Performance Limits of Microwave and Dual Microwave/Millimeter Wave Band Networks

by

Subhajit Majhi

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

Waterloo, Ontario, Canada, 2019

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner: Dr. Aria Nosratinia
Professor, Dept. of Electrical Engineering,
University of Texas, Dallas

Supervisor: Dr. Patrick Mitran
Professor, Dept. of Electrical & Computer Engineering,
University of Waterloo

Internal Members: Dr. Ravi R. Mazumdar
Professor and University Research Chair,
Dept. of Electrical & Computer Engineering,
University of Waterloo

Dr. Guang Gong
Professor and University Research Chair,
Dept. of Electrical & Computer Engineering,
University of Waterloo

Internal-External Member: Dr. Henry Wolkowicz
Professor, Dept. of Combinatorics & Optimization,
University of Waterloo

Author's Declaration

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

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

Abstract

Traditionally, wireless networks communicate over the conventional microwave band (sub-6 GHz) as it supports reliable communication over a large geographic area. The ever increasing demand for bandwidth to support the rising number of consumers and services, however, is fast depleting the available microwave spectrum. As such, complementing the microwave spectrum with additional bandwidth from the *millimeter-wave* (mm-wave) band has been envisioned as a promising solution to this problem.

Since transmissions in the mm-wave band are typically achieved with highly directional steerable antenna arrays to counter the severe path-loss in mm-wave frequencies, the resulting mm-wave links are typically rendered highly directional, which can often be modeled as directional *point-to-point* links. However, mm-wave transmissions are inherently *unreliable* compared to those in the microwave band. Hence, communicating *simultaneously* over both bands in an *integrated mm-wave/microwave dual-band* setup is emerging as a promising new technology. In this dual-band setting, high-rate data traffic can be carried by relatively unreliable high-bandwidth mm-wave links, while control signals and moderate-bandwidth traffic can be communicated over the relatively reliable microwave band.

In this thesis, we first study two dual-band *multi-user networks* that model two important aspects of wireless communication: *inter-user interference* and *relay-cooperation*. The broad goal of this study is to characterize information-theoretical performance limits of such networks, which can then be used to obtain insights on the optimal encoding/decoding strategy, effective resource allocation schemes, etc.

In the first part of this thesis, we study a two-transmitter two-receiver dual-band Gaussian interference channel (IC) operating over an integrated mm-wave/microwave dual-band. This channel models a setting where a pair of single-transmitter single-receiver links communicate simultaneously, and thus mutually interfere. Here, transmissions in the underlying microwave band are modeled as a two-user conventional Gaussian IC (GIC). In contrast, a transmitter in the mm-wave band is assumed to be capable of communicating to either the desired destination or the interfered destination via a point-to-point *direct-link* or a *cross-link*, respectively. The dual-band IC is first classified into 3 classes according to the interference level in the underlying microwave GIC, and then *sufficient channel conditions* are obtained under which the capacity region of the 3 classes are characterized. For cases in which the sufficient conditions do not necessarily hold, *approximate capacity* results are obtained that characterizes the capacity region to within 1/2 bit per channel use per user.

The performance of the dual-band IC is likely to be impacted significantly by the point-to-point nature and large bandwidth of the mm-wave links, and specifically by whether the mm-wave spectrum is used as direct-links or cross-links. Transmitting in either the direct-links only or the cross-links only is not optimal for all channel conditions, and there exists a non-trivial trade-off between the two modes. To understand the impact of this trade-off on the performance of the dual-band IC, we study the power allocation scheme over the mm-wave direct and cross-links that maximizes the sum-rate of the channel. The resulting power allocation strategy is characterized in closed form, which possesses rich properties and reveals useful insights into the trade-offs in such networks.

In the second part of this thesis, we study a fading Gaussian multiple-access relay channel (MARC) over an integrated mm-wave/microwave dual-band, where two sources communicate to a destination with the help of a relay. In the dual-band MARC, transmission in the underlying microwave band is modeled as a conventional Gaussian MARC. However, similar to that in the dual-band IC, a mm-wave transmitter in this channel is modeled as being able to communicate to either the destination or the relay by creating a *direct-link* or a *relay-link*, respectively. For dual-band MARC, we characterize an achievable region and a set of rate upper bounds, and then obtain sufficient channel conditions under which its capacity region is characterized.

Similar to the dual-band IC, the performance of the dual-band MARC will likely be significantly affected by whether the mm-wave band is used as direct-links or relay-links, and a non-trivial trade-off between the two modes exists in this case as well. To understand this trade-off, we study the transmission power allocation scheme over the mm-wave direct and relay-links that maximizes the sum-rate of the dual-band MARC. The resulting power allocation scheme, characterized in closed form, is observed to have rich structural properties, which reveal insights into the trade-offs in relay cooperation in dual-band networks.

While dual-band communication is a promising technology, currently the bulk of the connectivity is still supported by the microwave band. However, the problem of interference mitigation for conventional microwave bands is still open even for the basic case of a two-user IC. Motivated by this, in the third part of the thesis, we study the performance limits of the multiple-access interference channel (MAIC) which models the interference during cellular uplink over the conventional single band. Focusing on the weak interference case, which provides a more realistic model of the inter-cell interference, we characterize an achievable strategy and 3 novel upper bounds on the sum-rate in the partially symmetric case, thereby providing improved sum-rate upper and lower bounds in these cases.

Acknowledgements

It has been a great privilege to pursue my doctoral studies in the University of Waterloo, and I would like to thank all the kind people who made my stay at Waterloo a rewarding experience. Acknowledging all of them for their help and support would be an insurmountable task, and hence I express my sincere gratitude to at least a few of them below.

First, I would like to express my sincere gratitude to my supervisor, Prof. Patrick Mitran, for his thoughtful guidance, constant support, immense patience, and significant contribution throughout this work. His insightful observations, helpful comments and undivided attention during countless meetings is deeply appreciated. His expert technical guidance have inculcated in me qualities which I will treasure throughout my life.

I am grateful to all members of my PhD defense committee, Prof. Guang Gong, Prof. Ravi R. Mazumdar and Prof. Henry Wolkowicz, for offering valuable suggestions during my comprehensive exam. In particular, I am thankful to Prof. Ravi R. Mazumdar for enhancing my knowledge of Queuing Theory, and Prof. Henry Wolkowicz for enriching my understanding of optimization theory. Furthermore, I would like to thank Prof. Aria Nosratinia for accepting to serve as the external examiner for my thesis defense.

I would like to render special thanks to my colleague Mr. Meysam Shahrbaf Motlagh for providing valuable suggestions and encouragement that go beyond academics, and for the fruitful collaboration on the Massive MIMO project. My heartfelt thanks goes to all my friends (in no particular order), Mohammad, Mostafa, Thiru, Shankha, Rajib, Arpan, Syamantak and Debarshi, who made my stay in Canada pleasant and memorable. I would like to offer my heartiest regards to my house-mate, Dr. Rajdip Nayek, for his help throughout and beyond my thesis work, for sharing moments of delights and grieves together, and for all the stimulating discussions on larger themes of life such as spirituality and morals that has in no doubt enriched my life.

Finally, I would like to express my deepest regards to my parents and my brother for their love, support and sacrifices. They have always trusted and encouraged me in my endeavors, and without my parents' unwavering love and nuggets of wisdom, and my brother's enduring support from across the world, this journey would not have been possible.

Dedication

To

my parents,

Sandhya Majhi and Amar Majhi,

and my brother,

Ranadip Majhi.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

GIC	Gaussian Interference Channel
MAC	Multiple Access Channel
HK	Han-Kobayashi
TIN	Treat Interference as Noise
MAIC	Multiple Access Interference Channel
AWGN	Additive White Gaussian Noise
DCLIC	Direct-and-Cross-Link Interference Channel
CLIC	Cross-Link Interference Channel
Z-CLIC	Z-Cross-Link Interference Channel
LGR	Link Gain Regime
WF	Waterfilling
BMF	Bandwidth Mismatch Factor
MARC	Multiple Access Relay Channel
DR-MARC	Destination-and-Relay-Linked Multiple Access Relay Channel
R-MARC	Relay-Linked Multiple Access Relay Channel

List of Notations

$a := b$	b is the definition of a
X^n	A vector $(X_1, X_2, \ldots, X_n)$ of length n
$\mathbf{0}$	The zero vector
$\boldsymbol{v} \succeq \mathbf{0}$	A vector $\boldsymbol{v} := (v_1, \ldots, v_m)$, in which each element is nonnegative
$\mathbb{R}$	The set of real numbers
$\mathbb{R}_+$	The set of non-negative real numbers
$\mathbb{C}$	The set of complex numbers
$\lvert Q \rvert$	The cardinality of a set Q
$\mathsf{Re}(c)$	The real part of a complex number c
$\mathcal{X}_1 \times \mathcal{X}_2$	The Cartesian product of two sets $\mathcal{X}_1$ and $\mathcal{X}_2$
$\emptyset$	The empty set
$[1 : N]$	The set $\{1, 2, \ldots, N\}$
$\Pr[A]$	The probability of an event A
$X \sim p_X(x)$	A discrete random variable X, distributed according to probability mass function (pmf) $p_X(x)$, or a continuous random variable X, distributed according to probability density function (pdf) $p_X(x)$
$p_{Y\lvert X}(y\lvert x)$	The conditional pmf/pdf of random variable Y given X
$X \sim \mathcal{N}(\mu, \sigma^2)$	A Gaussian random variable X with mean μ and variance σ^2
$Z \sim \mathcal{CN}(\mu_c, \sigma_c^2)$	A circularly symmetric complex Gaussian (CSCG) random variable Z with mean μ_c and variance σ_c^2
$\Theta \sim \mathcal{U}[0, 2\pi)$	A random variable Θ that is uniformly distributed in $[0, 2\pi)$
$\mathbb{E}[X]$	The statistical expectation of a random variable X
$H(W)$	The entropy of a discrete random variable W
$h(V)$	The differential entropy of a continuous random variable V
$I(X;Y)$	The mutual information between random variables X and Y
$I(X;Y\lvert Z)$	The conditional mutual information between X and Y given Z
$X \multimap Y \multimap Z$	The Markov chain for random variables (X, Y, Z) such that $X \multimap Y \multimap Z \iff p(x\lvert y, z) = p(x\lvert y)$
$X \perp\!\!\!\perp Y$	Random variables X and Y are independent
i.i.d.	Independently and identically distributed
$\log(x)$	$\log_2(x)$
$\mathsf{C}(x)$	$\frac{1}{2}\log_2(1 + x)$
$\lfloor i \rfloor$	The largest integer no larger than i

Chapter 1

Introduction

In the last decade, advances in wireless communication technology have played a pivotal role in improving and expanding public connectivity around the globe [1]. In particular, the 4th generation of cellular networks (i.e., 4G) have been responsible for dramatically improving user quality of service as measured by data-rate, the range of connectivity, reliability [2, 3], etc. In addition to improving user connectivity in general, the higher speeds available in 4G have also facilitated the introduction and growth of an array of Internet-based applications such as high definition (HD) video streaming, interactive online gaming, virtual reality experience [4,5], etc., resulting in explosive growth of mobile traffic. For instance, the average smartphone is expected to generate 11 GB of traffic per month by 2022, more than a 4.5-fold increase over the 2017 average of 2 GB per month [6].

Since the available bandwidth for wireless communication is limited, this bandwidth is typically shared among many devices that transmit simultaneously over this shared medium. In cellular communication settings, this leads to the paradigm of *multi-user communication* where multiple users communicate over a shared frequency band simultaneously. One of the major bottlenecks to the performance of such multi-user networks is *inter-user interference*: due to the nature of the shared medium, transmissions from a transmitter are typically received by all neighboring receivers, thereby deteriorating the overall performance of the network. The optimal strategy for mitigating interference is unknown in general, and highly non-trivial even for basic cases [7]. For example, simply increasing the transmission powers of all devices will not reduce the effect of interference as the power of interference then also increases proportionally. Therefore, interference management schemes have been at the forefront of research [8–11].

1.1 Multi-User Communication in Conventional Networks

1.1.1 Interference in Conventional Networks

In incumbent cellular communication settings, users communicate to the base station (BS) over the shared microwave band (i.e., sub-6 GHz band). Due to the nature of the microwave band, when a user in a cell communicates to its designated BS, it inadvertently interferes with the reception at the BSs in neighboring cells. This leads to performance loss in general if interference is not mitigated properly.

The classic two-user interference channel (IC) is a useful model of the interference-limited cellular communication between two single-transmitter and single-receiver links. Here, each user transmitting to its designated BS in the uplink phase is modeled as a point-to-point link, and the IC then models the interference between two such point-to-point links operating in two neighboring cells that interfere mutually.

Although the IC considers a basic two-cell setup, the capacity region or the optimal strategies for the channel are still unknown in general. In fact, the optimal transmission strategy, which has been identified for a few specific cases, depends on the relative interference power received at each BS, or equivalently on the relative *strength* of the interfering links (denoted as *cross channels*) as compared to the desired links. More precisely, depending on the strength of the cross channels, the Gaussian IC can be classified into three cases: the GIC with *strong* interference, where both cross channels are stronger than the direct channels, the GIC with *weak* interference, where both cross channels are weaker than the direct channels, and the GIC with *mixed* interference, where one of the cross channels is stronger, while the other is weaker. In the Gaussian IC with strong interference, the capacity region is known, and in this regime the optimal strategy is to decode the messages of the intended users as well as the interfering users at each BS.

In many practical cellular settings, such as those where the power of transmissions from the desired users in the current cell is larger than the power of interference from users located far away in the neighboring cell, interference is appropriately modeled by the weak Gaussian IC. Therefore, finding optimal strategies for this case may reveal important practical insights. However, in contrast to strong interference, in the weak interference regime the strategy of decoding interference is suboptimal, and in fact, treating the interfering

signal as noise provides optimal sum-rate when the cross channels are sufficiently weak. The Han-Kobayashi (HK) strategy [7] for the two-user IC allows for partially decoding the interference, and the resulting region was shown to achieve within a 1/2 bit/channel use gap of the capacity region of the Gaussian IC with real channel coefficients. Hence, even in the basic set-up of the two-user IC, the problem of interference mitigation is yet to be solved completely. Moreover, advanced encoding-decoding strategies that achieve the capacity of the IC approximately (i.e., within a constant gap) such as the HK scheme are quite complex.

Moreover, in the IC only one transmitter-receiver pair communicates in each cell, whereas in typical cellular settings, multiple transmitter-receiver pairs communicate simultaneously. For example, during cellular uplink, multiple users communicate to the BS simultaneously, thus forming a multiple-access channel (MAC). During the uplink phase, when two such MACs operating in two neighboring cells mutually interfere, the resulting channel is known as the multiple-access interference channel (MAIC) [12]. The MAIC thus models a more realistic setting of cellular communication, however only a few results on its performance are known, especially in the weak interference regime. Lack of optimal strategies for the MAIC and similar interference-limited channels, necessitates further study aiming at the development of new techniques for inter-cell interference mitigation.

1.1.2 Cooperation in Conventional Networks

In addition to interference, *multipath fading* in wireless communication also leads to performance degradation in cellular networks [13–15]. Relay cooperation [16, 17] serves as an effective technology to counter the ill-effects of multipath fading. The basic setup of relay cooperation is modeled by the relay channel introduced in [16], where the transmission from a single source (e.g., user) to a single destination (e.g., base station) is aided by a single relay. More specifically, the relay receives the transmission from the user, and then reinforces the user signal at the destination by transmitting either a compressed/quantized or a re-encoded version of the same signal. Although the relay channel in [16] models a basic setup, its capacity region is unknown in general, whereas the decode-and-forward relaying scheme was shown to achieve within 1/2 bit of the capacity region [17].

While the relay channel models a cellular setting where the relay cooperates with only a single user, typically multiple users communicate to the base station in the uplink. The multiple-access relay channel (MARC) models such a realistic setup where a single relay

helps the transmission of multiple users [18, 19]. However, similar to the relay channel, the capacity region of the MARC is also unknown in general. Hence, it is crucial to investigate optimal strategies for the MARC, which may lead to useful relaying techniques in practical cellular networks.

1.2 5G and the Millimeter Wave Communication

While 4G [20] ushered in an era of vastly improved mobile communication compared to 3G, as the number of subscribers and demand for high data-rate applications has rapidly increased in recent years, due in part to the limited amount of bandwidth available in the microwave band, 4G is expected to become unable to deliver the required quality of service in the future [21]. In fact, with the introduction of new technologies such as the Internet-of-things (IoT) and machine-to-machine communication [22], which are expected to increase the number of connected devices exponentially in the future, 4G is expected to suffer from a serious problem of spectrum-scarcity.

To address this imminent problem, the fifth generation of cellular networks (i.e., 5G) was recently proposed that aims to improve user quality-of-experience, and meet market demands. While the broad goals of 5G include increasing the cell capacity and user data rates dramatically, as well as reducing latency and energy expenditure [22], among others, scarcity of spectrum in the microwave band poses a serious impediment to realizing such goals. To address this problem, many new disruptive technologies are being investigated.

One of the most promising solution to the problem of spectrum crunch in the sub-6 GHz microwave band is to integrate and provide cellular access over frequencies *above* 6 GHz, more specifically in the spectrum from 27 GHz to 80 GHz. While the spectrum from the 30 GHz to the 300 GHz band is formally referred to as the millimeter wave (mm-wave) band, with a slight abuse of notation, the 27-80 GHz band is also referred to as the mm-wave band, where most current research effort on 5G is concentrated [22, 23].

1.2.1 Transmission in the Millimeter Wave Band

Omnidirectional transmission in the mm-wave band, similar to those in the conventional microwave band, suffers from heavy attenuation due to its smaller wavelength and increased atmospheric and molecular absorption at such higher carrier frequencies [24, 25].

Such propagation losses can be overcome by beamforming with co-phased directional antennas arrays [26] deployed at the transmitters and the receivers, which typically produce unidirectional mm-wave links. In addition, high absorption in this band makes scattering sparse [26, 27], thus rendering the mm-wave links point-to-point. Hence, such mm-wave links are often modeled as point-to-point additive white Gaussian noise (AWGN) links in a multitude of studies ranging from studies on the network layer [28–30, 30–36] to those on the physical layer [37–40].

While conventional microwave transmission is able to diffract around large object such as buildings, mm-wave links fail to do so owing to their highly directional nature and relatively smaller wavelength. As a result, mm-wave transmissions suffer from blockage due to urban features [29, 41–43], and even human movements [44, 45], thus leading to intermittent connectivity. Moreover, measurement campaigns on mm-wave cellular networks (e.g., in the 28 GHz band in Brooklyn and Manhattan [41, 46], in the 38 GHz band in Austin [47–49], in the 28 GHz and 73 GHz bands in New York City [50], etc.) reveal that such networks provide connectivity within only $\sim$ 200 m distance from the base station, and a significant number of users beyond this range are typically in outage due to blocking.

1.2.2 The Integrated Millimeter-Wave/Microwave Dual-Band

While the mm-wave band offers plenty of bandwidth resources, transmissions in this band suffer from heavy attenuation, absorption and blocking, thus leading to unreliable connections. However, instead of transmitting over the unreliable mm-wave links only, communicating simultaneously over multiple radio bands including the mm-wave band and the conventional microwave band (called sub-6 GHz band) can compensate for this shortcoming [51–54]. NTT-Docomo proposed a heterogeneous network architecture in 2014, called the Phantom cell [55], where the macro-cells and small-cells operate over over the conventional microwave band and the mm-wave band, respectively [56–58].

Transmissions in the conventional microwave band complement those in the mm-wave band. More specifically, in contrast to mm-wave links, which provide high speed intermittent connectivity to a relatively small region, conventional networks support medium data rates and provides reliable cellular coverage over a much larger geographical area. As such, simultaneously (jointly) transmitting over an *integrated mm-wave/microwave dual-band* can provide high speed, reliable communication. In this setting, high speed data would be communicated via the high-bandwidth and somewhat unreliable mm-wave links, while

transmissions in the conventional microwave band would provide more reliable coverage and carry medium rate traffic and control signals [26, 41, 59–61].

Such integrated mm-wave/microwave dual-band networks have been subject to much investigation recently. For example, studies from the network layer perspective focus on how to improve metrics such as the number of connected users and throughput, etc., by using the high-bandwidth, highly directional mm-wave links along with incumbent microwave networks [26, 55, 59–66]. Unlike the microwave band, where inter-user interference are typically managed by interference-avoidance schemes such as scheduling, the directional nature of mm-links are particularly useful in simplifying the design of interference management schemes [63, 64]. Moreover, recent works from the physical layer perspective [23, 66–70] also indicate favorable results in such a dual-band setting. In fact, commercial production of dual-band modems such as those in Qualcomm [71] and Intel [67] illustrate the potential of such networks. Hence, simultaneous transmission in the microwave and mm-wave dual-bands is an attractive technology for future cellular access.

In Chapter 2, a more detailed survey of integrated mm-wave/microwave dual-band networks is presented.

1.2.3 Interference in Dual-Band Networks

In integrated mm-wave/microwave dual-band networks, while the transmissions in the mm-wave band are point-to-point in nature, transmissions in the conventional microwave band cause interference to unwanted destinations receiving in the same band. More specifically, many proposed mm-wave network architectures rely on communicating via narrow and directed beams produced by phased antenna arrays [55, 57], thereby reducing interference to unwanted receivers. However, in a dual-band setting, the underlying microwave transmissions will still suffer from the same interference problem as in conventional networks. Therefore, it is necessary to characterize effective interference management strategies for such dual-band networks.

1.3 Motivation and Overview of Contribution

While mm-wave band communication is emerging as one of the most promising technologies in 5G, mm-wave communication needs to be complemented by transmissions in the conventional microwave band to offset the unreliable nature of mm-wave transmissions. Such

dual-band networks have been studied from the network and wireless communication perspectives, however few studies have been reported on information-theoretic limits of such networks, and especially for multi-user networks over such a dual-band. Such studies are crucial in identifying the limits of achievable rates, simplified encoding schemes, effective resource allocation and interference management strategies, etc., for practical dual-band networks. Hence, in this thesis, we study interference mitigation and cooperation strategies for multi-user networks operating over an integrated mm-wave/microwave dual-band.

Transmissions in the mm-wave band, produced by phased-antenna arrays, are highly directional, and are typically modeled as point-to-point AWGN links. Moreover, such point-to-point mm-wave beams can be digitally *steered* towards a specific receiver as needed. Therefore, a transmitter in the mm-wave band can be modeled as being able to transmit towards an *intended* receiver and cause negligible to no interference to unwanted receivers [30, 72], while being able to *switch* to another receiver as needed.

It is worth noting that phased-antenna arrays for mm-wave beamforming are already available in practice, e.g., the AWA-0142 by Anokiwave [73], a 256-element reconfigurable phased-antenna array operating in the 24.25-27.5 GHz band. Moreover, in contrast to a co-phased antenna array where all antenna elements are used together to beamform to *one* particular receiver, in *hybrid* antenna array systems, the set of available antenna elements can be reconfigured into *multiple* smaller and independent antenna arrays, each of which can then be digitally controlled to produce a separate beam to carry *independent signals* to *multiple receivers* [72–75]. Hence, if a hybrid antenna array is used, a mm-wave transmitter is able to communicate independent information to *multiple* receivers *in parallel*.

Compared to the mm-wave band, in the microwave band transmitters and receivers are assumed to use a single antenna. Moreover, due to the nature of this band, a receiver receives the transmissions from both desired and undesired transmitters.

In Chapter 2, we first present a detailed survey on the integrated mm-wave/microwave dual-band, and then introduce multi-user information theoretic terminologies and state of the art on the channels studied in the thesis.

In Chapter 3, we first study the performance of a two-transmitter two-receiver dual-band interference channel (IC), where two sources communicate to their respective destinations over the integrated mm-wave/microwave dual-band, while interfering mutually. We characterize the performance limits of this channel by first classifying the channel depending on whether the underlying Gaussian IC in the microwave band has strong, weak,

or mixed interference, and then deriving capacity or approximate capacity results for the different variations of the channel.

The throughput of the dual-band IC will likely be affected significantly by the point-to-point mm-wave links [60]. For example, the mm-wave links can be used to either convey fresh information to the intended receiver, or forward *interference* information to the non-designated receiver. Neither of the two modes is optimal for all channel conditions, and thus there exists a non-trivial trade-off between the two modes. Hence, to characterize the optimal performance of the dual-band IC over the parameters in the mm-wave band, in Chapter 3, we also study the power allocation scheme over the mm-wave band that maximizes the sum-rate of the channel. We derive the optimal power allocation strategy in closed form, which possesses rich properties and reveals useful insights into the trade-offs in such networks.

In Chapter 4, we study the performance of the two-user dual-band fading Gaussian multiple-access relay channel (MARC), where two sources communicate to a destination with the help of a relay over the integrated mm-wave/microwave dual-band. Relay aided communication already plays a key role in microwave networks, and it will likely play a vital role in dual-band networks as well, especially to offset impairments arising from blockage and fading in the mm-wave band [37,40,64,76]. For example, the dual-band MARC models uplink scenarios in 5G such as the fixed wireless access [77], which is expected to overhaul last mile connectivity by replacing wired connections to end users with directional mm-wave links.

For the conventional MARC that operates over the microwave band, if the MARC is subject to Rayleigh and phase fading, its capacity is known in closed form for the case where the source-relay channel gains are stronger than the source-destination channel gains [77]. For the dual-band MARC, we consider the case where these strong source-relay conditions *do not necessarily hold* in the underlying microwave band, and then derive channel conditions over both bands under which its capacity is characterized in closed form.

Similar to the dual-band IC, each source in the MARC can utilize the point-to-point mm-wave links to transmit to either the destination or the relay, depending on which the performance (e.g., sum-rate) of the dual-band MARC is affected significantly. In fact, there exist a non-trivial trade-off between transmitting solely to the relay and the destination. To understand this trade-off, we study the transmission power allocation scheme over the mm-wave band that maximizes the sum-rate of the MARC. The resulting

power allocation scheme is then characterized in closed form, which is shown to have rich structural properties and reveal interesting insights into the trade-offs in relay cooperation in dual-band networks.

While stand-alone mm-wave and dual-band networks are still in their infancy, the incumbent microwave networks still provide the bulk of the connectivity. In fact, most research initiatives into high speed mm-wave networks are motivated by the need to support high bandwidth demands in locations with high user density such as urban centers, university campuses, etc. In contrast, for low data-rate demanding areas such as rural regions, it may be advantageous to continue using the already available microwave networks which provide reliable, long-range connectivity with moderate speed. Therefore, incumbent microwave networks are still indispensable, especially since commercial mm-wave networks are currently being rolled out only in a handful of urban regions [78], and more time is needed for experimentation and full-scale deployment. However, as already discussed, conventional microwave networks suffer from inter-user interference and the optimal strategies for most interference-networks are unknown in general.

Therefore, in Chapter 5, we study the multiple-access interference channel (MAIC) operating over the conventional microwave band only, which models the uplink cellular interference in conventional networks. Here, a pair of two-user multiple access channels operating in two neighboring cells mutually interfere. The capacity of the MAIC is unknown in general. In particular, for the MAIC with weak interference, which provides a model of the inter-cell interference in practice, existing achievable strategies are suboptimal. Hence, for this model, we first characterize an achievable strategy that provides better performance by allowing for partial interference decoding. Then, focusing on the sum-rate of the MAIC for the partially symmetric case, we derive 3 novel sum-rate upper bounds. Numerical examples illustrate that one of the upper bounds is quite close to the achievable sum-rate for a significant range of cross channel gains, which thus provides improved sum-rate upper and lower bounds in this range. Moreover, we characterize a set of rate bounds that serves as an outer to the capacity region of the MAIC irrespective of whether the MAIC has strong or weak interference.

Finally, in Chapter 6, conclusions are drawn and potential future works are outlined.

1.3.1 Organization of the Thesis

The thesis contains chapters that are organized as follows:

- **Chapter 1** provides a brief introduction to conventional networks and integrated mm-wave/microwave dual-band networks, and briefly outlines the motivation for the thesis and contribution of the thesis.
- **Chapter 2** provides background on existing works on the dual-band networks, and present a brief summary of information-theoretic terminologies for related multi-user channels. We also identify specific research objectives.
- **Chapter 3** presents the capacity and approximate capacity results as well as the study of a resource allocation strategy for the dual-band interference channel.
- **Chapter 4** presents the capacity results and the the study of a resource allocation strategy for the dual-band multiple-access relay channel.
- **Chapter 5** presents the results on the multiple-access interference channel over the conventional band.
- **Chapter 6** provides conclusions derived from the thesis, where several potential future studies are also outlined.

Chapter 2

Literature Survey

2.1 Integrated Millimeter-Wave/Microwave Dual-Band Networks

Explosive growth in the number of connected devices and the ever increasing demand for high-bandwidth applications in the future has motivated the use of the mm-wave band for cellular communication. As discussed in Chapter 1, transmissions in the mm-wave band are relatively unreliable due to their susceptibility to blockage. Hence, the paradigm of using the mm-wave band along with the conventional microwave band in an integrated dual-band setup has emerged as a promising technology in 5G.

In this chapter, we first discuss the current state of the art on integrated mm-wave/microwave dual-band communication, and then introduce the basics and state of the art of relevant multi-user information theoretic models.

2.1.1 Characteristics of Millimeter Wave Transmission

Transmissions in the mm-wave band suffer from heavy attenuation due to increased propagation losses including losses from rain, atmospheric and molecular absorption [24,25,29]. Such losses can be overcome by deploying beamforming at the transmitters and the receivers with co-phased-antennas arrays [26, 73]. Such antenna-arrays are already commercially available, e.g., the AWA-0142 by Anokiwave [73], which is a 256-element reconfigurable co-phased antenna-array operating in the 24.25-27.5 GHz band.

Communicating with directional antenna-arrays causes the mm-wave links to become highly directional. In addition, atmospheric and oxygen absorption in this band makes scattering sparse [26, 27], thus effectively producing *point-to-point* links in the mm-wave band. For example, transmitting with circularly polarized directional antenna-arrays and receiving through narrow beamwidth antennas, multipath reflections in the 60 GHz band can be suppressed to a great extent [34, 35, 79]. Such sparsity, coupled with narrow-beam directed communication, explains the directed nature mm-wave communication.

Numerous studies on mm-wave networks have focused on characterizing schemes that utilize the directed nature of mm-wave transmissions to obtain high network throughput. For instance, communicating with narrow-beam directional antennas in the 60 GHz mm-wave band is shown to produce highly directional links in indoor network settings [33–35] as well as outdoor network settings [30, 36]. Such links typically have a strong line-of-sight (LoS) component, whereas the non-LoS components are found intermittently at best, if any, and only appear in highly cluttered environments with powers an order of magnitude less than that of the LoS component. As such, the study in [30] motivated a "*pseudo-wired*" abstraction of directional mm-wave links, based on which a cross-layer multihop MAC protocol was proposed to counter blockage in an outdoor network. In [63], a hybrid mm-wave and 4G microwave protocol was proposed to communicate over a mm-wave/microwave dual-band. This protocol was specifically designed with the aim of utilizing the directional mm-wave links to reduce inter-user interference and achieve high data throughput in a device-to-device communication setting. Similar works such as those in [80–82] also use the directed nature of mm-wave links to design network layer protocols that provide high throughput in stand-alone mm-wave networks.

While transmissions in the microwave band are able to combat blockage by diffracting around large objects such as buildings, transmissions in the mm-wave band are typically unable to do so due to the small wavelength and directed nature of such links, resulting in increased absorption and blockage [29, 41, 42]. For example, building materials such as bricks can attenuate mm-wave transmissions by as much as 80 dB [43]. Even the human body can attenuate mm-wave transmissions by 20-30 dB [44], and nominal indoor human movement can block mm-wave transmissions for about 1% to 2% of time in realistic settings [45]. Moreover, since the carrier frequency in the mm-wave band is much higher compared to that in the microwave band, the Doppler effect [83] increases considerably. Thus the mm-wave channel typically changes in a few hundred micro seconds, making it difficult to estimate the channel in a timely manner. This, coupled with blockage, can

result in intermittent connectivity of mm-wave links.

In a nutshell, the mm-wave links are able to support *high data rates* due to their large bandwidth but can be *unreliable*.

2.1.2 Millimeter Wave Networks and their Limitations

Early studies on mm-wave networks mostly focused on investigating the propagation characteristics of indoor stand-alone mm-wave networks. In particular, the transmission characteristics of the 60 GHz band was studied for specific target applications such as HDTV and UHDV streaming [84], WiGiG, Wireless HD and ECMA 387 [1], etc. Studies on indoor cellular connectivity reveal that transmitting via omnidirectional antennas offers more resilience against blockage via multipath propagation, but its coverage range is quite small [85]. This range is considerably extended by communicating with a directional antenna-array which creates a strong directional line-of-sight (LoS) link [33]. A similar study in [34] with a directive antenna with a 3-dB beamwidth of $\sim 10^{\circ}$ was also shown to produce directional mm-wave links. Similar directional mm-wave links were the used to improve performance via new multihop MAC protocols [86], or through macroscopic diversity [35, 87], etc.

Since preliminary studies on mm-wave propagation provided favorable results in indoor environments, similar studies for outdoor networks were also conducted. For instance, it was shown in [36] that similar to indoor settings, communicating with directional antennas in the 60 GHz band typically results in strong LoS links in many cases. Similar results were found in the study in [30] where mm-wave links were typically directional and misaligned beams and lack of coordination were the primary bottleneck to network performance. Moreover, studies in [80–82] from the network layer perspective also concluded that beamforming with directed antenna-arrays is instrumental in reducing inter-user interference in the mm-wave band to a great extent compared to the microwave band.

Since mm-wave communication has shown promising results in a few outdoor settings, the next set of studies naturally focused on understanding whether cellular communication over the mm-wave band is viable. The studies in [41, 46] conducted in the 28 GHz band in Brooklyn and Manhattan reveal that while both LoS and non-LoS paths exist, non-LoS paths suffer from significantly heavier attenuation than LoS paths (as expected), and beyond 200 m a significant number of users (e.g., 57 % in Brooklyn) were in outage. A similar study in [88] in Manhattan found that unlike in [41, 46], the non-LoS multi-path

components, possibly created due to very dense urban features, were quite strong and helped reduce blockage. However, reliable connections were mostly limited to only 78 m, compared to 200 m in [41, 46]. A related study on the 38 GHz band in Austin [47–49] found propagation characteristics that are qualitatively similar to those found in earlier studies in [41, 46]. Such studies, along with a more detailed study in [50] indicate that stand-alone mm-wave cellular networks are expected to provide an order of magnitude more cell-capacity than existing 4G networks. However, the effective size of mm-wave cells will likely be limited to ∼200 m.

The Need for Integrated Millimeter-Wave/Microwave Dual-Band Networks

In cellular settings, the relatively smaller range of mm-wave coverage may be useful in reducing interference between neighboring cells, thus increasing frequency reuse. However, smaller cell sizes need to be compensated with much denser cell deployments which increases the overall network capacity significantly. While dense mm-wave cell deployment would be expensive, the increased capacity from such dense cell deployment can potentially be utilized in an efficient manner in high bandwidth-demanding regions such as urban city centers, university campuses, tourist spots, major transport hubs, etc. However, in areas with moderate to low bandwidth demand such as suburban or rural regions, capacity of such dense mm-wave networks will likely be severely under utilized.

Moreover, due to relatively smaller size of mm-wave cells, mobile users or vehicles in transit will need to undergo frequent hand-overs between adjacent cells. Furthermore, as discussed earlier, mm-wave channels are harder to estimate reliably as they change significantly in a few hundred micro-seconds due to increased Doppler effect. Therefore, such frequent hand-overs and channel estimation will interrupt connectivity, and increase channel estimation overhead significantly. In addition, blockage continues to be a significant bottleneck to mm-wave connectivity, especially in dense urban settings.

As such, while stand-alone mm-wave cellular networks may provide a drastic increase in network throughput due to their vast bandwidth, such networks may suffer from unreliability, along with other issues such as network underutilization, and increased mobility management overhead. These issues suggest that stand-alone mm-wave cellular systems are unlikely to provide uniform and reliable coverage with robust high capacity links to users across different types of deployments. Hence, it has been envisioned [61,89] that mm-wave cells will coexist with microwave bands in an integrated dual-band setup, which may

serve as an effective solution to the bandwidth-scarcity problem in the short-to-medium term future.

Dual-band connectivity has thus emerged as one of the most promising technologies for 5G cellular access in recent times, beginning with the introduction of the Phantom cell architecture by NTT-Docomo in 3GPP Release 12 [55] in 2014, through the introduction of the dual connectivity (EN-DC) configuration in the recent 3GPP Release 15 [90] in 2019. The key idea is to integrate additional spectrum from the mm-wave band while continuing to use the incumbent microwave spectrum in the sub-6 GHz band. Transmissions over the mm-wave band are expected to handle high-throughput data sessions, while transmissions in the conventional microwave band are likely to carry control signals, and provide connectivity of moderate speed data when the mm-wave network connectivity is disrupted or unavailable [56–58].

It is interesting to note that the transmission characteristics in the mm-wave band are complementary to those in the microwave band. To combat the propagation loss in higher mm-wave frequencies, mm-wave links need to be used with directional antenna-arrays. This renders mm-wave links highly directional, and as discussed earlier, such links are prone to blockage and mobility management issues, thereby limiting cellular coverage severely [29, 41, 42]. In contrast, conventional communication in the microwave band provides extended coverage and far less blockage [29, 63, 91]. Such complementary characteristics of the two bands thus call for an integrated dual-band access scheme where users can fall back on the conventional microwave cell when mm-wave coverage is lost [41, 60].

Dual-Band Networks from the Network Layer Perspective

In [62], a joint mm-wave and microwave network with time division multiplexing-based MAC structure was studied as a prospective candidate for 5G. Here, important control functions and moderate bandwidth traffic are communicated over the existing and reliable microwave network, whereas the high capacity mm-wave links are used to offload traffic from the macro cells and provide better services for high bandwidth demanding traffic.

Unlike the microwave band, where inter-user interference is typically managed by interference-avoidance schemes, e.g., scheduling, the directional nature of mm-links are particularly useful in simplifying design. For instance, in the related work in [63] an optimal resource allocation problem for a cellular network over the integrated mm-wave/microwave

dual-band is studied. The proposed technique, which utilizes the directional nature of the mm-wave links, is shown to vastly improve performance metrics such as the number of supported users and the link connection probability, etc.

Similarly in [60], several variations of the heterogeneous network over the existing microwave band and the mm-wave band were studied, with the objective of achieving high-capacity access and backhaul. In this setting, control traffic is carried by microwave links, while high speed services are supported by mm-wave links, leading to a high-throughput and reliable system. Studies such as [64, 92] also provide similar results, while that in [32] focuses on minimizing the energy consumption of such networks.

Recently, [38] studied the role of spatial re-use in enhancing resource management of a multi-hop heterogeneous network operating over an integrated mm-wave/microwave dual-band, where directed antenna-arrays are used to create the mm-wave links. In this work, solutions were proposed that improve the throughput and latency performance of the network. In a similar work in [93], a heterogeneous network with directional point-to-point mm-wave links and conventional microwave links are considered for device-to-device communication, and a resource allocation study is then shown to achieve significant improvement over stand-alone mm-wave transmission schemes.

In a related work in [64], a heterogeneous network was considered over two mm-wave bands, the 60 GHz band and the 70/80 GHz band. Due to its relatively higher absorption and blockage, the 60 GHz band was used to create short-distance point-to-point links, while the 70/80 GHz band was used to establish relatively longer-range links. A novel paradigm was then introduced that leverages the point-to-point links in the 60 GHz and the more reliable links of the 70-80 GHz bands to achieve significant throughput enhancement over the conventional networks.

Dual-Band Networks from the Physical Layer Perspective

Recent works have also focused on the physical layer performance metrics of integrated mm-wave/microwave dual-band networks. In [70], a cellular network over such a dual-band is studied where mm-wave small cells have been densely deployed with conventional sub-6 GHz microwave cells. The study of opportunistic uplink and downlink user-association in this network reveals that, while directional beamforming in the mm-wave band improves the user association overall, users are more likely to connect to sub-6 GHz cells in the uplink, thereby indicating the efficacy of a dual-band connectivity. Similar works in [23, 94] also

conclude that simultaneous transmission over the integrated mm-wave/microwave dual-band is one of the most promising solutions that addresses the problem of unreliability in mm-wave transmission. In these works, scheduling algorithms were proposed that utilizes side information from both bands in the network to provide optimized quality of service. In related works [68, 69], the channel information in the sub-6 GHz band is shown to be useful in estimating the mm-wave channels in such dual-band networks.

In a related work in [95], the authors studied the performance of traffic hot-spots that operate over an integrated mm-wave/microwave dual-band. They conclude that this dual-band network outperforms both the conventional microwave only network and the stand-alone mm-wave network in terms the coverage probability. They also show that there exists a non-trivial trade-off between the deployment parameters of the mm-wave and the microwave base stations that achieves the optimal performance.

More importantly, developments on dual-band communication such as those in [66, 67, 71] have demonstrated the immense potential of such networks in practice. For example, the authors in [66] designed a queue-based strategy that transmits from a single transmitter to a single receiver simultaneously over the dual 3 GHz microwave and 30 GHz mm-wave bands. Successful practical experiments conducted in this dual-band setup favorably indicate the feasibility and the effectiveness of dual-band transmission schemes. In addition, Intel [67] and Qualcomm [71] have announced the production of a dual-band modem that supports both sub-6 GHz and 28 GHz bands.

Finally, recent works in [37, 39, 40] are a few of the most relevant works on mm-wave networks from the physical layer perspective. Here fundamental *information-theoretic* performance limits of a class of stand-alone mm-wave Gaussian relay networks are established. In these works, the authors focused on Gaussian 1-2-1 networks, where a single transmitter communicates with a single receiver with the help of a set of half-duplex relays. In particular, while [37, 39] focuses on establishing approximate capacity results for such channels, thereby characterizing good achievable strategies for the network, [40] focuses on secure communication over the channel. It is worth noting that mm-wave transmissions in this network are also modeled as point-to-point additive white Gaussian noise channels, an assumption that aligns well with those in many studies from the network and physical layer perspectives and from practical measurement campaigns.

Connection to the Thesis

The preceding discussions show that while transmitting over the mm-wave spectrum is a promising technology for next generation wireless networks, stand-alone mm-wave networks inherently suffer from unreliability of mm-wave links. Therefore, communicating over an integrated mm-wave/microwave dual-band, which complements the transmissions in the high-bandwidth, unreliable mm-wave links with transmissions in the moderate-bandwidth, reliable microwave links, provides an attractive solution for the near-to-medium term future. Hence, networks over such an integrated dual-band have been subject to much investigation recently from the physical and network layer perspective, as discussed above.

However, very few studies, if any, focus on characterizing the information-theoretic performance limits of such dual-band networks. Such studies, especially for multi-user networks, are instrumental in finding practical achievable performance limits, good encoding-decoding strategies, effective resource allocation schemes, etc. While the works in [37,39,40] characterize information-theoretic performance limits of a specific stand-alone mm-wave channel, they consider neither dual-band networks nor the aspect of inter-cell interference.

In this thesis, we make progress in this front by studying two multi-user channels over the integrated mm-wave/microwave dual-band that analyze the aspects of interference and relay-cooperation.

2.2 Multi-User Information Theoretic Models

In the pioneering work in [96], Shannon first considered the problem of communicating from a single source to a single destination in a point-to-point channel under a probabilistic framework. It was shown that data can be reliably transmitted over the channel if the data rate is less than the so-called capacity of the channel (C), thus providing a fundamental limit on data rates.

Specifically, a discrete memoryless channel (DMC) is defined by the tuple $(\mathcal{X}, \mathcal{Y}, p_{Y|X}(y|x))$, where $\mathcal{X}$ and $\mathcal{Y}$ are the finite, discrete alphabets, and input $x \in \mathcal{X}$ produces output $y \in \mathcal{Y}$ according to the channel transition probability $p_{Y|X}(y|x)$.

The fundamental result in [96] shows that an input message with rate R can be reliably transmitted over the DMC if

$$R < C := \max_{p_X} I(X;Y),$$

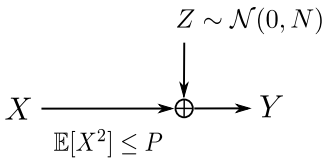


Figure 2.1: The point-to-point additive white Gaussian noise (AWGN) channel.

where $I(X;Y)$, the so-called *mutual information* between random variables X and Y, is defined as

$$I(X;Y) = \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p_X(x) p_{Y|X}(y|x) \log \frac{p_{Y|X}(y|x)}{p_Y(y)},$$

where X and Y are from finite discrete alphabets $\mathcal{X}$ and $\mathcal{Y}$, and they are distributed as $X \sim p_X(x)$ and $Y \sim p_Y(y) := \sum_{x' \in \mathcal{X}} p_{Y|X}(y|x') p_X(x')$. The input distribution p_X^* that maximizes $I(X;Y)$ is called the capacity achieving or optimal input distribution.

The point-to-point Additive White Gaussian Noise (AWGN) channel, depicted in Figure 2.1, is a continuous-alphabet counterpart of the DMC and a popular model that effectively captures noisy wireless communication. The system model of the standard AWGN is given by

$$Y_i = X_i + Z_i, \quad i = 1, \ldots, n,$$

where X_i and Y_i are the real input and output symbols at i^{th} channel, i.e., $X_i, Y_i \in \mathbb{R}$, and Z_i are additive white Gaussian noise samples which are identically and independently distributed (i.i.d.) as $Z_i \sim \mathcal{N}(0, N)$. In real-world transmissions, the inputs must satisfy a transmit power budget P in that over a block of n symbols, input X^n must satisfy $\frac{1}{n} \sum_{i=1}^n \mathbb{E}[X_i^2] \leq P$.

The capacity of the AWGN channel is given by

$$C = \frac{1}{2} \log \left(1 + \frac{P}{N} \right), \tag{2.1}$$

where the so-called signal-to-noise ratio is denoted by $\mathsf{SNR} := \frac{P}{N}$, and the optimal distribution is $X \sim \mathcal{N}(0, P)$.

Multiple-Access Channel

While the AWGN channel models single-user communication between a single transmitter-receiver pair, it is common to have multi-user communication in cellular settings, where multiple transmitter-receiver pairs communicate with each other over a shared medium.

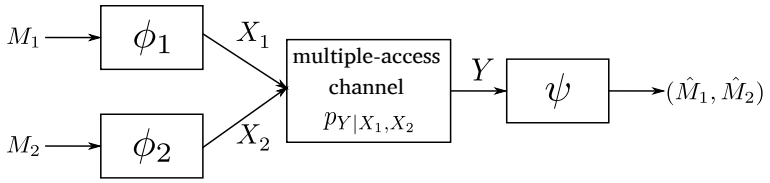


Figure 2.2: The two-user discrete memoryless multiple-access channel (DM-MAC).

One of the most widely studied multi-user channel is the multiple-access channel (MAC) where two or more sources simultaneously communicate to a single destination, which then attempts to estimate the messages from all sources.

The two-user discrete memoryless MAC [97, Chapter 4], denoted DM-MAC, is depicted in Figure 2.2 and defined as follows:

Definition 2.1 (The DM-MAC)**.** *The DM-MAC is defined by the tuple* $(\mathcal{X}_1, \mathcal{X}_2, \mathcal{Y}, p_{Y|X_1,X_2}(y|x_1,x_2))$ *where* $\mathcal{X}_1$ *and* $\mathcal{X}_2$ *are two finite, discrete input alphabets,* $\mathcal{Y}$ *is a finite, discrete output alphabet, and* $p_{Y|X_1,X_2}(y|x_1,x_2)$ *is the channel transition law.*

Here, $\mathcal{X}_1$ and $\mathcal{X}_2$ are *input* alphabets for the two sources S_1 and S_2, while $\mathcal{Y}$ is the *output* alphabet at the destination D, as depicted in Figure 2.2. Since the channel is memoryless, the channel probability mass function (pmf) after n channel uses decomposes as

$$p(y^n|x_1^n, x_2^n,) = \prod_{i=1}^{n} p(y_i|x_{1i}, x_{2i}).$$

We now introduce information theoretic terminologies for the MAC which generalizes to other multi-user channels as well. The problem of communicating over the DM-MAC is formalized by defining a *code* and an *achievable rate pair* [97, Chapter 4] as given below.

Definition 2.2 (Code for the DM-MAC)**.** *A* $(2^{nR_1}, 2^{nR_2}, n)$ *code for the DM-MAC consists of (i) two independent, uniformly distributed message sets* $\mathcal{M}_k := \{1, 2, \ldots, 2^{nR_k}\}$, $k = 1, 2$, *one for each of the sources* S_1 *and* S_2*; (ii) two encoders* $\phi_k : \mathcal{M}_k \to \mathcal{X}_k^n \in \mathcal{X}_k^n, k = 1, 2$, *one for each of* S_1 *and* S_2*; and (iii) a decoder for the destination* D $\psi : \mathcal{Y}^n \to \mathcal{M}_1 \times \mathcal{M}_2$.

To communicate messages M_1 and M_2, they are encoded into codewords $X_1^n(M_1) = \phi_1(M_1)$ and $X_2^n(M_2) = \phi_2(M_2)$, and transmitted over the channel. The destination receives a superimposed signal Y^n from which it decodes both messages. The probability of decoding error at the destination is given by

$$P_e^n := \Pr\left[\psi(Y^n) \neq (M_1, M_2)\right],$$

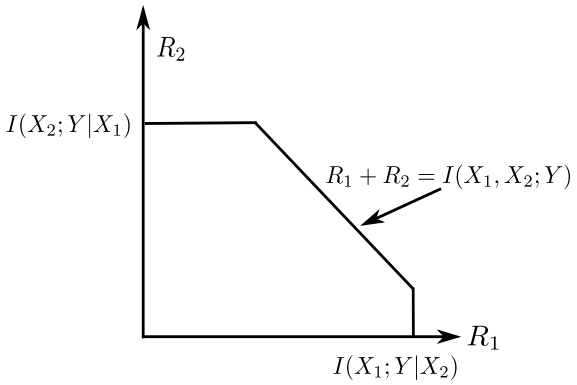


Figure 2.3: An example of the capacity region of the DM-MAC (neglecting Q).

where $(M_1, M_2) \in \mathcal{M}_1 \times \mathcal{M}_2$ are uniformly distributed.

The objective here is to design encoding/decoding strategies such that both messages can be decoded with vanishing error probability (P_e^n) as $n \to \infty$. These results are typically described in terms of achievable rates and capacity of the channel.

Definition 2.3 (Achievable rate pair and Capacity region for the DM-MAC)**.** *A rate pair (R_1, R_2) is said to be achievable for the DM-MAC if there exists a sequence of $(2^{nR_1}, 2^{nR_2}, n)$ codes such that $P_e^n \to 0$, as $n \to \infty$. The capacity region of the DM-MAC is defined as the closure of the set of all nonnegative achievable rate pairs.*

The capacity region of a multi-user channel provides a fundamental rate limit for the sources beyond which its message cannot be *reliably* decoded. Characterizing the capacity region of a channel typically involves two steps: (a) a converse part, which follows from Fano's inequality and specifies rate constraints (outer bounds) that is valid for arbitrarily correlated codewords; (b) an achievable part, which entails identifying a specific encoding/decoding strategy and codeword distribution that achieves the outer bounds.

For the DM-MAC, the *capacity region* is given by the set of all non-negative (R_1, R_2) that satisfy

$$\begin{aligned} R_1 &\leq I(X_1; Y|X_2, Q) \\ R_2 &\leq I(X_2; Y|X_1, Q) \\ R_1 + R_2 &\leq I(X_1, X_2; Y|Q), \end{aligned} \tag{2.2}$$

for some joint distribution $(X_1, X_2, Q) \sim p(q)p(x_1|q)p(x_2|q)$, where $Q \in \mathcal{Q}$ is the time-sharing variable with cardinality $|\mathcal{Q}| \leq 4$. In Figure 2.3, an example of the capacity region of the DM-MAC is provided (neglecting the time sharing variable Q).

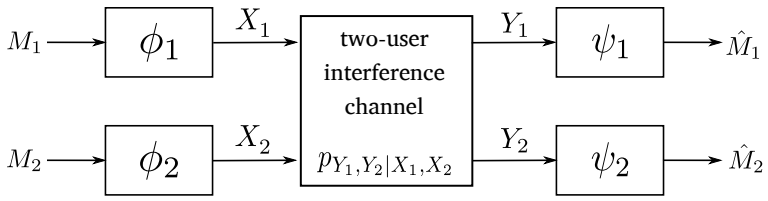


Figure 2.4: The two-user discrete memoryless interference channel (DM-IC).

The terminologies and the results for two-user DM-MAC naturally generalize to the K-user DM-MAC [97].

Similar to the DMC case, the DM-MAC has a counterpart that models the wireless multi-user access: the Gaussian MAC [97, Chapter 4.6]. In the two-user Gaussian MAC, the sources communicate their messages via real codewords X_1^n and X_2^n that must satisfy block power constraints P_1 and P_2. Notably, similar to the AWGN channel, the capacity region of the Gaussian MAC is achieved by using independent codewords $X_k^n \sim \mathcal{N}(0, P_k), k = 1, 2$, i.i.d., and without time-sharing.

2.2.1 Interference Channel

The interference channel (IC) is a fundamental model for multi-user *interference-limited* communication, where two or more source-destination pairs communicate over a shared wireless medium, and thus mutually interfere. Unlike the MAC, where the destination attempts to decode both desired messages, a destination in the IC is interested in its desired message only. Since interference is not useful for the destination and difficult to mitigate, it can drastically reduce the source-destination rate or network throughput if interference is not mitigated judiciously [97, Chapter 6].

The two-user discrete memoryless interference channel (DM-IC) [97, Chapter 6], depicted in Figure 2.4, is a fundamental model which provides insight into the functioning of the general interference-limited networks. The DM-IC consists of two source-destination pairs where source S_1 communicates with destination D_1 and source S_2 communicates with destination D_2 while interfering mutually.

Definition 2.4 (The DM-IC). *The DM-IC is defined by the tuple* $(\mathcal{X}_1, \mathcal{X}_2, \mathcal{Y}_1, \mathcal{Y}_2, p_{Y_1,Y_2|X_1,X_2}(y_1, y_2|x_1, x_2))$ *where* $\mathcal{X}_1$ *and* $\mathcal{X}_2$ *are two finite, discrete input alphabet sets,* $\mathcal{Y}_1$ *and* $\mathcal{Y}_2$ *are two finite, discrete output alphabet sets and* $p_{Y_1,Y_2,X_1,X_2}(y_1, y_2|x_1, x_2)$ *is the channel transition law.*

The transmission from the two sources are uncoordinated, and thus the transmitted messages are assumed to be independent. Since the channel is discrete memoryless, the n-letter channel pmf decomposes as

$$p(y_1^n, y_2^n|x_1^n, x_2^n) = \prod_{i=1}^{n} p(y_{1i}, y_{2i}|x_{1i}, x_{2i}). \tag{2.3}$$

The problem of reliably communicating over the IC is formalized by defining a code and an achievable rate pair as follows:

Definition 2.5 (Code for the DM-IC)**.** *A $(2^{nR_1}, 2^{nR_2}, n)$ code for the 2-DM-IC consists of (i) two independent, uniformly distributed message sets $\mathcal{M}_k := \{1, 2, \ldots, 2^{nR_k}\}$, $k = 1, 2$, one each for sources S_1 and S_2; (ii) two encoders, $\phi_k : \mathcal{M}_k \to \mathcal{X}_k^n, k = 1, 2$, one each for S_1 and S_2; and (iii) two decoders for destinations D_1 and D_2, $\psi_k : \mathcal{Y}_k^n \to \mathcal{M}_k, k = 1, 2$.*

To communicate messages $M_1 \in \mathcal{M}_1$ and $M_2 \in \mathcal{M}_2$ from S_1 and S_2, they are encoded into codeword $X_1^n(M_1)$ and $X_2^n(M_2)$ for transmission. Due to mutual interference, both codewords superimpose at each destination, and each destination attempts to estimate its designated message only. The decoding probability of error for the code is defined as

$$P_e^n = \Pr[\psi_1(Y_1^n) \neq M_1 \quad \cup \quad \psi_2(Y_2^n) \neq M_2],$$

where the average is taken over uniform distribution of $(M_1, M_2) \in \mathcal{M}_1 \times \mathcal{M}_2$.

Definition 2.6 (Achievable rate for the DM-IC)**.** *A rate pair (R_1, R_2) is said to be achievable for the DM-IC if there exists a sequence of $(2^{nR_1}, 2^{nR_2}, n)$ codes such that $P_e^n \to 0$, as $n \to \infty$.*

The *capacity region* of the DM-IC is defined as the closure of the set of all nonnegative achievable rate pairs. Unlike the DM-MAC, the closed form expression of capacity region of the DM-IC is *unknown in general* except for the strong interference case.

Definition 2.7 (Strong interference for the DM-IC)**.** *The DM-IC is said to have strong interference [98] if*

$$I(X_1; Y_1|X_2) \leq I(X_1; Y_2|X_2), \quad I(X_2; Y_2|X_1) \leq I(X_2; Y_1|X_1) \tag{2.4}$$

for all input distribution $p(x_1, x_2)$ on X_1 and X_2 with $p(x_1, x_2) = p(x_1)p(x_2)$.

The capacity region of the DM-IC under strong interference is given by the set of all nonnegative (R_1, R_2) that satisfy

$$R_1 \leq I(X_1; Y_1|X_2, Q)$$

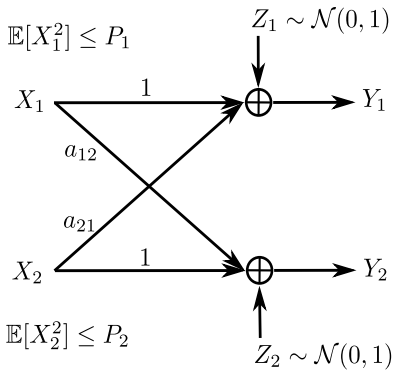


Figure 2.5: The two-user Gaussian interference channel (GIC).

$$R_2 \leq I(X_2; Y_2|X_1, Q)$$
$$R_1 + R_2 \leq \min\left(I(X_1, X_2; Y_1, Q), I(X_1, X_2; Y_2, Q)\right),$$

for some $(X_1, X_2, Q) \sim p(q)p(x_1|q)p(x_2|q)$, where $Q \in \mathcal{Q}$ is the time-sharing variable.

Compared to the DM-MAC, where the destination attempts to decode both messages, each destination in a DM-IC only decodes its own message. However, under strong interference, decoding both the *desired user* as well as the *interfering user* at each destination as in a DM-MAC is optimal. Therefore, in a strong DM-IC, each destination can deploy the DM-MAC decoder without additional rate penalty. Since, the sources effectively form a MAC at each of the destinations, the overall rate region is obtained by taking the intersection of these two MAC rate regions.

Note that the optimal encoding strategy for DM-ICs without strong interference, especially those with weak interference where the inequalities in (2.4) is reversed, is unknown.

Two User Gaussian Interference Channel

The two-user real Gaussian interference channel (denoted GIC) can be regarded as a wireless counterpart of a DM-IC, where the transmitted signals are continuous and are corrupted by AWGN at the receivers. The GIC thus effectively captures the interference-limited multi-user communication in wireless medium. Unlike the single-user AWGN channel, where increasing the user transmit power is helpful in improving the user rate, in the GIC increasing the transmit powers does not necessarily improve the performance as the resulting power in interference also increases proportionally [10]. Fortunately, interference often has useful structure that can be exploited to find strategies leading to superior throughput [10].

Without loss of generality, the GIC can be described by its standard form [99] as depicted in Figure 2.5: the channel outputs at destination D_1 and destination D_2 at time t are given by

$$Y_{1,t} = X_{1,t} + a_{21}X_{2,t} + Z_{1,t}, \tag{2.5}$$

$$Y_{2,t} = a_{12}X_{1,t} + X_{2,t} + Z_{2,t}, \quad t = 1, \ldots, n, \tag{2.6}$$

where the input symbols $X_{k,t} \in \mathbb{R}$ and the output symbols $Y_{k,t} \in \mathbb{R}$ are real-valued, $Z_{k,t} \sim \mathcal{N}(0,1)$, are i.i.d., AWGN samples, and $a_{k\ell} \in \mathbb{R}$ is the cross channel coefficients from source S_k to D_ℓ, $k \neq \ell \in \{1,2\}$. Similar to the AWGN channel, the input codewords X_k^n must satisfy a transmit power budget: $\frac{1}{n}\sum_{i=1}^n \mathbb{E}[X_{k,i}^2] \leq P_k, k = 1, 2$.

The terminologies for the GIC such as the definitions of a $(2^{nR_1}, 2^{nR_2}, n)$ code, the average probability of error and the achievable rates follows from those for the DM-IC after imposing the transmit power constraints, and hence are not repeated here.

Capacity Results for the Gaussian IC

Similar to the two-user DM-IC, the capacity region of GIC is not known in general except for certain regimes of channel parameters. In particular, the GIC is classified according to the relative strength of the power of interfering signal compared to the power of desired signals, which can be equivalently described by *regimes of cross channel gains* a_{12}^2 and a_{21}^2 [100] as follows. As such, the capacity region of the GIC has been characterized in closed form for the *very strong* interference and the *strong* interference regimes, while in the *weak* interference regime only few results exist.

Definition 2.8 (Very strong interference for the GIC [99])**.** *The GIC is said to have very strong interference if*

$$a_{12}^2 \geq 1 + P_2, \quad a_{21}^2 \geq 1 + P_1.$$

In this case, the capacity region is given by the set of $(R_1, R_2) \in \mathbb{R}_+^2$ such that

$$R_1 \leq \mathsf{C}(P_1)\,,$$
$$R_2 \leq \mathsf{C}(P_2)\,,$$

where $\mathsf{C}(x) := \frac{1}{2}\log_2(1+x)$.

In this case, each source encodes its message into codewords $X_k^n \sim \mathcal{N}(0, P_k), k = 1, 2$, i.i.d., whereas each destinations deploys *successive interference decoding*: a destination first

decodes the interference by treating its desired signal as noise, removes the interference from the received signal, and then decodes the desired message from the resulting "interference free channel". The very strong interference condition ensures that such a successive decoding does not incur any rate penalty.

Definition 2.9 (Strong interference for the GIC [101]). *The GIC is said to have strong interference if*

$$a_{12}^2 \geq 1, \quad a_{21}^2 \geq 1.$$

The capacity region of the strong GIC is given by the set of $(R_1, R_2) \in \mathbb{R}_+^2$ such that

$$\begin{aligned} R_1 &\leq \mathsf{C}(P_1) \\ R_2 &\leq \mathsf{C}(P_2) \\ R_1 + R_2 &\leq \min\left(\mathsf{C}\big(P_1 + a_{21}^2 P_2\big)\,,\ \mathsf{C}\big(P_2 + a_{12}^2 P_1\big)\right). \end{aligned}$$

Similar to the very strong case, the capacity region is obtained by transmitting codewords $X_k^n \sim \mathcal{N}(0, P_k)$, i.i.d., for $k = 1, 2$. In contrast, the destinations now deploy *simultaneous unique decoding* to decode both signals in a MAC decoding fashion.

Definition 2.10 (Mixed interference for the GIC [102]). *The GIC is said to have mixed interference if*

$$\textit{either} \quad a_{12}^2 \geq 1, \ \ a_{21}^2 < 1, \quad \textit{or} \quad a_{12}^2 < 1, \ \ a_{21}^2 \geq 1.$$

While the capacity region for this case is not known entirely, the sum capacity for the mixed IC with $a_{12}^2 \geq 1, \ \ a_{21}^2 < 1$ is given by [102]

$$C_{sum} = \mathsf{C}(P_2) + \min\left(\mathsf{C}\left(\frac{P_1}{1 + a_{21}^2 P_2}\right),\ \ \mathsf{C}\left(\frac{a_{12}^2 P_1}{1 + P_2}\right)\right).$$

For the case with $a_{12}^2 \geq 1, \quad a_{21}^2 < 1$, the sum-capacity is achieved by transmitting i.i.d. Gaussian codewords, and then decoding the interference at D_1 (where interference is strong) and treating the interference as noise at D_2 (where the interference is weak). The sum capacity for second set of condition is found by interchanging the roles of D_1 and D_2.

Definition 2.11 (Weak interference for the GIC [102]). *The GIC is said to have weak interference if*

$$a_{12}^2 < 1, \quad a_{21}^2 < 1.$$

Similar to the mixed GIC, the capacity region for the weak GIC is not fully known. However, the sum-capacity is known for the so-called *noisy interference* regime [11,102,103],

where

$$|a_{21}(1+a_{12}^2P_1)|+|a_{12}(1+a_{21}^2P_2)|\leq 1$$

holds. For the noisy interference regime, the sum-capacity is given by

$$C_{sum}=\mathsf{C}\left(\frac{P_1}{1+a_{21}^2P_2}\right)+\mathsf{C}\left(\frac{P_2}{1+a_{12}^2P_1}\right),$$

which is achieved by transmitting inputs $X_k^n\sim\mathcal{N}(0,P_k)$, i.i.d., for $k=1,2$, and treating interference as noise at both destinations.

Approximate Capacity Results for the Gaussian IC

Since the capacity region of the GIC is unknown in the weak interference regime except for the sum-capacity result for the noisy interference regime, a considerable effort has been directed at characterizing the so-called *approximate capacity* results, also called *constant gap* results, for this regime. Constant gap results typically entail characterizing an achievable rate region and an outer bound to the capacity region of the channel such that the difference between the two differ only by a small *constant* that is independent of all channel parameters, i.e., channel gains and transmit powers.

Definition 2.12 (Constant gap for the GIC)**.** *An achievable region is said to be within δ bit/channel use of the capacity region if for a rate pair (R_1,R_2) on the boundary of the achievable region, the rate pair $(R_1+\delta,R_2+\delta)$ is outside the outer bound to the capacity region, where $\delta\in\mathbb{R}_+$ is a constant and not a function of channel parameters $a_{12}^2,a_{21}^2,P_1,P_2$.*

Intuitively, an achievable scheme with a constant gap of δ achieves an achievable rate region from which the capacity region is at most δ bits away per dimension (user). As such, if δ is reasonably small, this rate gap is effectively negligible as compared to the rates achieved with moderate transmit powers. Hence, research efforts are typically geared towards determining a "good" achievable scheme and a "good" outer bound such that they differ only by a small constant.

The HK scheme [7] is the most versatile achievable scheme for the GIC, and provides the best known achievable rates that contains all other known schemes such as the joint decoding (optimal in the strong interference regime), or treating interference as noise (optimal in the noisy interference regime), etc., as special cases. In this scheme, each input message is allowed to be divided into multiple independent parts, each regarded as a virtual user, that are superimposed in a layer-by-layer fashion to generate a single codeword

for transmission. As such, the total transmit power budget is divided into the transmit power of those virtual users, also known as power-splitting. The decoder then decodes all virtual users from its desired user with a range of possibilities: it may choose to decode a set of desired virtual users in one of the many possible orders, whereas for the interfering virtual users, it may choose to decode only a subset of them depending on the strength of the received interference power. In addition, the power-splitting operation can be varied to achieve a wide range of achievable regions.

Such *versatility* of approaches in the HK scheme results in a significant *complexity*. In fact, the HK achievable region, found by taking union of rate regions over all possible parameters such as the power-splitting, the distribution of codewords used, the number of virtual users, etc., has not yet been characterized in closed form. This motivates the need to *simplify* the HK scheme such as those in [10, 100]. The scheme of [10] is of particular interest as it achieves the capacity of the weak GIC within $\delta = \frac{1}{2}$ bits/channel use per user.

In this scheme, each message is divided into only *two* independent parts, a common and a private part, and the total power budget P_k is split into two *fixed* parts, P_{kc} and P_{kp} for the common and the private parts, that depend on the cross channel gains as follows

$$P_{kp} = \begin{cases} \dfrac{1}{a_{k\ell}^2}, & \text{if } a_{k\ell}^2 P_k \geq 1, \quad \ell \neq k \in \{1,2\} \\ P_k, & \text{otherwise, and} \end{cases}$$

$$P_{kc} = P_k - P_{kp}. \tag{2.7}$$

Moreover, only i.i.d., Gaussian codewords are used for encoding, and time-sharing is not used. Finally, each destination decodes the intended private and common messages *uniquely*, the *interfering common* message *non-uniquely*, and treats the interfering private message as *noise*. These simplifications result in an achievable region in closed form, given by the set of rate tuples $(R_1, R_2) \in \mathbb{R}_+^2$ that satisfy [7, Lemma 2]

$$\begin{aligned} R_1 &\leq I(X_1; Y_1|W_2) \\ R_2 &\leq I(X_2; Y_2|W_1) \\ R_1 + R_2 &\leq I(X_2, W_1; Y_2) + I(X_1; Y_1|W_1, W_2) \\ R_1 + R_2 &\leq I(X_1, W_2; Y_1) + I(X_2; Y_2|W_1, W_2) \\ R_1 + R_2 &\leq I(X_1, W_2; Y_1|W_1) + I(X_2, W_1; Y_2|W_2) \\ R_1 + 2R_2 &\leq I(X_2, W_1; Y_2) + I(X_2; Y_2|W_1, W_2) + I(X_1, W_2; Y_1|W_1) \\ 2R_1 + R_2 &\leq I(X_1, W_2; Y_1) + I(X_1; Y_1|W_1, W_2) + I(X_2, W_1; Y_2|W_2), \end{aligned} \tag{2.8}$$

where the $I(.;.)$ terms are computed with distributions $W_k \sim \mathcal{N}(0, P_{kc})$, i.i.d., $U_k \sim \mathcal{N}(0, P_{kp})$, i.i.d., with $X_k := W_k + U_k \sim \mathcal{N}(0, P_k)$, and Y_k are defined by (2.5)-(2.6) with $X_k \sim \mathcal{N}(0, P_k), k = 1, 2$.

To evaluate how close the achievable region in (2.8) is to the capacity, closed form outer bounds are needed, which are difficult to derive in the weak interference regime in general. To that end, considerable progress has been made in [10, 11] where *genie-signal aided* techniques were used to characterize useful closed form outer bounds. Specifically, the closed form outer bound derived in [10] is of interest, as the achievable region in (2.8) was shown to differ from these outer bounds only by at most 1/2 bit/channel use per user.

The Parallel Gaussian IC

The traditional two-user IC models the interference between two source-destination pairs that communicate over a *frequency-flat* channel: in frequency-flat channels, the attenuation of a signal is assumed to be constant (i.e., flat) over the entire bandwidth of the channel. In contrast, in many wireless channels, the signal attenuation exhibits different characteristics in different ranges of frequencies. Such channel are referred to as *frequency-selective* channels. A common approach to analyze frequency-selective channels is to divide the entire bandwidth of the channel into a number of *sub-channels* such that each sub-channel behaves as a frequency-flat channel [104, Chapter 11]. Such a set of sub-channels can be realized via frequency-band separation, and thus these sub-channels do not interfere with one another, and hence are referred to as *parallel channels*.

The parallel Gaussian IC (PGIC) models two-user interference-limited communication over a frequency selective channel [105, 106]. In contrast to the GIC over a single frequency-flat channel, communication in the PGIC is over a set of parallel channels that do not interfere with each other. As such, the PGIC can be regarded as a set of Gaussian ICs that operate over several orthogonal parallel channels (sub-channels): a conventional (single-band) GIC operates in each sub-channel without interfering with those in other sub-channels [105].

Similar to the conventional GIC, the capacity region of the PGIC is known only when the GIC in each sub-channel has *strong interference* [107], while its sum-capacity is known for the case with noisy interference in each band [106]. Specifically, the capacity region in the strong interference case is achieved by encoding *jointly* over all sub-channels (as opposed to encoding separately in each sub-channel), and decoding messages from both transmitters

from the signals received from *all* sub-channels. Hence, while it may apparently seem reasonable to perform encoding/decoding separately over each sub-channel of a PGIC, such separate strategy is *suboptimal* in general [108, 109].

The ergodic fading GIC (EF-GIC) [110] models the effect of fading in wireless interference-limited communication. In contrast to the GIC with fixed gains, in the EF-GIC the channel gains are subject to fading that vary according to some fading distribution, e.g., Rayleigh fading. The achievable rates for the EF-GIC is found by taking expectation of the rates over the fading distribution: this expectation can be interpreted as taking a *weighted sum* of rates achievable for each *realization* of channel gains (fading state), with the *weights* denoting the probability of occurrence of that fading state. Hence, the EF-GIC can be regarded as a PGIC as follows: each fading state of an EF-GIC corresponds to a sub-channel in a PGIC, and the achievable rates for the EF-IC is found by a weighted average of the rates achievable in each sub-channel of the PGIC, where the weights are the probability of realization of the fading state. As such, the capacity of the ergodic fading GIC was characterized in all cases for which the capacity of the PGIC is known [110].

Resource Allocation for the Interference Channel

The overall performance of multi-user channels are typically characterized by optimizing a specific design-metric. Such metrics include the sum-rate, aimed at maximizing the total throughput of all users, the worst-user rate, aimed at providing a fair rate to users, etc. For interference channels in particular, resource allocation techniques that maximize the sum-rate have revealed insights into how resource should be allotted to increase the throughput of practical microwave networks [105, 111, 112]. In [105], the GIC over a frequency selective channel, modeled as the PGIC, is studied. For this model, the sub-optimal resource allocation scheme of iterative waterfilling was proposed, which was shown to provide close-to-optimum performance while possessing a distributed coding and power allocation structure, which are attractive in practice. Similar studies on optimal resource allocation have also been conducted from a game-theoretic perspective such as that for the frequency selective GIC in [111], and the GIC with quality-of-service and RF energy harvesting constraints in [112], which reveal useful insights that improve the performance in the corresponding settings.

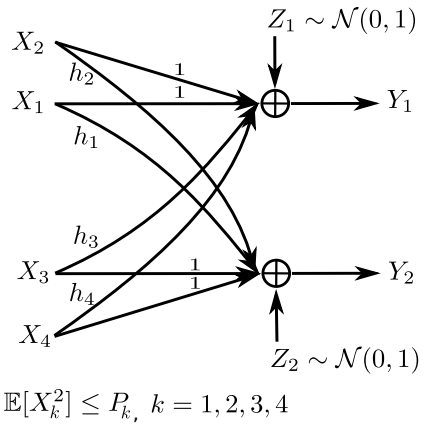


Figure 2.6: The Gaussian multiple-access interference channel (GMAIC) as defined in [12].

Connection to the Thesis

In Chapter 3, we study the two-user interference channel over the integrated mm-wave/microwave dual-band. We characterize fundamental performance limits of the dual-band IC in the form of capacity results and approximate capacity results for different variations of the channel. We then study the problem of resource allocation over the mm-wave links of the channel that optimizes the sum-rate of the dual-band IC, which reveals useful practical insights.

2.2.2 The Multiple Access Interference Channel

In cellular communication settings, the two-user IC models the interference between two point-to-point links that operates in two neighboring cells. Each such link can model the communication from a single source (e.g., user) to a single destination (e.g., the base station). However, as opposed to only a single transmitter-receiver pair, typically *multiple terminals* communicate within a cell. For example, during the cellular uplink phase, multiple users communicate to a single base station (BS) in the cell, which is effectively modeled as a multi-access channel (MAC). When two such MACs, operating in two neighboring cells, transmit over the same shared microwave band and mutually interfere, and the resulting channel is modeled as the multiple-access interference channel (MAIC) [12].

The MAIC can thus be regarded as a generalization of the two-user IC, as it models a more practical interference phenomena in cellular networks. Hence similar to the two-user IC, understanding how to address the interference problem in the MAIC is crucial in characterizing useful strategies for practical cellular networks.

In [12], the Gaussian MAIC (GMAIC), depicted in Figure 2.6, was studied where each interfering MAC is modeled as a Gaussian MAC. The authors in [12] classified the GMAIC based on the strength of interference in the *cross channels*, i.e., the channel from a source to the *interfered destination* located in the other cell. In particular, the GMAIC is said to have strong interference if the gains in the cross channels are larger than that of the respective direct channels, while it is said to have *weak interference* if the cross channel gains are smaller. Such a characterization serves as a natural extension of the strong and weak interference cases for the GIC [100].

The capacity of the GMAIC was then determined for a subset of the strong interference case in [12], where *jointly decoding* the interfering messages along with the intended messages as in a 4-user MAC is the optimal decoding strategy, if the channel gains satisfy a certain set of conditions in addition to the strong interference conditions. In contrast, for the weak interference regime, the capacity and the optimal achievable strategy is unknown in general. However, some progress were made in [12]: a sum-rate upper bound was derived for the weak interference regime, and then it was compared to the sum-rate achieved via the simple strategy of treating interference as noise (TIN). Empirical examples show that the gap between these two bounds are small when the cross channel gains are very small (close to zero). However, as either the transmit powers increase, or the cross channel gains increase within the weak interference regime, this gap increases consistently. This behavior calls for more sophisticated upper bounds and achievable strategies that perform well over a wide range of channel parameters.

In [113], another upper bound on the sum-capacity was derived via an interference-alignment approach. They also proposed an achievable scheme for the GMAIC based on lattice codes, and showed that the resulting sum-rate is within a constant gap to the sum-capacity. It should be noted that while [113] considers the same system model of the GMAIC as in [12], the model for channel gains, the definition of weak interference, and the upper bounding approaches are all different from those in [12].

First, in contrast to the channel gain of [12], which can model any real number, the notation adopted for channel gains in [113] is only able to model gains larger than 1. Second, while the definition of strong and weak interference in [12] can be regarded as a natural extension of the same for the GIC as in [100], which simply depends on the relative magnitude of the cross and direct channel gains and not on transmit powers, the weak interference in [113] is defined by a more complex relation involving both channel gains and transmit powers. Such a definition is motivated by the use of the specific outer

bounding technique in [113], and does not naturally extend from the same in the GIC.

Finally, the sum-capacity upper bound presented in [113, Theorem 7] is shown to be valid for a subset of channel gains. However, since the bound was derived using an existence result in number theory, it was only shown to exist but the exact subset of channel gains for which this bound holds was not specified. Moreover, the constant gap between this upper bound and the achievable sum-rate is presented as a function of some *unknown constants* that have *not* been characterized. Therefore, the bound is *not expressed* in closed form, nor is the exact value of the constant gap completely expressed.

Connection to the Thesis

The preceding discussion reveals that only a few results exist on the Gaussian MAIC. Particularly, in the weak interference regime, the upper bound of [12] becomes loose compared to the achievable sum-rate, while that in [113] is not characterized in closed form, as noted above.

In Chapter 5, we study the Gaussian MAIC over the conventional microwave band and characterize new performance bounds. The channel model and information-theoretic terminologies for the GMAIC are presented in Chapter 5, and hence are not repeated here. Note that since the notation for the GMAIC in [12] naturally extends from the same for the GIC, these notations are adopted in Chapter 5.

2.2.3 Relay Cooperation

Wireless channels are inherently unreliable due to obstacles, multipath channel fading, and power loss due to wave propagation over wireless medium. Relay cooperation is one of the most effective techniques to combat the ill-effects of wireless channels and boost the service range of cellular communication. The basic setup of relay cooperation, introduced as the relay channel in [16], is a widely studied topic which models the communication from a single source to a single destination aided by a single relay node.

The discrete memoryless relay channel (DM-RC) [16], depicted in Figure 2.7, and a code for such a channel are defined as follows:

Definition 2.13 (The DM-RC)**.** *The DM-RC is defined by the tuple* $(\mathcal{X}_1, \mathcal{X}_2, \mathcal{Y}_2, \mathcal{Y}_3, p_{Y_2,Y_3|X_1,X_2}(y_2, y_3|x_1, x_2))$*, where* $\mathcal{X}_1$ *and* $\mathcal{X}_2$ *are two finite, discrete input alphabets,* $\mathcal{Y}_2$ *and*

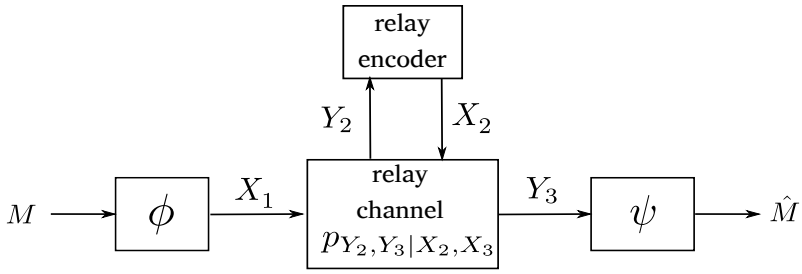


Figure 2.7: The discrete memoryless relay channel (DM-RC).

$\mathcal{Y}_3$ *are two finite, discrete output alphabets, and* $p_{Y_2,Y_3|X_1,X_2}(y_2, y_3|x_1, x_2)$ *is the channel transition law.*

Definition 2.14 (A code for the DM-RC)**.** *A* $(2^{nR}, n)$ *code for the DM-RC consists of (i) an uniformly distributed message set* $\mathcal{M} := \{1, 2, \ldots, 2^{nR}\}$*; (ii) and encoder for the source* $\phi : \mathcal{M} \to \mathcal{X}_1^n$*; (iii) a causal relay encoder* $\phi_{\mathsf{R}} : \mathcal{Y}_2^{i-1} \to \mathcal{X}_2$ *for each* $i \in \{1, 2, \ldots, n\}$*; and (iv) a decoder for the destination* $\psi : \mathcal{Y}_3^n \to \mathcal{M}$*.*

In essence, source S transmits a message $M \in \mathcal{M}$ via codeword $X^n(M)$, and due to the nature of conventional microwave band, this transmission is received by both relay R and destination D. After channel use $i \in \{1, 2, \ldots, n\}$, the relay encodes the information received in all previous time slots, i.e., Y_2^{i-1}, causally into a codeword symbol $X_{2,i}(Y_2^{i-1})$, and transmits it, thereby making the signal from the source and the relay correlated. The transmission from the source and the relay then superimpose at the destination, which then attempts to decode message M from its received signal.

