, except at $\alpha = 1$, the normalized sum-capacity $\frac{1}{K} d_\Sigma^{(\mathrm{IFC})}(\alpha;K)$ does not depend on K.

At the other extreme of message cognition, consider the case where all users are cognitive of all messages. In this case the channel is equivalent to a MIMO-BC with K transmit antennas and K single-antenna receivers. The system may zero-force the interference to obtain

$$d_\Sigma^{(\mathrm{BC})}(\alpha;K) = K \max\{1,\alpha\}, \tag{3.13}$$

except for a discontinuity at $\alpha = 1$ where $d_\Sigma^{(\mathrm{BC})}(\alpha;K) = 1$, since in this case all the receivers are statistically equivalent and TDMA is optimal. Also in this case, except at $\alpha = 1$, the normalized sum-capacity $\frac{1}{K} d_\Sigma^{(\mathrm{BC})}(\alpha;K)$ does not depend on K.

The sum-capacity of the symmetric LDC K-CIFC-CMS is given by (Equation 3.11), which is a function of K even after normalization by K, i.e.,

$$\frac{1}{K} d_\Sigma^{(\mathrm{CIFC-CMS})}(\alpha;K) = \max\{1,\alpha\} - \frac{\alpha}{K}. \tag{3.14}$$

This has the interesting interpretation that CMS looses α/K with respect to $d_\Sigma^{(\mathrm{BC})}(\alpha;K)/K$. In other words, as the number of cognitive users increases the CMS sum-capacity approaches the sum-capacity of a fully coordinated broadcast channel, which is intuitive.

Figure 8 shows the sum-capacity normalized by the number of users for different channel models; we do not show the discontinuity at $\alpha = 1$. We note the increase in performance in all

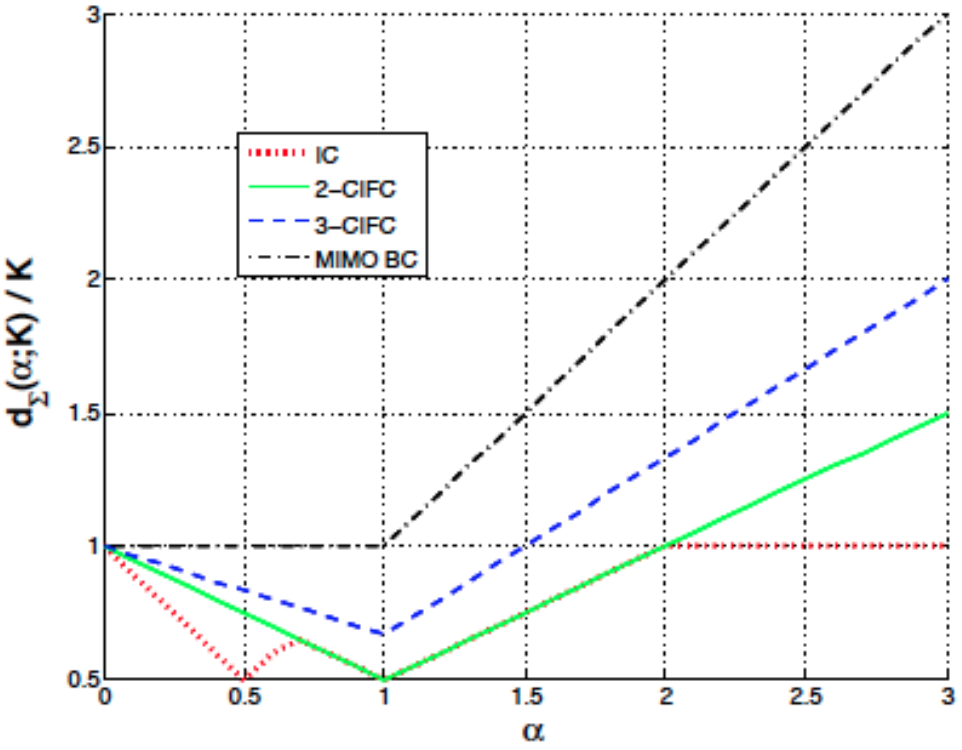


Figure 8. Normalized Generalized Degrees of Freedom of the K-user MIMO broadcast, interference, CIFC-CMS channels (normalized by the number of transmitters).

interference regimes when compared to that of the 2-user CIFC-CMS and the K-user interference channel, but a loss with respect to the K-user broadcast channel (BC) with K transmit antennas and K single antenna receivers.

3.5 Sum-Capacity for the Gaussian K-CIFC-CMS to within a constant gap

In this section we derive the sum-capacity for the symmetric Gaussian channel with an arbitrary number of users to within a constant gap. For notational convenience we denote the direct link gains as $|h_d|$ and the interference link gains as h_i, so that the channel in (Equation 3.1) can be rewritten as

$$Y_\ell = \Big(|h_d| - h_i\Big)X_\ell + h_i\Big(\sum_{j=1}^{K} X_j\Big) + Z_\ell,\ \ \ell \in [1:K].$$

The main results of this section are

Theorem 3.5.1. *The generalized Degrees-of-Freedom of the symmetric K-user Gaussian noise channel are*

$$d(\alpha) = K\max\{1,\alpha\} - \alpha,$$

with a discontinuity at $\alpha = 1$ in the special case where all channel gains are the same (in modulo and phase), in which case $d(1) = 1$.

Theorem 3.5.2. *The sum-capacity bound for $i = 1$ in* (Equation 3.7b) *is achievable for the symmetric Gaussian K-CIFC-CMS to within 6 bits per channel use for $K = 3$ and to within $(K-2)\log_2(K-2) + 3.88$ bits per channel use for $K \geq 4$.*

Theorem 3.5.3. *The sum-capacity bound for* $i = 1$ *in* (Equation 3.7b) *is achievable to within a factor* K *by beamforming to the primary user.*

3.5.1 Sum-capacity outer bound for the K-user case and symmetric channel gains

For the K-user symmetric Gaussian channel with $|h_d| \neq h_i$ the bound for $i = 1$ in (Equation 3.7b) can be further bounded as follows. We note that we may tighten the bound by choosing the "worst noise covariance matrix", but for simplicity, here we use independent noises.

$$\begin{aligned}
\sum_{k=1}^{K} R_k &\leq \sum_{u=1}^{K} I\Big(X_u, \cdots, X_K; Y_u \Big| X_1, Y_1, \cdots, X_{u-1}, Y_{u-1}\Big) = I\Big(X_1, \cdots, X_K; |h_d| X_1 + h_i \sum_{i=2}^{K} X_i + Z_1\Big) \\
&+ \sum_{u=2}^{K-1} I\Big(X_u, \cdots, X_K; |h_d| X_u + h_i \sum_{i=u+1}^{K} X_i + Z_u \Big| X_\ell,\ h_i \sum_{i=u}^{K} X_i + Z_\ell,\ \ell \in [1 : u-1]\Big) \\
&+ I\Big(X_K; |h_d| X_K + Z_K \Big| X_\ell,\ h_i X_K + Z_\ell,\ \ell \in [1 : K-1]\Big) \\
&\leq h\Big(|h_d| X_1 + h_i \sum_{i=2}^{K} X_i + Z_1\Big) - h(Z_1) + \sum_{u=2}^{K-1} h\Big([|h_d| - h_i] X_u + Z_u - Z_{u-1}) - h(Z_u) \\
&+ h\Big(|h_d| X_K + Z_K \Big| h_i X_K + \frac{1}{K-1} \sum_{\ell=1}^{K-1} Z_\ell\Big) - h(Z_K).
\end{aligned}$$

Finally, by the "Gaussian maximizes entropy" principle, we obtain

$$\sum_{k=1}^{K} R_k \leq \log\left(1 + \Big(|h_d| + (K-1)|h_i|\Big)^2\right) \tag{3.15a}$$

$$+ (K-2)\log(2) + (K-2)\log\left(1 + \frac{\big||h_d| - h_i\big|^2}{2}\right) \tag{3.15b}$$

$$+ \log\left(1 + \frac{|h_d|^2}{1 + (K-1)|h_i|^2}\right). \tag{3.15c}$$

For $h_i = |h_d|$ all received signals are statistically equivalent, therefore the K-CIFC-CMS is equivalent to a K-user Multiple Access Channel, whose sum-capacity is

$$\sum_{k=1}^{K} R_k \leq I(X_1, \ldots, X_K; |h_d| \sum_{i=1}^{K} X_i + Z_1) \leq \log(1 + K^2 |h_d|^2).$$

In the limit for high SNR and with the channel parameterization as in (Equation 3.2), the above outer bound can be further bounded

$$\sum_{k=1}^{K} R_k \leq \log(K^2) + (K-1)\log(2) + (K-1)\log\left(1 + \max\{|h_d|^2, |h_i|^2\}\right) + \log\left(1 + \frac{|h_d|^2}{1 + (K-1)|h_i|^2}\right),$$

to obtain the following gDoF outer bound

$$d(\alpha) \leq (K-1)\max\{1, \alpha\} + [1-\alpha]^+ = K\max\{1, \alpha\} - \alpha.$$

This gDoF remains valid for $\alpha = 1$ as long as $h_i = |h_d| \exp(j\theta)$ for $\exp(j\theta) \neq 1$; when $\exp(j\theta) = 1$ the K-user MAC sum-capacity gives $d(\alpha = 1) = 1$. This proves the converse part of Th. 3.5.1.

3.5.2 Achievable Rate Region K-CIFC with CMS

We now present a scheme which will be used in Section 3.5.3 to show that the symmetric outer bound derived in Section 3.5.1 is achievable to within a constant gap. Inspired by the capacity achieving strategy for the Gaussian MIMO-BC, we introduce a scheme that uses Dirty Paper Coding (DPC) with encoding order $1 \to 2 \to 3 \to \cdots K$. We denote by Σ_ℓ the covariance

matrix corresponding to the message intended for decoder ℓ, $\ell \in [1:K]$, as transmitted across the K antennas/transmitters. The overall input covariance matrix is

$$\mathrm{Cov}[X_1,\ldots,X_K] = \sum_{\ell=1}^{K} \Sigma_\ell : \left[\sum_{\ell=1}^{K} \Sigma_\ell\right]_{k,k} \le 1,\ k \in [1:K], \tag{3.16a}$$

where the constraints on the diagonal elements correspond to the input power constraints. Moreover, since message ℓ can only be broadcasted by transmitters with index larger than ℓ, we further impose

$$\left[\Sigma_\ell\right]_{k,k} = 0 \text{ for all } 1 \le k < \ell \le K. \tag{3.16b}$$

The achievable rate region is then the set of non-negative rates $(R_1,\ldots,R_K)$ that satisfy

$$R_\ell \le \log\left(1 + \frac{\mathbf{h}_\ell^\dagger \Sigma_\ell \mathbf{h}_\ell}{\mathbf{h}_\ell^\dagger \left(\sum_{k=\ell+1}^{K} \Sigma_k\right) \mathbf{h}_\ell}\right),\ \mathbf{h}_\ell^\dagger := [h_{\ell,1} h_{\ell,2} \ldots h_{\ell,K}],\ \ell \in [1:K], \tag{3.17}$$

for all possible $\mathrm{Cov}[X_1,\ldots,X_K]$ complying with (Equation 3.16), with the convention that $\sum_{k=K+1}^{K} \Sigma_k = 0$.

In particular we consider the transmit signals

$$\begin{aligned}
X_1 &= \alpha_1 U_1,\\
X_j &= \gamma_j U_j + \beta_j U_j^{(ZF)} + \alpha_j U_1,\ j \in [2:K-1],\\
X_K &= \gamma_K U_K - \beta_K \sum_{j=2}^{K-1} U_j^{(ZF)} + \alpha_K U_1,
\end{aligned}$$

where $U_\ell, U_\ell^{(ZF)}$ are i.i.d. $\mathcal{N}(0,1), \ell \in [1:K]$, and the coefficients $\{\alpha_1, \alpha_j, \beta_j, \gamma_j\}_{j\in[2:K]}$ satisfy

$$
\begin{aligned}
&|\alpha_1|^2 \le 1,\\
&|\gamma_j|^2 + |\beta_j|^2 + |\alpha_j|^2 \le 1,\ j \in [2:K-1],\\
&|\gamma_K|^2 + |\beta_K|^2(K-2) + |\alpha_K|^2 \le 1,
\end{aligned}
$$

in order to satisfy the power constraints. Notice the negative sign for β_K, which we shall use to implement zero-forcing of the aggregate interference $\sum_{j=2}^{K-1} U_j^{(ZF)}$. Moreover, all transmitters cooperate in beam forming U_1 to receiver 1. These two facts can be easily seen by observing that for $\beta_1 = \ldots = \beta_K := \beta$

$$
\sum_{\ell=1}^{K} X_\ell \Big|_{\beta_1=\ldots=\beta_K} = \sum_{\ell=1}^{K} \gamma_\ell U_\ell, \quad \gamma_1 := \sum_{\ell=1}^{K} \alpha_\ell.
$$

With these choices the message covariance matrices are

$$
\begin{aligned}
&\Sigma_1 = \mathbf{a}\mathbf{a}^\dagger,\ \mathbf{a} := [\alpha_1, \ldots, \alpha_K]^T,\\
&\Sigma_j = |\gamma_j|^2\ \mathbf{e}_j\mathbf{e}_j^\dagger + |\beta|^2\ (\mathbf{e}_j - \mathbf{e}_K)(\mathbf{e}_j - \mathbf{e}_K)^\dagger,\ j \in [2:K],
\end{aligned}
$$

where $\mathbf{e}_j$ indicates a length-K vector of all zeros except for a one in position j, $j \in [1:K]$, $\dagger$ indicates the Hermitian transpose, and where $\beta = \beta_1 = \ldots = \beta_K$.

We next express the channel vectors $\mathbf{h}_\ell$ for the symmetric Gaussian channel as

$$\mathbf{h}_\ell = (|h_d| - h_i)\ \mathbf{e}_\ell + h_i\ \left(\sum_{k=1}^{K} \mathbf{e}_k\right),\ \ell \in [1:K].$$

By noticing that

$$\mathbf{h}_\ell \mathbf{e}_j^\dagger = \delta[\ell - j](|h_d| - h_i) + h_i,\ \ell \in [1:K],$$

where $\delta[k]$ is the Kronecker's delta function, the following rates are achievable

$$R_1 = \log\left(1 + \frac{\left||h_d| + |h_i| \sum_{j=2}^{K} \alpha_j\right|^2}{1 + |h_i|^2 \sum_{k=2}^{K} |\gamma_k|^2}\right), \tag{3.18a}$$

$$R_j = \log\left(1 + \frac{\left||h_d| - h_i\right|^2 |\beta|^2 + |h_d|^2\ |\gamma_j|^2}{1 + |h_i|^2 \sum_{k=j+1}^{K} |\gamma_k|^2}\right), j \in [2:K-1], \tag{3.18b}$$

$$R_K = \log\left(1 + |h_d|^2\ |\gamma_K|^2\right), \tag{3.18c}$$

where we chose $\alpha_1 = \exp(j\angle h_i)$ to allow coherent combining at receiver 1 of the different signals carrying U_1, i.e., all users beamform to the primary receiver.

3.5.3 Additive Constant Gap Results for the symmetric Gaussian Channel

We now choose the parameters in (Equation 3.18) so as to match the outer bound given in (Equation 3.15). Due to the presence of the term $(K-1)|h_i|^2$ in the denominator of the equivalent SNR for receiver K, one might be tempted to suggest that the bound in (Equation 3.15c) would mean that the most cognitive user should treat all the other signals as noise. However we recall that user K is the most cognitive user and can therefore "pre-code" the whole interference

seen at its receiver using DPC; by doing so, receiver K would not have anything to treat as noise besides the Gaussian noise itself. We therefore interpret the term $\frac{1}{1+(K-1)|h_i|^2} \leq 1$ as the fraction of power transmitter K dedicates to the transmission of its own signal. This amounts to setting

$$|\gamma_K|^2 = \frac{1}{1+(K-1)|h_i|^2}$$

in (Equation 3.18c). This choice guarantees that the achievable rate for user K exactly matches the term in (Equation 3.15c) in the upper bound.

Next we would like to match the upper bound term in (Equation 3.15b) to the achievable rates in (Equation 3.18b) by setting

$$\gamma_j = 0,\ j \in [2, K-1], \quad \frac{1}{2} = \frac{|\beta|^2}{1+|h_i|^2|\gamma_K|^2}.$$

However, from the power constraint for user K, we must satisfy

$$|\beta|^2 \leq \frac{1-|\gamma_K|^2}{K-2},$$

which imposes the following condition

$$\frac{K-4}{K-2} + \left(|h_i|^2 + \frac{2}{K-2}\right)|\gamma_K|^2 \leq 0.$$

The above condition cannot be satisfied for $K \geq 4$; for $K = 3$ it requires that

$$|\gamma_3|^2 = \frac{1}{1+2|h_i|^2} \leq \frac{1}{|h_i|^2+2}$$

which can be satisfied by $|h_i|^2 \geq 1$. Therefore, in the following we shall assume $|h_i|^2 \geq 1$ and set $\gamma_j = 0,\ j \in [2, K-1]$ and

$$|\beta|^2 = |\beta|^2 = \begin{cases} \frac{1-|\gamma_K|^2}{K-2} = \frac{1}{K-2}\left(1 - \frac{1}{1+(K-1)|h_i|^2}\right) & K \geq 4 \\ \frac{1+|h_i|^2|\gamma_3|^2}{2} = \frac{1+3|h_i|^2}{2(1+2|h_i|^2)} & K = 3 \end{cases},$$

which implies

$$|\alpha_K|^2 = \begin{cases} 0 & K \geq 4 \\ 1 - |\beta|^2 - |\gamma_K|^2 = \frac{-1+|h_i|^2}{2(1+2|h_i|^2)} & K = 3 \end{cases}.$$

Finally, for $j \in [2 : K-1]$

$$|\alpha_j|^2 = 1 - |\beta_j|^2 = \begin{cases} \frac{K-3}{K-2} + \frac{1}{K-2}\ \frac{1}{1+(K-1)|h_i|^2} & K \geq 4 \\ \frac{1+|h_i|^2}{2(1+2|h_i|^2)} & K = 3 \end{cases}.$$

The rates then become: for $K \geq 4$ and for $j \in [2 : K-1]$

$$R_K = \log\left(1 + \frac{|h_d|^2}{1+(K-1)|h_i|^2}\right)$$

$$R_j = \log\left(1 + \left||h_d| - h_i\right|^2 \frac{\frac{1}{K-2}\ \frac{(K-1)|h_i|^2}{1+(K-1)|h_i|^2}}{1 + \frac{|h_i|^2}{1+(K-1)|h_i|^2}}\right) \geq \log\left(1 + \frac{\left||h_d| - h_i\right|^2}{K-2}\frac{K-1}{K+1}\right), \text{since } |h_i|^2 \geq 1,$$

$$R_1 = \log\left(1 + \frac{\left||h_d| + |h_i|\sqrt{(K-3)(K-2) + \frac{K-2}{1+(K-1)|h_i|^2}}\right|^2}{1 + \frac{|h_i|^2}{1+(K-1)|h_i|^2}}\right)$$
$$\geq \log\left(1 + \frac{\left||h_d| + |h_i|\sqrt{(K-3)(K-2)}\right|^2}{2}\right), \text{since } |h_i|^2 \geq 1,$$

and for K=3, the rates are

$$R_3 = \log\left(1 + \frac{|h_d|^2}{1+2|h_i|^2}\right)$$
$$R_2 = \log\left(1 + \left||h_d| - h_i\right|^2 \frac{1}{2}\right)$$
$$R_1 = \log\left(1 + \frac{\left||h_d| + |h_i|\ \left(\sqrt{\frac{1+|h_i|^2}{2(1+2|h_i|^2)}} + \sqrt{\frac{-1+|h_i|^2}{2(1+2|h_i|^2)}}\right)\right|^2}{1 + \frac{|h_i|^2}{1+2|h_i|^2}}\right)$$
$$\geq \log\left(1 + \frac{\left||h_d| + |h_i|\frac{1}{2}\right|^2}{2}\right), \text{ since } |h_i|^2 \geq 1.$$

By taking the difference between the outer bound in (Equation 3.15) and the lower bounds on the derived achievable rates we find that the gap is upper bounded by: for $K \geq 4$

$$GAP \leq (K-2)\log(2) + (K-2)\left(\log\left(1 + \frac{\left||h_d| - h_i\right|^2}{2}\right) - \log\left(1 + \frac{\left||h_d| - h_i\right|^2}{K-2}\frac{K-1}{K+1}\right)\right)$$
$$+ \log\left(1 + \left(|h_d| + (K-1)|h_i|\right)^2\right) - \log\left(1 + \frac{\left||h_d| + |h_i|\sqrt{(K-3)(K-2)}\right|^2}{2}\right)$$
$$\leq (K-2)\log(2) + (K-2)\log\left(\frac{(K+1)(K-2)}{2(K-1)}\right) + \log\left(\frac{2(K-3)(K-2)}{(K-1)^2}\right)$$
$$\leq (K-2)\log\left(K-2\right) + \log(2\exp(2)),$$

(where we used $\mathsf{K}\log_e(1+1/\mathsf{K}) \le 1$) and for $\mathsf{K} \ge 3$

$$\begin{aligned}\mathsf{GAP} &\le \log(2) + \log\left(1 + \left(|h_d| + 2|h_i|\right)^2\right) - \log\left(1 + \frac{\left||h_d| + |h_i|\frac{1}{2}\right|^2}{2}\right) \\ &\le 6\log(2).\end{aligned}$$

For $|h_i|^2 < 1$, we set $\beta_j = \alpha_j = 0, \gamma_j = 1$ for $j \in [2:\mathsf{K}]$ to obtain

$$\sum_{\ell=1}^{\mathsf{K}} \mathsf{R}_\ell = \sum_{\ell=1}^{\mathsf{K}} \log\left(1 + \frac{|h_d|^2}{1 + (\mathsf{K}-\ell)|h_i|^2}\right).$$

The gap to the upper bound is at most

$$\mathsf{GAP} \le (\mathsf{K}-2)\log(2) + 2\log(\mathsf{K}-1) + \sum_{\ell=2}^{\mathsf{K}-1} \log\left(\frac{\mathsf{K}-\ell}{2}\right),$$

which is smaller than the gap previously obtained for $|h_i|^2 \ge 1$. This proves Th. 3.5.2 and implies the direct part of Th. 3.5.1.

3.5.4 Multiplicative Constant Gap for the symmetric Gaussian Channel

In order to provide a complete characterization of the sum-capacity of the symmetric Gaussian channel we next consider approximating the sum-capacity to within a multiplicative gap, which is more relevant at low SNR than an additive gaps. To this end, note that the rate of user j is upper bounded by $\mathsf{C}_j := \log(1 + (|h_d| + (\mathsf{K}-j)|h_i|)^2), j \in [1:\mathsf{K}]$ which in turn is upper bounded by $\mathsf{K} \times \mathsf{C}_1$. Consider an achievability scheme in which all users beamform to user 1:

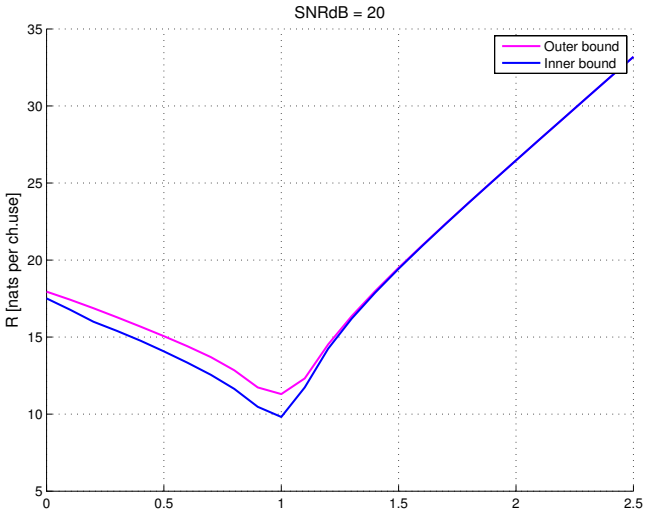


Figure 9. Comparison of the numerically optimized inner and outer bounds for $K = 3$ users at SNR$= 20$dB as a function of $\alpha = \frac{\log(|h_d|)}{\log(|h_i|)}$; notice a smaller gap than the worst case predicted 6 bits per channel use.

this achieves the sum-rate $R_1 + \cdots R_K = C_1$. This is to within a factor K of the upper bound, proving Th. 3.5.3.