While the relay channel has been subject to much investigation [16,17,114], its capacity region is still unknown in general. In [16], an upper bound to the rate was found using the *cutset* technique, and two important relaying schemes were introduced, the decode-and-forward (DF) scheme and the quantize-and-forward (QF) scheme. In DF relaying, the relay *decodes* the signal transmitted from the source, re-encodes it, and then transmits it to the destination. Due to such positive reinforcement of the transmitted signal, the overall transmission rate due to DF relaying improves if the quality of the source-relay link is good. Specifically, if the relay is located closer to the source, i.e., the resulting source-relay link is strong, the source message can be decoded at the relay without imposing additional constraint, thereby producing net gain in rate at the destination.

On the other hand, if the source-relay link is of bad quality, fully decoding the source signal at the relay imposes additional rate constraint, and hence DF relaying is not as useful

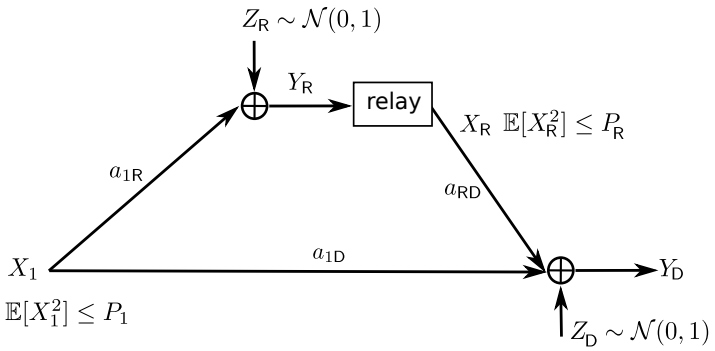


Figure 2.8: The Gaussian relay channel (GRC).

as in the former case. In such cases, quantize-and-forward (QF) relaying performs better, where the relay just compresses the received signal and forwards it to the destination, instead of trying to decode it fully.

The Gaussian relay channel (GRC) [16], depicted in Figure 2.8, is a wireless counterpart of the DM-RC. The channel output of the GRC at the i-th time-step are defined as

$$\begin{aligned} Y_{\mathsf{R}} &= a_{1\mathsf{R}} X_{1i} + Z_{\mathsf{R}i}, \\ Y_{\mathsf{D}} &= a_{1\mathsf{D}} X_{1i} + a_{\mathsf{RD}} X_{\mathsf{R}i} + Z_{\mathsf{D}i}, \quad i = 1, \ldots, n, \end{aligned}$$

where $a_{1\mathsf{R}}, a_{1\mathsf{D}}$ and a_{RD} are channel gains between the nodes as depicted in Figure 2.8, and $Z_{\mathsf{R}i} \sim \mathcal{N}(0,1)$, and $Z_{\mathsf{D}i} \sim \mathcal{N}(0,1)$ are independent noise samples.

Similar to the DM-RC, the capacity of the GRC is also unknown except for the special case of the degraded relay channel [16], where the signal received at the destination is assumed to be a degraded (i.e., noisier) version of the signal received at the relay. For degraded relay channels, capacity was shown to be achieved by the DF relaying scheme that uses correlated source and relay codewords.

Since the capacity of the relay channel is still unknown, considerable efforts have been directed towards characterizing approximate capacity results similar to the GIC. For the classic GRC, the DF relaying scheme was shown to achieve to within 1/2 bit of the capacity [17]. Similarly, in [114] it was shown that the compress-and-forward scheme achieves to within 1/2 bit of the capacity, whereas the simpler strategy of amplify-and-forward achieves within 1 bit of the capacity.

It is worth noting that typically a good achievable strategy for the relay channel will require the signals transmitted by the source and the relay to be correlated, and a part of the difficulty in obtaining capacity results lies in finding optimal distributions for the

codewords and optimal correlations between the source and relay codewords.

Multipath Fading and the Relay Channel

Relaying plays an important role in combating one of the most serious impediments in wireless communication: multipath fading. It is a small-scale propagation effect that arises when the transmitted signal reaches the receiver via multiple different propagation paths, each with distinct propagation delays, amplitudes, and phase shifts. Such indirect paths are formed, for example, by the reflection of the transmitted signal from surrounding obstacles such as buildings, cars or other urban features. When these multipath components combine at the receiver destructively, the receiver power is reduced to small levels, thereby causing *fading* that leads to failure of reliable communication [115–118].

Since relaying helps against multipath fading, fading relay channels have been subject to much investigation with the aim of characterizing optimal cooperation strategy and performance limits of such networks [13–15,119]. Two such fading models, namely Rayleigh and phase fading, have been studied in detail, as they are able to effectively model a number of communication impairments. For instance, Rayleigh fading effectively models the phenomena of multipath fading [120], whereas phase fading models the effects of phase uncertainties and synchronization errors in carrier oscillators [121].

In [77] the Gaussian relay channel subject to phase and Rayleigh fading was studied under the practical assumption of availability of channel state information (CSI) at the receiver. The capacity of this faded GRC was then characterized for the case when the source-relay links were *stronger* than the source-destination links in terms of data rate, or in other words, when the relay is *closer* to the source than the destination. Note that unlike [16], where the capacity of the GRC was derived under a degradedness assumption, the capacity result for the fading GRC in [77] does not need such an assumption. More specifically, due to the availability of CSI at receiver and the presence of phase or Rayleigh fading, *uncorrelated* source and relay signals were found to be optimal, resulting in a closed form capacity result.

Multiple Access Relay Channel

The multiple-access relay channel (MARC) is a natural generalization of the single source single destination relay channel of [16]. Instead of a single source, in the MARC multiple

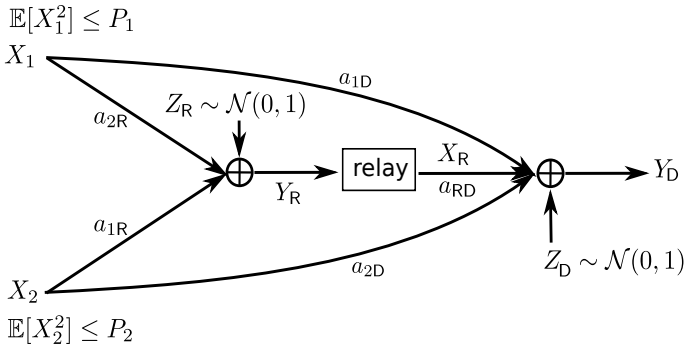


Figure 2.9: The Gaussian multiple-access relay channel (MARC).

sources communicate with a single destination as in a multiple access channel, while a single relay aids the communication of all users. Similar to the Gaussian relay channel, the Gaussian MARC models the MARC in wireless settings, and thus it has been studied extensively [18, 19, 122–124]. In Figure 2.9, the Gaussian MARC (G-MARC) with two sources is depicted.

Similar to the relay channel, the capacity of the G-MARC is also unknown in general, however some progress have been made in characterizing its performance limits. In [18], outer bounds to the capacity region of the G-MARC were characterized based on the cutset bounding technique, while an achievable rate region was derived after extending the DF relaying strategy of [16]. Similar to the degraded Gaussian relay channel in [16], the sum-capacity of the degraded G-MARC was characterized in [19].

Since the relay in the G-MARC is beneficial in combating the adverse effects of wireless fading as in the GRC, fading G-MARCs have been subject to investigation recently. In [124], the G-MARC subject to phase fading was studied, where conditions for joint source-channel communication over this channel were obtained. Moreover, similar to the fading GRC, the capacity region of the phase and Rayleigh faded G-MARC was also characterized under the assumption that the CSI is available at the receivers [77, 122, 123]. These results show that when the *near conditions* hold in the G-MARC, i.e., when the relay is located *near* the sources in the sense of [77, Theorem 9], its capacity region can be characterized in closed form.

Resource Allocation for the Multiple Access Relay Channel

Similar to other multi-user channels, such as those for the interference channels in [105, 111, 112], the overall performance of the MARC is typically characterized by optimizing a

design-specific goal such as those in [123, 125–127]. For instance, the sum-rate serves as an effective metric for the total throughput of the network, while the worst-user rate can serve as a fair user-rate in the channel. In [123], the optimal resource allocation for the ergodic fading orthogonal G-MARC was studied, where the relay is half-duplex and transmits and receives on orthogonal channels. Then, under the assumption of availability of CSI at all nodes, the optimal power allocations were derived that achieves the maximum sum-rate. In [127], a similar study was conducted on the G-MARC in an OFDMA framework, where several power and sub-carrier allocations were proposed, which were then shown to achieve close to optimal rates.

Connection to the Thesis

In Chapter 4, we study the dual-band fading Gaussian MARC, which operates over the integrated mm-wave/microwave dual-band. We consider a general ergodic fading, which specializes to phase and Rayleigh fading, and thus is consistent with similar works on the G-MARC in [77, 122, 123]. We characterize the fundamental performance limits of this dual-band MARC in terms of achievable rates and capacity results. We then study a resource allocation problem over the mm-wave links of this channel with the aim of optimizing the achievable sum-rate. The resulting power allocation scheme is shown to reveal useful insights for such dual-band networks.

Chapter 3

The Dual-Band Interference Channel

In this chapter[1], we study the performance of a two-source two-destination dual-band interference channel, denoted the dual-band IC. Here, a source communicates to its respective destination simultaneously over the microwave and the mm-wave bands. In the microwave band of this channel, the source-destination pairs mutually interfere, and thus form a conventional single-band Gaussian interference channel (GIC) [101].

The mm-wave channels, however, are operated as highly directional point-to-point links created by beamforming with co-phased antenna arrays [30]. Hence, a source in this band is well-modeled as being able to choose whether to transmit towards its desired destination only, the other destination only, or time-share between the two destinations. This raises the question of to which destination a source in the mm-wave band should transmit. In this dual-band IC, the transmitters in the mm-wave band can transmit in one of the following configurations: (a) the first source (S_1) to the first destination (D_1), and the second source (S_2) to the second destination (D_2); (b) S_1 to D_2, and S_2 to D_1; (c) both S_1 and S_2 to D_2; or (d) both S_1 and S_2 to D_1.

In this work, we focus on the first two cases. Therefore, the transmitters in the mm-wave band can either transmit (a) from S_1 to D_1, and from S_2 to D_2, via the S_1-D_1 and S_2-D_2 *direct-links*, (b) from S_1 to D_2, and from S_2 to D_1, via the S_1-D_2 and S_2-D_1 *cross-links*, or (c) divide the spectrum between the direct- and cross-modes via time/frequency division. The channel resulting from the first mode, where the sources transmit only in the direct-links, is denoted by the Direct-Link IC (DLIC), whereas the same for the second mode, where the sources transmit only in the cross-links, is denoted by the Cross-Link IC

[1]The results of this chapter (except for Section 3.3) have been published in [65] and [128].

(CLIC). Finally, in the third mode the sources share the spectrum between the two model, and hence the resulting channel is denoted by the Direct-and-Cross-Link IC (DCLIC).

In this chapter, we study the capacity of the DCLIC, the CLIC and several of its variations. Importantly, we first show that the capacity region of the DCLIC can be decomposed into the capacity region of the *underlying CLIC* and the set of direct-links. As such, the capacity of the DCLIC can then be established if the capacity of the underlying CLIC is characterized, and thus we focus solely on the CLIC thereon.

We observe that capacity of the CLIC depends largely on whether the destinations posses strong or weak interference in the microwave band. Based on this, we classify the CLIC in the following classes: (i) the strong CLIC, where both destinations have strong interference, (ii) the mixed CLIC where only one destination has strong interference while the other has weak interference, and (iii) the weak CLIC, where both destinations have weak interference. We then characterize sufficient channel conditions under which the capacity regions for the three classes of CLICs are established. Moreover, if these particular sufficient channel conditions do not hold, the capacity region of the CLICs, especially those of the the weak and mixed CLICs, are unknown. Therefore, we then characterize approximate capacity regions for them.

An important special case of the conventional microwave GIC is the Z-interference channel (ZIC) where one of the destinations do not suffer from interference, i.e., the microwave cross-link between from one source to this destination does not exist and taken to be zero [100]. We consider the dual-band counterpart of the ZIC, denoted Z-CLIC: here the underlying microwave GIC is a ZIC, and depending on whether both or one mm-wave cross-link(s) are present, we identify three kinds of Z-CLICs: Z-CLICs of type-0, type-1, and type-2, as detailed in the next section. We then characterize capacity and approximate capacity results for such channels.

For multi-user channels, it is common to characterize optimal resource allocation aimed at a specific design-metric, such as the sum-rate aimed at maximizing the total throughput, or the worst-user rate aimed at providing fair rates to all users, etc., which can serve as an effective tool for allocating resources in practice. For example, in the case of the interference channel, resource allocation schemes that maximize the sum-rate have been investigated thoroughly [105, 111, 112].

The DCLIC models a basic multiuser network over the microwave and mm-wave dual-band, and a similar resource allocation study will shed light on important trade-offs in

such networks. For example, in the dual-band IC the point-to-point mm-wave links can be used to either convey *desirable* user-information to the designated receiver, or forward *interference* information to the non-designated receiver. Neither of these two modes of transmission for the mm-wave links provides the optimal performance for all channel conditions, and thus there exists a non-trivial trade-off between the two modes of use for the mm-wave links. Since the mm-wave links will significantly affect the performance of the dual-band IC due to their point-to-point nature and relatively higher bandwidth, it is useful to understand how to optimize the performance of the DCLIC over the parameters in the mm-wave band. We therefore study the power allocation scheme over the four mm-wave links that maximizes the sum-rate of the channel. The optimal power allocation scheme, derived in closed form, possesses rich properties and reveals useful insights.

The contributions of this chapter are summarized as follows:

Capacity Results (presented in Section 3.2)

- Decomposition of the capacity of the DCLIC: We show that the capacity region of the DCLIC can be decomposed into the capacity region of the *underlying CLIC* and the two *direct-links* in the mm-wave band. As such, the direct-links improve the rates of individual users, whereas the cross-links play a non-trivial role in characterizing the capacity of the CLIC.

- Strong CLIC: The capacity of the strong CLIC is characterized, and the strong interference condition in the microwave band is shown to be sufficient to characterize the capacity.

- Weak CLIC: We characterize combined weak-very strong channel conditions over the cross-channel gains in the microwave and the mm-wave bands, under which its capacity is established. This condition shows that even if the GIC in the microwave band has weak interference, the mm-wave cross-links with adequately strong interference are able to drive the combined interference to a very strong regime, thus producing the capacity result.

- Mixed CLIC: We characterize the capacity of the mixed CLIC under a set of channel conditions. This condition, which holds over channel gains of the microwave and mm-wave links, can be interpreted as being able to drive the mixed interference regime to a strong one, thus characterizing the capacity.

- Z-CLIC: We identify three kinds of Z-CLICs, denoted as type-0, type-1, and type-2, depending on the presence of mm-wave cross-links. We first show that the capacity of the Z-CLIC of type-0 is the same as that of the Z-CLIC of type-1.

  Then the capacity of the Z-CLIC of type-1 is characterized under two sets of channel conditions, which can interpreted as being combined strong and very strong interference conditions over both bands.

  Finally, the capacity of the Z-CLIC of type-2 was shown to be the same as that of the Z-IC: the Z-IC is an one-sided IC operating in the microwave band only, and its capacity results are well known in literature.

Approximate Capacity Results (presented in Section 3.3)

- Weak CLIC: We obtain a constant gap result which characterizes the capacity of the weak CLIC within 1/2 bit/channel use per user for the entire weak interference regime.
- Mixed CLIC: We characterize the capacity region of the mixed CLIC within 1/2 bit/channel use per user which is valid for the entire regime of the mixed interference.
- Z-CLIC of type-1: For this model, we derive a constant gap result valid for the entire regime that characterizes its capacity within 1/2 bit/channel use per user.

Resource Allocation for the DCLIC (presented in Section 3.4)

We derive the optimal power allocation scheme over the mm-wave direct-links and cross-links of the DCLIC that maximizes the sum-rate of channel. For notational convenience, this optimal power allocation scheme for the the DCLIC is denoted by the IC-OPA. We observe that depending on the value of the power budget (P) and certain channel conditions, power is either allocated to a only specific subset of the 4 links, or is shared between all links. This behavior is equivalent to partitioning the entire range of the power budget P into several *link-gain regimes* (LGR), such that in each LGR, the IC-OPA allocates power to the links in distinct *modes*. We characterize all such LGRs which reveal useful insights. Specifically, we observe the following:

- The IC-OPA allocates the power budget among (P) the direct- and cross-links following two properties: a *waterfilling-like* property and a *saturation* property.

- When P is sufficiently small, the IC-OPA assigns power to only a subset of the links; the power in the allocated links increase piece-wise linearly as P increases, while the power in the remaining links are zero.

- Due to the saturation property, the IC-OPA imposes a maximum limit on the cross-link powers. When P exceeds a certain *saturation threshold* ($\mathsf{P}_{\mathsf{sat}}$), the cross-link powers reach their peak value, and all additional increments to P are added only to the direct-links that do not have such limits.

- If the underlying GIC in the microwave band has very strong interference, the IC-OPA assigns the power budget entirely to the direct-links.

- If the channel parameters satisfy one of the following criteria, then transmitting only in the direct-links is approximately optimal, in the sense that the difference between the sum-rates resulting from allocating to only direct-links and that from allocating optimally in all links is negligibly small: (a) the transmit powers in the underlying GIC in the microwave band is small; (b) the mm-wave cross-link gains are large.

- When the mm-wave bandwidth is *large*, the optimal power allocation for the symmetric case simplifies as follows: (a) if the direct-links are stronger than the cross-links, allocating the budget P entirely to the direct-links is optimal for all $P \geq 0$; (b) alternatively, if the cross-links are stronger, for P smaller than the saturation threshold, allocating P entirely to the cross-links is optimal, while for P larger than the threshold, the cross-links are saturated, and all increments of P are allotted to the direct-links.

The rest of the chapter is organized as follows: We define the system models in Section 3.1. We present all capacity results in Section 3.2. In Section 3.3, all constant gap results are presented. The resource allocation problem for the DCLIC is presented in Section 3.4. Finally, a summary of results is provided in Section 3.5.

3.1 System Model

We now define the system models of all the channels studied in this chapter. First, we define the discrete memoryless (DM) DCLIC, depicted in Figure 3.1, from which the Gaussian DCLIC is defined. We thus establish the notations needed to define the other models.

In the dual-band IC, the number of channel uses in the microwave band and the mm-wave band will likely be different as the mm-wave band has more bandwidth than the microwave band. We model this bandwidth difference with a *bandwidth mismatch factor* (BMF) α between the first (microwave) band and the second (mm-wave) band: during n accesses of the first band, the second band is accessed $\hat{n}(n) := \lfloor \alpha n \rfloor$ times in total, where $\lfloor x \rfloor$ denotes the largest integer no larger than x.

Note that a transmitter in the mm-wave band is assumed to be capable of creating either a direct-link or a cross-link by beamforming to the relevant destination via a co-phased antenna-array. Since the two links must be time-shared, the total BMF of α for the mm-wave band needs to be divided into two parts, α_1 for the cross-links and α_2 for the direct-links, with $\alpha = \alpha_1 + \alpha_2$. More specifically, during n accesses of the first band, the second band is accessed $n_1(n) := \lfloor \alpha_1 n \rfloor$ times as cross-links, and $n_2(n) := \lfloor \alpha_2 n \rfloor$ times as direct-links. For notational convenience, throughput this chapter we denote $n_1(n)$ and $n_2(n)$ by n_1 and n_2 respectively.

In an alternative setting where a mm-wave transmitter uses a *hybrid* antenna array instead of a single co-phased antenna arrays, in principle, available antenna elements in the hybrid antenna array can be reconfigured into two smaller and independent antenna arrays, and then be used to communicate to both destinations *simultaneously* [72–74], as discussed in Section 1.3. Hence, the direct- and cross-links from a mm-wave transmitter can then be realized simultaneously without any time-sharing, and in essence, both the direct- and cross-links are able to utilize the full BMF α. This setting can be regarded as a special case of the setting considered in this chapter by doubling the BMF α and then allocating equal bandwidths for the direct- and cross-links, i.e., $\alpha_1 = \alpha_2$, with a suitable reduction in the path-loss to account for the reduced array gain.

The discrete memoryless DCLIC

Definition 3.1 (The DM-DCLIC)**.** *The DM-DCLIC is defined by the tuple* $((\mathcal{X}_k, \hat{\mathcal{X}}_k, \bar{\mathcal{X}}_k)_{k=1}^2, (\mathcal{Y}_k, \hat{\mathcal{Y}}_k, \bar{\mathcal{Y}}_k)_{k=1}^2, p(.))$*, where* $\mathcal{X}_k$ *and* $\mathcal{Y}_k$ *are finite, discrete input and output alphabets of the interference channel in the first band,* $\bar{\mathcal{X}}_k$ *and* $\bar{\mathcal{Y}}_k$ *are finite, discrete input and output alphabets for the* S_k *to* D_k *direct-links in the second band,* $k \in \{1, 2\}$*,* $\hat{\mathcal{X}}_k$ *and* $\hat{\mathcal{Y}}_\ell$ *are finite, discrete input and output alphabets for the* S_k *to* D_ℓ *cross-link in the second band,* $k \neq \ell \in \{1, 2\}$*, and* $p(.)$ *is the channel transition law that decomposes as follows*

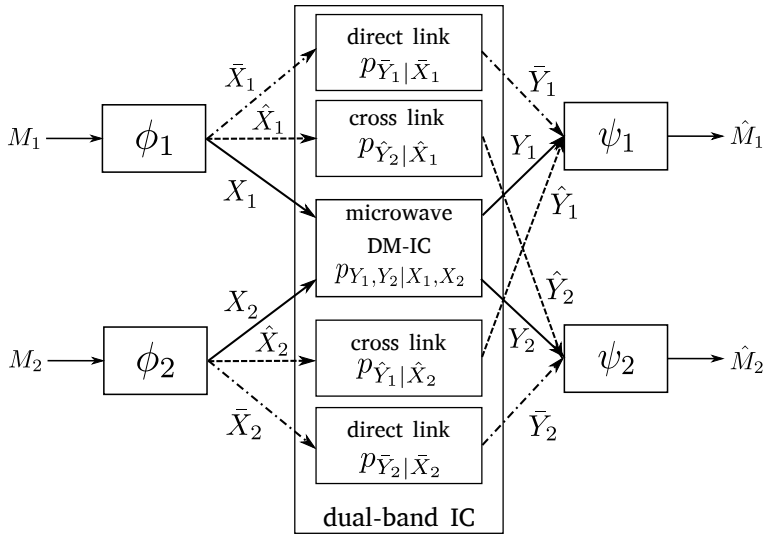


Figure 3.1: System model of the discrete memoryless DCLIC. It consists of an underlying DM-IC in the microwave band and the set of direct- and cross-links in the mm-wave band.

$$
\begin{aligned}
&p(y_1^n, y_2^n, \hat{y}_1^{n_1}, \hat{y}_2^{n_1}, \bar{y}_1^{n_2}, \bar{y}_2^{n_2} | x_1^n, x_2^n, \hat{x}_1^{n_1}, \hat{x}_2^{n_1}, \bar{x}_1^{n_1}, \bar{x}_2^{n_1}) \\
&= \prod_{\ell=1}^{n} p(y_{1\ell}, y_{2\ell} | x_{1\ell}, x_{2\ell}) \prod_{\ell=1}^{n_1} p(\hat{y}_{1\ell} | \hat{x}_{2\ell}) \prod_{\ell=1}^{n_1} p(\hat{y}_{2\ell} | \hat{x}_{1\ell}) \prod_{\ell=1}^{n_2} p(\bar{y}_{1\ell} | \bar{x}_{1\ell}) \prod_{\ell=1}^{n_2} p(\bar{y}_{2\ell} | \bar{x}_{2\ell}).
\end{aligned}
$$

We now define a code and an achievable rate pair for the DM-DCLIC.

Definition 3.2 (A code for the DM-DCLIC)**.** *A* $(2^{nR_1}, 2^{nR_2}, n, \alpha_1, \alpha_2)$ *code for the DCLIC to consist of (i) two independent, uniformly distributed message sets* $\mathcal{M}_k := \{1, 2, \ldots, 2^{nR_k}\}$, $k = 1, 2$, *one each for sources* S_1 *and* S_2*; (ii) two encoders* $\phi_k : \mathcal{M}_k \to \mathcal{X}_k^n \times \hat{\mathcal{X}}_k^{n_1} \times \bar{\mathcal{X}}_k^{n_2}$, $k = 1, 2$, *one each for* S_1 *and* S_2*; and (iii) two decoders* $\psi_k : \mathcal{Y}_k^n \times \hat{\mathcal{Y}}_k^{n_1} \times \bar{\mathcal{Y}}_k^{n_2} \to \mathcal{M}_k, k = 1, 2$, *one each for destinations* D_1 *and* D_2*, with* $n_k \leq \alpha_k n$*, and* $\alpha_k = \lim_{n\to\infty} n_k/n, k = 1, 2$.

As depicted in Figure 3.1, S_1 and S_2 communicate messages $M_1 \in \mathcal{M}_1$ and $M_2 \in \mathcal{M}_2$ by encoding them into codewords $\phi_k(M_k) = (X_k^n(M_k), \hat{X}_k^{n_1}(M_k), \bar{X}_k^{n_2}(M_k)), k = 1, 2$, for transmission. Then, $(X_1^n(M_1), X_2^n(M_2))$ are transmitted through the first band which forms a DMIC. Sender S_1 forwards the "interference information" to D_2 by transmitting $\hat{X}_1^{n_1}(M_1)$ via the S_1-D_2 cross-link in the second band, whereas $\hat{X}_2^{n_2}(M_2)$ is sent through the S_2-D_1 cross-link for interference forwarding from S_2. Finally, S_1 sends $\bar{X}_1^{n_2}(M_1)$ to D_1 via the S_1-D_1 direct-link in the second band to communicates the "desired information", while S_2 does the same via the S_2-D_2 direct-link.

The decoders estimate the input messages from the received signals. As such, the

decoding probability of error for the code is defined as

$$P_e^{(n)} = \Pr\left[\psi_1(Y_1^n, \hat{Y}_1^{n_1}, \bar{Y}_1^{n_2}) \neq M_1 \quad \cup \quad \psi_2(Y_2^n, \hat{Y}_2^{n_1}, \bar{Y}_2^{n_2}) \neq M_2\right],$$

where the average is taken over uniform distribution of $(M_1, M_2) \in \mathcal{M}_1 \times \mathcal{M}_2$.

Definition 3.3 (Achievable rate for the DMIC)**.** *A rate pair (R_1, R_2) is said to be achievable for the DM-DCLIC if there exists a sequence of $(2^{nR_1}, 2^{nR_2}, n, \alpha_1, \alpha_2)$ codes such that $n_1 \leq n\alpha_1, n_2 \leq n\alpha_2$, and $P_e^{(n)} \to 0$, as $n, n_1, n_2 \to \infty$.*

Finally, the capacity region of the DM-DCLIC is defined as the closure of the set of all achievable rate tuples.

Note that in the special case of BMF $\alpha_1 = \alpha_2 = 1$, i.e., where the microwave IC, the cross-links and the direct-links all have the same bandwidth, the DM-DCLIC can be modeled as an *equivalent* discrete memoryless interference channel where the users can communicate with an *augmented* alphabet. More specifically, this equivalent channel consists of 3 *parallel* sub-channels such that the transmissions in the 3 sub-channels are modeled as a conventional DMIC, a set of cross-links only, and a set of direct-links only, respectively. Since the BMF is 1, the codeword lengths for all sub-channels are the same. Thus the users can use an *augmented* alphabet (e.g., the Cartesian product of the 3 alphabets for the 3 sub-channels) to communicate over this equivalent channel.

However, for the general case of $\alpha_1 \neq 1, \alpha_2 \neq 1$, modeling the DM-DCLIC as a single equivalent channel as described above is not possible. When $\alpha_1, \alpha_2 \geq 1$ the DM-DCLIC can instead be modeled as a *state-dependent* interference channel where (a) in the first state, the channel consists of the single equivalent interference channel with 3 parallel sub-channels described in the paragraph above, (b) in the second state, only the cross-links are present, and (c) in the third state only the direct-links are present. The general BMF $\alpha_1, \alpha_2 \geq 1$ can then be realized by using 3 states of the channel appropriate number of times.

The Gaussian Models

We now define the Gaussian counterpart of the discrete memoryless DCLIC, and few of its variations studied here. In this chapter, we will focus on these Gaussian models, and thus reference to a channel model should be assumed to be a reference to the *Gaussian* model by default, unless otherwise specified. For example, by the DCLIC we refer to the

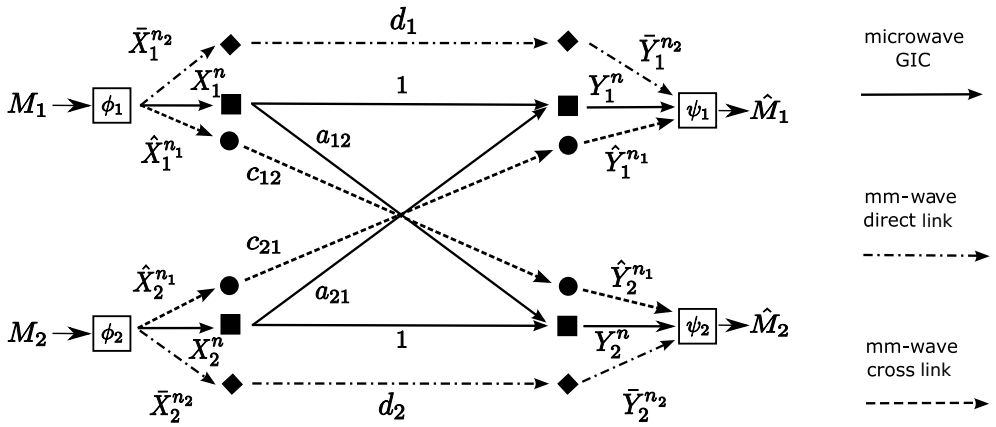


Figure 3.2: System model of the Gaussian DCLIC. It consists of an underlying GIC in the microwave band and the set of direct-links and cross-links in the mm-wave band.

Gaussian DCLIC, and the discrete memoryless DCLIC is specifically referred to as the DM-DCLIC.

The Gaussian DCLIC

The Gaussian DCLIC (denoted the DCLIC), depicted in Figure 3.2, consists of a GIC in the underlying microwave band and the set of direct- and cross-links in the mm-wave band. The GIC in the first band is modeled as in [101], where the signals received at D_1 and D_2 at the i^{th} channel use are given by

$$Y_{1i} = X_{1i} + a_{21}X_{2i} + Z_{1i}, \tag{3.1}$$

$$Y_{2i} = X_{2i} + a_{12}X_{1i} + Z_{2i}, \quad i = 1, \ldots, n. \tag{3.2}$$

Here, $X_{ki}, Y_{ki} \in \mathbb{R}$, $a_{k\ell} \in \mathbb{R}$ are cross-channel coefficients from S_k to D_ℓ, $k \neq \ell \in \{1, 2\}$ (a_{kk}^2 are normalized to 1 as in [101]), and $Z_{ki} \sim \mathcal{N}(0,\ 1)$, i.i.d., are noise samples. In addition, the codewords now satisfy the average power constraint, $\frac{1}{n}\sum_{i=1}^{n} X_{ki}^2 \leq P_k, k = 1, 2$.

The cross-links in the second band are point-to-point, and are modeled as

$$\hat{Y}_{2i} = c_{12}\hat{X}_{1i} + \hat{Z}_{2i}, \tag{3.3}$$

$$\hat{Y}_{1i} = c_{21}\hat{X}_{2i} + \hat{Z}_{1i}, \quad i = 1, \ldots, n_1, \tag{3.4}$$

where $\hat{X}_{ki}, \hat{Y}_{ki} \in \mathbb{R}$, $c_{k\ell} \in \mathbb{R}$ are coefficients of the S_k-D_ℓ cross-links, $k \neq \ell \in \{1, 2\}$, and $\hat{Z}_{ki} \sim \mathcal{N}(0,\ 1)$, i.i.d., are noise samples. The codewords satisfy the average power constraint, $\frac{1}{n_1}\sum_{i=1}^{n_1} \hat{X}_{ki}^2 \leq \hat{P}_k, k = 1, 2$. The direct-links are similarly modeled as

$$\bar{Y}_{1i} = d_1\bar{X}_{1i} + \bar{Z}_{1i}, \tag{3.5}$$

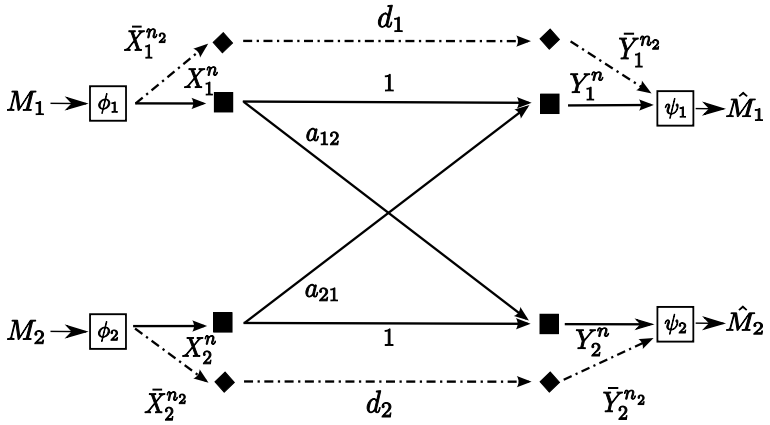


Figure 3.3: System model of the Gaussian DLIC. It consists of an underlying GIC in the microwave band and only the set of direct-links in the mm-wave band.

$$\bar{Y}_{2i} = d_2 \bar{X}_{2i} + \bar{Z}_{2i}, \quad \ell = 1, \ldots, n_2, \tag{3.6}$$

where $\bar{X}_{ki}, \bar{Y}_{ki} \in \mathbb{R}$, $d_k \in \mathbb{R}$ are the direct-link coefficients, $\bar{Z}_{k\ell} \sim \mathcal{N}(0,1)$, i.i.d., are noise samples, and codewords satisfy the average power constraint, $\frac{1}{n_2}\sum_{i=1}^{n_2} \bar{X}_{ki}^2 \leq \bar{P}_k, k = 1, 2$.

A $(2^{nR_1}, 2^{nR_2}, n, \alpha_1, \alpha_2)$ code and an achievable rate pair for the DCLIC are defined from those of the DM-DCLIC given in Def. 3.1 and Def. 3.2, respectively, by choosing all codeword alphabets to be $\mathbb{R}$ and imposing average power constraints on the codewords $X_k^n, \hat{X}_k^{n_1}$ and $\bar{X}_k^{n_2}, k = 1, 2$, as defined above.

Note that for the special case of BMF $\alpha_1 = \alpha_2 = 1$, i.e., where the microwave IC, the cross-links and the direct-links all have the same bandwidth, the DCLIC can be considered a special case of the parallel Gaussian IC [110] and the vector/MIMO gaussian IC [129].

The Gaussian DLIC

The Gaussian DLIC (denoted the DLIC) is defined from the DCLIC by taking $\hat{\mathcal{X}}_k = \hat{\mathcal{Y}}_k = \emptyset, k = 1, 2$, i.e., by deleting the two cross-links. The DLIC is depicted in Figure 3.3. Hence, the DLIC is defined by (3.1)–(3.2) and (3.5)–(3.6). The definition of a code and an achievable pair of the DLIC follow from those of the DCLIC after setting $\hat{\mathcal{X}}_k = \hat{\mathcal{Y}}_k = \emptyset, k = 1, 2$, hence not repeated here.

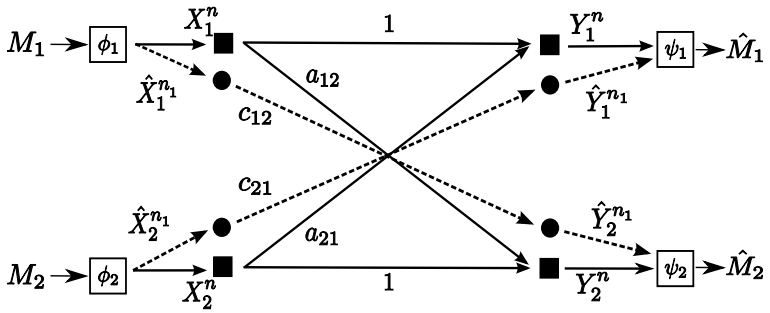


Figure 3.4: System model of the Gaussian CLIC. It consists of an underlying GIC in the microwave band and only the set of cross-links in the mm-wave band.

The Gaussian CLIC

The Gaussian CLIC (denoted the CLIC), depicted in Figure 3.4, is defined from the DCLIC by taking $\bar{\mathcal{X}}_k = \bar{\mathcal{Y}}_k = \emptyset, k = 1, 2$, i.e., by deleting the two direct-links. Hence, the CLIC is defined by (3.1)–(3.4). The definition of a code and an achievable pair of the CLIC follows from those of the DCLIC, hence not repeated.

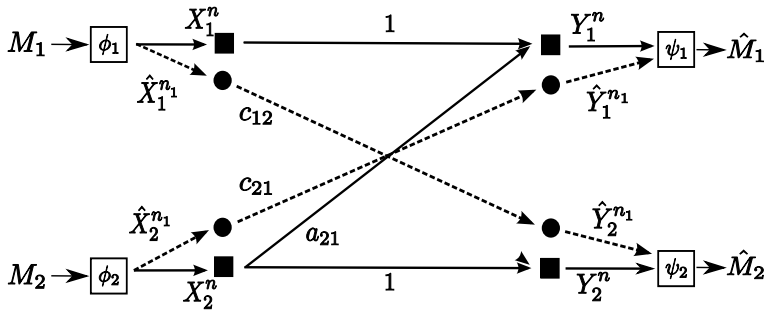
For ease of exposition, the CLIC is classified into 3 classes based on the interference of the underlying GIC with cross-channel gains a_{12}^2 and a_{21}^2. A CLIC is said to be a

1. *strong* CLIC, if $a_{12}^2 \geq 1$ and $a_{21}^2 \geq 1$, i.e., the underlying GIC is strong;
2. *weak* CLIC, if $a_{12}^2 < 1$ and $a_{21}^2 < 1$, i.e., the underlying GIC is weak;
3. *mixed* CLIC, if $a_{12}^2 < 1$ and $a_{21}^2 \geq 1$, i.e., the underlying GIC is mixed.

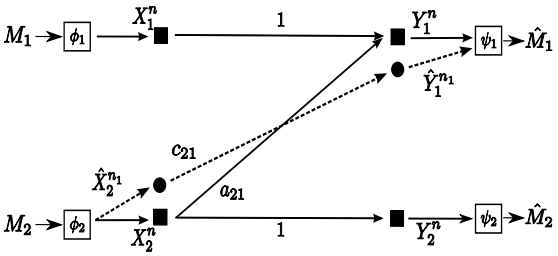
Note that the mixed CLIC can also be defined under the alternative condition $a_{12}^2 \geq 1$ and $a_{21}^2 < 1$, i.e., the role of the strong and weak destinations are exchanged. The results for the latter mixed CLIC can be trivially obtained from those for the former mixed CLIC, and hence not repeated.

The Gaussian Z-CLIC

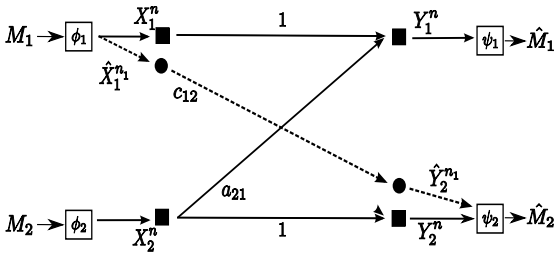
The Z-interference channel (ZIC) is an important subclass of the conventional GIC, where only one of the two destinations is affected by interference, while the other operates in an interference-free manner. For example, if D_2 in the GIC is unaffected by interference from S_1, the resulting ZIC can be regarded as a special case of the GIC after taking the corresponding cross-channel gain $a_{12} = 0$ and keeping all other links intact.



(a) The Gaussian Z-CLIC of type-0



(b) The Gaussian Z-CLIC of type-1



(c) The Gaussian Z-CLIC of type-2

Figure 3.5: (a) System model of the Gaussian Z-CLIC of type-0. In the underlying microwave ZIC, we have $a_{12} = 0$, but both mm-wave cross-links are present.

In this chapter, we consider the dual-band counterpart of the ZIC, denoted the Z-CLIC, where the underlying microwave GIC is a ZIC with $a_{12} = 0$ and $a_{21} \neq 0$. In addition to the microwave ZIC, the Z-CLIC also consists of two mm-wave cross-links, and based on which of the cross-links are present, the Z-CLIC can be classified into 3 cases: (i) Z-CLIC of type-0, depicted in Figure 3.5a, where both mm-wave cross-links are present; (ii) Z-CLIC of type-1, depicted in Figure 3.5b, where the mm-wave S_1-D_2 cross-link is removed, i.e., $c_{12} = 0$; and (iii) Z-CLIC of type-2, depicted in Figure 3.5c, where the mm-wave S_2-D_1 cross-link is removed, i.e., $c_{21} = 0$.

Note that while the Z-CLIC and the ZIC considered in this chapter has $a_{12} = 0$ and $a_{21} \neq 0$ in the underlying GIC, the alternate version of the Z-CLIC and the ZIC can be obtained by taking $a_{12} \neq 0$ and $a_{21} = 0$. However, the capacity results for this alternative setting follows trivially from those presented here by exchanging the roles of the sources and the destinations, and hence are not repeated.

Note that while the results included in this chapter are for the dual-band interference channel with *real* channel coefficients, these results can be extended to the case with *complex* channel coefficients following similar techniques.

3.2 Capacity Results

3.2.1 Decomposition of the Capacity of the DCLIC

Recall that in the DCLIC with BMFs α_1 and α_2, for n channel uses in the first band, the second band is used $n_1 = \lfloor \alpha_1 n \rfloor$ times as cross-links and $n_2 = \lfloor \alpha_2 n \rfloor$ times as direct-links. We show below that the capacity of the DCLIC can be decomposed into the capacity of the underlying CLIC, complemented by the direct-links that are used to transmit individual user information to their respective receivers.

In the following theorem, we present this decomposition result.

Theorem 3.1. *The capacity region of the DCLIC with BMFs α_1 and α_2 is given by the set of all non-negative rate tuples (R_1, R_2) that satisfy the decomposition*

$$\begin{aligned} R_1 &\leq r_1 + \alpha_2 \mathsf{C}\left(d_1^2 \bar{P}_1\right) \\ R_2 &\leq r_2 + \alpha_2 \mathsf{C}\left(d_2^2 \bar{P}_2\right), \end{aligned} \tag{3.7}$$

where (r_1, r_2) is an achievable rate tuple in the underlying CLIC with BMF α_1.

Proof. The proof is relegated to Appendix B.1. ■

Theorem 3.1 shows that an achievable rate pair in the DCLIC can be equivalently represented by a rate pair achievable in the underlying CLIC with BMF α_1, complemented by the capacity of the direct-links, $\left(\alpha_2 \mathsf{C}\left(d_1^2 \bar{P}_1\right), \alpha_2 \mathsf{C}\left(d_2^2 \bar{P}_2\right)\right)$, as depicted in Figure 3.6. Hence, the capacity of the DCLIC can be characterized from that of the underlying CLIC. Hence, hereon we focus on the CLIC instead of the DCLIC. In addition, this decomposition result shows that it is optimal to deploy the encoding/decoding for the direct-links *separately* from those of the CLIC. This clearly simplifies the encoding/decoding operation for the channel.

Decomposition of the Capacity of the DLIC

Before delving into the CLIC, we first consider the DLIC, depicted in Figure 3.3, which consists of an underlying GIC in the microwave band, and the two direct-links in the mm-wave band.

Similar to the DCLIC, the DLIC also admits a capacity decomposition as follows:

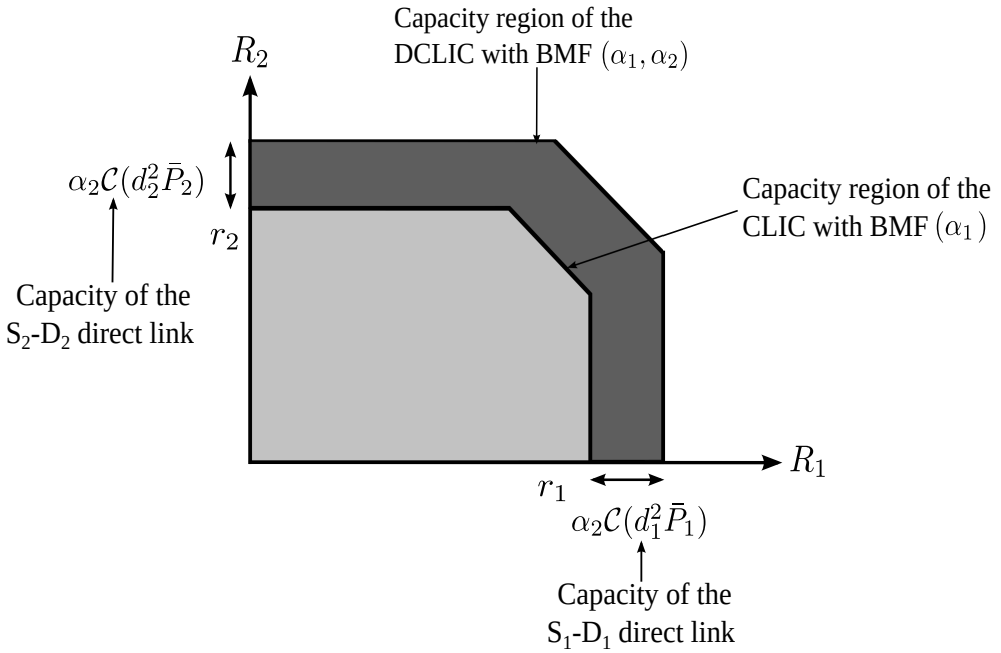


Figure 3.6: Decomposition of the Capacity of the DCLIC.

Lemma 3.1. *The capacity region of the DLIC with BMF α_2 is given by the set of all non-negative rate tuples (R_1, R_2) that satisfy the decomposition*

$$\begin{aligned} R_1 &\leq r_1 + \alpha_2 \mathsf{C}\left(d_1^2 \bar{P}_1\right) \\ R_2 &\leq r_2 + \alpha_2 \mathsf{C}\left(d_2^2 \bar{P}_2\right), \end{aligned} \tag{3.8}$$

where (r_1, r_2) is an achievable rate tuple in the underlying GIC in the microwave band.

The proof of this lemma follows that of Theorem 3.1, and hence is omitted. Similar to the result for the DCLIC, this result shows that a rate pair in the DLIC can be achieved by a rate pair achievable in the underlying microwave GIC, and then complemented by the direct-links. Therefore, the capacity of the DLIC can be found easily for cases in which a capacity result exists for the conventional GIC. For example, the capacity of the DLIC can be characterized when the underlying GIC has strong interference [100], whereas only sum-capacity results for the DLIC can be characterized if the underlying GIC has mixed [100] or noisy interference [108].

3.2.2 The Strong CLIC

Hereon, we focus entirely on the CLIC and its variants. In the strong CLIC, where the underlying GIC in the first band has strong interference (i.e., $a_{12}^2 \geq 1$ and $a_{21}^2 \geq 1$), the

capacity region is given below.

Lemma 3.2 (The strong interference case for the CLIC)**.** *The capacity region of the strong CLIC with BMF α_1 is given by the set of all non-negative rate tuples (R_1, R_2) that satisfy*

$$R_1 \leq \mathsf{C}(P_1)\,, \tag{3.9}$$

$$R_2 \leq \mathsf{C}(P_2)\,, \tag{3.10}$$

$$R_1 + R_2 \leq \mathsf{C}\big(P_1 + a_{21}^2 P_2\big) + \alpha_1 \mathsf{C}\Big(c_{21}^2 \hat{P}_2\Big)\,, \tag{3.11}$$

$$R_1 + R_2 \leq \mathsf{C}\big(P_2 + a_{12}^2 P_1\big) + \alpha_1 \mathsf{C}\Big(c_{12}^2 \hat{P}_1\Big)\,. \tag{3.12}$$

The proof of Lemma 3.2 follows from the proof for the strong interference case in [110] in a straightforward manner. Hence, we omit the proof and discuss only the key idea. In the strong CLIC, the GIC in the first band has strong interference. Additionally, in the second band, the cross-link gains are positive while the direct-link gains are zero, which results in a GIC with strong interference. Hence, the strong CLIC is a parallel GIC with strong interference in both bands, for which the capacity results is found in [110]. As such, the outer bound to the capacity region is characterized in a single-letter form under the sufficient condition $a_{12}^2 \geq 1$ and $a_{21}^2 \geq 1$. For the achievability part, each message is encoded *jointly* over both bands using Gaussian codewords, and each destination decodes messages from both transmitters in a multi-access channel fashion.

Since the capacity of the strong CLIC with BMF α_1 is characterized in closed form in Lemma 3.2, the capacity of the DCLIC with an underlying strong CLIC, which follows from Theorem 3.1, is also found in closed form as below

$$R_1 \leq \mathsf{C}(P_1) + \alpha_2 \mathsf{C}\big(d_1^2 \bar{P}_1\big) \tag{3.13}$$

$$R_2 \leq \mathsf{C}(P_2) + \alpha_2 \mathsf{C}\big(d_2^2 \bar{P}_2\big) \tag{3.14}$$

$$R_1 + R_2 \leq \mathsf{C}\big(P_1 + a_{21}^2 P_2\big) + \alpha_1 \mathsf{C}\Big(c_{21}^2 \hat{P}_2\Big) + \alpha_2 \mathsf{C}\big(d_1^2 \bar{P}_1\big) + \alpha_2 \mathsf{C}\big(d_2^2 \bar{P}_2\big) \tag{3.15}$$

$$R_1 + R_2 \leq \mathsf{C}\big(P_2 + a_{12}^2 P_1\big) + \alpha_1 \mathsf{C}\Big(c_{12}^2 \hat{P}_1\Big) + \alpha_2 \mathsf{C}\big(d_1^2 \bar{P}_1\big) + \alpha_2 \mathsf{C}\big(d_2^2 \bar{P}_2\big)\,. \tag{3.16}$$

While this result is trivial, expression of the rate region of the DCLIC will prove useful in Section 3.4.

3.2.3 The Weak CLIC

As defined in Section 3.1, in the weak CLIC, the underlying GIC has weak interference, i.e., $a_{12}^2 < 1$ and $a_{21}^2 < 1$. Hence, unlike the strong CLIC, where having strong cross-channel

gains in the underlying microwave band was sufficient to characterize the capacity, in the weak case additional conditions on the channel parameters of both bands are needed to characterize the capacity.

We say that the weak CLIC satisfies the combined weak-very strong interference condition, if the channel conditions in (3.17)-(3.18) hold. For such CLICs, we have the following capacity result

Theorem 3.2 (The combined weak-very strong interference case for the CLIC)**.** *If the channel parameters of the CLIC with BMF α_1 satisfy $a_{12}^2 < 1$ and $a_{21}^2 < 1$, as well as the conditions*

$$(1+P_1)(1+P_2) \leq (1+P_2+a_{12}^2 P_1)(1+c_{12}^2 \hat{P}_1)^{\alpha_1} \tag{3.17}$$

$$(1+P_1)(1+P_2) \leq (1+P_1+a_{21}^2 P_2)(1+c_{21}^2 \hat{P}_2)^{\alpha_1}, \tag{3.18}$$

then the capacity of the CLIC is given by the set of non-negative tuples (R_1, R_2) that satisfy

$$R_1 \leq \mathsf{C}(P_1) \tag{3.19}$$

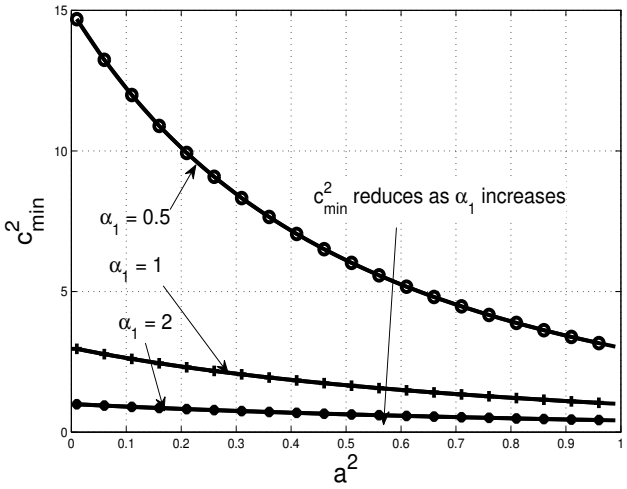
$$R_2 \leq \mathsf{C}(P_2)\,. \tag{3.20}$$

Proof. The proof, which follows along the lines of the proof for the ergodic very strong interference result in [110], is relegated to Appendix B.2. ■

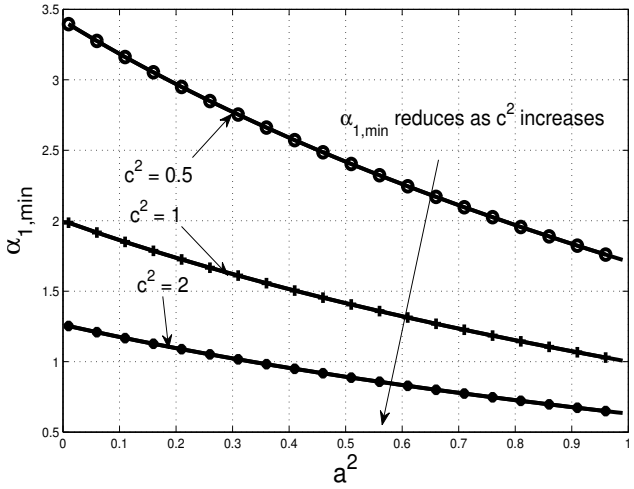
In this case, the capacity is obtained by first decoding the interfering message, removing its effect from the receives signal, and then decoding the desired message. Here, although the GIC in the first band is weak, under conditions (3.17)-(3.18), the cross-links in the second band are *sufficiently strong* in the sense that the *combined interference* forwarded through the cross-channels in both bands are enough to drive the CLIC to a very strong interference regime. Once the CLIC has very strong interference, it is optimal to first decode the interfering message by treating the desired user as noise, and then decode the desired message interference free.

Note that for the special case of BMF $\alpha_1 = \alpha_2 = 1$, i.e., where the microwave IC, the cross-links and the direct-links all have the same bandwidth, the capacity results for the strong CLIC and the weak CLIC can be recovered from the strong and the very strong interference cases of the parallel Gaussian IC [110] and the vector/MIMO gaussian IC [129].

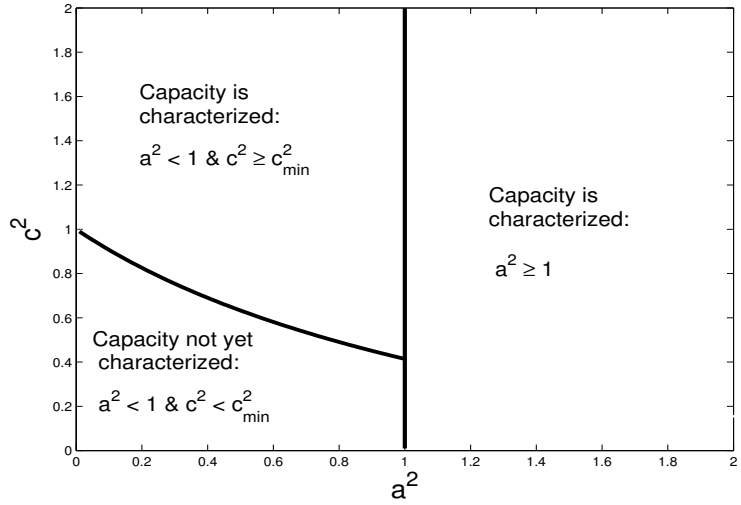
To illustrate the relationship between the channel parameters that satisfy the combined weak-very strong interference condition (3.17)-(3.18), we consider a symmetric weak CLIC



(a) c^2_{min} vs a^2



(b) $\alpha_{1,\mathsf{min}}$ vs a^2



(c) Partitioning of the set of (a^2, c^2) of a symmetric CLIC.

Figure 3.7: (a) The plot of c^2_{min} against a^2 for $\alpha_1 \in \{0.5, 1, 2\}$. (b) The plot of c^2_{min} against $\alpha_{1,\mathsf{min}}$ for $c^2 \in \{0.5, 1, 2\}$. (c) Partitioning of the set of channel gains (a^2, c^2) of a symmetric CLIC based on whether its capacity region has been characterized.

where the cross-channel gains as well as the transmit powers of the two sources are the same: in the microwave band, we have $a^2 := a^2_{12} = a^2_{21}$ and $P := P_1 = P_2$, and in the mm-wave band, we have $c^2 := c^2_{12} = c^2_{21}$ and $\hat{P} := \hat{P}_1 = \hat{P}_2$. For such a CLIC, the two conditions (3.17) and (3.18) coincide. For ease of exposition, we define

$$c^2_{\mathsf{min}} := \frac{1}{\hat{P}} \left(\left(\frac{(1+P)^2}{1+a^2 P} \right)^{1/\alpha_1} - 1 \right), \qquad \alpha_{1,\mathsf{min}} := \frac{\log\left(\frac{(1+P)^2}{1+a^2 P}\right)}{\log(1+c^2\hat{P})}, \tag{3.21}$$

where c^2_{min} represents the minimum c^2 required to satisfy (3.17), whereas $\alpha_{1,\mathsf{min}}$ denotes the minimum α_1 required for (3.17) to hold.

We now show the interplay between a^2, c^2_{min} and $\alpha_{1,\mathsf{min}}$ for $P = \hat{P} = 1$. In Figure 3.7a, we plot c^2_{min} against $a^2 \in (0,\ 1)$ for $\alpha_1 \in \{0.5, 1, 2\}$, and note that c^2_{min} reduces monotonically as a^2 or α_1 increases. This follows since, if a^2 increases then the first band forwards more interference, and if α_1 increases, then the pre-log factor of the cross-link capacity increases. In either case, smaller c^2_{min} is required to satisfy (3.17). Similarly in Figure 3.7b, we plot $\alpha_{1,\mathsf{min}}$ against $a^2 \in (0, 1)$ for $c^2 \in \{0.5, 1, 2\}$, and note that $\alpha_{1,\mathsf{min}}$ reduces as either a^2 or c^2 increases. Finally, in Figure 3.7c, we depict the set of cross-channel gains (a^2, c^2) of a symmetric CLIC with $P = \hat{P} = 1, \alpha_1 = 2$, and partition it depending on whether the capacity has been characterized in each subset. We note that while the capacity is known for the strong interference regime (i.e., $a^2 \geq 1$) and the combined weak-very strong interference regime (i.e., where (3.17) holds), the capacity for the remaining subset, where $a^2 < 1$ and $c^2 < c^2_m$ hold, has not been characterized yet.

3.2.4 The Mixed CLIC

In the mixed CLIC, the underlying GIC has mixed interference, i.e., $a^2_{12} < 1$ and $a^2_{21} \geq 1$. For the mixed CLIC, we derive a capacity result as presented below.

Theorem 3.3. *If the channel gains of the mixed CLIC with BMF α_1 satisfy $a^2_{12} < 1$, $a^2_{21} \geq 1$ and the following conditions*

$$(1 + P_1) \leq (1 + a^2_{12}P_1)(1 + c^2_{12}\hat{P}_1)^{\alpha_1} \tag{3.22}$$

$$(1 + P_1 + a^2_{21}P_2)(1 + c^2_{21}\hat{P}_2)^{\alpha_1} \leq (1 + P_2 + a^2_{12}P_1)(1 + c^2_{12}\hat{P}_1)^{\alpha_1}, \tag{3.23}$$

its capacity is given by the set of non-negative rate tuples (R_1, R_2) that satisfy

$$\begin{aligned} R_1 &\leq \mathsf{C}(P_1) \\ R_2 &\leq \mathsf{C}(P_2) \\ R_1 + R_2 &\leq \mathsf{C}\big(P_1 + a^2_{21}P_2\big) + \alpha_1\mathsf{C}\Big(c^2_{21}\hat{P}_2\Big). \end{aligned} \tag{3.24}$$

Proof. The proof is relegated to Appendix B.3. ■

In the mixed CLIC, the underlying GIC has $a^2_{21} \geq 1$ and $a^2_{12} < 1$, i.e., D_1 has strong interference while D_2 has weak interference. Hence, it is optimal for D_1 to decode both the intended and the interfering messages, whereas the same strategy may be sub-optimal

at D_2 due to the reception of weak interference as compared to D_1. Now, if the channel conditions (3.22)-(3.23) hold, sufficient interference is forwarded jointly through the S_1-D_2 cross-link and the weak microwave link that pushes D_2 to the strong interference regime and enables it to decode both messages without incurring rate loss.

Note that a similar capacity result can be characterized for the mixed CLIC where $a_{12}^2 \geq 1$ and $a_{21}^2 < 1$ by changing the roles of the sources and destinations.

3.2.5 The Z-CLIC

In the Z-CLIC, the underlying microwave GIC is a ZIC with $a_{12} = 0$. As presented in Section 3.1, the Z-CLIC can be classified into three cases: the Z-CLIC of type-0, where both mm-wave cross-links are present, the Z-CLIC of type-1, where only S_2-D_1 cross-link is present, and the Z-CLIC of type-2, where only S_1-D_2 cross-link is present, as depicted in Figure 3.5a - Figure 3.5c.

The following lemma shows that the capacity of the Z-CLIC of type-0 is the same as that of type-1.

Lemma 3.3. *The capacity of the Z-CLIC of type-0 is the same as that of the Z-CLIC of type-1.*

Proof. The proof is relegated to Appendix B.4. ■

In essence, D_2 in the Z-CLIC is not interfered by S_1, i.e., $Y_2^n \perp\!\!\!\perp \hat{Y}_2^{n_1}$. Therefore, the mm-wave S_1-D_2 cross-link does not provide additional help to D_2 in decoding its intended message. Hence, removing this cross-link does not result in the reduction of rate achievable at D_2. Removing this cross-link from the Z-CLIC of type-0 then produces the Z-CLIC of type-1 where the capacity region of the latter is the same as that for the former.

Therefore, hereon we focus on the Z-CLIC of type-1.

Capacity of the Z-CLIC of Type-1: For the Z-CLIC of type-1, we consider three cases: the combined very strong interference, the strong interference, and the weak interference. We characterize the capacity in the first two cases, while establishing a constant gap result for the third case in the next subsection.

Lemma 3.4 (Combined weak-very strong interference for the Z-CLIC of type-1)**.** *If the Z-CLIC of type-1 satisfies*

$$(1+P_1)(1+P_2) \leq (1+P_1+a_{21}^2P_2)(1+c_{21}^2\hat{P}_2)^{\alpha_1}, \tag{3.25}$$

its capacity region is given by the set of all non-negative tuples (R_1, R_2) *that satisfy*

$$\begin{aligned} R_1 &\leq \mathsf{C}(P_1) \\ R_2 &\leq \mathsf{C}(P_2)\,. \end{aligned}$$

The proof is similar to that of Theorem 3.2, and hence is not repeated. Note that in the Z-CLIC of type-1, we have $a_{12} = 0$ and $c_{12} = 0$, i.e., destination D_2 operates interference free, while D_1 still receives interference from source S_2. However, if (3.25) holds, then first decoding the interference at D_1 by treating the desired signal as noise, and then decoding the desired signal after removing the effect of interference, does not incur a rate penalty, which produces the rectangular rate region in Lemma 3.4.

Next, we consider two cases where the very strong interference condition in (3.25) does not hold, and thus sequential decoding as in Lemma 3.4 is not optimal anymore. First, we present the capacity region for the strong case where $a_{21}^2 \geq 1$.

Lemma 3.5 (Strong Interference for the Z-CLIC of type-1)**.** *If the Z-CLIC of type-1 satisfies* $a_{21}^2 \geq 1$ *and*

$$(1+P_1)(1+P_2) > (1+P_1+a_{21}^2P_2)(1+c_{21}^2\hat{P}_2)^{\alpha_1}, \tag{3.26}$$

its capacity region is given by the set of all non-negative tuples (R_1, R_2) *that satisfy*

$$\begin{aligned} R_1 &\leq \mathsf{C}(P_1) \\ R_2 &\leq \mathsf{C}(P_2) \\ R_1 + R_2 &\leq \mathsf{C}\big(P_1 + a_{21}^2P_2\big) + \alpha_1\mathsf{C}\Big(c_{21}^2\hat{P}_2\Big)\,. \end{aligned}$$

The proof is similar to that of Lemma 3.2, and hence is omitted. Similar to the case of Lemma 3.4, in the Z-CLIC of type-1, destination D_2 is interference free, while D_1 is interfered by source S_2. In contrast, now condition (3.26) holds, and hence sequential interference decoding at D_1 is sub-optimal. Instead, since D_1 has strong interference, both the desired and the interfering message is jointly decoded at D_1 as in multi-access channel, which is optimal.

Capacity of the Z-CLIC of Type-2: For the sake of completeness, we also consider the Z-CLIC of type-2, and show that its capacity is the same as that of the underlying ZIC

without any mm-wave link.

Lemma 3.6. *The capacity of the Z-CLIC of type-1 is the same as that of the conventional ZIC with* $a_{12} = 0$.

Proof. The proof follows along the lines of Lemma 3.3, and hence not repeated. ■

For a Z-CLIC of type-2 where $a_{12} = 0$ and $c_{21} = 0$, due to $a_{12} = 0$ destination D_2 operates interference free in the underlying GIC. Hence, the S_1-D_2 mm-wave cross-link does not provide any additional information, and thus can be removed without changing the capacity. This is equivalent to taking $c_{12} = 0$, which results in the conventional ZIC with $a_{12} = 0$ for which the capacity is known [100].

Note that for the capacity results for the alternate version of the Z-CLIC, where $a_{21} = 0$, can be obtained from the results of the current Z-CLIC by just changing the source and destination indexes appropriately.

3.3 Approximate Capacity Results

While the capacity region of the strong CLIC has been characterized in closed form, for the weak CLIC, mixed CLIC and the Z-CLIC, the capacity region were obtained only *under certain channel conditions*. Hence, we now derive approximate capacity results (i.e., constant gap results) for these channels that are valid for their entire regimes of definition.

As presented in Definition 2.12, a constant gap result characterizes an outer bound to the capacity region and an achievable rate region for the channel such that the gap between the two is a positive constant that does not depend on the channel parameters. Deriving constant gaps, therefore, typically entails characterizing a "good" achievable strategy and a "good" capacity outer bound such that they differ by only a constant. Hence, the original capacity region, which is sandwiched between the two bounds, is closely approximated by the achievable region, thus revealing an encoding/decoding scheme that is "approximately optimal".

3.3.1 The Weak CLIC

For the weak CLIC, the capacity region is unknown for the range of channel parameters that do not satisfy (3.17)-(3.18). In this case, we first derive an outer bound to the capacity

region with the help of a set of *genie-signals*, and then an achievable region that is within a constant gap from the outer bounds.

Outer Bounds to the Capacity Region

We derive the outer bound first. Note that in the weak interference regime, the destinations do not have sufficient interference information, which makes it hard to computable the single-letter rate upper bounds. This difficulty can be alleviated by providing a destination with the interfering message, but it leads to loose outer-bounds that fails to effectively capture the rates achievable in this regime. We strike a balance between the two approaches by providing *genie-signals* along with the interfering signals that produces computable rate upper bounds which effectively captures the achievable rates.

For the two destinations in the CLIC defined by (3.1)–(3.4), we define the following genie-signals in the microwave band

$$S_{1i} = a_{12}X_{1i} + Z_{2i}, \tag{3.27}$$

$$S_{2i} = a_{21}X_{2i} + Z_{1i}, \quad i = 1, \ldots, n, \tag{3.28}$$

and the following genie-signals in the mm-wave band

$$\hat{S}_{1i} = c_{12}\hat{X}_{1i} + \hat{Z}_{2i}, \tag{3.29}$$

$$\hat{S}_{2i} = c_{21}\hat{X}_{2i} + \hat{Z}_{1i}, \quad i = 1, \ldots, n_1. \tag{3.30}$$

Signals $(S_1^n, \hat{S}_1^{n_1})$ is provided to destination D_1, which can be regarded as a noisy version of its intended codewords that reinforces the desired signal. Similarly, $(S_2^n, \hat{S}_2^{n_1})$ is provided to destination to D_2. Also, from (3.1)–(3.4), it is apparent that $(Y_1^n, \hat{Y}_1^{n_1})|(X_1^n, \hat{X}_1^{n_1})$ has the same distribution as $(S_2^n, \hat{S}_2^{n_1})$, while $(Y_2^n, \hat{Y}_2^{n_1})|(X_2^n, \hat{X}_2^{n_1})$ has the same distribution as $(S_1^n, \hat{S}_1^{n_1})$.

For notational convenience, we define the following parameters

$$\begin{aligned} \mathsf{SNR}_1 &:= P_1, \quad \mathsf{INR}_1 := a_{21}^2 P_2, \\ \mathsf{SNR}_2 &:= P_2, \quad \mathsf{INR}_2 := a_{12}^2 P_1, \end{aligned} \tag{3.31}$$

where SNR_k can be interpreted as the ratio of power of the desired signal from S_k to D_k and noise power (i.e., 1 here), and INR_k represents the ratio of power of interference signal from S_ℓ to D_k and noise power, $k \neq \ell \in \{1, 2\}$. We now present the following outer bound to the capacity region.

Theorem 3.4 (An outer bound to the capacity region of the weak CLIC)**.** *The capacity region of the weak CLIC with BMF α_1 is contained within the set of non-negative rate pairs (R_1, R_2) that satisfy*

$$R_1 \leq \mathsf{C}(\mathsf{SNR}_1) \tag{3.32}$$

$$R_2 \leq \mathsf{C}(\mathsf{SNR}_2) \tag{3.33}$$

$$R_1 + R_2 \leq \mathsf{C}(\mathsf{SNR}_1) + \mathsf{C}\left(\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_2}\right) + C_{12} \tag{3.34}$$

$$R_1 + R_2 \leq \mathsf{C}(\mathsf{SNR}_2) + \mathsf{C}\left(\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_1}\right) + C_{21} \tag{3.35}$$

$$R_1 + R_2 \leq \mathsf{C}\left(\mathsf{INR}_2 + \frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}\right) + \mathsf{C}\left(\mathsf{INR}_1 + \frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right) + C_{12} + C_{21} \tag{3.36}$$

$$R_1 + 2R_2 \leq \mathsf{C}(\mathsf{INR}_2 + \mathsf{SNR}_2) + \mathsf{C}\left(\frac{\mathsf{SNR}_1}{(1+\mathsf{INR}_1)(1+\mathsf{INR}_2)}\right) + \mathsf{C}(\mathsf{SNR}_2) + C_{12} + C_{21} \tag{3.37}$$

$$2R_1 + R_2 \leq \mathsf{C}(\mathsf{INR}_1 + \mathsf{SNR}_2) + \mathsf{C}\left(\frac{\mathsf{SNR}_2}{(1+\mathsf{INR}_1)(1+\mathsf{INR}_2)}\right) + \mathsf{C}(\mathsf{SNR}_1) + C_{12} + C_{21} \tag{3.38}$$

where $C_{12} := \alpha_1 \mathsf{C}\left(c_{12}^2 \hat{P}_1\right)$ and $C_{21} := \alpha_1 \mathsf{C}\left(c_{21}^2 \hat{P}_2\right)$, with $\mathsf{C}(x) = \frac{1}{2}\log(1+x)$.

Proof. We present the key steps of the proof here while details are relegated to Appendix B.5. Following the outer bounding techniques in [10], each bound is derived by providing an appropriate additional signal, either a genie-signal defined in (3.27)-(3.30) or the full interference signal, to one or both receivers such that the resulting negative entropy terms either cancel out or can be bounded by applying the worst additive noise (WAN) technique [130]. The remaining positive entropy terms are then single-letterized by using Gaussian codewords as in [131, Lemma 1]. ■

Achievable Region

Unlike the strong interference case where decoding the interfering message along with the desired message, or the very strong case where decoding the interfering message and then the desired message in a sequential manner, in the weak interference regime, decoding the interference in its entirety is not optimal in general. Moreover, treating the interference as noise is also not optimal unless the interference is very weak. Therefore, we need a flexible framework that fills the gap between the two approaches of addressing the interference problem.

The Han-Kobayashi (HK) scheme [7] provides such a flexible framework, and achieves

the best known rates for the conventional IC [98]. In this scheme, the message from each source is first partitioned into a public and a private part, and both parts superimpose to generate a single codeword for transmission. In particular, each public message is first encoded into a "cloud center" which carries "coarse information", upon which a private message is superimposed to create a "satellite codeword" carrying the full information. The total transmission power is divided into those for the public and private parts. Such partitioning enables a destination to decode only a part (e.g., the public part) of the interference only, and relieves it from the constraint of either fully decoding the interference or treating it entirely as noise.

Note that in the HK scheme, the distribution of the codewords and the power-splitting between the public and private parts can be varied to adapt to channel conditions. Moreover, all such variations can be time-shared to produce an overall rate region which makes it prohibitively complex to compute the HK region in closed form as the time-sharing variable can have large cardinality. The need for an easily computable region motivated the simplified HK scheme of [10] for the conventional GIC, where only Gaussian codewords and a fixed power-splitting were used without time-sharing. The resulting rate region was able to achieve within 1/2 bit/channel use of the capacity for the real GIC.

In the following Lemma, we present a computable achievable rate region for the CLIC based on a scheme adapted from [10].

Lemma 3.7 (An achievable region for the CLIC based on message splitting)**.** *For the CLIC with BMF α_1 defined in (3.1)-(3.4) with transmit powers $P_1, P_2, \hat{P}_1$ and $\hat{P}_2$, let $\mathcal{P}^*$ be the set vectors $\boldsymbol{p} := (P_{U1}, P_{W1}, P_{U2}, P_{W2}, \hat{P}_1, \hat{P}_2) \succeq \mathbf{0}$ such that*

$$\mathcal{P}^* = \{\boldsymbol{p} : P_{U1} + P_{W1} = P_1, \;\; P_{U2} + P_{W2} = P_2\}. \tag{3.39}$$

For i.i.d. random variables $W_k \sim \mathcal{N}(0, P_{Wk}), \hat{X}_k \sim \mathcal{N}(0, \hat{P}_k)$ and $U_k \sim \mathcal{N}(0, P_{Uk}), k = 1, 2$, define the $X_k := W_k + U_k$ such that Y_k is the resulting outputs in (3.1)-(3.2), and $\hat{Y}_k$ be the resulting outputs in (3.3)-(3.4). Let $\mathcal{R}(\boldsymbol{p})$ be the set of non-negative rate tuples (R_1, R_2) that satisfy

$$\begin{aligned}
R_1 &\le I(X_1; Y_1|W_2) \\
R_2 &\le I(X_2; Y_2|W_1) \\
R_1 + R_2 &\le I(X_1; Y_1|W_1, W_2) + I(X_2, W_1; Y_2) + C_{12} \\
R_1 + R_2 &\le I(X_2; Y_2|W_1, W_2) + I(X_1, W_2; Y_1) + C_{21} \\
R_1 + R_2 &\le I(X_1, W_2; Y_1|W_1) + I(X_2, W_1; Y_2|W_2) + C_{12} + C_{21}
\end{aligned}$$

$$R_1 + 2R_2 \leq I(X_2, W_1; Y_2) + I(X_2; Y_2|W_1, W_2) + I(X_1, W_2; Y_1|W_1) + C_{12} + C_{21}$$
$$2R_1 + R_2 \leq I(X_1, W_2; Y_1) + I(X_1; Y_1|W_1, W_2) + I(X_2, W_1; Y_2|W_2) + C_{12} + C_{21}$$

where $C_{12} = \alpha_1 I(\hat{X}_1; \hat{Y}_2)$ *and* $C_{21} = \alpha_1 I(\hat{X}_2; \hat{Y}_1)$*. Then, an achievable region for the CLIC is given by* $\mathcal{R} = \cup_{\boldsymbol{p} \in \mathcal{P}^*} \mathcal{R}(\boldsymbol{p})$*.*

Proof. The proof of this lemma follows from extending the Han-Kobayashi encoding scheme for the conventional IC in [7] to the CLIC. The details are relegated to Appendix B.6. ■

Constant Gap Result

The achievable region in Theorem 3.7 is still not computable as a union still needs to be taken over all possible power-splits $\boldsymbol{p} \in \mathcal{P}^*$. Hence, to characterize a closed form, we now restrict the power-splits for the two sources (P_{U1}, P_{W1}) and (P_{U2}, P_{W2}) to a specific form that depends on the SNR and INR values of the two destinations defined in (3.31) as in [10]. In particular, the private message power P_{U2} and public message power P_{W2} for source S_2, defined as a function of INR_1, is given by

$$P_{U2} = \begin{cases} 1/a_{21}^2 & \text{if } \mathsf{INR}_1 \geq 1 \\ P_2 & \text{if } \mathsf{INR}_1 < 1, \end{cases} \qquad P_{W2} = \begin{cases} P_2 - 1/a_{21}^2 & \text{if } \mathsf{INR}_1 \geq 1 \\ 0 & \text{if } \mathsf{INR}_1 < 1. \end{cases} \tag{3.40}$$

Similarly, the private message power P_{U1} and public message power P_{W1} for source S_1, defined as a function of INR_2, is given by

$$P_{U1} = \begin{cases} 1/a_{12}^2 & \text{if } \mathsf{INR}_2 \geq 1 \\ P_1 & \text{if } \mathsf{INR}_2 < 1. \end{cases} \qquad P_{W1} = \begin{cases} P_1 - 1/a_{12}^2 & \text{if } \mathsf{INR}_2 \geq 1 \\ 0 & \text{if } \mathsf{INR}_2 < 1. \end{cases} \tag{3.41}$$

The resulting region in Lemma 3.7 can thus be parameterized by $(\mathsf{INR}_1, \mathsf{INR}_2)$ and denoted by $\mathcal{R}(\mathsf{INR}_1, \mathsf{INR}_2)$. It turns out that $\mathcal{R}(\mathsf{INR}_1, \mathsf{INR}_2)$ provides a closed form achievable region that it is within a constant gap to the capacity region of the weak CLIC.

Theorem 3.5 (The constant gap result for the weak CLIC)**.** *The achievable region, denoted by* $\mathcal{R}(\mathsf{INR}_1, \mathsf{INR}_2)$ *and found by restricting the region of Lemma 3.7 to the private-public message powers defined in* (3.40)-(3.41)*, is within* $\delta = \frac{1}{2}$ *bit/channel use per user of the capacity region of the weak CLIC with BMF* α_1*.*

Proof. The proof borrows the techniques of the constant gap results for the weak GIC in [10], and is detailed in Appendix B.7. In essence, the set of $(\mathsf{INR}_1, \mathsf{INR}_2)$ is first partitioned into four sets depending on the values of $(\mathsf{INR}_1, \mathsf{INR}_2)$, and the private-public message

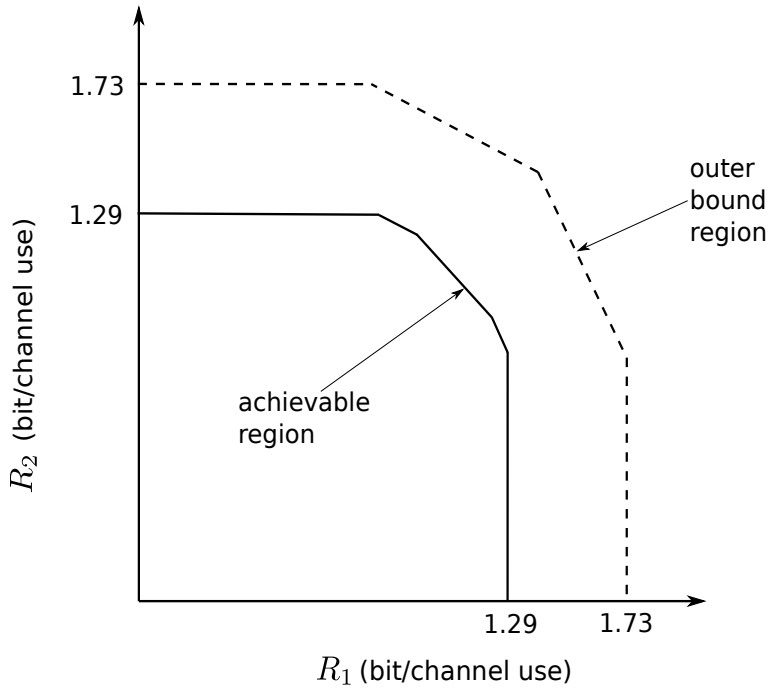


Figure 3.8: An example of the constant gap result for the CLIC. The gap between the achievable region and the outer bound region is within 1/2 bit/channel use per dimension.

powers are then adapted according to (3.40)-(3.41). This results in closed form expressions for the inner bound $\mathcal{R}(\mathsf{INR}_1, \mathsf{INR}_2)$ and outer bound in Theorem 3.4, which are then shown to be within a distance of $\delta = \frac{1}{2}$ bit per user. ■

Finally, we illustrate an example of the constant gap result for the CLIC in Figure 3.8 with parameters $\mathsf{SNR}_1 = \mathsf{SNR}_2 = 10, \mathsf{INR}_1 = \mathsf{INR}_2 = 2$, and $C_{12} = C_{21} = 0.1$. Here, while all seven constraints are active in the achievable region, only six constraints are active in the outer bound region, nevertheless, the gap between the two regions is within 1/2 bit/channel use in each dimension.