3.6 Numerical optimization of inner and outer bounds

Figure 9 shows the proposed upper and lower bounds for the symmetric channel with $K = 3$ users at SNR$= 20$dB. In this case the outer and lower bounds where optimized numerically so as to obtain a larger achievable rate and a tighter outer bound than those used for the analytical evaluation of the gap. We notice that the gap between the bounds is much less than the theoretical gap of 6 bits. In particular, for strong interference the bounds are extremely close to one another, showing again that the analytically provable gap of 6 bits is a worst case scenario,

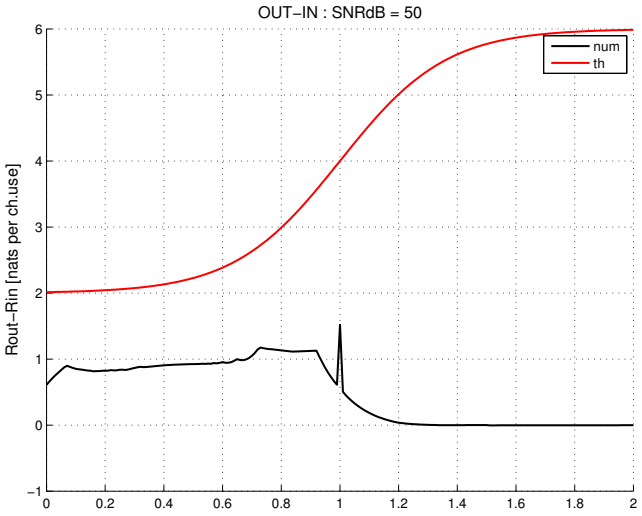
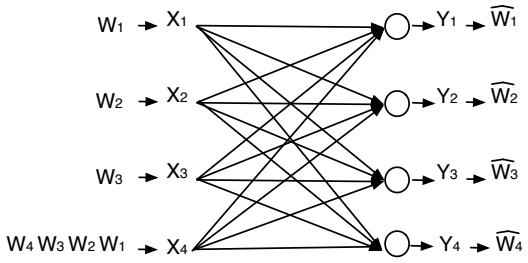


Figure 10. Analytical and numerical additive gaps for $\mathsf{K} = 3$ users at SNR$= 50$dB.

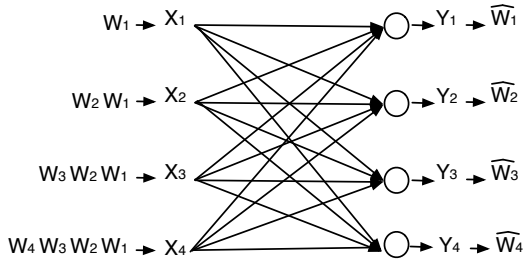
which is the result of crude bounding techniques rather than a poor achievability scheme. Figure 10 shows the additive gap for $\mathsf{K} = 3$ users at SNR$= 50$dB; notice the gap between the analytical upper and lower bounds (curve labeled 'th') converging to 6 bits for large α while the gap between the numerically optimized upper and lower bounds (curve labeled 'num') going to zero in the same regime; the largest gap is at $\alpha = 1$ where the channel matrix becomes rank deficient; overall the gap is at most around 1 bit, 5 bits smaller than the analytically provable gap.

3.7 K user Cognitive Interference Channel in Strong Interference

In this section we consider - in addition to the (K-CIFC-CMS) - the cognitive interference channel with cognitive only message sharing (K-CIFC-CoMS) consisting of $\mathsf{K}-1$ primary users



(a) K-CIFC-CoMS with $\mathsf{K} = 4$.



(b) K-CIFC-CMS with $\mathsf{K} = 4$.

and one cognitive user that has non-causal message knowledge of the all the primary users. The two message assignments $(\mathcal{M}_1, \ldots, \mathcal{M}_\mathsf{K})$ can be described as follows:

1. CoMS: $\mathcal{M}_\mathsf{i} = \{\mathsf{i}\}$, $\mathsf{i} \in [1 : \mathsf{K} - 1]$, and $\mathcal{M}_\mathsf{K} = [1 : \mathsf{K}]$,
2. CMS: $\mathcal{M}_\mathsf{i} = [1 : \mathsf{i}]$, $\mathsf{i} \in [1 : \mathsf{K}]$.

The general memoryless (K-CIFC-CoMS) is formally defined by

- channel inputs $\mathsf{X}_\mathsf{i} \in \mathcal{X}_\mathsf{i}$,
- channel outputs $\mathsf{Y}_\mathsf{i} \in \mathcal{Y}_\mathsf{i}$, $\mathsf{i} \in [1 : \mathsf{K}]$,
- and a memoryless channel $\mathbb{P}(\mathsf{Y}_1, \ldots, \mathsf{Y}_\mathsf{K} | \mathsf{X}_1, \ldots, \mathsf{X}_\mathsf{K})$.
- Messages W_i which are known to user K non causally.

A code with non-negative rate vector $(\mathsf{R}_1, \ldots, \mathsf{R}_\mathsf{K})$ and block-length N is defined by:

- messages W_i, uniformly distributed over $[1 : 2^{\mathsf{N}\mathsf{R}_\mathsf{i}}]$ and independent of all other random variables, where $\mathsf{N} \in \mathbb{N}^+$ is the codeword length and $\mathsf{R}_\mathsf{i} \in \mathbb{R}^+$ is the rate expressed in bits per channel use.

- encoding functions $X_i^N(W_{\mathcal{M}_i})$ with $\mathcal{M}_i \subseteq [1:K]$, $f_i^{(N)} : [1 : 2^{NR_i}] \to \mathcal{X}_i^N$ such that $X_i^N := f_i^{(N)}(W_i)$, $i \in [1 : K-1]$,
- and decoding functions $\widehat{W}_i(Y_i^N)$, for $i \in [1:K]$. $g_i^{(N)} : \mathcal{Y}_i^N \to [1 : 2^{NR_i}]$ such that $\widehat{W}_i = g_i^{(N)}(Y_i^N)$, $i \in [1:K]$,

The capacity region is the set of all rate tuples $(R_1, \ldots, R_K)$ for which there exists a sequence of codes indexed by the block-length N such that $P_e^{(N)} := \max_{i\in[1:K]} \mathbb{P}[\widehat{W}_i \neq W_i] \to 0$ as $N \to +\infty$.

The CoMS and CMS channels are shown in Fig. 11(a) and Fig. 11(b), respectively, for the case of $K = 4$ users.

3.7.1 Outer Bound under strong interference conditions

Our first result is a sum-capacity outer bound under a set of strong interference conditions:

Theorem 3.7.1. *For the* K*-CIFC-CoMS and the* K*-CIFC-CMS satisfying the following condition* $\forall j \in [2:K]$

$$I(X_{[j:K]}; Y_j | X_{[1:j-1]}) \leq I(X_{[j:K]}; Y_{j-1} | X_{[1:j-1]}), \tag{3.19}$$

the sum-capacity is upper bounded by

$$\sum_{j=1}^{K} R_j \leq I(X_{[1:K]}; Y_1), \tag{3.20}$$

for all input distributions $P_{X_1,\ldots,X_K}$ *that factor as*

1. *CoMS:* $\prod_{j=1}^{K-1} P_{X_j} \; P_{X_K|X_1,\ldots,X_{K-1}}$,

2. CMS: $P_{X_1,\ldots,X_K}$.

Proof. The proof can be found in the Appendix F. □

Remarks:

1. The condition in (Equation 3.19) intuitively says that for all $j \in [2:K]$, given that the signals $(X_1,\ldots,X_{j-1})$ have been removed, receiver j can decode the remaining signals $(X_j,\ldots,X_K)$ at a lower rate than receiver $j-1$, which somehow implies that receiver $j-1$ can 'better decode' signal X_j than the indented receiver j. Of all the receivers, receiver 1 is the 'most powerful' and the sum-capacity in (Equation 3.20) can be interpreted as 'joint decoding' of all transmit signals at receiver 1.

2. For $K=2$ the CoMS and the CMS models coincide and the condition in (Equation 3.19) reduces to [18, Eq. (93)], that is, $I(X_2;Y_2|X_1) \leq I(X_2;Y_1|X_1)$ for all P_{X_1,X_2}.

3. In [2] we derived the following sum-capacity upper bound for CMS case without any restriction

$$\sum_{j=1}^{K} R_j \leq \sum_{j=1}^{K} I(Y_j;X_{[j:K]}|X_{[1:j-1]},Y_{[1:j-1]}). \tag{3.21}$$

We notice that (Equation 3.21) and Theorem 3.7.1 coincide for channels that satisfy the following *degradedness* condition

$$X_{[j:K]} \to Y_{j-1} \to Y_j \text{ given } X_{[1:j-1]} \forall j \in [2:K] \text{ and for all possible } P_{X_1,\ldots,X_K}. \tag{3.22}$$

For the 3-user Gaussian channel, we shall see that the condition in (Equation 3.19) is less restrictive than (Equation 3.22).

3.7.2 Inner Bounds

3.7.2.1 Achievability scheme for K-CIFC-CMS

With CMS, message W_1 is known to all users. Thus all users may cooperate in sending message W_1 to receiver 1. In order to achieve the sum-outer bound, all users *beam form* to receiver 1 as in a MISO channel to achieve

$$R_1 = I(X_{[1:K]}; Y_1),$$

$$R_2 = \ldots R_K = 0,$$

for some $P_{X_1,\ldots,X_K}$. Hence, when the condition in (Equation 3.19) is satisfied for all input distributions, the sum-capacity of the K-CIFC-CMS is given by (Equation 3.20).

3.7.2.2 Achievability scheme for K-CIFC-CoMS

With CoMS, messages W_1 through W_{K-1} are known at transmitter K. Here we propose a simple achievable scheme where users 1 through $K-1$ use independent i.i.d coding (like in point-to-point channels) and user K superposes its own message to the codewords generated by the other users. All destinations are required to decode all messages, where non-intended messages are decoded non-uniquely. The achievable region is therefore the intersection of K

multiple access channels (with the difference that X_K can be correlated to all other mutually independent inputs) given by

$$R_{\mathcal{S}} + R_K \leq \min_{j \in [1:K]} I(X_{\mathcal{S}}; Y_j | X_{\mathcal{S}^c}),$$

for all $\mathcal{S} \subseteq [1 : K-1] \backslash \emptyset$ and

$$R_K \leq I(X_K; Y_K | X_{[1:K-1]}).$$

The achievable sum-rate is therefore obtained by $\mathcal{S} = [1 : K-1]$. To meet the sum-rate outer bound in (Equation 3.20) we need to impose the extra condition that

$$I(X_{[1:K]}; Y_1) \leq I(X_{[1:K]}; Y_j), \ j \in [2 : K], \tag{3.23}$$

for the distribution that attains the largest value in the sum-rate upper bound. Hence, when in addition to the condition in (Equation 3.19) being satisfied for the set of input distributions with the prescribed factorization, also the condition in (Equation 3.23) is satisfied, the sum-capacity of K-CIFC-CoMS is given by (Equation 3.20).

3.7.3 The Gaussian noise case

The power-constrained complex-valued single-antenna K-user Gaussian noise channel is described by the following input/output relationship

$$Y_j = \sum_{k\in[1:K]} h_{jk}X_k + Z_j,\ j \in [1:K], \tag{3.24}$$

where the channel gains $h_{jk} \in \mathbb{C}$, $(i,j) \in [1:K]^2$, are constant and known to all terminals, the noises are without loss of generality zero mean, unit variance proper-complex Gaussian random variables (their correlation does not matter as the receivers do not cooperate), and the inputs are subject to power constraint $\mathbb{E}[|X_k|^2] \leq 1$, $k \in [1:K]$.

3.7.3.1 Evaluation of the Outer bound for the Gaussian Noise Channel

In the next subsections we aim to evaluate Theorem 3.7.1 for the channel in (Equation 3.24). To do so, we need to identify the channels for which the strong interference condition which are given in (Equation 3.19) holds for all input distributions with the proper factorization depending on the message sharing mechanism. Next we argue that for the power-constrained Gaussian channel, the strong interference condition must be verified only for those input distributions that meet the power constraint with equality for each user. The idea is that all other distributions are suboptimal in the sense that one can find another distribution with provably better performance. The proof is by contradiction. Assume that there is an optimal input distribution for which user $k \in [1:K]$ uses $\mathbb{E}[|X_k|^2] = P_k \leq 1$ with a user k^* such that $P_{k^*} < 1$. Consider now a new communication scheme in which user k^* sends $X_{k^*,new} = X_{k^*} + X'_{k^*}$ where X_{k^*} is the signal that

was assumed optimal with power $P_{k^*} < 1$ and $X'_{k^*} \sim \mathcal{N}(0, 1 - P_k)$ is independent of everything else and has rate

$$R'_{k^*} = \log\left(1 + \min_{j\in[1:K]} \frac{|h_{jk^*}|^2(1 - P_{k^*})}{1 + (\sum_{\ell\in[1:K]} \sqrt{|h_{j\ell}|P_\ell})^2}\right) > 0$$

Now, the rate of the new message is such that X'_{k^*} can be decoded by all users by treating the signals assumed optimal as noise (no matter what their correlation structure is); after that, X'_{k^*} is removed for the received signal and the system is equivalent to the one assumed optimal. Now, since the rate of user k^* can be increased by $R'_{k^*} > 0$ we reached a contradiction. This shows that all users must use their full power. Therefore, for the power constrained Gaussian channel, one can repeat the same steps of the converse by considering only those input distributions that meet the power constraint with equality for all users. This implies that the strong interference condition must be verified only by these distributions (rather than all possible input distributions).

3.7.3.2 Sum-Capacity for the K-CIFC-CMS

The sum-capacity upper bound in (Equation 3.20), by the 'Gaussian maximizes entropy' theorem, yields

$$\begin{aligned}\sum_{k=1}^{K} R_k &\le \max_{\Sigma_x} \log\left(1 + \mathbf{h}_1^H \Sigma_x \mathbf{h}_1\right) \\ &= \log\left(1 + \Big(\sum_{k\in[1:K]} |h_{1k}|\Big)^2\right), \end{aligned} \tag{3.25}$$

since we can consider any input covariance matrix Σ_x. The sum-capacity in (Equation 3.25) is valid under the condition in (Equation 3.19), which amounts to verifying

$$h(Y_j|X_{[1:j-1]}) \leq h(Y_{j-1}|X_{[1:j-1]}) \; j \in [2:K], \tag{3.26}$$

over all proper-complex Gaussian distribution that meet the power constraint with equality [16–18]. The sum-capacity is achieved by beam forming (see Section 3.7.2.1) with

$$X_k = \exp\{-j\angle h_{1k}\} \, U, \; k \in [1:K], \; U \sim \mathcal{N}(0,1). \tag{3.27}$$

3.7.3.3 Sum-Capacity for the K-CIFC-CoMS

Consider an input covariance matrix:

$$\Sigma_x = \begin{bmatrix} \mathbf{I}_{K-1} & \rho \\ \rho^H & 1 \end{bmatrix} : \|\rho\|^2 \leq 1 \tag{3.28}$$

where $\rho \in \mathbb{C}^{K-1\times 1}$ is a vector of correlation coefficients. The sum-capacity upper bound in (Equation 3.20), by the 'Gaussian maximizes entropy' theorem [49], is maximized by a

jointly Gaussian input with covariance (Equation 3.28). We therefore obtain the following sum-capacity upper bound, for $\mathbf{h}_1^H = [h_{11}, h_{12}, \ldots, h_{1K}]$,

$$\sum_{j=1}^{K} R_j \leq \max_{\Sigma_x \text{ in eq.(Equation 3.28)}} \log\left(1 + \mathbf{h}_1^H \Sigma_x \mathbf{h}_1\right) = \log\left(1 + \left(|h_{1K}| + \sqrt{\sum_{j\in[1:K-1]} |h_{1j}|^2}\right)^2\right) \tag{3.29}$$

attained by $\rho_j = \lambda h_{1j}^*$ for $\lambda : \|\rho\|^2 = 1$. This optimal choice of correlation coefficients implies $R_K = 0$. The sum-capacity in (Equation 3.29) is valid under condition (Equation 3.19), which amounts to verifying

$$h(Y_j|X_{[1:j-1]}) \leq h(Y_{j-1}|X_{[1:j-1]}) \; j \in [2:K], \tag{3.30}$$

for all proper-complex Gaussian distributions with covariance matrix as in (Equation 3.28) [16–18]. The sum-rate in (Equation 3.29) is achievable by (see also Section 3.7.2.2)

$$X_j = T_j \text{ i.i.d. } \mathcal{N}(0,1), \; j \in [1:K-1], \tag{3.31a}$$

$$X_K = \sum_{j=1}^{K-1} T_j \rho_j : |\rho_j| \propto |h_{1j}|, \; j \in [1:K-1], \tag{3.31b}$$

under the condition in (Equation 3.23), that is,

$$\mathbf{h}_1^H \Sigma_x \mathbf{h}_1 \leq \mathbf{h}_j^H \Sigma_x \mathbf{h}_j, \; \forall j \in [2:K-1], \tag{3.32}$$

$$\mathbf{h}_j^H := [h_{j1}, h_{j2}, \ldots, h_{jK}] \tag{3.33}$$

for the choice of correlation coefficients implied by (Equation 3.31). Since $R_K = 0$, the condition in (Equation 3.32) need not to hold for $j = K$ as receiver K does not have anything to decode.

3.7.3.4 The case $K = 2$

For the 2-user case CMS and CoMS coincide. The sum-capacity is given by (Equation 3.25), i.e., $R_1 = \log(1+(|h_{11}|+|h_{12}|)^2)$ and $R_2 = 0$, under the condition in (Equation 3.26) for $K = j = 2$, which is equivalent to

$$\log(1 + |h_{22}|^2) \leq \log(1 + |h_{12}|^2) \Longleftrightarrow |h_{22}|^2 \leq |h_{12}|^2.$$

The achievability condition in (Equation 3.32) does not play a role for $K = 2$ because $R_2 = 0$.

Remark: The strong interference condition $|h_{22}|^2 \leq |h_{12}|^2$ is equivalent to [18, eq.(87)]. However, the strong interference capacity region in [18, Theorem 5] also requires [18, eq.(88)]. This is the case since in order to determine the sum-capacity only less restrictive conditions are needed compared to the case where the whole capacity region must be characterized.

3.7.3.5 The case $K = 3$ with CMS

For CMS and $K = 3$ we consider all jointly Gaussian inputs with covariance matrix given by the following

$$\mathrm{Cov}\begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} = \begin{bmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1^* & 1 & \rho_3 \\ \rho_2^* & \rho_3^* & 1 \end{bmatrix} : |\rho_i| \leq 1, i = 1, 2, 3,$$

$$|\rho_3 - \rho_1 \rho_2^*|^2 \leq (1 - |\rho_1|^2)(1 - |\rho_2|^2).$$

For CMS the optimal sum-rate in (Equation 3.25) is obtained for $R_1 = \log(1 + (|h_{11}| + |h_{12}| + |h_{13}|)^2), R_2 = R_3 = 0$ by beam forming. The condition in (Equation 3.26) is: for $j = 3$, by proceeding similarly to the case $K = j = 2$ discussed previously in Section 3.7.3.4, we have

$$|h_{33}|^2 \leq |h_{23}|^2, \tag{3.34}$$

and for $j = 2$ we must find the channel gains that satisfy

$$\log(1 + \mathbf{h}_2^H \mathbf{S} \mathbf{h}_2) \leq \log(1 + \mathbf{h}_1^H \mathbf{S} \mathbf{h}_1)$$

$$\mathbf{h}_2^* := \begin{bmatrix} h_{22} \\ h_{23} \end{bmatrix}, \mathbf{h}_1^* := \begin{bmatrix} h_{12} \\ h_{13} \end{bmatrix}, \mathbf{S} := \begin{bmatrix} 1 - |\rho_1|^2 & \rho_3 - \rho_1 \rho_2^* \\ \rho_3^* - \rho_1^* \rho_2 & 1 - |\rho_2|^2 \end{bmatrix},$$

which is equivalent to

$$\begin{bmatrix} h_{22} \\ h_{23} \end{bmatrix} = \xi \begin{bmatrix} h_{12} \\ h_{13} \end{bmatrix} : |\xi| \leq 1. \tag{3.35}$$

Remark: The condition in (Equation 3.35) corresponds to the 'degraded channel condition when conditioning on X_1' in (Equation 3.22). Given the message structure of CMS, there are so many coding possibilities at the transmitters that the channel conditions under which joint decoding of all messages at the least cognitive receiver is optimal only includes a form of 'degraded channel'. This suggests that for CMS and generic channel gains, other decoding strategies are sum-capacity optimal, see for example the symmetric sum-capacity result in [2]. Notice that here we did not ask for the conditions under which joint decoding of all messages at all receivers is optimal, i.e., when the channel reduces to compound MAC. If we were to ask for which channel gains joint decoding of all messages at a 'more cognitive receiver' than receiver 1 is optimal, we would generally find different conditions than the one in (Equation 3.35).

3.7.3.6 The case $K = 3$ with CoMS

For CoMS and $K = 3$ we consider all jointly Gaussian inputs with covariance matrix given by the following

$$\mathrm{Cov} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & \rho_2 \\ 0 & 1 & \rho_3 \\ \rho_2^* & \rho_3^* & 1 \end{bmatrix} : |\rho_2|^2 + |\rho_3|^2 \leq 1.$$

For CoMS the optimal sum-rate is obtained when $R_1+R_2 = \log(1+(|h_{13}|+\sqrt{|h_{11}|^2+|h_{12}|^2})^2), R_3 = 0$. The condition in (Equation 3.26) for $j = 3$ is as (Equation 3.34), while for $j = 2$ is

$$|h_{22}|^2 + |h_{23}|^2 + 2|h_{22}h_{23}^* - h_{12}h_{13}^*| \leq |h_{12}|^2 + |h_{13}|^2 \tag{3.36}$$

which includes the 'degraded condition' in (Equation 3.35).

The condition for achievability in (Equation 3.32) evaluated for $j = 2$ imposes that the channel gains satisfy

$$\begin{aligned} &\mathbf{h}_1^H \mathbf{S} \mathbf{h}_1 \leq \mathbf{h}_2^H \mathbf{S} \mathbf{h}_2, \ \mathbf{h}_1^* := \begin{bmatrix} h_{11} \\ h_{12} \\ h_{13} \end{bmatrix}, \mathbf{h}_2^* := \begin{bmatrix} h_{21} \\ h_{22} \\ h_{23} \end{bmatrix}, \\ &\mathbf{S} := \mathrm{Cov} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} \text{ with } \begin{array}{l} \rho_2 = \frac{h_{11}^*}{\sqrt{|h_{11}|^2+|h_{12}|^2}}, \\ \rho_3 = \frac{h_{12}^*}{\sqrt{|h_{11}|^2+|h_{12}|^2}}. \end{array} \end{aligned} \tag{3.37}$$

Remark: Interestingly, the condition in (Equation 3.36) is equivalent to [18, eq.(88)].

3.8 K user Cognitive Interference Channel in Weak Interference

The K-CIFC-CoMS consists of $K-1$ primary (each have an independent message) and one cognitive (with knowledge of all messages) transmitters and K receivers (each interested in one message only). The Gaussian channel is described by the following input-output relationship

$$Y_j = \sum_{k\in[1:K]} h_{jk}X_k + Z_j,\ j \in [1:K], \tag{3.38}$$

with the usual assumptions: the complex-valued channel gains are constant and known to all terminals, the inputs are subject to power constraints $\mathbb{E}[|X_i|^2] \leq 1,\ i \in [1:K]$, the outputs are subject to Gaussian noise $Z_i \sim \mathcal{N}(0,1),\ i \in [1:K]$.

Since the general model in (Equation 3.38) is described by K^2 parameters, in the following we shall focus on the *symmetric* case defined by: for $j \in [1:K-1]$,

$$h_{jj} = |h_d|,\ \text{(primary direct links)}, \tag{3.39a}$$

$$h_{jK} = h_c,\ \text{(secondary}\rightarrow\text{primary links)}, \tag{3.39b}$$

$$h_{jk} = h_i,\ k \notin \{j,K\}\ \text{(primary interfering links)}. \tag{3.39c}$$

Note that our "symmetric setting" in (Equation 3.39) makes the primary users completely equivalent but does not impose any restriction on the channel gains of the cognitive receiver (i.e., $h_{Ki},\ i \in [1:K]$) which are kept general.

The 3-CIFC-CoMS is shown in Fig. 11(a) and the 3-CIFC-CMS in Fig. 11(b). Notice the different message structure at the transmitters, clearly the capacity of the 3-CIFC-CMS is an outer bound to the capacity of the 3-CIFC-CoMS.

$$R_1 + R_2 + R_3 \leq 2\log\left(1 + (|h_d| + |h_c|)^2\right) + \log\left(1 + \frac{|h_{33}|^2}{1 + |h_c|^2}\right), \tag{3.40a}$$

$$R_1 + R_2 + R_3 \leq \log\left(1 + (|h_d| + |h_i| + |h_c|)^2\right) + \log(2) + \log\left(1 + \frac{\left||h_d| - h_i\right|^2}{2}\right) + \log\left(1 + \frac{|h_{33}|^2}{1 + 2|h_c|^2}\right). \tag{3.40b}$$

For the 3-user case, the condition in (Equation 3.39) means that the two primary direct links are equal and are denoted by $|h_d|$ (real valued without loss of generality), the interference cross gains are equal and denoted by h_i, and the "cooperative" cross channel gains from the cognitive transmitter are equal and denoted by h_c.

3.8.1 $K = 3$ user case

For the 3-user case, the outer bound in (Equation 3.7b) gives the sum-rate in (Equation 3.40). Note that although the channel model imposes an input distribution that factors as $P_{X_2,X_1,X_3} = P_{X_1} P_{X_2} P_{X_3|X_2,X_1}$, our bound has to be evaluated over all possible joint input distributions. For the Gaussian noise channel, the outer bound in (Equation 3.7b) is exhausted by jointly Gaussian inputs – by the "Gaussian maximizes entropy" principle – and can be further upper bounded as in (Equation 3.40) at the top of the page, as shown in [2]. Note that in the symmetric

case, the bound in (Equation 3.40b) is to within a constant gap from the sum-capacity of the 3-CIFC-CMS. Our main result is as follows

Theorem 3.8.1. *For the 3-CIFC-CoMS, let*

$$\textbf{\textit{Case 1:}} \quad |h_{33}|^2 \leq |h_c|^2, \quad 2|h_i|^2 \leq |h_c|^2 \leq |h_d|^2, \tag{3.41}$$

$$\textbf{\textit{Case 2:}} \quad |h_{33}|^2 > |h_c|^2, \quad 2|h_i|^2 \leq |h_c|^2 \leq |h_d|^2. \tag{3.42}$$

The sum-capacity bound in (Equation 3.40b) *is achievable to within 8.4 bits when* (Equation 3.41) *is satisfied, and to within 13 bits when* (Equation 3.42) *is satisfied.*

Proof. We consider the following transmit signals

$$X_1 = \alpha_1 T_{1ZF} + \gamma_1 T_{1p}, \tag{3.43a}$$

$$X_2 = \beta_2 T_{2ZF} + \gamma_2 T_{2p}, \tag{3.43b}$$

$$X_3 = -\alpha_3 T_{1ZF} - \beta_3 T_{2ZF} + 0 \cdot T_{3ZF} + \gamma_3 T_{3p}, \tag{3.43c}$$

where T_{iZF}, T_{ip} are independent $\mathcal{N}(0,1)$ random variables for $i \in [1:3]$ and the coefficients should satisfy the following conditions

$$|\alpha_1|^2 + |\gamma_1|^2 \leq 1, \tag{3.43d}$$

$$|\beta_2|^2 + |\gamma_2|^2 \leq 1, \tag{3.43e}$$

$$|\alpha_3|^2 + |\beta_3|^2 + |\gamma_3|^2 \leq 1, \tag{3.43f}$$

in order to satisfy the power constraints. With (Equation 3.43), the received signals are

$$Y_j = (h_{j1}\alpha_1 - h_{j3}\alpha_3)T_{1ZF} + h_{j1}\gamma_1 T_{1p} + (h_{j2}\beta_2 - h_{j3}\beta_3)T_{2ZF} + h_{j2}\gamma_2 T_{2p} + h_{j3}\gamma_3 T_{3p} + Z_j$$
$$j \in [1:3].$$

When $|h_{33}|^2 \leq |h_c|^2$, the terms in (Equation 3.40) depending on h_{33}, which give the rate for the cognitive user, are bounded by $\log\left(1 + \frac{|h_{33}|^2}{1+2|h_c|^2}\right) \leq \log\left(1 + \frac{|h_{33}|^2}{1+|h_c|^2}\right) \leq \log(2)$. This seems to suggest that it is optimal, to within a constant gap, to have $R_3 = 0$ when $|h_{33}|^2 \leq |h_c|^2$, that is, to have the cognitive user use all its resources to help the primary users / behave as a *cognitive relay*. The symmetric capacity region of a 2-user interference channel with cognitive relay was recently characterized to within a constant gap in [50] for almost all parameter regimes. The power splits are chosen to be

$$\gamma_1 = \gamma_2 = \gamma_3 = 0, \tag{3.44a}$$
$$\alpha_1 = \beta_2 = \alpha_3 = \beta_3 = \frac{h_i}{h_c}. \tag{3.44b}$$

Having the interference completely neutralized, the following rates are achievable

$$R_1 \geq \log\left(1 + \left||h_d| - h_i\right|^2\right), \tag{3.45a}$$
$$R_2 \geq \log\left(1 + \left||h_d| - h_i\right|^2\right). \tag{3.45b}$$

Note that we can further lower bound R_1 using (Equation 3.41) by

$$R_1 \geq \log\left(1 + \left(|h_d| - |h_i|\right)^2\right) \geq \log\left(1 + (1 - \frac{1}{\sqrt{2}})^2 |h_d|^2\right).$$

We next use the condition in (Equation 3.41) to further upper bound the first expression of the sum-capacity upper bound in (Equation 3.40b) as

$$\log\left(1 + (|h_d| + |h_i| + |h_c|)^2)\right) \leq \log\left(1 + (2 + \frac{1}{\sqrt{2}})^2 |h_d|^2\right).$$

Finally, by taking the difference between upper and lower bounds we get (note that $R_3 \leq \log_2(2)$)

$$\begin{aligned} \text{gap} &\leq \log\left(1 + (2 + \frac{1}{\sqrt{2}})^2 |h_d|^2\right) + \log\left(1 + \frac{\left||h_d| - h_i\right|^2}{2}\right) \\ &- \log\left(1 + (1 - \frac{1}{\sqrt{2}})^2 |h_d|^2\right) - \log\left(1 + \left||h_d| - h_i\right|^2\right) \\ &+ 2\log(2) \leq \log\left(4 \frac{(2 + \frac{1}{\sqrt{2}})^2}{(1 - \frac{1}{\sqrt{2}})^2}\right) \approx 8.4 \text{ bits.} \end{aligned}$$

When $|h_{33}|^2 > |h_c|^2$, the outer bound in (Equation 3.40) suggests that the intended signal at the cognitive receiver is strong enough to support $R_3 > 0$. We focus on the regime identified by (Equation 3.42). User $i \in [1:2]$ splits its message into two parts: private information (to be kept below the noise floor at the non-intended primary receiver) and zero-forced information (to be zero-forced by the cognitive transmitter at the non-intended primary receiver). The cognitive transmitter pre-codes against the whole interference seen at its receiver by using Dirty Paper Coding (DPC) so that its receiver does not experience interference from the primary users.

Next we choose the power splits so as to match the upper bound. In order to zero-force / neutralize the interference of a primary user at the non-intended primary receiver we set $h_{21}\alpha_1 = h_{23}\alpha_3$ and $h_{12}\beta_2 = h_{13}\beta_3$. Moreover, since the interfering T_{jp}'s are considered as noise at the primary receivers we set $|h_{12}\gamma_2| \leq 1,\ |h_{13}\gamma_3| \leq 1,\ |h_{21}\gamma_1| \leq 1,\ |h_{23}\gamma_3| \leq 1$. For the symmetric channel, under the condition in (Equation 3.42) this can be accomplished by setting

$$|\gamma_1|^2 = 1 - |\alpha_1|^2 = |\gamma_2|^2 = 1 - |\beta_2|^2 = |\gamma_3|^2 = \frac{1}{1 + 2|h_c|^2}, \tag{3.46a}$$

$$\alpha_3 = \beta_3 = \frac{h_i}{h_c}\sqrt{1 - \frac{1}{1 + 2|h_c|^2}}. \tag{3.46b}$$

With (Equation 3.46) and under the condition in (Equation 3.42), the variance of the overall noise (i.e., actual noise plus the interfering signals treated as noise) is upper bounded as

$$1 + \frac{|h_c|^2 + |h_i|^2}{1 + 2|h_c|^2} \leq 1 + \frac{(1 + 1/2)|h_c|^2}{1 + 2|h_c|^2} \leq 1.75. \tag{3.47}$$

With (Equation 3.47), we see that the rate of the primary users satisfy

$$\begin{aligned} R_1 &= R_2 \\ &\geq \log\left(1 + \frac{(1 - \frac{1}{1+2|h_c|^2})\big||h_d| - h_i\big|^2 + \frac{|h_d|^2}{1+2|h_c|^2}}{1.75}\right) \\ &\geq \log\left(1 + \frac{(1 - \frac{1}{\sqrt{2}})^2}{1.75}|h_d|^2\right), \end{aligned} \tag{3.48}$$

while for the cognitive user, who DPCs against the whole interference due to the primary users, we have

$$R_3 = \log\left(1 + \frac{|h_{33}|^2}{2|h_c|^2 + 1}\right). \tag{3.49}$$

We next use the condition in (Equation 3.42) to further upper bound the first two expressions of the sum-capacity outer bound in (Equation 3.40b) as

$$\log\left(1 + (|h_d| + |h_i| + |h_c|)^2)\right) \leq \log\left(1 + (2 + \frac{1}{\sqrt{2}})^2|h_d|^2\right),$$
$$\log\left(1 + \left||h_d| - h_i\right|^2\right) \leq \log\left((1 + \frac{1}{\sqrt{2}})^2\frac{|h_d|^2}{2}\right)$$

Finally, by taking the difference between upper and lower bounds we get (note that the terms depending on $|h_{33}|^2$ match exactly)

$$\begin{aligned}
\mathsf{gap} &\leq \log\left(1 + (2 + \frac{1}{\sqrt{2}})^2|h_d|^2\right) + \log\left((1 + \frac{1}{\sqrt{2}})^2\frac{|h_d|^2}{2}\right) \\
&- 2\log\left(1 + \frac{(1 - \frac{1}{\sqrt{2}})^2}{1.75}|h_d|^2\right) + \log(2) \\
&\leq \log(2) + \log\left(\frac{(2 + \frac{1}{\sqrt{2}})^2}{\left(\frac{(1-\frac{1}{\sqrt{2}})^2}{1.75}\right)}\right) + \log\left(\frac{(1 + \frac{1}{\sqrt{2}})^2}{2\left(\frac{(1-\frac{1}{\sqrt{2}})^2}{1.75}\right)}\right) \\
&\approx 13 \text{ bits.}
\end{aligned}$$

□

CHAPTER 4

COVERAGE ANALYSIS FOR BASE STATION COOPERATION IN MILLIMETER WAVE CELLULAR NETWORKS

One of the fundamental goals for 5G is a radical increase in data rates [51]. It is anticipated that higher data rates will be achieved by extreme densification of base stations, massive multiple-input-multiple-output (MIMO), increased data rate and/or base station cooperation [51]. However, prime microwave wireless spectrum has become severely limited, with little unassigned bandwidth available for emerging wireless products and services. Therefore, to fulfill the need for increased bandwidth, millimeter wave (mmWave) spectrum between 30 and 300 GHz have been considered for future 5G wireless mobile networks. Until recently, mmWave frequency bands were presumed to be unreliable for cellular communication due to blockage, absorption, diffraction, and penetration, resulting in outages and unreliable cellular communications [52]. However, the advances in CMOS radio-frequency circuits, along with the very small wavelength of mmWave signals, allows for the packing of large antenna arrays at both the transmit and receive ends, thus providing highly directional beam forming gains and acceptable signal-to-noise ratio (SNR) [52], [53]. This directionality will also lead to reduced interference when compared to microwave networks [52]. It is thus anticipated that mmWave spectrum holds tremendous potential for increasing spectral efficiency in upcoming cellular systems [54].

To further address the demand for higher data rates, cooperation between macro, pico and femto base stations has been proposed to enable a uniform broadband user experience across the

network. The dynamic coordination across several base stations - known as coordinated multi-point (CoMP) - will limit the intercell interference thus increasing throughput and enhancing performance at cell borders [55].

We consider a stochastic geometry based model is considered as in [19–23, 56], to study coverage in CoMP heterogenous mmWave network. To do so we need to incorporate key factors specific to a mmWave channel model. These specific mmWave characteristics are: a realistic mmWave channel model, highly directional channel gains and sensitivity to blockages. This is accomplished as follows:

The base stations are assumed to have a uniform linear array and we account for the corresponding directionality accordingly. The receivers are equipped with a uniform linear array. Coverage probabilities are derived for the case of a single antenna receiver (typical user). We use concepts from [19–21] to incorporate blockage, interference and different fading distributions (Rayleigh and Nakagami) in our analysis. The joint distribution of the cooperating base stations to the typical user in the presence of blockage is also derived.

The downlink CoMP mmWave heterogenous network model, the beamsteering at the base stations and the decoding at the typical user are explained in Section 4.1. The coverage probability in the absence of blockage, and with Rayleigh fading is derived in Section 4.1.8. In Section 4.2, we consider the heterogenous CoMP with Rayleigh fading mmWave network with a blockage parameter at each tier, and the coverage probability is derived accordingly. In Section 4.3, we derive the coverage probability for the same network model with blockage but consider a different fading distribution on the direct links, more precisely, Nakagami fading is

used to model the fading distribution on the direct links of the cooperating base stations while keeping the Rayleigh fading assumption for the interfering channel gains. Each section has numerical examples of the corresponding coverage probability. Proofs may be found in the Appendices. Table II, Table III, Table IV, Table V summarize all the notations used throughout this chapter.

TABLE II. POISSON POINT PROCESS VARIABLES

Notation	Description
K	Total number of tiers
$\Phi_{mmw,k}$	Homogenous Poisson Point Process (PPP) indexed by $k \in [1:K]$
λ_k	Intensity of the PPP $\Phi_{mmw,k}$
P_k	Available power at each base station that belongs to tier $k \in [1:K]$
v	Points on 2D plane representing location of base stations
$\|v\|$	Distance from point v to the typical user located at the origin
α	Pathloss exponent assumed equal for all tiers
$\Theta_k = \{\frac{\|v\|^\alpha}{P_k}, v \in \Phi_{mmw,k}\}$	Normalized pathloss between each base station in $\Phi_{mmw,k}$ and the typical user
$\lambda_k(v)$	Intensity of Θ_k
$\Theta = \cup_{k=1}^{K}\Theta_k$	Process representing the union of non-homogenous PPP, elements are indexed in increasing order WLOG
$\lambda(v) = \sum_{k=1}^{K}\lambda_k(v)$	Intensity of Θ
$\gamma_i' = \frac{\|v_i\|^\alpha}{P_k}$	Normalized pathloss
$\gamma' = \{\gamma_1', \cdots, \gamma_n'\}$	Set of normalized pathloss of the cooperating base stations
$f_{\Gamma'}(\gamma')$	Joint distribution of γ'

TABLE III. GENERAL CHANNEL MODEL VARIABLES

Notation	Description
N_t, N_r	Number of antennas at each base station and at the receiver
H_ν	MIMO channel from base station at location ν to typical user
h_ν	Small scale fading
L_ν	Number of channel clusters
ϕ_ν^t	Path angle at the transmitter
ϕ_ν^r	Path angle at the receiver
$f(\nu)$	Function that returns the index to which a base station at ν belongs to
$\gamma_\nu = \frac{P_{f(\nu)}}{\|\nu\|^\alpha}$	Pathloss
$\mathbf{a_{t(r)}}(.)$	Uniform linear array vector representation at the transmitter (receiver)
$\Delta_{t(r)}$	Normalized transmit (receive) antenna separation
L_t	Normalized length of the transmit antenna array
$\mathbf{n}$	Noise vector of i.i.d $\mathcal{CN}(0, \sigma_n^2)$

TABLE IV. CHANNEL VARIABLES FROM COOPERATING BASE STATIONS

Notation	Description
$\mathcal{T}$	Set of cooperating base stations with cardinality $\|\mathcal{T}\| = n$
$\nu_i, \quad i \in [1 : \|\mathcal{T}\|]$	Points on the 2D plane corresponding to cooperating base stations location (sometimes indexed by j instead of i)
$\mathbf{H}_{\nu_i}$	MIMO channel from the cooperating base stations
$h_{\nu_i}, \phi_{\nu_i}^t, \phi_{\nu_i}^r, \gamma_{\nu_i}$	Channel parameters of the interfering links as defined in Table III
$\Omega_{\phi_{\nu_i}^r}$	Directional cosine given by $\cos(\phi_{\nu_i}^r)$
$\mathbf{X}_{\nu_i}$	Transmit signal from cooperating base stations

4.1 Coverage Probability with no Blockage

4.1.1 Network Model

Consider a K tier heterogenous network where each tier is an independent two-dimensional homogenous Poisson point process (PPP). We denote the base station location process of tier $k \in [1 : K]$ by $\Phi_{\text{mmw},k}$ with density λ_k. The mmWave base stations that belong to the same tier k transmit with the same power P_k for $k \in [1 : K]$. We consider performance metrics as

TABLE V. CHANNEL VARIABLES FROM INTERFERING BASE STATIONS

Notation	Description
$l_i, \quad i \in [1 : \lvert\mathcal{T}^c\rvert]$	Points on the 2D plane corresponding to interfering base stations locations
$\mathbf{H}_{l_i}$	MIMO channel from the interfering base stations
$h_{l_i}, \phi^t_{l_i}, \phi^r_{l_i}, \gamma_{l_i}$	Channel parameters of the interfering links as defined in Table III
$\Omega_{\phi^r_{l_i}}$	Directional cosine given by $\cos(\phi^r_{l_i})$
$\theta^t_{l_i}$	Angle used by interfering base station at position l_i to beamsteer to a user other than typical user
$\Omega_{\theta^t_{l_i}}$	Directional cosine given by $\cos(\theta^t_{l_i})$
$\mathbf{X}_{l_i}$	Transmit signal from interfering base stations

experiences by the typical user located at the origin, and denote the set of cooperating base stations, which jointly transmit to the typical user, by $\mathcal{T} \in \cup_{k=1}^{K} \Phi_{\text{mmw},k}$. We assume that $|\mathcal{T}| = n$, and that these n base-stations correspond to those with the strongest received power at the typical user receiver. In the rest of the section, we first describe the channel model and then derive the output signal at the typical user receiver.

4.1.2 Simplified Clustered Channel Model

A clustered channel model, [53], [57], is used to model the wireless channel between the base stations and the typical user. We assume all base stations have the same number of transmit antennas N_t, while the receiver has N_r receive antennas. The $N_r \times N_t$ channel matrix $\mathbf{H}_v$, between a base station located at $v \in \mathbb{R}^2$ and the typical user is the sum of L_v clusters and is expressed as

$$\mathbf{H}_v = \frac{\sqrt{N_t N_r}}{L_v} \sum_{l=1}^{L_v} \sqrt{\gamma_{v,l}} h_{v,l} \mathbf{a}_r(\phi^r_{v,l}) \; \mathbf{a}_t(\phi^t_{v,l})^*, \tag{4.1}$$

where

- $\gamma_{\nu,l} = \frac{P_{f(\nu)}}{\|\nu\|^{\alpha}}$ is the pathloss,
- $f(\nu)$ is a function that returns the index k of the tier to which the base station at location ν belongs to,
- α is the pathloss exponent,
- $\|\nu\|$ is the distance from the base station at location ν to the user at the origin,
- $h_{\nu,l}$ is the complex fading channel gain.
- The vectors $\mathbf{a}_{t(r)}(\phi_{\nu}^{t(r)})$ are the normalized uniform linear array (ULA) transmit and receive array response and are given by [58, Eq. (7.21), Eq. (7.25)]

$$\mathbf{a}_{t(r)}(\phi_{\nu}^{t(r)}) = \frac{1}{\sqrt{N_{t(r)}}}[1, e^{-jA}, e^{-j2A}, \cdots, e^{-j(N_{t(r)}-1)A}]^T \tag{4.2}$$

  where $A = 2\pi\Delta_{t(r)}\cos(\phi_{\nu}^{t(r)})$ and $\Delta_{t(r)}$ is the normalized transmit (receive) antenna separation (normalized to the unit of the carrier wavelength), at a path angle $\phi_{\nu}^{t(r)}$ of departure (arrival) from the base station ν.

In the following, for simplicity, we shall consider the case $L_{\nu} = 1$.

4.1.3 Received Signal at the Typical User

In this section we will further divide the points $\nu \in \mathbb{R}^2$ into a set of points ν_i and l_i to differentiate between the location of the cooperating and interfering base stations respectively. The $N_r \times N_t$ desired channel matrices are denoted by $\mathbf{H}_{\nu_i}$ for $i \in [1 : |\mathcal{T}| = n]$, where n is a non-negative constant, while the interfering channel matrices are denoted by $\mathbf{H}_{l_i}$ for $i \in [1 : |\mathcal{T}^c|]$.

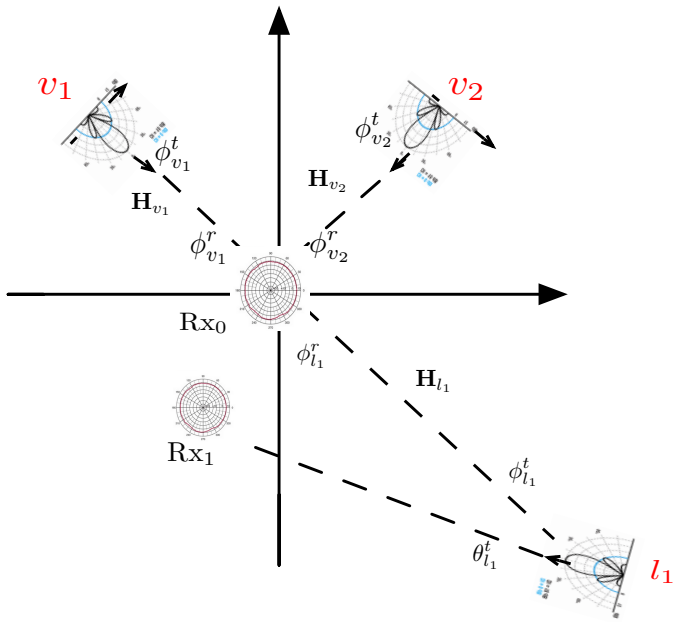


Figure 11. A typical user is served by two cooperating base stations at locations v_1 and v_2, while being interfered by base station at location l_1.