3.3.2 The Mixed CLIC

Recall that the capacity of the mixed CLIC was characterized only under conditions (3.22)-(3.23), beyond which the capacity is unknown. Hence, this motivates the study of a constant gap results for the mixed CLIC.

Outer Bound to the Capacity Region

First, we characterize an outer bound to the capacity region of the mixed CLIC.

Theorem 3.6 (An outer bound to the capacity region of the mixed CLIC)**.** *The capacity region of the mixed CLIC with BMF α_1 that satisfies $a_{21}^2 \geq 1$ and $a_{12}^2 < 1$ is contained within the set of non-negative rate pairs (R_1, R_2) that satisfy*

$$R_1 \leq \mathsf{C}(\mathsf{SNR}_1) \tag{3.42}$$

$$R_2 \leq \mathsf{C}(\mathsf{SNR}_2) \tag{3.43}$$

$$R_1 + R_2 \leq \mathsf{C}(\mathsf{SNR}_1) + \mathsf{C}\left(\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_2}\right) + C_{12} \tag{3.44}$$

$$R_1 + R_2 \leq \mathsf{C}(\mathsf{SNR}_1 + \mathsf{INR}_1) + C_{21} \tag{3.45}$$

$$R_1 + 2R_2 \leq \mathsf{C}(\mathsf{SNR}_2 + \mathsf{INR}_2) + \mathsf{C}\left(\mathsf{INR}_1 + \frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right) + \mathsf{C}\left(\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}\right) + C_{12} + C_{21} \tag{3.46}$$

where $C_{12} = \alpha_1 I(\hat{X}_1; \hat{Y}_2) = \alpha_1 \mathsf{C}(c_{12}^2 P_1)$ and $C_{21} = \alpha_1 I(\hat{X}_2; \hat{Y}_1) = \alpha_1 \mathsf{C}(c_{21}^2 P_2)$.

Proof. The proof is relegated to Appendix B.8. ■

Achievable Region

We characterize a closed form achievable region for the mixed CLIC by adapting the region in Lemma 3.7 to the following private-public message power split: the transmission power from source S_1 is split according to (3.41) depending on INR_2, whereas the total transmission power from source S_2 is allotted entirely to the common message, i.e., $P_{W2} = P_2$ and $P_{U2} = 0$, irrespective of the value of INR_1. For notational convenience, we denote the resulting achievable region by $\mathcal{R}(\mathsf{INR}_2, *)$ where $*$ in place of INR_1 denotes the insensitivity of the scheme to INR_1 values.

Intuitively, since strong interference from S_2 to D_1 holds (i.e., $a_{21}^2 \geq 1$), decoding the message M_2 entirely at D_1 does not incur any rate loss, and hence all of P_2 is allocated to the common message. In contrast, we have weak interference from S_1 to D_2 (i.e., $a_{12}^2 < 1$), where message splitting as in the weak CLIC proves to be beneficial.

Constant Gap Result

We now have the following constant gap result which shows that the achievable region $\mathcal{R}(\mathsf{INR}_2, *)$ is within 1/2 bit of the capacity region.

Theorem 3.7 (The constant gap result for the mixed CLIC)**.** *For the mixed CLIC with BMF α_1 that satisfies $a_{21}^2 \geq 1$ and $a_{12}^2 < 1$, the achievable region $\mathcal{R}(\mathsf{INR}_2, *)$ is within $\delta = 1/2$ bit/channel use of the capacity region.*

Proof. Similar to the weak CLIC, we first partition the range of INR_2 into two sets $\mathsf{INR}_2 \geq 1$ and $\mathsf{INR}_2 < 1$. Then for each set, we simplify the achievable region $\mathcal{R}(\mathsf{INR}_2, *)$ and show it to be within $\delta = 1/2$ bit of the outer bound in Theorem 3.6. The details are relegated to Appendix B.9. ■

Note that for the alternative mixed CLIC, where the destinations with strong and weak interference are exchanged (i.e., $a_{21}^2 < 1$ and $a_{12}^2 \geq 1$), a similar constant gap result can be obtained from Theorem 3.7 by exchanging the roles of the sources and destinations.

3.3.3 The Z-CLIC

Finally, consider the Z-CLIC and recall that its capacity was characterized for cases with very strong interference condition (3.25) or strong interference condition $a_{21}^2 \geq 1$. If neither of these two conditions hold, neither successive interfering decoding nor joint decoding of interference at destination D_1 is optimal. As such, based on the HK region in Lemma 3.7, we characterize an achievable region which is then shown to be within constant gap of the outer bound.

Theorem 3.8 (The constant gap result for the weak Z-CLIC of type-1)**.** *For the weak Z-CLIC of type-1 with BMF α_1 that satisfies $a_{21}^2 < 1$, the achievable region is within $\delta = 1/2$ bit/channel use of the capacity region.*

Proof. The proof is similar to that of Theorem 3.7 and is relegated to Appendix B.10. ■

3.4 Resource Allocation in the DCLIC

The mm-wave links will significantly impact the performance of the DCLIC due to the point-to-point nature of mm-wave transmissions and relatively larger bandwidth, and thus

it is important to quantify the impact of the mm-wave band on resource allocation in the DCLIC. As such, we take the sum-rate to be the figure of performance of the DCLIC, which captures the total throughput of the network, and study the optimal power allocation scheme over the mm-wave direct- and cross-links that maximizes the sum-rate. For this study, we solely focus on the strong DCLIC, as the capacity region for this case, given in (3.13)-(3.16), is known without any condition on the mm-wave link channel parameters.

Recall that the normalized bandwidth (α) of the second band is shared between the direct-links (α_2) and the cross-links (α_1). We denote the fraction of α alloted to the direct-links by $\beta := \alpha_2/\alpha$ and the same for the cross-links by $\bar{\beta} := 1 - \beta = \alpha_1/\alpha$, where $\beta, \bar{\beta} \in (0, 1)$. Thus, β provides a trade-off between the bandwidths in the direct- and cross-links.

We formulate the problem for a class of DCLICs that satisfies the following assumptions:

- [A1] the DCLIC has *strong interference* but not very strong interference, i.e., $1 \leq a_{12}^2 < 1 + P_1$ and $1 \leq a_{21}^2 < 1 + P_2$;
- [A2] the DCLIC satisfies the symmetry condition $P_2 + a_{12}^2 P_1 = P_1 + a_{21}^2 P_2$;
- [A3] $\beta > 0$ and $\bar{\beta} > 0$ are fixed a priori;
- [A4] the transmit power in the direct-links (p_k) and cross-link (q_k) from source S_k are constrained by the total power budget P, i.e., $\beta p_k + \bar{\beta} q_k = P, k = 1, 2$.

In [A1], we assume that the DCLIC has strong interference, as the power allocation for the very strong interference is trivial as shown in Section 3.4.5. For ease of exposition, we assume in [A2] that the underlying GIC receives equal power in both its receivers: this symmetry condition helps to ease the exposition while preserving the defining structures of the optimal power allocation scheme which is then extended to the asymmetric case in Section 3.4.4. Note that the class of DCLICs that satisfies [A2] contains as a special case the DCLIC where the underlying microwave GIC has symmetric cross-channel gains and power constraints.

In [A3], β is assumed to be fixed a priori and known. This models practical constraints in many wireless networks where dynamically allocating the bandwidth may not be feasible or straightforward [123]. In [A4], we assume that the total power budget P in both sources are the same. This causes no loss of generality as the relative difference between the power budgets of the two sources, if any, can be absorbed into the gains of mm-wave the direct- and cross-links without altering the treatment of the problem.

3.4.1 Problem Formulation and Solution

For a fixed power allocation (p_1, q_1, p_2, q_2) in the mm-wave channels, the achievable sum-rate for the DCLIC is given by the minimum of rates Σ_1, Σ_2 and Σ: Σ_1 and Σ_2 are the sum-rates achievable at destinations D_1 and D_2, given by (3.15) and (3.16), respectively, while Σ is the interference-free sum-rate, given by the sum of individual rates in (3.13) and (3.14). For convenience, the three sum-rate expressions are presented below

$$\Sigma_1 = A_1 + \frac{\alpha\bar{\beta}}{2}\log\left(1 + c_{21}^2 q_2\right) + \frac{\alpha\beta}{2}\left(\log\left(1 + d_1^2 p_1\right) + \log\left(1 + d_2^2 p_2\right)\right), \tag{3.47}$$

$$\Sigma_2 = A_2 + \frac{\alpha\bar{\beta}}{2}\log\left(1 + c_{12}^2 q_1\right) + \frac{\alpha\beta}{2}\left(\log\left(1 + d_1^2 p_1\right) + \log\left(1 + d_2^2 p_2\right)\right) \tag{3.48}$$

$$\Sigma = A + \frac{\alpha\beta}{2}\left(\log\left(1 + d_1^2 p_1\right) + \log\left(1 + d_2^2 p_2\right)\right), \tag{3.49}$$

where

$$\begin{aligned} A_1 &:= \frac{1}{2}\log\left(1 + P_1 + a_{21}^2 P_2\right) \\ A_2 &:= \frac{1}{2}\log\left(1 + P_2 + a_{12}^2 P_1\right) \\ A &:= \frac{1}{2}\log\left(1 + P_1\right)\left(1 + P_2\right) \end{aligned} \tag{3.50}$$

are defined for notational convenience. Note that under assumptions [A1] and [A2], we have $A_1 = A_2 < A$.

For a power allocation vector (p_1, q_1, p_2, q_2), a necessary and sufficient condition for R to be an achievable sum-rate of the DCLIC is $R \leq \min\{\Sigma_1, \Sigma_2, \Sigma\}$. Hence, the optimization problem that maximizes R over the transmit powers (p_1, q_1, p_2, q_2) is as follows

$$[\mathcal{P}1] \quad \text{maximize} \quad R \tag{3.51}$$

$$\text{subject to:} \quad R \leq \Sigma_1 \tag{3.52}$$

$$R \leq \Sigma_2 \tag{3.53}$$

$$R \leq \Sigma \tag{3.54}$$

$$\beta p_1 + \bar{\beta} q_1 = P \tag{3.55}$$

$$\beta p_2 + \bar{\beta} q_2 = P \tag{3.56}$$

$$(p_1, q_1, p_2, q_2, R) \succeq \mathbf{0}. \tag{3.57}$$

Solution of Problem $[\mathcal{P}1]$

The solution of problem $[\mathcal{P}1]$ has been relegated to Appendix B.11. Specifically, note that $[\mathcal{P}1]$ is a convex optimization problem: its objective function R is linear, the equality constraints (3.55) and (3.56) are affine, and the inequality constraints (3.52)-(3.54) are convex and differentiable. Moreover, $[\mathcal{P}1]$ satisfies the Slater constraint qualification condition [132, Chapter 5.2.3]. Hence, $[\mathcal{P}1]$ can be solved using the Karush–Kuhn–Tucker (KKT) conditions [132, Chapter 5.5.3], as detailed in Appendix B.11.

Constraint Qualification for Problem $[\mathcal{P}1]$

In Appendix B.12, it was shown that problem $[\mathcal{P}1]$ satisfies the Mangasarian-Fromovitz constraint qualification.

3.4.2 Optimal Power Allocation and Link Gain Regimes

To gain insights into the optimal power allocation scheme (IC-OPA), we characterize the solution of problem $[\mathcal{P}1]$ in closed form. As revealed shortly, link transmission powers are allocated in different *modes* that depend on certain conditions on the *channel parameters* and the power budget P. Such conditions lead to partitioning the set of all channel parameters and P into a few mutually exclusive *link-gain regimes* (LGR): in each LGR, the IC-OPA allocates the link transmission powers in different modes. Such a characterization turns out to be beneficial in deriving practical insights.

The LGRs are obtained while solving the convex problem $[\mathcal{P}1]$ using KKT conditions: we solve for the optimal primal variables (i.e., the optimal link transmit powers) and the optimal dual variables (i.e., the optimal Lagrange multipliers or OLM). For convenience, we first partition the set of OLMs associated with the inequality constraints (3.52), (3.53) and (3.54) into a few subsets depending on whether the OLMs are *positive or zero*, i.e., whether the associated constraint is *tight or not* (detailed in Appendix B.11). As such, each resulting subset of OLMs are mutually exclusive.

Then, for each such subset of OLMs, we characterize the optimal link transmit powers in closed form. Since the conditions that define these subsets of OLMs are still expressed in terms of the OLMs, we then express these conditions explicitly in terms of channel

Table 3.1: Definition of LGRs and optimal link powers for problem $[\mathcal{P}1]$ in terms of channel parameters $\boldsymbol{a} = (d_1, d_2, c_{12}, c_{21}, \gamma, P)$. Table 3.2 provides the threshold powers in terms of $(d_1, d_2, c_{12}, c_{21}, \gamma)$, and γ is defined in (3.58).

Definition of LGR	**Optimal Power Allocation**	
$\mathcal{A}_{\mathsf{d,d}} := \{\boldsymbol{a} : 0 \le P \le \bar{\mathsf{P}}_1\}$	$p_1 = \dfrac{P}{\beta},$	$q_1 = 0,$
	$p_2 = \dfrac{P}{\beta},$	$q_2 = 0,$
$\mathcal{A}_{\mathsf{c,cd}} := \{\boldsymbol{a} : 0 \le P \le \min(\bar{\mathsf{P}}_2, \bar{\mathsf{P}}_4)\}$	$p_1 = 0,$	$q_1 = \dfrac{P}{\bar{\beta}},$
	$p_2 = \dfrac{P(c_{21}^2 - c_{12}^2)}{\beta c_{21}^2},$	$q_2 = \dfrac{c_{12}^2}{c_{21}^2}\dfrac{P}{\bar{\beta}},$
$\mathcal{A}_{\mathsf{cd,cd}} := \{\boldsymbol{a} : \max(\bar{\mathsf{P}}_1, \bar{\mathsf{P}}_2) < P < \bar{\mathsf{P}}_3\}$	$p_1 = \dfrac{P - \bar{\beta} q_1}{\beta},$	$q_1 = \dfrac{F(P)}{c_{12}^2},$
	$p_2 = \dfrac{P - \bar{\beta} q_2}{\beta},$	$q_2 = \dfrac{c_{12}^2}{c_{21}^2} q_1,$
$\mathcal{S}_{\mathsf{cd,cd}} := \{\boldsymbol{a} : \max(\bar{\mathsf{P}}_3, \bar{\mathsf{P}}_4) \le P\}$	$p_1 = \dfrac{P}{\beta} - \dfrac{\bar{\beta}(\gamma - 1)}{\beta c_{12}^2},$	$q_1 = \dfrac{\gamma - 1}{c_{12}^2},$
	$p_2 = \dfrac{P}{\beta} - \dfrac{\bar{\beta}(\gamma - 1)}{\beta c_{21}^2},$	$q_2 = \dfrac{\gamma - 1}{c_{21}^2},$

parameters $(d_1, d_2, c_{12}, c_{21})$, power budget P, and a parameter γ, defined as

$$\gamma := \left(\frac{(1 + P_1)(1 + P_2)}{1 + P_1 + a_{21}^2 P_2} \right)^{1/\alpha\bar{\beta}}, \tag{3.58}$$

which models the effect of microwave band channel parameters.

As a result, the set of $(d_1, d_2, c_{12}, c_{21}, \gamma, P)$-tuples is now partitioned into a few subsets (i.e., LGRs), each corresponding to one and only one subset of OLMs. Moreover, the defining condition of each LGR is further simplified by expressing these condition as upper and lower bounds *threshold powers* on the power budget P that are functions of $(d_1, d_2, c_{12}, c_{21}, \gamma)$. Hence, for a given set of $(d_1, d_2, c_{12}, c_{21}, \gamma)$, this procedure results in *partitioning* the range of power budget $P \ge 0$ into a few intervals, each corresponding to an LGR and defined by the associated threshold powers.

LGRs in the Case with $c_{21}^2 > c_{12}^2$: Without loss of generality, we study here the optimal power allocation for the case with $c_{21}^2 > c_{12}^2$. In Table 3.1, the LGRs and the corresponding optimal powers are defined, where threshold powers $\bar{\mathsf{P}}_1, \bar{\mathsf{P}}_2, \bar{\mathsf{P}}_3$ and $\bar{\mathsf{P}}_4$ are expressed in Table 3.2 below, with $\varrho[\mathsf{p}(x)]$ denoting the non-negative square-root of polynomial $\mathsf{p}(x)$, and

$$
\begin{aligned}
F(P) &:= \frac{1}{2\bar{\beta}(1+\beta)}\left(E_1 + E_2 - \sqrt{(E_1 - E_2)^2 + 4\beta^2 E_1 E_2}\right) - 1, \text{ with} \\
E_1 &= \left(P c_{12}^2 + \bar{\beta} + \beta \frac{c_{12}^2}{d_1^2}\right), \\
E_2 &= \left(P c_{21}^2 + \bar{\beta} + \beta \frac{c_{21}^2}{d_2^2}\right).
\end{aligned} \tag{3.59}
$$

Note that in this case, only 4 LGRs are needed to characterize the IC-OPA, $\mathcal{A}_{\mathsf{d,d}}$, $\mathcal{A}_{\mathsf{c,cd}}$, $\mathcal{A}_{\mathsf{cd,cd}}$, and $\mathcal{S}_{\mathsf{cd,cd}}$. 3

As discussed in detail shortly, the IC-OPA follows two distinct properties, the *Waterfilling-like* property and the *saturation* property: the LGRs associated with the Waterfilling-like property are denoted by $\mathcal{A}_{(.,.)}$, while LGR $\mathcal{S}_{\mathsf{cd,cd}}$ is associated with the saturation property. We relegate the details of the derivation to Appendix B.11.

Also note that while we study the IC-OPA for the case with $c_{21}^2 > c_{12}^2$, the IC-OPA under $c_{21}^2 < c_{12}^2$ can be readily obtained from Table 3.1 by swapping the indexes $1 \leftrightarrow 2$, while the case with $c_{21}^2 = c_{12}^2$ is studied in Section 3.4.4.

Notation for the LGRs: The notation of the LGRs can be interpreted as follows. For LGR $\mathcal{A}_{\mathsf{x,y}}$ with $\mathsf{x}, \mathsf{y} \in \{\mathsf{d}, \mathsf{c}, \mathsf{cd}\}$, entries x and y denote the transmission status in the direct- and cross-links from source S_1 and S_2, respectively: for a given source, entry d, c and cd denotes that power is allocated to the direct-link only, the cross-link only, and in both links, respectively. For example, in LGR $\mathcal{A}_{\mathsf{c,cd}}$ the total power budget for source S_1 is allocated to the S_1-D_2 cross-links only, the budget for source S_2 is shared between both the cross- and direct-link originating from S_2.

Since the LGRs are defined from mutually exclusive sets of OLMs as discussed above, they are inherently mutually exclusive. For example, given a set of channel parameters $\boldsymbol{a} = (d_1, d_2, c_{12}, c_{21}, \gamma, P)$, due to the mutual exclusiveness of the LGRs, the threshold powers $\bar{\mathsf{P}}_1$ of $\mathcal{A}_{\mathsf{d,d}}$ and $\bar{\mathsf{P}}_2$ of $\mathcal{A}_{\mathsf{c,cd}}$ have the following relation: if $\bar{\mathsf{P}}_1 > 0$, then $\bar{\mathsf{P}}_2 < 0$, and conversely, if $\bar{\mathsf{P}}_1 < 0$, then $\bar{\mathsf{P}}_2 > 0$. In particular, if $\bar{\mathsf{P}}_1 > 0$, there exists $P \geq 0$ such that $0 \leq P \leq \bar{\mathsf{P}}_1$, i.e., transmitting as in $\mathcal{A}_{\mathsf{d,d}}$ is optimal for this range of P. In addition, since

Table 3.2: Definition of the threshold powers for LGRs defined in Table 3.1 with γ defined in (3.58).

$$\bar{\mathsf{P}}_1 := \varrho\left[\frac{d_1^2}{c_{12}^2(1+xd_1^2/\beta)} + \frac{d_2^2}{c_{21}^2(1+xd_2^2/\beta)} - 1\right]$$

$$\bar{\mathsf{P}}_2 := \varrho\left[\left(1+c_{12}^2 x/\bar{\beta}\right)\left(\frac{d_1^2}{c_{12}^2} + \frac{1}{x(c_{21}^2-c_{12}^2)/\beta + c_{21}^2/d_2^2}\right) - 1\right]$$

$$\bar{\mathsf{P}}_3 := \varrho\left[\frac{\beta\gamma}{xc_{21}^2+\bar{\beta}+\beta c_{21}^2/d_2^2-\bar{\beta}\gamma} + \frac{\beta\gamma}{xc_{12}^2+\bar{\beta}+\beta c_{12}^2/d_1^2-\bar{\beta}\gamma} - 1\right]$$

$$\bar{\mathsf{P}}_4 := \frac{\bar{\beta}(\gamma-1)}{c_{12}^2}$$

$\bar{\mathsf{P}}_1 > 0 \implies \bar{\mathsf{P}}_2 < 0$, one condition of $\mathcal{A}_{\mathsf{c,cd}}$, $0 \le P \le \bar{\mathsf{P}}_2$, is not satisfied by any $P \ge 0$, and hence transmitting as in $\mathcal{A}_{\mathsf{c,cd}}$ is not optimal for any $P \ge 0$. In this case, the direct-links can be interpreted as being stronger than cross-links. Similarly, the complementary case, where $\bar{\mathsf{P}}_2 > 0 \implies \bar{\mathsf{P}}_1 < 0$, can be interpreted as the cross-links being stronger than the direct-links.

3.4.3 Properties of the Optimal Power Allocation

It is well known that optimal power allocation in parallel Gaussian AWGN channels, where a source communicates to a single destination via parallel point-to-point links (sub-channels), follows the Waterfilling (WF) property: if the power budget is sufficiently small, it is allocated entirely to the "strongest" sub-channel, and, as the power budget is increased, power is allocated to the other sub-channels, in addition to the strongest one [97, Chapter 3.4.3]. In the ensuing discussion, it becomes clear that the IC-OPA has two distinct properties: a *WF-like* property, due to which it assigns power to the cross- and direct-links following a WF-like allocation, and a *saturation* property, due to which there exists a peak power constraint on the cross-links, while no such limit exists for the direct-links.

These two properties are effectively captured by a certain *saturation threshold* $\bar{\mathsf{P}}_{\mathsf{sat}}$ on the total power budget P such that the IC-OPA allocates power in the links as follows:

- For $P < \bar{\mathsf{P}}_{\mathsf{sat}}$: if P is *sufficiently* small, for each source, the IC-OPA allocates the power budget either entirely in the *strongest* of the direct- and cross-links (e.g., only in both direct-links as in $\mathcal{A}_{\mathsf{d,d}}$), or in the strongest link of one source and both links

of the other source (as in $\mathcal{A}_{\mathsf{c,cd}}$); as P increases further within $P < \bar{\mathsf{P}}_{\mathsf{sat}}$, power is eventually allocated to all four links in a WF fashion as in $\mathcal{A}_{\mathsf{cd,cd}}$.

- For $P \geq \bar{\mathsf{P}}_{\mathsf{sat}}$: the transmit powers allocated to both cross-links are *saturated* in the sense that as P increases further, the power in the cross-links remain unchanged, and all such increments in P are allocated only to the direct-links.

The Waterfilling-like Property

As already mentioned, condition $\bar{\mathsf{P}}_1 > 0$ in $\mathcal{A}_{\mathsf{d,d}}$ implies that the direct-links are "stronger" than the cross-links in that, for sufficiently small P, allocating P only to the direct-links results in a larger sum-rate than that achieved from allocating power to any other subset of links. Thus, when $P \leq \bar{\mathsf{P}}_1$, following its WF-like property, IC-OPA allocates P entirely to the direct-links, i.e., $p_1 = p_2 = \frac{P}{\beta}$, and $q_1 = q_2 = 0$. In $\mathcal{A}_{\mathsf{d,d}}$, the direct-link powers thus increase linearly with P, while the cross-link powers are zero.

On the other hand, condition $\bar{\mathsf{P}}_2 > 0$ in $\mathcal{A}_{\mathsf{c,cd}}$ implies that the cross-links are "stronger" than the direct-links in that, for sufficiently small P, the sum-rate achieved by allocating P to three links as in $\mathcal{A}_{\mathsf{c,cd}}$ is larger than that achieved in $\mathcal{A}_{\mathsf{d,d}}$. This power allocation remains optimal for $P \leq \bar{\mathsf{P}}_2$, and the powers in two cross-links and the direct-link from S_2 increase linearly with P. Moreover, condition $P < \bar{\mathsf{P}}_4$ ensures that power in the cross-links are not saturated.

Note that when the direct-links are stronger as in $\mathcal{A}_{\mathsf{d,d}}$, allocating P entirely to the both direct-links results in the optimal sum-rate $R = \Sigma_2 = \Sigma_1$. In contrast, when cross-links are stronger, due to the assumption $c_{21}^2 > c_{12}^2$, transmitting only in two cross-links results in a sub-optimal sum-rate $R = \Sigma_2 < \Sigma_1$. In this case, the sum-rate is limited to $R = \Sigma_2$, and allocating full power to the S_2-D_1 cross-link is not necessary as the larger rate Σ_1 is not realized. Hence, the IC-OPA allocates power to both cross-links as well as *a direct-link* by sharing P between the cross- and direct-links from source S_2 to produce the optimal sum-rate $R = \Sigma_2 = \Sigma_1$.

As P is increased, the marginal benefits of transmitting in a particular subset of links (e.g., only direct-links in $\mathcal{A}_{\mathsf{d,d}}$, or cross- and direct-links as in $\mathcal{A}_{\mathsf{c,cd}}$) begins to diminish. Hence, for $\max(\bar{\mathsf{P}}_1, \bar{\mathsf{P}}_2) < P \leq \bar{\mathsf{P}}_3$, the IC-OPA allocates power to all links in a WF fashion as in $\mathcal{A}_{\mathsf{cd,cd}}$. This also achieves the maximum sum-rate $R = \Sigma_2 = \Sigma_1$.

The Saturation Property

While the IC-OPA follows the WF property for $P < \bar{\mathsf{P}}_{\mathsf{sat}}$, once $P \geq \bar{\mathsf{P}}_{\mathsf{sat}}$ it follows the saturation property: both cross-link powers are simultaneously constrained at $q_1 = \frac{\gamma-1}{c_{12}^2}$ and $q_2 = \frac{\gamma-1}{c_{21}^2}$, and as P increases further, the cross-link powers remain unchanged and all additional increments in P are allocated only to the direct-links, as in LGR $\mathcal{S}_{\mathsf{cd,cd}}$. The cross-links are thus said to be saturated in that allocating more cross-link powers does not improve the sum-rate. Such limits on the cross-link powers is unlike the WF-only allocation in [97].

The cross-links become saturated due to the objective of maximizing $\min(\Sigma_1, \Sigma_2, \Sigma)$. The analysis of the KKT conditions reveals that for small enough P, it is optimal to allocate powers that achieve only $\Sigma_1 = \Sigma_2 < \Sigma$, while having a gap between $\Sigma_1 = \Sigma_2$ and Σ, as has been achieved in LGRs $\mathcal{A}_{\mathsf{d,d}}$, $\mathcal{A}_{\mathsf{c,cd}}$, and $\mathcal{A}_{\mathsf{cd,cd}}$. Note that the link powers in LGRs $\mathcal{A}_{(.,.)}$ increase piece-wise linearly as P increases: the increases in p_1 and p_2 results in an equal increase of Σ_1, Σ_2 and Σ, whereas an increase in q_1 only increases Σ_2, while an increase in q_2 only increases Σ_1. Therefore, as P increases, the gap between $\Sigma_1 = \Sigma_2$ and Σ reduces, and at $P = \bar{\mathsf{P}}_{\mathsf{sat}}$, the IC-OPA achieves $\Sigma_1 = \Sigma_2 = \Sigma$, with the corresponding cross-link powers reaching their saturation levels. However, as P increases beyond $\bar{\mathsf{P}}_{\mathsf{sat}}$, if any more power is allocated to either q_1 or q_2, a suboptimal sum-rate $R = \Sigma < \min\{\Sigma_1, \Sigma_2\}$ will result. Therefore, it is optimal to maintain q_1 and q_2 at their saturation levels, and to divert all additional increments of P to the direct-links that lead to sum-rate $R = \Sigma = \Sigma_1 = \Sigma_2$.

3.4.4 Evolution of Link Gain Regimes with the Power Budget

In the preceding section, we observed that for a given power budget P and a set of channel parameters, one of the four LGRs is *active* in the sense that the IC-OPA allocates the link powers according to the active LGR. As the budget P increases, the active LGR changes from one to another, thereby producing a sequence of active LGRs, denoted by *LGR path*.

As already mentioned, given a set of parameters $\boldsymbol{a} = (d_1, d_2, c_{12}, c_{21}, \gamma, P)$, we have:

- if $\bar{\mathsf{P}}_1 > 0$, then direct-links are said to be *stronger* than the cross-links in the sense that for the range $0 \leq P \leq \bar{\mathsf{P}}_1$, transmitting as in LGR $\mathcal{A}_{\mathsf{d,d}}$ is sum-rate optimal, whereas there exists no $P > 0$ for which transmitting as in $\mathcal{A}_{\mathsf{c,cd}}$ is optimal;

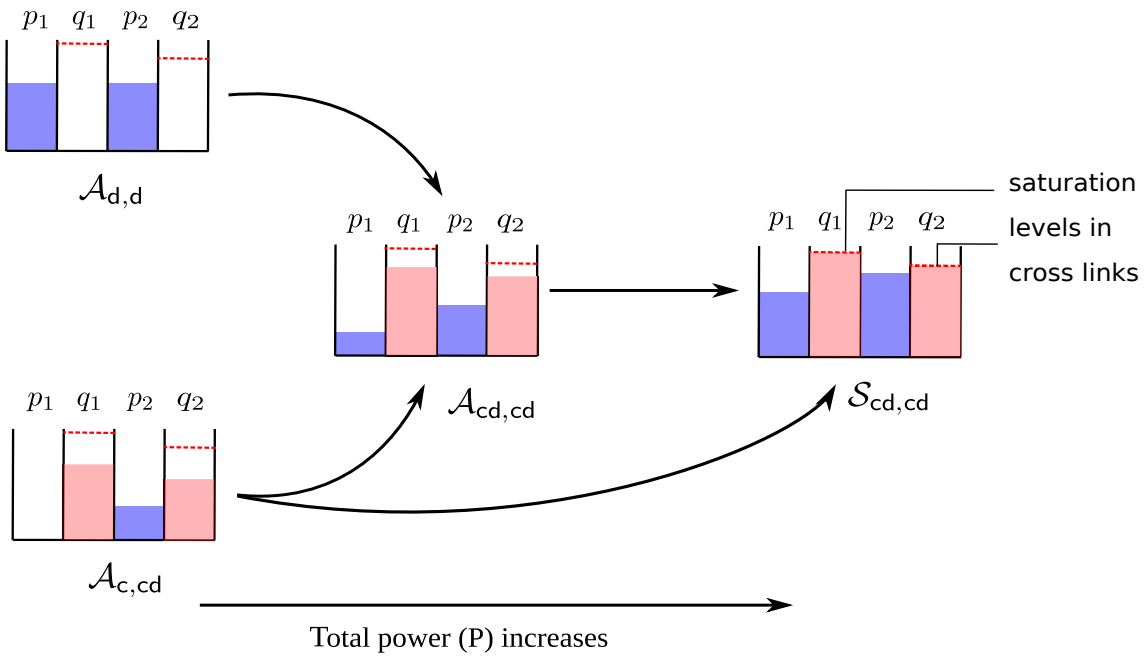


Figure 3.9: A pictorial representation of the four LGRs in the context of the Waterfilling and saturation properties. Here, allocation of power to a link is depicted via the presence of "water-level" in the link, while an absence of water-level denotes that power is not allocated to a link. Saturation levels of the cross-links are also depicted.

- conversely, if $\bar{\mathsf{P}}_2 > 0$, cross-links are *stronger* than the direct-links in the sense that for the range $0 \leq P \leq \min(\bar{\mathsf{P}}_2, \bar{\mathsf{P}}_4)$, transmitting as in LGR $\mathcal{A}_{\mathsf{c,cd}}$ is sum-rate optimal, whereas there exists no $P > 0$ for which transmitting as in $\mathcal{A}_{\mathsf{d,d}}$ is optimal;
- moreover, if $0 < \bar{\mathsf{P}}_3 < \bar{\mathsf{P}}_4 < \bar{\mathsf{P}}_2$, the cross-links are said to be *much stronger* in the sense that the IC-OPA continues transmitting as in $\mathcal{A}_{\mathsf{c,cd}}$ until saturation, beyond which it transmits as in $\mathcal{S}_{\mathsf{cd,cd}}$ skipping $\mathcal{A}_{\mathsf{cd,cd}}$.

LGR-paths for the IC-OPA

We now characterize all LGR paths for the IC-OPA: according to whether the cross-links are weaker, stronger, or much stronger than the direct-links, the IC-OPA follows paths [S1], [S2], and [S3], respectively. The conditions of the paths, and the intervals in which each LGR in the path is active, are given in Table 3.3, and an example is depicted in Figure 3.9. Note that the saturation threshold $\bar{\mathsf{P}}_{\mathsf{sat}}$ on P beyond which the cross-links saturate, depends on the specific LGR path followed as specified below.

Table 3.3: LGR paths for the IC-OPA. Table 3.2 provides the threshold powers in terms of link gains and γ. Each path originates from one of two initial LGRs $\mathcal{A}_{\mathsf{d,d}}$ and $\mathcal{A}_{\mathsf{c,cd}}$, and they terminate at the final LGR $\mathcal{S}_{\mathsf{cd,cd}}$.

LGR Path	**Condition**	**Interval of P in each LGR respectively**
[S1]: $\mathcal{A}_{\mathsf{d,d}} \to \mathcal{A}_{\mathsf{cd,cd}} \to \mathcal{S}_{\mathsf{cd,cd}}$	$\bar{\mathsf{P}}_1 > 0$	$[0,\ \bar{\mathsf{P}}_1),\ \ [\bar{\mathsf{P}}_1,\ \bar{\mathsf{P}}_3),\ \ [\bar{\mathsf{P}}_3,\ \infty)$
[S2]: $\mathcal{A}_{\mathsf{c,cd}} \to \mathcal{A}_{\mathsf{cd,cd}} \to \mathcal{S}_{\mathsf{cd,cd}}$	$\bar{\mathsf{P}}_2 > 0$	$[0,\ \min(\bar{\mathsf{P}}_2, \bar{\mathsf{P}}_4)),\ [\min(\bar{\mathsf{P}}_2, \bar{\mathsf{P}}_4),\ \max(\bar{\mathsf{P}}_3, \bar{\mathsf{P}}_4)),$ $[\max(\bar{\mathsf{P}}_3, \bar{\mathsf{P}}_4),\ \infty)$
[S3]: $\mathcal{A}_{\mathsf{c,cd}} \to \mathcal{S}_{\mathsf{cd,cd}}$	$0 < \bar{\mathsf{P}}_3 < \bar{\mathsf{P}}_4 < \bar{\mathsf{P}}_2$	$[0,\ \bar{\mathsf{P}}_4),\ \ [\bar{\mathsf{P}}_4,\ \infty)$

- If $\bar{\mathsf{P}}_1 > 0$, the IC-OPA follows path [S1]. Since the direct-links are now *stronger*, the IC-OPA allocates P entirely to the direct-links as in $\mathcal{A}_{\mathsf{d,d}}$ for $P < \bar{\mathsf{P}}_1$. However, as P increases, the additional benefit from transmitting only in the direct-links decreases. Hence, when $P \geq \bar{\mathsf{P}}_1$, the IC-OPA begins transmitting in both cross- and direct-links as in $\mathcal{A}_{\mathsf{cd,cd}}$ following its *WF-like* property, which remains optimal for all $\bar{\mathsf{P}}_1 \leq P \leq \bar{\mathsf{P}}_3$. Finally, when $P > \bar{\mathsf{P}}_3$, the *saturation* property comes into effect due to which the cross-links become saturated, and thus the IC-OPA follows the allocation in $\mathcal{S}_{\mathsf{cd,cd}}$. Note that, in this case, the saturation threshold for P is $\bar{\mathsf{P}}_{\mathsf{sat}} = \bar{\mathsf{P}}_3$.

- On the other hand, the IC-OPA follows path [S2] if $\bar{\mathsf{P}}_2 > 0$ but $0 < \bar{\mathsf{P}}_3 < \bar{\mathsf{P}}_4 < \bar{\mathsf{P}}_2$ is not satisfied, i.e., the cross-links are *stronger but not much stronger* than the direct-links. Hence, for all $P < \min\{\bar{\mathsf{P}}_2, \bar{\mathsf{P}}_4\}$, the IC-OPA now transmits in both the cross-links and the S_2-D_2 direct-link as in $\mathcal{A}_{\mathsf{c,cd}}$ following its *WF-like* property. Similar to path [S1], as P increases, the benefit of transmitting only in this subset of links diminishes, and thus $P \geq \min\{\bar{\mathsf{P}}_2, \bar{\mathsf{P}}_4\}$, the IC-OPA transmits in all links as in $\mathcal{A}_{\mathsf{cd,cd}}$. Finally, when $P > \max\{\bar{\mathsf{P}}_3, \bar{\mathsf{P}}_4\}$, the cross-links become *saturated*, and the IC-OPA follows the allocation in $\mathcal{S}_{\mathsf{cd,cd}}$ thereon. in this case, the saturation threshold for P is $\bar{\mathsf{P}}_{\mathsf{sat}} = \max(\bar{\mathsf{P}}_3, \bar{\mathsf{P}}_4)$.

  Note that, whenever $\bar{\mathsf{P}}_2 > 0$, the IC-OPA follows path [S2] irrespective of how $\bar{\mathsf{P}}_2, \bar{\mathsf{P}}_3$, and $\bar{\mathsf{P}}_4$ compare, except in two cases: (a) $0 < \bar{\mathsf{P}}_3 < \bar{\mathsf{P}}_4 < \bar{\mathsf{P}}_2$, in which case the IC-OPA follows path [S3], described next, and (b) $0 < \bar{\mathsf{P}}_4 < \bar{\mathsf{P}}_3 < \bar{\mathsf{P}}_2$, which is infeasible as they violate the mutual exclusiveness of $\mathcal{A}_{\mathsf{c,cd}}$ and $\mathcal{S}_{\mathsf{cd,cd}}$.

- Finally, when $0 < \bar{\mathsf{P}}_3 < \bar{\mathsf{P}}_4 < \bar{\mathsf{P}}_2$, the IC-OPA follows path [S3]. In this case, the cross-links are *much stronger* than the direct-links, and thus similar to [S2], for $0 \leq P < \bar{\mathsf{P}}_4$ the IC-OPA allocates power to both cross-links and a direct-link as in $\mathcal{A}_{\mathsf{c,cd}}$. However, as P increases, due to having much stronger cross-links, as much power is allocated to the cross-links as possible until they become saturated at $P \geq \bar{\mathsf{P}}_4$, and then the IC-OPA begins assigning power to all channels as in $\mathcal{S}_{\mathsf{cd,cd}}$. Hence, compared to [S2], $\mathcal{A}_{\mathsf{cd,cd}}$ is skipped. Note that in this case, the saturation threshold is $\bar{\mathsf{P}}_{\mathsf{sat}} = \bar{\mathsf{P}}_4$.

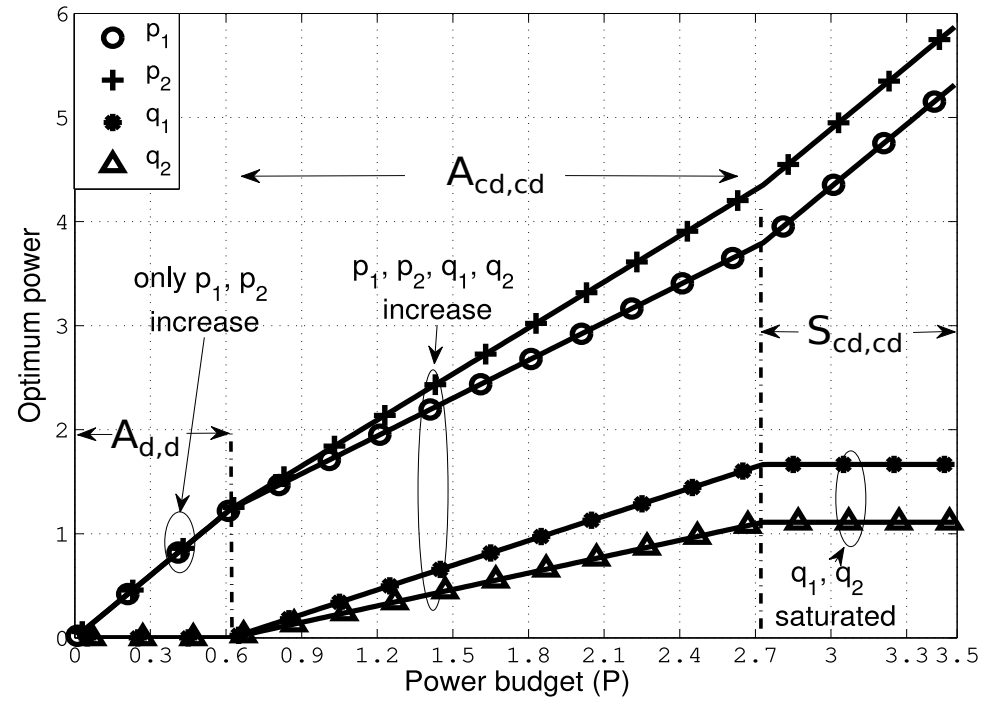
Numerical Examples

We illustrate the characteristics of the IC-OPA with numerical examples of paths [S1] and [S2] for the following parameters, $a_{12}^2 = a_{21}^2 = 1.5, P_1 = P_2 = 5, \alpha = 2, \beta = 0.5$. Note that in each plot, an analytical expression (of power or sum-rate), which is given by marker-line, is observed to match its counterpart computed numerically using the `CVX` program [133], given by solid line.

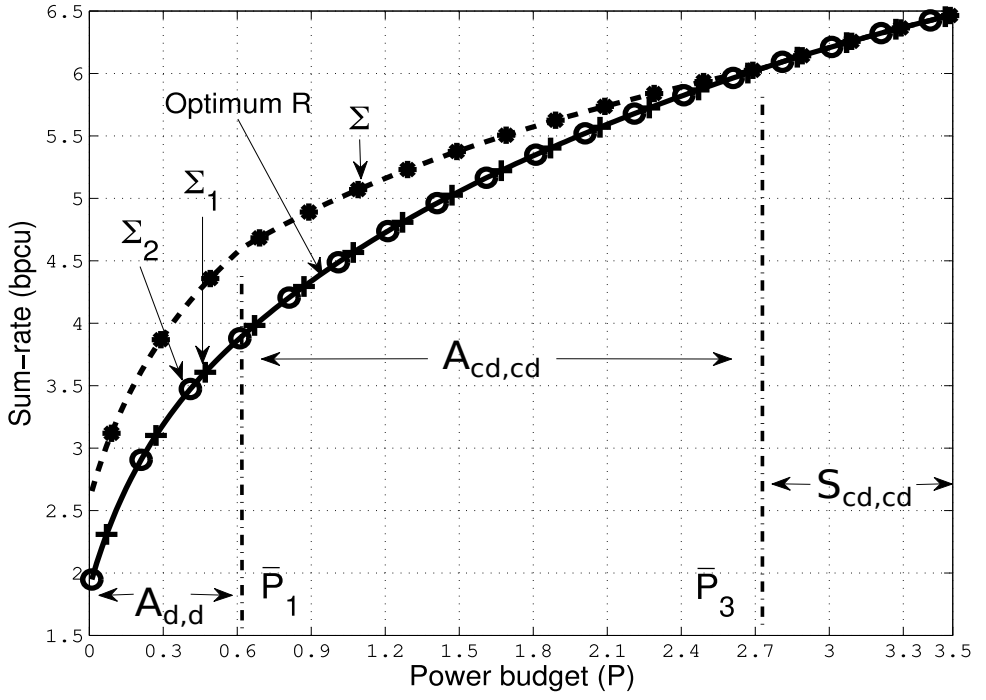
First, we illustrate an example of [S1] by choosing $d_1^2 = 2, d_2^2 = 3, c_{12}^2 = 1, c_{21}^2 = 1.5$, such that the direct-links are stronger than the cross-links in the sense that $\bar{\mathsf{P}}_1 = 0.619$ and $\bar{\mathsf{P}}_2 = -0.75 \notin \mathbb{R}_+$. In Figure 3.10a, we plot the optimal powers against the power budget P, and by labeling the power allocations according to LGRs, we verify that the IC-OPA indeed follows path [S1]: (i) when $0 \leq P < \bar{\mathsf{P}}_1 = 0.619$, the IC-OPA allocates P entirely to the direct-links as in $\mathcal{A}_{\mathsf{d,d}}$, and thus p_1 and p_2 increase with P, and $q_1 = q_2 = 0$; (ii) when $\bar{\mathsf{P}}_1 \leq P \leq \bar{\mathsf{P}}_3 = 2.73$, power is allocated to all links as in $\mathcal{A}_{\mathsf{cd,cd}}$, and thus q_1 and q_2 also increase with P; (iii) finally, when $P > \bar{\mathsf{P}}_3$, the IC-OPA follows $\mathcal{S}_{\mathsf{cd,cd}}$ where the cross-links become saturated simultaneously, and all increments of P are added to p_1 and p_2.

We depict the resulting constraints Σ_1, Σ_2, and Σ in Figure 3.10b. First, note that the IC-OPA preserves $R = \Sigma_1 = \Sigma_2$ for all P. However, there exists a gap between $\Sigma_1 = \Sigma_2$ and Σ in $\mathcal{A}_{\mathsf{d,d}}$ and $\mathcal{A}_{\mathsf{cd,cd}}$. Specifically, in $\mathcal{A}_{\mathsf{d,d}}$, the gap remains constant $(A - A_1)$; in $\mathcal{A}_{\mathsf{cd,cd}}$, it reduces gradually as the IC-OPA transmits in the cross-links, and, in $\mathcal{S}_{\mathsf{cd,cd}}$, it becomes zero as the IC-OPA achieves $R = \Sigma_1 = \Sigma_2 = \Sigma$, as expected.

Next, we illustrate an example of [S2] with the channel gains $d_1^2 = 0.5, d_2^2 = 1, c_{12}^2 = 1.5, c_{21}^2 = 3$, such that the cross-links are stronger than the direct-links in the sense that $\bar{\mathsf{P}}_2 = 0.22$ and $\bar{\mathsf{P}}_1 = -0.21 \notin \mathbb{R}_+$. In Figure 3.11a, we verify that the IC-OPA indeed follows path [S2] by plotting the powers against P and labeling them with LGRs: (i) when $P < \bar{\mathsf{P}}_2 = 0.22$, the IC-OPA follows the allocation in $\mathcal{A}_{\mathsf{c,cd}}$, and thus p_2, q_1 and q_2 increase

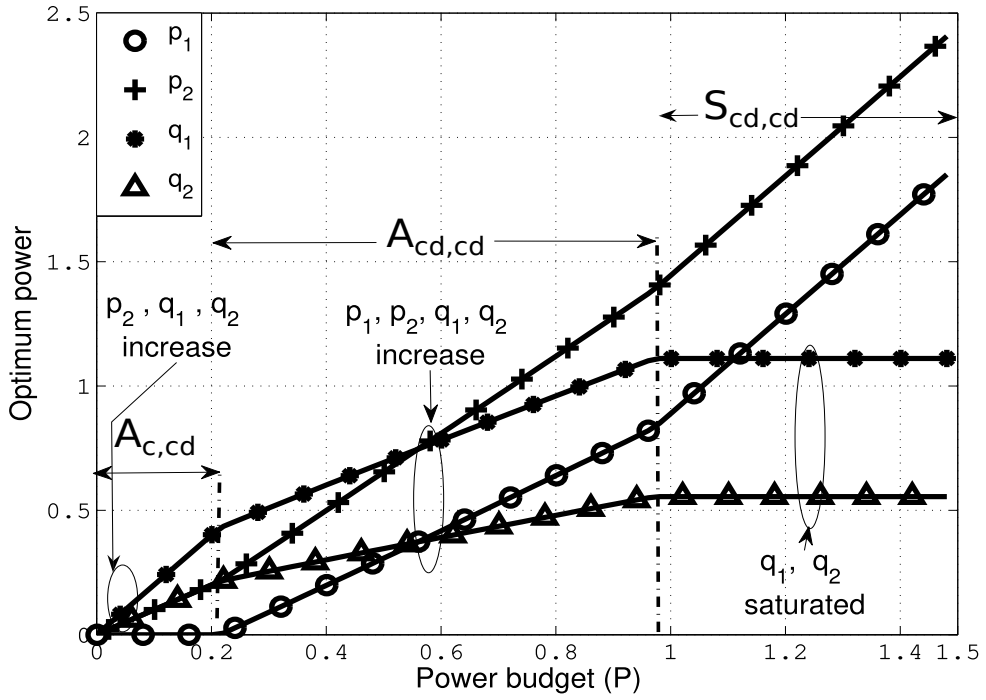


(a)

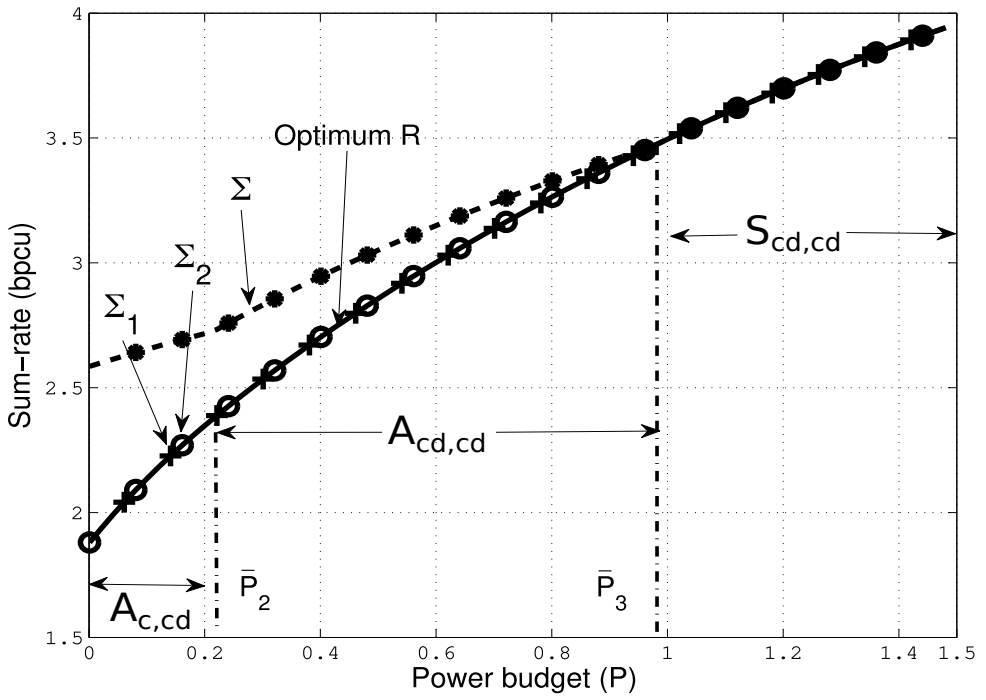


(b)

Figure 3.10: Example of path [S1] $\mathcal{A}_{\mathsf{d,d}} \rightarrow \mathcal{A}_{\mathsf{cd,cd}} \rightarrow \mathcal{S}_{\mathsf{cd,cd}}$: (a) The optimum power allocation. (b) The resulting sum-rate constraints.



(a)



(b)

Figure 3.11: Example of path [S2] $\mathcal{A}_{\mathsf{c,cd}} \rightarrow \mathcal{A}_{\mathsf{cd,cd}} \rightarrow \mathcal{S}_{\mathsf{cd,cd}}$. (a) The optimum power allocation. (b) The resulting sum-rate constraints.

with P, whereas $p_1 = 0$; (ii) when $\bar{\mathsf{P}}_2 \leq P \leq \bar{\mathsf{P}}_3 = 0.97$, the IC-OPA allocates power to all links as in $\mathcal{A}_{\mathsf{cd,cd}}$, and thus p_1 now increases with P as well; (iii) finally, when $P > \bar{\mathsf{P}}_3$, the IC-OPA follows $\mathcal{S}_{\mathsf{cd,cd}}$, and thus the cross-links become saturated simultaneously, as expected in [S2].

In Figure 3.11b, we plot the sum-rate constraints, and note that the IC-OPA achieves $R = \Sigma_1 = \Sigma_2$ in all the sets. In addition, the gap between $\Sigma_1 = \Sigma_2$ and Σ is gradually offset as the IC-OPA transmits in the cross-links in $\mathcal{S}_C$ and $\mathcal{A}_{\mathsf{cd,cd}}$, and it finally becomes zero in $\mathcal{S}_{\mathsf{cd,cd}}$, as expected.

The example of [S3] (i.e., $\mathcal{S}_C \to \mathcal{S}_{\mathsf{cd,cd}}$) is similar to [S2] with $\mathcal{A}_{\mathsf{cd,cd}}$ skipped, as hence omitted.

The Optimal Power Allocation for the Asymmetric DCLIC

The optimal power allocation scheme for the asymmetric case, where condition $A_1 = A_2$ in (3.50) does not necessarily hold, is presented in Appendix B.13.

The Optimal Power Allocation for the Completely Symmetric DCLIC

For more intuition, we now adapt the optimal power allocation scheme to the completely symmetric DCLIC, where the underlying microwave GIC is symmetric, i.e.,

$$a^2 := a_{12}^2 = a_{21}^2, \quad \text{and} \quad P_1 = P_2, \tag{3.60}$$

and the mm-wave cross- and direct-links are also symmetric, i.e.,

$$c^2 := c_{12}^2 = c_{21}^2, \quad \text{and} \quad d^2 := d_1^2 = d_2^2. \tag{3.61}$$

Note that the resulting γ now simples to

$$\tilde{\gamma} := \left(\frac{(1+P_1)^2}{1+P_1+a^2 P_1} \right)^{1/\alpha\bar{\beta}} > 1. \tag{3.62}$$

Due to symmetry, considering only symmetric power allocation of the form (p, q, p, q) is sufficient, and does not cause loss of generality. Moreover, for any feasible (symmetric) power allocation, we have $\Sigma_1 = \Sigma_2$, rendering the constraint $R \leq \Sigma_2$ in (3.53) redundant. Hence, the sum-rate optimization problem now is given by problem $[\mathcal{P}1]$ after deleting constraints (3.53) and (3.56). The resulting problem is solved in the same manner as $[\mathcal{P}1]$, and hence the details are omitted.

Table 3.4: Definition of LGRs and optimal link powers for the completely symmetric DCLIC in terms of channel parameters $\boldsymbol{a} = (d, c, \tilde{\gamma}, P)$, with the threshold powers defined in (3.63) and $\tilde{\gamma}$ in (3.62).

Definition of LGR	**Optimal Power Allocation**	
$\tilde{\mathcal{A}}_{\mathsf{d,d}} := \{\boldsymbol{a} : P \leq \tilde{\mathsf{P}}_1\}$	$p = \frac{P}{\beta}$,	$q = 0$,
$\tilde{\mathcal{A}}_{\mathsf{c,c}} := \{\boldsymbol{a} : P \leq \min(\tilde{\mathsf{P}}_2, \tilde{\mathsf{P}}_4)\}$	$p = 0$,	$p = \frac{P}{\bar{\beta}}$
$\tilde{\mathcal{A}}_{\mathsf{cd,cd}} := \{\boldsymbol{a} : \max(\tilde{\mathsf{P}}_1, \tilde{\mathsf{P}}_2) < P < \tilde{\mathsf{P}}_3\}$	$p = \frac{2}{1+\beta}\Big(P + \frac{\bar{\beta}}{c^2} - \frac{\bar{\beta}}{2d^2}\Big)$,	$q = \frac{1}{1+\beta}\Big(P + \frac{\beta}{d^2} - \frac{2\beta}{c^2}\Big)$
$\tilde{\mathcal{S}}_{\mathsf{cd,cd}} := \{\boldsymbol{a} : \max(\tilde{\mathsf{P}}_3, \tilde{\mathsf{P}}_4) \leq P\}$	$p = \frac{P}{\beta} - \frac{\bar{\beta}(\tilde{\gamma}-1)}{\beta c^2}$,	$q = \frac{\tilde{\gamma}-1}{c^2}$,

In Table 3.4, the 4 LGRs needed to characterize the optimal power allocation for this case are defined along with the optimal transmission powers in terms of parameters $\boldsymbol{a} = (d, c, \tilde{\gamma}, P)$, where the threshold powers are given by

$$\begin{aligned}
\tilde{\mathsf{P}}_1 &:= \beta\big(\frac{2}{c^2} - \frac{1}{d^2}\big), \\
\tilde{\mathsf{P}}_2 &:= \frac{\bar{\beta}}{2}\big(\frac{1}{d^2} - \frac{2}{c^2}\big), \\
\tilde{\mathsf{P}}_3 &:= \frac{(1+\beta)\tilde{\gamma}}{c^2} - \frac{\beta}{d^2} - \frac{\bar{\beta}}{c^2}, \\
\tilde{\mathsf{P}}_4 &:= \frac{\bar{\beta}(\tilde{\gamma}-1)}{c^2}.
\end{aligned} \tag{3.63}$$

The 4 LGRs in Table 3.4, $\tilde{\mathcal{A}}_{\mathsf{d,d}}, \tilde{\mathcal{A}}_{\mathsf{c,c}}, \tilde{\mathcal{A}}_{\mathsf{cd,cd}}$, and $\tilde{\mathcal{S}}_{\mathsf{cd,cd}}$, can be regarded as counterparts of the four LGRs $\mathcal{A}_{\mathsf{d,d}}, \mathcal{A}_{\mathsf{c,cd}}, \mathcal{A}_{\mathsf{cd,cd}}$ and $\mathcal{S}_{\mathsf{cd,cd}}$ defined in Table 3.1.

As expected, the optimal power allocation scheme (IC-OPA) in this case also follows the *Waterfilling-like* (WF-like) and the *saturation* properties as in Section 3.4.3. Moreover, the IC-OPA follows one of the three LGR paths, presented in Table 3.5, but due to symmetry, the channel conditions for the LGR paths can now be more clearly interpreted as detailed below:

- If $c^2 < 2d^2$ (i.e., the cross-links are *weaker* than the direct-links): the IC-OPA follows path [S1]. In this case, the IC-OPA first follows the WF property: for $P \leq \tilde{\mathsf{P}}_1$, it transmits only in the direct-links as in $\tilde{\mathcal{A}}_{\mathsf{d,d}}$, and then for $\tilde{\mathsf{P}}_1 < P < \tilde{\mathsf{P}}_3$, it transmits in all four links as in $\tilde{\mathcal{A}}_{\mathsf{cd,cd}}$. Thus, in $\tilde{\mathcal{A}}_{\mathsf{d,d}}$ and $\tilde{\mathcal{A}}_{\mathsf{cd,cd}}$, the link powers increase piece-wise

Table 3.5: LGR paths for the completely symmetric DCLIC with the threshold powers defined in (3.63). Each path originates from one of two initial LGRs $\tilde{\mathcal{A}}_{\mathsf{d,d}}$ and $\tilde{\mathcal{A}}_{\mathsf{c,cd}}$, and they terminate at the final LGR $\tilde{\mathcal{S}}_{\mathsf{cd,cd}}$.

LGR Path	**Condition**	**Interval of P in each LGR respectively**
[S1]: $\tilde{\mathcal{A}}_{\mathsf{d,d}} \to \tilde{\mathcal{A}}_{\mathsf{cd,cd}} \to \tilde{\mathcal{S}}_{\mathsf{cd,cd}}$	$c^2 < 2d^2$	$[0,\ \tilde{\mathsf{P}}_1),\ [\tilde{\mathsf{P}}_1, \tilde{\mathsf{P}}_3),\ [\tilde{\mathsf{P}}_3, \infty)$
[S2]: $\tilde{\mathcal{A}}_{\mathsf{c,c}} \to \tilde{\mathcal{A}}_{\mathsf{cd,cd}} \to \tilde{\mathcal{S}}_{\mathsf{cd,cd}}$	$2d^2 \leq c^2 < 2d^2\tilde{\gamma}$	$[0, \tilde{\mathsf{P}}_2),\ [\tilde{\mathsf{P}}_2, \min(\tilde{\mathsf{P}}_3, \tilde{\mathsf{P}}_4)),\ [\min(\tilde{\mathsf{P}}_3, \tilde{\mathsf{P}}_4), \infty)$
[S3]: $\tilde{\mathcal{A}}_{\mathsf{c,c}} \to \tilde{\mathcal{S}}_{\mathsf{cd,cd}}$	$c^2 \geq 2d^2\tilde{\gamma}$	$[0, \tilde{\mathsf{P}}_4),\ , [\tilde{\mathsf{P}}_4, \infty)$

linearly with P. Finally, for $P \geq \tilde{\mathsf{P}}_3$ saturation sets in, and the cross-link powers are saturated while all increments of P are allotted only to the direct-links as in $\tilde{\mathcal{S}}_{\mathsf{cd,cd}}$.

- If $2d^2 \leq c^2 < 2d^2\tilde{\gamma}$ (i.e., the cross-links are *stronger* than the direct-links): the IC-OPA now follows path [S2]. Similar to path [S1] above, due to the WF property, the IC-OPA transmits only in the cross-links as in $\tilde{\mathcal{A}}_{\mathsf{c,c}}$ when $P \leq \tilde{\mathsf{P}}_2$, and then transmits in all four links as in $\tilde{\mathcal{A}}_{\mathsf{cd,cd}}$ for $\tilde{\mathsf{P}}_2 < P < \min\{\tilde{\mathsf{P}}_3, \tilde{\mathsf{P}}_4\}$. When $P \geq \min\{\tilde{\mathsf{P}}_3, \tilde{\mathsf{P}}_4\}$, finally saturation occurs as in $\tilde{\mathcal{S}}_{\mathsf{cd,cd}}$.

- If $c^2 \geq 2d^2\tilde{\gamma}$ (i.e., the cross-links are *much stronger* than the direct-links): in this case, the IC-OPA follows path [S3]. Similar to [S2] above, for $P \leq \tilde{\mathsf{P}}_4$, the IC-OPA transmit in the cross-links as in $\tilde{\mathcal{A}}_{\mathsf{c,c}}$. Since the cross-links are much stronger, it continues to transmit only in the cross-links until they become saturated at $P = \tilde{\mathsf{P}}_4$, and beyond that it follows $\tilde{\mathcal{S}}_{\mathsf{cd,cd}}$. Hence, compared to [S2], LGR $\tilde{\mathcal{A}}_{\mathsf{cd,cd}}$ is skipped.

In Figure 3.12, we consider a completely symmetric DCLIC with parameters $a^2 = 1.5, P_1 = 5, \alpha = 2, \beta = 0.5$, and illustrate an example of how the set of cross- and direct-link gains (i.e., the set of (c^2, d^2)) can be partitioned depending on whether the set of cross-links or direct-links are stronger.

The Optimal Power Allocation in the Large Millimeter-Wave Bandwidth Regime

It is interesting to study how the optimal power allocation behaves as the mm-wave bandwidth becomes large. For this study, we consider the completely symmetric DCLIC for which the optimal allocation is given in Table 3.4.

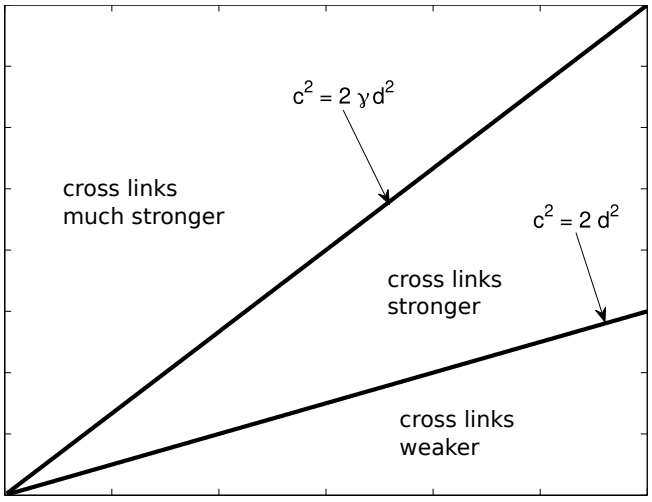


Figure 3.12: Example of the partition of the set of (c^2, d^2) of a completely symmetric DCLIC based on whether the cross-links are much stronger, stronger, or weaker than the direct-links.

Recall that the direct- and cross-links in the DCLIC are modeled as point-to-point AWGN channels with a bandwidth-mismatch factor (BMF) of α. For example, consider the S_1-D_1 direct-link with transmit power p, link gain d_0, and noise variance N, for which the achievable rate is given by $\frac{\alpha\beta}{2}\log\left(1+\frac{d_0^2 p}{N}\right)$ as in (3.47). In the previous settings, since the mm-wave bandwidth (BW) was considered to be a fixed constant, the BMF α and the noise variance N were also taken to be constant, e.g., N was taken to be $N=1$ in (3.6).

When the mm-wave BW increases, the noise variance N as well as the BMF α increases proportionally, which can be modeled by defining $N := N_0 w$ and $\alpha := \alpha_0 w$ for some $N_0, \alpha_0, w > 0$. Then, the large BW regime is modeled by having $w \to \infty$. For the S_1-D_1 direct-link described above, the achievable rate is now given by

$$\frac{\alpha_0 w\beta}{2}\log\left(1+\frac{d_0^2 p}{wN_0}\right) = \frac{\alpha_0 w\beta}{2}\log\left(1+d^2\frac{p}{N_0}\right), \tag{3.64}$$

with the identification $\alpha = \alpha_0 w$ and $d^2 = \frac{d_0^2}{w}$, which capture the effect of varying BW. For the other direct-link and two cross-links, after similarly capturing the effect of varying BW, the resulting link gains and BMF are given by

$$\alpha = \alpha_0 w, \quad d^2 = \frac{d_0^2}{w}, \quad c^2 = \frac{c_0^2}{w}. \tag{3.65}$$

We now examine the effect of large BW, i.e., $w \to \infty$, on the optimal power allocation of Table 3.4. First, parameter $\tilde{\gamma}$, defined in (3.62), simplifies as follows

$$\tilde{\gamma} = \left(\frac{(1+P_1)^2}{1+P_1+a^2 P_1}\right)^{1/\alpha_0\bar{\beta}w}$$

$$= \Gamma^{1/w}$$
$$= e^{\frac{1}{w}\log_e \Gamma}$$
$$\approx 1 + \frac{1}{w}\log_e \Gamma, \quad \text{for large } w, \tag{3.66}$$

where $\Gamma = \left(\frac{(1+P_1)^2}{1+P_1+a^2P_1}\right)^{1/\alpha_0\bar{\beta}} > 1$. As a result, accounting for the link gains in (3.65), for large w, threshold power $\tilde{\mathsf{P}}_4$ simplifies to

$$\tilde{\mathsf{P}}_4 = \frac{\bar{\beta}}{c^2}(\tilde{\gamma} - 1)$$
$$\approx \frac{\bar{\beta}w}{c_0^2}(1 + \frac{1}{w}\log_e \Gamma - 1) = \frac{\bar{\beta}}{c_0^2}\log_e \Gamma > 0. \tag{3.67}$$

Similarly, the other threshold powers simplify to

$$\tilde{\mathsf{P}}_1 = w\beta\left(\frac{2}{c_0^2} - \frac{1}{d_0^2}\right), \quad \tilde{\mathsf{P}}_2 = w\frac{\bar{\beta}}{2}\left(\frac{1}{d_0^2} - \frac{2}{c_0^2}\right),$$
$$\tilde{\mathsf{P}}_4 \approx w\beta\left(\frac{2}{c_0^2} - \frac{1}{d_0^2}\right) + \frac{1+\beta}{c_0^2}\log_e \Gamma.$$

As presented in Table 3.4, the optimal power allocation in the symmetric case is described by 4 LGRs, $\tilde{\mathcal{A}}_{\mathsf{d,d}}, \tilde{\mathcal{A}}_{\mathsf{c,c}}, \tilde{\mathcal{A}}_{\mathsf{cd,cd}}$, and $\tilde{\mathcal{S}}_{\mathsf{cd,cd}}$. Simple algebraic manipulations reveal that as $w \to \infty$, LGR $\tilde{\mathcal{A}}_{\mathsf{cd,cd}}$ approaches the empty set. This leaves only 3 LGRs, and thus the optimal power allocation for large BW, where $w \to \infty$, simplifies as follows:

- if $2d_0^2 \geq c_0^2$, the power budget P is entirely allocated to the direct-links for all $P \geq 0$.
- if $2d_0^2 < c_0^2$, when $P \leq \tilde{\mathsf{P}}_4 \approx \frac{\bar{\beta}}{c_0^2}\log_e \Gamma$, the budget P is entirely allocated to the cross-links; instead, when $P > \tilde{\mathsf{P}}_4 \approx \frac{\bar{\beta}}{c_0^2}\log_e \Gamma$, the cross-links are saturated with power $q \approx \frac{1}{c_0^2}\log_e \Gamma$, and all additional increments of P are allotted to the direct-links only.

In the large BW regime, allocating power in *both* direct and cross-links in Waterfilling-like fashion as in LGR $\tilde{\mathcal{A}}_{\mathsf{cd,cd}}$ is *suboptimal*. In this regime, when transmitting in either the set of direct-links or the set of cross-links, due to $w \to \infty$, the marginal return from allocating additional power to the same set of links remains *constant*, as opposed to reducing as in the regime with finite w. Thus, as P increases, it is beneficial to continue transmitting in the same set of links, and it is suboptimal to allocate a fraction P to the other type of link as in $\tilde{\mathcal{A}}_{\mathsf{cd,cd}}$. For example, when $2d_0^2 \geq c_0^2$, i.e., when budget P is allocated entirely to the direct-links, and thus $p = P$, the achievable rate in the S_1-D_1 direct-link in (3.64) simplifies to the following, as $w \to \infty$

$$\frac{\alpha_0 w\beta}{2}\log\left(1 + \frac{d_0^2 P}{wN_0}\right) \to \frac{\alpha_0 w\beta}{2}\frac{d_0^2 P\log_e 2}{wN_0} = \frac{\alpha_0\beta d_0^2\log_e 2}{2N_0}\ P.$$

Therefore, the marginal return from allocating additional increments of P to the direct-links is constant, and thus for all $P \geq 0$, it is optimal to transmit in the direct-links only as in $\tilde{\mathcal{A}}_{\mathsf{d},\mathsf{d}}$. Alternatively, when $2d_0^2 < c_0^2$, the achievable rates for the cross-links behave similarly as $w \to \infty$, and thus transmitting only in the cross-links is optimal for P smaller than the saturation threshold. However, for P larger than the threshold, the cross-links become saturated, and all additional increments of P are alloted to the direct-links, as in $\tilde{\mathcal{S}}_{\mathsf{cd},\mathsf{cd}}$.

3.4.5 Special Cases and Further Insights

In the following, we present a discussion on some useful insights obtained from the optimal power allocation. Unless otherwise specified, we refer to the power allocation scheme for the original problem [$\mathcal{P}1$], which is described by 4 LGRs in Table 3.1.

The case with very strong interference in the GIC: Recall that the DCLIC consists of an underlying GIC in the microwave band, and the mm-wave cross- and direct-links. In the optimal power allocation for the DCLIC with *strong* underlying GIC presented in Table 3.1, we observed that, if the mm-wave direct-links are stronger than the mm-wave cross-links, then for sufficiently small power budget P, it is optimal to allocate P entirely to the direct-links only (the Waterfilling-like property). However, as P increases such allocation becomes sub-optimal, and sharing the power budget should among all other links are optimal.

In contrast, if the underlying GIC has *very strong interference*, i.e., $a_{12}^2 \geq 1 + P_1$ and $a_{21}^2 \geq 1 + P_2$, it is optimal to allocate P entirely to the direct-links *for all values* of P, irrespective of whether the direct-links are stronger or weaker than the cross-links. Under very strong interference, any feasible power allocation results in $\Sigma \leq \min\{\Sigma_1, \Sigma_2\}$, and therefore, the sum-rate is maximized by maximizing Σ, which is achieved by allocating all of P to the direct-links.

The case with small/large transmit powers in the GIC: If the transmit powers in the underlying GIC with the strong CLIC are small, i.e., $P_1 \approx 0, P_2 \approx 0$, then allocating P entirely to the direct-links as in $\mathcal{A}_{\mathsf{d},\mathsf{d}}$ is approximately optimal for all values of P, in the sense that the difference between the sum-rates achieved with the allocation in $\mathcal{A}_{\mathsf{d},\mathsf{d}}$ and that achieved with the optimal scheme is small. This behavior can be attributed to

the fact that the GIC with strong interference (i.e., $a_{12}^2 \geq 1, a_{21}^2 \geq 1$) and small transmit powers (i.e., $P_1 \approx 0, P_2 \approx 0$) mimics the very strong interference regime (i.e., $a_{12}^2 \geq 1 + P_1$, $a_{21}^2 \geq 1 + P_2$), where allocating P entirely to the direct-links as in $\mathcal{A}_{\mathsf{d,d}}$ is optimal for all P.

On the other hand, if the transmit powers in the GIC are large, i.e., $P_1, P_2 \gg 1$, then it is optimal to allocate power in all links as in $\mathcal{A}_{\mathsf{cd,cd}}$ for all but a relatively short range of P with small values. From Table 3.1, recall that the IC-OPA transmits in a subset of all links such as only direct-links in $\mathcal{A}_{\mathsf{d,d}}$ or two cross- and a direct-link in $\mathcal{A}_{\mathsf{c,cd}}$ when P is sufficiently small, and beyond that it transmits in all links as in $\mathcal{A}_{\mathsf{cd,cd}}$. Recall that in $\mathcal{A}_{\mathsf{cd,cd}}$, the sum-rate is $R = \Sigma_1 = \Sigma_2 < \Sigma$, and as P increases the gap between Σ_1 and Σ decreases and gradually reduces to zero at $P = \bar{\mathsf{P}}_{\mathsf{sat}}$ as saturation sets in. However, if $P_1, P_2 \gg 1$, this gap is large, and a large P is needed to reach saturation. Therefore, for all moderate values of P, allocation in $\mathcal{A}_{\mathsf{cd,cd}}$ remain optimal.

The case with small/large mm-wave cross-link gains: If the mm-wave cross-link gains are large, allocating P entirely to the direct-links as in $\mathcal{A}_{\mathsf{d,d}}$ is approximately optimal for all P, in the sense described in the point immediately above. With large c_{12}^2 and c_{21}^2, the powers needed to saturate the cross-links, namely $q_1 = \frac{\gamma - 1}{c_{12}^2}$ and $q_2 = \frac{\gamma - 1}{c_{21}^2}$, is very small. Thus, the resulting optimal allocation scheme allots a very small fraction of P to the cross-links to saturate them, and redirects the remaining of P (almost all of P) to the direct-links, and this allocation closely resembles that in $\mathcal{A}_{\mathsf{d,d}}$.

The case with limiting values of β: Note that $\beta \in (0, 1)$. When $\beta \approx 1$, the DCLIC can be approximated by the DLIC, and thus allocating P entirely to only the direct-links is approximately optimal. In this case, as P increases, the sum-rate continues to increase with P since the direct-link rate continue to increase.

On the other hand, when $\beta \approx 0$, the DCLIC can be approximated by the CLIC where allocating P entirely to the cross-links is approximately optimal. However, in this case, as P increases, and reaches the saturation threshold, the sum-rate saturates to Σ. With in the cross-links already saturated, and no direct-links (i.e., $\beta \approx 0$) additional increments of P does not result in the increase of the sum-rate anymore.

3.5 Summary

In this chapter, we studied the performance of a two-user interference channel over the integrated mm-wave/microwave dual-band, where the mm-wave links from a source to its designated and the non-designated destinations are modeled as AWGN direct-link and cross-link respectively. We first showed that the capacity region of a DCLIC can be *decomposed* into the capacity region of the underlying CLIC and that of the direct-links, and hence the direct-links can be operated independently of the CLIC without compromising optimal rates. Next, we characterized the capacity region of the CLIC with strong underlying GIC, and then obtained *sufficient channel conditions* under which the capacity region of the CLIC with weak and mixed underlying GIC is characterized. Then, for the weak and mixed CLICs, an *approximate capacity result* was obtained that characterized the capacity region to within 1/2 bits per channel use per user.

We then characterized the *optimal power allocation* over the mm-wave direct-links and cross-links that maximizes the sum-rate of the DCLIC. The optimal power allocation scheme, which was characterized in closed form, allocates power to the links in *different modes* based on whether certain conditions on the link gains and budget P hold (i.e., based on which link gain regime is active). In particular, the optimal allocation has a *Waterfilling-like* property due to which, when the power budget P is small, it assigns the budget entirely to either both direct-links, or both cross-links and at most one direct-link. As the budget P is increased, the optimal allocation eventually assigns power to all links. As the power budget is increased further, due to the *saturation* property, a maximum limit is imposed on the power of the cross-links, while the power allocated to the direct-links increase linearly with the budget P. Moreover, in the *large* mm-wave bandwidth regime, if the direct-links are stronger than the cross-links, budget P should be allocated entirely to the direct-links for all $P \geq 0$. In contrast, if the cross-links are stronger, P should be allocated entirely to the cross-links until they saturate, and then all subsequent increments of P should be allotted to the direct-links only.

Chapter 4

The Dual-Band Multiple Access Relay Channel

In this chapter[1], we study the performance of a two-source single-relay dual-band multiple-access relay channel (MARC), denoted the dual-band MARC. As discussed in Section 2.2.3, relay cooperation plays a key role in improving the performance of microwave networks [18, 77], and will likely play a vital role in the dual-band networks as well, especially to offset mm-wave link impairments such as blockage and limited range [37, 40, 64, 76].

In the immediate future, the dual-band MARC can model uplink scenarios such as the *fixed wireless access* [77] which is expected to replace last-mile wired connections to end users with high speed wireless connection in 5G. Since dual-band communication typically requires a more complex hardware than microwave-only communication, in the immediate future, while base stations and fixed access-points are likely to be equipped with hardware capable of communicating over a dual-band, a substantial fraction of end user devices (e.g., mobile handsets) may be capable of conventional microwave-only communication. In this setting, the base station can provide high data-rate connectivity to fixed access-points due to the additional spectrum in the mm-wave band, while the fixed access points, likely to be located outside a building, can then provide high data-rate access to end users, e.g., users inside the building. In this setting, the dual-band MARC can model *relay-assisted uplink* from two such fixed access-points located in nearby buildings.

In the near to medium future, when mobile handsets are also equipped with dual-band communication capable hardware, the dual-band MARC can model relay-aided cellular

[1]The results of this chapter (except for Section 4.3) have been published in [134].

uplink from mobile users to the base station [67, 71, 90].

In the dual-band MARC studied in this chapter, two sources (S_1 and S_2) communicate to a destination (D) with the help of a relay (R) over an integrated mm-wave/microwave dual-band. In the underlying microwave band, transmissions from both sources are *superimposed* at the relay and at the destination as in a conventional MARC, denoted c-MARC, which operates over the microwave band only [77].

In contrast, transmissions in the mm-wave band are produced by using highly directional co-phased antenna arrays [30], and thus the resulting mm-wave links are modeled as AWGN links as in the dual-band IC in Chapter 2. Moreover, as discussed in Section 1.3, if *hybrid* antenna array systems are used for mm-wave communication, the set of available antenna elements can be reconfigured into *multiple* smaller and independent antenna arrays, each of which can then be digitally controlled to produce a separate beam to communicate to a specific receiver [72–74]. Hence, using hybrid antenna arrays, a mm-wave transmitter is able to communicate independent information to *multiple* receivers *in parallel*.

In the dual-band MARC, a mm-wave transmitter is assumed to deploy a *hybrid* antenna array, and thus it is able to create two *parallel* non-interfering links to communicate with both the relay and the destination *simultaneously* [72, 74, 135]. Similarly, a mm-wave receiver is modeled as being able to simultaneously receive transmissions from multiple mm-wave transmitters via *separate* mm-wave links with negligible inter-link interference [136].

It is therefore natural to ask whether a source in the mm-wave band should transmit to the relay, the destination, or both simultaneously. Depending on whether each of the two sources transmits to only the relay, only the destination, both, or none, 16 different models of the dual-band MARC are possible. The general model that includes all microwave and mm-wave links is referred to as the destination-and-relay-linked MARC (DR-MARC).

In the DR-MARC, sources S_1 and S_2 simultaneously communicate to the destination D via the mm-wave S_1-D and S_2-D *direct-links*, as well as the relay R via the mm-wave S_1-R and S_2-R *relay-links*. All other models with varying mm-wave link connectivity can be obtained from the DR-MARC by setting the relevant transmit powers and link gains to zero, and hence they are not defined explicitly. However, one specific model, where only the mm-wave S_1-R and S_2-R relay-links are active in the mm-wave band, is an important variant, and hence is referred to as the relay-linked MARC (R-MARC). Clearly, the R-MARC can be obtained from the DR-MARC by setting the transmit powers in both S_1-D and S_2-D mm-wave direct-links to zero.