Figure 11 shows an example of a network model, where two base stations at locations ν_1 and ν_2, jointly transmit to the typical receiver located at the origin (indicated as Rx_0) in the presence of a single interfering base station at location l_1. The MIMO channel matrices between the cooperating base stations and the typical user are given by $\mathbf{H}_{\nu_1}$ and $\mathbf{H}_{\nu_2}$. The channel matrix between the interfering base station and the typical user is denoted by $\mathbf{H}_{l_1}$. The angles, $\phi^t_{\nu_i}$ and $\phi^r_{\nu_i}$, are the cluster's angle of departure and arrival respectively from the base station ν_i, $i \in [1:2]$, to the typical receiver. The angle $\phi^t_{l_1}$ is the angle of departure of the cluster from the interfering base station. The base station at l_1 uses a beamsteering angle θ_{l_1} to transmit data to some other user (not the typical user) indicated as Rx_1. The received signal at the typical user can be expressed as

$$\begin{aligned}\mathbf{y} &= \sum_{i=1}^{|\mathcal{T}|=n} \mathbf{H}_{\nu_i}\mathbf{X}_{\nu_i} + \sum_{i=1}^{|\mathcal{T}^c|} \mathbf{H}_{l_i}\mathbf{X}_{l_i} + \mathbf{n} \\ &= \sum_{i=1}^{|\mathcal{T}|=n} \sqrt{N_t N_r}\sqrt{\gamma_{\nu_i}} h_{\nu_i}\ \mathbf{a}_r(\phi^r_{\nu_i})\ \mathbf{a}_t(\phi^t_{\nu_i})^*\mathbf{X}_{\nu_i} + \sum_{i=1}^{|\mathcal{T}^c|} \sqrt{N_t N_r}\sqrt{\gamma_{l_i}} h_{l_i}\ \mathbf{a}_r(\phi^r_{l_i})\ \mathbf{a}_t(\phi^t_{l_i})^*\mathbf{X}_{l_i} + \mathbf{n}\end{aligned} \tag{4.3}$$

where the first sum in (Equation 4.3) is the desired signal from the mmWave cooperating base stations, while the second sum contains the signals from the interfering base stations.

The user associates with a set of cooperating base stations $\mathcal{T}$, that provide the strongest average received power as in [23]. Specifically,

$$\mathcal{T} = \arg \max_{\{\nu_1 \cdots \nu_n\} \subset \cup_{k=1}^{K} \Phi_{\mathrm{mmw},k}} \sum_{i=1}^{n} \frac{P_{f(\nu_i)}}{\|\nu_i\|^{\alpha}} \tag{4.4}$$

and $\mathcal{T}^c := \cup_{k=1}^{K} \Phi_{mmw,k} \setminus \mathcal{T}$.

The path angles, $\phi^t_{\nu_i}$, $i \in [1 : |\mathcal{T}|]$, and $\phi^t_{l_i}$, $i \in [1 : |\mathcal{T}^c|]$, represent the angle of departure of the desired and interfering paths respectively, while $\phi^r_{\nu_i}$, $i \in [1 : |\mathcal{T}|]$, and $\phi^r_{l_i}$, $i \in [1 : |\mathcal{T}^c|]$, represent the angle of arrival of the received path from the cooperating and interfering base stations respectively. The transmit signals, $\mathbf{X}_{\nu_i}$, $i \in [1 : |\mathcal{T}|]$ and $\mathbf{X}_{l_i}$, $i \in [1 : |\mathcal{T}^c|]$, represent the signal from the cooperating and interfering base stations within $\mathcal{T}$ and $\mathcal{T}^c$ respectively. $\mathbf{n}$ is the noise vector of i.i.d $\mathcal{CN}(0, \sigma^2)$ components.

4.1.4 Beamsteering

The base stations in $\mathcal{T}$ jointly send the same data to the receiver. Each base station beamsteers to the typical user, therefore the transmitted signal is

$$\mathbf{X}_{\nu_i} = \mathbf{a}_t(\phi^t_{\nu_i})s \tag{4.5}$$

for $i \in [1 : |\mathcal{T}| = n]$, where s is channel input symbol transmitted by the cooperating base stations to the typical receiver. The signals transmitted by the interfering base stations are

$$\mathbf{X}_{l_i} = \mathbf{a}_t(\theta^t_{l_i})s_{l_i} \tag{4.6}$$

for $i \in [1 : |\mathcal{T}^c|]$, where s_{l_i} is the channel input symbol transmitted by the interfering base stations, while the angle $\theta^t_{l_i}$ is the angle used by base station l_i to beamsteer to a user other than the typical user, and is different from $\phi^t_{l_i}$. We assume that s and s_{l_i} are independent zero mean and unit variance random variables.

Assumption 1: We assume that the cooperating base stations have perfectly beamsteered to the typical receiver: notice that the angles in (Equation 4.5), used by the base station to beamsteer, are equal to the clusters' angles of departure in the desired channel in (Equation 4.3).

Remark 1: The beamsteering used is compliant with analog beam forming and requires no digital base band processing [53].

4.1.5 Decoding

The receiver uses a single vector $\mathbf{w} \in \mathbb{C}^{N_r \times 1}$ to detect the scalar transmit symbol, that is, the processed received signal is given by

$$\hat{y} = \mathbf{w}^* \mathbf{y} \tag{4.7}$$

$$\mathbf{w} = \sum_{j=1}^{n} \mathbf{a}_r(\phi_{v_j}^r) \tag{4.8}$$

Remark 2: The choice of $\mathbf{w}$ in (Equation 4.8) is one choice of a decoder that can be implemented readily using phase shifters in the RF domain (analog processing), in fact if one wants to consider a near optimal performance, then the work in [53], which finds a hybrid MIMO receiver combining algorithm and minimizes the mean-square-error between the transmitted and received signals under a set of RF hardware constraints for the resulting point-to-point channel should be generalized to finding a suitable algorithm for the downlink cooperative channel.

Assumption 2: We assume perfect CSI of the path angles at the decoders since these angles vary slowly. However, we assume that the phases of the complex channel gains, $h_{v_i}, i \in [1 : |\mathcal{T}|]$, are not available at the terminals as they change very quickly on the order of a wavelength and

thus cannot be tracked. Because of the perfect CSI assumption, the performance here should be considered as an upper bound on the performance of the more realistic case with imperfect CSI.

4.1.6 Output Signal

The output signal at the typical user under the previously stated assumptions is given by

$$\hat{y} = \mathbf{w}^* \mathbf{y} = w * n + \left(\sum_{j=1}^{n} \mathbf{a}_r(\phi_{\nu_j}^r)^* \right) \times$$
$$\left(\sqrt{N_t N_r} \sum_{i=1}^{n} \sqrt{\gamma_{\nu_i}} h_{\nu_i} \mathbf{a}_r(\phi_{\nu_i}^r) \mathbf{a}_t(\phi_{\nu_i}^t)^* \mathbf{a}_t(\phi_{\nu_i}^t) s + \sqrt{N_t N_r} \sum_{i=1}^{|\mathcal{T}^c|} \sqrt{\gamma_{l_i}} h_{l_i} \mathbf{a}_r(\phi_{l_i}^r) \mathbf{a}_t(\phi_{l_i}^t)^* \mathbf{a}_t(\theta_{l_i}^t) s_{l_i} \right) \tag{4.9a}$$

$$= \sqrt{N_t N_r} \sum_{j=1}^{n} \sum_{i=1}^{n} \sqrt{\gamma_{\nu_i}} h_{\nu_i} G_r(\Omega_{\phi_{\nu_j}^r} - \Omega_{\phi_{\nu_i}^r}) G_t(\Omega_{\phi_{\nu_i}^t} - \Omega_{\phi_{\nu_i}^t}) s \tag{4.9b}$$

$$+ \sqrt{N_t N_r} \sum_{j=1}^{n} \sum_{i=1}^{|\mathcal{T}^c|} \sqrt{\gamma_{l_i}} h_{l_i} G_r(\Omega_{\phi_{\nu_j}^r} - \Omega_{\phi_{l_i}^r}) G_t(\Omega_{\phi_{l_i}^t} - \Omega_{\theta_{l_i}^t}) s_{l_i} + z \tag{4.9c}$$

where $z := \mathbf{w}^* n \sim \mathcal{C}N(0, \sigma_n^2)$, and where we introduced the antenna-array-gain functions

$$G_x(y) := e^{j\pi\Delta_x(N_x-1)y} \frac{\sin(\pi\Delta_x N_x y)}{N_x \sin(\pi\Delta_x y)} : |G_x(y)| \leq 1, \quad x \in \{t, r\}, \tag{4.10}$$

$$\mathbf{a}_x(\phi_1)^* \mathbf{a}_x(\phi_2) = G_x(\Omega_{\phi_1} - \Omega_{\phi_2}), \quad x \in \{t, r\}, \tag{4.11}$$

with $\Omega_\phi := \cos(\phi)$ and Δ_x being the antenna separation.

4.1.7 SINR Expression

Based on (Equation 4.9c), the instantaneous SINR is then given by

$$\text{SINR} = \frac{\left| \sum_{i=1}^{n} \sqrt{\gamma_{v_i}} h_{v_i} C_{v_i} \right|^2}{\frac{\sigma_n^2}{N_t N_r} + \sum_{i=1}^{|\mathcal{T}^c|} \gamma_{l_i} |h_{l_i}|^2 |D_{l_i}|^2 \left| G_t(\Omega_{\phi_{l_i}^t} - \Omega_{\theta_{l_i}^t}) \right|^2}, \tag{4.12a}$$

$$C_{v_i} := \sum_{j=1}^{n} G_r(\Omega_{\phi_{v_j}^r} - \Omega_{\phi_{v_i}^r}), \tag{4.12b}$$

$$D_{l_i} := \sum_{j=1}^{n} G_r(\Omega_{\phi_{v_j}^r} - \Omega_{\phi_{l_i}^r}). \tag{4.12c}$$

Assuming a single antenna receiver with $N_r = 1$ ($C_{v_i} = D_{l_i} = n$ and $\sigma_n^2 = n^2\sigma^2$), the SINR in (Equation 4.12a) simplifies to

$$\text{SINR} = \frac{\left| \sum_{i=1}^{n} \sqrt{\gamma_{v_i}} h_{v_i} \right|^2}{\frac{\sigma^2}{N_t} + \sum_{i=1}^{|\mathcal{T}^c|} \gamma_{l_i} |h_{l_i}|^2 \left| G_t(\Omega_{\phi_{l_i}^t} - \Omega_{\theta_{l_i}^t}) \right|^2}, \tag{4.13}$$

In the following section, we will derive the coverage probability for the typical user with SINR as in (Equation 4.13) under the assumption that all angles are independent and uniformly distributed between $[-\pi, +\pi]$. We will first assume that the receiver is present in a rich scattering environment (Rayleigh fading assumption), coverage probability is given in Th. 4.1.1. The case where each tier experiences blockage is then considered and the coverage probability is derived accordingly and is given in Th. 4.2.1. Keeping the assumption of blockage, the Nakagami fading distribution is then used to model the improved fading distribution on the direct cooperating links due to beamsteering and two upper bounds on the coverage probability are

then derived and are given in Th. 4.3.1 and Th. 4.3.2, and shown to be numerically tight when the network experiences high blockage. Future work includes deriving the coverage probability for all the different cases described above using (Equation 4.12a), i.e, multiple antennas at the receivers.

4.1.8 Performance Analysis

Theorem 4.1.1. *sec:performance analysis no blockage ULA analysis The coverage probability for the typical user, with a single antenna, in a downlink mmWave heterogenous network with* K *tiers, with base stations having ULA with* N_t *antennas, of which* n *jointly transmit to it is, given by*

$$\mathbb{P}(\text{SINR} > T) = \int\limits_{0<\gamma'_1<\cdots<\gamma'_n<+\infty} \mathcal{L}_I\left(\frac{T}{\sum_{i\leq n}\gamma_i'^{-1}}\right)\mathcal{L}_N\left(\frac{T}{\sum_{i\leq n}\gamma_i'^{-1}}\right) f_{\Gamma'}(\gamma')\mathrm{d}\gamma' \tag{4.14}$$

where $\gamma'_i = \frac{\|v_i\|^\alpha}{P_{f(v_i)}}$ *for* $i \in [1:n]$*, and the Laplace transform of the interference and the noise are given by*

$$\mathcal{L}_I(s) = \exp\left(-\int_{\gamma'_n}^{\infty}\left[1-\int_{-2}^{+2}\left(\frac{1}{1+s|G_t(\varepsilon)|^2 v^{-1}}\right) f_\Upsilon(\varepsilon)\ \mathrm{d}\varepsilon\right]\lambda(v)\,\mathrm{d}v\right), \tag{4.15}$$

$$\mathcal{L}_N(s) = e^{-s\sigma^2/N_t}, \tag{4.16}$$

where the antenna array gain $G_t(\varepsilon)$ *is given by* (Equation 4.10) *and the probability density function of* $\Upsilon_i = \Omega_{\phi^t_{l_i}} - \Omega_{\theta^t_{l_i}}$ *is*

$$f_\Upsilon(\varepsilon) = \int_{\max\{-1,-1-\varepsilon\}}^{\min\{1,1-\varepsilon\}} \left(\frac{1}{\pi^2 \sqrt{1-(\varepsilon+y)^2}} \frac{1}{\sqrt{1-y^2}} \right) dy, \tag{4.17}$$

and $f_{\Gamma'}(\gamma')$ *is the joint distribution of* $\gamma' = [\gamma'_1, \cdots, \gamma'_n]$ *and is given by*

$$f_{\Gamma'}(\gamma') = \prod_{i=1}^{n} \lambda(\gamma'_i) e^{-\Lambda(\gamma'_n)} \tag{4.18}$$

while the intensity and intensity measure are given by

$$\lambda(\nu) = \sum_{k=1}^{K} \lambda_k \frac{2\pi}{\alpha} P_k^{\frac{2}{\alpha}} \nu^{\frac{2}{\alpha}-1}, \tag{4.19}$$

$$\Lambda(\gamma'_n) = \sum_{k=1}^{K} \pi \lambda_k P_k {\gamma'_n}^{\frac{2}{\alpha}} \tag{4.20}$$

Proof. Please refer to Appendix G for the proof. □

4.1.9 Numerical Results

In this section we numerically evaluate Th. 4.1.1. We compute the coverage probability for the typical user in a mmWave CoMP heterogenous network and compare it to the case with no base station cooperation. We consider a two tier network, $K = 2$ with parameters given in Table VI. The noise is given by $\sigma^2(\text{dBm}) = -174 + 10\log_{10}(\text{BW}) + \text{NF (dB)}$. In Figure 12 the coverage probability in (Equation 4.14) for $n = 2$ and $n = 1$ is plotted. In the absence of

blockage, the numerical results show that the coverage probability with cooperation for the case of $N_t = 16$ antennas exceeds that for the case with no cooperation by almost 11 percent. A smaller increase is recognized for the case when $N_t = 64$ antennas and for the same threshold. As expected an increase in the number of antennas at the base stations increases the coverage probability. For example, for the same threshold $T = 0$ dB, the coverage probability with cooperation and with $N_t = 32$ is approximately 0.92 while that for $N_t = 16$ is 0.8. Thus numerical results suggest when mmWave networks do not experience blockage, a substantial increase in coverage probability is still attained with cooperation.

TABLE VI. TIER 1 and TIER 2 PARAMETERS

Parameter	Value
Intensity	$\lambda_1 = (150^2\pi)^{-1}$, $\lambda_2 = (50^2\pi)^{-1}$
Power	$P_1 = 1$ W (30 dBm) and $P_2 = 0.25$ W
Path Loss	$\alpha = 3$
Antennas	$N_t = 8, 16, 32, 64$
Noise Figure (NF)	10 dB
Bandwidth (BW)	1 GHz

4.2 Coverage probability with Blockage

4.2.1 Network model

In this section we again consider a K-tier heterogenous network where each tier is an independent two-dimensional homogenous Poisson point process (PPP). The base station location process of each tier is denoted by $\Phi_{\text{mmw},k}$ with density λ_k for $k \in [1 : K]$. Each tier is charac-

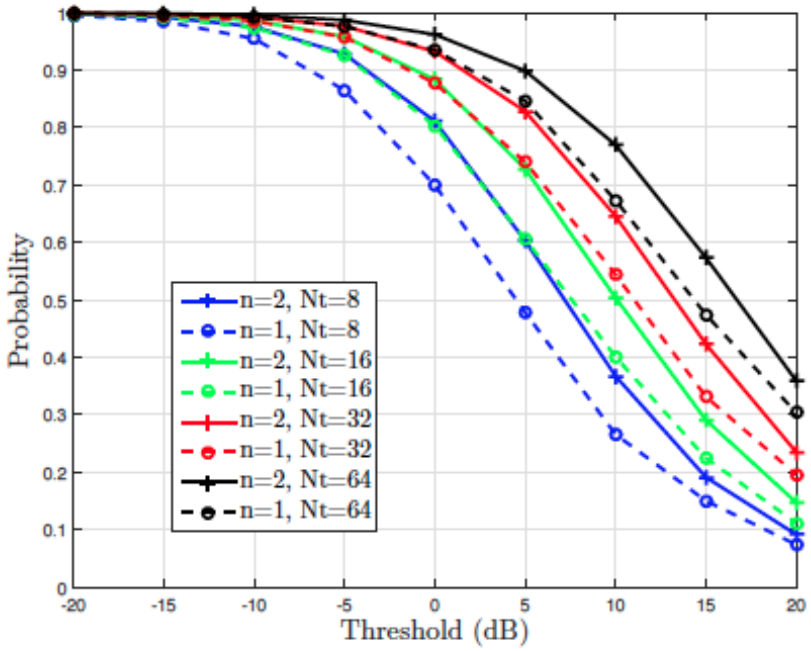


Figure 12. Coverage probability in (Equation 4.14) for a two-tier network with parameters in Table VI with two cooperating base stations ($n = 2$) and without base station cooperation ($n = 1$) and for different number of antennas.

terized by a non-negative blockage constant β_k for $k \in [1 : K]$ (determined by the density and average size of objects within the tier) as used in [20], [21]. Consequently, after defining the parameter β_k for $k \in [1 : K]$, we have that the probability of the communication link being a LOS link (no blockage on the link) within tier k is $\mathbb{P}(\text{LOS}_k) = e^{-\beta_k r}$, where r represents the length of the communication link, while the probability of a link being NLOS is $\mathbb{P}(\text{NLOS}_k) = 1 - \mathbb{P}(\text{LOS}_k)$. The LOS and NLOS links will have different pathloss exponents, α_1 and α_2, respectively and same for all $k \in [1 : K]$. The mmWave base stations that belong to the same tier k transmit

with the same power P_k for $k \in [1:K]$. With the assumption of blockage the following have to be re-derived (when compared to those in Th. 4.1.1).

- The Laplace transform of the interference has to be re-derived given the fact that the density of each tier is a function of two pathloss exponents and of the blockage parameter.
- The joint distribution of the base stations with the strongest received power at the typical user has to be re-derived for the same reason mentioned above.

In the following section we give the coverage probability for a multi-tier mmWave network with blockage when base stations cooperate in sending the same message to the typical receiver. The numerical results suggest that joint transmission increases the probability of coverage due to the increase in the number of signal path a receiver experiences.

4.2.2 Performance Analysis

Theorem 4.2.1. *The coverage probability for the typical user, with a single antenna, in a downlink mmWave heterogenous network with* K *tiers, and where each tier has a blockage parameter* β_k*, with* n *base stations having ULA with* N_t *antennas, jointly transmitting to it is given by* (Equation 4.14)*,* (Equation 4.15)*,* (Equation 4.16) *but where now the intensity* $\lambda(\nu)$ *in* (Equation 4.15) *is given by*

$$\lambda(\nu) = \sum_{k=1}^{K} A_k \nu^{\delta_1 - 1} e^{-a_k \nu^{\frac{\delta_1}{2}}} + B_k \nu^{\delta_2 - 1}(1 - e^{-b_k \nu^{\frac{\delta_2}{2}}}) \tag{4.21}$$

where $A_k = \pi\lambda_k\delta_1 P_k^{\delta_1}, a_k = \beta_k P_k^{\frac{\delta_1}{2}}, b_k = \beta_k P_k^{\frac{\delta_2}{2}}, B_k = \pi\lambda_k\delta_2 P_k^{\delta_2}, \delta_1 = 2/\alpha_1, \delta_2 = 2/\alpha_2$ *and the distribution of the distance of* n *closest base stations is given by* (Equation 4.18) *but where*

$$\Lambda(\gamma_n') = \sum_{k=1}^{K} \frac{2\pi\lambda_k}{\beta_k^2}\left(1 - e^{-\beta_k(\gamma_n' P_k)^{\delta_1/2}}(1+\beta_k(\gamma_n' P_k)^{\frac{\delta_1}{2}})\right) + \pi\lambda_k(\gamma_n' P_k)^{\delta_2} - \frac{2\pi\lambda_k}{\beta_k^2}\left(1 - e^{-\beta_k(\gamma_n' P_k)^{\frac{\delta_2}{2}}}(1+\beta_k(\gamma_n' P_k)^{\frac{\delta_2}{2}})\right) \tag{4.22}$$

Proof. Please refer to Appendix H for detailed proof. □

4.2.3 Numerical Results

In this section we numerically evaluate Th. 4.2.1. We compute the coverage probability for the typical user in a mmWave CoMP heterogenous network in the presence of blockage. In this example, we compare the coverage probability in Th. 4.2.1 for a one tier mmWave network with parameters given in Table VII with and without joint transmission. The examples provided illustrate scenarios when cooperation is beneficial (in terms of increasing the coverage probability) and examples when the increase is not substantial. Numerical results suggest that the former is in fact the case when the mmWave network is dense (captured by the tier radius and consequently its intensity) - a feature expected in millimeter wave networks [51], [20]. This can be interpreted as follows, with extreme densification, the number of LOS interfering base stations increases and thus interference increases, a remark also noted in [20]. Therefore, the increase in coverage with base station cooperation is substantial. In Figure 13 we consider the example with parameters given in Table VII. A tier of radius 80 meters is considered with

blockage parameters $\beta = 0.003, 0.006$. The increase in coverage probability for both cases is approximately an increase of 0.11 in probability for a threshold $T = 5, 10, 15, 20$ dB. Moreover, it is interesting to note that a tier blockage parameter $\beta = 0.006$ would result in higher coverage probability than that of $\beta = 0.003$. This can be interpreted as follows: an increase in the blockage parameter increases the probability of blockage of the interfering LOS base stations, resulting in higher coverage probabilities.

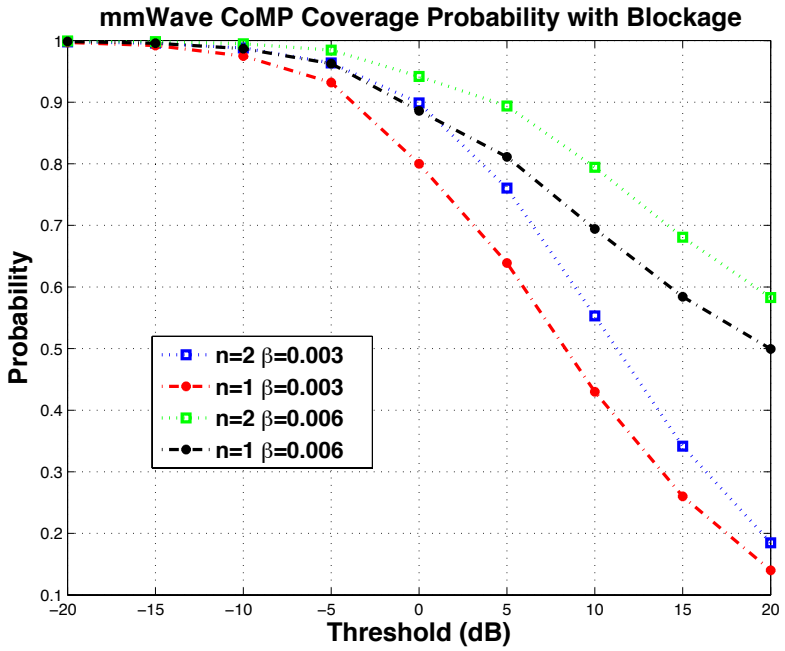


Figure 13. Coverage probability in Th. 4.2.1 for a one-tier network with parameters in Table VII with two cooperating base stations ($n = 2$) and without base station cooperation ($n = 1$) and for different blockage parameters.

A tier with a larger radius of 250 meters and with tier parameters given in Table VIII and with a blockage parameter $\beta = 0.025$ (high probability of blockage) is plotted in Figure 14. The coverage probability for this one tier CoMP mmWave network with $n = 2$ and with base station power available $P = 1$W is compared with the following cases, Case 1) a one tier mmWave network with no base station cooperation ($n = 1$), with a base station transmit power $P = 1$ W, Case 2) a one tier mmWave network with no base station cooperation $n = 1$, but with base station transmit power equal to the sum of transmit power if two base stations were to cooperate $P = 2$ W. The increase in coverage probability due to cooperation in this case is minimal and is approximately 0.05 for threshold $T = -10, -5$ dB. This can be interpreted as follows: 1) a tier with dense blockages will also block interfering signals and 2) when the density of base stations is not too dense, the n strongest base stations are not *too strong* to cause a substantial increase due to the fact that distance at which these cooperating base stations are located increases too (thus received power decreases). As seen in Figure 14, increasing the power at the base station (but with no cooperation) provides higher coverage probability than the case with base station cooperation.

4.3 Coverage Probability with Nakagami fading and blockage

4.3.1 Network Model

In this section we consider the same network model as in Section 4.2.1 but choose a different fading distribution on the channel gains from the strongest cooperating base stations, similar to [19]. In particular we consider Nakagami fading with parameter m, while keeping the same assumption of Rayleigh fading for the interfering channel gains.

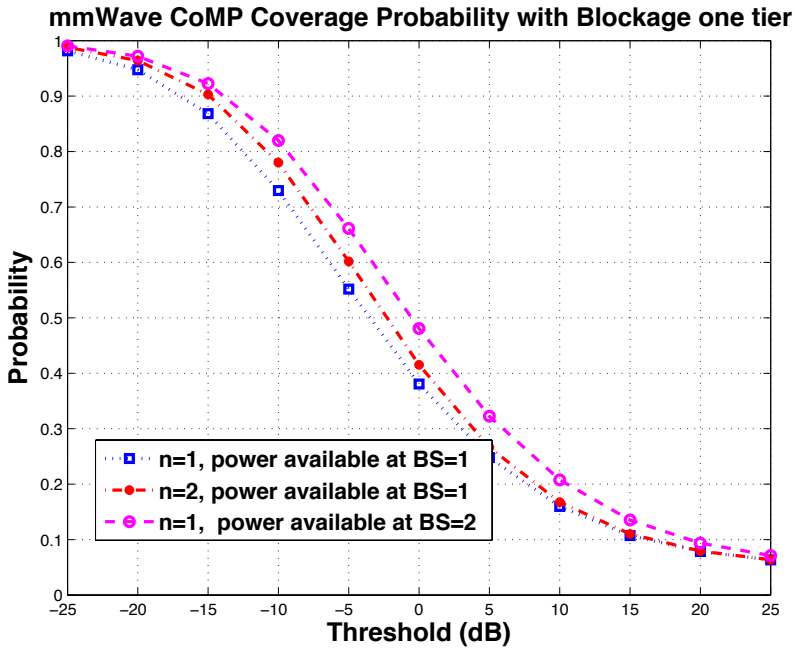


Figure 14. Coverage probability in Th. 4.2.1 for a one-tier network with parameters in Table VIII with two cooperating base stations ($n = 2$) and without base station cooperation ($n = 1$).

Using the coverage probability expression for a general fading distribution [59, Eq. 2.11], we are then able to derive an upper bound on the coverage probability for this network. We then consider another upper bound by evaluating the network in the absence of interference. The coverage probability is then defined as the probability that the signal-to-noise-ratio (SNR) is greater than a certain threshold. Using complex analysis methods of integration we give a closed form expression of the integral which reduces the time for numerical execution.

TABLE VII. TIER PARAMETERS WITH BLOCKAGE FOR NUMERICAL EXAMPLE 1

Parameter	Value
Intensity	$\lambda = (80^2\pi)^{-1}$
Power	$P_1 = 1$ W
Path Loss	$\alpha_1 = 2, \alpha_2 = 4$
Antennas	$N_t = 16$
Noise Figure	5 dB
Blockage	$\beta = 0.006, 0.003$
Bandwidth	1 GHz

TABLE VIII. TIER PARAMETERS WITH BLOCKAGE FOR NUMERICAL EXAMPLE 2

Parameter	Value
Intensity	$\lambda = (250^2\pi)^{-1}$
Power	$P_1 = 1$ W
Path Loss	$\alpha_1 = 2, \alpha_2 = 4$
Antennas	$N_t = 16$
Noise Figure	5 dB
Blockage	$\beta = 0.025$
Bandwidth	1 GHz

4.3.2 Performance Analysis

Theorem 4.3.1. *An upper bound on the coverage probability for the typical user, in a downlink mmWave heterogenous network with blockage with* K *tiers, where each tier has a blockage*

parameter β_k*, and with* n *base stations having ULA with* N_t *antennas jointly transmitting to it, with the assumption of Nakagami fading on the cooperating channel gains is given by*

$$\mathbb{P}(\text{SINR} > T) = \int\limits_{0<\gamma_1'<\cdots<\gamma_n'<+\infty} f_{\Gamma'}(\gamma') \int_{-\infty}^{\infty} \mathcal{L}_I(2j\pi T's)\mathcal{L}_N(2j\pi T's)\frac{\mathcal{L}_S^{UP}(-2j\pi s)-1}{2j\pi s} ds \; d\gamma' \tag{4.23}$$

where $\gamma_i' = \frac{\|v_i\|^\alpha}{P_{f(v_i)}}$ *for* $i \in [1:n]$ *and* $T' = \frac{T}{\sum_{i\leq n}\gamma_i'^{-1}}$ *while the joint distribution of* $\gamma' = [\gamma_1', \cdots, \gamma_n']$ *is given by* (Equation 4.18) *and where*

$$\lambda(\gamma_i') = \sum_{k=1}^{K} A_k \gamma_i'^{\delta_1 - 1} e^{-a_k \gamma_i'^{\frac{\delta_1}{2}}} + B_k \gamma_i'^{\delta_2 - 1}(1 - e^{-b_k \gamma_i'^{\frac{\delta_2}{2}}}) \tag{4.24}$$

and

$$\begin{aligned}\Lambda(\gamma_n') = \sum_{k=1}^{K} \frac{2\pi\lambda_k}{\beta_k^2}\left(1 - e^{-\beta_k(\gamma_n' P_k)^{\delta_1/2}}(1+\beta_k(\gamma_n' P_k)^{\frac{\delta_1}{2}})\right) \\ + \pi\lambda_k(\gamma_n' P_k)^{\delta_2} - \frac{2\pi\lambda_k}{\beta_k^2}\left(1 - e^{-\beta_k(\gamma_n' P_k)^{\frac{\delta_2}{2}}}(1+\beta_k(\gamma_n' P_k)^{\frac{\delta_2}{2}})\right)\end{aligned} \tag{4.25}$$

where $A_k = \pi\lambda_k\delta_1 P_k^{\delta_1}$, $a_k = \beta_k P_k^{\frac{\delta_1}{2}}$, $b_k = \beta_k P_k^{\frac{\delta_2}{2}}$, $B_k = \pi\lambda_k\delta_2 P_k^{\delta_2}$, $\delta_1 = 2/\alpha_1$, *and* $\delta_2 = 2/\alpha_2$*, and where the Laplace transform of* I *(assuming Rayleigh fading on the interfering links) is given by*

$$\mathcal{L}_I(s) = \exp\left(-\int_{\gamma_n'}^{\infty}\left[1 - \int_{-2}^{+2}\left(\frac{1}{1+s|G_t(\varepsilon)|^2\nu^{-1}}\right) f_\Upsilon(\varepsilon)\; d\varepsilon\right]\lambda(\nu)\; d\nu\right)$$

where

$$\lambda(\nu) = \sum_{k=1}^{K} A_k \nu^{\delta_1 - 1} e^{-a_k \nu^{\frac{\delta_1}{2}}} + B_k \nu^{\delta_2 - 1}(1 - e^{-b_k \nu^{\frac{\delta_2}{2}}}) \tag{4.26}$$

and where

$$\mathcal{L}_S^{UP}(s) = \frac{1}{(1 + s/m)^{nm}} \tag{4.27}$$

$$\mathcal{L}_N(s) = e^{-s\sigma^2/N_t}. \tag{4.28}$$

Proof. Please refer to Appendix I for detailed proof. □

Remark 3: A lower bound on the coverage probability would be given by that in Th. 4.2.1 with the Rayleigh fading assumption on the direct links.

Remark 4: For the case of no base station cooperation $n = 1$, the coverage probability in (Equation 4.23) above is exact and is not an upper bound.

4.3.3 Performance Analysis in the Absence of Interference

Theorem 4.3.2. *An upper bound on the coverage probability in the absence of interference for the typical user, in a downlink mmWave heterogenous network with blockage with* K *tiers, where each tier has a blockage parameter* β_k*, and with* n *base stations having ULA with* N_t *antennas*

jointly transmitting to it, with the assumption of Nakagami fading on the cooperating channel gains is given by

$$\mathbb{P}(\text{SNR} > T) = \int\limits_{0<\gamma'_1<\cdots<\gamma'_n<+\infty} f_{\Gamma'}(\gamma') \left(\frac{g^{(nm-1)}(z^*)}{(nm-1)!} \right) d\gamma'. \tag{4.29}$$

where $z^* = \frac{m}{2\pi j}$, $\gamma'_i = \frac{\|v_i\|^\alpha}{P_{f(v_i)}}$ *for* $i \in [1:n]$ *and* $T' = \frac{T}{\sum_{i\leq n} \gamma'^{-1}_i}$ *while the joint distribution of* $\gamma' = [\gamma'_1, \cdots, \gamma'_n]$ *is given by* (Equation 4.18) *and where*

$$g(z) = (-1)^{nm} \frac{1 - (1 - \frac{2\pi j z}{m})^{nm}}{(\frac{2\pi j}{m})^{nm}(2\pi j z)} e^{-2\pi j z \frac{T'\sigma^2}{N_t}} \tag{4.30}$$

and where $g^{nm-1}(z)$ *is the* $(nm-1)$ *derivative of the function* $g(z)$.

Proof. Please refer to Appendix I for the proof. □

A similar remark to that made in Remark 4 can be made for the coverage probability (in absence of interference) in (Equation 4.29), i.e., the probability of the SNR being greater than a certain threshold is exact for the case of no base station cooperation.

4.3.4 Numerical Results

We consider a one tier network, $K = 1$, with two cooperating base stations $n = 2$ and with tier parameters given in Table IX. The coverage probability in (Equation 4.23) for the case of $m = 1, 3$ (corresponding to Rayleigh, Nakagami) with and without base station cooperation are plotted. The purpose of this numerical example is to show that for tiers with high probability of blockage, in this case taken to be $\beta = 0.025$, evaluation of the coverage probability of the network

TABLE IX. TIER PARAMETERS WITH NAKAGAMI FADING

Parameter	Value
Intensity	$\lambda_1 = (200^2\pi)^{-1}$
Power	$P_1 = 1$ W (30 dBm)
Path Loss	$\alpha_1 = 2$ and $\alpha_2 = 4$
Antennas	$N_t = 16$
Noise Figure	5 dB
Bandwidth	1 GHz
Blockage	0.025

with and with the absence of interference yields almost exact numerical results. Therefore, we fix $n = 2$ and we plot the coverage probability in (Equation 4.23) and (Equation 4.29) for Case 1) $m = 3$ and $n = 2$. While we use Th. 4.2.1 to plot the coverage probability for the Rayleigh fading Case 2) $m = 1$ and $n = 2$ (exact lower bound). The two curves shown in Figure 15 for each of the cases corresponding to the Rayleigh and Nakagami fading almost exactly overlap.

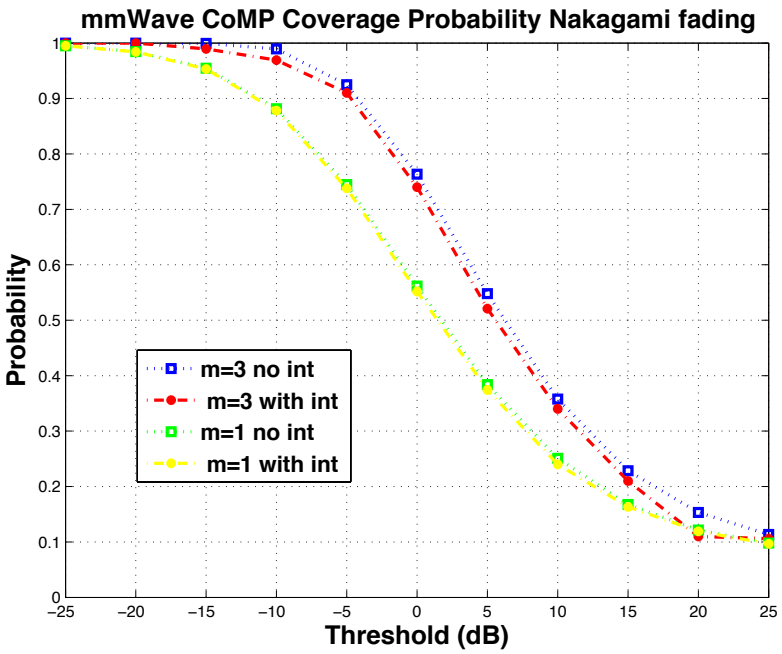


Figure 15. Upperbounds on coverage probability in Th. 4.3.1 and Th. 4.3.2 for a one tier network with parameters in Table IX and for $\mathfrak{n} = 2$. The lower bounds are plotted using from Th. 4.2.1

.

CHAPTER 5

CONCLUSIONS

We have studied the EGCIFC and developed the sum-capacity of a fading cognitive interference channel. A separable scheme (i.e. power allocation depends only on the current fading state and must not be done across states) was shown to be optimal. This is in contrast to the classical interference where considering a separable scheme is proved to be sub-optimal. It is known that encoding separately is optimal for a one-sided fading Gaussian interference channel (Z-IC) in which one of the receivers experiences no interference. One can think of the EGCIFC as a (Z-IC) since the cognitive transmitter is capable of precoding against the interference caused by the primary transmitter given that it has non-causal knowledge of its message. In developing the ergodic sum-capacity we have derived a new outer bound, provided the optimal separable achievability scheme along with the optimal power allocation policy, and developed a result of independent interest: the ergodic capacity of the MISO channel with per antenna power constraints.

As for the K-user cognitive interference channel with cumulative message sharing; a computable, general outer bound valid for any number of users and any memoryless channel is obtained. For the linear deterministic approximation of the Gaussian channel at high SNR we obtained the sum-capacity for all channel gains in the case of three users, and the symmetric sum-capacity for any K. For the Gaussian channel, we provided a unified achievability scheme which achieves the sum-capacity to within a constant additive and multiplicative gap. In the

linear deterministic channel, the sum-capacity was achieved by a scheme which only required cognition at one single user.

We have also considered the problem of base station cooperation in mmWave heterogenous networks. Using stochastic geometry, coverage probabilities were derived at the typical user, accounting for directionality at the base stations, blockage, interference and different fading distributions (Rayleigh and Nakagami). Numerical results suggest that coverage with cooperation rival that with no cooperation especially in dense mmWave networks. Future work includes deriving the coverage probability at a multi-antenna typical receiver, accounting for possible errors due to beam steering at the base stations and providing a comparison between achievable rates with and without cooperation.

APPENDICES

Appendix A

COPYRIGHT PERMISSION

This Appendix includes the copyright permission granted from the IEEE to use published work in thesis. The following statement has been copied from the CopyRight Clearance Center (Rightslink)

The IEEE does not require individuals working on a thesis to obtain a formal reuse license, however, you may print out this statement to be used as a permission grant:

1) The following IEEE copyright/ credit notice should be placed prominently in the references: [year of original publication] IEEE. Reprinted, with permission, from [author names, paper title, IEEE publication title, and month/year of publication].

2) Only the accepted version of an IEEE copyrighted paper can be used when posting the paper or your thesis on-line.

3) In placing the thesis on the author's university website, please display the following message in a prominent place on the website: In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of [university/educational entity's name goes here]'s products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution

Appendix A (Continued)

Thesis / Dissertation Reuse

The IEEE does not require individuals working on a thesis to obtain a formal reuse license, however, you may print out this statement to be used as a permission grant:

Requirements to be followed when using any portion (e.g., figure, graph, table, or textual material) of an IEEE copyrighted paper in a thesis:

1) In the case of textual material (e.g., using short quotes or referring to the work within these papers) users must give full credit to the original source (author, paper, publication) followed by the IEEE copyright line © 2011 IEEE.
2) In the case of illustrations or tabular material, we require that the copyright line © [Year of original publication] IEEE appear prominently with each reprinted figure and/or table.
3) If a substantial portion of the original paper is to be used, and if you are not the senior author, also obtain the senior author's approval.

Requirements to be followed when using an entire IEEE copyrighted paper in a thesis:

1) The following IEEE copyright/ credit notice should be placed prominently in the references: © [year of original publication] IEEE. Reprinted, with permission, from [author names, paper title, IEEE publication title, and month/year of publication]
2) Only the accepted version of an IEEE copyrighted paper can be used when posting the paper or your thesis on-line.
3) In placing the thesis on the author's university website, please display the following message in a prominent place on the website: In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of [university/educational entity's name goes here]'s products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to http://www.ieee.org/publications_standards/publications/rights/rights_link.html to learn how to obtain a License from RightsLink.

Appendix B

PROOF OF THEOREM 2.3.2

The MISO capacity with PerPC and arbitrary number of transmit antennas $i \in [1 : n]$ is the solution of the following optimization problem:

$$\max_{P_i(\mathbf{h})\geq 0:\mathbb{E}[P_i(\mathbf{h})]\leq \bar{P}_i, i\in[1:n]} \mathbb{E}\left[\log\left(1+\Big(\sum_i |h_i|\sqrt{P_i(\mathbf{h})}\Big)^2\right)\right]. \tag{B.1}$$

By Lagrange duality for $\lambda_i \geq 0$,

$$\mathcal{L} = \log\left(1+\Big(\sum_i |h_i|\sqrt{P_i(\mathbf{h})}\Big)^2\right) - \left(\sum_i \lambda_i P_i(\mathbf{h})\right) \tag{B.2}$$

$$\frac{\partial \mathcal{L}}{\partial P_i(\mathbf{h})} = \frac{\theta}{1+\theta^2}\frac{|h_i|}{\sqrt{P_i}} - \lambda_i = 0,\ \theta := \sum_i |h_i|\sqrt{P_i(\mathbf{h})}, \quad \forall i \tag{B.3}$$

$$\iff \sqrt{P_i(\mathbf{h})} = \frac{\theta}{1+\theta^2}\frac{|h_i|}{\lambda_i}, \quad \forall i \tag{B.4}$$

$$\theta = \sum_i |h_i|\sqrt{P_i(\mathbf{h})} = \frac{\theta}{1+\theta^2}\sum_i \frac{|h_i|^2}{\lambda_i} \tag{B.5}$$

$$\iff 1+\theta^2 = \sum_i \frac{|h_i|^2}{\lambda_i} \geq 1 \tag{B.6}$$

$$\iff \sqrt{P_j(\mathbf{h})} = \frac{\sqrt{\left(\sum_i \frac{|h_i|^2}{\lambda_i} - 1\right)^+}}{\sum_i \frac{|h_i|^2}{\lambda_i}}\frac{|h_j|}{\lambda_j} \tag{B.7}$$

$$\iff \lambda_j P_j(\mathbf{h}) = \left(1 - \frac{1}{\sum_i \frac{|h_i|^2}{\lambda_i}}\right)^+ \frac{\frac{|h_j|^2}{\lambda_j}}{\sum_i \frac{|h_i|^2}{\lambda_i}}. \tag{B.8}$$

Appendix B (Continued)

The rate becomes

$$\mathbb{E}\left[\log\left(1+\left[\sum_i \frac{|h_i|^2}{\lambda_i}-1\right]^+\right)\right] = \mathbb{E}\left[\log^+\left(\sum_i \frac{|h_i|^2}{\lambda_i}\right)\right] \tag{B.9}$$

and the Lagrange multipliers solve the non-linear system of equations

$$\lambda_j\overline{P_j} = \mathbb{E}\left[\left[1-\frac{1}{\sum_i \frac{|h_i|^2}{\lambda_i}}\right]^+ \frac{1}{\sum_i \frac{|h_i|^2}{\lambda_i}}\,\frac{|h_j|^2}{\lambda_j}\right]. \tag{B.10}$$

Appendix C

PROOF OF THEOREM 2.3.4

The sum-capacity of the EGCIFC in (Equation 2.4) can be re-written as the following optimization problem:

$$\max_{P_i(\mathbf{h})\geq 0:\mathbb{E}[P_i(\mathbf{h})]\leq \bar{P}_i, |\rho(\mathbf{h})|\leq 1, i\in[1:2]} \{\ \mathbb{E}\Bigg[\log\left(1+|h_{21}|^2P_1(\mathbf{h})+|h_{22}|^2P_2(\mathbf{h})+2|\rho(\mathbf{h})|\sqrt{|h_{21}|^2P_1(\mathbf{h})\ |h_{22}|^2P_2(\mathbf{h})}\right) + \log^+\left(\frac{1+(1-|\rho(\mathbf{h})|^2)\max\{|h_{11}|^2,|h_{21}|^2\}P_1(\mathbf{h})}{1+(1-|\rho(\mathbf{h})|^2)|h_{21}|^2P_1(\mathbf{h})}\right)\}\Bigg] \tag{C.1}$$

When the EGCIFC is in strong interference ($|h_{21}|^2 \geq |h_{11}|^2$), then the sum-capacity in (Equation C.1) is maximized by $|\rho^\star(\mathbf{h})| = 1$. In this case the sum-capacity becomes the solution of the following optimization problem:

$$\max_{P_i\geq 0:\mathbb{E}[P_i(\mathbf{h})]\leq \bar{P}_i} \mathbb{E}\left[\log\left(1+\left(\sum_i |h_i|\sqrt{P_i(\mathbf{h})}\right)^2\right)\right]. \tag{C.2}$$

This is the same problem as that of finding the optimal power allocation for a point to point MISO with PerPC and the solution presented in Appendix B.