Note that in addition to the 4 mm-wave links, the dual-band MARC also consists of an underlying conventional MARC (c-MARC) in the microwave band [77]. When such an individual c-MARC is subject to phase and Rayleigh fading, the capacity of the *near* c-MARC is characterized in closed form [77, Theorem 9], as detailed in Section 2.2.3. As such, for consistency, we assume that the underlying microwave c-MARC of the dual-band MARC is subject to a general ergodic fading where the phase of the fading coefficients are i.i.d. uniform in $[0, 2\pi)$, similar to phase and Rayleigh fading. This general fading contains phase and Rayleigh fading as special cases, and thus can model a wide range of channel impairments. For instance, phase fading models the effect of oscillator phase noise in high-speed time-invariant communications [121], the effect of phase-change due to slight transmitter-receiver misalignment in links with a strong LoS component [124], etc., while Rayleigh fading models the effect of rich scattering [120].

First, we consider the DR-MARC where all 4 mm-wave links are present, i.e., the sources simultaneously transmit in both the mm-wave relay-links and direct-links. We show that its capacity can be *decomposed* into the capacity of the *underlying* R-MARC and the two mm-wave direct-links. Hence, thereon we focus on the R-MARC.

Note that the R-MARC consists of an underlying c-MARC and two mm-wave relay-links. As detailed in Section 2.2.3, the capacity of *near* c-MARCs was characterized in closed form in [77, Theorem 9], where the relay is *near* the sources in the sense that the source-to-relay achievable rates are larger than the source-to-destination achievable rates. Similar to the near c-MARCs, we focus on the class of *jointly-near* R-MARCs that satisfy certain sufficient channel conditions over both bands (precisely defined in Section 4.2.4), under which the capacity of R-MARCs are characterized in closed form. The jointly-near conditions are more general than the near conditions in [77, Theorem 9] in the sense that when the R-MARC satisfies the jointly-near conditions, the underlying c-MARC need not necessarily satisfy its near conditions.

Moreover, the DR-MARC is a basic building block for future dual-band multiuser networks that captures the effect of relay-cooperation. Since, its performance will be significantly affected by the mm-wave links due to their point-to-point nature and relatively larger bandwidths [23,60,128], it is useful to understand how allocating the mm-wave band resources optimizes the performance, similar to that for the dual-band IC in Section 3.4 and other multiuser networks [13, 37, 128]. Therefore, to quantify the impact of the mm-wave spectrum on the performance of the DR-MARC, we study the power allocation strategy for the mm-wave direct-links and relay-links (subject to a power budget) that maximizes

the *achievable sum-rate* of the DR-MARC.

The contributions of this chapter are summarized as follows:

Capacity Results (presented in Section 4.2 and Section 4.3)

- We show that the capacity region of the DR-MARC can be *decomposed* into the capacity region of the *underlying* R-MARC and the two mm-wave *direct-links*. This shows that operating the R-MARC independently of the direct-links is optimal, which simplifies the encoding/decoding operation. Hence, the direct-links are able to improve the rates of individual users, while the relay-links play a non-trivial role in characterizing the capacity of the R-MARC.
- We then characterize an achievable region for the R-MARC based on a block-Markov encoding scheme, which performs joint encoding over both bands.
- For the R-MARC, we obtain a set of sufficient channel conditions, denoted the *jointly-near* conditions, under which the capacity of the R-MARC is characterized by the aforementioned achievable scheme. These jointly-near conditions show that even when the microwave source-relay channels are not strong enough, if the mm-wave source-relay links are sufficiently strong, then a closed form capacity result is characterized.
- We also extend these results to the K-user version of the DR-MARC and the R-MARC, where $K > 2$.

Resource Allocation in the DR-MARC (presented in Section 4.4)

We characterize the optimal power allocation scheme over the mm-wave direct-links and relay-links of the DR-MARC that maximizes the sum-rate achievable on the channel with the aforementioned achievable scheme, where the two mm-wave links from each source are subject to a total power budget P. For notational convenience, this optimal power allocation scheme for the DR-MARC is denoted the MARC-OPA. For intuition, we partition the entire range of the power budget P into several *link gain regimes* (LGR) such that that the MARC-OPA allocates link powers in different *modes* in each LGR. We obtain all such LGRs and modes of power allocation which reveal useful insights.

We observe that for a class of DR-MARCs, where the *jointly-near* condition in the underlying R-MARC is satisfied by the source-relay *microwave* channel gains only, allocating the entire budget P to the direct-links is optimal for all $P \geq 0$.

In contrast, for DR-MARCs where the jointly-near condition in the underlying R-MARC cannot be satisfied by the microwave channel gains only, and hence must be satisfied jointly by *both bands*, the optimal power allocation (the MARC-OPA) has the following properties:

- when P is *smaller* than a certain *saturation threshold* ($\mathsf{P}_{\mathsf{sat}}$), for the direct and relay-links of each source, the MARC-OPA allocates powers following a Waterfilling (WF) approach. Specifically, for sufficiently small P, the MARC-OPA allocates P entirely to the strongest of the direct-link and the relay-link of each source, and as P increases, it eventually allocates power to the remaining links. Thus, for $P < \mathsf{P}_{\mathsf{sat}}$, each link-power either increases piecewise linearly with P, or remains zero.

- when $P \geq \mathsf{P}_{\mathsf{sat}}$, *saturation* occurs where the power in both relay-links are *constrained* to satisfy a certain saturation condition. As P increases beyond $\mathsf{P}_{\mathsf{sat}}$, the relay-link powers vary with P as follows: there exists a *final* threshold $\mathsf{P}_{\mathsf{fin}} \geq \mathsf{P}_{\mathsf{sat}}$, such that

  1. if one relay-link is *significantly stronger* than the other (in a sense to be defined later), then for all $P \geq \mathsf{P}_{\mathsf{fin}}$, power in the stronger relay-link remains fixed at a constant level, while power in the weaker relay-link remains zero.
  2. in contrast, if one relay-link is only *stronger* but not significantly stronger than the other (in a sense to be defined later), for all $P \geq \mathsf{P}_{\mathsf{fin}}$, power in the stronger and the weaker relay-links monotonically increase and decrease respectively, and they approach positive constant levels as P grows.

  However, as opposed to the relay-links, as P increases beyond $\mathsf{P}_{\mathsf{sat}}$, the direct-link powers grow unbounded with P.

- when the mm-wave bandwidth is large, the optimal power allocation for the symmetric case simplifies as follows: (a) if the direct-links are stronger than the relay-links, allocating P entirely to the direct-links is optimal for all $P \geq 0$; (b) alternatively, if the relay-links are stronger, for P smaller than the saturation threshold, P should be allocated entirely to the relay-links, whereas for P larger than the threshold, the relay-links become saturated, and all increments of P should be allotted to the direct-links.

The rest of the chapter is organized as follows: The system model is presented in Section 4.1. The capacity results and some numerical examples are presented in Section 4.2. The capacity results are then extended to the K-user case in Section 4.3. In Section 4.4, the optimal resource allocation problem is presented in along with some discussion on insights derived from the optimal power allocation. Finally, a summary of results is provided in Section 4.5.

4.1 System Model

In Figure 4.1a, a cellular uplink scenario is depicted where two users (sources) communicate to a base station (destination) with the help of a relay. This two-user relay-assisted uplink is modeled by a Gaussian DR-MARC, denoted hereon by DR-MARC and depicted in Figure 4.1b.

Recall that a bandwidth mismatch factor (BMF) α may exist between the microwave and the mm-wave bands. Hence, for n uses of the microwave channels, the mm-wave links are used $n_1(n) := \lfloor \alpha n \rfloor$ times, where $\lfloor x \rfloor$ denotes the largest integer no larger than x. Moreover, since in the mm-wave band the sources are able to communicate to both the destination and the relay *simultaneously* via the mm-wave direct- and relay-links, for n uses of the microwave channels, both the mm-wave relay-links and direct-links are assumed to be used $n_1(n) = \lfloor \alpha n \rfloor$ times.

Channel Model

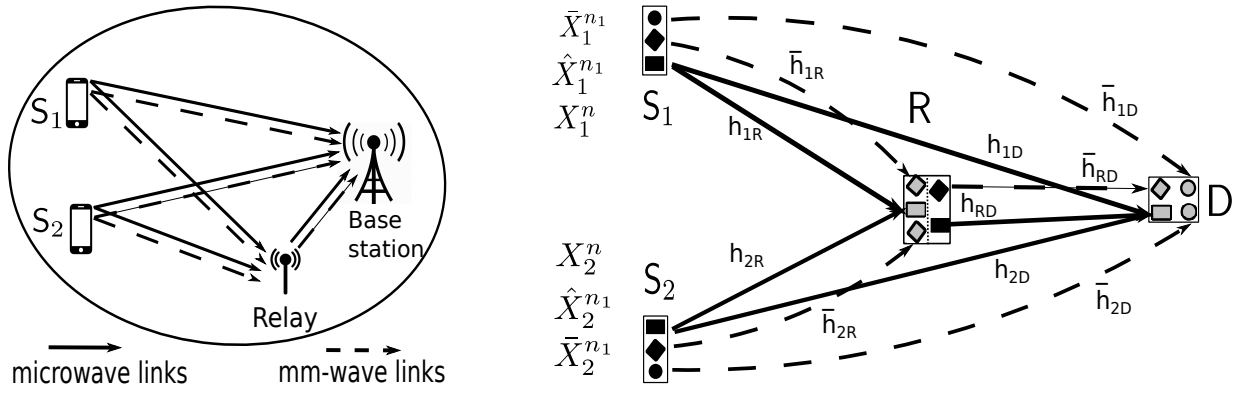
We now define the channel model of the DR-MARC. In the microwave (first) band, the outputs at destination D and relay R at the i-th use of the channel are given by [77]

$$Y_{\mathsf{D},i} = H_{1\mathsf{D},i}X_{1,i} + H_{2\mathsf{D},i}X_{2,i} + H_{\mathsf{RD},i}X_{\mathsf{R},i} + Z_{\mathsf{D},i} \tag{4.1}$$

$$Y_{\mathsf{R},i} = H_{1\mathsf{R},i}X_{1,i} + H_{2\mathsf{R},i}X_{2,i} + Z_{\mathsf{R},i}, \qquad i = 1, \ldots, n, \tag{4.2}$$

where $H_{k\mathsf{D},i} \in \mathbb{C}$ are channel fading coefficients from source S_k to destination D, $H_{k\mathsf{R},i} \in \mathbb{C}$ are the same from source S_k to relay R , $k \in \{1, 2\}$, and $H_{\mathsf{RD},i} \in \mathbb{C}$ are the same from R to D. The input symbols $X_{m,i} \in \mathbb{C}$ are block power constrained, $\frac{1}{n}\sum_{i=1}^{n} \mathbb{E}[|X_{m,i}|^2] \leq P_m, m \in \{1, 2, \mathsf{R}\}$. Also, the noise RVs are $Z_{\mathsf{R},i} \sim \mathcal{CN}(0, 1)$, i.i.d., and $Z_{\mathsf{D},i} \sim \mathcal{CN}(0, 1)$, i.i.d.

The outputs of the mm-wave S_1-R and S_2-R *relay-links* from sources S_1 and S_2 to relay



(a) (b)

Figure 4.1: (a) Example of the DR-MARC in cellular uplink: sources S_1 and S_2 communicate to the base station with the help of a relay. (b) System model of the Gaussian DR-MARC: solid lines denote microwave links while dashed lines denote mm-wave links.

R are modeled as

$$\bar{Y}_{\mathsf{1R},\ell} = \bar{H}_{\mathsf{1R},\ell}\hat{X}_{1,\ell} + \bar{Z}_{\mathsf{1R},\ell}, \tag{4.3}$$

$$\bar{Y}_{\mathsf{2R},\ell} = \bar{H}_{\mathsf{2R},\ell}\hat{X}_{2,\ell} + \bar{Z}_{\mathsf{2R},\ell}, \qquad \ell = 1, \ldots, n_1, \tag{4.4}$$

respectively, where $\bar{H}_{k\mathsf{R},\ell}$ are the fading coefficients of the mm-wave S_k-R relay-links. Here, the input symbols $\hat{X}_{k,\ell} \in \mathbb{C}$ are block power constrained, $\frac{1}{n_1}\sum_{\ell=1}^{n_1}\mathbb{E}[|\hat{X}_{k,\ell}|^2] \leq \hat{P}_k$, and the noise RVs are $\bar{Z}_{k\mathsf{R},\ell} \sim \mathcal{CN}(0,1), k \in \{1,2\}$, i.i.d.

Similarly, the outputs of the mm-wave S_1-D and S_2-D *direct-links* from sources S_1 and S_2 to destination D are modeled as

$$\bar{Y}_{\mathsf{1D},\ell} = \bar{H}_{\mathsf{1D},\ell}\bar{X}_{1,\ell} + \bar{Z}_{\mathsf{1D},\ell}, \tag{4.5}$$

$$\bar{Y}_{\mathsf{2D},\ell} = \bar{H}_{\mathsf{2D},\ell}\bar{X}_{2,\ell} + \bar{Z}_{\mathsf{2D},\ell}, \qquad \ell = 1, \ldots, n_1, \tag{4.6}$$

while the outputs of the mm-wave R-D link is modeled as

$$\bar{Y}_{\mathsf{RD},\ell} = \bar{H}_{\mathsf{RD},\ell}\bar{X}_{\mathsf{R},\ell} + \bar{Z}_{\mathsf{RD},\ell}, \qquad \ell = 1, \ldots, n_1, \tag{4.7}$$

where $\bar{H}_{k\mathsf{D},\ell}$ and $\bar{H}_{\mathsf{RD},\ell}$ are the fading coefficients of the mm-wave S_k-D direct-links and the mm-wave R-D link, respectively. The input symbols $\bar{X}_{m,\ell} \in \mathbb{C}$ are block power constrained, $\frac{1}{n_1}\sum_{\ell=1}^{n_1}\mathbb{E}[|\bar{X}_{m,\ell}|^2] \leq \bar{P}_m$, while the noise RVs are $\bar{Z}_{m\mathsf{D},\ell} \sim \mathcal{CN}(0,1), m \in \{1,2,\mathsf{R}\}$, i.i.d.

In the DR-MARC, to communicate a message M_k from source S_k, it is encoded into three codewords, $X_k^n(M_k), \hat{X}_k^{n_1}(M_k)$ and $\bar{X}_k^{n_1}(M_k)$, of lengths n, n_1 and n_1 respectively. Then, $X_k^n(M_k)$ is transmitted towards D by using the microwave channel n times. Due to

the nature of this band, $X_1^n(M_1)$ and $X_2^n(M_2)$ superimpose at D and at R as in the c-MARC [77], described by (4.1)-(4.2). Meanwhile, in the mm-wave band, codeword $\hat{X}_k^{n_1}(M_k)$ is transmitted to R through the S_k-R relay-link, while codeword $\bar{X}_k^{n_1}(M_k)$ is transmitted to D via the S_k-D direct-link, *simultaneously* by using both links n_1 times. The relay aids in communication by creating codewords X_{R}^n and $\bar{X}_{\mathsf{R}}^{n_1}$ from its received signals, and transmitting them to D via the microwave and mm-wave bands respectively.

We formally define the encoding/decoding procedure shortly.

Fading Model and Channel State Information

We assume that the DR-MARC is subject to an ergodic fading process where the fading coefficients across channel uses are i.i.d., such that the *phase* of the fading coefficients are i.i.d., random variables in $[0, 2\pi)$, while the magnitude of the fading coefficients are i.i.d., non-negative random variables. Specifically, the channel fading coefficients $H_{mt,i}$ in (4.1)-(4.2) for the first band are defined as

$$H_{mt,i} := \sqrt{G_{mt,i}} e^{j\Theta_{mt,i}}, \tag{4.8}$$

where the magnitude of fading coefficients are $G_{mt,i} \in \mathbb{R}_+$, i.i.d., and the phase of the fading coefficients $\Theta_{mt,i}$ are uniformly distributed in $[0, 2\pi)$, i.e., $\Theta_{mt,i} \sim \mathcal{U}[0, 2\pi)$, i.i.d., with $j := \sqrt{-1}$, $m \in \{1, 2, \mathsf{R}\}, t \in \{\mathsf{R}, \mathsf{D}\}, m \neq t$. Moreover, the channel fading coefficients $\bar{H}_{mt,\ell}$ in the second band are defined as

$$\bar{H}_{mt,\ell} := \sqrt{\bar{G}_{mt,\ell}} e^{j\bar{\Theta}_{mt,\ell}}, \tag{4.9}$$

where the magnitude of fading coefficients are $\bar{G}_{mt,\ell} \in \mathbb{R}_+$, i.i.d., and the phase of the fading coefficients $\bar{\Theta}_{mt,\ell} \in [0, 2\pi)$ are i.i.d., $m \in \{1, 2, \mathsf{R}\}, t \in \{\mathsf{R}, \mathsf{D}\}, m \neq t$.

Note that unlike phase $\Theta_{mt,i}$ in the microwave band which is *uniformly distributed* in $[0, 2\pi)$, phase $\bar{\Theta}_{mt,\ell} \in [0, 2\pi)$ in the mm-wave band can be distributed according to any appropriate distribution. Moreover. the magnitude of fading coefficients $G_{mt,i}$ and $\bar{G}_{mt,\ell}$ depend on the distance d_{mt} between transmitter node m and receiver node t, as well as pathloss exponent β_1 (for the first band) and β_2 (for the second band). This general fading specializes to cases where both bands are under phase or Rayleigh fading as follows:

- it specializes to phase fading by taking fading magnitudes $G_{mt,i}$ and $\bar{G}_{mt,\ell}$ to be constants that depend on the geometry of the network and path-loss coefficients, as well as taking mm-wave fading phases to be uniform, i.e., $\bar{\Theta}_{mt,\ell} \sim \mathcal{U}[0, 2\pi)$, i.i.d.;

- it specializes to Rayleigh fading by taking fading magnitudes $G_{mt,i} \sim \mathsf{exp}(\mu_{mt})$, and $\bar{G}_{mt,\ell} \sim \mathsf{exp}(\bar{\mu}_{mt})$, i.i.d., in which $\mathsf{exp}(\mu)$ is an exponential distribution with mean μ, and μ_{mt} and $\bar{\mu}_{mt}$ are constants that depend on the network-geometry and path-loss coefficients, as well as taking mm-wave fading phases to be $\bar{\Theta}_{mt,\ell} \sim \mathcal{U}[0, 2\pi)$, i.i.d..

We also assume the following on the availability of *channel state information* (CSI) at the receiver: (i) the long term parameters, i.e., the distances between transmitters and receivers, and the pathloss exponents, are known at all nodes; (ii) the instantaneous CSI, i.e., the phase and magnitude of the fading coefficients, are *not available* to any transmitter; and (iii) each receiver knows the instantaneous CSI on its *incoming* channels only, but has no CSI on other channels. These assumptions model practical scenarios where CSI feedback to a transmitter is unavailable, while a receiver can reliably estimate the CSI [13, 137].

Channel Code

We now define a code and an achievable rate pair for the DR-MARC. Note that given a BMF α, for n uses of the microwave band, the mm-wave band is used $n_1(n) := \lfloor \alpha n \rfloor$ times, while for n_1 uses in the mm-wave band, the microwave band is used $n(n_1) := \lfloor n/\alpha \rfloor$ times.

Definition 4.1 (A code for the DR-MARC)**.** *A $(2^{nR_1}, 2^{nR_2}, n, \alpha)$ code for the DR-MARC consists of (i) two independent, uniformly distributed message sets $\mathcal{M}_k = \{1, \ldots, 2^{nR_k}\}, k \in \{1, 2\}$, one for each source S_1 and S_2; (ii) two encoders ϕ_1 and ϕ_2 such that $\phi_k : \mathcal{M}_k \to \mathbb{C}^n \times \mathbb{C}^{n_1(n)} \times \mathbb{C}^{n_1(n)}, k \in \{1, 2\}$, one each for S_1 and S_2; (iii) a set of causal relay encoding functions, $\{f_i\}_{i=1}^n$ and $\{\bar{f}_\ell\}_{\ell=1}^{n_1(n)}$, such that*

$$\begin{aligned} x_{\mathsf{R},i} &= f_i(y_{\mathsf{R}}^{i-1}, \{h_{k\mathsf{R}}^{i-1}, \bar{y}_{k\mathsf{R}}^{n_1(i-1)}, \bar{h}_{k\mathsf{R}}^{n_1(i-1)}\}_{k=1}^2), \text{ and} \\ \bar{x}_{\mathsf{R},\ell} &= \bar{f}_\ell(y_{\mathsf{R}}^{n(\ell-1)}, \{\bar{y}_{k\mathsf{R}}^{\ell-1}, \bar{h}_{k\mathsf{R}}^{\ell-1}, h_{k\mathsf{R}}^{n(\ell-1)}\}_{k=1}^2), \end{aligned} \tag{4.10}$$

with $x_{\mathsf{R},i}, \bar{x}_{\mathsf{R},\ell} \in \mathbb{C}$; and (iv) a decoder ψ at D such that $\psi : \mathbb{C}^n \times \mathbb{C}^{3n_1(n)} \times \mathbb{C}^{3n} \times \mathbb{C}^{3n_1(n)} \to \mathcal{M}_1 \times \mathcal{M}_2$, with $n_1(n) := \lfloor \alpha n \rfloor$ and $n(n_1) := \lfloor n/\alpha \rfloor$.

Here, the relay helps by computing codewords $\{x_{\mathsf{R},i}\}_{i=1}^n$ and $\{\bar{x}_{\mathsf{R},\ell}\}_{\ell=1}^{n_1(n)}$ causally by applying relay-encoding functions $\{f_i\}_{i=1}^n$ and $\{\bar{f}_\ell\}_{\ell=1}^{n_1(n)}$ on its past received signals and CSI in both bands as in (4.10), and then transmitting them to destination D. The destination then estimates the input messages from the received signals. The decoding probability of error at D for the code is defined as

$$P_e^{(n)} := \Pr\left[\psi\left(Y_{\mathsf{D}}^n, \{\bar{Y}_{m\mathsf{D}}^{n_1}, H_{m\mathsf{D}}^n, \bar{H}_{m\mathsf{D}}^{n_1}\}_{m\in\{1,2,\mathsf{R}\}}\right) \neq (M_1, M_2)\right],$$

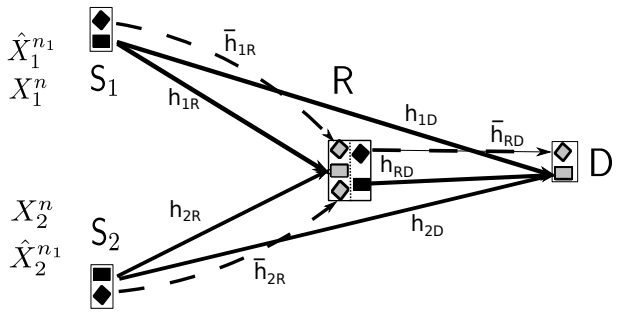


Figure 4.2: The R-MARC. Compared to the DR-MARC in Figure 4.1b, in the R-MARC the mm-wave direct-links are not present.

where the average is taken over uniform distribution of $(M_1, M_2) \in \mathcal{M}_1 \times \mathcal{M}_2$.

Definition 4.2 (Achievable rate for the DR-MARC)**.** *A rate tuple* (R_1, R_2) *is said to be achievable if there exists a sequence of* $(2^{nR_1}, 2^{nR_2}, n, \alpha)$ *codes such that* $n_1 = \lfloor n\alpha \rfloor$ *and* $P_e^{(n)} \to 0$*, as* $n \to \infty$*.*

The *capacity region* of the DR-MARC is defined as the closure of the set of all achievable rate tuples.

System Model for the R-MARC

The R-MARC, depicted in Figure 4.2, consists of an underlying c-MARC in the first band, and the mm-wave S_1- R and S_2- R relay-links and the R-D link in the second band. Hence, the system model of the R-MARC is defined by (4.1)-(4.2) in the first band, and (4.3)-(4.4) and (4.7) in the second band.

A $(2^{nR_1}, 2^{nR_2}, n, \alpha)$ code and achievable rate pair for the R-MARC are defined from those of the DR-MARC in Def. 4.1 and Def. 4.2 after removing the mm-wave S_1-D and S_2-D direct-links, i.e., setting $\bar{X}_{k\mathsf{D},l} = \bar{Y}_{k\mathsf{D},l} = \bar{H}_{k\mathsf{D},l} = \emptyset, k = 1, 2$, and modifying the encoding and decoding functions accordingly.

System Model for the K-User MARCs

In the K-user DR-MARC, K distinct users $\mathsf{S}_1,\ldots,\mathsf{S}_K$ communicate to the destination D over the integrated mm-wave/microwave dual-band with the help of a single relay R. In the first band, the channel outputs at D and R at the i-th channel use are respectively

given by

$$Y_{\mathsf{D},i} = H_{1\mathsf{D},i}X_{1,i} + H_{2\mathsf{D},i}X_{2,i} + \ldots + H_{K\mathsf{D},i}X_{K,i} + H_{\mathsf{RD},i}X_{\mathsf{R},i} + Z_{\mathsf{D},i}$$
$$Y_{\mathsf{R},i} = H_{1\mathsf{R},i}X_{1,i} + H_{2\mathsf{R},i}X_{2,i} + \ldots + H_{K\mathsf{R},i}X_{K,i} + Z_{\mathsf{R},i}, \qquad i = 1, \ldots, n,$$

where the channel fading coefficients $H_{mt,i} \in \mathbb{C}$, the input signals $X_{m,i} \in \mathbb{C}$ that have average power constraints P_m, $m \neq t, m \in \{1, 2, \ldots, K, \mathsf{R}\}, t \in \{\mathsf{R}, \mathsf{D}\}$, and the noise RVs $Z_{\mathsf{R},i} \sim \mathcal{CN}(0, 1)$ and $Z_{\mathsf{D},i} \sim \mathcal{CN}(0, 1)$, are all defined in the same manner as in (4.1)-(4.2).

Similarly, in the second band, the outputs of the S_k-R relay-links at the relay R are modeled as

$$\bar{Y}_{k\mathsf{R},\ell} = \bar{H}_{k\mathsf{R},\ell}\hat{X}_{k,\ell} + \bar{Z}_{k\mathsf{R},\ell}, \qquad k \in \{1, 2, \ldots, K\}, \qquad \ell = 1, \ldots, n_1,$$

while the outputs of the S_k-D direct-links and the R-D link at the destination D are modeled respectively as

$$\bar{Y}_{m\mathsf{D},\ell} = \bar{H}_{m\mathsf{D},\ell}\bar{X}_{m,\ell} + \bar{Z}_{m\mathsf{D},\ell}, \qquad m \in \{1, 2, \ldots, K\}, \qquad \ell = 1, \ldots, n_1,$$
$$\bar{Y}_{\mathsf{RD},\ell} = \bar{H}_{\mathsf{RD},\ell}\bar{X}_{\mathsf{R},\ell} + \bar{Z}_{\mathsf{RD},\ell}, \qquad \ell = 1, \ldots, n_1,$$

where the fading coefficients ($\bar{H}_{k\mathsf{R},\ell}$, $\bar{H}_{k\mathsf{D},\ell}$, $\bar{H}_{\mathsf{RD},\ell}$), the input signals $\hat{X}_{k,\ell} \in \mathbb{C}$ and $\bar{X}_{m,\ell} \in \mathbb{C}$ that have average power constraints $\hat{P}_m$ and $\bar{P}_k$, and the noise RVs are defined similarly to (4.3)-(4.7).

We also assume that the fading model and the assumption on the CSI of the K-User MARCs are the same as those for the two-user case detailed in Section 4.1. The other terminologies such as the code and the achievable region for the K-user DR-MARC are defined in a similar manner to those for the DR-MARC in Def. 4.1 and Def. 4.2, and hence not repeated.

Finally, the K-user R-MARC is defined from the K-user DR-MARC by removing the mm-wave S_k-D direct-links, $k = 1, \ldots, K$. Moreover, a code and achievable rates for the K-user R-MARC is defined in a similar fashion as for the R-MARC.

4.2 Capacity Results

4.2.1 Decomposition of the Capacity of the DR-MARC

We show that the capacity of the DR-MARC with BMF α can be decomposed into the capacity region of the underlying R-MARC and the capacity of the mm-wave direct-links.

The decomposition result is presented below.

Theorem 4.1. *The capacity of the DR-MARC with BMF α is given by the set of all rate tuples $(R_1, R_2) \in \mathbb{R}_+^2$ that satisfy*

$$\begin{aligned} R_1 &\leq r_1 + \alpha\mathbb{E}[\tilde{\mathsf{C}}(\bar{G}_{1\mathsf{D}}\bar{P}_1)], \\ R_2 &\leq r_2 + \alpha\mathbb{E}[\tilde{\mathsf{C}}(\bar{G}_{2\mathsf{D}}\bar{P}_2)], \end{aligned} \tag{4.11}$$

where $(r_1, r_2) \in \mathbb{R}_+^2$ is an achievable rate tuple in the underlying R-MARC, and the expectations are taken over channel gains $\bar{G}_{1\mathsf{D}}$ and $\bar{G}_{2\mathsf{D}}$, with $\tilde{\mathsf{C}}(x) = \log_2(1+x)$.

Proof. The proof is relegated to Appendix C.1. ■

This result shows that any achievable rate pair (R_1, R_2) in the DR-MARC can be obtained by achieving (r_1, r_2) in the underlying R-MARC, and then supplementing it with the capacity of the direct-links. Hence, operating the direct-links independently of the R-MARC remains optimal, which simplifies the encoding/decoding strategy. Since the capacity region of the DR-MARC can thus be determined from that of the underlying R-MARC, it is sufficient to focus on the R-MARC, which is considered next. While separating the operation of the mm-wave direct-links from the underlying R-MARC is optimal as noted in Theorem 4.1, in the R-MARC separating the operation of the mm-wave relay-links from the underlying c-MARC is suboptimal in general.

4.2.2 Achievable Region for the R-MARC

We first characterize an *achievable rate region* for the R-MARC.

Theorem 4.2. *An achievable region of the R-MARC with BMF α is given by the set of all rate tuples $(R_1, R_2) \in \mathbb{R}_+^2$ that satisfy*

$$R_1 < \mathbb{E}\left[\tilde{\mathsf{C}}\left(G_{1\mathsf{R}}P_1\right)\right] + \alpha\mathbb{E}\left[\tilde{\mathsf{C}}\left(\bar{G}_{1\mathsf{R}}\hat{P}_1\right)\right], \tag{4.12}$$

$$R_2 < \mathbb{E}\left[\tilde{\mathsf{C}}\left(G_{2\mathsf{R}}P_2\right)\right] + \alpha\mathbb{E}\left[\tilde{\mathsf{C}}\left(\bar{G}_{2\mathsf{R}}\hat{P}_2\right)\right], \tag{4.13}$$

$$R_1 + R_2 < \mathbb{E}\left[\tilde{\mathsf{C}}\left(G_{1\mathsf{R}}P_1 + G_{2\mathsf{R}}P_2\right)\right] + \alpha\mathbb{E}\left[\tilde{\mathsf{C}}\left(\bar{G}_{1\mathsf{R}}\hat{P}_1\right)\right] + \alpha\mathbb{E}\left[\tilde{\mathsf{C}}\left(\bar{G}_{2\mathsf{R}}\hat{P}_2\right)\right], \tag{4.14}$$

$$R_1 < \mathbb{E}\left[\tilde{\mathsf{C}}\left(G_{1\mathsf{D}}P_1 + G_{\mathsf{RD}}P_{\mathsf{R}}\right)\right] + \alpha\mathbb{E}\left[\tilde{\mathsf{C}}\left(\bar{G}_{\mathsf{RD}}\bar{P}_{\mathsf{R}}\right)\right], \tag{4.15}$$

$$R_2 < \mathbb{E}\left[\tilde{\mathsf{C}}\left(G_{2\mathsf{D}}P_2 + G_{\mathsf{RD}}P_{\mathsf{R}}\right)\right] + \alpha\mathbb{E}\left[\tilde{\mathsf{C}}\left(\bar{G}_{\mathsf{RD}}\bar{P}_{\mathsf{R}}\right)\right], \tag{4.16}$$

$$R_1 + R_2 < \mathbb{E}\left[\tilde{\mathsf{C}}\left(G_{1\mathsf{D}}P_1 + G_{2\mathsf{D}}P_2 + G_{\mathsf{RD}}P_{\mathsf{R}}\right)\right] + \alpha\mathbb{E}\left[\tilde{\mathsf{C}}\left(\bar{G}_{\mathsf{RD}}\bar{P}_{\mathsf{R}}\right)\right], \tag{4.17}$$

where the expectations are taken over the channel gains G_{mt} *and* $\bar{G}_{mt}$, $m \in \{1,2,\mathsf{R}\}, t \in \{\mathsf{R},\mathsf{D}\}, m \neq t$, *and* $\tilde{\mathsf{C}}(x) = \log_2(1+x)$.

Proof. The achievable region is obtained by performing a block Markov encoding scheme [122] as follows: (a) the source messages are partitioned into multiple blocks that are encoded and transmitted by the sources; (b) the relay deploys forward decoding to decode the source messages in the current block, and then re-encodes them for transmission in the next block; (c) the destination uses backward decoding to estimate the source messages in the reverse order of transmission. The details are relegated to Appendix C.2. ■

We emphasize that instead of sending independent messages through the microwave c-MARC and the mm-wave relay-links, the same message is *jointly* encoded into two codewords, one for the microwave c-MARC and the other for the mm-wave relay-link, which are then transmitted jointly over the two bands. Also bounds (4.15)-(4.17) can be interpreted as an achievable region for the multi-access channel (MAC) from the two sources to the destination aided by the relay, while bounds (4.12)-(4.14) can be regarded as the same for the MAC from the sources to the relay.

4.2.3 Capacity Region Outer Bounds for the R-MARC

We now characterize a set of upper bounds on the rates of the R-MARC.

Theorem 4.3 (Upper bounds on the rates of the R-MARC)**.** *The capacity region of the R-MARC with BMF* α *is contained within the set of rate tuples* $(R_1, R_2) \in \mathbb{R}_+^2$ *that satisfy*

$$R_1 < \mathbb{E}\left[\tilde{\mathsf{C}}\left(G_{1\mathsf{D}}P_1 + G_{\mathsf{RD}}P_{\mathsf{R}}\right)\right] + \alpha\mathbb{E}\left[\tilde{\mathsf{C}}\left(\bar{G}_{\mathsf{RD}}\bar{P}_{\mathsf{R}}\right)\right], \tag{4.18}$$

$$R_2 < \mathbb{E}\left[\tilde{\mathsf{C}}\left(G_{2\mathsf{D}}P_2 + G_{\mathsf{RD}}P_{\mathsf{R}}\right)\right] + \alpha\mathbb{E}\left[\tilde{\mathsf{C}}\left(\bar{G}_{\mathsf{RD}}\bar{P}_{\mathsf{R}}\right)\right], \tag{4.19}$$

$$R_1 + R_2 < \mathbb{E}\left[\tilde{\mathsf{C}}\left(G_{1\mathsf{D}}P_1 + G_{2\mathsf{D}}P_2 + G_{\mathsf{RD}}P_{\mathsf{R}}\right)\right] + \alpha\mathbb{E}\left[\tilde{\mathsf{C}}\left(\bar{G}_{\mathsf{RD}}\bar{P}_{\mathsf{R}}\right)\right], \tag{4.20}$$

with the expectations taken over gains G_{mt}, $\bar{G}_{mt}$, $m \neq t, m \in \{1,2,\mathsf{R}\}, t \in \{\mathsf{R},\mathsf{D}\}$.

Proof. The derivation of the outer-bounds follow from a cut-set bounding technique (see [138, Chapter 14.10]), and are relegated to Appendix C.3. ■

The key steps in the proof of is in steps (j) of (C.5): the cross-correlation coefficients between the source and relay codeword symbols are set to zero. Since instantaneous CSI are not available to the transmitters, and the phase of the fading coefficients are distributed as $\mathcal{U}[0, 2\pi)$, i.i.d., setting this cross-correlation coefficients to zero proves optimal.

4.2.4 Capacity of a class of R-MARCs

We now characterize a set of sufficient channel conditions, denoted the *jointly-near* conditions, under which the capacity of the R-MARC is established.

Theorem 4.4 (Capacity region of the jointly-near R-MARC)**.** *If the R-MARC with BMF α satisfies*

$$\mathbb{E}\left[\tilde{\mathsf{C}}\left(G_{1\mathsf{D}}P_1+G_{\mathsf{RD}}P_{\mathsf{R}}\right)\right]+\alpha\mathbb{E}\left[\tilde{\mathsf{C}}\left(\bar{G}_{\mathsf{RD}}\bar{P}_{\mathsf{R}}\right)\right]\leq\mathbb{E}\left[\tilde{\mathsf{C}}\left(G_{1\mathsf{R}}P_1\right)\right]+\alpha\mathbb{E}\left[\tilde{\mathsf{C}}\left(\bar{G}_{1\mathsf{R}}\hat{P}_1\right)\right] \quad (4.21)$$

$$\mathbb{E}\left[\tilde{\mathsf{C}}\left(G_{2\mathsf{D}}P_2+G_{\mathsf{RD}}P_{\mathsf{R}}\right)\right]+\alpha\mathbb{E}\left[\tilde{\mathsf{C}}\left(\bar{G}_{\mathsf{RD}}\bar{P}_{\mathsf{R}}\right)\right]\leq\mathbb{E}\left[\tilde{\mathsf{C}}\left(G_{2\mathsf{R}}P_2\right)\right]+\alpha\mathbb{E}\left[\tilde{\mathsf{C}}\left(\bar{G}_{2\mathsf{R}}\hat{P}_2\right)\right] \quad (4.22)$$

$$\mathbb{E}\left[\tilde{\mathsf{C}}\left(\sum_{k=1}^{2}G_{k\mathsf{D}}P_k+G_{\mathsf{RD}}P_{\mathsf{R}}\right)\right]+\alpha\mathbb{E}\left[\tilde{\mathsf{C}}\left(\bar{G}_{\mathsf{RD}}\bar{P}_{\mathsf{R}}\right)\right]\leq$$
$$\mathbb{E}\left[\tilde{\mathsf{C}}\left(G_{1\mathsf{R}}P_1+G_{2\mathsf{R}}P_2\right)\right]+\alpha\mathbb{E}\left[\tilde{\mathsf{C}}\left(\bar{G}_{1\mathsf{R}}\hat{P}_1\right)\right]+\alpha\mathbb{E}\left[\tilde{\mathsf{C}}\left(\bar{G}_{2\mathsf{R}}\hat{P}_2\right)\right], \quad (4.23)$$

its capacity region is given by the set of rate tuples $(R_1,R_2)\in\mathbb{R}_+^2$ that satisfy (4.18)-(4.20), where the expectations are over gains G_{mt}, $\bar{G}_{mt}$, $m\neq t, m\in\{1,2,\mathsf{R}\}, t\in\{\mathsf{R},\mathsf{D}\}$.

Proof. The proof follows from Theorem 4.1 and Theorem 4.2 in a straightforward manner. First, Theorem 4.2 provides upper bounds on rates R_1, R_2 and R_1+R_2 given in (4.18), (4.19), and (4.20), respectively. Next, under conditions (4.21)-(4.23), the rates achievable for the sources at the relay, given by (4.12)-(4.14), is larger than the rates achievable at the destination aided by the relay, given by (4.15)-(4.17). Thus, the achievable rates for the R-MARC are given by (4.15)-(4.17), which matches the upper bounds in Theorem 4.2, thus characterizing the capacity region. ■

The Jointly-Near Conditions: From the discussion in Section 2.2.3, one may recall that the capacity of an individual c-MARC operating solely in the microwave band (i.e., where mm-wave links do not exist) is known under the *near conditions* in [77, Theorem 9]: the relay is *near* the sources in the sense that the source-relay achievable rates are larger than the source-destination achievable rates. Similar conditions for the R-MARC, under which its capacity region is characterized in closed form, are given by (4.21)-(4.23). In relation to the near conditions for the individual c-MARC, conditions (4.21)-(4.23) for the R-MARC are denoted the *jointly-near* conditions, as the achievable scheme of Theorem 4.1 now performs *joint* encoding over the two bands of the R-MARC.

Intuitively, under (4.21)-(4.23) the relay achieves better rates than the destination, and thus can decode both messages without becoming a bottleneck to the achievable rates.

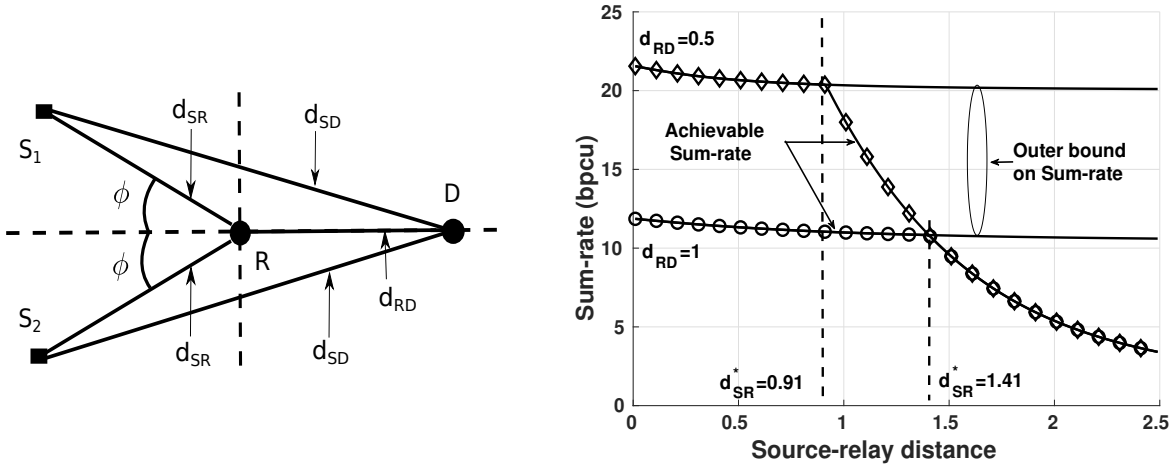
More specifically, under (4.21), the rate achievable from source S_1 to relay (on the r.h.s) is larger than the rate achievable from source S_1 to destination (on the l.h.s); (4.22) provides similar conditions for the rate of source S_2, while (4.22) provides similar conditions for the sum-rate.

Note that due to joint encoding over both bands, the source-relay-links over the microwave and mm-wave bands *together* characterize the jointly-near conditions in (4.21)-(4.23). If the underlying c-MARC in the R-MARC has strong enough microwave source-relay channel gains to satisfy the jointly-near conditions by themselves, then the mm-wave relay-links are not needed to characterize the closed form capacity result. In the complementary setting, however, the benefits of joint encoding becomes clear. Due to joint encoding over both bands, the rates achieved at the relay are augmented by the rate of the mm-wave relay-links, i.e., $\alpha\mathbb{E}\left[\tilde{\mathsf{C}}\left(\bar{G}_{1\mathsf{R}}\hat{P}_1\right)\right]$ and $\alpha\mathbb{E}\left[\tilde{\mathsf{C}}\left(\bar{G}_{2\mathsf{R}}\hat{P}_2\right)\right]$. Therefore, even when the microwave source-relay channel gains are not strong enough to satisfy the jointly-near conditions by themselves, if the mm-wave relay-links are sufficiently strong, then the jointly-near conditions may hold jointly over both the bands, thereby providing a closed form capacity result for the R-MARC.

Phase and Rayleigh Fading Cases: Note that the decomposition result for the DR-MARC in Theorem 4.1, the achievable region and outer bound results for the R-MARC in Theorem 4.2 and Theorem 4.3, as well as the capacity result in Theorem 4.4 all specialize directly to phase and Rayleigh fading cases. Specifically, for phase fading the microwave channel gains G_{mt} and mm-wave channel gains $\bar{G}_{mt}$ are taken to be constants that depend on the geometry of the network and path-loss coefficients, and thus the expectations in the theorems are not needed. Moreover, for Rayleigh fading the expectations are over $G_{mt,i} \sim \mathsf{exp}(\mu_{mt})$, and $\bar{G}_{mt,\ell} \sim \mathsf{exp}(\bar{\mu}_{mt})$, i.i.d., where the means μ_{mt} and $\bar{\mu}_{mt}$ are constants that depend on the network-geometry and path-loss coefficients.

4.2.5 Numerical Examples

We illustrate the impact of mm-wave spectrum on the capacity of the R-MARC by considering a two-dimensional topology as in Figure 4.3a. Here R and D are located on the x-axis at $(0,0)$ and $(0,\mathsf{d}_{\mathsf{RD}})$, and S_1 and S_2 are located symmetrically at $(-\mathsf{d}_{\mathsf{SR}}\cos\phi, \pm\mathsf{d}_{\mathsf{SR}}\sin\phi)$, with ϕ being the angle between a source and R and $\mathsf{d}_{\mathsf{SD}} = (\mathsf{d}_{\mathsf{SR}}^2 + \mathsf{d}_{\mathsf{RD}}^2 + 2\mathsf{d}_{\mathsf{SR}}\mathsf{d}_{\mathsf{RD}}\cos\phi)^{1/2}$ the resulting source-destination distance.



(a) (b)

Figure 4.3: (a) A 2-D geometry of the DR-MARC. The relay and the destination are located on the x-axis, and the sources are located symmetrically on either side of the x-axis at a distance d_{SD} from the destination, and at a distance d_{SR} from the relay. The relay-destination distanced is d_{RD}. (b) The achievable sum-rate matches the sum-rate outer bound if $\mathsf{d}_{\mathsf{SR}} \leq \mathsf{d}^*_{\mathsf{SR}}$ for both cases of d_{RD}.

For ease of exposition, (a) we take the underlying microwave band in the R-MARC to be under phase fading with constant fading magnitudes $G_{mt,i} = 1/\mathsf{d}_{mt}^{\beta_1}$ that depends on the inter-node distance d_{mt} and path-loss coefficient β_1 in a manner similar to [77], and (b) we similarly take constant mm-wave fading magnitudes $\bar{G}_{mt,\ell} := 1/\mathsf{d}_{mt}^{\beta_2}$ where β_2 is the path-loss coefficient for the mm-wave band, $m \in \{1, 2, \mathsf{R}\}, t \in \{\mathsf{D}, \mathsf{R}\}, m \neq t$. Therefore, expectations in conditions (4.21)-(4.23) and Theorem 4.2 are not needed, and observations can be directly interpreted in terms of inter-node distances and pathloss coefficients. Moreover, all power constraints in the R-MARC are set to 10 and pathloss coefficients are taken to be $\beta_1 = 2, \beta_2 = 4$.

In the first example, we illustrate the interdependencies between the jointly-near conditions in (4.21)-(4.23), the channel parameters $\mathsf{d}_{\mathsf{SR}}, \mathsf{d}_{\mathsf{SD}}, \mathsf{d}_{\mathsf{RD}}, \phi$ and BMF α in the setting of Figure 4.3a, and the achievable rates in Theorem 4.2. For simplicity, we focus on the achievable sum-rate (denoted the ASR), given by the minimum of r.h.s. of (4.14) and (4.17), and the sum-rate outer bound (denoted the OB), given by the r.h.s. of (4.17). As such note that under (4.23), the OB matches the ASR. For ease of exposition, we fix $\mathsf{d}_{\mathsf{RD}}, \phi$ and BMF α. Hence, condition (4.23) is now equivalent to $\mathsf{d}_{\mathsf{SR}} \leq \mathsf{d}^*_{\mathsf{SR}}(\mathsf{d}_{\mathsf{RD}}, \phi, \alpha)$ for

some threshold source-destination distance $\mathsf{d}^*_{\mathsf{SR}}(\mathsf{d}_{\mathsf{RD}}, \phi, \alpha)$. It is thus expected that for all $\mathsf{d}_{\mathsf{SR}} \leq \mathsf{d}^*_{\mathsf{SR}}(\mathsf{d}_{\mathsf{RD}}, \phi, \alpha)$, the ASR and the OB will match, while for $\mathsf{d}_{\mathsf{SR}} > \mathsf{d}^*_{\mathsf{SR}}(\mathsf{d}_{\mathsf{RD}}, \phi, \alpha)$, the ASR will be strictly less than the OB.

In Figure 4.3b, we verify this for a fixed $\phi = \pi/4$ and $\alpha = 2$ and the two cases $\mathsf{d}_{\mathsf{RD}} \in \{1, 0.5\}$ by plotting the ASR and the OB as functions of $\mathsf{d}_{\mathsf{SR}} \in (0, 2.5]$. We observe that the ASR matches the OB if $\mathsf{d}_{\mathsf{SR}} \leq \mathsf{d}^*_{\mathsf{SR}}$ with $\mathsf{d}^*_{\mathsf{SR}} \approx 1.41$ for $\mathsf{d}_{\mathsf{RD}} = 1$, and $\mathsf{d}^*_{\mathsf{SR}} \approx 0.91$ for $\mathsf{d}_{\mathsf{RD}} = 0.5$, otherwise the ASR is strictly smaller, as expected. Moreover, as d_{RD} reduces from 1 to 0.5, for condition (4.23) to hold, $\mathsf{d}^*_{\mathsf{SR}}$ also *reduces* from $\mathsf{d}^*_{\mathsf{SR}} \approx 1.41$ to ≈ 0.91.

In the next example, we illustrate the impact of the higher mm-wave bandwidth on the performance of the R-MARC. Specifically, in Figure 4.4, we depict the source locations relative to the relay and the destination in the setting of Figure 4.3a, for which the scheme of Theorem 4.2 achieves the capacity region of the R-MARC, i.e., source locations for which all of conditions (4.21)-(4.23) are satisfied. As such, we fix $\mathsf{d}_{\mathsf{RD}} = 1$, vary $\phi \in (0, \pi)$ and $\mathsf{d}_{\mathsf{SR}} \in (0, 2)$ to vary the location of the sources, and then plot the resulting regions. To capture the impact of BMF α, we overlay the region for the case without mm-wave links (i.e., $\alpha = 0$) on those with mm-wave links for $\alpha \in \{2, 4, 10\}$, as well as $\alpha \to \infty$.

First, for the case without mm-wave links (i.e., $\alpha = 0$), conditions (4.21)-(4.23) hold only when sources are within the innermost black region in Figure 4.4. It should be noted that for each ϕ, the resulting threshold distance $\mathsf{d}^*_{\mathsf{SR}}(\phi)$ is at the boundary of this region, and as ϕ increases from $\phi = 0$ to $\phi = \pi$, $\mathsf{d}^*_{\mathsf{SR}}(\phi)$ decreases monotonically from ≈ 1.2 to ≈ 0.47. We thus observe that conditions (4.21)-(4.23) hold for much larger threshold distance $\mathsf{d}^*_{\mathsf{SR}}$ when sources are located far away from destination (i.e., $\phi \approx 0$), and $\mathsf{d}^*_{\mathsf{SR}}$ reduces considerably when sources are closer to the destination (i.e., $\phi \approx \pi$).

We note that the above trends continue to hold when mm-wave links are used (i.e., $\alpha > 0$), however, the resulting region (union of the inner black and outer gray regions) now extends much closer to the destination. For example, for the region with $\alpha = 2$, $\mathsf{d}^*_{\mathsf{SR}}$ reduces to only ≈ 0.96 near the destination, compared to ≈ 0.47 with $\alpha = 0$. Moreover, while the resulting region grows with α, the growth saturates for higher values of α, with $\alpha = 10$ producing almost the same region as that for $\alpha \to \infty$.

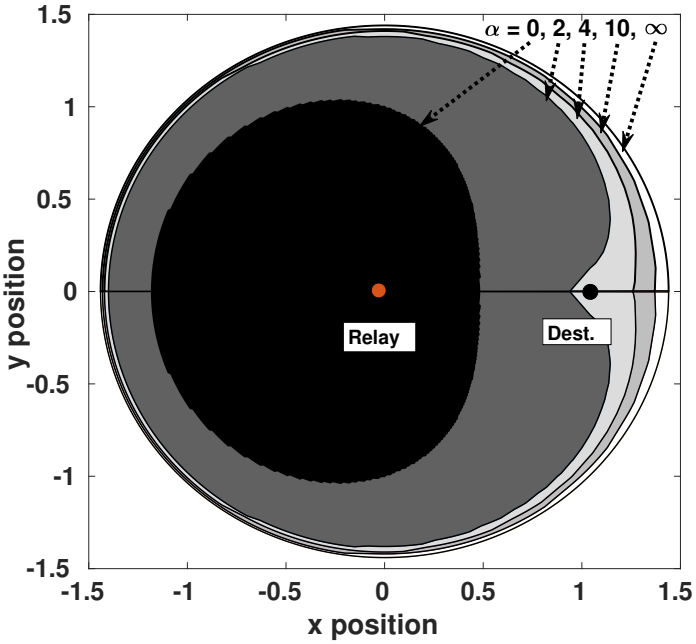


Figure 4.4: The source locations for which the scheme of Theorem 4.2 achieves the capacity of the R-MARC (i.e., the locations at coordinates $(x, \pm y)$ in the shaded regions).

4.3 Extension to the K-user Case

We now extend the achievability and capacity results for the two-user MARCs to the K-user MARCs where $K > 2$ sources communicate to a destination with the help of a relay. The system model and other terminologies of the K-user DR-MARC and K-user R-MARC are defined in Section 4.1. We also assume that the K-user MARCs follows the same fading model and the same assumptions on the channel state information (CSI) as in Section 4.1.

Extension of Theorem 4.1

Theorem 4.5 (Decomposition result for the K-user DR-MARC)**.** *The capacity of the K-user DR-MARC with BMF α is given by the set of all rate tuples $(R_1, \ldots, R_K) \in \mathbb{R}_+^K$ that satisfy*

$$R_k \leq r_k + \alpha \mathbb{E}[\tilde{\mathsf{C}}(\bar{G}_{k\mathsf{D}} \bar{P}_k)], \quad k \in \{1, \ldots, K\}, \tag{4.24}$$

where $(r_1, \ldots, r_K) \in \mathbb{R}_+^K$ is an achievable rate tuple in the underlying K-user R-MARC, and the expectations are taken over channel gains $\bar{G}_{1\mathsf{D}}, \ldots, \bar{G}_{K\mathsf{D}}$, with $\tilde{\mathsf{C}}(x) = \log(1 + x)$.

Proof Sketch. The proof of this result follows directly from the proof of Theorem 4.1 presented in Appendix C.1. More specifically, the technique used to upper bound rate R_1 can

be applied to each of rates $R_2, \ldots, R_K$. ■

Extension of Theorem 4.2

Theorem 4.6 (Achievable region for the K-user R-MARC)**.** *An achievable region of the K-user R-MARC with BMF α is given by the set of all rate tuples $(R_1, \ldots, R_K) \in \mathbb{R}_+^K$ that satisfy*

$$\sum_{k\in S} R_k < \mathbb{E}\left[\sum_{k\in S} \tilde{\mathsf{C}}\left(G_{k\mathsf{R}} P_k\right)\right] + \alpha \sum_{k\in S} \mathbb{E}\left[\tilde{\mathsf{C}}\left(\bar{G}_{k\mathsf{R}} \hat{P}_k\right)\right], \tag{4.25}$$

$$\sum_{k\in S} R_k < \mathbb{E}\left[\sum_{k\in S} \tilde{\mathsf{C}}\left(G_{k\mathsf{D}} P_k + G_{\mathsf{RD}} P_{\mathsf{R}}\right)\right] + \alpha \mathbb{E}\left[\tilde{\mathsf{C}}\left(\bar{G}_{\mathsf{RD}} \bar{P}_{\mathsf{R}}\right)\right], \tag{4.26}$$

for all $S \subseteq \{1, 2, \ldots, K\}$, $S \neq \emptyset$, where expectations are over channel gains G_{mt} and $\bar{G}_{mt}$, $m \neq t, m \in \{1, \ldots, K, \mathsf{R}\}, t \in \{\mathsf{R}, \mathsf{D}\}$.

Proof Sketch. The achievable region is obtained by performing block Markov encoding, and deploying backward decoding for the destination and forward decoding for the relay. The procedure follows that of the proof of Theorem 4.2 and naturally generalize to K users. Note that forward decoding at the relay results in rates (4.25), while backward decoding at the destination results in (4.26). ■

Extension of Theorem 4.3

Theorem 4.7 (Outer bounds to the capacity region of the K-user R-MARC)**.** *The capacity region of the R-MARC with BMF α is contained within the set of all rate tuples $(R_1, \ldots, R_K) \in \mathbb{R}_+^K$ that satisfy*

$$\sum_{k\in S} R_k < \mathbb{E}\left[\sum_{k\in S} \tilde{\mathsf{C}}\left(G_{k\mathsf{D}} P_k + G_{\mathsf{RD}} P_{\mathsf{R}}\right)\right] + \alpha \mathbb{E}\left[\tilde{\mathsf{C}}\left(\bar{G}_{\mathsf{RD}} \bar{P}_{\mathsf{R}}\right)\right], \tag{4.27}$$

for all $S \subseteq \{1, 2, \ldots, K\}$, $S \neq \emptyset$, where expectations are over channel gains $G_{m\mathsf{D}}, m \in \{1, \ldots, K, \mathsf{R}\}$ and $\bar{G}_{\mathsf{RD}}$.

Proof Sketch. The proof of this theorem follows from that of Theorem 4.3 in a straightforward manner by taking all subsets $\mathsf{U} \subseteq \{1, 2, \ldots, K\}, \mathsf{U} \neq \emptyset$, and hence is not repeated. Specifically, the crucial step of (j) of (C.5), where setting the cross-correlation between source and relay codeword symbols to zero is optimal, extends to the K-user case. ■

Extension of Theorem 4.4

Theorem 4.8 (Capacity region of a class of K-user R-MARCs)**.** *If the channel parameters of the K-user R-MARC with BMF α satisfies*

$$\mathbb{E}\left[\sum_{k\in S}\tilde{\mathsf{C}}\left(G_{k\mathsf{D}}P_k+G_{\mathsf{RD}}P_{\mathsf{R}}\right)\right]+\alpha\mathbb{E}\left[\tilde{\mathsf{C}}\left(\bar{G}_{\mathsf{RD}}\bar{P}_{\mathsf{R}}\right)\right]$$
$$\leq\mathbb{E}\left[\sum_{k\in S}\tilde{\mathsf{C}}\left(G_{k\mathsf{R}}P_k\right)\right]+\alpha\sum_{k\in S}\mathbb{E}\left[\tilde{\mathsf{C}}\left(\bar{G}_{k\mathsf{R}}\hat{P}_k\right)\right], \tag{4.28}$$

for all $S\subseteq\{1,2,\ldots,K\}$, $S\neq\emptyset$, where the expectations are over channel gains G_{mt}, $\bar{G}_{mt}$, $m\neq t, m\in\{1,\ldots,K,\mathsf{R}\}, t\in\{\mathsf{R},\mathsf{D}\}$, then its capacity region is given by the set of all rate tuples $(R_1,\ldots,R_K)\in\mathbb{R}_+^K$ that satisfy (4.27).

Proof Sketch. It is evident that under conditions (4.28), the achievable rate region in Theorem 4.6 simplifies to that given by (4.26), which then matches the outer bound region in Theorem 4.7, and thus the capacity region is characterized. ■

4.4 The Resource Allocation in the DR-MARC

The performance of the DR-MARC can be significantly affected by the mm-wave links due to their point-to-point nature and relatively larger bandwidths. To understand the impact of mm-wave spectrum on the performance of the DR-MARC, we study how the sum-rate achievable on the DR-MARC (with the scheme of Theorem 4.2) is maximized by optimally allocating power to the mm-wave direct-links and relay-links. We observe that the resulting scheme allocates power to the mm-wave links in different *modes* depending on whether certain channel conditions hold. This characterization reveals insights into the nature of the optimal scheme, and can serve as an effective resource allocation strategy for such dual-band networks in practice.

From Section 4.2 recall that the fading magnitudes of the mm-wave links are taken to be non-negative i.i.d., random variables. In this section for ease of exposition, we assume that these fading magnitudes are *constants* that depend on the network-geometry and path-loss parameter. This modified fading is an appropriate model for mm-wave links such as those in [30], as it can be viewed as a special case of the general fading model of [117] when the diffuse component associated with the non-LoS propagation is not present. Furthermore, this simplification reveals useful insights into the optimal power allocation, while keeping

the analysis tractable. In contrast, the microwave band is assumed to be under the same general fading as in Section 4.2.

We limit this study to the two-user DR-MARC only, where the problem of optimal power allocation is posed on the four mm-wave links (two direct-links and two relay-links) as such a study with an arbitrary number of users (say K) will involve $2K$ number of mm-wave links and will soon become intractable.

For notational convenience, we have the following:

- the mm-wave S_1-D and S_2-D direct-links are referred to as $\mathtt{DL}_1$ and $\mathtt{DL}_2$ respectively, while the mm-wave S_1-R and S_2-R relay-links are referred to as $\mathtt{RL}_1$ and $\mathtt{RL}_2$.
- as discussed above, under the simplified fading model for the mm-wave links, the gain of the direct-link $\mathtt{DL}_k$, $\bar{G}_{k\mathsf{D}}$, and the gain of the relay-link $\mathtt{RL}_k$, $\bar{G}_{k\mathsf{R}}$, are network-geometry dependent constants. For notational simplicity, the gain of $\mathtt{DL}_k$ is denoted by $d_k := \bar{G}_{k\mathsf{D}}$, while the gain of $\mathtt{RL}_k$ is denoted by $r_k := \bar{G}_{k\mathsf{R}}, \;\; k \in \{1, 2\}$.

We formulate the problem for the DR-MARCs under the following assumptions:

- [A1] the transmit power in $\mathtt{DL}_k$ (i.e., p_k) and $\mathtt{RL}_k$ (i.e., q_k) from source S_k are constrained by the total power budget P, i.e., $p_k + q_k = P, \;\; k = 1, 2$.
- [A2] The BMF $\alpha > 0$ is fixed a priori and known.

Note that assuming the same power budget for the two sources as in [A1] does not incur any loss in generality, as any difference between two different power budgets can be absorbed into the link gains without altering the treatment of the problem. In [A2], BMF $\alpha > 0$ is assumed to be fixed a priori and known.

4.4.1 Problem Formulation and Solution

For a fixed power allocation (p_1, q_1, p_2, q_2), R is an achievable sum-rate of the DR-MARC iff

$$R \leq \min\{\Sigma_{\mathsf{R}}, \Sigma_{\mathsf{D}}\}, \tag{4.29}$$

where Σ_{R} and Σ_{D} denote the sum-rates achievable at the relay and destination, given by

$$\Sigma_{\mathsf{R}} := \sigma_{\mathsf{R}} + \alpha \sum_{k=1}^{2} \log\left(1 + r_k q_k\right) + \log\left(1 + d_k p_k\right), \tag{4.30}$$

$$\Sigma_{\mathsf{D}} := \sigma_{\mathsf{D}} + \alpha \sum_{k=1}^{2} \log\left(1 + d_k p_k\right), \tag{4.31}$$

with

$$\begin{aligned} \sigma_{\mathsf{D}} &:= \mathbb{E}[\tilde{\mathsf{C}}(G_{1\mathsf{D}} P_1 + G_{2\mathsf{D}} P_2 + G_{\mathsf{RD}} P_{\mathsf{R}})] + \alpha \tilde{\mathsf{C}}(\bar{G}_{\mathsf{RD}} \bar{P}_{\mathsf{R}}), \quad \text{and} \\ \sigma_{\mathsf{R}} &:= \mathbb{E}[\tilde{\mathsf{C}}(G_{1\mathsf{R}} P_1 + G_{2\mathsf{R}} P_2)] \end{aligned} \tag{4.32}$$

that are defined in Theorem 4.2.

Note that rates Σ_{R} and Σ_{D} are obtained as follows. First, from the decomposition result in Theorem 4.1 it follows that for given direct-link powers p_1 and p_2, the sum-rate of the DR-MARC is given by the sum of the sum-rate of the underlying R-MARC and the rates of the two direct-links, i.e., $\alpha \log(1 + d_1 p_1)$ and $\alpha \log(1 + d_2 p_2)$. Now, for given relay-link powers q_1 and q_2, the sum-rate of the R-MARC is given by the minimum of r.h.s. of (4.14) and (4.17). It then follows that Σ_{R} is given by the sum of the r.h.s. of (4.14) and $\alpha \sum_{k=1}^{2} \log(1 + d_k p_k)$ as expressed in (4.30), while Σ_{D} is given by the sum of the r.h.s. of (4.17) and $\alpha \sum_{k=1}^{2} \log(1 + d_k p_k)$ as given in (4.31).

The problem of maximizing R over the transmit powers (p_1, q_1, p_2, q_2) is then

$$\begin{aligned} [\mathcal{P}2] \quad & \textit{maximize} && R && (4.33) \\ & \textit{subject to} && R \leq \Sigma_{\mathsf{R}}, && (4.34) \\ & && R \leq \Sigma_{\mathsf{D}}, && (4.35) \\ & && p_1 + q_1 = P, && (4.36) \\ & && p_2 + q_2 = P, && (4.37) \\ & && (p_1, q_1, p_2, q_2, R) \succeq \mathbf{0}. && (4.38) \end{aligned}$$

Solution of Problem $[\mathcal{P}2]$

The details of the solution of problem $[\mathcal{P}2]$ has been relegated to Appendix C.4. Specifically, note that $[\mathcal{P}2]$ is a convex optimization problem: its objective function R is linear, the equality constraints (4.36)-(4.37) are affine, and the inequality constraints (4.34)-(4.35) are convex and differentiable. Moreover, $[\mathcal{P}2]$ satisfies the Slater constraint qualification condition [132, Chapter 5.2.3]. Hence, $[\mathcal{P}2]$ can be solved using the Karush–Kuhn–Tucker (KKT) conditions [132, Chapter 5.5.3], as detailed in Appendix C.4.

Constraint Qualification for Problem $[\mathcal{P}2]$

In Appendix C.5, problem $[\mathcal{P}2]$ is shown to satisfy the Mangasarian-Fromovitz constraint qualification.

4.4.2 Link Gain Regimes and Optimal Power Allocation

To gain insights, we derive the optimal power allocation in closed form, and describe it in terms of *link-gain regimes* (LGR): LGRs are partitions of the set of all tuples of link gains and power budget P, which are found while solving the KKT conditions for $[\mathcal{P}2]$. Specifically, we derive the KKT conditions and solve for the optimal primal variables (i.e., link transmission powers) and the optimal Lagrange multipliers (OLM). To simplify the procedure, we consider the set of tuples of OLMs associated with *inequality* constraints in (4.34), (4.35) and (4.38), and partition this set into a few subsets based on whether the OLMs in the set are *positive* or *zero*, i.e., whether the associated primal constraints are *tight* or *not* (detailed in Appendix C.4).

For each resulting partition of the set of OLM tuples, we first derive the expression for the optimal powers in closed form. However, the conditions that define these partitions are still characterized in terms of the OLMs. Therefore, to express the optimal power allocation explicitly in terms of link gains (r_1, r_2, d_1, d_2) and power budget P, we express the conditions that partition the set of the OLM tuples in terms of link gains, P, and a parameter γ, defined as

$$\gamma := 2^{(\sigma_{\mathsf{D}} - \sigma_{\mathsf{R}})/\alpha} \tag{4.39}$$

which models the effect of microwave band parameters, with σ_{D} and σ_{R} defined in (4.32). As a result, the set of all $(r_1, r_2, d_1, d_2, \gamma, P)$-tuples are partitioned into a few subsets, each called an LGR and correspond to one and only one subset of OLM tuples. The conditions for each LGR is then simplified and expressed as upper and lower bounds (*threshold powers*) on power budget P where the threshold powers depend on $(r_1, r_2, d_1, d_2, \gamma)$. This results in an equivalent partitioning of the range of the power budget, $P \geq 0$, into a few intervals, each describing an LGR.

For the optimum sum-rate problem, we specifically consider the two cases $\sigma_{\mathsf{D}} \leq \sigma_{\mathsf{R}}$ and $\sigma_{\mathsf{D}} > \sigma_{\mathsf{R}}$, as the optimum power allocation for the two cases differ significantly. Moreover, when interpreting the optimum power allocation scheme, we need to often compare Σ_{D} and Σ_{R}. Substituting the expressions of Σ_{D} and Σ_{R} in (4.30) and (4.31), this comparison

is equivalent to that between σ_{D} and $\sigma_{\mathsf{R}}+\alpha\big(\log(1+r_1q_1)+\log(1+r_2q_2)\big)$, i.e., comparison between $(1+r_1q_1)(1+r_2q_2)$ and $2^{(\sigma_{\mathsf{D}}-\sigma_{\mathsf{R}})/\alpha}$. For convenience, we thus define $\gamma=2^{(\sigma_{\mathsf{D}}-\sigma_{\mathsf{R}})/\alpha}$ in (4.39). Moreover, cases $\sigma_{\mathsf{D}}\leq\sigma_{\mathsf{R}}$ and $\sigma_{\mathsf{D}}>\sigma_{\mathsf{R}}$ are now equivalent to cases $\gamma\leq 1$ and $\gamma>1$.

Case $\gamma\leq 1$:

When $\gamma\leq 1$, the set of all $(r_1,r_2,d_1,d_2,\gamma,P)$-tuples turn out to belong to a *single* LGR where the allocation $(p_1,q_1,p_2,q_2)=(P,0,P,0)$ is optimal for all $P\geq 0$.

First, note that for any feasible allocation (p_1,q_1,p_2,q_2), rates Σ_{D} and Σ_{R} in (4.30)-(4.31) are expressed as $\Sigma_{\mathsf{D}}=\sigma_{\mathsf{D}}+\alpha\sum_{k=1}^{2}\log\left(1+d_kp_k\right)$ and $\Sigma_{\mathsf{R}}=\sigma_{\mathsf{R}}+\alpha\sum_{k=1}^{2}\log\left(1+d_kp_k\right)+\alpha\sum_{k=1}^{2}\log\left(1+r_kq_k\right)$. Since in this case $\gamma\leq 1$, we have $\sigma_{\mathsf{D}}\leq\sigma_{\mathsf{R}}$ from (4.39). In addition, since for any feasible allocation $q_k\geq 0$, we have $\Sigma_{\mathsf{D}}\leq\Sigma_{\mathsf{R}}$. Therefore, for any feasible allocation (p_1,q_1,p_2,q_2), the sum-rate is $R=\Sigma_{\mathsf{D}}$. However, Σ_{D} is a function of p_1, p_2 only, and thus for a given power budget P, allocating $p_1=p_2=P$ and $q_1=q_2=0$ maximizes Σ_{D} (i.e., the achievable sum-rate $R=\Sigma_{\mathsf{D}}$) for all values of channel parameters and P.

Case $\gamma>1$:

In contrast to the previous case, for the case with $\gamma>1$, the optimal power allocation (the MARC-OPA) has much richer structure. As discussed in detail shortly, the MARC-OPA follows two distinct properties, the *Waterfilling-like* property and the *saturation* property. In this case, the set of all channel parameter-tuples $\boldsymbol{c}:=(r_1,r_2,d_1,d_2,\gamma,P)$ are partitioned into 14 LGRs: 9 LGRs are associated with the Waterfilling-like property, and the remaining 5 LGRs are associated with the saturation property.

In Table 4.1, the definition of LGRs associated with the Waterfilling-like property (denoted by $\mathcal{A}_{(.,.)}$) and the optimal link powers in these LGRs are presented, whereas the same for the LGRs associated with the saturation property (denoted by $\mathcal{S}_{(.,.)}$) are presented in Table 4.2. Note that for the LGRs in Table 4.2, the relay-link gains are denoted by $\boldsymbol{r}:=(r_1,r_2)$, and the set of all $\boldsymbol{r}$ are partitioned into 4 subsets

$$\begin{aligned}\mathcal{R}_{S1}&:=\{\boldsymbol{r}:r_1\geq\gamma r_2\},&\qquad\mathcal{R}_{1}&:=\{\boldsymbol{r}:\gamma r_2>r_1\geq r_2\},\\ \mathcal{R}_{S2}&:=\{\boldsymbol{r}:r_2\geq\gamma r_1\},&\qquad\mathcal{R}_{2}&:=\{\boldsymbol{r}:\gamma r_1>r_2>r_1\}.\end{aligned}\tag{4.40}$$

Finally, the threshold powers for the LGRs in Table 4.1 and 4.2 are defined in Table 4.3, with $\varrho[\mathsf{f}(x)]$ denoting the non-negative root of polynomial $\mathsf{f}(x)$.

Table 4.1: For the case with $\gamma > 1$, definition of LGRs and optimal link powers associated with the Waterfilling-like property of the MARC-OPA is given in terms of channel parameters $\boldsymbol{c} = (r_1, r_2, d_1, d_2, \gamma, P)$. Table 4.3 provides the threshold powers in terms of $(r_1, r_2, d_1, d_2, \gamma)$.

Definition of LGR	**Optimal Power Allocation**	
$\mathcal{A}_{\mathsf{d},\mathsf{d}} := \{\boldsymbol{c} : 0 \leq$	$p_1 = P,$	$q_1 = 0,$
$P \leq \min(\mathsf{P}_{\mathsf{d},\mathsf{d}}, \hat{\mathsf{P}}_{\mathsf{d},\mathsf{d}})\}$	$p_2 = P,$	$q_2 = 0$
$\mathcal{A}_{\mathsf{d},\mathsf{r}} := \{\boldsymbol{c} : 0 \leq$	$p_1 = P,$	$q_1 = 0,$
$P \leq \min(\hat{\mathsf{P}}'_{\mathsf{d},\mathsf{d}}, \mathsf{P}_{\mathsf{d},\mathsf{d}}, \mathsf{P}_{\mathsf{d},\mathsf{r}})\}$	$p_2 = 0,$	$q_2 = P$
$\mathcal{A}_{\mathsf{r},\mathsf{d}} := \{\boldsymbol{c} : 0 \leq$	$p_1 = 0,$	$q_1 = P,$
$P \leq \min(\mathsf{P}'_{\mathsf{d},\mathsf{d}}, \hat{\mathsf{P}}_{\mathsf{d},\mathsf{d}}, \mathsf{P}_{\mathsf{r},\mathsf{d}})\}$	$p_2 = P,$	$q_2 = 0$
$\mathcal{A}_{\mathsf{r},\mathsf{r}} := \{\boldsymbol{c} : 0 \leq$	$p_1 = 0,$	$q_1 = P,$
$P \leq \min(\mathsf{P}'_{\mathsf{d},\mathsf{d}}, \hat{\mathsf{P}}'_{\mathsf{d},\mathsf{d}}, \mathsf{P}_{\mathsf{r},\mathsf{r}})\}$	$p_2 = 0,$	$q_2 = P$
$\mathcal{A}_{\mathsf{rd},\mathsf{d}} := \{\boldsymbol{c} : \max(\mathsf{P}_{\mathsf{d},\mathsf{d}}, \mathsf{P}'_{\mathsf{d},\mathsf{d}}) <$	$p_1 = \frac{1}{2}\left(P + \frac{1}{r_1} - \frac{1}{d_1}\right),$	$q_1 = \frac{1}{2}\left(P - \frac{1}{r_1} + \frac{1}{d_1}\right),$
$P \leq \min(\hat{\mathsf{P}}_{\mathsf{d},\mathsf{d}}, \mathsf{P}_{\mathsf{rd},\mathsf{d}})\}$	$p_2 = P,$	$q_2 = 0$
$\mathcal{A}_{\mathsf{d},\mathsf{rd}} := \{\boldsymbol{c} : \max(\hat{\mathsf{P}}_{\mathsf{d},\mathsf{d}}, \hat{\mathsf{P}}'_{\mathsf{d},\mathsf{d}}) <$	$p_1 = P,$	$q_1 = 0,$
$P \leq \min(\mathsf{P}_{\mathsf{d},\mathsf{d}}, \mathsf{P}_{\mathsf{d},\mathsf{rd}})\}$	$p_2 = \frac{1}{2}\left(P + \frac{1}{r_2} - \frac{1}{d_2}\right),$	$q_2 = \frac{1}{2}\left(P - \frac{1}{r_2} + \frac{1}{d_2}\right)$
$\mathcal{A}_{\mathsf{r},\mathsf{rd}} := \{\boldsymbol{c} : \max(\hat{\mathsf{P}}_{\mathsf{d},\mathsf{d}}, \hat{\mathsf{P}}'_{\mathsf{d},\mathsf{d}}) <$	$p_1 = 0,$	$q_1 = P,$
$P \leq \min(\mathsf{P}'_{\mathsf{d},\mathsf{d}}, \mathsf{P}_{\mathsf{r},\mathsf{rd}})\}$	$p_2 = \frac{1}{2}\left(P + \frac{1}{r_2} - \frac{1}{d_2}\right),$	$q_2 = \frac{1}{2}\left(P - \frac{1}{r_2} + \frac{1}{d_2}\right)$
$\mathcal{A}_{\mathsf{rd},\mathsf{r}} := \{\boldsymbol{c} : \max(\mathsf{P}_{\mathsf{d},\mathsf{d}}, \mathsf{P}'_{\mathsf{d},\mathsf{d}}) <$	$p_1 = \frac{1}{2}\left(P + \frac{1}{r_1} - \frac{1}{d_1}\right),$	$q_1 = \frac{1}{2}\left(P - \frac{1}{r_1} + \frac{1}{d_1}\right),$
$P \leq \min(\hat{\mathsf{P}}'_{\mathsf{d},\mathsf{d}}, \mathsf{P}_{\mathsf{rd},\mathsf{r}})\}$	$p_2 = 0,$	$q_2 = P$
$\mathcal{A}_{\mathsf{rd},\mathsf{rd}} := \{\boldsymbol{c} : \max(\mathsf{P}_{\mathsf{d},\mathsf{d}}, \hat{\mathsf{P}}_{\mathsf{d},\mathsf{d}},$	$p_1 = \frac{1}{2}\left(P + \frac{1}{r_1} - \frac{1}{d_1}\right),$	$q_1 = \frac{1}{2}\left(P - \frac{1}{r_1} + \frac{1}{d_1}\right),$
$\mathsf{P}'_{\mathsf{d},\mathsf{d}}, \hat{\mathsf{P}}'_{\mathsf{d},\mathsf{d}}) < P \leq \mathsf{P}_{\mathsf{rd},\mathsf{rd}}\}$	$p_2 = \frac{1}{2}\left(P + \frac{1}{r_2} - \frac{1}{d_2}\right),$	$q_2 = \frac{1}{2}\left(P - \frac{1}{r_2} + \frac{1}{d_2}\right),$

Notation for the LGRs: For the LGRs in Table 4.1, which are associated with the Waterfilling-like property, $\mathcal{A}_{x,y}, x, y \in \{\mathsf{d}, \mathsf{r}, \mathsf{rd}\}$, x and y denote the transmission status in the mm-wave links of sources S_1 and S_2 respectively as follows: for each source, subscript d,

Table 4.2: For the case with $\gamma > 1$, definition of LGRs and optimal link powers associated with the Saturation property of the MARC-OPA is given in terms of channel parameters $\boldsymbol{c} = (r_1, r_2, d_1, d_2, \gamma, P)$. Table 4.3 provides the threshold powers in terms of $(r_1, r_2, d_1, d_2, \gamma)$.

Definition of LGR	**Optimal Power Allocation**	
$\mathcal{S}_{\mathsf{r,rd}} := \{\boldsymbol{c} : \max(\mathsf{P}_{\mathsf{r,r}}, \mathsf{P}_{\mathsf{r,rd}}) <$	$p_1 = 0,$	$q_1 = P,$
$P \leq \min(\bar{\mathsf{P}}_{\mathsf{r,rd}}, \mathsf{P}_{\mathsf{r,d}})\}$	$p_2 = P - q_2,$	$q_2 = \frac{1}{r_2}\frac{\gamma}{1+Pr_1} - \frac{1}{r_2}$
$\mathcal{S}_{\mathsf{rd,r}} := \{\boldsymbol{c} : \max(\mathsf{P}_{\mathsf{r,r}}, \mathsf{P}_{\mathsf{rd,r}}) <$	$p_1 = P - q_1,$	$q_1 = \frac{1}{r_1}\frac{\gamma}{1+Pr_2} - \frac{1}{r_1},$
$P \leq \min(\bar{\mathsf{P}}_{\mathsf{rd,r}}, \mathsf{P}_{\mathsf{d,r}})\}$	$p_2 = 0,$	$q_2 = P$
$\mathcal{S}_{\mathsf{rd,d}} := \{\boldsymbol{c} : \boldsymbol{r} \in \mathcal{R}_{S1},$	$p_1 = P - \frac{\gamma-1}{r_1},$	$q_1 = \frac{\gamma-1}{r_1},$
$\max(\mathsf{P}_{\mathsf{r,d}}, \mathsf{P}_{\mathsf{rd,d}}, \bar{\mathsf{P}}_{\mathsf{rd,d}}) < P\}$	$p_2 = P,$	$q_2 = 0$
$\cup\{\boldsymbol{c} : \boldsymbol{r} \in (\mathcal{R}_1 \cup \mathcal{R}_2 \cup \mathcal{R}_{S2}),$		
$\max(\mathsf{P}_{\mathsf{r,d}}, \mathsf{P}_{\mathsf{rd,d}}) < P < \bar{\mathsf{P}}_{\mathsf{rd,d}}\}$		
$\mathcal{S}_{\mathsf{d,rd}} := \{\boldsymbol{c} : \boldsymbol{r} \in \mathcal{R}_{S2},$	$p_1 = P,$	$q_1 = 0,$
$\max(\mathsf{P}_{\mathsf{d,r}}, \mathsf{P}_{\mathsf{d,rd}}, \bar{\mathsf{P}}_{\mathsf{d,rd}}) < P\}$	$p_2 = P - \frac{\gamma-1}{r_2},$	$q_2 = \frac{\gamma-1}{r_2}$
$\cup\{\boldsymbol{c} : \boldsymbol{r} \in (\mathcal{R}_1 \cup \mathcal{R}_2 \cup \mathcal{R}_{S1}),$		
$\max(\mathsf{P}_{\mathsf{d,r}}, \mathsf{P}_{\mathsf{d,rd}}) < P < \bar{\mathsf{P}}_{\mathsf{d,rd}}\}$		
$\mathcal{S}_{\mathsf{rd,rd}} := \{\boldsymbol{c} : \boldsymbol{r} \in (\mathcal{R}_1 \cup \mathcal{R}_2),$	$p_1 = P - q_1,$	
$\max(\bar{\mathsf{P}}_{\mathsf{rd,d}}, \bar{\mathsf{P}}_{\mathsf{d,rd}}, \bar{\mathsf{P}}_{\mathsf{rd,r}}, \bar{\mathsf{P}}_{\mathsf{r,rd}}, \mathsf{P}_{\mathsf{rd,rd}}) < P\}$	$q_1 = \frac{1}{r_1}\Big(\frac{\gamma(Pr_1 + r_1 d_1^{-1} + 1)}{Pr_2 + r_2 d_2^{-1} + 1}\Big)^{1/2} - \frac{1}{r_1},$	
$\cup\{\boldsymbol{c} : \boldsymbol{r} \in \mathcal{R}_{S1},$	$p_2 = P - q_2,$	
$\max(\bar{\mathsf{P}}_{\mathsf{d,rd}}, \bar{\mathsf{P}}_{\mathsf{rd,r}}, \bar{\mathsf{P}}_{\mathsf{r,rd}}, \mathsf{P}_{\mathsf{rd,rd}}) < P \leq \bar{\mathsf{P}}_{\mathsf{rd,d}}\}$	$q_2 = \frac{1}{r_2}\Big(\frac{\gamma(Pr_2 + r_2 d_2^{-1} + 1)}{Pr_1 + r_1 d_1^{-1} + 1}\Big)^{1/2} - \frac{1}{r_2}$	
$\cup\{\boldsymbol{c} : \boldsymbol{r} \in \mathcal{R}_{S2},$		
$\max(\bar{\mathsf{P}}_{\mathsf{rd,d}}, \bar{\mathsf{P}}_{\mathsf{rd,r}}, \bar{\mathsf{P}}_{\mathsf{r,rd}}, \mathsf{P}_{\mathsf{rd,rd}}) < P \leq \bar{\mathsf{P}}_{\mathsf{d,rd}}\}$		

r and rd denotes that the MARC-OPA transmits in the direct-link only, the relay-link only and both the direct- and relay-links originating from the source, respectively. For example, in LGR $\mathcal{A}_{\mathsf{rd,d}}$ the MARC-OPA transmits in both the direct- and relay-links of source S_1, while it transmits only in the direct-link of source S_2. The LGRs $\mathcal{S}_{(.,.)}$ in Table 4.2, which are associated with the saturation property, can be similarly interpreted. Moreover, the threshold powers $\mathsf{P}_{(.,.)}, \hat{\mathsf{P}}_{(.,.)}, \mathsf{P}'_{(.,.)}, \hat{\mathsf{P}}'_{(.,.)}$ and $\bar{\mathsf{P}}_{(.,.)}$ follow the same notation as the LGRs,

Table 4.3: Definition of the threshold powers in LGRs defined in Table 4.1 and Table 4.2.

$\mathsf{P}_{\mathsf{d},\mathsf{d}} := r_1^{-1} - d_1^{-1}$,	$\mathsf{P}'_{\mathsf{d},\mathsf{d}} := -\mathsf{P}_{\mathsf{d},\mathsf{d}}$,
$\hat{\mathsf{P}}_{\mathsf{d},\mathsf{d}} := r_2^{-1} - d_2^{-1}$,	$\hat{\mathsf{P}}'_{\mathsf{d},\mathsf{d}} := -\hat{\mathsf{P}}_{\mathsf{d},\mathsf{d}}$,
$\mathsf{P}_{\mathsf{r},\mathsf{d}} := (\gamma - 1) r_1^{-1}$,	$\mathsf{P}_{\mathsf{d},\mathsf{r}} := (\gamma - 1) r_2^{-1}$,
$\mathsf{P}_{\mathsf{r},\mathsf{r}} := \varrho[(1 + xr_1)(1 + xr_2) - \gamma]$,	
$\mathsf{P}_{\mathsf{rd},\mathsf{d}} := (2\gamma - 1) r_1^{-1} - d_1^{-1}$,	$\mathsf{P}_{\mathsf{d},\mathsf{rd}} := (2\gamma - 1) r_2^{-1} - d_2^{-1}$,
$\mathsf{P}_{\mathsf{r},\mathsf{rd}} := \varrho[(1 + r_2 d_2^{-1} + xr_2)(1 + xr_1) - 2\gamma]$,	$\mathsf{P}_{\mathsf{rd},\mathsf{r}} := \varrho[(1 + r_1 d_1^{-1} + xr_1)(1 + xr_2) - 2\gamma]$,
$\mathsf{P}_{\mathsf{rd},\mathsf{rd}} := \varrho[(1 + r_1 d_1^{-1} + xr_1)$ $(1 + r_2 d_2^{-1} + xr_2) - 4\gamma]$,	
$\bar{\mathsf{P}}_{\mathsf{r},\mathsf{rd}} := \varrho[(1 + r_2/d_2 + xr_2)(1 + xr_1)^2$ $-\gamma(1 + r_1/d_1 + xr_1)]$,	$\bar{\mathsf{P}}_{\mathsf{rd},\mathsf{r}} := \varrho[(1 + r_1/d_1 + xr_1)(1 + xr_2)^2$ $-\gamma(1 + r_2/d_2 + xr_2)]$,
$\bar{\mathsf{P}}_{\mathsf{rd},\mathsf{d}} := \dfrac{\gamma - 1 + \gamma r_2 d_2^{-1} - r_1 d_1^{-1}}{r_1 - \gamma r_2}$,	$\bar{\mathsf{P}}_{\mathsf{d},\mathsf{rd}} := \dfrac{\gamma - 1 + \gamma r_1 d_1^{-1} - r_2 d_2^{-1}}{r_2 - \gamma r_1}$.

while the following relationships hold: $\mathsf{P}'_{(.,.)} := -\mathsf{P}_{(.,.)}$, and $\hat{\mathsf{P}}'_{(.,.)} = \hat{\mathsf{P}}_{(.,.)}$. Finally, threshold powers $\bar{\mathsf{P}}_{(.,.)}$ are used for LGRs $\mathcal{S}_{(.,.)}$ only, while all other threshold powers are used for both type of LGRs $\mathcal{A}_{(.,.)}$ and $\mathcal{S}_{(.,.)}$.

LGRs are Mutually Exclusive: Note that all LGRs in Table 4.1 and Table 4.2 are mutually exclusive in that, for a given tuple of channel parameters $\boldsymbol{c} = (r_1, r_2, d_1, d_2, \gamma, P)$, the condition for *one and only one* LGR holds. For example, suppose a tuple $\boldsymbol{c} \in \mathcal{A}_{\mathsf{d},\mathsf{d}}$, hence it satisfies $\min(\mathsf{P}_{\mathsf{d},\mathsf{d}}, \hat{\mathsf{P}}_{\mathsf{d},\mathsf{d}}) \geq P \geq 0$. From Table 4.3, since $\mathsf{P}'_{\mathsf{d},\mathsf{d}} := -\mathsf{P}_{\mathsf{d},\mathsf{d}}$, $\hat{\mathsf{P}}'_{\mathsf{d},\mathsf{d}} := -\hat{\mathsf{P}}_{\mathsf{d},\mathsf{d}}$ the condition $(\mathsf{P}_{\mathsf{d},\mathsf{d}}, \hat{\mathsf{P}}_{\mathsf{d},\mathsf{d}}) \succeq \mathbf{0}$ for $\mathcal{A}_{\mathsf{d},\mathsf{d}}$ requires $(\mathsf{P}'_{\mathsf{d},\mathsf{d}}, \hat{\mathsf{P}}'_{\mathsf{d},\mathsf{d}}) \preceq \mathbf{0}$, which implies that $\mathcal{A}_{\mathsf{r},\mathsf{r}} = \mathcal{A}_{\mathsf{r},\mathsf{d}} = \mathcal{A}_{\mathsf{d},\mathsf{r}} = \mathcal{A}_{\mathsf{r},\mathsf{rd}} = \mathcal{A}_{\mathsf{rd},\mathsf{r}} = \emptyset$. Next, if $\boldsymbol{c} \in \mathcal{A}_{\mathsf{d},\mathsf{d}}$, then $\boldsymbol{c} \notin \mathcal{A}_{\mathsf{rd},\mathsf{d}}$ as condition $\mathsf{P}_{\mathsf{d},\mathsf{d}} < P$ for $\mathcal{A}_{\mathsf{rd},\mathsf{d}}$ violates condition $\mathsf{P}_{\mathsf{d},\mathsf{d}} \geq P$ for $\mathcal{A}_{\mathsf{d},\mathsf{d}}$. Similarly, if $\boldsymbol{c} \in \mathcal{A}_{\mathsf{d},\mathsf{d}}$, then $\boldsymbol{c} \notin \mathcal{A}_{\mathsf{d},\mathsf{rd}}$ and $\boldsymbol{c} \notin \mathcal{A}_{\mathsf{rd},\mathsf{rd}}$. Also, when $\boldsymbol{c} \in \mathcal{A}_{\mathsf{d},\mathsf{d}}$, $\boldsymbol{c} \notin \mathcal{S}_{\mathsf{rd},\mathsf{d}}$ as condition $\mathsf{P}_{\mathsf{rd},\mathsf{d}} < P$ for $\mathcal{S}_{\mathsf{rd},\mathsf{d}}$ violates $\mathsf{P}_{\mathsf{d},\mathsf{d}} > P$ for $\mathcal{A}_{\mathsf{d},\mathsf{d}}$ since $\mathsf{P}_{\mathsf{d},\mathsf{d}} < \mathsf{P}_{\mathsf{rd},\mathsf{d}}$; similarly when $\boldsymbol{c} \in \mathcal{A}_{\mathsf{d},\mathsf{d}}$, $\boldsymbol{c} \notin \mathcal{S}_{\mathsf{d},\mathsf{rd}}$. We can also show that for $\boldsymbol{c} \in \mathcal{A}_{\mathsf{d},\mathsf{d}}$, we have $\boldsymbol{c} \notin \mathcal{S}_{\mathsf{r},\mathsf{rd}}$, $\boldsymbol{c} \notin \mathcal{S}_{\mathsf{rd},\mathsf{r}}$ and $\boldsymbol{c} \notin \mathcal{S}_{\mathsf{rd},\mathsf{rd}}$ via simple algebraic manipulations. Therefore, if $\boldsymbol{c} \in \mathcal{A}_{\mathsf{d},\mathsf{d}}$, $\boldsymbol{c}$ does not belong to any other LGR. Similarly any other LGR-pair can be shown to be mutually exclusive.

4.4.3 Properties of the Optimal Power Allocation

We observe that the MARC-OPA has two underlying properties: a *Waterfilling-like* property and a *saturation* property. First, there exists a certain *saturation threshold* $\mathsf{P}_{\mathsf{sat}}$ such that for power budget $P < \mathsf{P}_{\mathsf{sat}}$, the MARC-OPA allocates powers as follows:

- if P is *sufficiently small* such that P satisfies the condition of one of the 4 LGRs $\mathcal{A}_{\mathsf{d,d}}, \mathcal{A}_{\mathsf{r,d}}, \mathcal{A}_{\mathsf{d,r}}$ or $\mathcal{A}_{\mathsf{r,r}}$, then for each source the MARC-OPA transmits only in the *strongest* of the relay-link and direct-link originating from that source.
- as P increases, for at least one source, the MARC-OPA transmits in *both* the relay-link and direct-link originating from that source, thereby transmitting in 3 of the 4 mm-wave links. As P increases further, depending on the relative magnitudes of the link gains, the MARC-OPA may eventually transmit all 4 links. Thus, for $P < \mathsf{P}_{\mathsf{sat}}$, all link powers are either zero, or increase piecewise linearly with P.

The Waterfilling-like Property

This property of the MARC-OPA resembles the Waterfilling (WF) [138, Chapter 10.4] property for parallel AWGN channels, and thus it is referred to as the *WF-like* property. All LGRs with power allocations satisfying this property are denoted by LGRs $\mathcal{A}_{x,y}, x, y \in \{\mathsf{d}, \mathsf{r}, \mathsf{rd}\}$. Specifically, depending on the gains of the direct-links ($\mathtt{DL}_1$ and $\mathtt{DL}_2$) and the gains of the relay-links ($\mathtt{RL}_1$ and $\mathtt{RL}_2$), the MARC-OPA transmits in one of the following sets of links: (i) $\mathtt{DL}_1$ and $\mathtt{DL}_2$ if $d_1 \geq r_1$, $d_2 \geq r_2$, (ii) $\mathtt{RL}_1$ and $\mathtt{RL}_2$ if $r_1 > d_1$, $r_2 > d_2$, (iii) $\mathtt{DL}_1$ and $\mathtt{RL}_2$ if $d_1 \geq r_1$, $r_2 > d_2$, and (iv) $\mathtt{DL}_2$ and $\mathtt{RL}_1$ if $d_2 \geq r_2$, $r_1 > d_1$. Note that the corresponding LGRs are $\mathcal{A}_{\mathsf{d,d}}$, $\mathcal{A}_{\mathsf{r,r}}$, $\mathcal{A}_{\mathsf{d,r}}$ and $\mathcal{A}_{\mathsf{r,d}}$.

Since the marginal return from transmitting only in the strongest link of each source diminishes as P increases, for sufficiently large P (that is still below $\mathsf{P}_{\mathsf{sat}}$) the MARC-OPA transmits in one additional link. The chosen link, however, depends on the relative magnitude of the link gains, and provides the maximum increase in the sum-rate among all inactive links.

For example, consider a given (r_1, r_2, d_1, d_2, P)-tuple such that for $P < \min(\mathsf{P}_{\mathsf{d,d}}, \hat{\mathsf{P}}_{\mathsf{d,d}})$, the MARC-OPA transmits in links $\mathtt{DL}_1$ and $\mathtt{DL}_2$ only, i.e., as in $\mathcal{A}_{\mathsf{d,d}}$. Now, if $\mathsf{P}_{\mathsf{d,d}} < \hat{\mathsf{P}}_{\mathsf{d,d}}$ holds, and P increases to $\mathsf{P}_{\mathsf{d,d}} \leq P \leq \min(\hat{\mathsf{P}}_{\mathsf{d,d}}, \mathsf{P}_{\mathsf{rd,d}})$, then the MARC-OPA transmits in the relay-link $\mathtt{RL}_1$, in addition to the two direct-links, following the allocation in LGR $\mathcal{A}_{\mathsf{rd,d}}$.

Note that through LGRs $\mathcal{A}_{\mathsf{d,d}}$ and $\mathcal{A}_{\mathsf{rd,d}}$, the powers p_1 and q_1 increase piecewise linearly with P, while $p_2 = P$ increasing linearly with P and $q_2 = 0$, as per the WF-like property.

Based on the intuition for LGR $\mathcal{A}_{\mathsf{rd,d}}$ above, the other LGRs that follow the WF-like property, i.e., $\mathcal{A}_{\mathsf{d,rd}}$, $\mathcal{A}_{\mathsf{r,rd}}$ and $\mathcal{A}_{\mathsf{rd,r}}$, can be explained as well. Specifically, the intuition behind LGR $\mathcal{A}_{\mathsf{d,rd}}$ follows by swapping the roles of the sources compared to $\mathcal{A}_{\mathsf{rd,d}}$, whereas the intuition behind $\mathcal{A}_{\mathsf{r,rd}}$ and $\mathcal{A}_{\mathsf{rd,r}}$ follow from $\mathcal{A}_{\mathsf{d,rd}}$ and $\mathcal{A}_{\mathsf{rd,d}}$ respectively by exchanging the roles of the relay-links as well as the direct-links. Finally, in $\mathcal{A}_{\mathsf{rd,rd}}$ the MARC-OPA transmits in all 4 links following a Waterfilling allocation.

The Saturation Property

While for $P < \mathsf{P}_{\mathsf{sat}}$, the MARC-OPA follows the WF-like property, for $P \geq \mathsf{P}_{\mathsf{sat}}$, the MARC-OPA limits the relay-link powers such that the saturation condition holds, which is expressed as

$$(1 + r_1 q_1)(1 + r_2 q_2) = \gamma. \tag{4.41}$$

Thus, as P increases beyond $\mathsf{P}_{\mathsf{sat}}$, q_1 and q_2 can *no longer both increase* with P. In contrast, the direct-link powers $p_k = P - q_k$, increase unbounded with P. This property is referred to as the saturation property and is clearly unlike WF-like property.

The 5 LGRs satisfying this property are denoted by $\mathcal{S}_{(.,.)}$ in Table 4.2. Given a set of parameters $(r_1, r_2, d_1, d_2, \gamma)$, as P increases, saturation *first* occurs in one of LGRs $\mathcal{S}_{(.,.)}$, called the *saturation LGR*. The specific saturation LGR depends on the given set of parameters $(r_1, r_2, d_1, d_2, \gamma)$, or equivalently on how the resulting threshold powers compare. In either case, the saturation threshold $\mathsf{P}_{\mathsf{sat}}$ is given by the lower bound on P in the respective LGRs $\mathcal{S}_{(.,.)}$ in Table 4.2. For example, if the saturation LGR is $\mathcal{S}_{\mathsf{r,rd}}$, then $\mathsf{P}_{\mathsf{sat}} = \max(\mathsf{P}_{\mathsf{r,r}}, \mathsf{P}_{\mathsf{r,rd}})$.

While saturation may occur in one of the 5 possible LGRs $\mathcal{S}_{(.,.)}$ in Table 4.2 that depend on the given set of parameters $(r_1, r_2, d_1, d_2, \gamma)$, the general intuition behind all 5 saturation cases are the same. We explain below the intuition behind the saturation phenomena by observing how the so-called sum-rate-gap, defined below in (4.42), behaves as the power budget P increases from $P = 0$ to the saturation threshold $\mathsf{P}_{\mathsf{sat}}$. Recall that the objective of the MARC-OPA is to maximize $R = \min\{\Sigma_{\mathsf{R}}, \Sigma_{\mathsf{D}}\}$. For a given $P \geq 0$ and a feasible power allocation (p_1, q_1, p_2, q_2), we define the sum-rate-gap as follows

$$\Delta R(p_1, q_1, p_2, q_2) := \Sigma_{\mathsf{D}} - \Sigma_{\mathsf{R}} = \sigma_{\mathsf{D}} - \sigma_{\mathsf{R}} - \alpha \left(\log\left(1 + r_1 q_1\right) + \log\left(1 + r_2 q_2\right) \right), \tag{4.42}$$

where Σ_{R} and Σ_{D} are expressed in (4.30)-(4.31). Clearly, ΔR is not a function of p_k, but it is a decreasing function of q_k. Also, since the sum-rate is $R = \min\{\Sigma_{\mathsf{R}}, \Sigma_{\mathsf{D}}\}$, so long as $\Delta R > 0$, only $R = \Sigma_{\mathsf{R}} < \Sigma_{\mathsf{D}}$ is achieved.

We can now explain saturation by observing how ΔR changes as P increases from $P = 0$. First, note that at $P = 0$, the resulting allocation is $p_k = q_k = 0$, and therefore the resulting rates are $\Sigma_{\mathsf{D}} = \sigma_{\mathsf{D}}$ and $\Sigma_{\mathsf{R}} = \sigma_{\mathsf{R}}$, with σ_{R} and σ_{D} defined in (4.32). Since in the current case we have $\gamma > 1$, which implies $\sigma_{\mathsf{D}} > \sigma_{\mathsf{R}}$ from (4.39), we have the sum-rate-gap $\Delta R = \sigma_{\mathsf{D}} - \sigma_{\mathsf{R}} > 0$. Hence, at $P = 0$ only $R = \Sigma_{\mathsf{R}} < \Sigma_{\mathsf{D}}$ is achieved.

Next, recall that as P increases from $P = 0$, due to the Waterfilling-like property, link powers p_k and q_k increase piecewise linearly with P. Since ΔR is not a function of p_k, if p_k increases with P, ΔR remains *unaffected*. In contrast, if q_k increases with P, ΔR generally *reduces* from $\Delta R|_{P=0} = \sigma_{\mathsf{D}} - \sigma_{\mathsf{R}} > 0$. More specifically, except for LGR $\mathcal{A}_{\mathsf{d,d}}$, for all other LGRs in Table 4.1, as P increases, either q_1 or q_2, or both increase with P, and thus in all these LGRs ΔR monotonically decreases with P. As such, when P is sufficiently increased, q_1 and q_2 are alloted enough power to reduce ΔR to $\Delta R = 0$. This is achieved at $P = \mathsf{P}_{\mathsf{sat}}$, the so-called saturation threshold, and the resulting condition $\Delta R = 0 \iff (1 + r_1 q_1)(1 + r_2 q_2) = \gamma$, is called the saturation condition given in (4.41). In terms of rates, now the sum-rate $R = \Sigma_{\mathsf{R}} = \Sigma_{\mathsf{D}}$ is achieved.

If P is increased beyond $\mathsf{P}_{\mathsf{sat}}$, it turns out to be *optimal* to constrain q_1 and q_2 such that $\Delta R = 0$ continues to hold. Therefore, as P is increased beyond $\mathsf{P}_{\mathsf{sat}}$, all additional increase in P needs to be allocated to the direct links, i.e., $p_k = P - q_k$. As such, for all $P \geq \mathsf{P}_{\mathsf{sat}}$, the sum-rate $R = \Sigma_{\mathsf{R}} = \Sigma_{\mathsf{D}}$ is achieved.

Optimal Powers in Saturation LGRs: As noted earlier, for a given set of parameter $(r_1, r_2, d_1, d_2, \gamma)$ saturation can first occurs in one of 5 LGRs $\mathcal{S}_{(.,.)}$, depending on which the optimal powers vary differently with P. Specifically, in $\mathcal{S}_{\mathsf{r,rd}}$, as P increases, $q_1 = P$ increases *linearly* with P, and thus $p_1 = P - q_1 = 0$. However, due to saturation, since $q_1 = P$ we have $q_2 = (\frac{\gamma}{1+Pr_1} - 1)/r_2$, which *decreases non-linearly* with P. The same trend is found in the complementary LGR $\mathcal{S}_{\mathsf{rd,r}}$ after the role of the two sources are swapped compared to $\mathcal{S}_{\mathsf{r,rd}}$. In $\mathcal{S}_{\mathsf{rd,rd}}$, as P increases, if $r_1 \geq r_2$ (resp. $r_1 < r_2$), q_1 and q_2 (resp. q_2 and q_1) monotonically increase and decrease non-linearly with P, while both p_1 and p_2 increase non-linearly. Finally, in $\mathcal{S}_{\mathsf{rd,d}}$, as P increases, $q_1 = \frac{\gamma-1}{r_1}$ and $q_2 = 0$ remain fixed, and all additional increments of P are allotted to the direct-links, whereas in the complementary LGR $\mathcal{S}_{\mathsf{d,rd}}$, the same trend is followed with the roles of the sources exchanged.

Saturation LGR and Final LGR: For a given set of parameters $(r_1, r_2, d_1, d_2, \gamma)$ as P increases, while saturation *first* occurs in one of LGRs $\mathcal{S}_{(.,.)}$ for $P \geq \mathsf{P}_{\mathsf{sat}}$ associated with that LGR, as P increases further, one or more other LGRs $\mathcal{S}_{(.,.)}$ may become optimal where *saturation continues* to hold. More specifically, there exists a threshold $\mathsf{P}_{\mathsf{fin}} \geq \mathsf{P}_{\mathsf{sat}}$ such that for all $P \geq \mathsf{P}_{\mathsf{fin}}$, a specific LGR $\mathcal{S}_{(.,.)}$ remain optimal, denoted the *final* LGR.

To describe this more precisely, recall that in (4.40), we partitioned the relay-link gains $\boldsymbol{r} := (r_1, r_2)$ into subsets $\mathcal{R}_{S1} := \{\boldsymbol{r} : r_1 \geq \gamma r_2\}$, $\mathcal{R}_1 := \{\boldsymbol{r} : \gamma r_2 > r_1 \geq r_2\}$, $\mathcal{R}_2 := \{\boldsymbol{r} : \gamma r_1 > r_2 > r_1\}$, and $\mathcal{R}_{S2} := \{\boldsymbol{r} : r_2 \geq \gamma r_1\}$. Intuitively, in $\mathcal{R}_{S2}$, relay-link $\mathtt{RL}_2$ is *significantly stronger* than $\mathtt{RL}_1$ (i.e., $r_2 \geq \gamma r_1$) while in $\mathcal{R}_2$, it is only stronger (i.e., $r_2 > r_1$) but not significantly stronger (i.e., $r_2 < \gamma r_1$). The intuitions for $\mathcal{R}_{S1}$ and $\mathcal{R}_1$ follow similarly. We observe that for a given set of parameters $(r_1, r_2, d_1, d_2, \gamma)$, if

- $\boldsymbol{r} \in \mathcal{R}_1$ or $\boldsymbol{r} \in \mathcal{R}_2$: $\mathsf{P}_{\mathsf{fin}} = \max(\bar{\mathsf{P}}_{\mathsf{rd,d}}, \bar{\mathsf{P}}_{\mathsf{d,rd}}, \bar{\mathsf{P}}_{\mathsf{rd,r}}, \bar{\mathsf{P}}_{\mathsf{r,rd}}, \mathsf{P}_{\mathsf{rd,rd}})$, and the final LGR is $\mathcal{S}_{\mathsf{rd,rd}}$.
- $\boldsymbol{r} \in \mathcal{R}_{S1}$: $\mathsf{P}_{\mathsf{fin}} = \max(\mathsf{P}_{\mathsf{r,d}}, \mathsf{P}_{\mathsf{rd,d}}, \bar{\mathsf{P}}_{\mathsf{rd,d}})$, and the final LGR is $\mathcal{S}_{\mathsf{rd,d}}$.
- $\boldsymbol{r} \in \mathcal{R}_{S2}$: $\mathsf{P}_{\mathsf{fin}} = \max(\mathsf{P}_{\mathsf{d,r}}, \mathsf{P}_{\mathsf{d,rd}}, \bar{\mathsf{P}}_{\mathsf{d,rd}})$, and the final LGR is $\mathcal{S}_{\mathsf{d,rd}}$.

Naturally, for some link gain tuples, the saturation and the final LGRs are the same, and thus $\mathsf{P}_{\mathsf{fin}} = \mathsf{P}_{\mathsf{sat}}$.

4.4.4 Evolution of Link Gain Regimes with the Power Budget

In Table 4.1 and Table 4.2, the LGRs are defined as partitions of the set of the power budget P. Since the threshold powers in Table 4.3 are functions of link gains, for a given $(r_1, r_2, d_1, d_2, \gamma)$-tuple and P, it is easy to determine which LGR is *active* (i.e., the condition of which LGR holds). It is evident that, as P increases, the active LGR may change as well, and thus the MARC-OPA follows a *set of active* LGRs, called a *LGR-path*, which reveals useful insights on the optimal power allocation.

Recall that for a given $(r_1, r_2, d_1, d_2, \gamma)$-tuple, saturation can occur in either of 5 LGRs, $\mathcal{S}_{\mathsf{r,rd}}, \mathcal{S}_{\mathsf{rd,r}}$, $\mathcal{S}_{\mathsf{rd,rd}}, \mathcal{S}_{\mathsf{rd,d}}$ and $\mathcal{S}_{\mathsf{d,rd}}$, which leads to a vast number of LGR-paths and makes it difficult to interpret insights. Hence, to simplify the exposition, we now assume the gains of the direct-links to be the same, i.e., $d := d_1 = d_2$. Although this causes some loss of generality, the resulting LGR-paths are simplified. For example, under this assumption, for $\boldsymbol{r} \in \mathcal{R}_2$, LGRs $\mathcal{A}_{\mathsf{r,d}} = \mathcal{A}_{\mathsf{rd,d}} = \mathcal{A}_{\mathsf{r,rd}} = \mathcal{S}_{\mathsf{r,rd}} = \mathcal{S}_{\mathsf{rd,d}} = \mathcal{S}_{\mathsf{d,rd}} = \emptyset$, and thus saturation can

now occur only in either $\mathcal{S}_{\mathsf{rd,r}}$ or $\mathcal{S}_{\mathsf{rd,rd}}$. Nonetheless, the paths for the case with $d_1 \neq d_2$ can be similarly characterized.

In this section, we discuss the paths for $\boldsymbol{r} \in \mathcal{R}_2$ and $\boldsymbol{r} \in \mathcal{R}_{S2}$ only, as the paths for $\boldsymbol{r} \in \mathcal{R}_1$ and $\boldsymbol{r} \in \mathcal{R}_{S1}$ can be derived from those for $\boldsymbol{r} \in \mathcal{R}_2$ and $\boldsymbol{r} \in \mathcal{R}_{S2}$, respectively, by exchanging the roles of relay-links $\mathtt{RL}_2$ and $\mathtt{RL}_1$ as well as those of direct-links $\mathtt{DL}_2$ and $\mathtt{DL}_1$.

Case $\boldsymbol{r} \in \mathcal{R}_2$:

In this case, we have 7 LGR-paths denoted $[S1], \ldots, [S7]$ which are presented in Table 4.4 along with their underlying conditions, and the interval of P for each LGR in the path.

Note that the conditions in Table 4.4 are mutually exclusive for $\boldsymbol{r} \in \mathcal{R}_2$. Moreover, for ease of exposition, we have omitted the boundary cases in which one or more inequalities in the condition of an LGR-path hold with equality. In such boundary cases, one or more LGRs in the LGR-path becomes empty. The LGR-path is then simplified by omitting the empty LGRs, and the simplified LGR-path may turn out to be the same as some other LGR-path. For example, consider paths $[S2]$ and $[S3]$ in Table 4.4. If the channel parameters are such that the two threshold powers $\mathsf{P}_{\mathsf{r,r}}$ and $\mathsf{P}'_{\mathsf{d,d}}$ are equal, i.e., $\mathsf{P}_{\mathsf{r,r}} = \mathsf{P}'_{\mathsf{d,d}}$, then simple algebraic manipulation reveals that LGR $\mathcal{A}_{\mathsf{rd,rd}}$ is empty. The simplified path, obtained by omitting this empty LGR from $[S2]$, turns out to be the same as path $[S3]$.

Initial LGR: While paths $[S1]$, $[S2]$, $[S3]$ originate from the *initial* LGR $\mathcal{A}_{\mathsf{r,r}}$, path $[S4]$ originates from $\mathcal{A}_{\mathsf{d,d}}$, and paths $[S5], [S6], [S7]$ from $\mathcal{A}_{\mathsf{d,r}}$. The initial LGRs vary based on how the direct-link gain d compares to the relay-link gains r_1 and r_2. For example, if $d \geq r_2 \geq r_1 \iff 0 \leq \hat{\mathsf{P}}_{\mathsf{d,d}} \leq \mathsf{P}_{\mathsf{d,d}}$ (i.e., each $\mathtt{DL}_k$ is stronger than its corresponding $\mathtt{RL}_k$), for sufficiently small P, following the WF-like property the MARC-OPA transmits only in the direct-links as in LGR $\mathcal{A}_{\mathsf{d,d}}$, and thus $\mathcal{A}_{\mathsf{d,d}}$ serves as the initial LGR.

On the other hand, if $r_2 \geq r_1 > d \iff 0 < \mathsf{P}'_{\mathsf{d,d}} \leq \hat{\mathsf{P}}'_{\mathsf{d,d}}$ (i.e., each $\mathtt{RL}_k$ is stronger than its corresponding $\mathtt{DL}_k$), for sufficiently small P, following the WF-like property the MARC-OPA transmits only in the relay-links as in $\mathcal{A}_{\mathsf{r,r}}$. Furthermore, depending on how threshold powers $\mathsf{P}'_{\mathsf{d,d}}, \hat{\mathsf{P}}'_{\mathsf{d,d}}$ and $\mathsf{P}_{\mathsf{r,r}}$ compare, the MARC-OPA follows one of the paths $[S1], [S2], [S3]$, as given in Table 4.4. Similarly, for the case of $r_2 > d > r_1 \iff \mathsf{P}_{\mathsf{d,d}} > 0, \hat{\mathsf{P}}'_{\mathsf{d,d}} > 0$, the MARC-OPA transmits in the two stronger links $\mathtt{RL}_2$ and $\mathtt{DL}_1$ as in $\mathcal{A}_{\mathsf{d,r}}$. Then, based on how threshold powers $\mathsf{P}_{\mathsf{d,d}}, \hat{\mathsf{P}}'_{\mathsf{d,d}}$ and $\mathsf{P}_{\mathsf{rd,r}}$ compare, one of paths $[S5], [S6], [S7]$ is followed.

Table 4.4: LGR paths for $\boldsymbol{r} \in \mathcal{R}_2$. Table 4.3 provides the threshold powers in terms of link gains and γ. Each path originates from one of three initial LGRs $\mathcal{A}_{\mathsf{r,r}}, \mathcal{A}_{\mathsf{d,d}}$ or $\mathcal{A}_{\mathsf{d,r}}$, and terminate at the final LGR $\mathcal{S}_{\mathsf{rd,rd}}$. If for certain values of link gains and γ, one or more LGRs in the path becomes empty, the path is simplified by omitting the empty LGRs.

LGR path	**Condition**	**Interval of P in each LGR**
$[S1]: \mathcal{A}_{\mathsf{r,r}} \to \mathcal{A}_{\mathsf{rd,r}} \to \mathcal{A}_{\mathsf{rd,rd}} \to \mathcal{S}_{\mathsf{rd,rd}}$	$0 < \mathsf{P}'_{\mathsf{d,d}} < \hat{\mathsf{P}}'_{\mathsf{d,d}} < \mathsf{P}_{\mathsf{r,r}}$	$[0, \mathsf{P}'_{\mathsf{d,d}})$, $[\mathsf{P}'_{\mathsf{d,d}}, \hat{\mathsf{P}}'_{\mathsf{d,d}})$, $[\hat{\mathsf{P}}'_{\mathsf{d,d}}, \mathsf{P}_{\mathsf{rd,rd}})$, $[\mathsf{P}_{\mathsf{rd,rd}}, \infty)$
$[S2]: \mathcal{A}_{\mathsf{r,r}} \to \mathcal{A}_{\mathsf{rd,r}} \to \mathcal{S}_{\mathsf{rd,r}} \to \mathcal{S}_{\mathsf{rd,rd}}$	$0 < \mathsf{P}'_{\mathsf{d,d}} < \mathsf{P}_{\mathsf{r,r}} < \hat{\mathsf{P}}'_{\mathsf{d,d}}$	$[0, \mathsf{P}'_{\mathsf{d,d}})$, $[\mathsf{P}'_{\mathsf{d,d}}, \mathsf{P}_{\mathsf{rd,r}})$, $[\mathsf{P}_{\mathsf{rd,r}}, \bar{\mathsf{P}}_{\mathsf{rd,r}})$, $[\bar{\mathsf{P}}_{\mathsf{rd,r}}, \infty)$
$[S3]: \mathcal{A}_{\mathsf{r,r}} \to \mathcal{S}_{\mathsf{rd,r}} \to \mathcal{S}_{\mathsf{rd,rd}}$	$0 < \mathsf{P}_{\mathsf{r,r}} < \mathsf{P}'_{\mathsf{d,d}} < \hat{\mathsf{P}}'_{\mathsf{d,d}}$	$[0, \mathsf{P}_{\mathsf{r,r}})$, $[\mathsf{P}_{\mathsf{r,r}}, \bar{\mathsf{P}}_{\mathsf{rd,r}})$, $[\bar{\mathsf{P}}_{\mathsf{rd,r}}, \infty)$
$[S4]: \mathcal{A}_{\mathsf{d,d}} \to \mathcal{A}_{\mathsf{d,rd}} \to \mathcal{A}_{\mathsf{rd,rd}} \to \mathcal{S}_{\mathsf{rd,rd}}$	$0 < \hat{\mathsf{P}}_{\mathsf{d,d}} < \mathsf{P}_{\mathsf{d,d}}$	$[0, \hat{\mathsf{P}}_{\mathsf{d,d}})$, $[\hat{\mathsf{P}}_{\mathsf{d,d}}, \mathsf{P}_{\mathsf{d,d}})$, $[\mathsf{P}_{\mathsf{d,d}}, \mathsf{P}_{\mathsf{rd,rd}})$, $[\mathsf{P}_{\mathsf{rd,rd}}, \infty)$
$[S5]: \mathcal{A}_{\mathsf{d,r}} \to \mathcal{A}_{\mathsf{rd,r}} \to \mathcal{A}_{\mathsf{rd,rd}} \to \mathcal{S}_{\mathsf{rd,rd}}$	$0 < \mathsf{P}_{\mathsf{d,d}} < \hat{\mathsf{P}}'_{\mathsf{d,d}} < \mathsf{P}_{\mathsf{rd,r}}$	$[0, \mathsf{P}_{\mathsf{d,d}})$, $[\mathsf{P}_{\mathsf{d,d}}, \hat{\mathsf{P}}'_{\mathsf{d,d}})$ $[\hat{\mathsf{P}}'_{\mathsf{d,d}}, \mathsf{P}_{\mathsf{rd,rd}})$, $[\mathsf{P}_{\mathsf{rd,rd}}, \infty)$
$[S6]: \mathcal{A}_{\mathsf{d,r}} \to \mathcal{A}_{\mathsf{rd,r}} \to \mathcal{S}_{\mathsf{rd,r}} \to \mathcal{S}_{\mathsf{rd,rd}}$	$0 < \mathsf{P}_{\mathsf{d,d}} < \mathsf{P}_{\mathsf{rd,r}} < \hat{\mathsf{P}}'_{\mathsf{d,d}}$	$[0, \mathsf{P}_{\mathsf{d,d}})$, $[\mathsf{P}_{\mathsf{d,d}}, \mathsf{P}_{\mathsf{rd,r}})$, $[\mathsf{P}_{\mathsf{rd,r}}, \bar{\mathsf{P}}_{\mathsf{rd,r}})$, $[\bar{\mathsf{P}}_{\mathsf{rd,r}}, \infty)$
$[S7]: \mathcal{A}_{\mathsf{d,r}} \to \mathcal{A}_{\mathsf{d,rd}} \to \mathcal{A}_{\mathsf{rd,rd}} \to \mathcal{S}_{\mathsf{rd,rd}}$	$0 < \hat{\mathsf{P}}'_{\mathsf{d,d}} < \mathsf{P}_{\mathsf{d,d}}$	$[0, \hat{\mathsf{P}}'_{\mathsf{d,d}})$, $[\hat{\mathsf{P}}'_{\mathsf{d,d}}, \mathsf{P}_{\mathsf{d,d}})$, $[\mathsf{P}_{\mathsf{d,d}}, \mathsf{P}_{\mathsf{rd,rd}})$, $[\mathsf{P}_{\mathsf{rd,rd}}, \infty)$

Saturation Cases: In this case, saturation first occurs in either LGR $\mathcal{S}_{\mathsf{rd,rd}}$ or LGR $\mathcal{S}_{\mathsf{rd,r}}$ as follows.

Saturation occurs in $\mathcal{S}_{\mathsf{rd,rd}}$ if the condition of one of the paths $[S1], [S4], [S5]$ or $[S7]$ is met. Here, $\mathsf{P}_{\mathsf{sat}} = \mathsf{P}_{\mathsf{rd,rd}}$, and as P increases for all $P \geq \mathsf{P}_{\mathsf{rd,rd}}$, q_2 increases monotonically and q_1 decreases monotonically, and they approach constants $q_k \to \bar{q}_k := \sqrt{\frac{\gamma}{r_1 r_2}} - \frac{1}{r_k} > 0$, as $P \to \infty$. Intuitively, since in $\mathcal{S}_{\mathsf{rd,rd}}$, condition $(1 + r_1 q_1)(1 + r_2 q_2) = \gamma$ must hold, as P increases, q_1 and q_2 both cannot increase. Moreover, since in the case with $\boldsymbol{r} \in \mathcal{R}_2$, relay-link $\mathtt{RL}_2$ is stronger than $\mathtt{RL}_1$, as P increases, the MARC-OPA achieves the best rate by *increasing* q_2 and *decreasing* q_1. However, since $\mathtt{RL}_2$ is *not significantly* stronger than

$\mathtt{RL}_1$, instead of only in $\mathtt{RL}_2$, the MARC-OPA should transmit in *both* relay-links for all $P \geq \mathsf{P}_{\mathsf{rd,rd}}$. Thus, q_1 and q_2 both remain non-zero and approach constant levels as $P \to \infty$.

On the other hand, saturation first occurs in LGR $\mathcal{S}_{\mathsf{rd,r}}$ if the condition of one of the paths $[S2], [S3]$ or $[S6]$ holds. Here, $\mathsf{P}_{\mathsf{sat}} = \max(\mathsf{P}_{\mathsf{rd,r}}, \mathsf{P}_{\mathsf{r,r}})$, and $\mathcal{S}_{\mathsf{rd,r}}$ is active for only $\max(\mathsf{P}_{\mathsf{rd,r}}, \mathsf{P}_{\mathsf{r,r}}) \leq P \leq \bar{\mathsf{P}}_{\mathsf{rd,r}}$. When $\mathcal{S}_{\mathsf{rd,r}}$ is active, for source S_2, the MARC-OPA allocates $(p_2, q_2) = (0, P)$. In contrast, for source S_1, the MARC-OPA allocates $(p_1, q_1) = (P - q_1, \frac{1}{r_1}(\frac{\gamma}{1+Pr_2} - 1))$. Clearly, as P increases in this range, $q_2 = P$ *increases* and $q_1 = \frac{1}{r_1}(\frac{\gamma}{1+Pr_2} - 1)$ *decreases*, and hence the MARC-OPA follows the same trend as in $\mathcal{S}_{\mathsf{rd,rd}}$ discussed above.

Final LGR: For $P \geq \mathsf{P}_{\mathsf{fin}} = \max(\mathsf{P}_{\mathsf{rd,rd}}, \bar{\mathsf{P}}_{\mathsf{rd,r}})$, all paths terminate at the final LGR $\mathcal{S}_{\mathsf{rd,rd}}$.

LGR-paths: We first discuss path $[S2]$ in some detail, and interpret its underlying condition $0 < \mathsf{P}'_{\mathsf{d,d}} < \mathsf{P}_{\mathsf{r,r}} < \hat{\mathsf{P}}'_{\mathsf{d,d}}$ as follows:

- In this case, each $\mathtt{RL}_k$ is stronger than its corresponding $\mathtt{DL}_k, k = 1, 2$. In terms of channel gains, this is equivalent to $r_2 > r_1 > d \iff 0 < \mathsf{P}'_{\mathsf{d,d}} < \hat{\mathsf{P}}'_{\mathsf{d,d}}$. Hence, for $P \in [0, \mathsf{P}'_{\mathsf{d,d}})$, following the WF-like property the MARC-OPA allocates P entirely to $\mathtt{RL}_1$ and $\mathtt{RL}_2$ as in $\mathcal{A}_{\mathsf{r,r}}$, where $q_1 = q_2 = P$ increase with P, while $p_1 = p_2 = 0$.

- As P increases, the return from transmitting only in the relay-links decreases. Due to condition $0 < \mathsf{P}'_{\mathsf{d,d}} < \mathsf{P}_{\mathsf{r,r}}$, for the range of $P \in [\mathsf{P}'_{\mathsf{d,d}}, \mathsf{P}_{\mathsf{rd,r}})$, the MARC-OPA achieves the best rate by transmitting in *both* $\mathtt{DL}_1$ and $\mathtt{RL}_1$ as in LGR $\mathcal{A}_{\mathsf{rd,r}}$. Hence, for $P \in [\mathsf{P}'_{\mathsf{d,d}}, \mathsf{P}_{\mathsf{rd,r}})$, following the WF-like property the MARC-OPA allocates power as in $\mathcal{A}_{\mathsf{rd,r}}$ where p_1, q_1 and q_2 increase with P, and $p_2 = 0$.

- Due to condition $0 < \mathsf{P}_{\mathsf{r,r}} < \hat{\mathsf{P}}'_{\mathsf{d,d}}$, as P is increased further to $P = \mathsf{P}_{\mathsf{sat}} = \mathsf{P}_{\mathsf{rd,r}}$, saturation first occurs in LGR $\mathcal{S}_{\mathsf{rd,r}}$ as it becomes active. In this LGR, the MARC-OPA allocates $p_1 = P - q_1, q_1 = \frac{1}{r_1}\frac{\gamma}{1+Pr_2} - \frac{1}{r_1}$, and $p_2 = 0, q_2 = P$, which remains optimal for the range of $P \in [\mathsf{P}_{\mathsf{rd,r}}, \bar{\mathsf{P}}_{\mathsf{rd,r}})$.

- Finally, for $P \geq \bar{\mathsf{P}}_{\mathsf{rd,r}}$, LGR $\mathcal{S}_{\mathsf{rd,rd}}$ becomes active.

Path $[S1]$ is similar to $[S2]$ as the first two LGRs and the final LGR are the same. The only difference between the two is that condition $0 < \hat{\mathsf{P}}'_{\mathsf{d,d}} < \mathsf{P}_{\mathsf{r,r}}$ in $[S1]$ is opposite to that in $[S2]$, and due to this LGR $\mathcal{S}_{\mathsf{rd,r}}$ in $[S2]$ is swapped with LGR $\mathcal{A}_{\mathsf{rd,rd}}$ in $[S1]$. Hence in

$[S1]$, as P increases, instead of saturation, the MARC-OPA continues to allocate power according to LGR $\mathcal{A}_{\mathsf{rd,rd}}$ following the WF-like property. Finally, as P is increased further, saturation occurs in LGR $\mathcal{S}_{\mathsf{rd,rd}}$, which remains active for $P \geq \mathsf{P}_{\mathsf{sat}} = \mathsf{P}_{\mathsf{rd,rd}}$.

Path $[S3]$ is similar to $[S2]$ except that LGR $\mathcal{A}_{\mathsf{rd,r}}$ is now skipped. The difference between the two paths is that as opposed that in $[S2]$, we now have condition $\mathsf{P}_{\mathsf{r,r}} < \mathsf{P}'_{\mathsf{d,d}} < \hat{\mathsf{P}}'_{\mathsf{d,d}}$. Hence, as P is increased, $\mathcal{A}_{\mathsf{rd,r}}$ is skipped and instead $\mathcal{S}_{\mathsf{rd,r}}$ becomes active at $P = \mathsf{P}_{\mathsf{sat}} = \mathsf{P}_{\mathsf{r,r}}$ where saturation occurs first. Finally, for $P \geq \bar{\mathsf{P}}_{\mathsf{rd,r}}$, LGR $\mathcal{S}_{\mathsf{rd,rd}}$ is active.

Path $[S4]$ is complementary to $[S1]$ in that each $\mathtt{DL}_k$ is now stronger than its corresponding $\mathtt{RL}_k, k = 1, 2$. In terms of link gains, this is equivalent to $d \geq r_2 \geq r_1 \iff 0 \leq \hat{\mathsf{P}}_{\mathsf{d,d}} \leq \mathsf{P}_{\mathsf{d,d}}$. As a result, compared to the first two LGRs $\mathcal{A}_{\mathsf{r,r}}$ and $\mathcal{A}_{\mathsf{rd,r}}$ in $[S1]$, the first two in $[S4]$ are $\mathcal{A}_{\mathsf{d,d}}$ and $\mathcal{A}_{\mathsf{d,rd}}$, while the remaining two LGRs are the same as in $[S1]$. Similar to $[S1]$, in $[S4]$ for sufficiently small P, the MARC-OPA allocates power following the WF-like property as LGRs $\mathcal{A}_{\mathsf{d,d}}$, $\mathcal{A}_{\mathsf{d,rd}}$ and $\mathcal{A}_{\mathsf{rd,rd}}$ become active, and then saturation occurs in LGR $\mathcal{S}_{\mathsf{rd,rd}}$ for $P \geq \mathsf{P}_{\mathsf{sat}} = \mathsf{P}_{\mathsf{rd,rd}}$. The details are similar to those of $[S1]$, and hence are omitted.

Finally, for the case of $r_2 > d > r_1$, where $\mathtt{DL}_1$ is stronger than $\mathtt{RL}_1$ and $\mathtt{DL}_2$ is weaker than $\mathtt{RL}_2$, the MARC-OPA follows $[S5], [S6]$ and $[S7]$ in a manner similar to that of $[S1], [S2]$ and $[S4]$ respectively. While the exact thresholds of P and specific LGRs in a path vary depending on a given set of channel parameters $(r_1, d_1, r_2, d_2, \gamma)$, overall the the MARC-OPA allocates the link-powers in accordance to its WF-like and saturation properties. The details are omitted to avoid repetition.

Case $\boldsymbol{r} \in \mathcal{R}_{S2}$:

In this case, we have 10 paths, denoted $[T3], \ldots, [T7]$ and $[N1], \ldots, [N5]$ and given in Table 4.5. Note that paths $[T3], \ldots, [T7]$ are the counterparts of paths $[S3], \ldots, [S7]$ in Table 4.4 with $\mathcal{S}_{\mathsf{d,rd}}$ appended as the final LGR, and thus are denoted in this manner. Also, paths $[S1]$ and $[S2]$ do not have any counterparts here, and thus $[T1]$ and $[T2]$ are not defined. Moreover, paths $[N1], \ldots, [N5]$ are valid exclusively for $\boldsymbol{r} \in \mathcal{R}_{S2}$.

Similar to the case with $\boldsymbol{r} \in \mathcal{R}_2$, the conditions in Table 4.5 are mutually exclusive for $\boldsymbol{r} \in \mathcal{R}_{S2}$. Moreover, similar to those for the case with $\boldsymbol{r} \in \mathcal{R}_2$, we have omitted the boundary cases in which one or more inequalities in the condition of an LGR-path hold with equality. In such boundary cases one or more LGRs become empty, and the simplified LGR-path is then obtained by omitting the empty LGRs.

Table 4.5: LGR paths for $\boldsymbol{r} \in \mathcal{R}_{S2}$. Table 4.3 provides the threshold powers in terms of link gains and γ. Each path originates from one of three different LGRs $\mathcal{A}_{\mathsf{r,r}}, \mathcal{A}_{\mathsf{d,d}}$ or $\mathcal{A}_{\mathsf{d,r}}$, and they terminate at the final LGR $\mathcal{S}_{\mathsf{d,rd}}$. If for certain values of link gains and γ, some LGR in the path becomes empty, the path is simplified by omitting the empty LGRs.

LGR path	**Condition**	**Interval of P in each LGR**
$[T3]: \mathcal{A}_{\mathsf{r,r}} \to \mathcal{S}_{\mathsf{rd,r}} \to \mathcal{S}_{\mathsf{rd,rd}} \to \mathcal{S}_{\mathsf{d,rd}}$	$0 < \mathsf{P}_{\mathsf{r,r}} < \mathsf{P}'_{\mathsf{d,d}} < \hat{\mathsf{P}}'_{\mathsf{d,d}}$, and $\overline{\mathsf{P}}_{\mathsf{rd,r}} \leq \mathsf{P}_{\mathsf{d,r}}$	$[0, \mathsf{P}_{\mathsf{r,r}})$, $[\mathsf{P}_{\mathsf{r,r}}, \overline{\mathsf{P}}_{\mathsf{rd,r}})$, $[\overline{\mathsf{P}}_{\mathsf{rd,r}}, \overline{\mathsf{P}}_{\mathsf{d,rd}}), [\overline{\mathsf{P}}_{\mathsf{d,rd}}, \infty)$
$[N1]: \mathcal{A}_{\mathsf{r,r}} \to \mathcal{S}_{\mathsf{rd,r}} \to \mathcal{S}_{\mathsf{d,rd}}$	$0 < \mathsf{P}_{\mathsf{r,r}} < \mathsf{P}'_{\mathsf{d,d}} < \hat{\mathsf{P}}'_{\mathsf{d,d}}$, and $\overline{\mathsf{P}}_{\mathsf{rd,r}} > \mathsf{P}_{\mathsf{d,r}}$	$[0, \mathsf{P}_{\mathsf{r,r}})$, $[\mathsf{P}_{\mathsf{r,r}}, \mathsf{P}_{\mathsf{d,r}})$, $[\mathsf{P}_{\mathsf{d,r}}, \infty)$
$[T4]: \mathcal{A}_{\mathsf{d,d}} \to \mathcal{A}_{\mathsf{d,rd}} \to \mathcal{A}_{\mathsf{rd,rd}} \to \mathcal{S}_{\mathsf{rd,rd}} \to \mathcal{S}_{\mathsf{d,rd}}$	$0 < \hat{\mathsf{P}}_{\mathsf{d,d}} < \mathsf{P}_{\mathsf{d,d}} < \mathsf{P}_{\mathsf{d,rd}}$	$[0, \hat{\mathsf{P}}_{\mathsf{d,d}})$, $[\hat{\mathsf{P}}_{\mathsf{d,d}}, \mathsf{P}_{\mathsf{d,d}})$, $[\mathsf{P}_{\mathsf{d,d}}, \mathsf{P}_{\mathsf{rd,rd}})$, $[\mathsf{P}_{\mathsf{rd,rd}}, \overline{\mathsf{P}}_{\mathsf{d,rd}})$, $[\overline{\mathsf{P}}_{\mathsf{d,rd}}, \infty)$
$[N2]: \mathcal{A}_{\mathsf{d,d}} \to \mathcal{A}_{\mathsf{d,rd}} \to \mathcal{S}_{\mathsf{d,rd}}$	$0 < \hat{\mathsf{P}}_{\mathsf{d,d}} < \mathsf{P}_{\mathsf{d,rd}} < \mathsf{P}_{\mathsf{d,d}}$	$[0, \hat{\mathsf{P}}_{\mathsf{d,d}})$, $[\hat{\mathsf{P}}_{\mathsf{d,d}}, \mathsf{P}_{\mathsf{d,rd}})$, $[\mathsf{P}_{\mathsf{d,rd}}, \infty)$
$[T5]: \mathcal{A}_{\mathsf{d,r}} \to \mathcal{A}_{\mathsf{rd,r}} \to \mathcal{A}_{\mathsf{rd,rd}} \to \mathcal{S}_{\mathsf{rd,rd}} \to \mathcal{S}_{\mathsf{d,rd}}$	$0 < \mathsf{P}_{\mathsf{d,d}} < \hat{\mathsf{P}}'_{\mathsf{d,d}} < \mathsf{P}_{\mathsf{rd,r}} < \mathsf{P}_{\mathsf{d,r}}$	$[0, \mathsf{P}_{\mathsf{d,d}})$, $[\mathsf{P}_{\mathsf{d,d}}, \hat{\mathsf{P}}'_{\mathsf{d,d}})$, $[\hat{\mathsf{P}}'_{\mathsf{d,d}}, \mathsf{P}_{\mathsf{rd,rd}})$, $[\mathsf{P}_{\mathsf{rd,rd}}, \overline{\mathsf{P}}_{\mathsf{d,rd}})$, $[\overline{\mathsf{P}}_{\mathsf{d,rd}}, \infty)$
$[T6]: \mathcal{A}_{\mathsf{d,r}} \to \mathcal{A}_{\mathsf{rd,r}} \to \mathcal{S}_{\mathsf{rd,r}} \to \mathcal{S}_{\mathsf{rd,rd}} \to \mathcal{S}_{\mathsf{d,rd}}$	$0 < \mathsf{P}_{\mathsf{d,d}} < \mathsf{P}_{\mathsf{rd,r}} < \hat{\mathsf{P}}'_{\mathsf{d,d}} < \mathsf{P}_{\mathsf{d,r}}$, or $0 < \mathsf{P}_{\mathsf{d,d}} < \overline{\mathsf{P}}_{\mathsf{rd,r}} < \mathsf{P}_{\mathsf{d,r}} < \hat{\mathsf{P}}'_{\mathsf{d,d}}$	$[0, \mathsf{P}_{\mathsf{d,d}})$, $[\mathsf{P}_{\mathsf{d,d}}, \mathsf{P}_{\mathsf{rd,r}})$, $[\mathsf{P}_{\mathsf{rd,r}}, \overline{\mathsf{P}}_{\mathsf{rd,r}})$, $[\overline{\mathsf{P}}_{\mathsf{rd,r}}, \overline{\mathsf{P}}_{\mathsf{d,rd}})$, $[\overline{\mathsf{P}}_{\mathsf{d,rd}}, \infty)$
$[T7]: \mathcal{A}_{\mathsf{d,r}} \to \mathcal{A}_{\mathsf{d,rd}} \to \mathcal{A}_{\mathsf{rd,rd}} \to \mathcal{S}_{\mathsf{rd,rd}} \to \mathcal{S}_{\mathsf{d,rd}}$	$0 < \hat{\mathsf{P}}'_{\mathsf{d,d}} < \mathsf{P}_{\mathsf{d,d}} < \mathsf{P}_{\mathsf{d,r}}$, or $0 < \hat{\mathsf{P}}'_{\mathsf{d,d}} < \mathsf{P}_{\mathsf{d,r}} < \mathsf{P}_{\mathsf{d,d}} < \mathsf{P}_{\mathsf{d,rd}}$	$[0, \hat{\mathsf{P}}'_{\mathsf{d,d}})$, $[\hat{\mathsf{P}}'_{\mathsf{d,d}}, \mathsf{P}_{\mathsf{d,d}})$, $[\mathsf{P}_{\mathsf{d,d}}, \mathsf{P}_{\mathsf{rd,rd}})$, $[\mathsf{P}_{\mathsf{rd,rd}}, \overline{\mathsf{P}}_{\mathsf{d,rd}})$, $[\overline{\mathsf{P}}_{\mathsf{d,rd}}, \infty)$
$[N3]: \mathcal{A}_{\mathsf{d,r}} \to \mathcal{A}_{\mathsf{rd,r}} \to \mathcal{S}_{\mathsf{rd,r}} \to \mathcal{S}_{\mathsf{d,rd}}$	$0 < \mathsf{P}_{\mathsf{d,d}} < \mathsf{P}_{\mathsf{d,r}} < \overline{\mathsf{P}}_{\mathsf{rd,r}} < \hat{\mathsf{P}}'_{\mathsf{d,d}}$	$[0, \mathsf{P}_{\mathsf{d,d}})$, $[\mathsf{P}_{\mathsf{d,d}}, \mathsf{P}_{\mathsf{rd,r}})$, $[\mathsf{P}_{\mathsf{rd,r}}, \mathsf{P}_{\mathsf{d,r}})$, $[\mathsf{P}_{\mathsf{d,r}}, \infty)$
$[N4]: \mathcal{A}_{\mathsf{d,r}} \to \mathcal{A}_{\mathsf{d,rd}} \to \mathcal{S}_{\mathsf{d,rd}}$	$0 < \hat{\mathsf{P}}'_{\mathsf{d,d}} < \mathsf{P}_{\mathsf{d,r}} < \mathsf{P}_{\mathsf{d,rd}} < \mathsf{P}_{\mathsf{d,d}}$	$[0, \hat{\mathsf{P}}'_{\mathsf{d,d}})$, $[\hat{\mathsf{P}}'_{\mathsf{d,d}}, \mathsf{P}_{\mathsf{d,rd}})$, $[\mathsf{P}_{\mathsf{d,rd}}, \infty)$
$[N5]: \mathcal{A}_{\mathsf{d,r}} \to \mathcal{S}_{\mathsf{d,rd}}$	$0 < \mathsf{P}_{\mathsf{d,r}} < \min(\hat{\mathsf{P}}'_{\mathsf{d,d}}, \mathsf{P}_{\mathsf{d,d}})$	$[0, \mathsf{P}_{\mathsf{d,r}})$, $[\mathsf{P}_{\mathsf{d,r}}, \infty)$

Initial LGR: While $[T3]$ and $[N1]$ originate from the initial LGR $\mathcal{A}_{\mathsf{r,r}}$, $[T4]$ and $[N2]$ originate from LGR $\mathcal{A}_{\mathsf{d,d}}$, and $[T5], \dots [N5]$ originate from LGR $\mathcal{A}_{\mathsf{d,r}}$. The initial LGRs vary depending on how d compares to r_1 and r_2 as in the case of $\boldsymbol{r} \in \mathcal{R}_2$, hence is not repeated here.

Saturation Cases: Saturation first occurs in one of LGRs $\mathcal{S}_{\mathsf{rd,rd}}$, $\mathcal{S}_{\mathsf{d,rd}}$ and $\mathcal{S}_{\mathsf{rd,r}}$.

Saturation first occurs in $\mathcal{S}_{\mathsf{rd,rd}}$ if the condition of one of the paths $[T4], [T5]$ or $[T7]$ is met. Here, $\mathsf{P}_{\mathsf{sat}} = \max(\bar{\mathsf{P}}_{\mathsf{rd,r}}, \mathsf{P}_{\mathsf{rd,rd}})$. Unlike in case $\boldsymbol{r} \in \mathcal{R}_2$, LGR $\mathcal{S}_{\mathsf{rd,rd}}$ is now active only for the finite range of P, i.e., $\max(\bar{\mathsf{P}}_{\mathsf{rd,r}}, \mathsf{P}_{\mathsf{rd,rd}}) \leq P \leq \bar{\mathsf{P}}_{\mathsf{d,rd}}$. Intuitively, when $\boldsymbol{r} \in \mathcal{R}_{S2}$, relay-link $\mathtt{RL}_2$ is significantly stronger than $\mathtt{RL}_1$ (i.e., $r_2 > \gamma r_1$). Hence, transmitting in *both* relay-links as in $\mathcal{S}_{\mathsf{rd,rd}}$ is optimal only for a *finite* range of P. In fact, for all P sufficiently large, transmitting only in the significantly stronger link $\mathtt{RL}_2$ turns out to be optimal.

Saturation first occurs in $\mathcal{S}_{\mathsf{rd,r}}$ if the condition of one of the paths $[T3], [N1], [T6]$ or $[N3]$ hold. Here, $\mathsf{P}_{\mathsf{sat}} = \max(\mathsf{P}_{\mathsf{rd,r}}, \mathsf{P}_{\mathsf{r,r}})$, and $\mathcal{S}_{\mathsf{rd,r}}$ is active for the range $\max(\mathsf{P}_{\mathsf{rd,r}}, \mathsf{P}_{\mathsf{r,r}}) \leq P \leq \min(\bar{\mathsf{P}}_{\mathsf{rd,r}}, \mathsf{P}_{\mathsf{d,r}})$.

Finally, saturation first occurs in $\mathcal{S}_{\mathsf{d,rd}}$ when the condition of one of the paths $[N2], [N4]$ or $[N5]$ hold. Here, for all $P \geq \mathsf{P}_{\mathsf{sat}} = \max(\mathsf{P}_{\mathsf{d,r}}, \mathsf{P}_{\mathsf{d,rd}}, \bar{\mathsf{P}}_{\mathsf{d,rd}})$, LGR $\mathcal{S}_{\mathsf{d,rd}}$ is active. Note that in $\mathcal{S}_{\mathsf{d,rd}}$, as P increases, $q_2 = \frac{\gamma - 1}{r_2} > 0$ and $q_1 = 0$ are fixed, and all additional increments of P are allotted to the direct-links only. Intuitively, since $\mathtt{RL}_2$ is significantly stronger than $\mathtt{RL}_1$, for all $P \geq \mathsf{P}_{\mathsf{sat}}$, the best rate is achieved by transmitting only in $\mathtt{RL}_2$.

Final LGR: For $P \geq \mathsf{P}_{\mathsf{fin}} = \max(\bar{\mathsf{P}}_{\mathsf{d,rd}}, \mathsf{P}_{\mathsf{d,rd}}, \mathsf{P}_{\mathsf{d,r}})$, all paths terminate at final LGR $\mathcal{S}_{\mathsf{d,rd}}$.

LGR-paths: Since paths $[T3], \dots, [T7]$ can be interpreted similarly to paths $[S3], \dots, [S7]$, they are not detailed here. Hence, we only discuss paths $[N1], \dots, [N5]$ briefly.

Path $[N1]$ is similar to $[T3]$ with $\mathcal{S}_{\mathsf{rd,rd}}$ skipped. In contrast to $[T3]$, we now have $\mathsf{P}_{\mathsf{d,r}} < \bar{\mathsf{P}}_{\mathsf{rd,r}}$, and hence the allocation of $\mathcal{S}_{\mathsf{rd,rd}}$ is suboptimal. Therefore, as P increases and $P > \mathsf{P}_{\mathsf{d,r}}$, LGR $\mathcal{S}_{\mathsf{rd,rd}}$ is skipped, and the MARC-OPA allocates the link powers according to $\mathcal{S}_{\mathsf{d,rd}}$ for all $P \geq \mathsf{P}_{\mathsf{d,r}}$.

Path $[N2]$ is similar to $[T4]$ with $\mathcal{A}_{\mathsf{rd,rd}}$ and $\mathcal{S}_{\mathsf{rd,rd}}$ skipped. The conditions for $[N2]$ simplifies to $r_2 \geq r_1(2\gamma - 1)$. Due to this, $\mathtt{RL}_2$ is so much stronger than $\mathtt{RL}_1$ that, for all $P \geq 0$, the best rate is achieved by transmitting solely in $\mathtt{RL}_2$ and not transmitting in $\mathtt{RL}_1$

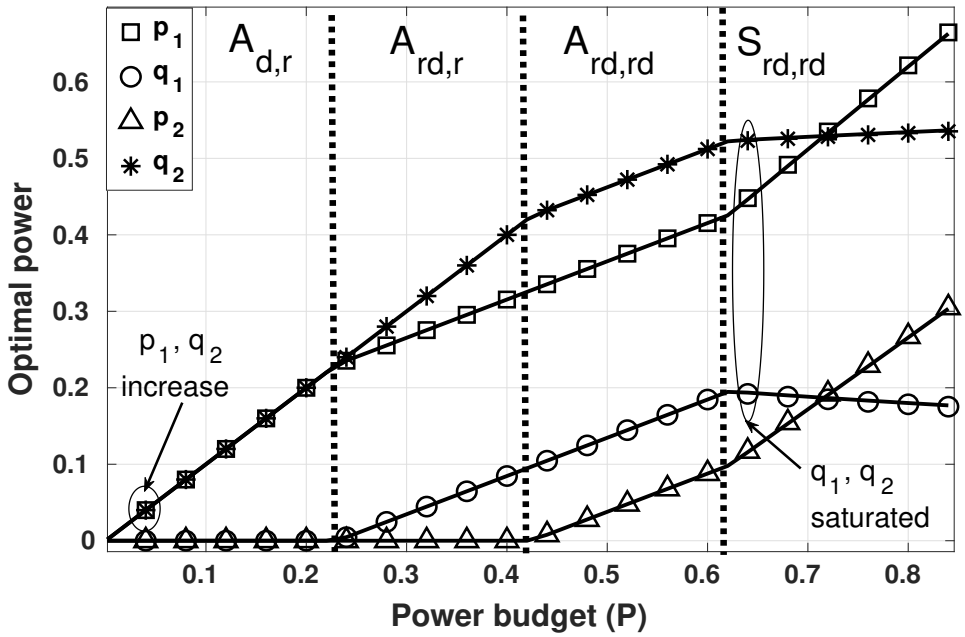


Figure 4.5: Path $[S5]$ with $(r_1, r_2, d, \gamma) = (1, 2.9, 1.3, 3)$. For $P < \mathsf{P}_{\mathsf{sat}} = 0.62$, all link powers follow the WF-like property. At $P = \mathsf{P}_{\mathsf{sat}}$, saturation occurs in LGR $\mathcal{S}_{\mathsf{rd,rd}}$ and it remains active for all $P \geq \mathsf{P}_{\mathsf{sat}}$.

at all. Thus, compared to $[T4]$ where non-zero power is allocated to $\mathtt{RL}_1$ in LGRs $\mathcal{A}_{\mathsf{rd,rd}}$ and $\mathcal{S}_{\mathsf{rd,rd}}$, these LGRs are now suboptimal and need to be skipped here.

Likewise, $[N3]$ is similar to $[T6]$ with $\mathcal{S}_{\mathsf{rd,rd}}$ skipped, $[N4]$ to $[T7]$ with $\mathcal{A}_{\mathsf{rd,rd}}$ and $\mathcal{S}_{\mathsf{rd,rd}}$ skipped, and $[N5]$ to $[N4]$ with $\mathcal{A}_{\mathsf{d,rd}}$ skipped. The conditions for these paths can be interpreted in a manner similar to those discussed above, and hence are not detailed.

Numerical Examples: We now illustrate examples of paths $[S5]$ and $[T5]$ in Figure 4.5 and Figure 4.6 respectively by plotting the optimal link powers against budget P for parameters (r_1, r_2, d, γ) as noted in the respective figures. In each example, the analytical expression of powers (marker-line) indeed match their numerically computed counterparts (solid line) using the `CVX` package [133]. We also verify that the MARC-OPA follows the respective paths by labeling the active LGRs in the relevant intervals.

In Figure 4.5, we verify path $[S5]$ where $\mathsf{P}_{\mathsf{sat}} = \mathsf{P}_{\mathsf{fin}} = \mathsf{P}_{\mathsf{rd,dr}} = 0.62$. Here, LGR $\mathcal{A}_{\mathsf{d,r}}$ is first active for $0 \leq P < \mathsf{P}_{\mathsf{d,d}}$, where $p_1 = q_2 = P$, while $q_1 = p_2 = 0$. Then, for $\mathsf{P}_{\mathsf{d,d}} \leq P < \hat{\mathsf{P}}_{\mathsf{d,d}}$, LGR $\mathcal{A}_{\mathsf{rd,r}}$ becomes active where, in addition to p_1 and q_2, q_1 increases with P as well. As P increases, for $\hat{\mathsf{P}}'_{\mathsf{d,d}} \leq P < \mathsf{P}_{\mathsf{sat}}$, LGR $\mathcal{A}_{\mathsf{rd,rd}}$ is active where all 4 powers

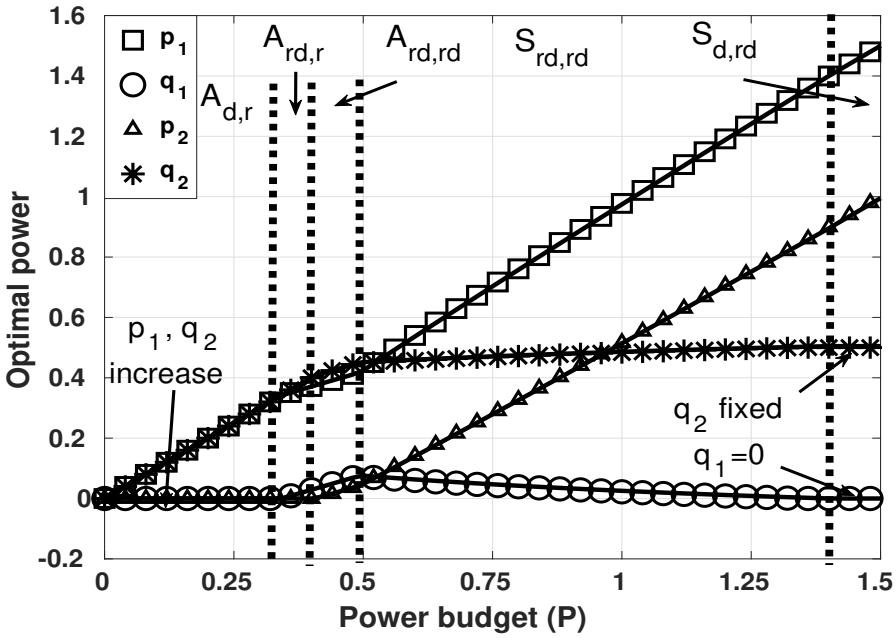


Figure 4.6: Path $[T5]$ with $(r_1, r_2, d, \gamma) = (1, 4, 1.52, 3)$. For $P < \mathsf{P}_{\mathsf{sat}} = 0.49$, all link powers follow the WF-like property. Saturation first occurs at $P = \mathsf{P}_{\mathsf{sat}}$ and LGR $\mathcal{S}_{\mathsf{rd,rd}}$ becomes active. For $P \geq \mathsf{P}_{\mathsf{fin}} = 1.34$, LGR $\mathcal{S}_{\mathsf{d,rd}}$ is active where $q_2 = 1/2$ and $q_1 = 0$ remain fixed.

increase with P. Finally, for $P \geq \mathsf{P}_{\mathsf{sat}}$, saturation occurs in $\mathcal{S}_{\mathsf{rd,rd}}$ where q_2 increases and q_1 decreases towards limits $\bar{q}_2 = 0.67$ and $\bar{q}_1 = 0.02$ (not shown in Figure 4.5), while p_1 and p_2 grow unbounded with P.

We similarly verify $[T5]$ in Figure 4.6 and omit the details since in $[T5]$, the first 4 LGRs are the same as those of $[S5]$ in Figure 4.5. Nevertheless, for $[T5]$ while saturation occurs at $\mathsf{P}_{\mathsf{sat}} = \mathsf{P}_{\mathsf{rd,rd}} = 0.49$ in $\mathcal{S}_{\mathsf{rd,rd}}$, unlike in $[S5]$, the final LGR is $\mathcal{S}_{\mathsf{d,rd}}$ where $q_2 = 0.5, q_1 = 0$ are fixed for all $P \geq \mathsf{P}_{\mathsf{fin}} = 1.34$.

4.4.5 Optimal Power Allocation for the Symmetric Case

For the symmetric case with $d = d_1 = d_2$ and $r = r_1 = r_2$, considering a symmetric power allocation of the form (p, q, p, q) is sufficient. The optimal power allocation scheme (the MARC-OPA) for this case is presented in Table 4.6. Due to symmetry, only 3 LGRs $\mathcal{A}_{\mathsf{d,d}}, \mathcal{A}_{\mathsf{r,r}}$ and $\mathcal{A}_{\mathsf{rd,rd}}$ are sufficient to describe the Waterfilling-like property, while only one LGR $\mathcal{S}_{\mathsf{rd,rd}}$ is needed to describe the saturation property.

Moreover, in this case only 3 LGR-paths, $[\tilde{S}1]$, $[\tilde{S}2]$ and $[\tilde{S}3]$, are valid which are presented in Table 4.7 along with the underlying conditions on channel gains and the

Table 4.6: For the symmetric case with $\gamma > 1$ definition of LGRs and optimal link powers are given in terms of channel parameters $\boldsymbol{c} = (r, d, \gamma, P)$, where γ is defined in (4.39).

Definition of LGR	**Optimal Power Allocation**	
$\mathcal{A}_{\mathsf{d},\mathsf{d}} := \{\boldsymbol{c} : 0 \le P \le \frac{1}{r} - \frac{1}{d}\}$	$p = P,$	$q = 0$
$\mathcal{A}_{\mathsf{r},\mathsf{r}} := \{\boldsymbol{c} : 0 \le P \le \min\left(\frac{1}{d} - \frac{1}{r}, \frac{\gamma^{1/2}-1}{r}\right)\}$	$p = 0,$	$q = P$
$\mathcal{A}_{\mathsf{rd},\mathsf{rd}} := \{\boldsymbol{c} : \max\left(\frac{1}{r} - \frac{1}{d}, \frac{1}{d} - \frac{1}{r}\right) < P < \frac{2\gamma^{1/2}-1}{r} - \frac{1}{d}\}$	$p = \frac{1}{2}\left(P + \frac{1}{r} - \frac{1}{d}\right),$	$q = \frac{1}{2}\left(P - \frac{1}{r} + \frac{1}{d}\right)$
$\mathcal{S}_{\mathsf{rd},\mathsf{rd}} := \{\boldsymbol{c} : \max\left(\frac{\gamma^{1/2}-1}{r}, \frac{2\gamma^{1/2}-1}{r} - \frac{1}{d}\right) \le P$	$p = P - \frac{\gamma^{1/2}-1}{r},$	$q = \frac{\gamma^{1/2}-1}{r}$

intervals of the LGRs. In particular, we observe the following

- if $d \ge r$ (i.e., direct-links are *stronger* than relay-links), path $[\tilde{S}3]$ is followed. Since direct-links are stronger, for $P \in [0, \frac{1}{r} - \frac{1}{d})$, the MARC-OPA transmits only in the direct-links as in $\mathcal{A}_{\mathsf{d},\mathsf{d}}$. Then, for $P \in [\frac{1}{r} - \frac{1}{d}, \frac{2\gamma^{1/2}-1}{r} - \frac{1}{d})$, all 4 links are allocated power in a WF-fashion as in $\mathcal{A}_{\mathsf{rd},\mathsf{rd}}$. Finally, for all $P \ge \mathsf{P}_{\mathsf{sat}} = \frac{2\gamma^{1/2}-1}{r} - \frac{1}{d}$, saturation occurs in $\mathcal{S}_{\mathsf{rd},\mathsf{rd}}$, where $q = \frac{\gamma^{1/2}-1}{r}$ remains fixed.

- if $d < r \le d\gamma^{1/2}$ (i.e., relay-links are *stronger but not significantly stronger* than direct-links), path $[\tilde{S}1]$ is followed. Since relay-links are stronger, for $P \in [0, \frac{1}{d} - \frac{1}{r})$, the MARC-OPA now transmits only in the relay-links as in $\mathcal{A}_{\mathsf{r},\mathsf{r}}$. Then, as P increases, LGRs $\mathcal{A}_{\mathsf{rd},\mathsf{rd}}$ and $\mathcal{S}_{\mathsf{rd},\mathsf{rd}}$ are followed as above for path $[\tilde{S}3]$.

- if $r > d\gamma^{1/2}$ (relay-links are *significantly* stronger), path $[\tilde{S}2]$ is followed. When $P \in [0, \frac{\gamma^{1/2}-1}{r})$, the MARC-OPA transmits only in relay-links as in $\mathcal{A}_{\mathsf{r},\mathsf{r}}$ until they saturate, and then for $P \ge \mathsf{P}_{\mathsf{sat}} = \frac{\gamma^{1/2}-1}{r}$, the relay-links saturate and $\mathcal{S}_{\mathsf{rd},\mathsf{rd}}$ becomes active.

The Optimal Power Allocation in the Large Millimeter-Wave Bandwidth Regime

It is interesting to study how the optimal power allocation for the symmetric case behaves as the mm-wave bandwidth becomes large. Recall that the direct- and relay-links in the

Table 4.7: LGR paths for the symmetric case. Each path originates from one of two initial LGRs $\mathcal{A}_{\mathsf{r,r}}$ or $\mathcal{A}_{\mathsf{d,d}}$, and terminate at the final LGR $\mathcal{S}_{\mathsf{rd,rd}}$.

LGR path	**Condition**	**Interval of P in each LGR**
$[\tilde{S}1] : \mathcal{A}_{\mathsf{r,r}} \to \mathcal{A}_{\mathsf{rd,rd}} \to \mathcal{S}_{\mathsf{rd,rd}}$	$d < r \leq d\gamma^{1/2}$	$[0, \frac{1}{d} - \frac{1}{r})$, $[\frac{1}{d} - \frac{1}{r}, \frac{2\gamma^{1/2}-1}{r} - \frac{1}{d})$, $[\frac{2\gamma^{1/2}-1}{r} - \frac{1}{d}, \infty)$
$[\tilde{S}2] : \mathcal{A}_{\mathsf{r,r}} \to \mathcal{S}_{\mathsf{rd,rd}}$	$r > d\gamma^{1/2}$	$[0, \frac{\gamma^{1/2}-1}{r})$, $[\frac{\gamma^{1/2}-1}{r}, \infty)$
$[\tilde{S}3] : \mathcal{A}_{\mathsf{d,d}} \to \mathcal{A}_{\mathsf{rd,rd}} \to \mathcal{S}_{\mathsf{rd,rd}}$	$d \geq r$	$[0, \frac{1}{r} - \frac{1}{d})$, $[\frac{1}{r} - \frac{1}{d}, \frac{2\gamma^{1/2}-1}{r} - \frac{1}{d})$, $[\frac{2\gamma^{1/2}-1}{r} - \frac{1}{d}, \infty)$

DR-MARC are modeled as point-to-point AWGN channels with a bandwidth-mismatch factor (BMF) of α. For example, consider the S_1-D_1 direct-link with transmit power p, link gain $d_0^{1/2}$, and noise variance N, for which the achievable rate is given by $\alpha \log\left(1 + \frac{d_0 p}{N}\right)$ as in (4.30). In the previous settings, since the mm-wave bandwidth (BW) was considered to be a fixed constant, the BMF α and the noise variance N were also taken to be constant, e.g., N was taken to be $N = 1$ in (4.7).