When the EGCIFC satisfies

$$|h_{11}|^2 > |h_{21}|^2, \tag{C.3}$$

Appendix C (Continued)

then the sum-capacity (Equation C.1) reduces to the following optimization problem:

$$\max_{P_i(\mathbf{h})\geq 0:\mathbb{E}[P_i(\mathbf{h})]\leq \bar{P_i},|\rho(\mathbf{h})|\leq 1} \mathbb{E}\Bigg[\log\left(1+|h_{21}|^2P_1(\mathbf{h})+|h_{22}|^2P_2(\mathbf{h})+2|\rho(\mathbf{h})|\sqrt{|h_{21}|^2P_1(\mathbf{h})\ |h_{22}|^2P_2(\mathbf{h})}\right) + \log\left(\frac{1+(1-|\rho(\mathbf{h})|^2)|h_{11}|^2P_1(\mathbf{h})}{1+(1-|\rho(\mathbf{h})|^2)|h_{21}|^2P_1(\mathbf{h})}\right)\Bigg]. \tag{C.4}$$

When (Equation C.3) holds, we will determine when the following $P_1(\mathbf{h}), P_2(\mathbf{h}), |\rho(\mathbf{h})|$ assignments are optimal by substituting them back into the KKT conditions for the following Lagrangian problem:

$$\max_{P_i(\mathbf{h})\geq 0,|\rho(\mathbf{h})|\leq 1,i\in[1:2]} \mathcal{L} = \mathbb{E}\Bigg[\log\left(1+|h_{21}|^2P_1(\mathbf{h})+|h_{22}|^2P_2(\mathbf{h})+2|\rho(\mathbf{h})|\sqrt{|h_{21}|^2P_1(\mathbf{h})\ |h_{22}|^2P_2(\mathbf{h})}\right) + \log\left(\frac{1+(1-|\rho(\mathbf{h})|^2)|h_{11}|^2P_1(\mathbf{h})}{1+(1-|\rho(\mathbf{h})|^2)|h_{21}|^2P_1(\mathbf{h})}\right)\Bigg] - \lambda_1\left(\mathbb{E}\left[P_1(\mathbf{h})\right]-\bar{P_1}\right)-\lambda_2\left(\mathbb{E}\left[P_2(\mathbf{h})\right]-\bar{P_2}\right)$$

Appendix C (Continued)

The KKT conditions for $i \in \{1, 2\}$ are then

$$\mu_i(\mathbf{h})P_i(\mathbf{h}) = 0 \tag{C.5a}$$

$$\gamma_1(\mathbf{h})|\rho(\mathbf{h})| = 0 \tag{C.5b}$$

$$\gamma_2(\mathbf{h})(1 - |\rho(\mathbf{h})|) = 0 \tag{C.5c}$$

$$P_i(\mathbf{h}) \geq 0 \tag{C.5d}$$

$$|\rho(\mathbf{h})| \leq 1 \tag{C.5e}$$

$$|\rho(\mathbf{h})| \geq 0 \tag{C.5f}$$

$$\gamma_i, \mu_i, \lambda_i \geq 0 \tag{C.5g}$$

$$\frac{\partial \mathcal{L}}{\partial P_1(\mathbf{h})} = \frac{|h_{21}|^2 + |\rho(\mathbf{h})|\sqrt{|h_{21}|^2|h_{22}|^2}\,\sqrt{\frac{P_2(\mathbf{h})}{P_1(\mathbf{h})}}}{1 + |h_{21}|^2P_1(\mathbf{h}) + |h_{22}|^2P_2(\mathbf{h}) + 2|\rho(\mathbf{h})|\sqrt{|h_{21}|^2P_1(\mathbf{h})\|h_{22}|^2P_2(\mathbf{h})}} + \frac{\left(|h_{11}|^2 - |h_{21}|^2\right)\,\left(1 - |\rho(\mathbf{h})|^2\right)}{\left(1 + (1 - |\rho(\mathbf{h})|^2)|h_{21}|^2P_1(\mathbf{h})\right)\left(1 + (1 - |\rho(\mathbf{h})|^2)|h_{11}|^2P_1(\mathbf{h})\right)} - (\lambda_1 - \mu_1(\mathbf{h})) = 0 \tag{C.5h}$$

$$\frac{\partial \mathcal{L}}{\partial P_2(\mathbf{h})} = \frac{|h_{22}|^2 + |\rho(\mathbf{h})|\sqrt{|h_{21}|^2|h_{22}|^2}\,\sqrt{\frac{P_1(\mathbf{h})}{P_2(\mathbf{h})}}}{1 + |h_{21}|^2P_1(\mathbf{h}) + |h_{22}|^2P_2(\mathbf{h}) + 2|\rho(\mathbf{h})|\sqrt{|h_{21}|^2P_1(\mathbf{h})|h_{22}|^2P_2(\mathbf{h})}} - (\lambda_2 - \mu_2(\mathbf{h})) = 0 \tag{C.5i}$$

$$\frac{\partial \mathcal{L}}{\partial |\rho(\mathbf{h})|} = \frac{\sqrt{|h_{21}|^2P_1(\mathbf{h})|h_{22}|^2P_2(\mathbf{h})}}{1 + |h_{21}|^2P_1(\mathbf{h}) + |h_{22}|^2P_2(\mathbf{h}) + 2|\rho(\mathbf{h})|\sqrt{|h_{21}|^2P_1(\mathbf{h})|h_{22}|^2P_2(\mathbf{h})}} - \frac{\left(|h_{11}|^2 - |h_{21}|^2\right)\,|\rho(\mathbf{h})|P_1}{\left(1 + (1 - |\rho(\mathbf{h})|^2)|h_{21}|^2P_1(\mathbf{h})\right)\left(1 + (1 - |\rho(\mathbf{h})|^2)|h_{11}|^2P_1(\mathbf{h})\right)} + (\gamma_1(\mathbf{h}) - \gamma_2(\mathbf{h})) = 0 \tag{C.5j}$$

Appendix C (Continued)

C.0.5 On the optimality of $|\rho^\star(\mathbf{h})| = 0$ (i.e., independent inputs):

$|\rho^\star(\mathbf{h})| = 0$ is an optimal solution if $\big($with $\gamma_2(\mathbf{h}) = 0$ and $\gamma_1(\mathbf{h}) \geq 0$ from (Equation C.5e) and (Equation C.5f) the following KKT conditions (evaluated at $|\rho^\star(\mathbf{h})| = 0$) are satisfied:

From (Equation C.5h)
$$\frac{|h_{21}|^2}{1+|h_{21}|^2P_1(\mathbf{h})+|h_{22}|^2P_2(\mathbf{h})} + \frac{\left(|h_{11}|^2-|h_{21}|^2\right)}{\left(1+|h_{21}|^2P_1(\mathbf{h})\right)\left(1+|h_{11}|^2P_1(\mathbf{h})\right)} = \lambda_1 - \mu_1(\mathbf{h}), \tag{C.6a}$$

From (Equation C.5i)
$$\frac{|h_{22}|^2}{1+|h_{21}|^2P_1(\mathbf{h})+|h_{22}|^2P_2(\mathbf{h})} = \lambda_2 - \mu_2(\mathbf{h}) \tag{C.6b}$$

From (Equation C.5j)
$$\frac{\sqrt{|h_{21}|^2P_1(\mathbf{h})|h_{22}|^2P_2(\mathbf{h})}}{1+|h_{21}|^2P_1(\mathbf{h})+|h_{22}|^2P_2(\mathbf{h})} = -\gamma_1(\mathbf{h}) \leq 0 \tag{C.6c}$$

Notice that from (Equation C.6c) we have that $P_1(\mathbf{h}) \cdot P_2(\mathbf{h}) = 0$, that is, the powers cannot be simultaneously strictly positive. We now proceed by finding the optimal power allocation.

Subcase 1: $P_1(\mathbf{h}) = 0$ and $P_2(\mathbf{h}) = 0$

From (Equation C.5a)
$$\text{if } P_1(\mathbf{h}) = 0 \rightarrow \mu_1(\mathbf{h}) \geq 0 \text{ and } P_2(\mathbf{h}) = 0 \rightarrow \mu_2(\mathbf{h}) \geq 0 \tag{C.7}$$

From (Equation C.6a)
$$\frac{\partial \mathcal{L}}{\partial P_2(\mathbf{h})} = |h_{22}|^2 - (\lambda_2 - \mu_2(\mathbf{h})) = 0 \rightarrow \frac{|h_{22}|^2}{\lambda_2} \leq 1 \tag{C.8}$$

From (Equation C.6b)
$$\frac{\partial \mathcal{L}}{\partial P_1(\mathbf{h})} = |h_{11}|^2 - (\lambda_1 - \mu_1(\mathbf{h})) = 0 \rightarrow \frac{|h_{11}|^2}{\lambda_1} \leq 1 \tag{C.9}$$

Conclusion 1: then $P_1^\star(\mathbf{h}) = 0$ and $P_2^\star(\mathbf{h}) = 0$ and $|\rho^\star(\mathbf{h})| = 0$ are optimal when:

$$\mathcal{R}_1 := \left\{ \frac{|h_{11}|^2}{\lambda_1}, \frac{|h_{21}|^2}{\lambda_1}, \frac{|h_{22}|^2}{\lambda_2} > 0 : \text{eq. (Equation C.3) holds}, \frac{|h_{11}|^2}{\lambda_1} \leq 1, \frac{|h_{22}|^2}{\lambda_2} \leq 1 \right\}.$$

Appendix C (Continued)

This can be interpreted as follows: since both direct channel gains are weak, the optimal scheme is to have the transmitters refrain from using any power and use power only when the channel gains are stronger.

Subcase 2: $P_1(\mathbf{h}) > 0$ and $P_2(\mathbf{h}) = 0$

From (Equation C.5a) $$\text{if } P_1(\mathbf{h}) > 0 \rightarrow \mu_1(\mathbf{h}) = 0 \text{ and } P_2(\mathbf{h}) > 0 \rightarrow \mu_2(\mathbf{h}) \geq 0 \tag{C.10}$$

From (Equation C.6a) $$\frac{\partial \mathcal{L}}{\partial P_1(\mathbf{h})} = \frac{|h_{21}|^2}{1+|h_{21}|^2 P_1(\mathbf{h})} + \frac{|h_{11}|^2 - |h_{21}|^2}{(1+|h_{21}|^2 P_1(\mathbf{h}))(1+|h_{11}|^2 P_1(\mathbf{h}))} - \lambda_1 = 0 \tag{C.11}$$

$$\rightarrow P_{1a}(\mathbf{h}) = \frac{1}{\lambda_1} - \frac{1}{|h_{11}|^2}$$

From (Equation C.6b) $$\frac{\partial \mathcal{L}}{\partial P_2(\mathbf{h})} = \frac{|h_{22}|^2}{1+|h_{21}|^2 P_1(\mathbf{h})} - (\lambda_2 - \mu_2(\mathbf{h})) = 0$$

$$\rightarrow P_{1b}(\mathbf{h}) = \frac{|h_{22}|^2}{(\lambda_2 - \mu_2(\mathbf{h}))|h_{21}|^2} - \frac{1}{|h_{21}|^2} \tag{C.12}$$

Evaluating $P_{1a}(\mathbf{h}) = P_{1b}(\mathbf{h})$ then we obtain

$$\frac{\frac{|h_{22}|^2}{\lambda_2}}{1 + \frac{|h_{21}|^2}{\lambda_1} - \frac{|h_{21}|^2}{|h_{11}|^2}} \leq 1$$

Conclusion 2: then $P_1^\star(\mathbf{h}) = P_{1WF}(h_{11}) = [\frac{1}{\lambda_1} - \frac{1}{|h_{11}|^2}]^+$ and $P_2^\star(\mathbf{h}) = 0$ and $|\rho^\star(\mathbf{h})| = 0$ are optimal when

$$\mathcal{R}_2 := \left\{ \frac{|h_{11}|^2}{\lambda_1}, \frac{|h_{21}|^2}{\lambda_1}, \frac{|h_{22}|^2}{\lambda_2} > 0 : \text{eq. (Equation C.3) holds}, \frac{|h_{11}|^2}{\lambda_1} > 1, \frac{\frac{|h_{22}|^2}{\lambda_2}}{1 + \frac{|h_{21}|^2}{\lambda_1} - \frac{|h_{21}|^2}{|h_{11}|^2}} \leq 1 \right\}.$$

Appendix C (Continued)

One can interpret this region as follows: since $|h_{11}|^2$ is strong enough the cognitive user water-fills over its direct link, while the primary transmitter refrains until it sees better channel gain conditions.

Subcase 3: $P_1(\mathbf{h}) = 0$ and $P_2(\mathbf{h}) > 0$

From (Equation C.5a) $$\text{if } P_1(\mathbf{h}) = 0 \rightarrow \mu_1(\mathbf{h}) \geq 0 \text{ and } P_2(\mathbf{h}) > 0 \rightarrow \mu_2(\mathbf{h}) = 0 \tag{C.13}$$

From (Equation C.6a) $$\frac{\partial \mathcal{L}}{\partial P_1(\mathbf{h})} = \frac{|h_{21}|^2}{1 + |h_{22}|^2 P_2(\mathbf{h})} + \left(|h_{11}|^2 - |h_{21}|^2\right) = (\lambda_1 - \mu_1(\mathbf{h})) \leq \lambda_1 \tag{C.14}$$

From (Equation C.6b) $$\frac{\partial \mathcal{L}}{\partial P_2(\mathbf{h})} = \frac{|h_{22}|^2}{1 + |h_{22}|^2 P_2(\mathbf{h})} - \lambda_2 = 0 \rightarrow P_2^\star(\mathbf{h}) = \frac{1}{\lambda_2} - \frac{1}{|h_{22}|^2} > 0$$

$$\rightarrow \frac{|h_{22}|^2}{\lambda_2} > 1 \tag{C.15}$$

Evaluating (Equation C.14) for $P_2^\star(\mathbf{h})$ gives

$$\frac{\frac{|h_{21}|^2}{\lambda_1}}{\frac{|h_{22}|^2}{\lambda_2}} + \frac{|h_{11}|^2}{\lambda_1} - \frac{|h_{21}|^2}{\lambda_1} \leq 1$$

Conclusion 3: then $P_1^\star(\mathbf{h}) = 0$ and $P_2^\star(\mathbf{h}) = P_{2\text{WF}}(h_{22}) = [\frac{1}{\lambda_2} - \frac{1}{|h_{22}|^2}]^+$ and $|\rho^\star(\mathbf{h})| = 0$ are optimal when

$$\mathcal{R}_3 := \left\{ \frac{|h_{11}|^2}{\lambda_1}, \frac{|h_{21}|^2}{\lambda_1}, \frac{|h_{22}|^2}{\lambda_2} > 0 : \text{eq. (Equation C.3) holds}, \frac{|h_{22}|^2}{\lambda_2} > 1, \frac{\frac{|h_{21}|^2}{\lambda_1}}{\frac{|h_{22}|^2}{\lambda_2}} + \frac{|h_{11}|^2}{\lambda_1} - \frac{|h_{21}|^2}{\lambda_1} \leq 1 \right\}.$$

Appendix C (Continued)

One can interpret this region as follows: since $|h_{22}|^2$ is strong enough the primary user water-fills over its direct link, while the cognitive transmitter refrains until it sees better channel gains.

C.0.6 On the optimality of $|\rho^\star(\mathbf{h})| = 1$ (i.e., identical inputs up to affine transformation):

$|\rho(\mathbf{h})| = 1$ is an optimal solution if $\gamma_1(\mathbf{h}) = 0$ and $\gamma_2(\mathbf{h}) \geq 0$.

From (Equation C.5h)
$$\frac{|h_{21}|^2 + \sqrt{|h_{21}|^2|h_{22}|^2}\,\sqrt{\frac{P_2(\mathbf{h})}{P_1(\mathbf{h})}}}{1 + |h_{21}|^2 P_1(\mathbf{h}) + |h_{22}|^2 P_2(\mathbf{h}) + 2\sqrt{|h_{21}|^2 P_1(\mathbf{h})|h_{22}|^2 P_2(\mathbf{h})}} = \lambda_1 - \mu_1(\mathbf{h}) \tag{C.16a}$$

From (Equation C.5i)
$$\frac{|h_{22}|^2 + \sqrt{|h_{21}|^2|h_{22}|^2}\,\sqrt{\frac{P_1(\mathbf{h})}{P_2(\mathbf{h})}}}{1 + |h_{21}|^2 P_1(\mathbf{h}) + |h_{22}|^2 P_2(\mathbf{h}) + 2\sqrt{|h_{21}|^2 P_1(\mathbf{h})|h_{22}|^2 P_2(\mathbf{h})}} = \lambda_2 - \mu_2(\mathbf{h}) \tag{C.16b}$$

From (Equation C.5j)
$$\frac{\sqrt{|h_{21}|^2 P_1(\mathbf{h})|h_{22}|^2 P_2(\mathbf{h})}}{1 + |h_{21}|^2 P_1 + |h_{22}|^2 P_2(\mathbf{h}) + 2\sqrt{|h_{21}|^2 P_1(\mathbf{h})|h_{22}|^2 P_2(\mathbf{h})}} \geq \left(|h_{11}|^2 - |h_{21}|^2\right) P_1(\mathbf{h}) \tag{C.16c}$$

Subcase 4: $P_1(\mathbf{h}) > 0,\ P_2(\mathbf{h}) > 0$

Appendix C (Continued)

If both powers are strictly positive we further have that $\mu_1(\mathbf{h}) = \mu_2(\mathbf{h}) = 0$; this implies that we need to solve

From (Equation C.16a)
$$\sqrt{|h_{21}|^2 P_1(\mathbf{h})}\frac{\sqrt{|h_{21}|^2 P_1(\mathbf{h})} + \sqrt{|h_{22}|^2 P_2(\mathbf{h})}}{1 + (\sqrt{|h_{21}|^2 P_1(\mathbf{h})} + \sqrt{|h_{22}|^2 P_2(\mathbf{h})})^2} = \lambda_1 P_1(\mathbf{h}),$$

$$P_1(\mathbf{h}) > 0, \tag{C.17a}$$

From (Equation C.16b)
$$\sqrt{|h_{22}|^2 P_2(\mathbf{h})}\frac{\sqrt{|h_{21}|^2 P_1(\mathbf{h})} + \sqrt{|h_{22}|^2 P_2(\mathbf{h})}}{1 + (\sqrt{|h_{21}|^2 P_1(\mathbf{h})} + \sqrt{|h_{22}|^2 P_2(\mathbf{h})})^2} = \lambda_2 P_2(\mathbf{h}),$$

$$P_2(\mathbf{h}) > 0, \tag{C.17b}$$

From (Equation C.16c)
$$\frac{\sqrt{|h_{21}|^2 P_1(\mathbf{h})\ |h_{22}|^2 P_2(\mathbf{h})}}{1 + (\sqrt{|h_{21}|^2 P_1(\mathbf{h})} + \sqrt{|h_{22}|^2 P_2(\mathbf{h})})^2} \geq \left(\frac{|h_{11}|^2}{\lambda_1} - \frac{|h_{21}|^2}{\lambda_1}\right)\lambda_1 P_1(\mathbf{h}).$$

(C.17c)

This is the same optimization problem as that in the MISO channel with PerPC. The optimal powers in this case are given by:

$$P_1^\star(\mathbf{h}) = \left(\frac{\frac{|h_{21}|^2}{\lambda_1} + \frac{|h_{22}|^2}{\lambda_2} - 1}{\left(\frac{|h_{21}|^2}{\lambda_1} + \frac{|h_{22}|^2}{\lambda_2}\right)^2}\right)\frac{|h_{21}|^2}{\lambda_1^2} \tag{C.18}$$

$$P_2^\star(\mathbf{h}) = \left(\frac{\frac{|h_{21}|^2}{\lambda_1} + \frac{|h_{22}|^2}{\lambda_2} - 1}{\left(\frac{|h_{21}|^2}{\lambda_1} + \frac{|h_{22}|^2}{\lambda_2}\right)^2}\right)\frac{|h_{22}|^2}{\lambda_1^2} \tag{C.19}$$

Appendix C (Continued)

We next need to verify that, for these optimizing powers, we satisfy (Equation C.17c); by doing so, we conclude that if

$$\frac{|h_{11}|^2}{\lambda_1} \leq \frac{|h_{21}|^2}{\lambda_1} + \frac{\frac{|h_{22}|^2}{\lambda_2}}{\frac{|h_{21}|^2}{\lambda_1} + \frac{|h_{22}|^2}{\lambda_2}} \tag{C.20}$$

$$\frac{|h_{21}|^2}{\lambda_1} + \frac{|h_{22}|^2}{\lambda_2} > 1, \tag{C.21}$$

Conclusion 4: then $P_1^\star(\mathbf{h}) =$ eq. (Equation C.18) and $P_2^\star(\mathbf{h}) =$ eq. (Equation C.19) and $|\rho^\star(\mathbf{h})| = 1$ are optimal when

$$\mathcal{R}_4 : \left\{ \frac{|h_{22}|^2}{\lambda_1}, \frac{|h_{21}|^2}{\lambda_1}, \frac{|h_{22}|^2}{\lambda_2} > 0 : \text{eq. (Equation C.3) holds}, \frac{|h_{11}|^2}{\lambda_1} \leq \frac{|h_{21}|^2}{\lambda_1} + \frac{\frac{|h_{22}|^2}{\lambda_2}}{\frac{|h_{21}|^2}{\lambda_1} + \frac{|h_{22}|^2}{\lambda_2}} \right\}.$$

The channel gain condition in $\mathcal{R}_4$ implies that $\frac{|h_{11}|^2}{\lambda_1} \leq \frac{|h_{21}|^2}{\lambda_1} + \frac{|h_{22}|^2}{\lambda_2}$ (given that $\frac{|h_{21}|^2}{\lambda_1} + \frac{|h_{22}|^2}{\lambda_2} > 1$). In this case the sum of the channel gains to the primary receiver is stronger than that of the direct gain to the cognitive receiver and so one would suspect that performing a MISO type of scheme is optimal which is in fact the case.

C.0.7 On the optimality $0 < |\rho(\mathbf{h})| < 1$ with $P_1(\mathbf{h}) > 0$ and $P_2(\mathbf{h}) > 0$:

When $0 < |\rho(\mathbf{h})| < 1$ is optimal (with $\gamma_1(\mathbf{h}) = 0$ and $\gamma_2(\mathbf{h}) = 0$ from (Equation C.5e) and (Equation C.5f)), we show that $P_1^\star(\mathbf{h}) > 0$ and $P_2^\star(\mathbf{h}) > 0$ are optimal by contradiction (whose exact values are to be find numerically).

Appendix C (Continued)

Suppose that $P_1(\mathbf{h}) = 0$ and $P_2(\mathbf{h}) = 0$ is optimal when

$$\mathcal{R}_5 : \left\{ \frac{|h_{11}|^2}{\lambda_1}, \frac{|h_{21}|^2}{\lambda_1}, \frac{|h_{22}|^2}{\lambda_2} > 0 : \text{eq. (Equation C.3) holds}, \left(\mathcal{R}_1 \cup \mathcal{R}_2 \cup \mathcal{R}_3 \cup \mathcal{R}_4 \cup \mathcal{R}_5\right)^c \right\} \tag{C.22}$$

Substituting this assumption back in the KKT conditions, gives the region defined by $\mathcal{R}_1$ which contradicts the presumption of the region.

Suppose that $P_1(\mathbf{h}) > 0$ and $P_2(\mathbf{h}) = 0$ is optimal when (Equation C.22). Substituting this assumption back in the KKT conditions, gives the region defined by $\mathcal{R}_2$ which contradicts the presumption of the region.

Suppose that $P_1(\mathbf{h}) = 0$ and $P_2(\mathbf{h}) > 0$ is optimal when (Equation C.22). Substituting this assumption back in the KKT conditions, gives the region defined by $\mathcal{R}_3$ which contradicts the presumption of the region.

Suppose that $P_1(\mathbf{h}) > 0$ and $P_2(\mathbf{h}) > 0$ and $|\rho^\star(\mathbf{h})| = 1$ is optimal when (Equation C.22). Substituting this assumption back in the KKT conditions, gives the region defined by $\mathcal{R}_4$ which contradicts the presumption of the region.

Conclusion 6: Thus, $P_1^\star(\mathbf{h}) > 0$ and $P_2^\star(\mathbf{h}) > 0$ and $0 < |\rho^\star(\mathbf{h})| < 1$ whose values are to be found numerically are optimal when

$$\mathcal{R}_5 := \left\{ \frac{|h_{11}|^2}{\lambda_1}, \frac{|h_{21}|^2}{\lambda_1}, \frac{|h_{22}|^2}{\lambda_2} > 0 : \text{eq. (Equation C.3) holds}, \left(\mathcal{R}_1, \mathcal{R}_2, \mathcal{R}_3, \mathcal{R}_4, \mathcal{R}_5\right)^c \right\}.$$

Appendix D

PROOF OF THEOREM 3.3.1

By Fano's inequality $H(W_i|Y_i^N) \leq N\epsilon_N$ with $\epsilon_N \to 0$ as $N \to \infty$ for all $i \in [1:3]$. The bounds in equation (Equation 3.6a) through (Equation 3.6c) are a simple application of the cut-set bound. The bound in (Equation 3.6d) is obtained as follows:

$$
\begin{aligned}
N(R_2 + R_3 - 2\epsilon_N) &\overset{(a)}{\leq} I(Y_2^N; W_2) + I(Y_3^N; W_3) \overset{(b)}{\leq} I(Y_2^N, W_1; W_2) + I(Y_3^N, Y_2^N, W_1, W_2; W_3) \\
&\overset{(c)}{=} I(Y_2^N; W_2|W_1) + I(Y_3^N, Y_2^N; W_3|W_1, W_2) \\
&\overset{(d)}{=} I(Y_2^N; W_2|W_1) + I(Y_2^N; W_3|W_1, W_2) + I(Y_3^N; W_3|W_1, W_2, Y_2^N) \\
&\overset{(e)}{=} I(Y_2^N; W_2, W_3|W_1) + I(Y_3^N; W_3|W_1, W_2, Y_2^N) \\
&\overset{(f)}{=} I(Y_2^N; W_2, W_3|W_1, X_1^N) + I(Y_3^N; W_3|W_1, W_2, Y_2^N, X_1^N, X_2^N) \\
&\overset{(g)}{\leq} \sum_{t=1}^{N} H(Y_{2,t}|X_{1,t}) - H(Y_{2,t}|X_{1,t}, X_{2,t}, X_{3,t}) + H(Y_{3,t}|X_{1,t}, X_{2,t}) - H(Y_{3,t}|X_{1,t}, X_{2,t}, X_{3,t}) \\
&\overset{(h)}{=} \sum_{t=1}^{N} I(Y_{2,t}; X_{2,t}, X_{3,t}|X_{1,t}) + I(Y_{3,t}; X_{3,t}|X_{1,t}, X_{2,t}),
\end{aligned}
$$

where (a) follows from Fano's inequality, (b) the non-negativity of mutual information, (c) from the independence of the messages, (d) and (e) from the chain rule (note the side information allows one to recombine different entropy terms), (f) because the inputs are deterministic functions of the messages, (g) follows since conditioning reduces entropy, and (h) definition of mutual information. Using similar steps (give enough messages to reconstruct the inputs, and

Appendix D (Continued)

give outputs to recombine terms by using the chain rule of mutual information) we obtain the bound in (Equation 3.6e). The main steps are:

$$
\begin{aligned}
& N(R_1 + R_2 + R_3 - 3\epsilon_N) \\
& \leq I(Y_1^N; W_1) + I(Y_2^N; W_2) + I(Y_3^N; W_3) \\
& \leq I(Y_1^N; W_1) + I(Y_2^N, Y_1^N, W_1; W_2) + I(Y_3^N, Y_1^N, W_1, Y_2^N, W_2; W_3) \\
& \leq I(Y_1^N; W_1, W_2, W_3) + I(Y_2^N; W_2, W_3 | Y_1^N, W_1) + I(Y_3^N; W_3 | Y_1^N, W_1, Y_2^N, W_2) \\
& \leq \sum_{t=1}^{N} I(Y_{1,t}; X_{1,t}, X_{2,t}, X_{3,t}) + I(Y_{2,t}; X_{2,t}, X_{3,t} | X_{1,t}, Y_{1,t}) + I(Y_{3,t}; X_{3,t} | X_{1,t}, X_{2,t}, Y_{1,t}, Y_{2,t}).
\end{aligned}
$$

Appendix E

PROOF OF THEOREM 3.3.2

By Fano's inequality $H(W_i|Y_i^N) \leq N\epsilon_N$ with $\epsilon_N \to 0$ as $N \to \infty$ for all $i \in [1 : K]$. For (Equation 3.7a) we have

$$
\begin{aligned}
& N(R_i - \epsilon_N) \\
& \leq I(Y_i^N; W_i) \\
& \leq I(Y_i^N; W_i | W_1, \ldots, W_{i-1}) \\
& = \sum_{t=1}^{N} h(Y_{i,t}|W_1, \ldots, W_{i-1}, Y_i^{t-1}) - h(Y_{i,t}|W_1, \ldots, W_i, Y_i^{t-1}) \\
& \leq \sum_{t=1}^{N} h(Y_{i,t}|W_1, \ldots, W_{i-1}) - h(Y_{i,t}|W_1, \ldots, W_K, Y_i^{t-1}) \\
& \leq \sum_{t=1}^{N} h(Y_{i,t}|X_{1,t}, \ldots, X_{i-1,t}) - h(Y_{i,t}|X_{1,t}, \ldots, X_{K,t}) \\
& = \sum_{t=1}^{N} I(Y_{i,t}; X_{i,t}, \ldots, X_{K,t}|X_{1,t}, \ldots, X_{i-1,t}).
\end{aligned}
$$

Appendix E (Continued)

For (Equation 3.7b) we have

$$
\begin{aligned}
N\sum_{j=i}^{K}(R_j-\epsilon_N) &\le \sum_{j=i}^{K} I(Y_j^N;W_j) \le \sum_{j=i}^{K} I(Y_j^N,\ \underbrace{W_1,\ldots,W_{i-1}}_{=\emptyset \text{ for } i=1},\ \underbrace{Y_i^N,W_i,\ldots Y_{j-1}^N,W_{j-1}}_{=\emptyset \text{ for } j=i};W_j)\\
&= \sum_{j=i}^{K} I(Y_i^N,\ldots Y_j^N;W_j|W_1,\ldots,W_{j-1})\\
&= \sum_{j=i}^{K}\sum_{k=i}^{j} I(Y_k^N;W_j|W_1,\ldots,W_{j-1},\ Y_i^N,\ldots,Y_{k-1}^N)\\
&= \sum_{k=i}^{K}\sum_{j=k}^{K} I(Y_k^N;W_j|W_1,\ldots,W_{j-1},\ Y_i^N,\ldots,Y_{k-1}^N)\\
&= \sum_{k=i}^{K} I(Y_k^N;W_k,\ldots,W_K|\underbrace{W_1,\ldots,W_{i-1}}_{=\emptyset \text{ for } i=1},\ \underbrace{W_i,Y_i^N,\ldots,W_{k-1},Y_{k-1}^N}_{=\emptyset \text{ for } k=i})\\
&\le \sum_{k=i}^{K}\sum_{t=1}^{N} I(Y_{k,t};X_{k,t},\ldots,X_{K,t}|X_{1,t},\ldots,X_{k-1,t},Y_{i,t},\ldots,Y_{k-1,t}).
\end{aligned}
$$

Appendix F

PROOF OF THEOREM 3.7.1

In order to obtain the sum-rate upper bound in (Equation 3.20) we first present the following Lemma, which is an extension to any K of [18, Lemma 5] (for CMS) and [16, Lemma 1] (for CoMS):

Lemma F.0.3. *If per-letter condition in* (Equation 3.19) *is satisfied for all prescribed input distributions then*

$$\mathrm{I}(\mathsf{X}_{[\mathsf{j}:\mathsf{K}]}^{\mathsf{N}};\mathsf{Y}_{\mathsf{j}}^{\mathsf{N}}|\mathsf{X}_{[1:\mathsf{j}-1]}^{\mathsf{N}}) \leq \mathrm{I}(\mathsf{X}_{[\mathsf{j}:\mathsf{K}]}^{\mathsf{N}};\mathsf{Y}_{\mathsf{j}-1}^{\mathsf{N}}|\mathsf{X}_{[1:\mathsf{j}-1]}^{\mathsf{N}}), \tag{F.1}$$

for all input distributions $\mathsf{P}_{\mathsf{X}_1^{\mathsf{N}},\ldots,\mathsf{X}_{\mathsf{K}}^{\mathsf{N}}}$, $\mathsf{N} \in \mathbb{N}$, *that factor as*

1. *PMS:* $\prod_{\mathsf{j}=1}^{\mathsf{K}-1} \mathsf{P}_{\mathsf{X}_{\mathsf{j}}^{\mathsf{N}}}\ \mathsf{P}_{\mathsf{X}_{\mathsf{K}}^{\mathsf{N}}|\mathsf{X}_1^{\mathsf{N}},\ldots,\mathsf{X}_{\mathsf{K}-1}^{\mathsf{N}}}$,
2. *CMS:* $\mathsf{P}_{\mathsf{X}_1^{\mathsf{N}},\ldots,\mathsf{X}_{\mathsf{K}}^{\mathsf{N}}}$.

Appendix F (Continued)

We are now ready to present the proof of Theorem 3.7.1: For $\epsilon_N > 0 : \epsilon_N \to 0$ as $N \to +\infty$

$$N\sum_{j=1}^{K}(R_j-\epsilon_N) \overset{(a)}{\leq} \sum_{j=1}^{K} I(W_j;Y_j^N) \overset{(b)}{\leq} \sum_{j=1}^{K} I(W_j;Y_j^N|W_{[1:j-1]}) \leq \sum_{j=1}^{K} I(X_j^N;Y_j^N|X_{[1:j-1]}^N)$$
$$\overset{(c)}{=} \sum_{j=1}^{K-1} I(X_j^N;Y_j^N|X_{[1:j-1]}^N) + I(X_K^N;Y_K^N|X_{[1:K-1]}^N) \overset{(d)}{\leq} \sum_{j=1}^{K-1} I(X_j^N;Y_j^N|X_{[1:j-1]}^N) + I(X_K^N;Y_{K-1}^N|X_{[1:K-1]}^N)$$
$$\overset{(e)}{=} \sum_{j=1}^{K-2} I(X_j^N;Y_j^N|X_{[1:j-1]}^N) + I(X_{[K-1:K]}^N;Y_{K-1}^N|X_{[1:K-2]}^N) \overset{(f)}{\leq} \sum_{j=1}^{K-2} I(X_j^N;Y_j^N|X_{[1:j-1]}^N) + I(X_{[K-1:K]}^N;Y_{K-2}^N|X_{[1:K-2]}^N)$$
$$\ldots \overset{(g)}{\leq} I(X_{[1:K]}^N;Y_1^N) \overset{(h)}{\leq} \sum_{t=1}^{N} I(X_{1t},\ldots,X_{Kt};Y_{1t}) \overset{(i)}{\leq} NI(X_{[1:K]};Y_1|Q) \overset{(j)}{\leq} NI(X_{[1:K]};Y_1),$$

where: (a) follows from Fano's inequalities $H(W_j|Y_j^N) \leq N\epsilon_N$, $\forall j \in [1:K]$, (c) from the independence of messages, the definition of encoding functions (for all $\mathcal{M}_j \subseteq [1:j]$, $j \in [1:K]$) and data processing inequality, (d), (f) and (g) from the condition in (Equation 3.19) for $j = K$, $j = K-1$, up to $j = 2$ and Lemma F.0.3, (h) from the chain rule of entropy and from the fact that conditioning reduces entropy, (i) by introducing a time-sharing random variable that is uniformly distributed $Q \sim \mathcal{U}nif[1:N]$, and (j) by conditioning reduces entropy.

Appendix G

PROOF OF THEOREM 4.1.1

The analysis of the coverage probability for the mmWave heterogeneous network is similar to that in [22, Appendix A] with two major differences. The first difference is the presence of multiple antennas at the transmitter. The second difference is that the interference is a function of i.i.d uniformly distributed random variables, assuming that the path angles are independent and uniformly distributed over $[-\pi, +\pi]$.

Let $\Theta_k = \{\frac{\|v\|^\alpha}{P_k}, v \in \Phi_{\text{mmw},k}\}$ for $k \in [1 : K]$. Its density can be derived using the Mapping theorem [60, Thm. 2.34] and is given by

$$\lambda_k(v) = \lambda_k \frac{2\pi}{\alpha} P_k^{\frac{2}{\alpha}} v^{\frac{2}{\alpha}-1}, \ k \in [1 : K]. \tag{G.1}$$

The process $\Theta = \cup_{k=1}^K \Theta_k$ has a density

$$\lambda(v) = \sum_{k=1}^K \lambda_k(v). \tag{G.2}$$

G.1 Distribution of closest base stations

We assume that the elements in the process Θ are indexed in increasing order. Let

$$\gamma_i' = \frac{\|v_i\|^\alpha}{P_{f(v_i)}}$$

Appendix G (Continued)

then $\gamma' = \{\gamma'_1, \cdots, \gamma'_n\}$ denotes the set of *normalized pathloss* of the cooperating base stations. We first present the distribution of the two nearest base stations by following similar steps as done in [61], then derive the distribution of n closest base stations. The distribution of the closest two base stations (assuming two cooperating base stations) is given by

$$f_{\Gamma'}(\gamma'_1, \gamma'_2) = f_{\Gamma'_2|\Gamma'_1}(\gamma'_2|\gamma'_1) f_{\Gamma'_1}(\gamma'_1) \tag{G.3}$$

where the distribution of the first closest base station from the null probability of a 2-D Poisson point process is

$$f_{\Gamma'_1}(\gamma'_1) = \lambda(\gamma'_1) e^{-\Lambda(\gamma'_1)} \tag{G.4}$$

while the conditional distribution is given by

$$f_{\Gamma'_2|\Gamma'_1}(\gamma'_2|\gamma'_1) = \lambda(\gamma'_2) e^{-\Lambda(\gamma'_2)+\Lambda(\gamma'_1)} \tag{G.5}$$

The joint distribution for the case of $n = 2$ base stations is obtained by substituting (Equation G.4) and (Equation G.5) in (Equation G.3). The result can be generalized to any number n of co-operating base stations

$$f_{\Gamma'}(\gamma') = \prod_{i=1}^{n} \lambda(\gamma'_i) e^{-\Lambda(\gamma'_n)}. \tag{G.6}$$

Appendix G (Continued)

G.2 Derivation of Coverage Probability

The SINR expression at the typical user with omnidirectional reception is given by (Equation 4.13) and is reported next for convenience

$$\text{SINR} = \frac{\big|\sum_{i=1}^{n}\sqrt{\gamma_{v_i}}h_{v_i}\big|^2}{\frac{\sigma^2}{N_t}+\sum_{i=1}^{|\mathcal{T}^c|}\gamma_{l_i}|h_{l_i}|^2\big|G_t(\Omega_{\phi^t_{l_i}}-\Omega_{\theta^t_{l_i}})\big|^2}. \tag{G.7}$$

We have assumed that the cooperating base stations have *normalized pathloss* γ'_i by $i \leq n$, then the desired signal power at the numerator of (Equation G.7) can be re-written (after dropping the index v_i and replacing it with just i) as

$$S = \big|\sum_{i=1}^{n}\sqrt{\gamma_{v_i}}h_{v_i}\big|^2 = \big|\sum_{i\leq n}\gamma_i'^{-1/2}h_i\big|^2.$$

We have that the interfering base stations are indexed with $i > n$, then the power of the interference I can be expressed (by replacing the index l_i with just i) as

$$\begin{aligned} I &= \sum_{i=1}^{|\mathcal{T}|^c}\gamma_{l_i}|h_{l_i}|^2|G_t(\Omega_{\phi^t_{l_i}}-\Omega_{\theta^t_{l_i}})|^2 \\ &= \sum_{i>n}\gamma_i'^{-1}|h_i|^2\big|G_t(\Upsilon_i)\big|^2 \end{aligned} \tag{G.8}$$

Appendix G (Continued)

The coverage probability for a threshold T, can be re-written as

$$\begin{aligned}
\mathbb{P}(\text{SINR} > T) &= \mathbb{P}\left(S > T(I + \frac{\sigma^2}{N_t})\right) \\
&= \mathbb{E}_{\gamma',I}\left[\mathbb{P}\left(\left|\sum_{i\leq n} \gamma_i'^{-1/2} h_i\right|^2 > T(I + \frac{\sigma^2}{N_t})\middle|\gamma', I\right)\right] \\
&\overset{(a)}{=} \mathbb{E}_{\gamma',I}\left[\exp\left(\frac{-T(I + \frac{\sigma^2}{N_t})}{\sum_{i\leq n}\gamma_i'^{-1}}\right)\right] \\
&\overset{(b)}{=} \mathbb{E}_{\gamma'}\left[\mathcal{L}_I\left(\frac{T}{\sum_{i\leq n}\gamma_i'^{-1}}\right)\mathcal{L}_N\left(\frac{T}{\sum_{i\leq n}\gamma_i'^{-1}}\right)\right. \\
&\overset{(c)}{=} \int_{0<\gamma_1'<\cdots<\gamma_n'<+\infty} \mathcal{L}_I\left(\frac{T}{\sum_{i\leq n}\gamma_i'^{-1}}\right)\mathcal{L}_N\left(\frac{T}{\sum_{i\leq n}\gamma_i'^{-1}}\right)
\end{aligned}$$

where (a) follows from the cumulative density function of the exponentially distributed random variable S (due to Rayleigh fading assumption) with mean $\sum_{i\leq n}\gamma_i'^{-1}$; (b) follows from the definition of the Laplace transform of I, $\mathcal{L}_I(s) = \mathbb{E}[e^{-sI}]$ and the Laplace transform of the noise, $\mathcal{L}_N(s) = \mathbb{E}[e^{-s\sigma^2/N_t}]$; (c) by definition of the expectation with respect to the distribution of γ'.

Next we evaluate the Laplace transform of the interference I, but before going into the details of the derivation we need to find the distribution of $\Upsilon_i := \Omega_{\phi^t_{l_i}} - \Omega_{\theta^t_{l_i}}$, since the interference in (Equation G.8) is a function of the beam forming gain function which in turn is a function of Υ_i. The beam forming gain is given by (Equation 4.10).

Appendix G (Continued)

With the assumption that the interfering path angles of departure and the beamsteering angle used by the interfering base stations are i.i.d $\sim \mathsf{U}([-\pi, +\pi])$, then the directional cosine $\Omega_{\phi^t_{l_i}}$ and $\Omega_{\theta^t_{l_i}}$ are random variables with the following common probability density function

$$f_\Omega(\omega) = \begin{cases} \frac{1}{\pi\sqrt{1-\omega^2}} & \text{if } -1 \leq \omega \leq 1; \\ 0 & \text{otherwise.} \end{cases}$$

then the distribution of $\Upsilon_i = \Omega_{\phi^t_{l_i}} - \Omega_{\theta^t_{l_i}}$ is the result of the convolution of the probability density functions of $\Omega_{\phi^t_{l_i}}$ and $\Omega_{\theta^t_{l_i}}$ and is given by

$$f_{\Upsilon_i}(\varepsilon_i) = \int_{\max\{-1,-1-\varepsilon_i\}}^{\min\{1,1-\varepsilon_i\}} \left(\frac{1}{\pi^2\sqrt{1-(\varepsilon_i+\omega)^2}} \frac{1}{\sqrt{1-\omega^2}} \right) dy \tag{G.9}$$

Then the Laplace transform of the interference can be derived

$$\begin{aligned} \mathcal{L}_I(s) &= \mathbb{E}\left[e^{-s\sum\limits_{i>n} \gamma_i'^{-1}|h_i|^2|G_t(\Upsilon_i)|^2} \right] \\ &= \mathbb{E}\left[\prod_{i>n} \left(e^{-s\gamma_i'^{-1}|h_i|^2|G_t(\Upsilon_i)|^2} \right) \right] \\ &\overset{(a)}{=} \mathbb{E}_{\{\Upsilon_i\},\Theta}\left[\prod_{i>n} \mathbb{E}_{|h|^2}\left(e^{-s\gamma_i'^{-1}|h|^2|G_t(\Upsilon_i)|^2} \right) \right] \\ &\overset{(b)}{=} \mathbb{E}_{\{\Upsilon_i\},\Theta}\left[\prod_{i>n} \left(\frac{1}{1+s|G_t(\Upsilon_i)|^2\gamma_i'^{-1}} \right) \right] \end{aligned}$$

Appendix G (Continued)

$$\stackrel{(c)}{=} \mathbb{E}_{\Theta}\left[\prod_{i>n}\mathbb{E}_{\Upsilon}\left(\frac{1}{1+s|G_t(\Upsilon)|^2\gamma_i'^{-1}}\right)\right] \tag{G.10}$$

$$\stackrel{(d)}{=} \mathbb{E}_{\Theta}\left[\prod_{i>n}\left(\int_{-2}^{+2}\left(\frac{1}{1+s|G_t(\varepsilon)|^2\gamma_i'^{-1}}\right)f_{\Upsilon}(\varepsilon)\ d\varepsilon\right)\right]$$

$$\stackrel{(e)}{=} \exp\left(-\int_{\gamma_n'}^{\infty}\left[1-\int_{-2}^{+2}\left(\frac{1}{1+s|G_t(\varepsilon)|^2\nu^{-1}}\right)f_{\Upsilon}(\varepsilon)\ d\varepsilon\right]\lambda(\nu)\,d\nu\right) \tag{G.11}$$

where (a) follows from the i.i.d distribution of $|h_i|^2$ and their independence from Θ and Υ_i; where (b) follows from the Rayleigh fading assumption and the moment generating function of an exponential random variable; where (c) follows from the i.i.d distribution of Υ_i and their independence from Θ; (d) from the taking the expectation with respect to the random variable Υ_i whose distribution is given by (Equation G.9); (e) follows from the probability generating function of poisson point process [60, Thm. 4.9] (as used in [23, Eq. (38)]) and where $\lambda(\nu)$ is given by (Equation G.1).

Next we give an approximation of the Laplace transform of the interference for easier numerical evaluations (by approximating the beam forming gain function by a piecewise linear

Appendix G (Continued)

function) and to compare our results with [23], the Laplace transform of the interference is then given by

$$\begin{aligned}
\mathcal{L}_I(s) &\overset{(d)}{=} \mathbb{E}_\Theta\left[\prod_{i>n}\left(\int_{-2}^{+2}\left(\frac{1}{1+s|G_t(\varepsilon)|^2\gamma_i'^{-1}}\right)f_\Upsilon(\varepsilon)\ d\varepsilon\ \right)\right] \\
&\overset{(e)}{\approx} \mathbb{E}_\Theta\left[\prod_{i>n}\left(\int_{-2}^{-1/L_t} f_\Upsilon(\varepsilon)\ d\varepsilon + \int_{-1/L_t}^{1/L_t}\frac{1}{1+s\gamma_i'^{-1}}f_\Upsilon(\varepsilon)\ d\varepsilon + \int_{1/L_t}^{2} f_\Upsilon(\varepsilon)\ d\varepsilon\right)\right] \\
&\overset{(f)}{=} \mathbb{E}_\Theta\left[\prod_{i>n}\left(1-\frac{c\ s\gamma_i^{-1}}{1+s\gamma_i^{-1}}\right)\right] \\
&\overset{(g)}{=} \exp\left(-\int_{\gamma_n'}^{\infty}\left[\frac{c\ s\nu^{-1}}{1+s\nu^{-1}}\right]\lambda(\nu)\ d\nu\right)
\end{aligned}$$

where in (e) an approximation of the gain function was used which is given

$$G_t(\varepsilon) = \begin{cases} 1 & \text{if } -\frac{1}{L_t} \le \varepsilon \le \frac{1}{L_t}, \quad L_t = N_t\Delta_t \\ 0 & \text{if otherwise.} \end{cases}$$

(f) defining $c := \int_{-1/L_t}^{+1/L_t} f_\Upsilon(\varepsilon)\ d\varepsilon$.