When the mm-wave BW increases, the noise variance N as well as the BMF α increases proportionally, which can be modeled by defining $N := N_0 w$ and $\alpha := \alpha_0 w$ for some $N_0, \alpha_0, w > 0$. Then, the large BW regime is modeled by having $w \to \infty$. For the S_1-D_1 direct-link described above, the achievable rate is now given by

$$\alpha_0 w \log\left(1 + \frac{d_0 p}{w N_0}\right) = \alpha_0 w \log\left(1 + d\frac{p}{N_0}\right), \tag{4.43}$$

with the identification $\alpha = \alpha_0 w$ and $d = \frac{d_0}{w}$, which capture the effect of varying BW. For the other direct-link, two relay-links, and the relay-destination link in the mm-wave band after similarly capturing the effect of varying BW, the resulting link gains and BMF are given by

$$\alpha = \alpha_0 w, \quad d = \frac{d_0}{w}, \quad r = \frac{r_0}{w}, \quad \bar{G}_{\mathsf{RD}} = \frac{\bar{G}_{\mathsf{RD},0}}{w}. \tag{4.44}$$

We now examine the effect of large BW, i.e., $w \to \infty$, on the optimal power allocation of Table 4.6. First, from the definition of σ_{D} and σ_{R} in (4.32), parameter $\gamma^{1/2}$ in (4.39) simplifies as follows

$$\gamma^{1/2} = 2^{\frac{(\sigma_{\mathsf{D}} - \sigma_{\mathsf{R}})}{2\alpha}}$$

$$
\begin{aligned}
&= 2^{(\mathbb{E}[\tilde{\mathsf{C}}(G_{1\mathsf{D}}P_1+G_{2\mathsf{D}}P_2+G_{\mathsf{RD}}P_{\mathsf{R}})]-\mathbb{E}[\tilde{\mathsf{C}}(G_{1\mathsf{R}}P_1+G_{2\mathsf{R}}P_2)])/2\alpha} \quad \cdot \quad 2^{(\log(1+\bar{G}_{\mathsf{RD}}\bar{P}_{\mathsf{R}}))/2} \\
&\overset{(a)}{=} 2^{\delta/\alpha} \quad \left(1+\bar{G}_{\mathsf{RD}}\bar{P}_{\mathsf{R}}\right)^{1/2} \\
&\overset{(b)}{=} 2^{\delta/\alpha_0 w} \quad \left(1+\frac{\bar{G}_{\mathsf{RD},0}}{w}\bar{P}_{\mathsf{R}}\right)^{1/2} \\
&\overset{(c)}{=} e^{\frac{\delta}{\alpha_0 w}\log_e 2} \quad \left(1+\frac{\bar{G}_{\mathsf{RD},0}}{w}\bar{P}_{\mathsf{R}}\right)^{1/2} \\
&\overset{(d)}{\approx} \left(1+\frac{\delta}{\alpha_0 w}\log_e 2\right) \quad \left(1+\frac{\bar{G}_{\mathsf{RD},0}}{2w}\bar{P}_{\mathsf{R}}\right) \\
&\overset{(e)}{\approx} 1+\frac{1}{w}\Gamma+o(1/w)
\end{aligned} \tag{4.45}
$$

where (a) follows from defining $\delta := 1/2(\mathbb{E}[\tilde{\mathsf{C}}(G_{1\mathsf{D}}P_1 + G_{2\mathsf{D}}P_2 + G_{\mathsf{RD}}P_{\mathsf{R}})] - \mathbb{E}[\tilde{\mathsf{C}}(G_{1\mathsf{R}}P_1 + G_{2\mathsf{R}}P_2)])$, (b) follows from expressing BMF α and link gains as in (4.44), (c) follows from expressing $2^{\delta/\alpha_0 w}$ as $e^{\frac{\delta}{\alpha_0 w}\log_e 2}$, (d) follows from $e^{\frac{\delta}{\alpha_0 w}\log_e 2} \approx 1+\frac{\delta}{\alpha_0 w}\log_e 2$, and $\left(1+\frac{\bar{G}_{\mathsf{RD},0}}{w}\bar{P}_{\mathsf{R}}\right)^{1/2} \approx \left(1+\frac{\bar{G}_{\mathsf{RD},0}}{2w}\bar{P}_{\mathsf{R}}\right)$, for large w, and finally in (e), the quadratic term is presented as $o(1/w)$ such that $\frac{o(1/w)}{1/w} \to 0$, as $w \to \infty$, and $\Gamma = \left(\frac{\delta \log_e 2}{\alpha_0} + \frac{\bar{G}_{\mathsf{RD},0}}{2}\bar{P}_{\mathsf{R}}\right)$.

As a result, accounting for the link gains in (4.44) and the parameter $\gamma^{1/2}$ in (4.45), for large w, i.e., $w \to \infty$, the threshold power $\frac{\gamma^{1/2}-1}{r}$ in Table 4.6 simplifies to

$$
\begin{aligned}
\frac{\gamma^{1/2}-1}{r} &\approx \frac{1+\Gamma/w-1}{r_0/w} \\
&= \frac{\Gamma}{r_0}.
\end{aligned} \tag{4.46}
$$

Similarly, the other threshold powers in Table 4.6 simplify to

$$
\begin{aligned}
\frac{1}{r}-\frac{1}{d} &= w\left(\frac{1}{r_0}-\frac{1}{d_0}\right), \\
\frac{1}{d}-\frac{1}{r} &= w\left(\frac{1}{d_0}-\frac{1}{r_0}\right), \\
\frac{2\gamma^{1/2}-1}{r}-\frac{1}{d} &\approx w\left(\frac{1}{r_0}-\frac{1}{d_0}\right)+\frac{2\Gamma}{r_0}.
\end{aligned}
$$

As presented in Table 4.6, the optimal power allocation in the symmetric case is described by 4 LGRs, $\mathcal{A}_{\mathsf{d,d}}, \mathcal{A}_{\mathsf{r,r}}, \mathcal{A}_{\mathsf{rd,rd}}$, and $\mathcal{S}_{\mathsf{rd,rd}}$. Simple algebraic manipulations reveal that as $w \to \infty$, LGR $\mathcal{A}_{\mathsf{rd,rd}}$ approaches the empty set. This leaves only 3 LGRs, and thus the optimal power allocation for large BW, where $w \to \infty$, simplifies as follows:

- if $d_0 \geq r_0$, the power budget P is entirely allocated to the direct-links for all $P \geq 0$.
- if $d_0 < r_0$, when $P \leq \frac{\gamma^{1/2}-1}{r} \approx \frac{\Gamma}{r_0}$, the budget P is entirely allocated to the relay-

links; instead, when $P > \frac{\Gamma}{r_0}$, the relay-links are saturated with power $q \approx \frac{\Gamma}{r_0}$, and all additional increments of P are allotted to the direct-links only.

In the large BW regime, allocating power in *both* direct and relay-links in Waterfilling-like fashion as in LGR $\mathcal{A}_{\mathsf{rd,rd}}$ is *suboptimal*. In this regime, when transmitting in either the set of direct-links or the set of relay-links, due to $w \to \infty$, the marginal return from allocating additional power to the same set of links remains *constant*, as opposed to reducing as in the regime with finite w. Thus, as P increases, it is beneficial to continue transmitting in the same set of links, and it is suboptimal to allocate a fraction P to the other type of link as in $\mathcal{A}_{\mathsf{rd,rd}}$. For example, when $d_0 \geq r_0$, i.e., when budget P is allocated entirely to the direct-links, and thus $p = P$, the achievable rate in the S_1-D_1 direct-link in (4.43) simplifies to the following, as $w \to \infty$

$$\frac{\alpha_0 w}{2} \log\left(1 + \frac{d_0 P}{w N_0}\right) \;\to\; \frac{\alpha_0 w}{2} \frac{d_0 P \log_e 2}{w N_0} \;=\; \frac{\alpha_0 d_0 \log_e 2}{2 N_0}\; P. \tag{4.47}$$

Therefore, the marginal return from allocating additional increments of P to the direct-links is constant, and thus for all $P \geq 0$, it is optimal to transmit in the direct-links only as in $\mathcal{A}_{\mathsf{d,d}}$. Alternatively, when $d_0 < r_0$, the achievable rates for the relay-links behave similarly as $w \to \infty$, and thus transmitting only in the relay-links is optimal for P smaller than the saturation threshold. However, for P larger than the threshold, the relay-links become saturated, and all additional increments of P are alloted to the direct-links, as in $\mathcal{S}_{\mathsf{rd,rd}}$.

Optimum power allocation in a 2-D topology

We now illustrate how the mode of optimal link powers varies as the source locations vary. For this, we consider the 2-D topology of Figure 4.3a, where R and D are located on the x-axis at $(0,0)$ and $(0, \mathsf{d}_{\mathsf{RD}})$, while the sources are located at $(-\mathsf{d}_{\mathsf{SR}} \cos\phi, \pm \mathsf{d}_{\mathsf{SR}} \sin\phi)$ with ϕ being the angle between the sources and the relay. Due to symmetric source placement, the resulting link gains are symmetric, i.e., $d = d_1 = d_2$ and $r = r_1 = r_2$, which simplifies the power allocation. Moreover, similar to the Section 4.2.5, the microwave band is assumed to be under phase fading, whereas the mm-wave fading gains are taken to be constant. More specifically, from node s to t, the microwave fading gains are $G_{st} = 1/\mathsf{d}_{\mathsf{st}}^{\beta_1}$, while the mm-wave relay-link gains are $r = 1/\mathsf{d}_{\mathsf{SR}}^{\beta_2}$ and the mm-wave direct-link gains are $d = 1/\mathsf{d}_{\mathsf{SD}}^{\beta_2}$, where $\mathsf{d}_{(.,.)}$ is the inter-node distance and β_1, β_2 are path-loss coefficients in the two bands.

For illustration, we take the following parameters $P_k = 10, k \in \{1, 2, \mathsf{R}\}, \bar{P}_{\mathsf{R}} = 1$,

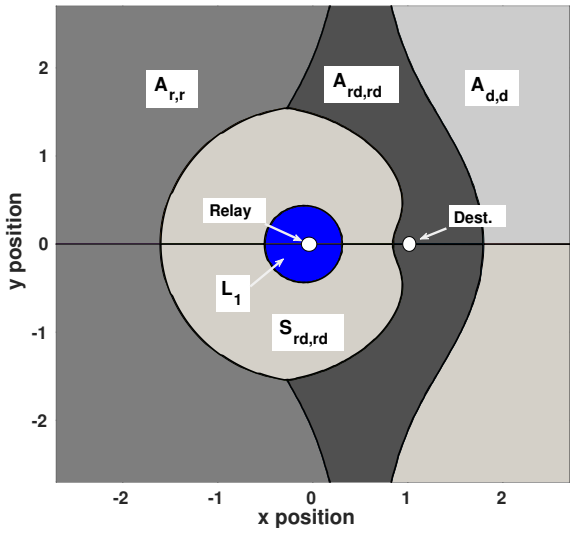


Figure 4.7: For the 2-D network topology of the DR-MARC in Figure 4.3a, the source locations are at coordinates $(x, \pm y)$. The source locations are partitioned into several regions, for each of which the optimal transmission mode for the mm-wave links are labeled.

$\beta_1 = 2, \beta_2 = 4$, $\alpha = 2$, while the power budget is $P = 10$. We then plot the source locations in Figure 4.7 by varying $\phi \in (0, \pi)$ and $\mathsf{d}_{\mathsf{SR}} \in (0, 4)$ for fixed $\mathsf{d}_{\mathsf{RD}} = 1$ unit, and partition this space based on which mode of mm-wave transmission is optimal. First, in region L_1, sources are much closer to the relay than the destination in that $\sigma_{\mathsf{R}} \geq \sigma_{\mathsf{D}}$ (i.e., $\gamma \leq 1$), with σ_{R}, σ_{D} and γ defined in (4.30), (4.31) and (4.39). Therefore, for sources located in L_1, it is optimal to transmit only in the direct-links for all $P \geq 0$.

All regions except L_1, correspond to the case of $\gamma > 1$, and depending on the budget P and source locations (i.e., the resulting direct-link and relay-link gains), the optimal transmission mode in different regions vary. For example, the sources in the region labeled $\mathcal{A}_{\mathsf{r,r}}$ are not as close to the relay as in L_1 but are sufficiently close to the relay such that $0 < P \leq \mathsf{d}_{\mathsf{SD}}^{\beta_2} - \mathsf{d}_{\mathsf{SR}}^{\beta_2}$ holds. Hence, for these source locations, allocating the budget P entirely to the relay-links is optimal. On the other hand, the sources in the region labeled $\mathcal{A}_{\mathsf{d,d}}$ are sufficiently close to the destination in that $0 < P \leq \mathsf{d}_{\mathsf{SR}}^{\beta_2} - \mathsf{d}_{\mathsf{SD}}^{\beta_2}$ holds. Hence, it is optimal to allocate the budget P entirely to the direct-links. As opposed to these two regions, the sources in the region labeled $\mathcal{A}_{\mathsf{rd,rd}}$ are at an intermediate distance from the relay and the destination in that $P < (2\gamma^{1/2} - 1)\mathsf{d}_{\mathsf{SR}}^{\beta_2} - \mathsf{d}_{\mathsf{SD}}^{\beta_2}$ holds. Here, transmitting in all 4 links as in $\mathcal{A}_{\mathsf{rd,rd}}$ is optimal. Finally, sources in the region $\mathcal{S}_{\mathsf{rd,rd}}$ are such that $P \geq (2\gamma^{1/2} - 1)\mathsf{d}_{\mathsf{SR}}^{\beta_2} - \mathsf{d}_{\mathsf{SD}}^{\beta_2}$ hold. Here, saturation occurs, and allocating power as in $\mathcal{S}_{\mathsf{rd,rd}}$ is optimal. Clearly, for fixed d_{SR}, d_{SD} and γ, as P increases, the region labeled $\mathcal{S}_{\mathsf{rd,rd}}$ grows.

4.5 Summary

In this chapter, we studied the performance of the fading MARC over the integrated mm-wave/microwave dual-band, where the mm-wave links to the relay and to the destination are modeled as non-interfering AWGN links. We showed that the capacity of the DR-MARC can be *decomposed* into the capacity of the underlying R-MARC and the two mm-wave direct-links, hence the direct-links can be operated independently of the R-MARC without compromising optimal rates. Next, we characterized an *achievable region* for the R-MARC, which is then shown to achieve the *capacity region* of the R-MARC under a set of sufficient channel conditions, denote the *jointly-near* conditions. This shows that even when the sources are not near in the underlying microwave c-MARC, for *sufficiently strong* source-relay mm-wave links, they become *jointly near* over both bands such that capacity is characterized in closed form.

Next, we studied the *optimal power allocation* over the mm-wave links that maximizes the achievable sum-rate of the DR-MARC for the case where the fading gains of the mm-wave links are geometry dependent constants. The resulting scheme allocates power in *different modes* depending on the power budget P and the link gains (i.e., according to the active link gain regime), and all such modes were characterized. When the budget P is sufficiently small, it is entirely allocated only to the strongest of the relay- and direct-links, and as P increases but remains below the saturation threshold, power is allocated to other links as in *Waterfilling* solution. However, for P above the so-called *saturation* threshold, if one relay-link is stronger but not significantly stronger than the other, power in the two links respectively increases and decreases with P and approach non-zero levels as $P \to \infty$. Otherwise, power in the *significantly* stronger relay-link is fixed at a constant while that in the other is zero.

Moreover, in the *large* mm-wave bandwidth regime, if the direct-links are stronger than relay-links, budget P should be allocated entirely to the direct-links for all $P \geq 0$. Alternatively, if the relay-links are stronger, P should be allocated entirely to the relay-links until they saturate, and then all subsequent increments of P should be allocated to the direct-links only.

These results illustrate the benefits of relay-cooperation and the impact of point-to-point mm-wave links on the performance of the dual-band MARC, and thus they can be useful in practical resource allocation in dual-band uplink scenarios.

Chapter 5

The Multiple Access Interference Channel

Inter-user interference, caused by transmission of other users in a shared medium, is a common source of performance bottleneck in wireless communication. In cellular communication settings, such interference is typically caused by base stations and users in neighboring cells that operate at the same carrier frequency. Since interference is a common impairment in wireless networks, means to mitigate interference has been widely studied in the literature [10, 12, 100, 139, 140].

In this chapter[1], we study the multiple-access interference channel (MAIC) operating solely in the microwave band. The MAIC models the interference-limited communication between a pair of two-user multiple-access channels that operate in two neighboring cells over the same shared band, and thus mutually interfere. In typical cellular communication, multiple transmitters communicate to a single receiver simultaneously. For example, in the cellular uplink, multiple users transmit to the base station and form a multiple-access channel (MAC). Hence, the MAIC models communication when two such MACs mutually interfere, and thus provides a more realistic model of cellular interference.

The performance of the MAIC is limited due to the presence of noise as well as interference. While increasing the transmit power of all users helps to safeguard transmissions against noise, the power in the resulting inter-user interference also increases proportionally. This increased interference can then constitute a possible major bottleneck to the performance. Hence, understanding how interference should be handled by, e.g., decoding

[1]The results of this chapter (except for Section 5.4) have been published in [140].

it, treating it as noise or some other scheme, is immensely helpful in characterizing useful strategies for the MAIC.

Similar to the two-user interference channel (IC), the capacity of the Gaussian MAIC (GMAIC) is known only in a few specific regimes [12,113,141]. For example, if the interference is strong, jointly decoding the two interfering messages along with the two intended messages as in a four-user Gaussian MAC proves optimal for certain subsets of the strong interference regime [12].

In contrast, the MAIC with *weak interference* models a more practical setting of cellular interference such as the following: when the users in a cell interferes with the base station of another cell, since the interfering users are located at a farther distance than the desired users, the resulting power in the interference is *typically smaller* than the power in the desired signals. While the weak MAIC effectively models such settings, neither jointly decoding interference nor treating interference as noise (TIN) is optimal in this regime.

Since characterizing the capacity region of the MAIC is highly non-trivial in general, recent studies have focused on characterizing the sum-rate performance of the GMAIC. In [12, eq.(13)], a sum-rate upper bound was characterized for the GMAIC which performs well for very small values of the cross-channel gains. However, numerical studies show that as the cross-channel gains increases within the weak interference regime, the gap between the upper bound and the TIN achievable sum-rate increase consistently and becomes large. Moreover, the sum-rate of the GMAIC has also been studied in [113]. However, as discussed in Section 2.2.2, the channel model, the exact definition of weak interference, and the upper bounding approach are all different from those in [12]. While a constant gap result for the sum-rate has been proposed in the study of [113], this gap has only been shown to exist, and neither the exact numerical value of the gap, nor the exact set of channel gains for which this gap holds, has been characterized.

In this chapter, we study the performance limits of the GMAIC, and focus primarily on the *weak* MAIC which is of more practical interest. Note that the usual terminologies of the MAIC such as the channel model and the definition of weak and strong interference regimes in [12] can be regarded as natural extensions of the same for the IC [100], and hence the notations for this chapter are adopted from [12].

First, we characterize an achievable region for the MAIC based on the Han-Kobayashi encoding (HK) scheme for the IC [7]. We then focus on the sum-rate, and characterize several genie-signal aided sum-rate upper bounds and an outer bound to the capacity region

of the channel.

The contributions of this chapter are summarized as follows:

- Achievable rate region for the discrete memoryless (DM) MAIC: In Section 5.2, we provide an achievable region for the DM-MAIC based on the HK strategy, where each transmitter splits its message into a private and a common part, and each receiver is allowed to decode the common part of the interfering message non-uniquely.
- Achievable rate region for the Gaussian MAIC: for the GMAIC, directly adapting the achievable scheme from the DM-MAIC results in a non-computable region due to a possible large cardinality of the time-sharing variable Q. We characterize a computable region by limiting $Q \in \mathcal{Q} = \{1, 2\}$ and adapting the operation at the transmitters depending on the state of Q.
- Sum-rate upper bounds for the *partially symmetric* GMAIC: Focusing on the partially symmetric case (to be defined in the next section), we derive three novel upper bounds on the sum-rate of the *weak* GMAIC by providing different *genie-signals* to the receivers. We also simplify the first upper bound to a closed form for the *completely symmetric* case where all cross-channel gains are the same. Numerical examples demonstrate that for a wide range of cross-channel gains, the achievable sum-rate and one of the sum-rate upper bounds differ only by a small numerical gap, thus providing good approximations to the sum-capacity in these cases.
- For the general GMAIC without any symmetry, we characterize a sum-rate upper bound for the *weak* interference case. We then characterize a set of outer bounds to the capacity region of the GMAIC that is valid for any interference regime.

The rest of the chapter is organized as follows: The system model is defined in Section 5.1. The achievable region is characterized in Section 5.2. In Section 5.3, the sum-rate upper bounds for the partially symmetric GMAIC are characterized, and numerical examples are presented. In Section 5.4, the capacity region outer bound is derived, and finally, a summary is provided in Section 5.5.

5.1 System Model

The MAIC consists of two interfering 2-user MACs as depicted in Figure 5.1. In the first MAC, sources S_1 and S_2 send independent messages M_1 and M_2 to destination D_1, while in

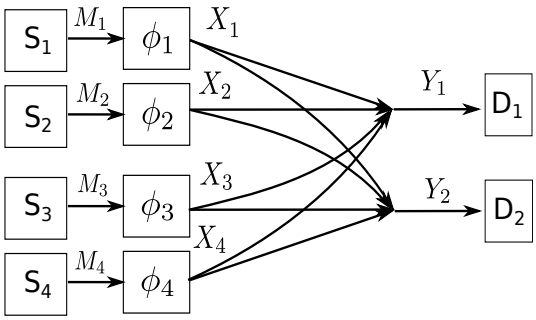


Figure 5.1: A depiction of the MAIC where the multiple-access channel from sources S_1 and S_2 to D_1 and the multiple-access channel from sources S_3 and S_4 to D_2 mutually interfere.

the second MAC, sources S_3 and S_4 communicate their independent messages M_3 and M_4 to destination D_2. However, due to the nature of the shared medium, transmissions from all sources superimpose at both destinations, and thus cause mutual interference between the two MACs.

First, we consider the discrete memoryless MAIC as depicted in Figure 5.2.

Definition 5.1 (The DM-MAIC)**.** *The DM-MAIC is defined by the tuple* $((\mathcal{X}_k)_{k=1}^4, \mathcal{Y}_1, \mathcal{Y}_2, p_{Y_1,Y_2|X_1,X_2,X_3,X_4}(y_1, y_2|x_1, x_2, x_3, x_4))$ *where* $\mathcal{X}_k, k \in \{1, 2, 3, 4\}$ *are finite, discrete input alphabets,* $\mathcal{Y}_\ell, \ell \in \{1, 2\}$ *are finite, discrete output alphabets, and the channel transition law is defined as* $p_{Y_1,Y_2|X_1,X_2,X_3,X_4}(y_1, y_2|x_1, x_2, x_3, x_4)$.

The transmission from the sources are uncoordinated, and hence the transmitted messages are assumed to be independent. Since the channel is discrete memoryless, the channel probability mass function (pmf) after n channel uses factors as

$$p(y_1^n, y_2^n|x_1^n, x_2^n, x_3^n, x_4^n) = \prod_{i=1}^{n} p(y_{1i}, y_{2i}|x_{1i}, x_{2i}, x_{3i}, x_{4i}). \tag{5.1}$$

Definition 5.2 (Code for the DM-MAIC)**.** *A* $(2^{nR_1}, 2^{nR_2}, 2^{nR_3}, 2^{nR_4}, n)$ *code for the DM-MAIC consists of (i) four uniformly distributed message sets,* $\mathcal{M}_k = \{1, \ldots, 2^{nR_k}\}$*, one for each source* $\mathsf{S}_k, k \in \{1, 2, 3, 4\}$*; (ii) an encoder for each source* $\phi_k : \mathcal{M}_k \to \mathcal{X}_k^n, k \in \{1, 2, 3, 4\}$*; and (iii) two decoders* ψ_1 *and* ψ_2 *for destinations* D_1 *and* D_2 *such that* $\psi_1 : \mathcal{Y}_1^n \to \mathcal{M}_1 \times \mathcal{M}_2$ *and* $\psi_2 : \mathcal{Y}_2^n \to \mathcal{M}_3 \times \mathcal{M}_4$.

The probability of decoding error for the DM-MAIC is defined as

$$P_e^{(n)} := \mathrm{P}[\psi_1(Y_1^n) \neq (M_1, M_2) \quad \cup \quad \psi_2(Y_2^n) \neq (M_3, M_4)],$$

where all messages are independent of each other and distributed uniformly.

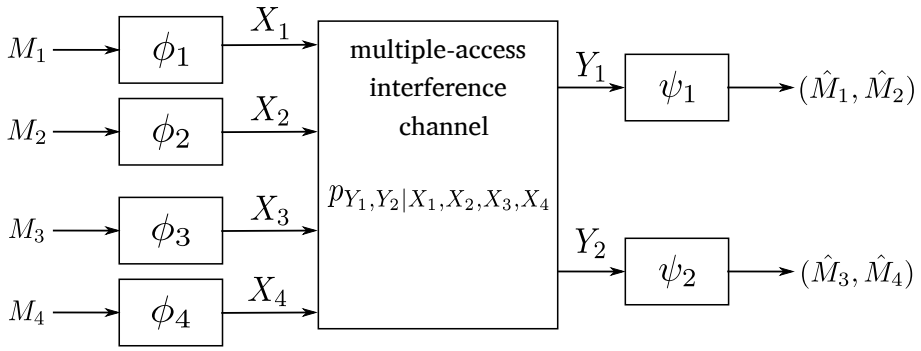


Figure 5.2: The discrete memoryless MAIC (DM-MAIC).

Definition 5.3 (Achievable rate for the DM-MAIC)**.** *A rate tuple* (R_1, R_2, R_3, R_4) *is said to be achievable for the DM-MAIC if there exists a sequence of* $(2^{nR_1}, 2^{nR_2}, 2^{nR_3}, 2^{nR_4}, n)$ *codes such that* $P_e^{(n)} \to 0$*, as* $n \to \infty$*.*

Moreover, the capacity region of the MAIC, denoted by $\mathcal{C}_{\mathsf{MAIC}}$, is defined as the closure of the set of all achievable rate tuples. Finally, the sum-capacity is defined as

$$C_{\mathsf{S}} := \max_{(R_1,R_2,R_3,R_4)\in\mathcal{C}_{\mathsf{MAIC}}} R_1 + R_2 + R_3 + R_4. \tag{5.2}$$

The Gaussian MAIC

In this work, we mostly focus on the real Gaussian MAIC (GMAIC), depicted in Figure 5.3, where the channel coefficients are real-valued, fixed and known throughout the network [12]. The channel outputs at D_1 and D_2 at the i-th channel use are given by

$$\begin{aligned} Y_{1,i} &= X_{1,i} + X_{2,i} + h_3 X_{3,i} + h_4 X_{4,i} + Z_{1,i}, \\ Y_{2,i} &= h_1 X_{1,i} + h_2 X_{2,i} + X_{3,i} + X_{4,i} + Z_{2,i}, \quad i = 1, \ldots, n \end{aligned} \tag{5.3}$$

where h_3 and h_4 are the cross-channel coefficients from S_3 and S_4 to D_1, whereas h_1 and h_2 are the cross-channel coefficients from S_1 and S_2 to D_2. The direct-channel coefficients from the sources to their designated destinations in the individual MACs, e.g., the coefficient for the channel from S_1 to D_1, are assumed to be 1. The input codewords $X_k^n \in \mathbb{R}^n$ are assumed to be block power constrained, i.e., $\frac{1}{n}\sum_{i=1}^n \mathbb{E}[X_{k,i}^2] \leq P_k, k \in \{1,2,3,4\}$. Finally, the noise samples are $Z_{1,i} \sim \mathcal{N}(0,1)$, i.i.d., and $Z_{2,i} \sim \mathcal{N}(0,1)$, i.i.d.

We focus mostly on the GMAIC with *weak* interference, where the magnitude of the cross-channel coefficients are smaller than that of the direct-channels, i.e.,

$$h_k^2 < 1, \quad k \in \{1,2,3,4\}. \tag{5.4}$$

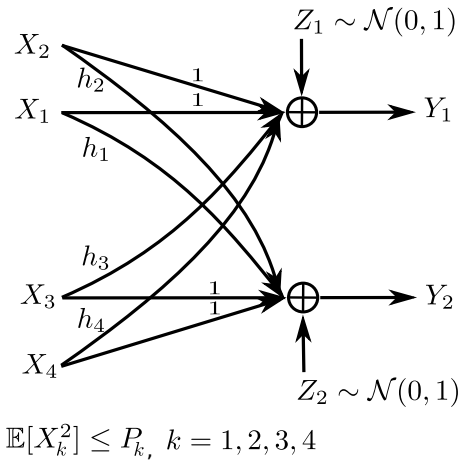


Figure 5.3: The Gaussian MAIC (GMAIC) as defined in [12].

A $(2^{nR_1}, 2^{nR_2}, 2^{nR_3}, 2^{nR_4}, n)$ code and an achievable rate pair for the GMAIC are defined from those of the DM-MAIC given in Def. 5.2 and Def. 5.3, respectively, by choosing all codeword alphabets to be $\mathbb{R}$ and imposing average power constraints on the codewords $X_k^n, k \in \{1,2,3,4\}$. Hence they are not repeated here.

Note that while the results included in this chapter are for the MAIC with *real* channel coefficients, these results can be extended to the case with *complex* channel coefficients following similar techniques.

Moreover, the MAIC can be considered a special case of the $k \times m$-user X channel [142, 143] after restricting k and m to $k = 4, m = 2$, and eliminating all messages from the first and second sources to the second destination and all messages from the third and fourth sources to the first destination. However, we have used a different achievable strategy as compared to those in [142,143] for the X channel. Moreover, the outer bounding techniques for the X channel in [142, 143] provide loose bounds for the MAIC in general, and thus we proposed novel genie-aided bounds that provide better results.

5.2 Achievable Region for the MAIC

First, we derive an achievable region for the DM-MAIC based on the HK scheme [7] for the IC. We then adapt it to the GMAIC.

5.2.1 An Achievable Region for the DM-MAIC

Let $\mathcal{P}$ be the set of joint pmfs p on the random variables $(Q, U_1, X_1, U_2, X_2, U_3, X_3, U_4, X_4)$ that factor as

$$p = p_Q(q) \prod_{k=1}^{4} p_{U_k, X_k|Q}(u_k, x_k|q). \tag{5.5}$$

We thus have the following achievable region for the DM-MAIC.

Theorem 5.1 (An achievable Region for the DM-MAIC)**.** *The region* $\mathcal{A} := \cup_{p\in\mathcal{P}} (\mathcal{R}_1(p) \cap \mathcal{R}_2(p))$ *is achievable for the DM-MAIC, where the union is over all* $p \in \mathcal{P}$ *that factors as* (5.5)*, and for a given* p*,*

$$\mathcal{R}_1(p) := \{(R_1, R_2, R_3, R_4) \succeq \mathbf{0} \mid \exists (R_{10}, R_{20}, R_{30}, R_{40}) \succeq \mathbf{0}, \tag{5.6}$$

such that

$$R_{10} < R_1, \tag{5.7}$$
$$R_{20} < R_2, \tag{5.8}$$
$$R_{30} < R_3, \tag{5.9}$$
$$R_{40} < R_4 \tag{5.10}$$
$$R_1 - R_{a0} < I(X_1; Y_1|U_a, X_2, U_3, U_4, Q) \tag{5.11}$$
$$R_2 - R_{b0} < I(X_2; Y_1|U_b, X_1, U_3, U_4, Q) \tag{5.12}$$
$$R_1 + R_2 - R_{a0} - R_{b0} < I(X_1, X_2; Y_1|U_a, U_b, U_3, U_4, Q) \tag{5.13}$$
$$R_1 - R_{a0} + R_{30} < I(X_1, U_3; Y_1|U_a, X_2, U_4, Q) \tag{5.14}$$
$$R_2 - R_{b0} + R_{30} < I(X_2, U_3; Y_1|U_b, X_1, U_4, Q) \tag{5.15}$$
$$R_1 + R_2 - R_{a0} - R_{b0} + R_{30} < I(X_1, X_2, U_3; Y_1|U_a, U_b, U_4, Q) \tag{5.16}$$
$$R_1 - R_{a0} + R_{40} < I(X_1, U_4; Y_1|U_a, X_2, U_3, Q) \tag{5.17}$$
$$R_2 - R_{b0} + R_{40} < I(X_2, U_4; Y_1|U_b, X_1, U_3, Q) \tag{5.18}$$
$$R_1 + R_2 - R_{a0} - R_{b0} + R_{40} < I(X_1, X_2, U_4; Y_1|U_a, U_b, U_3, Q) \tag{5.19}$$
$$R_1 - R_{a0} + R_{30} + R_{40} < I(X_1, U_3, U_4; Y_1|U_a, X_2, Q) \tag{5.20}$$
$$R_2 - R_{b0} + R_{30} + R_{40} < I(X_2, U_3, U_4; Y_1|U_b, X_1, Q) \tag{5.21}$$
$$R_1 + R_2 - R_{a0} - R_{b0} + R_{30} + R_{40} < I(X_1, X_2, U_3, U_4; Y_1|U_a, U_b, Q)\}, \tag{5.22}$$

for all $a \in \{\emptyset, 1\}$ *and* $b \in \{\emptyset, 2\}$*, with* $R_{\emptyset 0} := 0$ *and* $U_\emptyset := \emptyset$*, and* $I(.)$ *terms computed with the given distribution* $p \in \mathcal{P}$*. Region* $\mathcal{R}_2(p)$ *is defined by the region found by swapping indexes* $1 \leftrightarrow 3$*,* $2 \leftrightarrow 4$ *and signal* $Y_2 \leftrightarrow Y_1$ *in* $\mathcal{R}_1(p)$*.*

Proof. The proof is relegated to Appendix D.1. ■

While the proof is relegated to Appendix D.1, we briefly outline the key points here. First, the time-sharing variable Q is used to time-share between different achievable strategies. For example, each realization of Q may correspond to a different distribution of the associated random variables satisfying (5.5), each resulting in a specific achievable region, and the overall achievable region is then obtained by averaging over all such rate regions. However, irrespective of a specific instance of Q, the same encoding and decoding strategies are deployed. Hence, for brevity we omit the time sharing variable and explain the basic principles behind the encoding and decoding strategies.

The achievable region is derived using arguments similar to that for the Han-Kobayashi scheme in [7] for the IC. More specifically, here source S_k, $k \in \{1,2,3,4\}$ partitions its messages M_k (with rate R_k) into a common message $M_{k\mathsf{c}}$ (with rate R_{k0}) and a private message $M_{k\mathsf{p}}$ (with rate $R_k - R_{k0} > 0$). It then performs superposition encoding to superimpose both partitions of the message into a single codeword $X_k^n(M_{k\mathsf{c}}, M_{k\mathsf{p}})$ for transmission.

Destination D_1 then *uniquely* decodes both the private and common messages from its designated sources S_1 and S_2, i.e., $(M_{1\mathsf{c}}, M_{1\mathsf{p}}, M_{2\mathsf{c}}, M_{2\mathsf{p}})$, and *non-uniquely* decodes only the 2 common messages $(M_{3\mathsf{c}}, M_{4\mathsf{c}})$ from the non-designated sources S_3 and S_4, while treating the private messages $(M_{3\mathsf{p}}, M_{4\mathsf{p}})$ as noise. The decoding scheme follows from generalizing the scheme for the IC in [7]. The resulting rate region for destination D_1 is given by $\mathcal{R}_1(p)$ for some common message rates $0 < R_{k0} < R_k, k \in \{1,2,3,4\}$.

Note that the first 4 inequalities (5.7)-(5.10) on the common message rates are applicable to both destinations D_1 and D_2, whereas the remaining inequalities correspond to the rate constraints for non-unique decoding at destination D_1. Moreover, each set of 3 constraints, (5.11)-(5.13), (5.14)-(5.16), (5.17)-(5.19), and (5.20)-(5.22), succinctly represents a set of 8 constraints that are expressed concisely using binary variables $a \in \{\emptyset, 1\}$ and $b \in \{\emptyset, 2\}$. For example, the set of 3 constraints (5.11)-(5.13) actually represents 8 constraints when accounted for all combinations of a, b.

The decoding procedure for destination D_2 follows from that for D_1 after exchanging the roles of the two sets of sources as well as that of the two destinations. Hence, the rate constraints for D_2 can be found from that for D_1 by changing the respective variables as indicated in the theorem. Finally, the achievable region for a given distribution p satisfying (5.5) is obtained by taking the intersection $\mathcal{R}_1(p) \cap \mathcal{R}_2(p)$ for the two destinations, while

the overall achievable region is obtained by taking the union over all such distribution $\cup_{p\in\mathcal{P}} (\mathcal{R}_1(p) \cap \mathcal{R}_2(p))$.

5.2.2 Achievable Region for the Gaussian MAIC

Note that the rates in Theorem 5.1 need to be computed by taking expectation over a time-sharing variable Q, and thus many different distributions for the input codewords can potentially be considered, one for each realization of Q. Moreover, when the message-splitting operation is deployed in the Gaussian MAIC, it amounts to the total transmit power at the source being split between the private and public messages. Hence, for each realization of Q, a different power-split can be considered. Since the cardinality of Q can potentially be large, it results in a large number of choices for the encoding strategy, and hence the exact characterization of the rates in Theorem 5.1 becomes computationally hard for the Gaussian MAIC.

Therefore, we provide a computable region for the GMAIC by restricting the cardinality of Q, the input distribution and the power splitting at the sources as follows:

- only i.i.d. Gaussian codewords are considered for the inputs.
- Q is restricted to be binary, i.e., $Q \in \mathcal{Q} = \{0, 1\}$, with $\mathrm{P}[Q = 0] = \mathrm{P}[Q = 1] = \frac{1}{2}$.
- for $Q = 0$, the encoding operation is modified as follows:
  - sources S_3 and S_4 *do not deploy* message splitting: they transmit only *common* messages (i.e., $M_{\ell\mathsf{c}} = M_\ell, M_{\ell\mathsf{p}} = \emptyset, \ell = 3, 4$), which are then encoded into codewords $U_\ell^n \sim \mathcal{N}(0, P_\ell)$, i.i.d., $\ell = 3, 4$.
  - sources S_1 and S_2 *perform message splitting*: for a fixed power splitting $P_{k,\mathsf{p}} := \alpha_k P_k$ and $P_{k,\mathsf{c}} := (1 - \alpha_k)P_k$, with $0 \leq \alpha_k \leq 1, k = 1, 2$, they encode their private messages into codewords $V_k^n \sim \mathcal{N}(0, P_{k,\mathsf{p}})$, i.i.d., and common messages into codewords $U_k^n \sim \mathcal{N}(0, P_{k,\mathsf{c}})$, i.i.d., and transmit the superimposed codeword $X_k^n := U_k^n + V_k^n, k = 1, 2$.
- conversely, for $Q = 1$ the roles of S_1 and S_2 are swapped with those of S_3 and S_4.
- the decoding operations at destinations D_1 and D_2 are unchanged.

The resulting region is computable where each $I(.)$ term is now computed with Gaussian signals with appropriate powers that depend on $Q \in \{0,1\}$. For example, in constraint (5.14), i.e., $R_1 - R_{a0} + R_{30} < I(X_1, U_3; Y_1|U_a, X_2, U_4, Q)$, with $a \in \{\emptyset, 1\}$, the $I(.)$ term evaluates to $\frac{1}{2}\log(1 + \frac{P_1 - P_{a,\mathsf{c}} + h_3^2 P_3}{1}) + \frac{1}{2}\log(1 + \frac{P_1 - P_{a,\mathsf{c}} + h_3^2 P_{3,\mathsf{c}}}{1 + h_3^2 P_{3,\mathsf{p}} + h_4^2 P_{4,\mathsf{p}}})$, where $P_{\emptyset,\mathsf{c}} := 0$ and the first and the second terms correspond to $Q = 0$ and $Q = 1$, respectively.

Since the achievable region for the GMAIC is obtained from that in Theorem 5.1 following the restrictions listed above, it is not repeated here.

5.3 Partially Symmetric MAIC: Sum-rate Upper Bounds

We first consider the *partially-symmetric* GMAIC where the cross-channel gains for sources S_1 and S_2 are the same, i.e.,

$$h_1^2 = h_2^2,$$

as well the cross-channel gains for sources S_3 and S_4 are also the same, i.e.,

$$h_3^2 = h_4^2.$$

We then characterize three upper bounds for the sum-rate of the partially-symmetric GMAIC based on genie-signal aided upper bounding techniques.

5.3.1 A Genie-aided Sum-rate Upper Bound

In the weak case, both cross-channels are weak, i.e.,

$$h_1^2 = h_2^2 < 1, \quad h_3^2 = h_4^2 < 1.$$

We provide two types of side-information (genie-signals) to the receivers and characterize the bounds based on the worst additive noise technique [144]. First, consider the genie-signals S_1^n and S_2^n intended for destinations D_1 and D_2:

$$\begin{aligned} S_1^n &= h_1(X_1^n + X_2^n) + N_1^n, \\ S_2^n &= h_3(X_3^n + X_4^n) + N_2^n, \end{aligned} \tag{5.23}$$

where $N_k \sim \mathcal{N}(0, \sigma_{N_k}^2)$ is independent of $(X_1, \ldots, X_4)$ and correlated with Z_k, which is defined in (5.3), such that $\mathbb{E}(N_k Z_k) = \rho_{N_k}\sigma_{N_k}$, with $|\rho_{N_k}| \leq 1$, and $\sigma_{N_k}^2 \leq 1$, $k = 1, 2$. These signals provide a noisy version of the *intended* signals to the receivers.

Next, we consider genie-signals U_1^n and U_2^n intended for receivers D_1 and D_2, given by

$$\begin{aligned} U_1^n &= h_3(X_3^n + X_4^n) + W_1^n, \\ U_2^n &= h_1(X_1^n + X_2^n) + W_2^n, \end{aligned} \tag{5.24}$$

where $W_k \sim \mathcal{N}(0, \sigma_{W_k}^2)$ is independent of $(X_1, \ldots, X_4)$ and correlated with Z_k, which is defined in (5.3), such that $\mathbb{E}(Z_k W_k) = \rho_{W_k}\sigma_{W_k}$, with $|\rho_{W_k}| \leq 1$, and $\sigma_{W_k}^2 \leq 1, k = 1, 2$. As opposed to (5.23), (5.24) provide a noisy version of the *interfering* signals to the receivers.

We denote the set of parameters involved in (5.23)-(5.24) as

$$\mathbf{c} := (\sigma_{N_1}, \sigma_{N_2}, \sigma_{W_1}, \sigma_{W_2}, \rho_{N_1}, \rho_{N_2}, \rho_{W_1}, \rho_{W_2}) \tag{5.25}$$

such that $\sigma_{N_k}, \sigma_{W_k} \in [0, 1]$ and $\rho_{N_k}, \rho_{W_k} \in [-1, 1], k = 1, 2$.

For notational convenience, we also tabulate the variance $(\sigma_{(.)}^2)$ of a few relevant Gaussian random variables appearing in the expression of the upper bound as follows

$$\sigma_{W_2|Z_2-W_2}^2 = \sigma_{W_2}^2(1 - \rho_{W_2}^2)/(1 + \sigma_{W_2}^2 - 2\rho_{W_2}\sigma_{W_2}) \tag{5.26}$$

$$\sigma_{Z_2-W_2}^2 = 1 + \sigma_{W_2}^2 - 2\rho_{W_2}\sigma_{W_2} \tag{5.27}$$

$$\sigma_{Z_1-N_1/h_1}^2 = 1 + {\sigma_{N_1}}^2/{h_1}^2 - 2\rho_{N_1}\sigma_{N_1}/h_1 \tag{5.28}$$

$$\sigma_{Z_1|N_1}^2 = 1 - \rho_{N_1}^2. \tag{5.29}$$

We now have the following sum-rate upper bound.

Theorem 5.2 (Sum-rate upper bound for the partially symmetric weak GMAIC)**.** *For a set of parameters* $\mathbf{c}$ *in* (5.25)*, consider the following conditions:*

$$\text{[A1]}\ \sigma_{Z_1-N_1/h_1}^2 \leq 1, \tag{5.30}$$

$$\text{[A2]}\ \sigma_{Z_2-W_2}^2 \leq \sigma_{Z_1|N_1}^2/h_3^2, \tag{5.31}$$

$$\text{[A3]}\ \sigma_{W_2}^2 \leq \sigma_{W_2|Z_2-W_2}^2, \quad \sigma_{N_1}^2 \leq 1, \tag{5.32}$$

$$\text{[A4]}\ \ \sigma_{N_1}^2 \leq \sigma_{W_2|Z_2-W_2}^2, \quad \sigma_{W_2}^2 \leq 1. \tag{5.33}$$

If $\mathbf{c}$ *satisfies either the three conditions* [A1], [A2] *and* [A3]*, or the three conditions* [A1], [A2] *and* [A4]*, the sum-capacity of the partially symmetric weak GMAIC is upper bounded by*

$$C_{\mathsf{S}} \leq \frac{1}{4}\log \underbrace{\left(\frac{P_1 + P_2 + h_3^2(P_3 + P_4) + 1 - \frac{(h_1(P_1+P_2)+\rho_{N_1}\sigma_{N_1})^2}{h_1^2(P_1+P_2)+\sigma_{N_1}^2}}{h_3^2(P_3 + P_4) + 1 - \frac{\sigma_{N_1}^2(\rho_{N_1}-\sigma_{N_1}/h_1)^2}{h_1^2(P_1+P_2)+\sigma_{N_1}^2}} \right)}_{\mu_1}$$

$$+\frac{1}{4}\log\underbrace{\left(\frac{h_1^2(P_1+P_2)+\sigma_{N_1}^2}{h_1^2(P_1+P_2)+\sigma_{W_2|Z_2-W_2}^2}\right)}_{\mu_2}-\frac{1}{4}\log(h_3^2)+\frac{1}{4}\log\underbrace{\left(\frac{h_1^2(P_1+P_2)+\sigma_{W_2}^2}{h_1^2(P_1+P_2)+1}\right)}_{\mu_4}$$

$$+\frac{1}{4}\log\underbrace{\left(\frac{h_1^2(P_1+P_2)+P_3+P_4+1-\frac{(h_1^2(P_1+P_2)+\sigma_{W_2}\rho_{W_2})^2}{h_1^2(P_1+P_2)+\sigma_{W_2}^2}}{P_3+P_4+\frac{1-\rho_{N_1}^2}{h_3^2}-\frac{\sigma_{W_2}^2(\rho_{W_2}-\sigma_{W_2})^2}{h_1^2(P_1+P_2)+\sigma_{W_2}^2}}\right)}_{\mu_3}-\frac{1}{4}\log\left(\sigma_{N_1}^2\sigma_{Z_2-W_2}^2\right)$$

$$+\frac{1}{4}\log\left(1+P_1+P_2+h_3^2(P_3+P_4)\right)+\frac{1}{4}\log\left(1+P_3+P_4+h_1^2(P_1+P_2)\right). \tag{5.34}$$

Proof. We relegate the proof to Appendix D.2. ■

While the details are relegated to Appendix D.2, the key idea of the proof is given here: we provide side-information S_1^n to destination D_1 and U_2^n to destination D_2, and then single-letterize the resulting multi-letter entropy terms by applying the (conditional) worst additive noise (WAN) technique [130, 144].

The effectiveness of this bound is illustrated with some numerical examples in Section 5.3.3. Note that a complementary upper bound can also be found by providing U_1^n to D_1 and S_2^n to D_2. The resulting bound is obtained from that in Theorem 5.2 by exchanging indexes $1 \leftrightarrow 3$ and $2 \leftrightarrow 4$. Additionally, another bound can be obtained by providing only the noisy interfering signals U_1^n and U_2^n to D_1 and D_2, however, this bound turns out to be quite loose compared to that in Theorem 5.2, and hence is omitted.

The Completely Symmetric Case: Simplification of the Bound in Theorem 5.2

Note that the upper bound in (5.34) is valid for all channel parameters that satisfy the constraints of Theorem 5.2, and therefore, to characterize the tightest upper bound, (5.34) needs to be minimized over all choices of these parameters. This problem, however, is non-convex and thus difficult to solve. Hence, we simplify bound (5.34) to a closed form for the *completely-symmetric weak* GMAIC where all cross-channel gains as well as all power constraints are the same, i.e.,

$$h^2 := h_k^2 < 1, \quad \text{and} \quad P := P_k, \quad k \in \{1,2,3,4\}. \tag{5.35}$$

We now have the following upper bound on the sum-rate of the completely symmetric weak GMAIC.

Theorem 5.3. *The sum-capacity of the completely symmetric weak GMAIC is upper bounded by the minimum of (5.36) and (5.37) below:*

$$C_{\mathsf{S}} \le \frac{1}{2}\log\left(2|h|P + \frac{1+2P}{|h|}\right) + \frac{1}{4}\log\frac{4h^4}{\sigma_{N_1}^2(4h^2-\sigma_{N_1}^2)}, \quad \textit{where} \tag{5.36}$$

$$\sigma_{N_1}^2 = \begin{cases} 4h^2(1-h^2), & \textit{if } h^2 \le \frac{1}{2}, \\ 1, & \textit{otherwise}, \end{cases}$$

$$C_{\mathsf{S}} \le \frac{1}{2}\log\left(2|h|P + \frac{1+2P}{|h|}\right) + \frac{1}{2}\log\frac{4h^2+1}{4|h|}. \tag{5.37}$$

Proof. Bound (5.36) is obtained by taking conditions [A1], [A2] and the first condition in [A3], i.e., $\sigma_{W_2}^2 \le \sigma_{W_2|Z_2-W_2}^2$, with equality, and then minimizing the remaining terms over $\sigma_{N_1}^2 \le 1$. First, taking [A1] with equality results in $\mu_1 = 1$ in (5.34), and thus $\log(\mu_1) = 0$. This also results in $\sigma_{N_1} = 2\rho_{N_1}h$. Similarly, taking [A2] with equality simplifies μ_3 in (5.34) to $\mu_3 = 1$, and thus $\log(\mu_3) = 0$.

Next, taking the first condition in [A3] $\sigma_{W_2}^2 \le \sigma_{W_2|Z_2-W_2}^2$ with equality gives $\sigma_{W_2}^2 = \rho_{W_2}^2$, and this results in the numerator of μ_4 and denominator of μ_2 in (5.34) having the same value which cancel each other. Moreover, from (5.27), (5.29), and equality in [A1] and [A2], we solve for ρ_{W_2}, which gives us $\rho_{W_2} = h^{-1}(h^2 + \sigma_{N_1}^2/4h^2 - 1)^{1/2}$.

The remaining terms in (5.34) then simplify to

$$C_{\mathsf{S}} \le \min_{\sigma_{N_1}^2} \frac{1}{2}\log\left(1+2P+2h^2P\right) - \frac{1}{4}\log(h^2) + \frac{1}{4}\log\left(\frac{2h^2P+\sigma_{N_1}^2}{2h^2P+1}\right) + \frac{1}{4}\log\frac{4h^4}{\sigma_{N_1}^2(4h^2-\sigma_{N_1}^2)} \tag{5.38}$$

$$\text{subject to} \quad 4h^2(1-h^2) \le \sigma_{N_1}^2 \le \min(1, 4h^2). \tag{5.39}$$

In the constraint (5.39), the first upper bound $\sigma_{N_1}^2 \le 1$ follows directly as defined in (5.25), whereas the other upper bound $\sigma_{N_1}^2 \le 4h^2$ follows from the upper bound $\rho_{N_1}^2 \le 1$ in (5.25) since $\sigma_{N_1} = 2\rho_{N_1}h$. Moreover, the lower bound $4h^2(1-h^2) \le \sigma_{N_1}^2$ follows from the lower bound $\rho_{W_2}^2 \ge 0$ in (5.25) after substituting ρ_{W_2} via $\rho_{W_2} = h^{-1}(h^2 + \sigma_{N_1}^2/4h^2 - 1)^{1/2}$.

Note that only the last 2 terms in (5.38) depend on $\sigma_{N_1}^2$, and thus the problem reduces to that of minimizing these two terms subject to (5.39). This problem is non-convex as the term $\gamma(\sigma_{N_1}) := \log\left(\frac{2h^2P+\sigma_{N_1}^2}{2h^2P+1}\right)$ in (5.38) is non-convex. However, since for all valid σ_{N_1} (i.e., $\sigma_{N_1}^2 \in [0,1]$), we have $\gamma(\sigma_{N_1}) \le 0$, we simplify the problem by taking $\gamma(\sigma_{N_1}) = 0$. Thus, the problem reduces to

$$\text{minimize} \quad -\log\sigma_{N_1}^2(4h^2-\sigma_{N_1}^2) \tag{5.40}$$

$$\text{subject to} \quad 4h^2(1-h^2) \le \sigma_{N_1}^2 \le \min(1, 4h^2). \tag{5.41}$$

This problem is convex, and hence is solved using KKT conditions (see Appendix D.3 for the details). The optimum solution is given in closed form by

$$\sigma_{N_1}^2 = \begin{cases} 4h^2(1-h^2), & \text{if } h^2 \le \frac{1}{2}, \\ 1, & \text{otherwise,} \end{cases}$$

and, when this solution is applied to (5.36), it yields a closed form upper bound to the sum-rate.

Next, the bound in (5.37) is found by taking conditions [A1], [A2] and both conditions in [A4] with equality. Specifically, taking [A1] with equality results in $\log(\mu_1) = 0$ in (5.34), and taking [A2] with equality results in $\log(\mu_3) = 0$ in (5.34), as in the previous case. Then, taking the first and the second condition in [A4] with equality results in $\log(\mu_2) = 0$ and $\log(\mu_4) = 0$ in (5.34) respectively. Moreover, taking [A2] and [A4] with equality results in the following closed form solution of variables $\rho_{W_2} = \frac{4h^2-1}{4h^2+1}$ and $\sigma_{N_1}^2 = \frac{4h^2}{(4h^2+1)}$. Substituting the expressions for ρ_{W_2} and $\sigma_{N_1}^2$, the remaining terms simplify to (5.37). ■

5.3.2 Additional Genie-aided Sum-rate Upper Bounds

The sum-rate upper bound in Theorem 5.2 is generally better than existing bounds in the weak interference regime such as that in [12] when the values of the cross-channel gains are moderate (i.e., cross-channel gains not close to either 0 or 1). However, when the cross-channel gains are close to 1, the following sum-rate upper bound outperforms that of Theorem 5.2.

Theorem 5.4. *The sum-capacity of the partially symmetric GMAIC with $h_1^2 < 1$ is upper bounded by*

$$R_1 + R_2 \le \frac{1}{2}\log\left(1 + P_1'\right), \quad R_3 + R_4 \le \frac{1}{2}\log\left(1 + \frac{P_2'}{P_1' + 1/h_1^2}\right) \tag{5.42}$$

for some $P_1', P_2' > 0$ that satisfy

$$P_1' + P_2' = P_1 + P_2 + \frac{P_3 + P_4}{h_1^2}. \tag{5.43}$$

Conversely, an upper bound for the partially symmetric GMAIC with $h_3^2 < 1$ is found by swapping indexes $1 \leftrightarrow 3$ and $2 \leftrightarrow 4$.

Proof. The proof is given in the Appendix D.4. ■

We prove this result by transforming the GMAIC into a degraded 4-user broadcast channel (BC) in a step-by-step manner as detailed in Appendix D.4, which follows a similar approach for the IC in [139]. Intuitively, compared to the GMAIC where the individual power constraint of each source needs to be satisfied, in the degraded BC the powers of all sources can be shared. Since power-sharing does not reduce the sum-rate, the resulting sum-rate upper bound for the degraded BC provides an upper bound to the sum-rate of the GMAIC as well.

Note that instead of channel parameters $(P_1, \dots P_4)$, the above bound is expressed in terms variables (P_1', P_2') that satisfy (5.43). Hence, the best upper bound can be found by optimizing (5.42) subject to (5.43). However, the problem is non-convex, and hence difficult to solve. Therefore, to characterize a closed form of this bound, we choose the particular operating point on the (P_1', P_2')-plane where the two rate bounds in (5.42) are the same, i.e., $R_1 + R_2 = R_3 + R_4$ holds. Solving for P_1', P_2', and substituting them, the following closed form bound is obtained

$$C_{\mathsf{S}} \leq \log \left[1 - \frac{1}{h_1^2} + \left(\left(1 + \frac{1}{h_1^2} \right)^2 + 4 \left(P_1 + P_2 + \frac{P_3 + P_4}{h_1^2} \right) \right)^{1/2} \right] - 1. \tag{5.44}$$

Also note that unlike Theorem 5.2, where both cross-channels are weak, the bound in Theorem 5.4 is valid when at least one of the cross-channel gains is weak, e.g., for (5.44) to hold, channel gain $h_1^2 < 1$.

We now provide another upper bound on the sum-rate, which is valid irrespective of whether any of the cross links is weak or strong.

Theorem 5.5. *The sum-capacity of the partially symmetric GMAIC is upper bounded by*

$$C_{\mathsf{S}} \leq \frac{1}{2} \log \left[\frac{\left(1 + P_3 + P_4 + h_1^2 (P_1 + P_2)\right) (1 + P_1 + P_2)}{1 + \min(h_1^2, 1)(P_1 + P_2)} \right].$$

A second upper bound is found by swapping indexes $1 \leftrightarrow 3$ *and* $2 \leftrightarrow 4$.

Proof. The proof is given in the Appendix D.5 ■

We outline the key steps here. We provide destination D_2 a single genie-signal, and aim to upper bound the resulting terms as in Theorem 5.2. In contrast to Theorem 5.2, where two different genie-signals, a noisy version of the intended signals and a noisy version of the interfering signals, are provided, in this case a *single* genie-signal is provided which can be considered a noisy version of both the intended and the interfering signals. This

approach is similar to that of [139] for the IC. Under some conditions on the parameters of the genie-signal, destination D_2 is then shown to be able to decode all messages as in a 4-user MAC, thus providing a sum-rate upper bound. Finally, the obtained upper bound is minimized over the parameters of the genie-signal to characterize a tighter bound.

While the bound in Theorem 5.5 is valid in both the weak and strong interference regime, for the weak interference regime, the bound in Theorem 5.5 is outperformed by that in Theorem 5.4.

Lemma 5.1. *If $h_1^2 < 1$, the bound in Theorem 5.4 is tighter than that in Theorem 5.5.*

Proof. Suppose that $h_1^2 < 1$. We can choose $P_1' = P_1 + P_2$ and $P_2' = \frac{P_3+P_4}{h_1^2}$, and the resulting bound in (5.42) simplifies to that of Theorem 5.5. Therefore, the bound in Theorem 5.5 is a special case of that in Theorem 5.4, and is generally looser. ■

For the case with $h_3^2 < 1$, the counterparts of bounds in Theorem 5.4 and Theorem 5.5 are obtained by swapping the indexes $1 \leftrightarrow 3$ and $2 \leftrightarrow 4$. Then, following the principle of Lemma 5.1, the bound resulting from Theorem 5.4 can be shown to be tighter than that from Theorem 5.5.

5.3.3 Numerical Examples

We now present some numerical examples to demonstrate the effectiveness of the sum-rate upper bounds derived in the previous section. In Figure 5.4, we present numerical examples of the upper bounds in Theorem 5.2 through Theorem 5.4, and for comparison we also include the upper bound in [12, equation (13)]. For simplicity of exposition, we provide the examples for the completely symmetric case where all cross-channel gains are the same $h^2 = h_k^2$ as well as all transmit powers are the same $P = P_k, k \in \{1, \ldots, 4\}$. More specifically, we plot these bounds against cross-channel gain $h^2 \in (0, 1)$ in the weak interference regime with transmit power $P = 100$.

Note that the bound in Theorem 5.2 is subject to certain constraints on the parameters involved, and not available in closed form. Hence, we plot the numerically optimized version of this bound. Also recall that the bound of Theorem 5.3 was obtained by simplifying the bound of Theorem 5.2 for the completely symmetric case. Moreover, for the bound in Theorem 5.4, the closed form version provided in (5.44) is plotted.

We observe the following in Figure 5.4 :

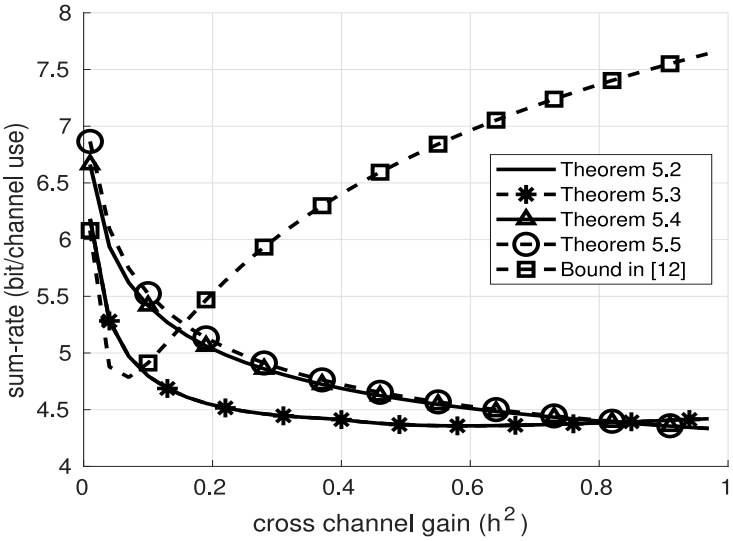


Figure 5.4: Sum-rate upper bounds for the completely symmetric weak GMAIC. The following sum-rate upper bounds are plotted against cross-channel gain $h^2 \in (0, 1)$ with transmit power $P = 100$: Theorem 5.2, its simplified version in Theorem 5.3, the closed form of Theorem 5.4 in (5.44), Theorem 5.5, and [12, equation (13)].

- Bound in [12] (square markers): this is the tightest bound for small h^2, i.e., $h^2 \leq 0.09$ approximately. However, as h^2 increases beyond $h^2 \approx 0.09$, the gap between this bound and other bounds increases consistently in $h^2 \in (0, 1)$.

- Theorem 5.2 (solid line): this bound is the tightest one for a wide range of h^2, i.e., $0.09 < h^2 \leq 0.83$ (approx.).

- Theorem 5.3 (solid star markers): even though this bound is the simplified version of the bound in Theorem 5.2, for this example, the former almost overlaps with the latter, and thus provides a tight bound.

- Theorem 5.4 (triangle markers): for a wide range of h^2, i.e., $h^2 \leq 0.83$ (approx.), this bound is outperformed by Theorem 5.2. However, for large values of h^2, i.e., $1 > h^2 > 0.83$, it provides the tightest bound.

- Theorem 5.4 (circle markers): for the entire range of weak channel gains, i.e., $0 < h^2 < 1$, this bound is outperformed by that in Theorem 5.4. In addition, for large values of h^2, i.e., $1 > h^2 > 0.84$, it outperforms Theorem 5.2.

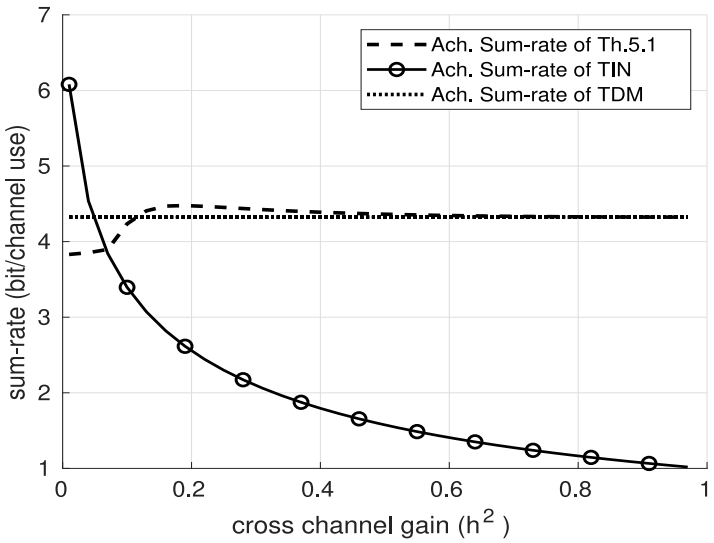


Figure 5.5: Achievable sum-rates for the completely symmetric weak GMAIC: the sum-rates from the scheme of Theorem 5.1, the TIN scheme, the TDM scheme. The bounds are plotted against cross-channel gain $h^2 \in (0, 1)$ with transmit power $P = 100$.

We now compare three achievable sum-rates, the sum-rate resulting from the scheme in Theorem 5.1, the sum-rate achieved by treating interference as noise (R_{TIN}), and the sum-rate achieved by time-division multiplexing with power control (R_{TDM}), as specified below.

Note that the rate region in Theorem 5.1 is expressed in terms of rates $(R_1, \ldots, R_4)$ as well as common rates $(R_{10}, \ldots, R_{40})$. Hence, we characterize the maximum sum-rate by first applying Fourier Motzkin elimination [100] to eliminate the common rates, and then finding the optimum private-common power splitting via a numerical grid search.

In the TIN scheme, the two interfering messages are treated entirely as noise at each receiver, and the resulting sum-rate is

$$R_{\mathsf{TIN}} = \log\left(1 + \frac{2P}{1 + 2h^2 P}\right).$$

Note that TIN can be regarded as a special case of the scheme in Theorem 5.1 if the time-sharing variable is restricted to only $Q = 0$, and all sources transmit only private messages, while both destinations treat the interfering messages as noise.

In the TDM with power control, for a fraction $\frac{1}{2}$ of the total transmission time, the sources in the first MAC transmit with double their average power limit, i.e., $2P$, while the sources in the second MAC remain silent. For the remaining $\frac{1}{2}$ of the total transmission

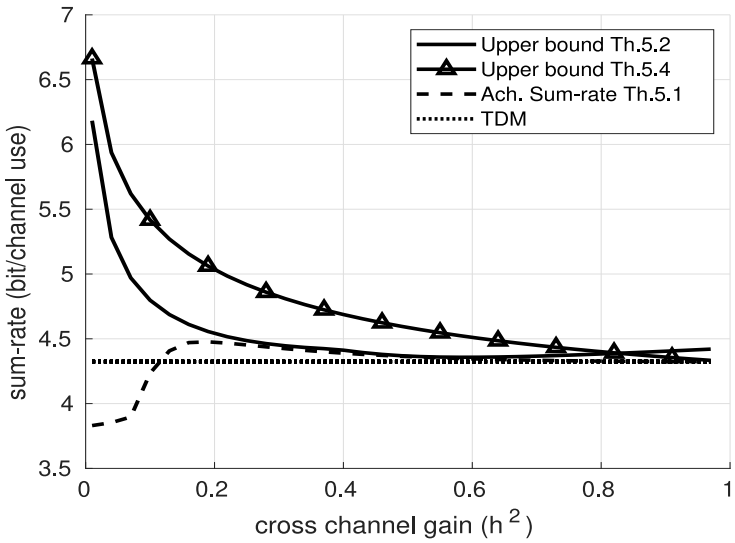


Figure 5.6: Comparison of sum-rate upper bounds of Theorem 5.2 and (5.44) with achievable sum-rates of Theorem 5.1 and TDM for the completely symmetric weak GMAIC. The bounds are plotted against cross-channel gain $h^2 \in (0,1)$ with transmit power $P = 100$.

time, the roles of the two MACs are swapped. Since each source transmit for only $\frac{1}{2}$ of the total transmission time with power $2P$, the average power constraint of P still holds. The resulting sum-rate is given by

$$R_{\mathsf{TDM}} = \frac{1}{2}\log(1+4P).$$

In Figure 5.5, we plot these three achievable sum-rates for the completely symmetric GMAIC against $h^2 \in (0,1)$ in the weak interference regime for $P = 100$, and observe that

- TIN: the TIN sum-rate performs well only for very small h^2, i.e., $h^2 \leq 0.05$ (approx.), and beyond that it performs poorly compared to the two other bounds.
- The scheme of Theorem 5.1: except for small values of h^2, i.e., $h^2 \leq 0.11$ (approx.), this sum-rate provides the best performance among the three, however, its performance almost overlap with the TDM bound for large enough h^2, i.e., $h^2 \geq 0.6$ (approx.).
- TDM: this sum-rate does not depend on h^2, and hence remains constant. Moreover, its performance becomes close to that of the scheme of Theorem 5.1 for large enough h^2, i.e., $h^2 \geq 0.6$ (approx.).

Finally, in Figure 5.6, we compare the two most relevant sum-rate upper bounds, i.e., the bound in Theorem 5.2 (optimized numerically as in Figure 5.4) and the bound in (5.44) derived from Theorem 5.4, with two most relevant achievable sum-rates, i.e., the sum-rate of Theorem 5.1 (optimized numerically as in Figure 5.5) and the TDM sum-rate. The sum-rates are plotted for the completely symmetric GMAIC against $h^2 \in (0,1)$ for the weak interference regime with $P = 100$, from which we observe that

- the achievable rate of Theorem 5.1 is quite close to the sum-rate upper bound in Theorem 5.2. For example, for $h^2 \geq 0.16$, the achievable rate is within 97% of the upper bound.
- the simple scheme of TDM is also quite close to the sum-rate upper bound in Theorem 5.2. For example, for $h^2 \geq 0.16$, the TDM rate is within 94% of the upper bound.

We thus observe that for the case of Figure 5.6, the sum-capacity of the symmetric weak GMAIC is closely approximated by the upper bound in Theorem 5.2 and the lower bound in Theorem 5.1 for a wide range of cross-channel gains.

5.4 Outer Bounds for the General Gaussian MAIC

In this section, we characterize a sum-rate upper bound and an outer bound to the capacity region of the GMAIC in the general case, where the channel gains are arbitrary and do not need to satisfy either the partially symmetric or the completely symmetric channel conditions of Section 5.3. Since symmetry conditions on the channel gain do not hold anymore, new upper bounding techniques are needed.

5.4.1 A Sum-rate Upper Bound

First, we present the following sum-rate upper bound, which is derived based on providing specific subsets of interfering signals to each receiver.

Theorem 5.6 (A sum-rate upper bound for the general MAIC)**.** *Let* $\Omega_1 := \{1,2\}$, $\Omega_2 := \{3,4\}$, *and* $X_{\Omega_1} := \{X_m\}_{m\in\Omega_1}$, $X_{\Omega_2} := \{X_m\}_{m\in\Omega_2}$. *The sum-capacity of the GMAIC is then upper bounded by*

$$C_{\mathsf{S}} \leq \frac{1}{2}\min\left(\, \mathsf{U}(X_3,X_2),\ \mathsf{U}(X_3,X_1),\ \mathsf{U}(X_4,X_2),\ \mathsf{U}(X_4,X_1)\,\right),$$

where, for $\ell \in \Omega_1$ *and* $k \in \Omega_2$*, we have*

$$\begin{aligned}\mathsf{U}(X_k, X_\ell) := {} & I(X_{\Omega_1}, X_{\Omega_2\setminus k}; Y_1|X_k) + I(X_{\Omega_1\setminus\ell}, X_{\Omega_2}; Y_2|X_\ell) \\ & + I(X_{\Omega_1\setminus\ell}; Y_1|X_{\Omega_2}, X_\ell) - I(X_{\Omega_1\setminus\ell}; Y_2|X_{\Omega_2}, X_\ell) \\ & + I(X_{\Omega_2\setminus k}; Y_2|X_{\Omega_1}, X_k) - I(X_{\Omega_2\setminus k}; Y_1|X_{\Omega_1}, X_k) \\ & + \min\big(I(X_\ell, X_k; Y_1|X_{\Omega_1\setminus\ell}, X_{\Omega_2\setminus k}) + I(X_k; Y_2|X_{\Omega_2\setminus k}, X_{\Omega_1}) - I(X_k; Y_1|X_{\Omega_2\setminus k}, X_{\Omega_1}), \\ & \qquad I(X_\ell, X_k; Y_2|X_{\Omega_1\setminus\ell}, X_{\Omega_2\setminus k}) + I(X_\ell; Y_1|X_{\Omega_2}, X_{\Omega_1\setminus\ell}) - I(X_\ell; Y_2|X_{\Omega_2}, X_{\Omega_1\setminus\ell})\big)\,,\end{aligned}$$

in which $X_i \sim \mathcal{N}(0, P_i), i \in \{1, 2, 3, 4\}$*, are independent of each other, and* (Y_1, Y_2) *are obtained from* (5.3)*.*

Proof. The proof is relegated to Appendix D.6. ■

While the proof is relegated to Appendix D.6, we provide the key idea here. Referring to Theorem 5.6, we take $\ell = 2, k = 3$, and provide destination D_1 the interfering signal X_k^n and destination D_2 the interfering signal X_ℓ^n. The resulting terms can then be expressed as either of the two types of terms which are upper bounded as follows: (a) *positive* multi-letter entropy terms, which are upper bounded by taking $X_i \sim \mathcal{N}(0, P_i)$, i.i.d., as in [131, Lemma 1], or (b) the *difference* of two multi-letter entropy terms, which are upper bounded by applying the *worst additive noise* technique [130], for which $X_i \sim \mathcal{N}(0, P_i)$, i.i.d., proves optimal. The bounds for all other cases of $\ell \in \Omega_1$ and $k \in \Omega_2$ are obtained similarly.

5.4.2 An Outer Bound to the Capacity Region

We now present an outer bound to the capacity region of the GMAIC. The outer bounding technique is based on a similar result on a related model, the MAC-IC-MAC channel [145] depicted in Figure 5.7. In contrast to the GMAIC where all users in one MAC interferes with the receiver of the other MAC, in the MAC-IC-MAC only one user from each MAC interferes with the receiver of the other MAC. Hence, in the the MAC-IC-MAC the two non-interfering users communicate with their desired receivers in a point-to-point fashion, while the two interfering users form a two-user IC as depicted in Figure 5.7. It can be noted that the MAC-IC-MAC can be obtained from the GMAIC by removing the interfering links from S_2 to D_2 and S_4 to D_1.

We now have the following outer bound result for the GMAIC.

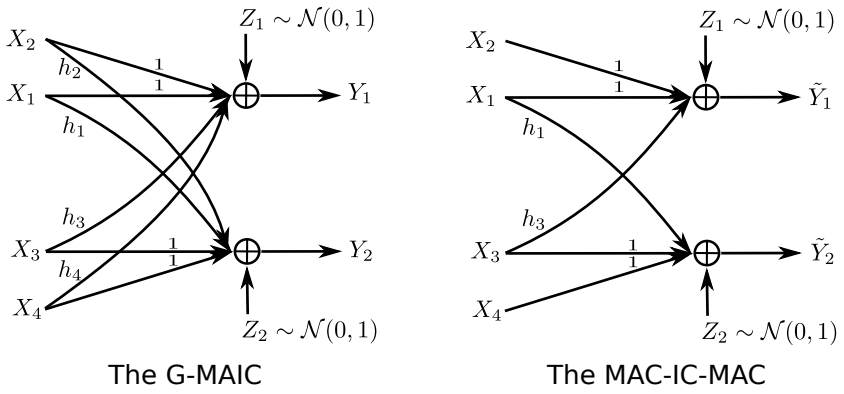


Figure 5.7: The GMAIC and the Gaussian MAC-IC-MAC of [145].

Theorem 5.7 (An outer bound to the capacity region of the general GMAIC)**.** *The capacity region of the GMAIC is contained within the set of non-negative rate tuples* (R_1, R_2, R_3, R_4) *that satisfy the following bounds*

$$\sum_{k\in\Omega_1} R_k \le \mathsf{C}\left(\sum_{k\in\Omega_1} P_k\right), \quad \Omega_1 \in 2^{\{1,2\}} \setminus \emptyset = \{\{1\},\{2\},\{1,2\}\}$$

$$\sum_{k\in\Omega_2} R_k \le \mathsf{C}\left(\sum_{k\in\Omega_2} P_k\right), \quad \Omega_2 \in 2^{\{3,4\}} \setminus \emptyset = \{\{3\},\{4\},\{3,4\}\}$$

$$\sum_{k\in\Gamma_1} R_k + \sum_{k\in\Omega_2} R_k \le \mathsf{C}\left(\frac{P_2}{1+h_2^2P_2} + \sum_{k\in\Gamma_1\setminus\{2\}} P_k\right) + \mathsf{C}\left(h_2^2P_2 + \sum_{k\in\Omega_2} P_k\right)$$

$$\Gamma_1 \in \{2\} \cup 2^{\{1,2\}\setminus\{2\}} = \{\{2\},\{1,2\}\}$$

$$\sum_{k\in\Omega_1} R_k + \sum_{k\in\Gamma_2} R_k \le \mathsf{C}\left(\frac{P_3}{1+h_3^2P_3} + \sum_{k\in\Gamma_2\setminus\{3\}} P_k\right) + \mathsf{C}\left(h_3^2P_3 + \sum_{k\in\Omega_1} P_k\right)$$

$$\Gamma_2 \in \{3\} \cup 2^{\{3,4\}\setminus\{3\}} = \{\{3\},\{3,4\}\}$$

$$\sum_{k\in\Gamma_1} R_k + \sum_{k\in\Gamma_2} R_k \le \mathsf{C}\left(\frac{P_2}{1+h_2^2P_2} + h_3^2P_3 + \sum_{k\in\Gamma_1\setminus\{2\}} P_k\right) + \mathsf{C}\left(\frac{P_3}{1+h_3^2P_3} + h_2^2P_2 + \sum_{k\in\Gamma_2\setminus\{3\}} P_k\right)$$

$$\sum_{k\in\Gamma_1} R_k + \sum_{k\in\Omega_1} R_k + \sum_{k\in\Gamma_2} R_k \le \mathsf{C}\left(\frac{P_2}{1+h_2^2P_2} + \sum_{k\in\Gamma_1\setminus\{2\}} P_k\right) + \mathsf{C}\left(h_3^2P_3 + \sum_{k\in\Omega_1} P_k\right)$$

$$+\mathsf{C}\left(\frac{P_3}{1+h_3^2P_3}+h_2^2P_2+\sum_{k\in\Gamma_2\setminus\{3\}}P_k\right)$$

$$\sum_{k\in\Gamma_1}R_k+\sum_{k\in\Omega_2}R_k+\sum_{k\in\Gamma_2}R_k\leq\mathsf{C}\left(\frac{P_3}{1+h_3^2P_3}+\sum_{k\in\Gamma_2\setminus\{3\}}P_k\right)+\mathsf{C}\left(h_2^2P_2+\sum_{k\in\Omega_2}P_k\right)$$
$$+\mathsf{C}\left(\frac{P_2}{1+h_2^2P_2}+h_3^2P_3+\sum_{k\in\Gamma_1\setminus\{2\}}P_k\right). \tag{5.45}$$

Proof. The proof is relegated to Appendix D.7. ■

The key idea is that in general, removing one or more interfering links from the GMAIC results in a channel whose capacity region is no less than the capacity region of the GMAIC. Intuitively, having less interference via eliminating interfering links may facilitate decoding of the intended signals with no lesser (and possibly larger) rates. Since the MAC-IC-MAC is obtained by eliminating two interfering links from the GMAIC, capacity region of the MAC-IC-MAC naturally serves as an outer bound to the capacity region of the GMAIC. While the capacity region of the MAC-IC-MAC is unknown in general, an outer bound to its capacity region has been characterized in [145, Theorem 7]. This outer bound for the MAC-IC-MAC, therefore, also serves as an outer bound to the capacity region of the GMAIC, which is presented in Theorem 5.7.

5.5 Summary

In this chapter, we studied the performance limits of the Gaussian MAIC (GMAIC). The GMAIC models the interference-limited communication between a pair of two-user multiple-access channels that operate solely over the shared conventional single band in two neighboring cells, and thus mutually interfere. For the GMAIC, we first characterized an achievable region based on the Han-Kobayashi strategy, where each transmitter splits its message into a private and a common part, and each decoder is allowed to non-uniquely decode the common interfering message. Numerical examples for the symmetric case illustrates that for a significant range of channel gains, the sum-rate resulting from the proposed achievable strategy outperforms that of traditional schemes such as the scheme of treating interference as noise.

Then, focusing on the *weak* interference case, which is a more relevant model for practical cellular interference, we characterized 3 novel genie-signal aided sum-rate upper bounds for the partially symmetric case, and one sum-rate upper bound for the case without any symmetry. Numerical examples for the symmetric case shows that the proposed sum-rate upper bounds outperform the existing one for a broad range of channel gains. They also show that for a wide range of channel gains, the achievable sum-rate and one of the upper bounds differ only by a small numerical gap, thus providing good approximations to the sum-capacity in these cases.

Chapter 6

Concluding Remarks

6.1 Summary of Contributions and Conclusions

Complementing the transmissions in the conventional microwave band with transmissions in the mm-wave spectrum has emerged as an attractive solution to the problem of spectrum scarcity in 5G. Therefore, in this thesis, we studied the performance of two multi-user networks over an integrated mm-wave/microwave dual-band. In this work, transmissions in the mm-wave band, which are established by beamforming via digitally steerable co-phased antenna arrays, are modeled as point-to-point AWGN links that can be steered towards a desired receiver with negligible to no interference to neighboring nodes. In contrast, due to the nature of the conventional microwave band, transmissions in this band interfere with the reception at neighboring nodes.

We studied two multi-user channels in such a dual-band setting that model two important aspects of wireless communication: inter-user interference and relay cooperation. In particular, we studied the two-user dual-band *interference channel*, and the two-user dual-band *multiple-access relay channel* with the overall goal of characterizing the fundamental limits of their performance. Moreover, since the performance of the two dual-band networks are significantly affected by the point-to-point mm-wave links, we then studied the problem of optimal resource allocation over the mm-wave links of these networks. In addition, motivated by the fact that the optimal performance of interference-limited networks over the conventional single band is still an open problem, we studied the *multiple-access interference channel* over a single band with the aim of characterizing its performance limits and optimal encoding/decoding strategies.

The following are the major contributions of the thesis:

In Chapter 3, we studied the performance limits of two significant variations of the dual-band two-user interference channel: (a) the direct-and-cross-link IC (DCLIC), which consists of an Gaussian interference channel (GIC) in the underlying microwave band as well as two direct-links and two cross-links in the mm-wave band; and (b) the cross-link IC (CLIC), which consists os the microwave GIC and the two mm-wave relay-links.

The major findings from this study are as follows:

1. First, we showed that the capacity region of the DCLIC can be decomposed into the capacity region of the underlying *CLIC* and the capacity of the two mm-wave *direct-links*. This illustrates that it is optimal to operate the mm-wave direct-links *independently* from the CLIC, which simplifies the encoding/decoding procedure. As such, the mm-wave direct-links can be used to improve the data-rates of individual users, whereas the mm-wave cross-links play a non-trivial role in characterizing the capacity of the CLIC.

2. Due to the decomposition result, the capacity region of the DCLIC can be characterized from the capacity region of the CLIC. Hence, we next focus on the CLIC.

   Depending on whether the underlying GIC of the CLIC has strong, weak, or mixed interference, we classified the CLIC into the strong CLIC, the weak CLIC, and the mixed CLIC. We then characterized *sufficient channel conditions* under which the capacity region of these 3 classes of the CLIC is obtained in closed form. More specifically, we observed the following:

   (a) For the *strong* CLIC, the capacity region was characterized in closed form for the entire regime. In this case, the strategy of *jointly decoding* the intended messages and the interfering messages is optimal.

   (b) For the *weak* CLIC, we characterized the capacity under the *combined weak-very strong* conditions. This condition shows that even if the GIC in the microwave band has weak interference, adequately strong mm-wave cross-links are able to forward enough interference information to drive the overall interference to the very strong regime, where the capacity region is characterized by *successive interference cancellation*.

   (c) For the mixed CLIC, the capacity region was characterized under a set of sufficient conditions, which can be interpreted as being able to drive the CLIC from

the mixed to the strong interference regime. Under this condition, the strategy of jointly decoding both the intended and the interfering messages is optimal.

3. For the weak and mixed CLICs that do not necessarily satisfy the sufficient conditions in items (b)-(c) above, we characterized an *approximate* capacity result. More specifically, an achievable strategy is characterized where a message is allowed to be *partitioned* into a private part and a public part, while a receiver is able to decode only the *interfering public* message non-uniquely. Then, by adapting the power allocated to the private and public parts according to the level of interference in the underlying GIC, a closed form achievable region was obtained. This achievable region was then shown to be within 1/2 bit/channel use per user for the entire regimes of the weak or mixed CLICs.

4. Next, we studied the problem of optimal power allocation (OPA) over the mm-wave direct- and cross-links that maximizes the sum-rate of the DCLIC under a total power budget P for each source. The resulting scheme was shown to allocate power to the mm-wave links in distinct *modes* depending on the power budget P and channel parameters. The OPA was derived in closed form and its characteristics were described using the *link-gain regimes* which reveal useful insights. In particular, depending on the value of P, the OPA was shown to allocate power to the mm-wave links following either a *Waterfilling-like* property or a *saturation* property. More specifically, we observed the following:

   - For sufficiently small P, the OPA assigns power to only a subset of the links that depends on the relative strength os the links. Then, as P increases but remains below a saturation threshold, the OPA allocates power to all 4 links. In this regime of P, the OPA follows a Waterfilling-like strategy, and thus the power in the allocated links increase piece-wise linearly, while that in the remaining links are zero.
   - When P is increased beyond the saturation threshold, saturation occurs, which imposes a *maximum limit* on the power in the cross-links. Hence, as P increases further, the cross-link powers remain limited to a peak value, and all additional increments to P are allocated only to the direct-links.

   We also observed that when the mm-wave bandwidth is large, the optimal power allocation for the symmetric case simplifies as follows: (a) if the direct-links are

stronger than the cross-links, allocating the budget P entirely to the direct-links is optimal for all $P \geq 0$; (b) on the other hand, if the cross-links are stronger, all P smaller than the saturation threshold should be allocated entirely to the cross-links, whereas for P larger than the threshold, the cross-links become saturated, and hence all increments of P should be allotted to the direct-links.

In Chapter 4, we studied the dual-band two-user fading multiple-access relay channel, and more specifically, the performance limits of the two important variants of the channel: (a) the direct-and-relay-link MARC (DR-MARC), which consists of a conventional Gaussian MARC in the underlying microwave band as well as two direct-links and two relay-links in the mm-wave band; and (b) the relay-link MARC (R-MARC), which consists of an underlying Gaussian MARC and only the two mm-wave relay-links.

We obtain the following findings from this study:

1. First, we showed that the capacity region of the DR-MARC can be *decomposed* into the capacity region of the underlying *R-MARC* and the capacity of the two mm-wave *direct links*. Hence, similar to the dual-band interference channel, it is optimal to operate the mm-wave direct links *independently* from the R-MARC, which simplifies the encoding/decoding operation. While the direct-links can be used to improve individual user-rates, the relay-links play a significant role in the performance characterization of the R-MARC.

2. Due to this decomposition, the capacity of the DR-MARC can be readily determined from the capacity of the R-MARC. Hence, we focused on the R-MARC next:

   - We characterized an achievable region for the R-MARC based on a block-Markov encoding scheme that performs encoding *jointly over both bands*, where the relay operates in the decode-and-forward mode.
   - We obtained sufficient channel conditions, denoted the *jointly-near* conditions, under which the capacity of the R-MARC is characterized by the aforementioned achievable scheme. This condition show that even when the underlying microwave source-relay channels are not strong enough, if the mm-wave relay-links are sufficiently strong, then adequate information can be sent to the relay such that the relay is not a bottleneck to achievable rates. As a result, under these conditions the capacity region of the R-MARC was derived in closed form.

Moreover, the achievability and capacity results were then shown to generalize to the K-user version of the DR-MARC and the R-MARC, where $K > 2$.

3. We then studied the problem of the optimal power allocation (OA) over the mm-wave direct and relay-links of the DR-MARC with the goal of maximizing its sum-rate achievable with the aforementioned achievable scheme, where the links are subject to a total power budget P for each source. The resulting scheme was shown to allocate the link powers in different *modes* based on the value of P and channel parameters. All such modes were characterized in closed form, and the OA was then described using the *link gain regimes*, which reveals useful insights.

   For example, for a class of DR-MARCs where the jointly-near condition in the underlying R-MARC is satisfied by the source-relay *microwave channel gains only*, it is optimal to allocate the budget P entirely to the direct-links for all values of $P \geq 0$.

   In contrast, for DR-MARCs where the jointly-near condition needs to be satisfied *jointly* by the channel gains of *both* the microwave and mm-wave source-relay links, the MARC-OPA has a richer structure. Depending on the value of P, the MARC-OPA was shown to follow either a *Waterfilling-like* or a *saturation* property. In particular, we observed that

   - when P is *smaller* than a certain *saturation threshold*, the MARC-OPA follows the Waterfilling-like strategy. Hence, for sufficiently small P, the MARC-OPA allocates P entirely to the strongest of the direct and relay-links of each source. As P increases further but remains below the saturation threshold, the MARC-OPA eventually allocates power to all the two remaining links. Thus, in this regime, each link-power either increases piecewise linearly with P, or remains zero.
   - for P *larger* than the saturation threshold, the relay-link powers are *constrained* to satisfy a certain saturation condition, and thus the power in both relay-links cannot increase with P. As P increases further, if a relay link is *only stronger* than the other relay-link, then the power of both relay-links vary with P monotonically and approach positive constant levels, whereas if the relay link is *significantly stronger*, its power remains constant, and the power in the weaker relay-link remains at zero. In contrast to the relay-links, as P increases, the direct-link powers grow unbounded.

Moreover, when the mm-wave bandwidth becomes large, the optimal power allocation scheme for the symmetric case simplifies as follows: (a) if the direct-links are stronger than the relay-links, the budget P should be allocated entirely to the direct-links for all $P \geq 0$; (b) in contrast, when the relay-links are stronger, if P is smaller than the saturation threshold, then P should be allocated entirely to the relay-links, whereas for P larger than the threshold, the relay-links saturate, and thus all increments of P should then be allotted to the direct-links.

Finally, in Chapter 5, we studied the performance limits of the Gaussian multiple-access interference channel (MAIC) over a conventional single band, where a pair of two-user multiple access channels mutually interfere while communicating over a shared medium. More specifically, we studied the weak Gaussian MAIC, which is a more relevant model in practice, and focused on the sum-rate performance of the channel, from which the following findings are obtained:

1. We characterized an achievable region for the Gaussian MAIC where each transmitter is allowed to *partition* its message into a private and a common part, and each receiver is allowed to decode the *common interfering* message non-uniquely. A computable achievable region was then obtained by limiting the time-sharing variable Q to be *binary* and adapting the transmitters' operation accordingly. Numerical examples showed that this scheme *outperforms* conventional achievable schemes over a significant range of channel gains for the weak Gaussian MAIC.

2. We then considered the *partially symmetric* weak Gaussian MAIC, and derived three *upper bounds* on the sum-rate by providing different *genie signals* to the receivers. Numerical examples were then used to illustrate effectiveness of these bounds. By comparing the upper bounds with the achievable sum-rates for the completely symmetric case, it was found that for a wide range of cross channel gains, one of the upper bounds was close to the sum-rate achievable with the aforementioned scheme, thereby providing improved lower and upper bounds on the sum-rate of the weak Gaussian MAIC.

3. Finally, we derived a sum-rate upper bound for the *general* weak Gaussian MAIC, and characterized a set of outer bounds to the capacity region that is valid for any interference regime of the Gaussian MAIC.

6.2 Directions for Future Work

In the following, we list a few potential future directions we plan to pursue. While a few of these are extensions or generalizations of the work presented in this thesis, for the other problems new approaches are needed.

Resource Allocation for Dual-Band Channels with Alternative Objectives

Recall that in Chapter 3 and Chapter 4, we studied two problems of mm-wave resource allocation with the goal of maximizing the *sum-rate* of the respective channels. The sum-rate was considered as it captures the aggregate performance of all users, an important figure of merit for multi-user networks. However, such a resource allocation may greedily assign more rate to users with stronger channels, while starving users with weaker channels.

In many settings, it may be desirable or necessary to provide a fair rate to all users, and hence optimizing the max-min rate [146,147] or the proportional-fair rate [148–150] may be more relevant in such cases. Hence, as an alternative to the sum-rate objective in the thesis, the resource allocation problems for the dual-band channels can be investigated with either the max-min rate or the proportional-fair rate objective functions. Such problems are of interest for wireless settings with fairness quality-of-service constraints, and the resulting power allocation schemes may lead to useful practical insights. While the exact nature of these optimal power allocation schemes will be revealed only after thorough analysis, we believe that characteristics of these power allocation schemes will still be governed by the Waterfilling and the saturations properties in some form.

Extension to Other Dual-Band Multi-User Channels

In this thesis, we studied multi-user channels that model inter-user interference and relay cooperation, two important aspects in wireless communication. In particular, the dual-band IC and the dual-band MARC models the two aspects in a mm-wave and microwave dual-band setting. As networks become denser, interference mitigation using relays will play a key role in improving the user performance, which can be effectively captured by the interference-relay channel (IRC), which is a more general model compared to the IC or the MARC.

In the basic model of the dual-band fading IRC, two sources communicate to their respective destinations over both bands with the help of a relay in the presence of fading. In

the dual-band IRC, the transmissions in the underlying microwave band mutually interfere at each destination, which can be captured by the conventional Gaussian IRC, while the transmissions in the mm-wave band can be modeled as point-to-point steerable links.

While it is apparent that characterizing the performance limits of the IRC may reveal useful practical insights, the capacity region of the single band fading IRC or the optimal encoding/decoding and relaying strategies are unknown in general except for a few specific cases [122,124]. Part of the difficulty in obtaining closed form capacity results lies in finding optimal source-relay codeword correlation in the presence of interference. Moreover, while conventional techniques from the literature on the Gaussian IC can be applied, it provides capacity for only a narrow set of channel parameters. Hence, new techniques are needed to characterize the performance limits of single band and dual-band IRC.

Full duplex communication, which has the potential of doubling the spectral efficiency for communication by communicating simultaneously in both directions, has been subject to much investigation [151], and is now gradually becoming viable in practice [152]. In the related two-way relay channel (TWRC) [153], two terminals communicate simultaneously in both directions with the help of one relay. The dual-band TWRC and its multi-terminal extension can effectively model a scenario of relay-assisted full-duplex communication between many pairs of interacting terminals over an integrated mm-wave/microwave dual-band. In such a dual-band TWRC, transmissions in the underlying microwave band can be modeled as a Gaussian TWRC, while those in the mm-wave band can be modeled as point-to-point links. It is worth noting that similar to the relay channel, the capacity of the Gaussian TWRC is also unknown in general, and hence novel techniques may be needed to characterize the capacity of single and dual-band TWRC.

The dual-band IRC and the dual-band TWRC are two basic models that incorporate relay-cooperation, interference mitigation, and full-duplex communication in a dual-band setting, all of which are important aspects in current wireless communication practices. Therefore, studying the performance limits of these channels is an important and interesting further step towards characterizing practical insights that are applicable to a wide range of wireless settings.

Optimal Encoding/Decoding Strategies for Single Band Channels

While the primary focus of the thesis was on dual-band channels, interference mitigation in conventional networks is still an important open problem. Hence, in Chapter 5 we

studied the multiple-access interference channel (MAIC) over the microwave band where two multiple-access channels mutually interfere. While our work provides improved bounds on the sum-rate performance of the MAIC, its capacity region is still unknown in general. Since the MAIC presents a more general setting compared to the two-user IC, optimal or approximately optimal encoding/decoding strategies for the MAIC are expected to provide valuable insights into effective interference mitigation strategies for practical microwave cellular networks. Therefore, characterizing capacity or approximate capacity results for the MAIC is another potentially important future work which may require revisiting the approaches for outer bounds and achievable strategy.

References

[1] T. S. Rappaport, J. N. Murdock, and F. Gutierrez, "State of the art in 60-GHz integrated circuits and systems for wireless communications," *Proceedings of the IEEE*, vol. 99, no. 8, pp. 1390–1436, Aug 2011.

[2] "Spatial channel model for multiple input multiple output (MIMO) simulations (Release 10)," *Standard 3GPP TR 25.996*, March 2011.

[3] "Guidelines for evaluation of radio interference technologies for IMT-Advanced," *Standard ITU-R M.2135*, 2008.

[4] P. Rysavy, "Transition to 4G: 3GPP Broadband Evolution to IMT-Advanced (4G)," *Rysavy Research*, 2010.

[5] "Long Term HSPA Evolution: Mobile broadband evolution beyond 3GPP Release 10, Espoo, Finland," *Nokia Siemens Networks*, 2010.

[6] "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 20172022 White Paper," Cisco, Tech. Rep., Feb. 2019.

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

[8] A. Dytso, D. Tuninetti, and N. Devroye, "Interference as noise: Friend or foe?" *IEEE Transactions on Information Theory*, vol. 62, no. 6, pp. 3561–3596, 2016.

[9] H. F. Chong, M. Motani, H. K. Garg, and H. E. Gamal, "On the Han-Kobayashi region for the interference channel,," *IEEE Transactions on Information Theory*, 2006.

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

[11] X. Shang, G. Kramer, and B. Chen, "A new outer bound and the noisy-interference sum-rate capacity for Gaussian interference channels," *IEEE Transactions on Information Theory*, vol. 55, no. 2, pp. 689–699, Feb 2009.

[12] A. Chaaban, A. Sezgin, B. Bandemer, and A. Paulraj, "On Gaussian multiple access channels with interference: Achievable rates and upper bounds," in *Multiple Access Communications*. Springer Berlin Heidelberg, 2011, pp. 87–96.

[13] L. Pinals, A. A. A. Haija, and M. Vu, "Link regime and power savings of decode-forward relaying in fading channels," *IEEE Transactions on Communications*, vol. 64, no. 3, pp. 931–946, Mar. 2016.

[14] R. U. Nabar, H. Bolcskei, and F. W. Kneubuhler, "Fading relay channels: performance limits and space-time signal design," *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 6, pp. 1099–1109, Aug 2004.

[15] A. S. Avestimehr and D. N. C. Tse, "Outage capacity of the fading relay channel in the low-SNR regime," *IEEE Transactions on Information Theory*, vol. 53, no. 4, pp. 1401–1415, April 2007.

[16] T. M. Cover and A. E. Gamal, "Capacity theorems for the relay channel," *IEEE Transactions on Information Theory*, vol. 25, no. 5, pp. 572–584, Sep. 1979.

[17] A. S. Avestimehr, S. N. Diggavi, and D. N. C. Tse, "Wireless network information flow: A deterministic approach," *IEEE Transactions on Information Theory*, vol. 57, no. 4, pp. 1872–1905, April 2011.

[18] G. Kramer and A. J. van Wijngaarden, "On the white Gaussian multiple-access relay channel," in *Proc. IEEE ISIT*, Sorrento, Italy, Jun. 2000, p. 40.

[19] L. Sankar, N. B. Mandayam, and H. V. Poor, "On the sum-capacity of degraded Gaussian multiple-access relay channels," *IEEE Transactions on Information Theory*, vol. 55, no. 12, pp. 5394–5411, Dec. 2009.

[20] "LTE-A 4G Solution, Stockholm, Sweden," *Ericsson.*, April 2011.

[21] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, "Millimeter wave mobile communications for 5G cellular: It will work!" *IEEE Access*, vol. 1, pp. 335–349, 2013.

[22] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What will 5G be?" *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1065–1082, June 2014.

[23] O. Semiari, W. Saad, and M. Bennis, "Joint millimeter wave and microwave resources allocation in cellular networks with dual-mode base stations," *IEEE Transactions on Wireless Communications*, vol. 16, no. 7, pp. 4802–4816, July 2017.

[24] Z. Qingling and J. Li, "Rain attenuation in millimeter wave ranges," in *2006 7th International Symposium on Antennas, Propagation EM Theory*, Oct 2006, pp. 1–4.

[25] R. J. Humpleman and P. A. Watson, "Investigation of attenuation by rainfall at 60 GHz," *Proceedings of the Institution of Electrical Engineers*, vol. 125, no. 2, pp. 85–91, February 1978.

[26] S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter-wave cellular wireless networks: Potentials and challenges," *Proceedings of the IEEE*, vol. 102, no. 3, pp. 366–385, March 2014.

[27] T. S. Rappaport, G. R. MacCartney, M. K. Samimi, and S. Sun, "Wideband millimeter-wave propagation measurements and channel models for future wireless communication system design," *IEEE Transactions on Communications*, vol. 63, no. 9, pp. 3029–3056, Sept 2015.

[28] "IEEE Standard for Information technology–Telecommunications and information exchange between systems–Local and metropolitan area networks–Specific requirements-Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 3: Enhancements for Very High Throughput in the 60 GHz Band," *IEEE Std 802.11ad-2012 (Amendment to IEEE Std 802.11-2012, as amended by IEEE Std 802.11ae-2012 and IEEE Std 802.11aa-2012)*, pp. 1–628, Dec 2012.

[29] S. Singh, F. Ziliotto, U. Madhow, E. Belding, and M. Rodwell, "Blockage and directivity in 60 GHz wireless personal area networks: from cross-layer model to multihop

mac design," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 8, pp. 1400–1413, October 2009.

[30] S. Singh, R. Mudumbai, and U. Madhow, "Interference analysis for highly directional 60-GHz mesh networks: The case for rethinking medium access control," *IEEE/ACM Transactions on Networking*, vol. 19, no. 5, pp. 1513–1527, Oct 2011.

[31] Y. Li, J. Luo, W. Xu, N. Vucic, E. Pateromichelakis, and G. Caire, "A joint scheduling and resource allocation scheme for millimeter wave heterogeneous networks," in *2017 IEEE Wireless Communications and Networking Conference (WCNC)*, March 2017, pp. 1–6.

[32] Y. Niu, C. Gao, Y. Li, L. Su, D. Jin, Y. Zhu, and D. O. Wu, "Energy-efficient scheduling for mmwave backhauling of small cells in heterogeneous cellular networks," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 3, pp. 2674–2687, March 2017.

[33] H. Xu, V. Kukshya, and T. S. Rappaport, "Spatial and temporal characteristics of 60-GHz indoor channels," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 3, pp. 620–630, Apr 2002.

[34] T. Manabe, Y. Miura, and T. Ihara, "Effects of antenna directivity and polarization on indoor multipath propagation characteristics at 60 GHz," *IEEE Journal on Selected Areas in Communications*, vol. 14, no. 3, pp. 441–448, Apr 1996.

[35] T. Manabe, K. Sato, H. Masuzawa, K. Taira, T. Ihara, Y. Kasashima, and K. Yamaki, "Polarization dependence of multipath propagation and high-speed transmission characteristics of indoor millimeter-wave channel at 60 GHz," *IEEE Transactions on Vehicular Technology*, vol. 44, no. 2, pp. 268–274, May 1995.

[36] E. Ben-Dor, T. S. Rappaport, Y. Qiao, and S. J. Lauffenburger, "Millimeter-wave 60 GHz outdoor and vehicle AOA propagation measurements using a broadband channel sounder," in *Global Telecommunications Conference (GLOBECOM 2011), 2011 IEEE*, Dec 2011, pp. 1–6.

[37] Y. H. Ezzeldin, M. Cardone, C. Fragouli, and G. Caire, "Gaussian 1-2-1 networks: Capacity results for mmwave communications," *CoRR*, vol. abs/1801.02553, 2018. [Online]. Available: http://arxiv.org/abs/1801.02553

[38] Y. Li, J. Luo, R. A. Stirling-Gallacher, and G. Caire, "Integrated access and backhaul optimization for millimeter wave heterogeneous networks," *CoRR*, vol. abs/1901.04959, 2019. [Online]. Available: http://arxiv.org/abs/1901.04959

[39] Y. H. Ezzeldin, M. Cardone, C. Fragouli, and G. Caire, "Polynomial-time capacity calculation and scheduling for half-duplex 1-2-1 networks," *CoRR*, vol. abs/1901.02933, 2019. [Online]. Available: http://arxiv.org/abs/1901.02933

[40] G. K. Agarwal, Y. H. Ezzeldin, M. Cardone, and C. Fragouli, "Secure communication over 1-2-1 networks," *CoRR*, vol. abs/1801.03061, 2018. [Online]. Available: http://arxiv.org/abs/1801.03061

[41] H. Zhao, R. Mayzus, S. Sun, M. Samimi, J. K. Schulz, Y. Azar, K. Wang, G. N. Wong, F. Gutierrez, and T. S. Rappaport, "28 GHz millimeter wave cellular communication measurements for reflection and penetration loss in and around buildings in New York City," in *Communications (ICC), 2013 IEEE International Conference on*. IEEE, 2013, pp. 5163–5167.

[42] A. V. Alejos, M. G. Sanchez, and I. Cuinas, "Measurement and analysis of propagation mechanisms at 40 GHz: Viability of site shielding forced by obstacles," *IEEE Transactions on Vehicular Technology*, vol. 57, no. 6, pp. 3369–3380, Nov 2008.

[43] C. R. Anderson and T. S. Rappaport, "In-building wideband partition loss measurements at 2.5 and 60 GHz," *IEEE Transactions on Wireless Communications*, vol. 3, no. 3, pp. 922–928, May 2004.

[44] J. S. Lu, D. Steinbach, P. Cabrol, and P. Pietraski, "Modeling human blockers in millimeter wave radio links," *ZTE Communications*, vol. 10, no. 4, pp. 23–28, Dec 2012.

[45] S. Collonge, G. Zaharia, and G. E. Zein, "Influence of the human activity on wideband characteristics of the 60 GHz indoor radio channel," *IEEE Transactions on Wireless Communications*, vol. 3, no. 6, pp. 2396–2406, Nov 2004.

[46] Y. Azar, G. N. Wong, K. Wang, R. Mayzus, J. K. Schulz, H. Zhao, F. Gutierrez, D. Hwang, and T. S. Rappaport, "28 GHz propagation measurements for outdoor cellular communications using steerable beam antennas in New York city," in *2013 IEEE International Conference on Communications (ICC)*, June 2013, pp. 5143–5147.

[47] T. S. Rappaport, F. Gutierrez, E. Ben-Dor, J. N. Murdock, Y. Qiao, and J. I. Tamir, "Broadband millimeter-wave propagation measurements and models using adaptive-beam antennas for outdoor urban cellular communications," *IEEE Transactions on Antennas and Propagation*, vol. 61, no. 4, pp. 1850–1859, April 2013.

[48] J. N. Murdock, E. Ben-Dor, Y. Qiao, J. I. Tamir, and T. S. Rappaport, "A 38 GHz cellular outage study for an urban outdoor campus environment," in *2012 IEEE Wireless Communications and Networking Conference (WCNC)*, April 2012, pp. 3085–3090.

[49] T. S. Rappaport, Y. Qiao, J. I. Tamir, J. N. Murdock, and E. Ben-Dor, "Cellular broadband millimeter wave propagation and angle of arrival for adaptive beam steering systems (invited paper)," in *Radio and Wireless Symposium (RWS), 2012 IEEE*, 2012, pp. 151–154.

[50] 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 Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1164–1179, June 2014.

[51] O. Galinina, S. Andreev, M. Gerasimenko, Y. Koucheryavy, N. Himayat, S. P. Yeh, and S. Talwar, "Capturing spatial randomness of heterogeneous cellular/WLAN deployments with dynamic traffic," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1083–1099, June 2014.

[52] T. Zhou, Y. Huang, W. Huang, S. Li, Y. Sun, and L. Yang, "QoS-aware user association for load balancing in heterogeneous cellular networks," in *2014 IEEE 80th Vehicular Technology Conference (VTC2014-Fall)*, Sept 2014, pp. 1–5.

[53] S. Corroy, L. Falconetti, and R. Mathar, "Cell association in small heterogeneous networks: Downlink sum rate and min rate maximization," in *2012 IEEE Wireless Communications and Networking Conference (WCNC)*, April 2012, pp. 888–892.

[54] S. Singh and J. G. Andrews, "Joint resource partitioning and offloading in heterogeneous cellular networks," *IEEE Transactions on Wireless Communications*, vol. 13, no. 2, pp. 888–901, February 2014.

[55] "5G Radio Access: Requirements, Concept and Technologies," NTT DOCOMO, INC, Tech. Rep., Jul. 2014.

[56] "3GPP RWS-120010: Requirements, candidate solutions & technology roadmap for LTE Rel-12 onward," *DOCOMO*, June 2012.

[57] Y. Kishiyama, A. Benjebbour, T. Nakamura, and H. Ishii, "Future steps of LTE-A: evolution toward integration of local area and wide area systems," *IEEE Wireless Communications*, vol. 20, no. 1, pp. 12–18, February 2013.

[58] H. Ishii, Y. Kishiyama, and H. Takahashi, "A novel architecture for LTE-B :C-plane/U-plane split and phantom cell concept," in *2012 IEEE Globecom Workshops*, Dec 2012, pp. 624–630.

[59] Y. Niu, Y. Li, D. Jin, L. Su, and A. V. Vasilakos, "A survey of millimeter wave communications (mmwave) for 5G: opportunities and challenges," *Wireless Networks*, vol. 21, no. 8, pp. 2657–2676, 2015.

[60] K. Zheng, L. Zhao, J. Mei, M. Dohler, W. Xiang, and Y. Peng, "10 Gb/s Hetsnets with Millimeter-wave communications: Access and networking - challenges and protocols," *IEEE Communications Magazine*, vol. 53, no. 1, pp. 222–231, Jan. 2015.

[61] Z. Pi and F. Khan, "An introduction to millimeter-wave mobile broadband systems," *IEEE Communications Magazine*, vol. 49, no. 6, pp. 101–107, June 2011.

[62] J. Qiao, X. S. Shen, J. W. Mark, Q. Shen, Y. He, and L. Lei, "Enabling device-to-device communications in millimeter-wave 5G cellular networks," *IEEE Communications Magazine*, vol. 53, no. 1, pp. 209–215, January 2015.

[63] ——, "Enabling device-to-device communications in millimeter-wave 5G cellular networks," *IEEE Communications Magazine*, vol. 53, no. 1, pp. 209–215, January 2015.

[64] H. Mehrpouyan, M. Matthaiou, R. Wang, G. K. Karagiannidis, and Y. Hua, "Hybrid millimeter-wave systems: a novel paradigm for hetnets," *IEEE Communications Magazine*, vol. 53, no. 1, pp. 216–221, January 2015.

[65] S. Majhi and P. Mitran, "On the capacity of a class of dual-band interference channels," in *Proceedings of the 2016 IEEE International Symposium on Information Theory, Barcelona, Spain*, Jul. 2016, pp. 2759–2763.

[66] M. Hashemi, C. E. Koksal, and N. B. Shroff, "Out-of-band millimeter wave beamforming and communications to achieve low latency and high energy efficiency in 5G

systems," *IEEE Transactions on Communications*, vol. 66, no. 2, pp. 875–888, Feb. 2018.

[67] Intel Corporation, "Intel Announces Worlds First Global 5G Modem," *Report*, Jan. 2017. [Online]. Available: https://newsroom.intel.com/newsroom/wp-content/uploads/sites/11/2017/01/5G-modem-fact-sheet.pdf

[68] A. Ali, N. Gonzlez-Prelcic, and R. W. Heath, "Estimating millimeter wave channels using out-of-band measurements," in *2016 Information Theory and Applications Workshop (ITA)*, Jan 2016, pp. 1–6.

[69] A. Ali, N. Gonzlez-Prelcic, and R. W. Heath, "Millimeter wave beam-selection using out-of-band spatial information," *IEEE Transactions on Wireless Communications*, vol. 17, no. 2, pp. 1038–1052, Feb 2018.

[70] H. Elshaer, M. N. Kulkarni, F. Boccardi, J. G. Andrews, and M. Dohler, "Downlink and uplink cell association with traditional macrocells and millimeter wave small cells," *IEEE Transactions on Wireless Communications*, vol. 15, no. 9, pp. 6244–6258, Sept 2016.

[71] Qualcomm Corporation, "Qualcomm Delivers Breakthrough 5G NR mmWave and Sub-6 GHz RF Modules for Mobile Devices," *Press Release*, Jul. 2018.

[72] W. Hong, Z. H. Jiang, C. Yu, J. Zhou, P. Chen, Z. Yu, H. Zhang, B. Yang, X. Pang, M. Jiang, Y. Cheng, M. K. T. Al-Nuaimi, Y. Zhang, J. Chen, and S. He, "Multibeam antenna technologies for 5G wireless communications," *IEEE Transactions on Antenna and Propagation*, vol. 65, no. 12, pp. 6231–6249, Dec. 2017.

[73] Intel Corporation, "Anokiwave Introduces 26 GHz 5G mmW Reconfigurable 256-Element Active Antenna Array," *Report*, Feb. 2018. [Online]. Available: https://www.anokiwave.com/company/company-news/releases/awa_0142.html

[74] J. Brady, N. Behdad, and A. M. Sayeed, "Beamspace MIMO for millimeter-wave communications: System architecture, modeling, analysis, and measurements," *IEEE Transactions on Antenna and Propagation*, vol. 61, no. 7, pp. 3814–3827, Jul. 2013.

[75] Q. Yang, Y. Ban, K. Kang, C. Sim, and G. Wu, "Siw multibeam array for 5G mobile devices," *IEEE Access*, vol. 4, pp. 2788–2796, Jun. 2016.

[76] Z. Wei, X. Zhu, S. Sun, and Y. Huang, "Energy-efficiency-oriented cross-layer resource allocation for multiuser full-duplex decode-and-forward indoor relay systems at 60 GHz," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 12, pp. 3366–3379, Dec. 2016.

[77] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Transactions on Information Theory*, vol. 51, no. 9, pp. 3037–3063, Sep. 2005.

[78] Verizon, "More Verizon 5G Ultra Wideband service in more places with more smartphones," *Report*, Sep. 2019. [Online]. Available: https://www.verizon.com/about/news/verizon-5g-ultra-wideband-more-places-smartphones

[79] M. R. Williamson, G. E. Athanasiadou, and A. R. Nix, "Investigating the effects of antenna directivity on wireless indoor communication at 60 GHz," in *Personal, Indoor and Mobile Radio Communications, 1997. Waves of the Year 2000. PIMRC '97., The 8th IEEE International Symposium on*, vol. 2, Sept 1997, pp. 635–639 vol.2.

[80] H. Shokri-Ghadikolaei, C. Fischione, G. Fodor, P. Popovski, and M. Zorzi, "Millimeter wave cellular networks: A MAC layer perspective," *IEEE Transactions on Communications*, vol. 63, no. 10, pp. 3437–3458, Oct 2015.

[81] H. Shokri-Ghadikolaei and C. Fischione, "The transitional behavior of interference in millimeter wave networks and its impact on medium access control," *IEEE Transactions on Communications*, vol. 64, no. 2, pp. 723–740, Feb 2016.

[82] H. Shokri-Ghadikolaei, C. Fischione, P. Popovski, and M. Zorzi, "Design aspects of short range millimeter wave networks: A MAC layer perspective," *arXiv preprint arXiv:1509.07538*, 2015.

[83] T. Rappaport, *Wireless communications: Principles and practice*, 2nd ed., ser. Prentice Hall communications engineering and emerging technologies series. Prentice Hall, 2002, includes bibliographical references and index.

[84] R. C. Daniels, J. N. Murdock, T. S. Rappaport, and R. W. Heath, "60 GHz wireless: Up close and personal," *IEEE Microwave Magazine*, vol. 11, no. 7, pp. 44–50, Dec 2010.

[85] T. Zwick, T. J. Beukema, and H. Nam, "Wideband channel sounder with measurements and model for the 60 GHz indoor radio channel," *IEEE Transactions on Vehicular Technology*, vol. 54, no. 4, pp. 1266–1277, July 2005.

[86] S. Singh, F. Ziliotto, U. Madhow, E. M. Belding, and M. Rodwell, "Blockage and directivity in 60 GHz wireless personal area networks," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 8, pp. 1400–1413, Oct. 2009.

[87] G. M. Stamatelos and D. D. Falconer, "Packet diversity techniques for a broadband atm-oriented radio indoor environment," in *Communications, 1994. ICC '94, SUPERCOMM/ICC '94, Conference Record, 'Serving Humanity Through Communications.' IEEE International Conference on*, May 1994, pp. 566–570 vol.1.

[88] M. Samimi, K. Wang, Y. Azar, G. N. Wong, R. Mayzus, H. Zhao, J. K. Schulz, S. Sun, F. Gutierrez, and T. S. Rappaport, "28 GHz angle of arrival and angle of departure analysis for outdoor cellular communications using steerable beam antennas in New York City," in *Vehicular Technology Conference (VTC Spring), 2013 IEEE 77th*, June 2013, pp. 1–6.

[89] M. Marcus and B. Pattan, "Millimeter wave propagation: spectrum management implications," *IEEE Microwave Magazine*, vol. 6, no. 2, pp. 54–62, June 2005.

[90] "3GPP Release 15 Overview," *IEEE Spectrum*, Apr. 2019. [Online]. Available: https://spectrum.ieee.org/telecom/wireless/3gpp-release-15-overview

[91] T. S. Rappaport, J. N. Murdock, and F. Gutierrez, "State of the art in 60-GHz integrated circuits and systems for wireless communications," *Proceedings of the IEEE*, vol. 99, no. 8, pp. 1390–1436, Aug 2011.

[92] S. Niknam, A. A. Nasir, H. Mehrpouyan, and B. Natarajan, "A multiband OFDMA heterogeneous network for millimeter wave 5G wireless applications," *IEEE Access*, vol. 4, pp. 5640–5648, 2016.

[93] Y. Chen, B. Ai, Y. Niu, R. He, Z. Zhong, and Z. Han, "Resource allocation for device-to-device communications in multi-cell multi-band heterogeneous cellular networks," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 5, pp. 4760–4773, May 2019.

[94] O. Semiari, W. Saad, and M. Bennis, "Context-aware scheduling of joint millimeter wave and microwave resources for dual-mode base stations," in *2016 IEEE International Conference on Communications (ICC)*, May 2016, pp. 1–6.

[95] M. Shi, K. Yang, Z. Han, and D. Niyato, "Coverage analysis of integrated sub-6 GHz-mmwave cellular networks with hotspots," *CoRR*, 2019. [Online]. Available: https://arxiv.org/pdf/1909.06975.pdf

[96] C. Shannon, "Two-way communication channels," *Mathematical Statistics and Probability, Proc. 4th Berkeley Symp. on*, vol. 1, no. 5, pp. 611–644, Sep 1961.

[97] A. E. Gamal and Y.-H. Kim, *Network Information Theory*. New York, NY, USA: Cambridge University Press, 2012.

[98] M. Costa and A. Gamal, "The capacity region of the discrete memoryless interference channel with strong interference," *IEEE Transactions on Information Theory*, vol. 33, no. 5, pp. 710–711, Sep. 1987.

[99] A. Carleial, "A case where interference does not reduce capacity," *IEEE Transactions on Information Theory*, vol. 21, no. 5, pp. 569–570, Sep. 1975.

[100] I. Sason, "On achievable rate regions for the Gaussian interference channel," *IEEE Transactions on Information Theory*, vol. 50, pp. 1345–1356, Jun 2004.

[101] H. Sato, "The capacity of the Gaussian interference channel under strong interference," *IEEE Transactions on Information Theory*, vol. 27, no. 6, pp. 786–788, Nov. 1981.

[102] A. Motahari and A. Khandani, "Capacity bounds for the Gaussian interference channel," *IEEE Transactions on Information Theory*, vol. 55, no. 2, pp. 620–643, Feb 2009.

[103] V. Annapureddy and V. Veeravalli, "Gaussian interference networks: Sum capacity in the low-interference regime and new outer bounds on the capacity region," *IEEE Transactions on Information Theory*, vol. 55, no. 7, pp. 3032–3050, July 2009.

[104] J. G. Proakis and M. Salehi, *Digital Communications*. New York, NY, USA: McGraw-Hill, 2008.

[105] S. T. Chung and J. Cioffi, "The capacity region of frequency-selective Gaussian interference channels under strong interference," *IEEE Transactions on Communications*, vol. 55, no. 9, pp. 1812–1821, Sep. 2007.

[106] X. Shang, B. Chen, G. Kramer, and H. Poor, "Noisy-interference sum-rate capacity of parallel Gaussian interference channels," *IEEE Transactions on Information Theory*, vol. 57, no. 1, pp. 210–226, Jan. 2011.

[107] S. W. Choi and S. Y. Chung, "On the separability of parallel Gaussian interference channels," in *In Proceedings of the IEEE International Symposium on Information Theory, Seoul, South Korea*, July 2009, pp. 2592–2596.

[108] V. Cadambe and S. Jafar, "Parallel Gaussian interference channels are not always separable," *IEEE Transactions on Information Theory*, vol. 55, no. 9, pp. 3983–3990, Sep. 2009.

[109] L. Sankar, X. Shang, E. Erkip, and H. Poor, "Ergodic two-user interference channels: Is separability optimal?" in *Communication, Control, and Computing, 2008 46th Annual Allerton Conference on*, Sep. 2008, pp. 723–729.

[110] L. Sankar, X. Shang, E. Erkip, and H. V. Poor, "Ergodic fading interference channels: Sum-capacity and separability," *IEEE Transactions on Information Theory*, vol. 57, no. 5, pp. 2605–2626, May 2011.

[111] W. Yu, G. Ginis, and J. M. Cioffi, "Distributed multiuser power control for digital subscriber lines," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 5, pp. 1105–1115, Jun 2002.

[112] H. Chen, Y. Ma, Z. Lin, Y. Li, and B. Vucetic, "Distributed power control in interference channels with QoS constraints and RF energy harvesting: A game-theoretic approach," *IEEE Transactions on Vehicular Technology*, vol. 65, pp. 10 063–10 069, 2016.

[113] R. Fritschek and G. Wunder, "On multiuser gain and the constant-gap sum capacity of the Gaussian interfering multiple access channel," *CoRR*, 2017. [Online]. Available: http://arxiv.org/abs/1705.04514

[114] W. Chang, S. Chung, and Y. H. Lee, "Gaussian relay channel capacity to within a fixed number of bits," *CoRR*, vol. abs/1011.5065, 2010. [Online]. Available: http://arxiv.org/abs/1011.5065

[115] A. Shah and A. M. Haimovich, "Performance analysis of optimum combining in wireless communications with Rayleigh fading and cochannel interference," *IEEE Transactions on Communications*, vol. 46, no. 4, pp. 473–479, 1998.

[116] W. Turin, R. Jana, C. Martin, and J. Winters, "Modeling wireless channel fading," in *IEEE 54th Vehicular Technology Conference. VTC Fall 2001. Proceedings (Cat. No. 01CH37211)*, vol. 3. IEEE, 2001, pp. 1740–1744.

[117] G. D. Durgin, T. S. Rappaport, and D. A. De Wolf, "New analytical models and probability density functions for fading in wireless communications," *IEEE Transactions on Communications*, vol. 50, no. 6, pp. 1005–1015, 2002.

[118] T. S. Rappaport *et al.*, *Wireless communications: principles and practice*. prentice hall PTR New Jersey, 1996, vol. 2.

[119] B. Rankov and A. Wittneben, "Spectral efficient protocols for half-duplex fading relay channels," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 2, pp. 379–389, February 2007.

[120] B. Sklar, "Rayleigh fading channels in mobile digital communication systems. I. Characterization," *IEEE Communications Magazine*, vol. 35, no. 9, pp. 136–146, Sep. 1997.

[121] S. Wu and Y. Bar-Ness, "OFDM systems in the presence of phase noise: Consequences and solutions," *IEEE Transactions on Communications*, vol. 52, no. 11, pp. 1988–1996, Nov. 2004.

[122] R. Dabora, "The capacity region of the fading interference channel with a relay in the strong interference regime," *IEEE Transactions on Information Theory*, vol. 58, no. 8, pp. 5172–5184, Aug. 2012.

[123] L. Sankar, Y. Liang, N. B. Mandayam, and H. V. Poor, "Fading multiple access relay channels: Achievable rates and opportunistic scheduling," *IEEE Transactions on Information Theory*, vol. 57, no. 4, pp. 1911–1931, Apr. 2011.

[124] H. E. Saffar, E. H. M. Alian, and P. Mitran, "Source-channel communication over phase-incoherent multiuser channels," *IEEE Transactions on Communications*, vol. 62, no. 8, pp. 2996–3003, Aug. 2014.

[125] J. Liang, X. Wang, and W. Zhang, "Capacity theorem and optimal power allocation for frequency division multiple-access relay channels," in *2011 IEEE Global Telecommunications Conference - GLOBECOM 2011*, Dec 2011, pp. 1–5.

[126] V. Tervo, X. Zhou, P. Lu, M. Juntti, and T. Matsumoto, "Power allocation for orthogonal multiple access relay channel allowing intra-link errors," in *European Wireless 2016; 22th European Wireless Conference*, May 2016, pp. 1–6.

[127] S. Schedler and V. Kühn, "Resource allocation for the multiple-access relay channels and OFDMA," *EURASIP Journal on Advances in Signal Processing*, vol. 2013, no. 1, p. 120, 2013.

[128] S. Majhi and P. Mitran, "On the capacity and the optimal sum-rate of a class of dual-band interference channels," *Entropy*, vol. 19, no. 9, 2017.

[129] X. Shang, B. Chen, G. Kramer, and H. V. Poor, "Capacity regions and sum-rate capacities of vector gaussian interference channels," *IEEE Transactions on Information Theory*, vol. 56, no. 10, pp. 5030–5044, 2010.

[130] S. N. Diggavi and T. M. Cover, "The worst additive noise under a covariance constraint," *IEEE Transactions on Information Theory*, vol. 47, no. 7, pp. 3072–3081, Nov 2001.

[131] J. Thomas, "Feedback can at most double Gaussian multiple access channel capacity (corresp.)," *IEEE Transactions on Information Theory*, vol. 33, no. 5, pp. 711–716, Sep. 1987.

[132] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge University Press, 2004.

[133] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming," http://cvxr.com/cvx, Mar. 2014.

[134] S. Majhi and P. Mitran, "Dual-band fading multiple access relay channels," *IEEE Transactions on Communications*, vol. 67, no. 9, pp. 6053–6070, Sep. 2019.

[135] B. Zhai, A. Tang, C. Huang, C. Han, and X. Wang, "Antenna subarray management for hybrid beamforming in millimeter-wave mesh backhaul networks," *Nano Communication Networks*, vol. 19, pp. 92 – 101, 2019.

[136] Q. Xue, X. Fang, and M. Xiao, "Beam management for millimeter wave beamspace MU-MIMO systems," in *2017 IEEE/CIC International Conference on Communications in China (ICCC)*, Oct. 2017, pp. 1–6.

[137] A. Sabharwal, D. Dash, and S. Diggavi, "Compound Gaussian multiple access channels with noisy feedback," in *2008 46th Annual Allerton Conference on Communication, Control, and Computing*, Sep. 2008, pp. 887–894.

[138] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley, 1991.

[139] G. Kramer, "Outer bounds on the capacity of Gaussian interference channels," *IEEE Transactions on Information Theory*, vol. 50, pp. 581–586, Mar 2004.

[140] S. Majhi and P. Mitran, "On the capacity of Gaussian multiple-access interference channels," in *Proceedings of the 2019 IEEE International Symposium on Information Theory, Paris, France*, Jul. 2019.

[141] N. Modi, "Capacity region of two interfering mimo multiple access channels with strong interference," Ph.D. dissertation, North Dakota State University, 2015.

[142] V. R. Cadambe and S. A. Jafar, "Interference alignment and the degrees of freedom of wireless x networks," *IEEE Transactions on Information Theory*, vol. 55, no. 9, pp. 3893–3908, 2009.

[143] C. Geng, H. Sun, and S. A. Jafar, "On the optimality of treating interference as noise: General message sets," *IEEE Transactions on Information Theory*, vol. 61, no. 7, pp. 3722–3736, 2015.

[144] J. Nam, "On the high-SNR capacity of the Gaussian interference channel and new capacity bounds," *IEEE Transactions on Information Theory*, vol. 63, no. 8, pp. 5266–5285, Aug 2017.

[145] Y. Pang and M. K. Varanasi, "A unified theory of multiple-access and interference channels via approximate capacity regions for the MAC-IC-MAC," *IEEE Transactions on Information Theory*, vol. 65, no. 3, pp. 1898–1920, March 2019.

[146] P. Wang, H. Jiang, W. Zhuang, and H. V. Poor, "Redefinition of max-min fairness in multi-hop wireless networks," *IEEE Transactions on Wireless Communications*, vol. 7, no. 12, p. 4786–4791, 2008.

[147] B. Radunovic and J.-Y. Le Boudec, "Why max-min fairness is not suitable for multi-hop wireless networks," Ecole Polytechnique Federale de Lausanne, Switzerland, Tech. Rep., 2003.

[148] T. Bu, L. Li, and R. Ramjee, "Generalized proportional fair scheduling in third generation wireless data networks," in *Proceedings IEEE INFOCOM 2006. 25TH IEEE International Conference on Computer Communications*, April 2006, pp. 1–12.

[149] V. K. Lau, "Proportional fair space-time scheduling for wireless communications," *IEEE Transactions on Communications*, vol. 53, no. 8, pp. 1353–1360, 2005.

[150] H. Kim and Y. Han, "A proportional fair scheduling for multicarrier transmission systems," *IEEE Communications letters*, vol. 9, no. 3, pp. 210–212, 2005.

[151] N. V. Shende, Ö. Gürbüz, and E. Erkip, "Half-duplex or full-duplex communications: Degrees of freedom analysis under self-interference," *IEEE Transactions on Wireless Communications*, vol. 17, no. 2, pp. 1081–1093, 2017.

[152] A. K. Khandani, "Two-way (true full-duplex) wireless," in *2013 13th Canadian Workshop on Information Theory*, June 2013, pp. 33–38.

[153] B. Rankov and A. Wittneben, "Achievable rate regions for the two-way relay channel," in *2006 IEEE International Symposium on Information Theory*, July 2006, pp. 1668–1672.

[154] J. Nachbar, "Finite dimensional optimization part I: The KKT theorem," St. Louis, Missouri, USA, March 2018.

[155] T. Hoheisel, "Mathematical programs with vanishing constraints," Ph.D. dissertation, Wurzburg, Germany, 2009.

[156] G. Wachsmuth, "On LICQ and the uniqueness of Lagrange multipliers," *Operations Research Letters*, vol. 41, no. 1, pp. 78 – 80, 2013. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0167637712001459

[157] M. Costa, "On the Gaussian interference channel," *IEEE Transactions on Information Theory*, vol. 31, no. 5, pp. 607–615, Sep. 1985.

[158] E. J. Violette, R. H. Espeland, and G. R. Hand, "Millimeter-wave urban and suburban propagation measurements using narrow and wide bandwidth channel probes," *NASA STI/Recon Technical Report N*, vol. 86, Nov. 1985.

[159] M. Rebato, F. Boccardi, M. Mezzavilla, S. Rangan, and M. Zorzi, "Hybrid spectrum sharing in mmwave cellular networks," *IEEE Transactions on Cognitive Communications and Networking*, vol. 3, no. 2, pp. 155–168, June 2017.

[160] O. Semiari, W. Saad, and M. Bennis, "Downlink cell association and load balancing for joint millimeter wave-microwave cellular networks," in *2016 IEEE Global Communications Conference (GLOBECOM)*, Dec 2016, pp. 1–6.

[161] J. Park, S. L. Kim, and J. Zander, "Tractable resource management with uplink decoupled millimeter-wave overlay in ultra-dense cellular networks," *IEEE Transactions on Wireless Communications*, vol. 15, no. 6, pp. 4362–4379, June 2016.

APPENDICES

Appendix A

List of Publications

The following is a list of journal publications and conference papers resulting from the work contained in this thesis:

Journal Articles

- **S. Majhi** and P. Mitran, On the capacity and the optimal sum-rate of a class of dual-band interference channels, Entropy, vol. 19, no. 9, 2017.
- **S. Majhi** and P. Mitran, Dual-band fading multiple access relay channels, IEEE Transactions on Communications, vol. 67, no. 9, pp. 6053-6070, 2019.
- **S. Majhi** and P. Mitran, On the high-SNR capacity of the Gaussian multiple-access interference channels, (in preparation).

Conference Proceedings

- **S. Majhi** and P. Mitran, On the capacity of a class of dual-band interference channels, in Proceedings of the 2016 IEEE International Symposium on Information Theory, Barcelona, Spain, Jul. 2016, pp. 2759-2763.
- **S. Majhi** and P. Mitran, On the capacity of Gaussian multiple-access interference channels, in Proceedings of the 2019 IEEE International Symposium on Information Theory, Paris, France, Jul. 2019.

Appendix B

Appendices for Chapter 3

B.1 Proof of Theorem 3.1

Outer bound: We derive the bounds on R_1 and R_2 as follows. If source S_1 transmits the message $M_k, k = 1, 2$, we have from Fano's inequality

$$\begin{aligned}
nR_1 &= H(M_1) \\
&= I(M_1; Y_1^n, \hat{Y}_1^{n_1}, \bar{Y}_1^{n_2}) + H(M_1|Y_1^n, \hat{Y}_1^{n_1}, \bar{Y}_1^{n_2}) \\
&\leq I(X_1^n, \hat{X}_1^{n_1}, \bar{X}_1^{n_2}; Y_1^n, \hat{Y}_1^{n_1}, \bar{Y}_1^{n_2}) + n\epsilon_n \\
&\overset{(a)}{=} I(X_1^n, \hat{X}_1^{n_1}; Y_1^n, \hat{Y}_1^{n_1}) + I(\bar{X}_1^{n_2}; Y_1^n, \hat{Y}_1^{n_1}|X_1^n, \hat{X}_1^{n_1}) \\
&\quad + I(\bar{X}_1^{n_2}; \bar{Y}_1^{n_2}|Y_1^n, \hat{Y}_1^{n_1}) + I(X_1^n, \hat{X}_1^{n_1}; \bar{Y}_1^{n_2}|\bar{X}_1^{n_2}, Y_1^n, \hat{Y}_1^{n_1}) + n\epsilon_n \\
&\overset{(b)}{=} I(X_1^n, \hat{X}_1^{n_1}; Y_1^n, \hat{Y}_1^{n_1}) + I(\bar{X}_1^{n_2}; \bar{Y}_1^{n_2}|Y_1^n, \hat{Y}_1^{n_1}) + n\epsilon_n \\
&= I(X_1^n, \hat{X}_1^{n_1}; Y_1^n, \hat{Y}_1^{n_1}) + \sum_{\ell=1}^{n_2} h(\bar{Y}_{1\ell}|\bar{Y}_1^{\ell-1}, Y_1^n, \hat{Y}_1^{n_1}) - h(\bar{Y}_{1\ell}|\bar{Y}_1^{\ell-1}, \bar{X}_1^{n_1}, Y_1^n, \hat{Y}_1^{n_1}) + n\epsilon_n \\
&\overset{(c)}{\leq} I(X_1^n, \hat{X}_1^{n_1}; Y_1^n, \hat{Y}_1^{n_1}) + \sum_{\ell=1}^{n_2} h(d_1\bar{X}_{1\ell} + \bar{Z}_{1\ell}) - h(\bar{Z}_{1\ell}) + n\epsilon_n \\
&\overset{(d)}{\leq} I(X_1^n, \hat{X}_1^{n_1}; Y_1^n, \hat{Y}_1^{n_1}) + n_2\mathsf{C}(d_1^2\bar{P}_1) + n\epsilon_n, \qquad \text{(B.1)}
\end{aligned}$$

where (a) follows from the chain rule; (b) follows since the second term of (a) is zero due to the Markov chain (MC) $\bar{X}_1^{n_2} \multimap (X_1^n, \hat{X}_1^{n_1}) \multimap (Y_1^n, \hat{Y}_1^{n_1})$, and the fourth term of (a) is zero due to the MC $(Y_1^n, \hat{Y}_1^{n_1}, X_1^n, \hat{X}_1^{n_1}) \multimap \bar{X}_1^{n_2} \multimap \bar{Y}_1^{n_2}$; (c) follows since unconditioning does not reduce entropy, and also from the memoryless Gaussian model of the direct-links; and (d) follows since $\bar{X}_{1\ell} \sim \mathcal{N}(0, \bar{P}_1)$, i.i.d., maximizes the mutual information. Bounding R_2

similarly, and since $n_2 \leq \alpha_2 n$, we have

$$R_1 \leq r_1 + \alpha_2 \mathsf{C}(d_1^2 \bar{P}_1)$$
$$R_2 \leq r_2 + \alpha_2 \mathsf{C}(d_2^2 \bar{P}_2),$$

where (r_1, r_2) is an achievable rate pair in the underlying CLIC with BMF α_1, which is defined as the set of all non-negative (r_1, r_2) that satisfy

$$r_1 := \sup_{X_1^n, \hat{X}_1^{n_1}} \frac{1}{n} I(X_1^n, \hat{X}_1^{n_1}; Y_1^n, \hat{Y}_1^{n_1}), \quad r_2 := \sup_{X_2^n, \hat{X}_2^{n_1}} \frac{1}{n} I(X_2^n, \hat{X}_2^{n_1}; Y_2^n, \hat{Y}_2^{n_1})$$

for product distribution $p(x_1^n, \hat{x}_1^{n_1}) p(x_2^n, \hat{x}_2^{n_1})$ and $n_1 \leq \alpha_1 n$.

Achievability: We will code over t blocks of symbols together. We choose (n, n_1, n_2), and define $\mathbf{U}_k := (X_k^n, \hat{X}_k^{n_1}), k = 1, 2$, where $p(\mathbf{U}_1, \mathbf{U}_2) = p(\mathbf{U}_1) p(\mathbf{U}_2)$, and define $\bar{\mathbf{U}}_k := \bar{X}_k^{n_2}, k = 1, 2$. Suppose source S_k transmits $M_k \in \mathcal{M}_k, k = 1, 2$. Thus, to encode M_k over t blocks for the underlying CLIC, we generate 2^{tnR_k} i.i.d. sequences $\mathbf{U}_k^t(M_k)$, one for each $M_k \in \mathcal{M}_k$, where $\mathbf{U}_k^t(M_k)$ are distributed according to $p(\mathbf{u}_k^t) = \prod_{\ell=1}^t p(\mathbf{u}_{k,\ell}) = \prod_{\ell=1}^t p\big(x_{k,(\ell-1)n+1}^{n\ell} \hat{x}_{k,(\ell-1)n_1+1}^{n_1 \ell}\big), k = 1, 2$. Similarly, for the two direct-links, we generate 2^{tnR_k} i.i.d. sequences $\bar{\mathbf{U}}_k^t(M_k)$, distributed as $\prod_{\ell=1}^{n_2 t} p(\bar{x}_{k\ell})$, where $\bar{x}_{k\ell} \sim \mathcal{N}(0, \bar{P}_k)$, for $k = 1, 2$, and $\ell = 1, \ldots, n_2$.

Thus, to transmit $M_k, k = 1, 2$, $\mathbf{U}_1^t(M_1)$ and $\mathbf{U}_2^t(M_2)$ are transmitted through the CLIC, and $\bar{\mathbf{U}}_1^t(M_1)$ and $\bar{\mathbf{U}}_2^t(M_2)$ are transmitted through the two direct-links in the second band. Upon receiving the sequences $(Y_k^{nt}, \hat{Y}_k^{n_1 t})$ and $\bar{Y}_k^{n_2 t}$, destination D_k employs joint typical decoding to estimate the transmitted message as in [97, Chapter 4.3]. The probability of decoding error vanishes with $t \to \infty$, if

$$R_k < \frac{1}{n} I(X_k^n, \hat{X}_k^{n_1}; Y_k^n, \hat{Y}_k^{n_1}) + \frac{n_2}{n} \mathsf{C}(d_k^2 \bar{P}_k), \quad k = 1, 2, \tag{B.2}$$

which matches the upper bound on R_1 and R_2, as $n_1 \leq \alpha_1 n, n_2 \leq \alpha_2 n$, and $n \to \infty$. Finally, the capacity region of the CLIC with BMF α_1 is defined as the closure of the union of all sets of achievable rate pairs (R_1, R_2) that satisfy (B.2), where the union is taken over all product distributions, $p(x_1^n, \hat{x}_1^{n_1}) p(x_2^n, \hat{x}_2^{n_1})$, and for all positive integers n, n_1, such that $n \leq \alpha_1 n$.

B.2 Proof of Theorem 3.2

Outer bound: We derive the bound on R_1 first. Assuming that source S_1 transmits message M_1, we have

$$\begin{aligned}
nR_1 &= H(M_1) \\
&= I(M_1; Y_1^n, \hat{Y}_1^n) + H(M_1|Y_1^n, \hat{Y}_1^n) \\
&\overset{(a)}{\leq} I(X_1^n, \hat{X}_1^n; Y_1^n, \hat{Y}_1^n) + n\epsilon_n \\
&\leq I(X_1^n, \hat{X}_1^n; Y_1^n, \hat{Y}_1^n, X_2^n, \hat{X}_2^{n_1}) + n\epsilon_n \\
&\overset{(b)}{=} I(X_1^n, \hat{X}_1^n; Y_1^n, \hat{Y}_1^n|X_2^n, \hat{X}_2^{n_1}) + n\epsilon_n \\
&\overset{(c)}{=} I(X_1^n, \hat{X}_1^n; Y_1^n|X_2^n, \hat{X}_2^{n_1}) + I(X_1^n, \hat{X}_1^n; \hat{Y}_1^{n_1}|X_2^n, \hat{X}_2^{n_1}, Y_1^n) + n\epsilon_n \\
&\overset{(d)}{=} I(X_1^n; Y_1^n|X_2^n) + n\epsilon_n \\
&\overset{(e)}{\leq} \sum_{i=1}^{n} I(X_{1i}; Y_{1i}|X_{2i}) + n\epsilon_n \\
&= \frac{n}{n} \sum_{i=1}^{n} I(X_{1i}; Y_{1i}|X_{2i}) + n\epsilon_n \\
&\overset{(f)}{=} nI(X_1; Y_1|X_2, Q) + n\epsilon_n \\
&\overset{(g)}{\leq} n\frac{1}{n} \sum_{i=1}^{n} \frac{1}{2} \log\left(1 + P_{1i}\right) + n\epsilon_n \\
&\overset{(h)}{\leq} n\mathsf{C}(P_1) + n\epsilon_n,
\end{aligned} \tag{B.3}$$

where the first term in (a) follows from Markov chain (MC) $M_1 \multimap (X_1^n, \hat{X}_1^{n_1}) \multimap (Y_1^n, \hat{Y}_1^{n_1})$, while the second term follows from Fano's inequality; (b) follows from $(X_1^n, \hat{X}_1^{n_1}) \perp\!\!\!\perp (X_2^n, \hat{X}_2^{n_1})$; (c) follows from the chain rule; (d) follows from the first term in (c) due to MCs $\hat{X}_2^{n_1} \multimap X_2^n \multimap Y_1^n$ and $(\hat{X}_1^{n_1}, \hat{X}_2^{n_1}) \multimap (X_1^n, X_2^n) \multimap Y_1^n$, and since the second term in (c) is zero due to MC $(X_1^n, \hat{X}_1^{n_1}, X_2^n, Y_1^n) \multimap \hat{X}_2^{n_1} \multimap \hat{Y}_1^{n_1}$; (e) follows since unconditioning does not reduce entropy, and also from the memoryless Gaussian model; and (f) follows by defining random variables $X_1 := X_{1Q}$ and $X_2 := X_{2Q}$ with $\mathbb{E}[X_{1i}^2] =: P_{1i}$ and $\mathbb{E}[X_{2i}^2] =: P_{2i}, i \in \{1, \ldots, n\}$, where $Q \sim \mathbb{U}[1, \ldots, n]$, $\frac{1}{n}\sum_{i=1}^{n} P_{1i} \leq P_1$, $\frac{1}{n}\sum_{i=1}^{n} P_{2i} \leq P_2$, and Y_{1Q} is the resulting output at D_1; (g) follows from the fact that Gaussian inputs maximize differential entropy; (h) follows by applying Jensen's inequality to the concave log function.

We similarly bound R_2 by $R_2 \leq \mathsf{C}(P_2)$.

Achievability: Fix block lengths n and n_1 such that $n_1 \leq n\alpha_1$. Sender S_k encodes message $M_k \in \mathcal{M}_k$ into two codewords, $X_k^n(M_k)$ and $\hat{X}_k^{n_1}(M_k)$ that are intended for the microwave and the mm-wave bands, respectively, and generated according to distributions $X_k^n(M_k) \sim \mathcal{N}(P_k)$, i.i.d., and $\hat{X}_k^{n_1}(M_k) \sim \mathcal{N}(\hat{P}_k)$, i.i.d., $k = 1, 2$. Of the generated codewords, $X_1^n(M_1)$ and $X_2^n(M_2)$ are transmitted through the GIC in the microwave band, while $\hat{X}_1^{n_1}(M_1)$ and $\hat{X}_1^{n_1}(M_1)$ are transmitted through the S_1-D_2 cross-link and the S_2-D_1 cross-link, respectively. Destination D_1 first estimates the interfering message M_2 from the signals observed over both bands while treating the intended message M_1 as noise as in a point-to-point channel. Similarity, destination D_2 estimates M_1 while treating M_2 as noise. Using standard random coding arguments [97, Chapter 6], the probability of decoding error can be shown to become arbitrarily small at D_1 and D_2, as $n \to \infty$, if the following conditions are satisfied, respectively

$$R_2 \leq \frac{1}{2}\log\left(1 + \frac{a_{21}^2 P_2}{1 + P_1}\right) + \alpha_1 \frac{1}{2}\log\left(1 + c_{21}^2 \hat{P}_2\right) =: r_{21}, \tag{B.4}$$

$$R_1 \leq \frac{1}{2}\log\left(1 + \frac{a_{12}^2 P_1}{1 + P_2}\right) + \alpha_1 \frac{1}{2}\log\left(1 + c_{12}^2 \hat{P}_1\right) =: r_{12}. \tag{B.5}$$

Once the interfering message has been decoded, each destination removes the interference and estimates the intended signal interference-free, which results in the rates at D_1 and D_2

$$R_1 \leq \frac{1}{2}\log\left(1 + P_1\right) =: r_{11} \tag{B.6}$$

$$R_2 \leq \frac{1}{2}\log\left(1 + P_2\right) =: r_{22}. \tag{B.7}$$

Note that r_{11} in (B.6) provides the maximum rate from S_1 to D_1. Now, under condition (3.17), we have $r_{11} \leq r_{12}$, i.e., decoding M_1 at D_2 does not incur any rate loss. Similarly, under condition (3.18), we have $r_{22} \leq r_{21}$, i.e., decoding M_2 at D_1 does not incur any rate loss as well. Hence, under conditions (3.17)-(3.18), the rates (B.6)-(B.7) are achievable, which matches the rate outer bounds.

B.3 Proof of Theorem 3.3

Outer bound: The outer bounds on R_1 and R_2 can be derived as in (B.3), hence, we derive the outer bound on the sum rate. Assuming that message M_k was transmitted from source S_k, $k = 1, 2$, we have

$$\begin{aligned} n(R_1 + R_2) &= nR_2 + nR_1 \\ &= H(M_2) + H(M_1) \end{aligned}$$

$$
\begin{aligned}
&\overset{(a)}{\leq} I(M_2; Y_2^n, \hat{Y}_2^{n_1}) + I(M_1; Y_1^n, \hat{Y}_1^{n_1}) + n\epsilon_n \\
&\overset{(b)}{\leq} I(X_2^n, \hat{X}_2^{n_1}; Y_2^n, \hat{Y}_2^{n_1}) + I(X_1^n, \hat{X}_1^{n_1}; Y_1^n, \hat{Y}_1^{n_1}) + n\epsilon_n \\
&\overset{(c)}{\leq} I(X_2^n, \hat{X}_2^{n_1}; Y_2^n, \hat{Y}_2^{n_1}, X_1^n, \hat{X}_1^{n_1}) + I(X_1^n, \hat{X}_1^{n_1}; Y_1^n) + I(X_1^n, \hat{X}_1^{n_1}; \hat{Y}_1^{n_1} | Y_1^n) + n\epsilon_n \\
&\overset{(d)}{=} I(X_2^n; Y_2^n | X_1^n) + h(Y_1^n) - h(Y_1^n | X_1^n) + h(\hat{Y}_1^{n_1} | Y_1^n) - h(\hat{Y}_1^{n_1} | Y_1^n, X_1^n, \hat{X}_1^{n_1}) + n\epsilon_n \\
&\overset{(e)}{\leq} I(X_2^n; Y_2^n | X_1^n) + h(Y_1^n) - h(Y_1^n | X_1^n) + h(\hat{Y}_1^{n_1}) - h(\hat{Y}_1^{n_1} | Y_1^n, X_1^n, \hat{X}_1^{n_1}, \hat{X}_2^{n_1}) + n\epsilon_n \\
&\overset{(f)}{\leq} I(X_2^n; Y_1^n | X_1^n) + h(Y_1^n) - h(Y_1^n | X_1^n) + h(\hat{Y}_1^{n_1}) - h(\hat{Y}_1^{n_1} | \hat{X}_2^{n_1}) + n\epsilon_n \\
&\overset{(g)}{=} I(X_1^n, X_2^n; Y_1^n) + I(\hat{X}_2^{n_1}; \hat{Y}_1^{n_1}) + n\epsilon_n \\
&\overset{(h)}{\leq} nI(X_{1G}, X_{2G}; Y_{1G}) + n_1 I(\hat{X}_{2G}; \hat{Y}_{1G}) + n\epsilon_n \\
&= n\mathsf{C}\big(P_1 + a_{12}^2 P_2\big) + n_1 \mathsf{C}\Big(c_{12}^2 \hat{P}_2\Big) + n\epsilon_n
\end{aligned} \tag{B.8}
$$

where (a) follows from Fano's inequality; (b) follows from $M_k \multimap (X_k^n, \hat{X}_k^{n_1}) \multimap (Y_k^n, \hat{Y}_k^{n_1}), k = 1, 2$; (c) follows by providing $(X_1^n, \hat{X}_1^{n_1})$ to D_2; (d) follows from the fact that $(X_1^n, \hat{X}_1^{n_1}) \perp\!\!\!\perp (X_2^n, \hat{X}_2^{n_1})$, then expressing mutual information in terms of differential entropies and applying $Y_1^n \multimap X_1^n \multimap \hat{X}_1^{n_1}$ to the third term; (e) follows from unconditioning steps $h(\hat{Y}_1^{n_1} | Y_1^n) \leq h(\hat{Y}_1^{n_1})$ and $-h(\hat{Y}_1^{n_1} | Y_1^n, X_1^n, \hat{X}_1^{n_1}) \leq -h(\hat{Y}_1^{n_1} | Y_1^n, X_1^n, \hat{X}_1^{n_1}, \hat{X}_2^{n_1})$; (f) follows from having $a_{21}^2 \geq 1$ which results in $I(X_2^n; Y_2^n | X_1^n) \leq I(X_2^n; Y_1^n | X_1^n)$, and also due to the Markov chain $\hat{Y}_1^{n_1} \multimap \hat{X}_2^{n_1} \multimap (X_1^n, \hat{X}_1^{n_1}, Y_1^n)$ which results in $-h(\hat{Y}_1^{n_1} | Y_1^n, X_1^n, \hat{X}_1^{n_1}, \hat{X}_2^{n_1}) = -h(\hat{Y}_1^{n_1} | \hat{X}_2^{n_1})$; (g) follows by consuming the first three terms in (f) into one; and (h) follows from maximizing the mutual information terms by using $X_1^n \sim \mathcal{N}(0, P_1)$, $X_2^n \sim \mathcal{N}(0, P_2)$, and $\hat{X}_2^n \sim \mathcal{N}(0, \hat{P}_2)$, i.i.d., as in [131, Lemma 1] and following similar steps to (e)-(h) of (B.3).

Achievability: We use an achievability scheme that is similar to that is Appendix B.2. Specifically, the encoding proceeds in exactly the same manner as in Appendix B.2, whereas the decoding scheme is different: instead of sequential decoding, at each destination both the desired and the interfering messages are uniquely decoded as in a multi-access channel. Standard random coding arguments show that the following rates achievable at D_2

$$
\begin{aligned}
R_1 &\leq \mathsf{C}\big(a_{12}^2 P_1\big) + \alpha_1 \mathsf{C}\Big(c_{12}^2 \hat{P}_1\Big) \\
R_2 &\leq \mathsf{C}(P_2) \\
R_1 + R_2 &\leq \mathsf{C}\big(P_2 + a_{12}^2 P_1\big) + \alpha_1 \mathsf{C}\Big(c_{12}^2 \hat{P}_1\Big),
\end{aligned} \tag{B.9}
$$

whereas at the achievable rates at D_1 are given below

$$\begin{aligned} R_1 &\leq \mathsf{C}(P_1) \\ R_2 &\leq \mathsf{C}\big(a_{21}^2 P_2\big) + \alpha_1 \mathsf{C}\Big(c_{21}^2 \hat{P}_2\Big) \\ R_1 + R_2 &\leq \mathsf{C}\big(P_1 + a_{21}^2 P_2\big) + \alpha_1 \mathsf{C}\Big(c_{21}^2 \hat{P}_2\Big). \end{aligned} \tag{B.10}$$

Since $a_{21}^2 \geq 1$, comparing the two bounds on R_2 in (B.9) and (B.10), we have $R_2 \leq \mathsf{C}(P_2) < \mathsf{C}(a_{21}^2 P_2) + \alpha_1 \mathsf{C}\Big(c_{21}^2 \hat{P}_2\Big)$. Next, condition (3.22) requires that $\mathsf{C}(P_1) \leq \mathsf{C}(a_{12}^2 P_1) + \alpha_1 \mathsf{C}\Big(c_{12}^2 \hat{P}_1\Big)$, from which we have $R_1 \leq \mathsf{C}(P_1)$. Finally, condition (3.23) requires $\mathsf{C}(P_1 + a_{21}^2 P_2) + \alpha_1 \mathsf{C}\Big(c_{21}^2 \hat{P}_2\Big) \leq \mathsf{C}(P_2 + a_{12}^2 P_1) + \alpha_1 \mathsf{C}\Big(c_{12}^2 \hat{P}_1\Big)$, and thus comparing the two bounds on $R_1 + R_2$, we have $R_1 + R_2 \leq \mathsf{C}(P_1 + a_{21}^2 P_2) + \alpha_1 \mathsf{C}\Big(c_{21}^2 \hat{P}_2\Big)$.

B.4 Proof of Lemma 3.3

We bound rates R_1 and R_2 for the Z-CLIC of type-0 as follows. Since the signals received as D_1 of both these channels are the same, the transmission rate of S_1 remains unchanged, and thus the bound on R_1 are the same for both channels.

We bound rate R_2 for the Z-CLIC of type-0 as

$$\begin{aligned} nR_2 &= H(M_2) \\ &= I(M_2; Y_2^n, \hat{Y}_2^{n_1}) + H(M_2 | Y_2^n, \hat{Y}_2^{n_1}) \\ &= I(M_2; Y_2^n) + I(M_2; \hat{Y}_2^{n_1} | Y_2^n) + H(M_2 | Y_2^n, \hat{Y}_2^{n_1}) \\ &= I(M_2; Y_2^n) + H(M_2 | Y_2^n) \end{aligned} \tag{B.11}$$

where the last step follows from having $I(M_2; \hat{Y}_2^{n_1} | Y_2^n) = 0$, and from $H(M_2 | Y_2^n, \hat{Y}_2^{n_1}) = H(M_2 | Y_2^n)$ as $(W_2, Y_2^n) \perp\!\!\!\perp \hat{Y}_2^{n_1}$. Finally, note that the constraint on R_2 in (B.11) is the same as that in the Z-CLIC of type-1.

B.5 Proof of Theorem 3.4

The bounds on the individual rates R_1 and R_2 in (3.32) and (3.33) are found by following the same steps as in (3.19)-(3.20) that are detailed in Appendix B.2.

The sum-rate bound in (3.34) is derived by providing interfering signals $(X_2^n, \hat{X}_2^{n_1})$ to

D_1 that completely cancels interference from source S_2 as follows

$$
\begin{aligned}
n(R_1+R_2) &= H(M_1)+H(M_2)\\
&\overset{(a)}{\leq} I(M_1;Y_1^n,\hat{Y}_1^{n_1})+I(M_2;Y_2^n,\hat{Y}_2^{n_1})+n\epsilon_n\\
&\overset{(b)}{\leq} I(X_1^n,\hat{X}_1^{n_1};Y_1^n,\hat{Y}_1^{n_1})+I(X_2^n,\hat{X}_2^{n_1};Y_2^n,\hat{Y}_2^{n_1})+n\epsilon_n\\
&\overset{(c)}{\leq} I(X_1^n,\hat{X}_1^{n_1};Y_1^n,\hat{Y}_1^{n_1},X_2^n,\hat{X}_2^{n_1})+I(X_2^n,\hat{X}_2^{n_1};Y_2^n,\hat{Y}_2^{n_1})+n\epsilon_n\\
&\overset{(d)}{=} I(X_1^n,\hat{X}_1^{n_1};Y_1^n,\hat{Y}_1^{n_1}|X_2^n,\hat{X}_2^{n_1})+h(Y_2^n,\hat{Y}_2^{n_1})-h(Y_2^n,\hat{Y}_2^{n_1}|X_2^n,\hat{X}_2^{n_1})+n\epsilon_n\\
&= I(X_1^n;Y_1^n|X_2^n)+h(Y_2^n)+h(\hat{Y}_2^{n_1}|Y_2^n)-h(Y_2^n|X_2^n,\hat{X}_2^{n_1})\\
&\quad -h(\hat{Y}_2^{n_1}|Y_2^n,X_2^n,\hat{X}_2^{n_1})+n\epsilon_n\\
&\overset{(e)}{\leq} I(X_1^n;Y_1^n|X_2^n)+h(Y_2^n)+h(\hat{Y}_2^{n_1})-h(Y_2^n|X_2^n)-h(\hat{Y}_2^{n_1}|Y_2^n,X_2^n,\hat{X}_2^{n_1},\hat{X}_1^{n_1})+n\epsilon_n\\
&\overset{(f)}{=} h(Y_2^n)-h(Z_1^n)+h(\hat{Y}_2^{n_1})-h(\hat{Z}_2^{n_1})+h(Y_1^n|X_2^n)-h(Y_2^n|X_2^n)+n\epsilon_n\\
&\overset{(g)}{\leq} nh(Y_{2G})-nh(Z_2)+n_1h(\hat{Y}_{2G})-n_1h(\hat{Z}_2)+n\log\left(\frac{1+P_1}{1+a_{12}^2P_1}\right)+n\epsilon_n\\
&= n\mathsf{C}\big(P_2+a_{12}^2P_1\big)+n_1\mathsf{C}\Big(c_{12}^2\hat{P}_1\Big)+n\log\left(\frac{1+P_1}{1+a_{12}^2P_1}\right)+n\epsilon_n
\end{aligned}
\tag{B.12}
$$

where (a) follows from Fano's inequality; (b) follows from $M_k \multimap (X_k^n,\hat{X}_k^{n_1}) \multimap (Y_k^n,\hat{Y}_k^{n_1}), k=1,2$; (c) follows by providing $(X_2^n,\hat{X}_2^{n_1})$ to D_1; (d) follows since $(X_1^n,\hat{X}_1^{n_1}) \perp\!\!\!\perp (X_2^n,\hat{X}_2^{n_1})$; (e) follows from unconditioning steps $h(\hat{Y}_2^{n_1}|Y_2^n)\leq h(\hat{Y}_2^{n_1})$ and $-h(\hat{Y}_2^{n_1}|Y_2^n,X_2^n,\hat{X}_2^{n_1})\leq -h(\hat{Y}_2^{n_1}|Y_2^n,X_2^n,\hat{X}_2^{n_1},\hat{X}_1^{n_1})$; (f) follows since $h(\hat{Y}_2^{n_1}|Y_2^n,X_2^n,\hat{X}_2^{n_1},\hat{X}_1^{n_1})=h(\hat{Y}_2^{n_1}|\hat{X}_1^{n_1})=h(\hat{Z}_2^{n_1})$ due to the Markov chain $\hat{Y}_2^{n_1} \multimap \hat{X}_1^{n_1} \multimap (X_2^n,\hat{X}_2^{n_1},Y_2^n)$; (g) follows from upper bounding $h(Y_2^n)\leq n\mathrm{h}(Y_{2G})$ using $X_1\sim\mathcal{N}(0,P_1)$ and $X_2\sim\mathcal{N}(0,P_2)$, and $h(\hat{Y}_2^{n_1})\leq n_1h(\hat{Y}_{2G})$ using $\hat{X}_1\sim\mathcal{N}(0,\hat{P}_1)$ as in [131, Lemma 1], and upper bounding $h(Y_1^n|X_2^n)-h(Y_2^n|X_2^n)=h(X_1^n+Z_1^n)-h(a_{12}X_1^n+Z_2^n)\leq n\log\left(\frac{1+P_1}{1+a_{12}^2P_1}\right)$ by applying the WAN technique.

The bound in (3.35) is derived by exchanging the roles of $\mathsf{S}_1\leftrightarrow\mathsf{S}_2$ and $\mathsf{D}_1\leftrightarrow\mathsf{D}_2$, and then following the steps in (B.12).