If $c = 1$ then the Laplace transform in step (g) simplifies to that in [23, Eq. (38)].

Appendix H

PROOF OF THEOREM 4.2.1

H.1 Intensity and Intensity Measure

Let $\Theta_k = \{\frac{\|v\|^\alpha}{P_k}, v \in \Phi_{\text{mmw},k}\}$ for $k \in [1:K]$ with intensity $\lambda_k(v)$ given in (Equation G.1). The pathloss α is a random variable that takes on values α_1 and α_2 with probability $e^{-\beta_k v}$ and $1-e^{-\beta_k v}$ respectively (note that we have dropped the $\|.\|$ of v for easier notation). Then the process $\Theta = \cup_{k=1}^{K}\Theta_k$ is a non-homogenous PPP with density $\lambda(v) = \sum_{k=1}^{K}\lambda_k(v)$. In the following we compute the intensity and intensity measure of Θ_k for $k \in [1:K]$. By using the Mapping Theorem [60, Thm. 2.34] the intensity measure and the intensity of each tier k, $k \in [1:K]$, are given by

$$\begin{aligned}\Lambda_k([0,r]) &= \int_0^{(rP_k)^{\frac{\delta_1}{2}}} 2\pi\lambda_k v e^{-\beta_k v}dv + \int_0^{(rP_k)^{\frac{\delta_2}{2}}} 2\pi\lambda_k v(1-e^{-\beta_k v})dv \\ &= \frac{2\pi\lambda_k}{\beta_k^2}\left(1-e^{-\beta_k(rP_k)^{\frac{\delta_1}{2}}}(1+\beta_k(rP_k)^{\frac{\delta_1}{2}})\right) + \pi\lambda_k(rP_k)^{\frac{\delta_2}{2}} \\ &\quad - \frac{2\pi\lambda_k}{\beta_k^2}\left(1-e^{-\beta_k(rP_k)^{\frac{\delta_2}{2}}}(1+\beta_k(rP_k)^{\frac{\delta_2}{2}})\right) \qquad \text{(H.1)}\\ \lambda_k(v) &= \frac{d\Lambda_k([0,v])}{dv} = A_k v^{\delta_1-1}e^{-a_k v^{\frac{\delta_1}{2}}} + B_k v^{\delta_2-1}(1-e^{-b_k v^{\frac{\delta_2}{2}}}) \qquad \text{(H.2)}\end{aligned}$$

with $\delta_1 = \frac{2}{\alpha_1}$, $\delta_2 = \frac{2}{\alpha_2}$, $A_k = \pi\lambda_k\delta_1 P_k^{\delta_1}$, $a_k = \beta_k P_k^{\frac{\delta_1}{2}}$, $b_k = \beta_k P_k^{\frac{\delta_2}{2}}$ and$B_k = \pi\lambda_k\delta_2 P_k^{\delta_2}$.

Appendix H (Continued)

The process $\Theta = \cup_{k=1}^{K} \Theta_k$ has the following intensity measure and intensity

$$\Lambda(\nu) = \sum_{k=1}^{K} \Lambda_k(\nu) \tag{H.3}$$

$$\lambda(\nu) = \sum_{k=1}^{K} \lambda_k(\nu) = \sum_{k=1}^{K} A_k \nu^{\delta_1 - 1} e^{-a_k \nu^{\frac{\delta_1}{2}}} + B_k \nu^{\delta_2 - 1}(1 - e^{-b_k \nu^{\frac{\delta_2}{2}}}). \tag{H.4}$$

Appendix I

PROOF OF THEOREM 4.3.1 AND THEOREM 4.3.2

I.1 Proof of Theorem 4.3.1

Let us re-consider a different distribution on the direct links - while keeping the same Rayleigh fading assumption on the interfering links - in particular let us consider that the fading is Nakagami with shape parameter m and scale parameter $\theta = 1$. In this case we will derive the distribution of an upper bound on the desired signal in particular the distribution of the following

$$S = \Big| \sum_{i \leq n} \gamma_i'^{-1/2} h_i \Big|^2 \leq \sum_{i \leq n} \gamma_i'^{-1} \sum_{i \leq n} |h_i|^2 = S^{UP}$$

But we have that

$$\mathbb{P}\left(S^{UP} \geq T(I + \frac{\sigma^2}{N_t}) \right) = \mathbb{P}\left(\sum_{i \leq n} |h_i|^2 \geq \frac{T(I + \frac{\sigma^2}{N_t})}{\sum_i \gamma_i^{-1}} \right)] = \mathbb{P}\left(S_{UP} \geq T'(I + \frac{\sigma^2}{N_t}) \right) \Big|_{T' = \frac{T}{\sum_i \gamma_i^{-1}}, S_{UP} = \sum_i |h_i|^2}$$

We have that

$$h_i \sim \text{Nakagami}(m, 1)$$

$$|h_i|^2 \sim \text{Gamma}(1, 1/m)$$

$$\sum_i |h_i|^2 \sim \text{Gamma}(nm, 1/m)$$

Appendix I (Continued)

Then the Laplace transform is

$$\mathcal{L}_{S^{UP}}(s) = \frac{1}{(1+s/m)^{nm}}.$$

We then have from [59, Eq. 2.11] that the coverage probability is

$$\begin{aligned}
\mathbb{P}(\text{SINR} > T) &= \mathbb{P}\left(S > T'(I + \frac{\sigma^2}{N_t})\right) \\
&= \int_{0<\gamma_1'<\cdots<\gamma_n'<+\infty} f_{\Gamma'}(\gamma')\mathbb{P}\left(S > T'(I + \frac{\sigma^2}{N_t})\right) d\gamma' \\
&= \int_{0<\gamma_1'<\cdots<\gamma_n'<+\infty} f_{\Gamma'}(\gamma') \int_{-\infty}^{\infty} \mathcal{L}_I(2j\pi T's)\mathcal{L}_N(2j\pi T's)\frac{\mathcal{L}_{S^{UP}}(-2j\pi s) - 1}{2j\pi s} ds \; d\gamma'
\end{aligned} \tag{I.1}$$

where the joint distribution of γ' is given by

$$f_{\Gamma'}(\gamma') = \prod_{i=1}^{n} \lambda(\gamma_i')e^{-\Lambda(\gamma_n')} \tag{I.2}$$

Next we have from (Equation G.11) the Laplace transform of the interference $\mathcal{L}_I(s)$ with an intensity $\lambda(\nu)$ given by (Equation H.4) while the Laplace transform of the noise is given by

$$\mathcal{L}_N(s) = \mathbb{E}[e^{-s\frac{\sigma^2}{N_t}}] = e^{-s\frac{\sigma^2}{N_t}}. \tag{I.3}$$

This concludes the proof.

Appendix I (Continued)

I.2 Proof of Theorem 4.3.2

The coverage probability in the absence of interference is given by (Equation 4.23) with $\mathcal{L}_I(s) = 1$ and is

$$\mathbb{P}(\text{SNR} > T) = \int\limits_{0<\gamma'_1<\cdots<\gamma'_n<+\infty} f_{\Gamma'}(\gamma') \int_{-\infty}^{\infty} \frac{\mathcal{L}_S^{UP}(-2j\pi s) - 1}{2j\pi s} \mathcal{L}_N(2j\pi T' s) ds \; d\gamma' \tag{I.4a}$$

$$= \int\limits_{0<\gamma'_1<\cdots<\gamma'_n<+\infty} f_{\Gamma'}(\gamma') \int_{-\infty}^{\infty} \frac{\frac{1}{(1-\frac{2j\pi s}{m})^{nm}} - 1}{2j\pi s} e^{-2j\pi s \frac{T'\sigma^2}{N_t}} ds \; d\gamma' \tag{I.4b}$$

$$= \int\limits_{0<\gamma'_1<\cdots<\gamma'_n<+\infty} f_{\Gamma'}(\gamma') \underbrace{\int_{-\infty}^{\infty} f(s)\, ds}_{Q} \; d\gamma'. \tag{I.4c}$$

In the following we seek to solve Q. Note that a pole of order nm exists in the integrand thus the integral Q can be solved using contour integration and is as follows

$$Q = \text{Res}_{z^* = \frac{m}{2\pi j}}[f(z)] = \lim_{z \to z^*} \frac{1}{(nm-1)!} \left(\frac{d}{dz}\right)^{nm-1} (z - z^*)^{nm} f(z). \tag{I.5}$$

The function $f(z)$ can be re-written in the following form

$$f(z) = \frac{g(z)}{(z - z^*)^{nm}} \tag{I.6}$$

Appendix I (Continued)

with

$$f(z) = \frac{\frac{1}{(1-\frac{2\pi jz}{m})^{nm}} - 1}{2\pi jz} e^{-2\pi jz\frac{T'\sigma^2}{N_t}} = \frac{(-1)^{nm}\frac{1-(1-\frac{2\pi jz}{m})^{nm}}{(2\pi j)^{nm}(2\pi jz)} e^{-2\pi jz\frac{T'\sigma^2}{N_t}}}{\left(z-\frac{m}{2\pi j}\right)^{nm}}, \tag{I.7a}$$

$$g(z) = (-1)^{nm}\frac{1-(1-2\pi jz)^{nm}}{(2\pi j)^{nm}(2\pi jz)} e^{-2\pi jz\frac{T'\sigma^2}{N_t}}. \tag{I.7b}$$

Then after substituting the functions in (Equation I.5), we can express the integral I as

$$Q = \frac{g^{(nm-1)}(z^*)}{(nm-1)!}. \tag{I.8}$$

This concludes the proof.

CITED LITERATURE

1. D. Maamari, N. Devroye, and D. Tuninetti, "The sum-capacity of the ergodic fading Gaussian cognitive interference channel," *IEEE Trans. Wireless Comm.*, vol. 14, no. 2, pp. 809–820, Sep. 2014.

2. D. Maamari, D. Tuninetti, and N. Devroye, "Approximate sum-capacity of K-user cognitive interference channels with cumulative message sharing," *IEEE J. Select. Areas Commun.*, vol. 31, no. 11, pp. 657–667, March 2013.

3. ——, "The capacity of the ergodic MISO channel with per-antenna power constraint and an application to the fading cognitive interference channel," in *Proc. IEEE Int. Symp. Inf. Theory*, Honolulu, Hawaii USA, June 2014.

4. ——, "The sum-capacity of different K-user cognitive interference channels in strong interference," in *Proc. IEEE Inf. Theory Workshop*, Sevilla, Sep. 2013.

5. ——, "On the K-user cognitive interference channel with cumulative message sharing sum-capacity," in *Proc. IEEE Int. Symp. Inf. Theory*, 2013, pp. 2034–2038.

6. ——, "The sum-capacity of the linear deterministic three-user cognitive interference channel," in *Proc. IEEE Int. Symp. Inf. Theory*, 2012, pp. 2107–2111.

7. A. Goldsmith, S. Jafar, I. Maric, and S. Srinivasa, "Breaking spectrum gridlock with cognitive radios: An information theoretic perspective," *Proceedings of the IEEE*, vol. 97, no. 5, pp. 894–914, 2009.

8. N. Devroye, P. Mitran, and V. Tarokh, "Achievable rates in cognitive radio channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 1813–1827, May 2006.

9. A. Jovicic and P. Viswanath, "Cognitive Radio: an information-theoretic perspective," *IEEE Trans. Inf. Theory*, vol. 55, no. 9, pp. 3945 – 3958, Sep. 2009.

10. W. Wu, S. Vishwanath, and A. Aripostathis, "Capacity of a class of cognitive radio channels: interference channels with degraded message sets," *IEEE Trans. Inf. Theory*, vol. 53, no. 11, pp. 4391–4399, Nov. 2007.

CITED LITERATURE (Continued)

11. I. Maric, R. D. Yates, and G. Kramer, "The strong interference channel with unidirectional cooperation," in *Proc. Workshop on Info. Theory and Applications*, La Jolla, Feb. 2006.

12. S. Rini, D. Tuninetti, and N. Devroye, "New inner and outer bounds for the memoryless cognitive interference channel and some new capacity results," *IEEE Trans. Inf. Theory*, vol. 57, no. 7, pp. 4087–4109, Jul. 2011.

13. ——, "Inner and outer bounds for the Gaussian cognitive interference channel and new capacity results ," *IEEE Trans. Inf. Theory*, vol. 58, no. 2, pp. 820–848, Jan. 2012.

14. K. G. Nagananda, P. Mohapatra, C. R. Murthy, and S. Kishore, "Multiuser cognitive radio networks: An information theoretic perspective," http://arxiv.org/abs/1102.4126.

15. K. Nagananda and C. Murthy, "Information theoretic results for three-user cognitive channels," in *Proc. IEEE Global Telecommun. Conf.*, 2009, pp. 1–6.

16. M. Mirmohseni, B. Akhbari, and M. R. Aref, "Three-user cognitive interference channel: Capacity region with strong interference," *IET, Communications*, vol. 6, pp. 2099–2107, 2013.

17. M. G. Kang and W. Choi, "The capacity of a three user interference channel with a cognitive transmitter in strong interference," in *Proc. IEEE Int. Symp. Inf. Theory*, 2012.

18. I. Maric, R. D. Yates, and G. Kramer, "Capacity of interference channels with partial transmitter cooperation," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3536–3548, Oct. 2007.

19. S. Akoum, O. El Ayach, and R. W. Heath, "Coverage and capacity in mmWave cellular system," in *Proc. Asilomar Conf. Signals, Systems and Computers*, Pacific Grove, CA USA, Nov. 2012, pp. 688–692.

20. T. Bai and R. W. Heath, "Coverage in dense millimeter wave cellular networks," in *Proc. Asilomar Conf. Signals, Systems and Computers*, Pacific Grove, CA USA, Nov. 2013, pp. 2062–2066.

21. ——, "Coverage and rate analysis for millimeter wave cellular networks," *IEEE Trans. Wireless Comm.*, vol. PP, no. 99, p. 1, Oct. 2014.

CITED LITERATURE (Continued)

22. G. Nigam, P. Minero, and M. Haenggi, "Coordinated multipoint in heterogeneous networks: A stochastic geometry approach," in *Proc. IEEE Global Telecommun. Conf.*, Atlanta, Dec. 2013.

23. ——, "Coordinated multipoint joint transmission in heterogeneous networks," *IEEE Trans. Commun.*, 2014.

24. R. A. Tannious and A. Nosratinia, "Cognitive radio protocols based on exploiting Hybrid ARQ retransmissions," *IEEE Trans. Wireless Comm.*, vol. 9, no. 9, pp. 2833 – 2841, 2010.

25. V. Annapureddy, A. El Gamal, and V. Veeravalli, "Degrees of freedom of interference channels with CoMP transmission and reception ," *IEEE Trans. Inf. Theory*, vol. 58, no. 9, pp. 5740–5760, Sep. 2012.

26. L. Sankar, X. Shang, and V. Poor, "Ergodic fading interference channels: sum-capacity and separability," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 2605 – 2626, May 2011.

27. D. Tuninetti, "Gaussian fading interference channels: Power control," in *Proc. Asilomar Conf. Signals, Systems and Computers*, Monterrey, CA, Oct. 2008, pp. 701 – 706.

28. V. R. Cadambe and S. A. Jafar, "Parallel Gaussian interference channels are not always separable," *IEEE Trans. Inf. Theory*, vol. 55, no. 9, pp. 3983 – 3990, Sep. 2009.

29. D. Tse and S. V. Hanly, "Multiaccess fading channels. I. Polymatroid structure, optimal resource allocation and throughput capacities," *IEEE Trans. Inf. Theory*, vol. 44, no. 7, pp. 2796 – 2815, Nov. 1998.

30. L. Li and A. Goldsmith, "Capacity and optimal resource allocation for fading broadcast channels-Part I: Ergodic capacity," *IEEE Trans. Inf. Theory*, vol. 47, no. 3, pp. 1083–1102, Mar. 2001.

31. A. Ghasemi and E. S. Sousa, "Fundamental limits of spectrum-sharing in fading environments," *IEEE Trans. Wireless Comm.*, vol. 6, no. 2, pp. 649 – 658, Feb. 2007.

32. X. Kang, Y.-C. Liang, A. Nallanathan, H. K. Garg, and R. Zhang, "Optimal power allocation for fading channels in cognitive radio networks: Ergodic capacity and outage capacity," *IEEE Trans. Wireless Comm.*, vol. 8, no. 2, pp. 940–950, Feb. 2009.

CITED LITERATURE (Continued)

33. R. Zhang, S. Cui, and Y.-C. Liang, "On ergodic sum capacity of fading cognitive multiple-access and broadcast channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 11, pp. 5161 – 5178, Nov. 2009.

34. M. Vu, " MISO capacity with per-antenna power constraint," *IEEE Trans. Commun.*, vol. 59, no. 5, pp. 1268–1274, May 2011.

35. T. Cover and J. Thomas, *Elements of Information Theory: Second Edition.* Wiley, 2006.

36. M. Costa, "Writing on dirty paper," *IEEE Trans. Inf. Theory*, vol. IT-29, pp. 439–441, May 1983.

37. R. Knopp and P. A. Humblet, "Information capacity and power control in single-cell multiuser communications," in *Proc. IEEE Int. Conf. Commun.*, vol. 1, Seattle, USA, June 1995, pp. 331–335.

38. K. Nagananda, C. Murthy, and S. Kishore, "Achievable rates in three-user interference channels with one cognitive transmitter," in *Proc. IEEE Int. Sig. Proc. Commun. (SPCOM)*, 2010, pp. 1–5.

39. M. Mirmohseni, B. Akhbari, and M. R. Aref, "Capacity bounds for the three-user cognitive Z-interference channel," in *the 12th Canadian Workshop on Inf. Theory*, Kelowna, British Columbia, Canada, May 2011, pp. 34–37.

40. S. Rini, D. Tuninetti, N. Devroye, and A. Goldsmith, "On the capacity of the interference channel with a cognitive relay," *IEEE Trans. Inf. Theory*, vol. 60, no. 4, pp. 2148–2179, Sept. 2014.

41. S. Sridharan, S. Vishwanath, S. A. Jafar, and S. Shamai, "On the capacity of cognitive relay assisted Gaussian interference channel," in *Proc. IEEE Int. Symp. Inf. Theory*, 2008, pp. 549–553.

42. T. Cover, "Broadcast channels," *IEEE Trans. Inf. Theory*, vol. IT-18, no. 1, pp. 2–14, Jan. 1972.

43. A. Avestimehr, S. Diggavi, and D. Tse, "A deterministic approach to wireless relay networks," *Proc. Allerton Conf. Commun., Control and Comp.*, Sep. 2007.

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

CITED LITERATURE (Continued)

45. C. Suh and D. Tse, "Feedback capacity of the Gaussian interference channel to within 2 Bits," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 2667–2685, May 2011.

46. V. Prabhakaran and P. Viswanath, "Interference channels with source cooperation," *IEEE Trans. Inf. Theory*, vol. 57, no. 1, pp. 156–186, 2011.

47. K. Marton, "A coding theorem for the discrete memoryless broadcast channel," *IEEE Trans. Inf. Theory*, vol. 25, no. 5, pp. 306–311, May 1979.

48. S. A. Jafar and S. Vishwanath, "Generalized degrees of freedom of the symmetric Gaussian K user interference channel," *IEEE Trans. Inf. Theory*, vol. 56, no. 7, pp. 3297 – 3303, 2010.

49. A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge University Press, 2011.

50. A. Dytso, S. Rini, N. Devroye, and D. Tuninetti, "On the capacity region of the two-user interference channel with a cognitive relay," *IEEE Trans. Wireless Comm.*, vol. 13, no. 12, pp. 6824–6838, Dec. 2014.

51. J. G. Andrews, S. Buzzi, W. Choi, S. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What will 5G be?" *IEEE J. Select. Areas Commun.*, vol. 32, pp. 1065–1082, Jun. 2014.

52. S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter-wave cellular wireless networks: Potentials and Challenges," *Proceedings of the IEEE*, vol. 102, no. 2, pp. 366–385, Mar. 2014.

53. O. El Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath., "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Trans. Wireless Comm.*, vol. 13, no. 3, pp. 1499–15 113, Jan. 2014.

54. T. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, "Millimeter wave communications for 5G cellular: It will work!" *IEEE Access*, vol. 1, no. 1, pp. 335–349, Aug. 2013.

55. L. Daweon, S. Hanbyul, B. Clerckx, E. Hardouin, D. Mazzarese, S. Nagata, and K. Sayana, "Coordinated multipoint transmission and reception in LTE-advanced: Deployment scenarios and operational challenges," *IEEE Communications Magazine*, vol. 50, no. 2, pp. 148–155, 2012.

CITED LITERATURE (Continued)

56. J. G. Andrews, F. Baccelli, and R. K. Ganti, "A tractable approach to coverage and rate in cellular networks," *IEEE Trans. Commun.*, vol. 59, no. 11, pp. 3122–3134, Jan. 2011.

57. M. R. Akdeniz, Y. Liu, M. K. Samimi, S. Sun, S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter wave channel modeling and cellular capacity evaluation," *IEEE J. Select. Areas Commun.*, vol. 32, no. 6, pp. 1164–1179, 2014.

58. D. Tse and P. Viswanath, *Fundamentals of Wireless Communication.* Cambridge University Press, 2005.

59. F. Baccelli, B. Blaszczyszyn, and P. Muhlethaler, "Stochastic analysis of spatial and opportunistic aloha," *IEEE J. Select. Areas Commun.*, vol. 27, no. 7, pp. 1105–1119, Sep. 2009.

60. M. Haenggi, *Stochastic Geometry for Wireless Networks.* Cambridge University Press, 2013.

61. D. Moltchanov, "Survey paper: Distance distributions in random networks," *Ad Hoc Networks*, vol. 10, no. 6, pp. 1146–1166, Aug. 2012.

VITA

NAME: Diana Maamari

EDUCATION: B.S., Electrical Engineering, University of Balamand, Lebanon, 2009

M.S., Electrical Engineering, University of Illinois at Chicago, Chicago, Illinois, 2011

Ph.D., Electrical and Computer Engineering, University of Illinois at Chicago, Chicago, Illinois, 2015

EXPERIENCE: Teaching Assistant, University of Illinois at Chicago, Chicago; Probability and Stochastic Processes, Spring 2012.

Research Assistant, University of Illinois at Chicago, Chicago; Fall 2012 -Spring 2015.

PUBLICATIONS: D. Maamari, D. Tuninetti, and N. Devroye, "Approximate sum-capacity of K-user cognitive interference channels with cumulative message sharing," *IEEE J. Select. Areas Commun.*, vol. 31, no. 11, pp. 657-667, March 2013.

D. Maamari, N. Devroye, and D. Tuninetti,"The sum-capacity of the ergodic fading Gaussian cognitive interference channel," *IEEE Trans. On Wireless Comm.*, vol. 14, no.2, pp. 809-820, Sep.214.

D. Maamari, N. Devroye, and D. Tuninetti, "The capacity of the ergodic MISO channel with per-antenna power constraint and an application to the fading cognitive interference channel," in *Proc. IEEE Int. Symp. Inf. Theory*, Honolulu, Hawaii USA, June 2014.

VITA (Continued)

D. Maamari, D. Tuninetti, and N. Devroye, "The sum-capacity of different K-user cognitive interference channels in strong interference," in *Proc. IEEE Inf. Theory Workshop*, Sevilla, Sep. 2013.

D. Maamari, N. Devroye, and D. Tuninetti, "On the K-user cognitive interference channel with cumulative message sharing sum-capacity," in *Proc. IEEE Int. Symp. Inf. Theory*, 2013, pp. 2034-2038.

D. Maamari, D. Tuninetti, and N. Devroye, "The sum-capacity of the linear deterministic three-user cognitive interference channel," in *Proc. IEEE Int. Symp. Inf. Theory*, 2012, pp. 2107-2111.