The sum-rate bound in (3.36) is derived as follows:

$$
\begin{aligned}
n(R_1+R_2) &= H(W_1)+H(W_2)\\
&= I(W_1;Y_1^n,\hat{Y}_1^{n_1})+I(W_2;Y_2^n,\hat{Y}_2^{n_1})\\
&\overset{(a)}{\leq} I(X_1^n,\hat{X}_1^{n_1};Y_1^n,\hat{Y}_1^{n_1})+I(X_2^n,\hat{X}_2^{n_1};Y_2^n,\hat{Y}_2^{n_1})\\
&\overset{(b)}{\leq} I(X_1^n,\hat{X}_1^{n_1};Y_1^n,\hat{Y}_1^{n_1},S_1^n,\hat{S}_1^{n_1})+I(X_2^n,\hat{X}_2^{n_1};Y_2^n,\hat{Y}_2^{n_1},S_2^n,\hat{S}_2^{n_1})
\end{aligned}
$$

$$
\begin{aligned}
&\overset{(c)}{=} I(X_1^n, \hat{X}_1^{n_1}; S_1^n, \hat{S}_1^{n_1}) + I(X_1^n, \hat{X}_1^{n_1}; Y_1^n, \hat{Y}_1^{n_1} | S_1^n, \hat{S}_1^{n_1}) \\
&\quad + I(X_2^n, \hat{X}_2^{n_1}; S_2^n, \hat{S}_2^{n_1}) + I(X_2^n, \hat{X}_2^{n_1}; Y_2^n, \hat{Y}_2^{n_1} | S_2^n, \hat{S}_2^{n_1}) \\
&\overset{(d)}{=} h(S_1^n, \hat{S}_1^{n_1}) - h(S_1^n, \hat{S}_1^{n_1} | X_1^n, \hat{X}_1^{n_1}) \\
&\quad + h(Y_1^n, \hat{Y}_1^{n_1} | S_1^n, \hat{S}_1^{n_1}) - h(Y_1^n, \hat{Y}_1^{n_1} | X_1^n, \hat{X}_1^{n_1}, S_1^n, \hat{S}_1^{n_1}) \\
&\quad + h(S_2^n, \hat{S}_2^{n_1}) - h(S_2^n, \hat{S}_2^{n_1} | X_2^n, \hat{X}_2^{n_1}) \\
&\quad + h(Y_2^n, \hat{Y}_2^{n_1} | S_2^n, \hat{S}_2^{n_1}) - h(Y_2^n, \hat{Y}_2^{n_1} | X_2^n, \hat{X}_2^{n_1}, S_2^n, \hat{S}_2^{n_1}) \\
&\overset{(e)}{=} h(a_{12} X_1^n + Z_2^n, c_{12} \hat{X}_1^{n_1} + \hat{Z}_2^{n_1}) - h(Z_2^n, \hat{Z}_2^{n_1}) \\
&\quad + h(Y_1^n, \hat{Y}_1^{n_1} | S_1^n, \hat{S}_1^{n_1}) - h(a_{21} X_2^n + Z_1^n, c_{21} \hat{X}_2^{n_1} + \hat{Z}_1^{n_1}) \\
&\quad + h(a_{21} X_2^n + Z_1^n, c_{21} \hat{X}_2^{n_1} + \hat{Z}_1^{n_1}) - h(Z_1^n, \hat{Z}_1^{n_1}) \\
&\quad + h(Y_2^n, \hat{Y}_2^{n_1} | S_2^n, \hat{S}_2^{n_1}) - h(a_{12} X_1^n + Z_2^n, c_{12} \hat{X}_1^{n_1} + \hat{Z}_2^{n_1}) \\
&\overset{(f)}{\leq} h(Y_1^n | S_1^n) - h(Z_1^n) + h(\hat{Y}_1^{n_1} | \hat{S}_1^{n_1}) - h(\hat{Z}_1^{n_1}) \\
&\quad + h(Y_2^n | S_2^n) - h(Z_2^n) + h(\hat{Y}_2^{n_1} | \hat{S}_2^{n_1}) - h(\hat{Z}_2^{n_1}) \\
&\overset{(g)}{\leq} n\mathrm{h}(Y_{1G} | S_{1G}) - n\mathrm{h}(Z_1) + n_1 \mathrm{h}(\hat{Y}_{1G}) - n_1 h(\hat{Z}_1) \\
&\quad + n\mathrm{h}(Y_{2G} | S_{2G}) - n\mathrm{h}(Z_2) + n_1 \mathrm{h}(\hat{Y}_{2G}) - n_1 h(\hat{Z}_2) \\
&\overset{(h)}{=} \frac{n}{2} \log \left(1 + a_{12}^2 P_1 + \frac{P_2}{1 + a_{21}^2 P_2} \right) + \frac{n}{2} \log \left(1 + a_{21}^2 P_2 + \frac{P_1}{1 + a_{12}^2 P_1} \right) \\
&\quad + \frac{n_1}{2} \log(1 + c_{12}^2 \hat{P}_1) + \frac{n_1}{2} \log(1 + c_{21}^2 \hat{P}_2)
\end{aligned}
$$

where (a) follows from $(Y_k^n, \hat{Y}_k^{n_1}) \multimap (X_k^n, \hat{X}_k^{n_1}) \multimap W_k$, for $k = 1, 2$; (b) follows since providing genie-signals $(S_1^n, \hat{S}_1^{n_1})$ to D_1 and $(S_2^n, \hat{S}_2^{n_1})$ to D_2 does not reduce the mutual information; (c) follows from expanding the $I(.)$ terms using chain rule; (d) follows from expressing express mutual informations in terms of differential entropies; (e) follows from expressing the variables via their respective definitions, for example, the third term follows as

$$
\begin{aligned}
&\mathrm{h}(Y_1^n, \hat{Y}_1^{n_1} | X_1^n, \hat{X}_1^{n_1}, S_1^n, \hat{S}_1^{n_1}) \\
&\quad = \mathrm{h}(X_1^n + a_{21} X_2^n + Z_1^n, c_{21} \hat{X}_2^{n_1} + \hat{Z}_1^{n_1} | X_1^n, \hat{X}_1^{n_1}, a_{12} X_1^n + Z_2^n, c_{12} \hat{X}_1^{n_1} + \hat{Z}_2^{n_1}) \\
&\quad = \mathrm{h}(a_{21} X_2^n + Z_1^n, c_{21} \hat{X}_2^{n_1} + \hat{Z}_1^{n_1} | X_1^n, \hat{X}_1^{n_1}, Z_2^n, \hat{Z}_2^{n_1}) \\
&\quad = \mathrm{h}(a_{21} X_2^n + Z_1^n, c_{21} \hat{X}_2^{n_1} + \hat{Z}_1^{n_1})
\end{aligned}
$$

(f) follows by first noticing that the first and the eighth terms, and the fourth and the fifth terms in (e) cancel each other, and then unconditioning terms $h(Y_1^n, \hat{Y}_1^{n_1} | S_1^n, \hat{S}_1^{n_1})$ and $h(Y_2^n, \hat{Y}_2^{n_1} | S_2^n, \hat{S}_2^{n_1})$; (g) follows by single letterizing the positive entropy terms in (f): (i) $h(\hat{Y}_k^{n_1} | \hat{S}_k^{n_1}) = h(\hat{Y}_k^{n_1})$, which is then single letterized $h(\hat{Y}_k^{n_1}) \leq n_1 h(\hat{Y}_{kG})$ by using

$\hat{X}_\ell^{n_1} \sim \mathcal{N}(0, P_\ell)$, i.i.d., for $k \neq \ell \in \{1, 2\}$; (ii) $h(Y_1^n|S_1^n)$ is single letterized $h(Y_1^n|S_1^n) \leq nh(Y_{1G}|S_{1G})$ following [131, Lemma 1], by choosing inputs $X_1^n \sim \mathcal{N}(0, P_1)$, i.i.d., and $X_2^n \sim \mathcal{N}(0, P_2)$, i.i.d., where Y_{1G} and S_{1G} are the resulting output variables; finally, (h) follows from computing the (conditional) entropy terms in (g) which is straightforward.

Next, the bound in (3.37) is derived as follows

$$
\begin{aligned}
n(R_1 + 2R_2) &= H(M_1) + H(M_2) + H(M_2) \\
&\overset{(a)}{\leq} I(M_1; Y_1^n, \hat{Y}_1^{n_1}) + I(M_2; Y_2^n, \hat{Y}_2^{n_1}) + I(M_2; Y_2^n, \hat{Y}_2^{n_1}) + n\epsilon_n \\
&\overset{(b)}{\leq} I(X_1^n, \hat{X}_1^{n_1}; Y_1^n, \hat{Y}_1^{n_1}) + I(X_2^n, \hat{X}_2^{n_1}; Y_2^n, \hat{Y}_2^{n_1}) + I(X_2^n, \hat{X}_2^{n_1}; Y_2^n, \hat{Y}_2^{n_1}) + n\epsilon_n \\
&\overset{(c)}{\leq} I(X_1^n, \hat{X}_1^{n_1}; Y_1^n, \hat{Y}_1^{n_1}, S_1^n, \hat{S}_1^{n_1}) \\
&\quad + I(X_2^n, \hat{X}_2^{n_1}; Y_2^n, \hat{Y}_2^{n_1}, X_1^n, \hat{X}_1^{n_1}) + I(X_2^n, \hat{X}_2^{n_1}; Y_2^n, \hat{Y}_2^{n_1}) + n\epsilon_n \\
&\overset{(d)}{=} I(X_1^n, \hat{X}_1^{n_1}; S_1^n, \hat{S}_1^{n_1}) + I(X_1^n, \hat{X}_1^{n_1}; Y_1^n, \hat{Y}_1^{n_1}|S_1^n, \hat{S}_1^{n_1}) \\
&\quad + I(X_2^n; Y_2^n|X_1^n) + I(X_2^n, \hat{X}_2^{n_1}; Y_2^n, \hat{Y}_2^{n_1}) + n\epsilon_n \\
&\overset{(e)}{=} h(a_{12}X_1^n + Z_2^n, c_{12}\hat{X}_1^{n_1} + \hat{Z}_2^{n_1}) - h(Z_2^n, \hat{Z}_2^{n_1}) \\
&\quad + h(Y_1^n, \hat{Y}_1^{n_1}|S_1^n, \hat{S}_1^{n_1}) - h(a_{21}X_2^n + Z_1^n, c_{21}\hat{X}_2^{n_1} + \hat{Z}_1^{n_1}) \\
&\quad + h(X_2^n + Z_2^n) - h(Z_2^n) + h(Y_2^n, \hat{Y}_2^{n_1}) - h(a_{12}X_1^n + Z_2^n, c_{12}\hat{X}_1^{n_1} + \hat{Z}_2^{n_1}) + n\epsilon_n \\
&\overset{(f)}{\leq} h(Y_2^n) + h(\hat{Y}_2^{n_1}) - h(Z_2^n) - h(\hat{Z}_2^{n_1}) + h(Y_1^n|S_1^n) + h(\hat{Y}_1^{n_1}|\hat{S}_1^{n_1}) + n\epsilon_n \\
&\quad + h(X_2^n + Z_2^n) - h(a_{21}X_2^n + Z_1^n) - h(Z_2^n) - h(c_{21}\hat{X}_2^{n_1} + \hat{Z}_1^{n_1}|a_{21}X_2^n + Z_1^n, \hat{X}_2^{n_1}) \\
&\overset{(g)}{\leq} nh(Y_{2G}) + n_1 h(\hat{Y}_{2G}) - nh(Z_2) - n_1 h(\hat{Z}_2) + nh(Y_{1G}|S_{1G}) + n_1 h(\hat{Y}_{1G}) \\
&\quad + nh(X_{2G} + Z_2) - nh(a_{21}X_{2G} + Z_1) - nh(Z_2) - n_1 h(\hat{Z}_1) + n\epsilon_n \\
&= n\mathsf{C}\big(P_2 + a_{12}^2 P_1\big) + n_1 \mathsf{C}\Big(c_{12}^2 \hat{P}_1\Big) + n_1 \mathsf{C}\Big(c_{21}^2 \hat{P}_2\Big) \\
&\quad + \mathsf{C}\left(a_{21}^2 P_2 + \frac{P_1}{1 + a_{12}^2 P_1}\right) + n \log\left(\frac{1 + P_1}{1 + a_{12}^2 P_1}\right) + n\epsilon_n
\end{aligned} \tag{B.13}
$$

where (a) follows from Fano's inequality; (b) follows from $M_k \multimap (X_k^n, \hat{X}_k^{n_1}) \multimap (Y_k^n, \hat{Y}_k^{n_1}), k = 1, 2$; (c) follows by providing signals $(S_1^n, \hat{S}_1^{n_1})$ to D_1 and $(X_1^n, \hat{X}_1^{n_1})$ to D_2; (d) follows from $(X_1^n, \hat{X}_1^{n_1}) \perp\!\!\!\perp (X_2^n, \hat{X}_2^{n_1})$; (e) follows by expanding the mutual information terms into those of differential entropy, and then using the definition of the variables involved; (f) follows by first noticing that the first and the last terms in (e) cancel, and then unconditioning $h(Y_2^n, \hat{Y}_2^{n_1}) \leq h(Y_2^n) + h(\hat{Y}_2^{n_1})$, and $h(Y_1^n, \hat{Y}_1^{n_1}|S_1^n, \hat{S}_1^{n_1}) \leq h(Y_1^n|S_1^n) + h(\hat{Y}_1^{n_1}|\hat{S}_1^{n_1})$; (g) follows from first unconditioning $-h(c_{21}\hat{X}_2^{n_1} + \hat{Z}_1^{n_1}|a_{21}X_2^n + Z_1^n, \hat{X}_2^{n_1}) \leq -h(c_{21}\hat{X}_2^{n_1} + \hat{Z}_1^{n_1}|a_{21}X_2^n + Z_1^n, \hat{X}_2^{n_1}, \hat{X}_2^{n_1}) = -h(\hat{Z}_1^{n_1})$, and then maximizing the remaining entropy terms with Gaus-

sian random variables as in step (g) of (B.12).

Finally, the bound in (3.38) is derived from (B.13) by exchanging the roles of $\mathsf{S}_1 \leftrightarrow \mathsf{S}_2$ and $\mathsf{D}_1 \leftrightarrow \mathsf{D}_2$.

B.6 Proof of Lemma 3.7

Encoding: Fix the block lengths n and n_1 such that $n_1 = \lfloor \alpha_1 n \rfloor$. For a given set of power constraints $P_1, P_2, \hat{P}_1$ and $\hat{P}_2$, take P_1 and P_2, and fix a power-splitting $(P_{U1}, P_{W1}, P_{U2}, P_{W2}) \succeq \mathbf{0}$ such that vector $\boldsymbol{p} = (P_{U1}, P_{W1}, P_{U2}, P_{W2}, \hat{P}_1, \hat{P}_2) \in \mathcal{P}^*$.

The message $M_1 \in [1 : 2^{nR_1}]$ from source S_1 is first divided into a public part $M_{10} \in [1 : 2^{nR_{10}}]$ and a private part $M_{11} \in [1 : 2^{nR_{11}}]$ such that $R_1 = R_{10} + R_{11}$. The message from source S_2 is similarly divided into public and private parts $M_{20} \in [1 : 2^{nR_{20}}]$ and $M_{21} \in [1 : 2^{nR_{21}}]$ such that $R_2 = R_{20} + R_{21}$. The public message M_{10} is then jointly encoded into two codewords $(W_1^n(M_{10}), \hat{X}_1^{n_1}(M_{10}))$, where the codewords are generated according to $W_1^n \sim \mathcal{N}(0, P_{W1})$ and $\hat{X}_1^{n_1} \sim \mathcal{N}(0, \hat{P}_1)$, i.i.d. Similarly, message M_{20} is jointly encoded into two codewords $(W_2^n(M_{20}), \hat{X}_2^{n_1}(M_{20}))$ for the two bands, where the codewords are generated according to $W_2^n \sim \mathcal{N}(0, P_{W2})$ and $\hat{X}_2^{n_1} \sim \mathcal{N}(0, \hat{P}_2)$, i.i.d. The private messages M_{11} and M_{21} are encoded into two codewords $U_1^n(M_{11})$ and $U_2^n(M_{21})$, generated according to $U_1^n \sim \mathcal{N}(0, P_{U1})$ and $U_2^n \sim \mathcal{N}(0, P_{U2})$, i.i.d.

For transmission of message $M_1 = (M_{10}, M_{11})$, the private codeword $U_1^n(M_{11})$ is superimposed on the cloud center $W_1^n(M_{10})$ to create a single codeword $X_1^n(M_{10}, M_{11}) := W_1^n(M_{10}) + U_1^n(M_{11})$. Similarly, to transmit message $M_2 = (M_{20}, M_{21})$, codewords $U_2^n(M_{21})$ is superimposed on $W_2^n(M_{20})$ to create a single codeword $X_2^n(M_{20}, M_{21}) := W_2^n(M_{20}) + U_2^n(M_{21})$.

Since in the mm-wave band, only cross-links only carry interference information, a message splitting is not necessary.

Decoding: The decoder at destination D_1 uniquely decodes both desired public and private messages (M_{10}, M_{11}) and non-uniquely decodes the interfering public message M_{20} while treating the interfering private message M_{21} as noise as in [7]. Using the standard random coding arguments as in [97, Chapter 6.5.1], such decoding at D_1 is successful as $n, n_1 \to \infty$, if the public and private message rates satisfy the following constraints

$$R_{11} < I(X_1; Y_1 | W_1, W_2)$$

$$
\begin{aligned}
R_{11} + R_{10} &< I(X_1; Y_1|W_2) \\
R_{11} + R_{20} &< I(X_1, W_2; Y_1|W_1) + I(X_2, W_1; Y_2) + C_{21} \\
R_{11} + R_{10} + R_{20} &< I(X_1, W_2; Y_2) + C_{21}.
\end{aligned} \tag{B.14}
$$

Similarly, destination D_2 uniquely decodes intended public and private messages (M_{20}, M_{21}), while non-uniquely decoding M_{10} and treating M_{11} as noise, resulting in the following rate constraints

$$
\begin{aligned}
R_{21} &< I(X_2; Y_2|W_1, W_2) \\
R_{21} + R_{20} &< I(X_2; Y_2|W_1) \\
R_{21} + R_{10} &< I(X_2, W_1; Y_2|W_2) + I(X_1, W_2; Y_1) + C_{12} \\
R_{21} + R_{20} + R_{10} &< I(X_2, W_1; Y_1) + C_{12}.
\end{aligned} \tag{B.15}
$$

The individual private and public message rates satisfy $R_1 = R_{10} + R_{11}$ and $R_2 = R_{20} + R_{21}$. To present the achievable region in terms of only R_1 and R_2, we eliminate the individual private and public message rates in (B.14) and (B.15) by applying Fourier-Motzkin elimination as in [97, App. D], which results in the region $\mathcal{R}(\boldsymbol{p})$ in Theorem 3.7. Finally, $\mathcal{R}(\boldsymbol{p})$ was obtained for a fixed power-splitting $\boldsymbol{p} \in \mathcal{P}^*$, hence the overall achievable region is found by taking union over all such power-splits $\mathcal{R} = \cup_{\boldsymbol{p}\in\mathcal{P}^*}\mathcal{R}(\boldsymbol{p})$.

B.7 Proof of Theorem 3.5

We partition the range of $(\mathsf{INR}_1, \mathsf{INR}_2)$ into four sets: (i) $S_1^{\text{inr}} = \{(\mathsf{INR}_1, \mathsf{INR}_2) : \mathsf{INR}_1 \geq 1, \mathsf{INR}_2 \geq 1\}$, (ii) $S_2^{\text{inr}} = \{(\mathsf{INR}_1, \mathsf{INR}_2) : \mathsf{INR}_1 < 1, \mathsf{INR}_2 \geq 1\}$, (iii) $S_3^{\text{inr}} = \{(\mathsf{INR}_1, \mathsf{INR}_2) : \mathsf{INR}_1 \geq 1, \mathsf{INR}_2 < 1\}$, and (iv) $S_4^{\text{inr}} = \{(\mathsf{INR}_1, \mathsf{INR}_2) : \mathsf{INR}_1 < 1, \mathsf{INR}_2 < 1\}$. In each set, depending on the values of $(\mathsf{INR}_1, \mathsf{INR}_2)$, the private-public message powers are adapted according to (3.40)-(3.41). This results in closed form expressions for the inner bound $\mathcal{R}(\mathsf{INR}_1, \mathsf{INR}_2)$ and outer bound in Theorem 3.4, which are then shown to be within a distance of $\delta = \frac{1}{2}$ bit per user.

First consider S_1^{inr} where $\mathsf{INR}_1 \geq 1$ and $\mathsf{INR}_2 \geq 1$. Adapting $\mathcal{R}(\mathsf{INR}_1, \mathsf{INR}_2)$ according to the powers in (3.40)-(3.41), achievable region $\mathcal{R}(S_1^{\text{inr}})$ is given by

$$R_1 \leq \mathsf{C}(1 + \mathsf{SNR}_1) - 1/2 \tag{B.16}$$

$$R_2 \leq \mathsf{C}(1 + \mathsf{SNR}_2) - 1/2 \tag{B.17}$$

$$R_1 + R_2 \leq \mathsf{C}(2\mathsf{INR}_2 + \mathsf{SNR}_1 - 1) + \mathsf{C}\left(\frac{1+\mathsf{SNR}_2}{\mathsf{INR}_2}\right) + C_{12} - 1 \tag{B.18}$$

$$R_1 + R_2 \leq \mathsf{C}(2\mathsf{INR}_1 + \mathsf{SNR}_2 - 1) + \mathsf{C}\left(\frac{1+\mathsf{SNR}_1}{\mathsf{INR}_1}\right) + C_{21} - 1 \tag{B.19}$$

$$R_1 + R_2 \leq \mathsf{C}\left(\mathsf{INR}_1 + \frac{\mathsf{SNR}_1}{\mathsf{INR}_2}\right) + \mathsf{C}\left(\mathsf{INR}_2 + \frac{\mathsf{SNR}_2}{\mathsf{INR}_1}\right) + C_{12} + C_{21} - 1 \tag{B.20}$$

$$R_1 + 2R_2 \leq \mathsf{C}(\mathsf{INR}_2 + \mathsf{SNR}_2) + \mathsf{C}\left(1 + \frac{\mathsf{SNR}_2}{\mathsf{INR}_1}\right) + \mathsf{C}\left(\mathsf{INR}_1 + \frac{\mathsf{SNR}_1}{\mathsf{INR}_2}\right) + C_{12} + C_{21} - 3/2 \tag{B.21}$$

$$2R_1 + R_2 \leq \mathsf{C}(\mathsf{INR}_1 + \mathsf{SNR}_1) + \mathsf{C}\left(1 + \frac{\mathsf{SNR}_1}{\mathsf{INR}_2}\right) + \mathsf{C}\left(\mathsf{INR}_2 + \frac{\mathsf{SNR}_2}{\mathsf{INR}_1}\right) + C_{12} + C_{21} - 3/2, \tag{B.22}$$

where C_{12} and C_{21} are given in Lemma 3.7.

Now consider the achievable region in (B.16)-(B.22) and the outer bound in (3.32)-(3.38): clearly, both are piece-wise linear, and only consist of straight lines with slopes $0, -1/2, -1, -2$, and ∞. Our aim is to bound the difference between two bounds with the same slope by a constant. As such, we define δ_{R_1} to be the difference between the bounds on R_1 in (3.32) and (B.16), and similarly define variables $\delta_{R_2}, \delta_{R_1+R_2}, \delta_{2R_1+R_2}$, and $\delta_{R_1+2R_2}$ for the subsequent bounds.

Now, for the achievable region in (B.16)-(B.22) to be within 1/2 bit of the capacity region, if (R_1, R_2) is an achievable rate tuple, then $(R_1 + 1/2, R_2 + 1/2)$ must be outside the outer bound region in (3.32)-(3.38). If (R_1, R_2) is on the boundary of the achievable region (B.16)-(B.22), it must be on one of the bounding straight lines. Hence, for (B.16)-(B.22) to be within 1/2 bit of the capacity region, the following must be satisfied

$$\begin{aligned} \delta_{R_1} &\leq 1/2 \\ \delta_{R_2} &\leq 1/2 \\ \delta_{R_1+R_2} &\leq 1 \\ \delta_{R_1+2R_2} &\leq 3/2 \\ \delta_{2R_1+R_2} &\leq 3/2. \end{aligned} \tag{B.23}$$

For δ_{R_1}, we write

$$\begin{aligned} \delta_{R_1} &= \mathsf{C}(\mathsf{SNR}_1) - \mathsf{C}(1 + \mathsf{SNR}_1) + \frac{1}{2} \\ &= \frac{1}{2}\log(1 + \mathsf{SNR}_1) - \frac{1}{2}\log(2 + \mathsf{SNR}_1) + \frac{1}{2} \end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{2}\log\left(\frac{1+\mathsf{SNR}_1}{2+\mathsf{SNR}_1}\right)+\frac{1}{2}\\
&< \frac{1}{2}\log\left(\frac{2+\mathsf{SNR}_1}{2+\mathsf{SNR}_1}\right)+\frac{1}{2}\\
&= \frac{1}{2}.
\end{aligned}
$$

Similarly, δ_{R_2} is bounded as $\delta_{R_2} < 1/2$. Next, denoting the three outer bounds on R_1+R_2 in (3.34)-(3.36) by a_1, a_2, a_3, respectively, and the three inner bounds on the same in (B.18)-(B.20) by b_1, b_2, b_3, respectively, we have [10, eq.(62)]

$$
\begin{aligned}
\delta_{R_1+R_2} &= \min\{a_1,a_2,a_3\} - \min\{b_1,b_2,b_3\}\\
&= \max\{a_1-b_1, a_2-b_2, a_3-b_3\},
\end{aligned}
$$

and thus $\delta_{R_1+R_2}$ is bounded by the maximum of the three bounds $a_k - b_k, k = 1,2,3$. In particular, $a_1 - b_1$ in (3.34) and (B.18) is bounded as follows

$$
\begin{aligned}
a_1 - b_1 &= \frac{1}{2}\log\left(1+\mathsf{SNR}_1\right)+\frac{1}{2}\log\left(1+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_2}\right)+C_{12}\\
&\quad -\frac{1}{2}\log\left(2\mathsf{INR}_2+\mathsf{SNR}_1\right)-\frac{1}{2}\log\left(1+\frac{1+\mathsf{SNR}_2}{\mathsf{INR}_2}\right)-C_{12}+1\\
&= \frac{1}{2}\log\left(\frac{1+\mathsf{SNR}_1}{2\mathsf{INR}_2+\mathsf{SNR}_1}\right)+\frac{1}{2}\log\left(\frac{\mathsf{INR}_2}{1+\mathsf{INR}_2}\right)+1\\
&< \frac{1}{2}\log\left(\frac{2\mathsf{INR}_2+\mathsf{SNR}_1}{2\mathsf{INR}_2+\mathsf{SNR}_1}\right)+\frac{1}{2}\log\left(\frac{1+\mathsf{INR}_2}{1+\mathsf{INR}_2}\right)+1\\
&= 1
\end{aligned}
$$

where the inequality is due to $1 < 2 < 2\,\mathsf{INR}_2$. The other two bounds are similarly found to be $a_2 - b_2 < 1$ and $a_3 - b_3 < 1$. Therefore, we have $\delta_{R_1+R_2} < 1$.

Next, we bound $\delta_{R_1+2R_2}$ in (3.37) and (B.21)

$$
\begin{aligned}
\delta_{R_1+2R_2} &= \frac{1}{2}\log\left(1+\mathsf{INR}_2+\mathsf{SNR}_2\right)+\frac{1}{2}\log\left(1+\frac{\mathsf{SNR}_1}{(1+\mathsf{INR}_1)(1+\mathsf{INR}_2)}\right)+\frac{1}{2}\log\left(1+\mathsf{SNR}_2\right)\\
&\quad -\frac{1}{2}\log\left(1+\mathsf{INR}_2+\mathsf{SNR}_2\right)-\frac{1}{2}\log\left(2+\frac{\mathsf{SNR}_2}{\mathsf{INR}_1}\right)-\frac{1}{2}\log\left(1+\mathsf{INR}_1+\frac{\mathsf{SNR}_1}{\mathsf{INR}_2}\right)+\frac{3}{2}\\
&= \frac{1}{2}\log\left(1+\mathsf{INR}_1+\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right)-\frac{1}{2}\log\left(1+\mathsf{INR}_1+\frac{\mathsf{SNR}_1}{\mathsf{INR}_2}\right)\\
&\quad +\frac{1}{2}\log\left(\frac{1+\mathsf{SNR}_2}{1+\mathsf{INR}_1}\right)-\frac{1}{2}\log\left(2+\frac{\mathsf{SNR}_2}{\mathsf{INR}_1}\right)+\frac{3}{2}\\
&< \frac{1}{2}\log\left(1+\mathsf{INR}_1+\frac{\mathsf{SNR}_1}{\mathsf{INR}_2}\right)-\frac{1}{2}\log\left(1+\mathsf{INR}_1+\frac{\mathsf{SNR}_1}{\mathsf{INR}_2}\right)
\end{aligned}
$$

$$+\frac{1}{2}\log\left(\frac{2\mathsf{INR}_1+\mathsf{SNR}_2}{\mathsf{INR}_1}\right)-\frac{1}{2}\log\left(2+\frac{\mathsf{SNR}_2}{\mathsf{INR}_1}\right)+\frac{3}{2}$$
$$=\frac{3}{2}$$

where the inequality is due to $1<2<2\,\mathsf{INR}_1$. The bound on $\delta_{2R_1+R_2}$ is similarly found to be $\delta_{2R_1+R_2}<3/2$.

Hence, the constant gap result is proved for the set S_1^{inr}. For the other three sets, we proceed similarly, and verify that constant gap result with the same constant gap of $\delta=1/2$ bit/channel use hold.

B.8 Proof of Theorem 3.6

The proof follows from steps similar to those of Theorem 3.7 and Theorem 3.3, except the bound for R_1+2R_2 in (3.46), hence only an outline is provided. First, the individual bounds in (3.42) and (3.43) are derived following the same steps as those for (3.32) and (3.33). Second, since the weak interference condition $a_{12}^2<1$ in the mixed CLIC is the same as in the weak CLIC, the sum-rate bound in (3.44) is the same as in (3.34). Third, due to the strong interference condition $a_{21}^2\geq 1$, the sum-rate bound in (3.45), which has the same expression as the bound (3.24) of Theorem 3.3, follows from the same steps as those for (3.24).

Next, the bound on R_1+2R_2 in (3.46) needs to be derived differently compared to that in (3.37) as the the strong interference condition $a_{21}^2\geq 1$ in the mixed CLIC is opposite to that in the weak CLIC. Assuming that source S_k transmits message $M_k, k=1,2$, bound (3.46) is derived as follows

$$\begin{aligned}
n(R_1+2R_2) &= H(M_1)+H(M_2)+H(M_2)\\
&\overset{(a)}{\leq} I(M_1;Y_1^n,\hat{Y}_1^{n_1})+I(M_2;Y_2^n,\hat{Y}_2^{n_1})+I(M_2;Y_2^n,\hat{Y}_2^{n_1})+n\epsilon_n\\
&\overset{(b)}{\leq} I(X_1^n,\hat{X}_1^{n_1};Y_1^n,\hat{Y}_1^{n_1})+I(X_2^n,\hat{X}_2^{n_1};Y_2^n,\hat{Y}_2^{n_1})+I(X_2^n,\hat{X}_2^{n_1};Y_2^n,\hat{Y}_2^{n_1})+n\epsilon_n\\
&\overset{(c)}{\leq} I(X_1^n,\hat{X}_1^{n_1};Y_1^n,\hat{Y}_1^{n_1},S_1^n,\hat{S}_1^{n_1})+I(X_2^n,\hat{X}_2^{n_1};Y_2^n,\hat{Y}_2^{n_1})\\
&\quad+I(X_2^n,\hat{X}_2^{n_1};Y_2^n,\hat{Y}_2^{n_1},X_1^n,\hat{X}_1^{n_1},S_2^n,\hat{S}_2^{n_1})+n\epsilon_n\\
&\overset{(d)}{=} I(X_1^n,\hat{X}_1^{n_1};S_1^n,\hat{S}_1^{n_1})+I(X_1^n,\hat{X}_1^{n_1};Y_1^n,\hat{Y}_1^{n_1}|S_1^n,\hat{S}_1^{n_1})+I(X_2^n,\hat{X}_2^{n_1};Y_2^n,\hat{Y}_2^{n_1})\\
&\quad+I(X_2^n,\hat{X}_2^{n_1};S_2^n,\hat{S}_2^{n_1}|X_1^n,\hat{X}_1^{n_1})+I(X_2^n,\hat{X}_2^{n_1};Y_2^n,\hat{Y}_2^{n_1}|X_1^n,\hat{X}_1^{n_1},S_2^n,\hat{S}_2^{n_1})+n\epsilon_n
\end{aligned}$$

$$
\begin{aligned}
&\overset{(e)}{=} \underbrace{h(S_1^n,\hat{S}_1^{n_1})}_{\beta_1} - h(S_1^n,\hat{S}_1^{n_1}|X_1^n,\hat{X}_1^{n_1}) + h(Y_1^n,\hat{Y}_1^{n_1}|S_1^n,\hat{S}_1^{n_1}) \\
&\quad - \underbrace{h(Y_1^n,\hat{Y}_1^{n_1}|S_1^n,\hat{S}_1^{n_1},X_1^n,\hat{X}_1^{n_1})}_{\beta_2} + h(Y_2^n,\hat{Y}_2^{n_1}) - \underbrace{h(Y_2^n,\hat{Y}_2^{n_1}|X_2^n,\hat{X}_2^{n_1})}_{\beta_3} \\
&\quad + \underbrace{h(S_2^n,\hat{S}_2^{n_1}|X_1^n,\hat{X}_1^{n_1})}_{\beta_4} - h(S_2^n,\hat{S}_2^{n_1}|X_1^n,\hat{X}_1^{n_1},X_2^n,\hat{X}_2^{n_1}) \\
&\quad + h(Y_2^n,\hat{Y}_2^{n_1}|X_1^n,\hat{X}_1^{n_1},S_2^n,\hat{S}_2^{n_1}) - h(Y_2^n,\hat{Y}_2^{n_1}|X_1^n,\hat{X}_1^{n_1},S_2^n,\hat{S}_2^{n_1},X_2^n,\hat{X}_2^{n_1}) + n\epsilon_n \\
&\overset{(f)}{\leq} h(Y_1^n|S_1^n) + h(\hat{Y}_1^{n_1}|\hat{S}_1^{n_1}) - h(Z_2^n) - h(\hat{Z}_2^{n_1}) + h(Y_2^n) - h(Z_1^n) + n\epsilon_n \\
&\quad + h(\hat{Y}_2^{n_1}) - h(\hat{Z}_1^{n_1}) + h(Y_2^n|X_1^n,S_2^n) + h(\hat{Y}_2^{n_1}|\hat{X}_1^{n_1},\hat{S}_2^{n_1}) - h(Z_2^n) - h(\hat{Z}_2^{n_1}) \\
&\overset{(g)}{=} h(Y_1^n|S_1^n) - h(Z_1^n) + h(\hat{Y}_1^{n_1}) - h(\hat{Z}_1^{n_1}) + h(Y_2^n) - h(Z_2^n) + n\epsilon_n \\
&\quad + h(\hat{Y}_2^{n_1}) - h(\hat{Z}_2^{n_1}) + h(Y_2^n|X_1^n,S_2^n) - h(Z_2^n) + h(\hat{Z}_2^{n_1}) - h(\hat{Z}_2^{n_1}) \\
&\overset{(h)}{\leq} n\left(h(Y_{1G}|S_{1G}) - h(Z_1) + h(Y_{2G}) - h(Z_2) + h(Y_{2G}|X_{1G},S_{2G}) - h(Z_2)\right) + n\epsilon_n \\
&\quad + n_1\left(h(\hat{Y}_{2G}) - h(\hat{Z}_2) + h(\hat{Y}_{1G}) - h(\hat{Z}_1)\right) \\
&= \mathsf{C}\big(P_2 + a_{12}^2 P_1\big) + \mathsf{C}\left(a_{21}^2 P_2 + \frac{P_1}{a_{12}^2 P_1 + 1}\right) + \mathsf{C}\left(\frac{P_2}{a_{21}^2 P_2 + 1}\right) + n\epsilon_n \qquad \text{(B.24)}
\end{aligned}
$$

where (a) follows from Fano's inequality; (b) follows from $M_k \multimap (X_k^n, \hat{X}_k^{n_1}) \multimap (Y_k^n, \hat{Y}_k^{n_1}), k = 1, 2$; (c) follows by providing $(S_1^n, \hat{S}_1^{n_1})$ to D_1 and $(X_1^n, \hat{X}_1^{n_1}, S_2^n, \hat{S}_2^{n_1})$ to D_2; (d) follows from expanding the mutual information terms using chain rule and since $(X_1^n, \hat{X}_1^{n_1}) \perp\!\!\!\perp (X_2^n, \hat{X}_2^{n_1})$; (e) follows by expression mutual information via differential entropies; (f) follows by first noticing that β_1 and β_3, as well as β_2 and β_4, cancel each other, and then applying unconditioning steps $h(Y_1^n, \hat{Y}_1^{n_1}|S_1^n, \hat{S}_1^{n_1}) \leq h(Y_1^n|S_1^n) + h(\hat{Y}_1^{n_1}|\hat{S}_1^{n_1})$, $h(Y_2^n, \hat{Y}_2^{n_1}) \leq h(Y_2^n) + h(\hat{Y}_2^{n_1})$ and $h(Y_2^n, \hat{Y}_2^{n_1}|X_1^n, \hat{X}_1^{n_1}, S_2^n, \hat{S}_2^{n_1}) \leq h(Y_2^n|X_1^n, S_2^n) + h(\hat{Y}_2^{n_1}|\hat{X}_1^{n_1}, \hat{S}_2^{n_1})$; (g) follows since $\hat{Y}_1^{n_1} \perp\!\!\!\perp \hat{S}_1^{n_1}$ and $h(\hat{Y}_2^{n_1}|\hat{X}_1^{n_1}, \hat{S}_2^{n_1}) = h(\hat{Z}_2^{n_1})$ due to the Markov chain $\hat{Y}_2^{n_1} \multimap \hat{X}_1^{n_1} \multimap \hat{S}_2^{n_1}$; finally, (h) follows from maximizing (conditional) entropy by choosing i.i.d. Gaussian inputs [131, Lemma 1] through steps similar to steps (e)-(h) of (B.3).

Finally, note that compared to the seven outer bounds in Theorem 3.7 for the weak CLIC, the mixed case has only five bounds. This is due to the fact that the two remaining bounds, a bound on $2R_1 + R_2$ and the third bound on $R_1 + R_2$, have the same expressions as those in (3.36) and (3.38) but are redundant.

B.9 Proof of Theorem 3.7

We partition the range of INR_2 into two sets $\mathsf{INR}_2 \geq 1$ and $\mathsf{INR}_2 < 1$, and then for each set, we simplify the achievable region $\mathcal{R}(\mathsf{INR}_2, *)$ and show it to be within $\delta = 1/2$ bit of the outer bound in Theorem 3.6.

First consider $\mathsf{INR}_2 \geq 1$, for which $\mathcal{R}(\mathsf{INR}_2, *)$ is characterized by rates

$$R_1 \leq \mathsf{C}(\mathsf{SNR}_1) \tag{B.25}$$

$$R_2 \leq \mathsf{C}(\mathsf{SNR}_2) - 1/2 \tag{B.26}$$

$$R_1 + R_2 \leq \mathsf{C}(\mathsf{INR}_2 + \mathsf{SNR}_2) + \mathsf{C}\left(\frac{\mathsf{SNR}_1}{\mathsf{INR}_2}\right) + C_{12} - 1/2 \tag{B.27}$$

$$R_1 + R_2 \leq \mathsf{C}(\mathsf{INR}_1 + \mathsf{SNR}_1) + C_{21} \tag{B.28}$$

$$R_1 + 2R_2 \leq \mathsf{C}(\mathsf{INR}_2 + \mathsf{SNR}_2) + \mathsf{C}\left(\mathsf{INR}_1 + \frac{\mathsf{SNR}_1}{\mathsf{INR}_2}\right) + C_{12} + C_{21} - 1/2 \tag{B.29}$$

Note that compared to seven inequalities for the weak CLIC, the achievable region for th mixed CLIC is defined by only five inequalities as one constraint on $R_1 + R_2$ and one on $2R_1 + R_2$ turn out to be redundant. Similar to (B.23) for the weak CLIC, for the achievable region (B.25)-(B.29) to be within 1/2 bit of the outer bound (3.42)-(3.46), the gap $\delta_{(.)}$ between two respective bounds should satisfy

$$\begin{aligned} \delta_{R_1} &\leq 1/2 \\ \delta_{R_2} &\leq 1/2 \\ \delta_{R_1+R_2} &\leq 1 \\ \delta_{R_1+2R_2} &\leq 3/2. \end{aligned} \tag{B.30}$$

First, it is clear from (3.42) and (B.25) that $\delta_{R_1} = 0$, and also from (3.43) and (B.26), that $\delta_{R_2} = 1/2$. Next, the difference between the bounds in (3.45) and (B.28) is $\delta_{R_1+R_2} = 0$, while the same between the bounds in (3.45) and (B.28) is

$$\begin{aligned} \delta_{R_1+R_2} =& \frac{1}{2}\log\left(1+\mathsf{SNR}_1\right) + \frac{1}{2}\log\left(1+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_2}\right) \\ & - \frac{1}{2}\log\left(1+\mathsf{INR}_2+\mathsf{SNR}_2\right) - \frac{1}{2}\log\left(1+\frac{\mathsf{SNR}_1}{\mathsf{INR}_2}\right) + \frac{1}{2} \\ =& \frac{1}{2}\log\frac{\left(1+\mathsf{SNR}_1\right)\mathsf{INR}_2}{\left(\mathsf{INR}_2+\mathsf{SNR}_1\right)\left(1+\mathsf{INR}_2\right)} + \frac{1}{2} \\ <& 1 \end{aligned}$$

where the inequality follows from simple algebraic manipulations. Similarly, it can be

shown that $\delta_{R_1+2R_2} < 3/2$ in a straightforward manner. Hence, the constant gap result is proved for $\mathsf{INR}_2 \geq 1$.

The proof for the other set $\mathsf{INR}_2 < 1$ follows similar steps, and hence not repeated.

B.10 Proof of Theorem 3.8

Proof. **Outer Bound:** For the Z-CLIC of type-1, the following set of tuples (R_1, R_2) provide an outer bound to the capacity region

$$R_1 \leq \mathsf{C}(\mathsf{SNR}_1) \tag{B.31}$$

$$R_2 \leq \mathsf{C}(\mathsf{SNR}_2) \tag{B.32}$$

$$R_1 + R_2 \leq \mathsf{C}\left(\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_1}\right) + \mathsf{C}(\mathsf{SNR}_2) + C_{21}, \tag{B.33}$$

where the individual rate bounds are the same as (3.32) and (3.33) and found as in (B.3), and the sum-rate bound is the same as (3.35) and is found as in Appendix B.5.

Achievability: The achievable region is characterized as follows. Since $a_{12} = 0$, D_2 is interference-free, and hence S_1 does not deploy message splitting and communicates only the private message by allocating $P_{U1} = P_1$ and $P_{W1} = 0$. On the other hand, since $a_{21}^2 \neq 0$, S_1 deploys message splitting and adapts the private-common message powers according to the interference level INR_1 following (3.40). The resulting achievable region $\mathcal{R}\left(\mathsf{INR}_1, \mathsf{INR}_2 = 0\right)$ is presented below for the two cases: $\mathsf{INR}_1 \geq 1$, and $\mathsf{INR}_1 < 1$.

If $\mathsf{INR}_1 \geq 1$, then $\mathcal{R}\left(\mathsf{INR}_1, \mathsf{INR}_2 = 0\right)$ is given by the set of tuples (R_1, R_2) that satisfy

$$R_1 \leq \mathsf{C}(\mathsf{SNR}_1) - 1/2 \tag{B.34}$$

$$R_2 \leq \mathsf{C}(\mathsf{SNR}_2) \tag{B.35}$$

$$R_1 + R_2 \leq \mathsf{C}(\mathsf{INR}_1 + \mathsf{SNR}_1) + \mathsf{C}\left(\frac{\mathsf{SNR}_2}{\mathsf{INR}_1}\right) + C_{21} - 1/2, \tag{B.36}$$

whereas for $\mathsf{INR}_1 < 1$, $\mathcal{R}\left(\mathsf{INR}_1, \mathsf{INR}_2 = 0\right)$ is given by

$$R_1 \leq \mathsf{C}\left(\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_1}\right) \tag{B.37}$$

$$R_2 \leq \mathsf{C}(\mathsf{SNR}_2) \tag{B.38}$$

$$R_1 + R_2 \leq \mathsf{C}\left(\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_1}\right) + \mathsf{C}(\mathsf{SNR}_2) + C_{21}. \tag{B.39}$$

Constant Gap: Now, for $\mathsf{INR}_1 < 1$, it is clear from (B.32) and (B.38) that $\delta_{R_2} = 0$, and from (B.33) and (B.39) that $\delta_{R_1+R_2} = 0$, while from (B.31) and (B.37) it follows that $\delta_{R_1} < 1/2$. Similarly, for $\mathsf{INR}_1 \geq 1$, it is apparent that $\delta_{R_2} = 0$, and simple algebraic manipulations of bounds in from (B.31) and (B.34) lead to $\delta_{R_1} < 1/2$, while those in (B.33) and (B.36) lead to $\delta_{R_1+R_2} < 1$. Thus, the constant gaps are obtained. ■

B.11 Derivation of the Optimal Power Allocation in Section 3.4.1

Convexity of Problem $[\mathcal{P}1]$**:** We first show that $[\mathcal{P}1]$ is convex. Denote by $\mathbf{x} := (p_1, q_1, p_2, q_2, R) \in \mathbb{R}_+^5$ a feasible point, satisfying the constraints in $[\mathcal{P}1]$. The objective of $[\mathcal{P}1]$ is equivalent to *minimizing* $-R$, which is linear. In addition, the equality constraints (3.55) and (3.56) are affine. Next, we note that the constraints (3.52)–(3.54) are convex. To illustrate, we consider constraint (3.52) denoted

$$g_1(\mathbf{x}) := R - A_1 - \frac{\alpha\bar{\beta}\kappa}{2}\ln(1+c_{21}^2 q_2) - \frac{\alpha\beta\kappa}{2}\ln(1+d_1^2 p_1) - \frac{\alpha\beta\kappa}{2}\ln(1+d_2^2 p_2),$$

with $\kappa := 1/\ln 2$, and derive its Hessian

$$\nabla^2 g_1(\mathbf{x}) := \frac{\alpha\kappa}{2}\ \mathrm{diag}\left[\frac{\beta d_1^4}{(1+d_1^2 p_1)^2},\ \ 0,\ \ \frac{\beta d_2^4}{(1+d_2^2 p_2)^2},\ \ \frac{\bar{\beta} c_{21}^4}{(1+c_{21}^2 q_2)^2},\ \ 0\right],$$

where $D = \mathrm{diag}\,[a_1, \ldots, a_m]$ is a diagonal matrix with elements $a_1, \ldots, a_m$. Note that $\nabla^2 g_1(\mathbf{x})$ is positive semidefinite, and thus (3.52) is convex. Likewise, constraints (3.53) and (3.54) are found to be convex. In addition, (3.55) and (3.56), and $\mathbf{x} \succeq \mathbf{0}$ imply that the feasible set is compact for given $P > 0$. Hence, $[\mathcal{P}1]$ is a convex problem over a compact set.

The KKT Conditions: Problem $[\mathcal{P}1]$ satisfies Slater's condition [132, Chapter 5.2.3]: it can be illustrated by considering the point $\tilde{\mathbf{x}} := \left(P - \bar{\beta}\epsilon/\beta,\ \ \epsilon,\ \ P - \bar{\beta}\epsilon/\beta,\ \ \epsilon,\ \ A_1\right)$ which is strictly feasible for sufficiently small $\epsilon > 0$. Therefore, $[\mathcal{P}1]$ can be solved using the KKT conditions [132, Chapter 5.5.3]. Moreover, $[\mathcal{P}1]$ also satisfies the Mangasarian-Fromovitz constraint qualification as illustrated in Appendix B.12.

We define the Lagrange multipliers $\{\lambda_i\}_{i=1}^3$, corresponding to constraints (3.52)–(3.54), $\{\mu_i\}_{i=1}^2$, corresponding to constraints (3.55) and (3.56), and $\{\rho_i\}_{i=1}^5$, corresponding to the

non-negativity constraints, $(p_1, q_1, p_2, q_2, R) \succeq \mathbf{0}$. The Lagrangian is then defined as

$$\begin{aligned}\mathcal{L} = &-R + \lambda_1(R - \Sigma_1) + \lambda_2(R - \Sigma_2) + \lambda_3(R - \Sigma) \\ &+ \mu_1\big(\beta p_1 + \bar{\beta} q_1 - P\big) + \mu_2\big(\beta p_2 + \bar{\beta} q_2 - P\big) - \rho_1 p_1 - \rho_2 q_1 - \rho_3 p_2 - \rho_4 q_2 - \rho_5 R,\end{aligned}$$

where Σ_1, Σ_2, and Σ are defined in (3.47)–(3.49) with $A_1 = A_2$.

With a slight abuse of notation, we denote the optimal primal variable by (p_1, q_1, p_2, q_2, R), and the optimal Lagrange multipliers by $(\lambda_1, \lambda_2, \lambda_3, \mu_1, \mu_2, \rho_1, \rho_2, \rho_3, \rho_4, \rho_5)$, which satisfy the KKT conditions below

$$\lambda_1 + \lambda_2 + \lambda_3 = 1, \tag{B.40}$$

$$\rho_1 = \beta\left(\mu_1 - \frac{\alpha\kappa}{2}\frac{d_1^2}{1 + d_1^2 p_1}\right), \tag{B.41}$$

$$\rho_2 = \bar{\beta}\left(\mu_1 - \frac{\alpha\kappa}{2}\frac{c_{12}^2\lambda_2}{1 + c_{12}^2 q_1}\right), \tag{B.42}$$

$$\rho_3 = \beta\left(\mu_2 - \frac{\alpha\kappa}{2}\frac{d_2^2}{1 + d_2^2 p_2}\right), \tag{B.43}$$

$$\rho_4 = \bar{\beta}\left(\mu_2 - \frac{\alpha\kappa}{2}\frac{c_{21}^2\lambda_1}{1 + c_{21}^2 q_2}\right), \tag{B.44}$$

$$\beta p_1 + \bar{\beta} q_1 = P, \tag{B.45}$$

$$\beta p_2 + \bar{\beta} q_2 = P, \tag{B.46}$$

$$R - \Sigma_1 \leq 0, \tag{B.47}$$

$$R - \Sigma_2 \leq 0, \tag{B.48}$$

$$R - \Sigma \leq 0, \tag{B.49}$$

$$\lambda_1(R - \Sigma_1) = 0, \tag{B.50}$$

$$\lambda_2(R - \Sigma_2) = 0, \tag{B.51}$$

$$\lambda_3(R - \Sigma) = 0, \tag{B.52}$$

$$\rho_1 p_1 = 0, \tag{B.53}$$

$$\rho_2 q_1 = 0, \tag{B.54}$$

$$\rho_3 p_2 = 0, \tag{B.55}$$

$$\rho_4 q_2 = 0, \tag{B.56}$$

$$(p_1, q_1, p_2, q_2, R, \lambda_1, \lambda_2, \lambda_3, \rho_1, \rho_2, \rho_3, \rho_4) \succeq \mathbf{0}. \tag{B.57}$$

Note that since $R \geq A_1 > 0$, the associated multiplier $\rho_5 = 0$. Next, in order to solve for the optimal variables from the KKT conditions, we assume that $A_1 = A_2$ (assumptions [A2]) and $c_{12}^2 < c_{21}^2$, and partition the set of optimal Lagrange multipliers.

Partitioning the Set of the Optimal Lagrange Multipliers: We now partition the set of optimal Lagrange multipliers, i.e., the set of all $(\boldsymbol{\rho}, \boldsymbol{\lambda})$, where $\boldsymbol{\rho} := (\rho_1, \rho_2, \rho_3, \rho_4) \succeq \mathbf{0}$ and $\boldsymbol{\lambda} := (\lambda_1, \lambda_2, \lambda_3) \succeq \mathbf{0}$, into 18 subsets. First, the set of (ρ_1, ρ_2)-tuples are partitioned into 3 subsets $\mathcal{I}_1 := \{(\rho_1, \rho_2) : \rho_1 > 0, \rho_2 = 0\}$, $\mathcal{I}_2 := \{(\rho_1, \rho_2) : \rho_1 = 0, \rho_2 > 0\}$, and $\mathcal{I}_3 := \{(\rho_1, \rho_2) : \rho_1 = 0, \rho_2 = 0\}$, since the other possible subset $\mathcal{I}_4 := \{(\rho_1, \rho_2) : \rho_1 > 0, \rho_2 > 0\}$ is empty as $(\rho_1, \rho_2) \in \mathcal{I}_4$ requires $p_1 = q_1 = 0$, which violates (B.45). Similarly, the set of (ρ_3, ρ_4)-tuples are partitioned into 3 subsets $\mathcal{J}_1 := \{(\rho_3, \rho_4) : \rho_3 > 0, \rho_4 = 0\}$, $\mathcal{J}_2 := \{(\rho_3, \rho_4) : \rho_3 = 0, \rho_4 > 0\}$, and $\mathcal{J}_3 := \{(\rho_3, \rho_4) : \rho_3 = 0, \rho_4 = 0\}$.

Next, the set of $\boldsymbol{\lambda} \succeq \mathbf{0}$ is partitioned into only 2 non-empty subsets $\mathcal{L}_1 := \{\boldsymbol{\lambda} : \lambda_1 > 0, \lambda_2 > 0, \lambda_3 = 0\}$, and $\mathcal{L}_2 := \{\boldsymbol{\lambda} : \lambda_1 > 0, \lambda_2 > 0, \lambda_3 > 0\}$, since all other subsets are empty as illustrated below. First, the remaining subsets of $\boldsymbol{\lambda}$ are denoted as $\mathcal{L}_3 := \{\boldsymbol{\lambda} : \lambda_1 > 0, \lambda_2 = 0, \lambda_3 \geq 0\}$, $\mathcal{L}_4 := \{\boldsymbol{\lambda} : \lambda_1 = 0, \lambda_2 > 0, \lambda_3 \geq 0\}$, and $\mathcal{L}_5 := \{\boldsymbol{\lambda} : \lambda_1 = 0, \lambda_2 = 0, \lambda_3 > 0\}$, whereas the subset $\mathcal{L}_6 := \{\boldsymbol{\lambda} : \lambda_1 = 0, \lambda_2 = 0, \lambda_3 = 0\}$ is empty as it violates (B.40).

Now, we examine below which $\boldsymbol{\rho} \in \mathcal{I}_k \cap \mathcal{J}_l$ and $\boldsymbol{\lambda} \in \mathcal{L}_j$ jointly satisfy the KKT conditions (i.e., compatible), or not (i.e., incompatible):

B1. For any $\boldsymbol{\rho} \in \mathcal{I}_k \cap \mathcal{J}_l$, $k \in \{1, 3\}, l \in \{1, 2, 3\}$, we have $\boldsymbol{\lambda} \notin \mathcal{L}_3, \mathcal{L}_5$. This follows from a specific relation between ρ_2 and λ_2: for any $\boldsymbol{\rho} \in \mathcal{I}_k \cap \mathcal{J}_l$ where $\rho_2 = 0$, and for any $\boldsymbol{\lambda} \notin \mathcal{L}_m$ where $\lambda_2 = 0$, we have $\mu_1 = 0$ from (B.42), resulting in $\rho_1 = -\alpha\kappa d_1^2/(2(1 + d_1^2 p_1)) < 0$ from (B.41), which violates $\rho_1 \geq 0$ in (B.57). Since $\rho_2 = 0$ in all $\boldsymbol{\rho} \in \mathcal{I}_k \cap \mathcal{J}_l$, $k \in \{1, 3\}, l \in \{1, 2, 3\}$, and $\lambda_2 = 0$ in all $\boldsymbol{\lambda} \notin \mathcal{L}_3, , \mathcal{L}_5$, these pairs of $\boldsymbol{\rho}$ and $\boldsymbol{\lambda}$ are incompatible.

B2. For any $\boldsymbol{\rho} \in \mathcal{I}_k \cap \mathcal{J}_l$, $k \in \{1, 2, 3\}, l \in \{1, 3\}$, we have $\boldsymbol{\lambda} \notin \mathcal{L}_4, \mathcal{L}_5$. This follows from the consequence of having $\rho_4 = 0$ in all $\boldsymbol{\rho} \in \mathcal{I}_k \cap \mathcal{J}_l$, $k \in \{1, 2, 3\}, l \in \{1, 3\}$, and $\lambda_1 = 0$ in all $\boldsymbol{\lambda} \notin \mathcal{L}_4, , \mathcal{L}_5$, as in [B1] above.

B3. For $\boldsymbol{\rho} \in \mathcal{I}_1 \cap \mathcal{J}_2$, we have $\boldsymbol{\lambda} \notin \mathcal{L}_4, \mathcal{L}_1, \mathcal{L}_2$. This follows since for $\boldsymbol{\rho} \in \mathcal{I}_1 \cap \mathcal{J}_2$, we have $\rho_1 > 0, \rho_2 = 0, \rho_3 = 0, \rho_4 > 0$, which requires $p_1 = 0, q_1 = P/\bar{\beta}, p_2 = P/\beta, q_2 = 0$, from (B.53)–(B.56), (B.45), and (B.46). This further implies $\Sigma_1 < \Sigma_2$, which contradicts the implications of $\boldsymbol{\lambda} \in \mathcal{L}_4$, that requires $\Sigma_1 > \Sigma_2$, or of $\boldsymbol{\lambda} \in \mathcal{L}_1, \mathcal{L}_2$ that requires $\Sigma_1 = \Sigma_2$.

B4. For $\boldsymbol{\rho} \in \mathcal{I}_2 \cap \mathcal{J}_1$, we have $\boldsymbol{\lambda} \notin \mathcal{L}_3, \mathcal{L}_1, \mathcal{L}_2$, which follows along the lines of [B3] by exchanging the roles of ρ_1, ρ_3, and Σ_1 with those of ρ_2, ρ_4, and Σ_2, respectively..

B5. For $\boldsymbol{\rho} \in \mathcal{I}_2 \cap \mathcal{J}_3$, we have $\boldsymbol{\lambda} \notin \mathcal{L}_3, \mathcal{L}_1, \mathcal{L}_2$. It follows from the fact that $\boldsymbol{\rho} \in \mathcal{I}_2 \cap \mathcal{J}_3$ satisfies $\rho_1 = 0, \rho_2 > 0, \rho_3 = 0, \rho_4 = 0$, which requires $p_1 = P/\beta, q_1 = 0, p_2 > 0, q_2 > 0$, following (B.53)–(B.56), (B.45), and (B.46). This assertion results in $\Sigma_1 > \Sigma_2$, which contradicts the implications of $\boldsymbol{\lambda} \in \mathcal{L}_3$, that requires $\Sigma_1 < \Sigma_2$, or of $\boldsymbol{\lambda} \in \mathcal{L}_1, \mathcal{L}_2$ that requires $\Sigma_1 = \Sigma_2$.

B6. For $\boldsymbol{\rho} \in \mathcal{I}_3 \cap \mathcal{J}_2$, we have $\boldsymbol{\lambda} \notin \mathcal{L}_4, , \mathcal{L}_1, \mathcal{L}_2$, which follows along the lines of [B5] by exchanging the roles of ρ_1, ρ_2, and Σ_1 with those of ρ_3, ρ_4, and Σ_2, respectively.

B7. For $\boldsymbol{\rho} \in \mathcal{I}_2 \cap \mathcal{J}_2$, we have $\boldsymbol{\lambda} \notin \mathcal{L}_3, \mathcal{L}_4, \mathcal{L}_5, \mathcal{L}_2$. This follows from the fact that $\boldsymbol{\rho} \in \mathcal{I}_2 \cap \mathcal{J}_2$ satisfies $\rho_1 = 0, \rho_2 > 0, \rho_3 = 0, \rho_4 > 0$, which requires $p_1 = P/\beta, q_1 = 0, p_2 = P/\beta, q_2 = 0$, from (B.53)–(B.56), (B.45), and (B.46). This results in $\Sigma_1 = \Sigma_2 < \Sigma$, which is consistent with the implication of $\boldsymbol{\lambda} \in \mathcal{L}_1$ only, and not with any other subset of $\boldsymbol{\lambda}$.

B8. For $\boldsymbol{\rho} \in \mathcal{I}_1 \cap \mathcal{J}_1$ or $\boldsymbol{\rho} \in \mathcal{I}_3 \cap \mathcal{J}_1$, we have $\boldsymbol{\lambda} \notin \mathcal{L}_1, \mathcal{L}_2$. This follows since in $\mathcal{I}_1 \cap \mathcal{J}_1$ or $\mathcal{I}_3 \cap \mathcal{J}_1$, $\boldsymbol{\rho}$ satisfies $\rho_3 > 0$ and $\rho_4 = 0$, which requires $q_2 = P/\bar{\beta}, p_2 = 0$. However, in $\mathcal{L}_1$ or $\mathcal{L}_2$, $\boldsymbol{\lambda}$ satisfies $\lambda_1 > 0, \lambda_2 > 0$, which requires $\Sigma_1 = \Sigma_2$ that results in $q_1 = \frac{c_{21}^2}{c_{12}^2} q_2 = \frac{c_{21}^2}{c_{12}^2} P/\bar{\beta}$. Under $c_{21}^2 > c_{12}^2$, the resulting $q_1 = \frac{c_{21}^2}{c_{12}^2} P/\bar{\beta} > P$ which violates the total power constraint.

The assertions of [B1]-[B8] above are summarized succinctly in Table B.1 below where the incompatibility between subsets $\mathcal{I}_k \cap \mathcal{J}_l$ and $\mathcal{L}_m$ is indicated by $\times$, whereas compatibility between the two, which corresponds to a specific link-gain regime (LGR), is labeled with the associated LGR.

Table B.1: Compatibility of subsets $\mathcal{I}_k \cap \mathcal{J}_l$ and $\mathcal{L}_m$.

$\mathcal{I}_k \cap \mathcal{J}_l$	$\mathcal{L}_3$	$\mathcal{L}_4$	$\mathcal{L}_5$	$\mathcal{L}_1$	$\mathcal{L}_2$
$\mathcal{I}_1 \cap \mathcal{J}_1$	$\times$	$\times$	$\times$	$\times$	$\times$
$\mathcal{I}_1 \cap \mathcal{J}_2$	$\times$	$\times$	$\times$	$\times$	$\times$
$\mathcal{I}_1 \cap \mathcal{J}_3$	$\times$	$\times$	$\times$	$\mathcal{A}_{\mathsf{c,cd}}$	$\mathcal{S}_{\mathsf{cd,cd}}$
$\mathcal{I}_2 \cap \mathcal{J}_1$	$\times$	$\times$	$\times$	$\times$	$\times$
$\mathcal{I}_2 \cap \mathcal{J}_2$	$\times$	$\times$	$\times$	$\mathcal{A}_{\mathsf{d,d}}$	$\times$
$\mathcal{I}_2 \cap \mathcal{J}_3$	$\times$	$\times$	$\times$	$\times$	$\times$
$\mathcal{I}_3 \cap \mathcal{J}_1$	$\times$	$\times$	$\times$	$\times$	$\times$
$\mathcal{I}_3 \cap \mathcal{J}_2$	$\times$	$\times$	$\times$	$\times$	$\times$
$\mathcal{I}_3 \cap \mathcal{J}_3$	$\times$	$\times$	$\times$	$\mathcal{A}_{\mathsf{cd,cd}}$	$\mathcal{S}_{\mathsf{cd,cd}}$

The Optimum Power Allocation: As illustrated in Table B.1, each LGR is associated with one or more mutually exclusive subsets of optimal Lagrange multipliers as follows: (i) $\mathcal{A}_{\mathsf{c,cd}}$ corresponds to $\boldsymbol{\rho} \in \mathcal{I}_1 \cap \mathcal{J}_3$ and $\boldsymbol{\lambda} \in \mathcal{L}_1$; (ii) $\mathcal{A}_{\mathsf{d,d}}$ corresponds to $\boldsymbol{\rho} \in \mathcal{I}_2 \cap \mathcal{J}_2$ and $\boldsymbol{\lambda} \in \mathcal{L}_1$; (iii) $\mathcal{S}_{\mathsf{cd,cd}}$ corresponds to $\boldsymbol{\rho} \in \mathcal{I}_3 \cap \mathcal{J}_3 \cup \mathcal{I}_1 \cap \mathcal{J}_3$ and $\boldsymbol{\lambda} \in \mathcal{L}_2$; and (iv) $\mathcal{A}_{\mathsf{cd,cd}}$ corresponds to $\boldsymbol{\rho} \in \mathcal{I}_3 \cap \mathcal{J}_3$ and $\boldsymbol{\lambda} \in \mathcal{L}_1$.

Next, we obtain the conditions of each LGR explicitly in terms of the channel parameters by characterizing the condition of the corresponding subsets of optimal Lagrange multipliers. In addition, we characterize the optimal power allocation in each LGR as well.

$\underline{\mathcal{A}_{\mathsf{c,cd}}}$: Since $\boldsymbol{\rho} \in \mathcal{I}_1 \cap \mathcal{J}_3$, it satisfies $\rho_1 > 0, \rho_2 = 0, \rho_3 = 0, \rho_4 = 0$, which requires $p_1 = 0, q_1 = P/\bar{\beta}$ from (B.45), (B.53), and (B.54), and also $p_2 > 0, q_2 > 0$ from (B.46), (B.55), and (B.56). In addition, $\boldsymbol{\lambda} \in \mathcal{L}_1$ implies $\Sigma_1 = \Sigma_2$ from (B.50) and (B.51), which, from the expressions of Σ_1 and Σ_2 in (3.47) and (3.48) gives $c_{21}^2 q_2 = c_{12}^2 P/\bar{\beta}$. Thus, we have $q_2 = P c_{12}^2/(c_{21}^2 \bar{\beta})$, and therefore, $p_2 = P(1 - c_{12}^2/c_{21}^2)/\beta$ from (B.46). Note that $P > 0$ and $c_{12}^2 < c_{21}^2$ are sufficient for $(q_1, p_2, q_2) \succ \mathbf{0}$, and any additional condition is not required for this non-negativity. Next, $\lambda_3 = 0$ is equivalent to $\Sigma_2 < \Sigma$, resulting in $P < \bar{\beta}(\gamma - 1)/c_{12}^2$, i.e., $P < \bar{\mathsf{P}}_4$ where $\gamma > 1$ is defined in (3.58). Then, from $\rho_3 = 0, \rho_4 = 0$, and (B.43)–(B.44), we have $\lambda_1 = (1 + P c_{12}^2/\bar{\beta})/(P(c_{21}^2 - c_{12}^2)/\beta + c_{12}^2/d_2^2)$. In addition, from $\rho_2 = 0$ and (B.41), the condition for $\rho_1 > 0$ is $\lambda_2 > d_1^2(1 + P c_{12}^2)/c_{12}^2$.

Since $\lambda_3 = 0$, from (B.40) we must satisfy $\lambda_2 + \lambda_1 = 1$, which subsequently gives $d_1^2(1 + Pc_{12}^2/\bar{\beta})/c_{12}^2 + (1 + Pc_{12}^2/\bar{\beta})/(P(c_{21}^2 - c_{12}^2)/\beta + c_{12}^2/d_2^2) < 1$, i.e., $P < \bar{\mathsf{P}}_2$ as in Table 3.1. Thus, the condition of $\mathcal{A}_{\mathsf{c,cd}}$ is

$$[\text{C1}]: \quad P < \bar{\mathsf{P}}_4, \quad P < \bar{\mathsf{P}}_2, \tag{B.58}$$

and the optimal transmit powers in $\mathcal{A}_{\mathsf{c,cd}}$ are as given in the second row of Table 3.1.

$\underline{\mathcal{A}_{\mathsf{d,d}}}$: Since $\boldsymbol{\rho} \in \mathcal{I}_2 \cap \mathcal{J}_2$, it satisfies $\rho_1 = 0, \rho_2 > 0, \rho_3 = 0, \rho_4 > 0$, which imply $p_1 = P/\beta, q_1 = 0, p_2 = P/\beta, q_2 = 0$, following (B.53)–(B.56) and (B.45) and (B.46). In addition, $\boldsymbol{\lambda} \in \mathcal{L}_1$ implies $\Sigma_1 = \Sigma_2 < \Sigma$ from (B.50)–(B.52), and assumption [A1] is sufficient for this. Next, using $\rho_1 = \rho_3 = 0$, and (B.41)–(B.44), the sufficient conditions for $\rho_2 > 0$ and $\rho_4 > 0$ are found to be $\lambda_2 < d_1^2/(c_{12}^2(1 + d_1^2 P/\beta))$ and $\lambda_1 < d_2^2/(c_{21}^2(1 + d_2^2 P/\beta))$, respectively. Since $\lambda_3 = 0$, and thus $\lambda_1 + \lambda_2 = 1$ from (B.40), the bounds on λ_1 and λ_2 are combined, which gives $d_1^2/(c_{12}^2(1 + d_1^2 P/\beta)) + d_2^2/(c_{21}^2(1 + d_2^2 P/\beta)) > 1$, i.e., $P < \bar{\mathsf{P}}_1$ as in Table 3.1. Finally, $P > 0$ is sufficient for $(p_1, p_2) \succ \mathbf{0}$. Therefore, the condition of $\mathcal{A}_{\mathsf{d,d}}$ is

$$[\text{C2}]: \quad P < \bar{\mathsf{P}}_1, \tag{B.59}$$

and the optimal transmit powers in $\mathcal{A}_{\mathsf{d,d}}$ are as given in the first row of Table 3.1.

$\underline{\mathcal{S}_{\mathsf{cd,cd}}}$: Since $\boldsymbol{\rho} \in \mathcal{I}_3 \cap \mathcal{J}_3$, it satisfies $\boldsymbol{\rho} = \mathbf{0}$, which imply $(p_1, q_1, p_2, q_2) \succ \mathbf{0}$, following (B.53)–(B.56). In addition, $\boldsymbol{\lambda} \in \mathcal{L}_2$ implies $\Sigma_1 = \Sigma$ and $\Sigma_2 = \Sigma$ following (B.50)–(B.52), which gives $q_1 = \frac{\gamma - 1}{c_{12}^2}$ and $q_2 = \frac{\gamma - 1}{c_{21}^2}$, respectively, with $\gamma > 1$. Thus, we also have $p_1 = P/\beta - \frac{\bar{\beta}(\gamma - 1)}{\beta c_{12}^2}$ and $p_2 = P/\beta - \frac{\bar{\beta}(\gamma - 1)}{\beta c_{21}^2}$ from (B.45) and (B.46). Note that the condition for $p_1 > 0$ is equivalent to $P > \bar{\beta}(\gamma - 1)/c_{12}^2$, i.e., $P > \bar{\mathsf{P}}_4$, which is also sufficient for $p_2 > 0$ due to $c_{21}^2 > c_{12}^2$.

Now, from $\boldsymbol{\rho} = \mathbf{0}$ and (B.41)–(B.44), and using the expressions of (p_1, q_1, p_2, q_2) as derived above, we find that $\lambda_1 = \beta\gamma/(Pc_{21}^2 + \beta c_{21}^2/d_2^2 - \bar{\beta}(\gamma - 1))$,
and $\lambda_2 = \beta\gamma/(Pc_{12}^2 + \beta c_{12}^2/d_1^2 - \bar{\beta}(\gamma - 1))$. We also note that $P > \bar{\mathsf{P}}_4$ is sufficient to ensure $\lambda_1 > 0, \lambda_2 > 0$. Next, to ensure $\lambda_3 > 0$, λ_1 and λ_2 must satisfy $\lambda_1 + \lambda_2 < 1$, which gives the condition $P > \bar{\mathsf{P}}_3$ as in $\mathcal{S}_{\mathsf{cd,cd}}$ in Table 3.1. In addition, note that $\gamma > 1$ is sufficient for $(q_1, q_2) \succ \mathbf{0}$.

Finally, the case with $\boldsymbol{\rho} \in \mathcal{I}_1 \cap \mathcal{J}_3$ only differ from that with $\mathcal{I}_3 \cap \mathcal{J}_3$ in that now $\rho_1 > 0$, and thus $p_1 = 0$. We note that $P = \bar{\mathsf{P}}_4$ is sufficient for $p_1 = 0$. We also note that the other sufficient condition, which follows from the conditions on λ_1 and λ_2 as in $\mathcal{I}_3 \cap \mathcal{J}_3$, is expressed by evaluating $g_3(P) < 1$ derived above at $P = \bar{\mathsf{P}}_4$. Therefore, the conditions of

$\mathcal{S}_{\mathsf{cd,cd}}$ are given by

$$[\text{C3}]: \quad P \geq \bar{\mathsf{P}}_4, \quad P > \bar{\mathsf{P}}_3, \tag{B.60}$$

whereas the optimal transmit powers are as given in the fourth row of Table 3.1.

$\underline{\mathcal{A}_{\mathsf{cd,cd}}}$: Since $\boldsymbol{\rho} \in \mathcal{I}_3 \cap \mathcal{J}_3$, it satisfies $\boldsymbol{\rho} = \mathbf{0}$ that implies $(p_1, q_1, p_2, q_2) \succ \mathbf{0}$, which follows from (B.53)–(B.56). Next, substituting $\boldsymbol{\rho} = \mathbf{0}$ in (B.41)–(B.44), (B.45) and (B.46), we find $\frac{\alpha\kappa}{2\mu_1}(\beta + \bar{\beta}\lambda_2) = P + \frac{\bar{\beta}}{c_{12}^2} + \frac{\beta}{d_1^2}$, and $\frac{\alpha\kappa}{2\mu_2}(\beta + \bar{\beta}\lambda_1) = P + \frac{\bar{\beta}}{c_{21}^2} + \frac{\beta}{d_2^2}$. In addition, $\boldsymbol{\lambda} \in \mathcal{L}_1$ implies $\Sigma_1 = \Sigma_2$, from which and from (3.47) and (3.48), we have $\lambda_2\mu_2 c_{12}^2 = \lambda_1\mu_1 c_{21}^2$. In addition, since $\lambda_3 = 0$ we also have $\lambda_1 = 1 - \lambda_2$. Combining these conditions, we get a quadratic equation of λ_2, $\hat{A}\lambda_2^2 - \hat{B}\lambda_2 + \hat{C} = 0$, where $\hat{A} := \bar{\beta}(E_1 - E_2)$, $\hat{B} := E_1 - E_2 + 2\beta E_2$, and $\hat{C} := \beta E_2$, with E_1, E_2 as defined in (3.59). One of its roots, $\lambda_2^{(1)} := (\hat{B} + \sqrt{\hat{B}^2 - 4\hat{A}\hat{C}})/2\hat{A}$, is infeasible as it violates (B.40) and the non-negativity of λ_2 for $\hat{A} > 0$ and $\hat{A} < 0$, respectively. Therefore, the valid solution is $\lambda_2 = (\hat{B} - \sqrt{\hat{B}^2 - 4\hat{A}\hat{C}})/2\hat{A}$. Next, from (B.42) and substituting λ_2 in $\frac{\alpha\kappa}{2\mu_1}(\beta + \bar{\beta}\lambda_2) = P + \frac{\bar{\beta}}{c_{12}^2} + \frac{\beta}{d_1^2}$, and with some algebraic simplification, we obtain $q_1 = F(P)/c_{12}^2$, where $F(P)$ is defined in (3.59). Finally, simple algebraic manipulations show that due to the mutual exclusiveness of the LGRs, the condition of $\mathcal{A}_{\mathsf{cd,cd}}$ is

$$[\text{C4}]: \quad \max(\bar{\mathsf{P}}_1, \bar{\mathsf{P}}_2) \leq P, \quad P \leq \bar{\mathsf{P}}_3, \tag{B.61}$$

and the optimal transmit powers are as given in the third row of Table 3.1.

B.12 Constraint Qualification for Problem $[\mathcal{P}1]$

The Mangasarian-Fromovitz Constraint Qualification

Consider an optimization problem $[\mathcal{P}']$ over variable $\boldsymbol{x} \in \mathbb{R}^v$ as follows

$$[\mathcal{P}'] \quad \text{minimize} \quad f(\boldsymbol{x}) \tag{B.62}$$

$$\text{subject to:} \quad g_i(\boldsymbol{x}) \leq 0, \quad i = 1, \ldots, m, \tag{B.63}$$

$$h_i(\boldsymbol{x}) = 0, \quad i = 1, \ldots, n, \tag{B.64}$$

where $f, g_i, h_i : \mathbb{R}^v \to \mathbb{R}$ are assumed to be continuously differentiable functions. For a feasible point $\boldsymbol{x}^*$ of problem $[\mathcal{P}']$, the active set is defined as

$$\mathcal{I}(\boldsymbol{x}^*) := \{i \in \{1, \ldots, m\} : g_i(\boldsymbol{x}^*) = 0\}. \tag{B.65}$$

The Mangasarian-Fromovitz constraint qualification conditions [155, Section 2.1.2], [154, Section 4.3] holds at $\boldsymbol{x}^*$ if

- $\nabla h_i(\boldsymbol{x}^*)$ are linearly independent for $i = 1, \ldots, n$, and
- there exists a direction vector $\boldsymbol{t} \in \mathbb{R}^v$ such that

  1. $\nabla h_i(\boldsymbol{x}^*)^T \boldsymbol{t} = 0, \quad i = 1, \ldots, n$, and
  2. $\nabla g_i(\boldsymbol{x}^*)^T \boldsymbol{t} < 0, \quad i \in \mathcal{I}(\boldsymbol{x}^*)$.

Constraints and Gradients in Problem $[\mathcal{P}1]$

Note that in problem $[\mathcal{P}1]$ defined in (3.51)-(3.57), all equality and inequality constraints are *differentiable*. Given a set of channel parameters $(d_1, d_2, c_{12}, c_{21}, \gamma, P)$, consider a *feasible point* $\boldsymbol{x} = (p_1, q_1, p_2, q_2, R)$ for problem $[\mathcal{P}1]$.

For convenience, we denote the inequality constraints (3.52), (3.53) and (3.54) by

$$
\begin{aligned}
g_1(\boldsymbol{x}) &= R - A_1 - \frac{\alpha\bar{\beta}}{2}\log\left(1 + c_{21}^2 q_2\right) - \frac{\alpha\beta}{2}\log\left(1 + d_1^2 p_1\right) - \frac{\alpha\beta}{2}\log\left(1 + d_2^2 p_2\right), \\
g_2(\boldsymbol{x}) &= R - A_2 - \frac{\alpha\bar{\beta}}{2}\log\left(1 + c_{12}^2 q_1\right) - \frac{\alpha\beta}{2}\log\left(1 + d_1^2 p_1\right) - \frac{\alpha\beta}{2}\log\left(1 + d_2^2 p_2\right), \\
g_3(\boldsymbol{x}) &= R - A - \frac{\alpha\beta}{2}\log\left(1 + d_1^2 p_1\right) - \frac{\alpha\beta}{2}\log\left(1 + d_2^2 p_2\right),
\end{aligned}
$$

where $A, A_1, A_2, \alpha, \beta, \bar{\beta}$ are non-negative constants that do not depend on the feasible point $\boldsymbol{x}$. The gradients of these functions, $\nabla g_1(\boldsymbol{x})$, $\nabla g_2(\boldsymbol{x})$ and $\nabla g_3(\boldsymbol{x})$, are given by

$$
\begin{aligned}
\nabla g_1(\boldsymbol{x}) &= \left(-\frac{\alpha\beta\kappa}{2}\frac{d_1^2}{1 + d_1^2 p_1}, \;\; 0, \;\; -\frac{\alpha\beta\kappa}{2}\frac{d_2^2}{1 + d_2^2 p_2}, \;\; -\frac{\alpha\bar{\beta}\kappa}{2}\frac{c_{21}^2}{1 + c_{21}^2 q_2}, \;\; 1\right)^T, \\
\nabla g_2(\boldsymbol{x}) &= \left(-\frac{\alpha\beta\kappa}{2}\frac{d_1^2}{1 + d_1^2 p_1}, \;\; -\frac{\alpha\bar{\beta}\kappa}{2}\frac{c_{12}^2}{1 + c_{12}^2 q_1}, \;\; -\frac{\alpha\beta\kappa}{2}\frac{d_2^2}{1 + d_2^2 p_2}, \;\; 0, \;\; 1\right)^T, \\
\nabla g_3(\boldsymbol{x}) &= \left(-\frac{\alpha\beta\kappa}{2}\frac{d_1^2}{1 + d_1^2 p_1}, \;\; 0, \;\; -\frac{\alpha\beta\kappa}{2}\frac{d_2^2}{1 + d_2^2 p_2}, \;\; 0, \;\; 1\right)^T,
\end{aligned}
$$

where $\mathbf{a}^T$ denotes the transpose of vector $\mathbf{a}$, and $\kappa = \log_2 e > 0$.

Next, the non-negativity constraints in (3.57) and their gradients are given as follows

$$
\begin{aligned}
g_4(\boldsymbol{x}) &= -p_1, & \nabla g_4(\boldsymbol{x}) &= (-1, 0, 0, 0, 0)^T, \\
g_5(\boldsymbol{x}) &= -q_1, & \nabla g_5(\boldsymbol{x}) &= (0, -1, 0, 0, 0)^T, \\
g_6(\boldsymbol{x}) &= -p_2, & \nabla g_6(\boldsymbol{x}) &= (0, 0, -1, 0, 0)^T, \\
g_7(\boldsymbol{x}) &= -q_2, & \nabla g_7(\boldsymbol{x}) &= (0, 0, 0, -1, 0)^T, \\
g_8(\boldsymbol{x}) &= -R, & \nabla g_8(\boldsymbol{x}) &= (0, 0, 0, 0, -1)^T.
\end{aligned}
$$

Finally, the equality constraints (3.55) and (3.56) are denoted as

$$h_1(\boldsymbol{x}) = \beta p_1 + \bar{\beta} q_1 - P, \tag{B.66}$$

$$h_2(\boldsymbol{x}) = \beta p_2 + \bar{\beta} q_2 - P, \tag{B.67}$$

and their gradients are given as follows

$$\nabla h_1(\boldsymbol{x}) = (\beta, \bar{\beta}, 0, 0, 0)^T \tag{B.68}$$

$$\nabla h_2(\boldsymbol{x}) = (0, 0, \beta, \bar{\beta}, 0)^T. \tag{B.69}$$

Problem $[\mathcal{P}1]$ **Satisfies the MFCQ**

First, note that $\nabla h_1(\boldsymbol{x})$ and $\nabla h_2(\boldsymbol{x})$ in (B.68)-(B.69) are linearly independent. Next, for notational convenience, the active set $\mathcal{I}(\boldsymbol{x})$ at $\boldsymbol{x}$ is denoted as

$$\begin{aligned} &\mathcal{I}(\boldsymbol{x}) = \mathcal{I}_1(\boldsymbol{x}) \cup \mathcal{I}_2(\boldsymbol{x}) \cup \mathcal{I}_3(\boldsymbol{x}), \quad \text{where} \\ &\mathcal{I}_1(\boldsymbol{x}) \subseteq \{1,2,3\} \setminus \emptyset, \quad \mathcal{I}_2(\boldsymbol{x}) \subseteq \{4,5\} \setminus (4,5), \quad \mathcal{I}_3(\boldsymbol{x}) \subseteq \{6,7\} \setminus (6,7). \end{aligned} \tag{B.70}$$

More specifically, $\mathcal{I}_1(\boldsymbol{x})$ is a subset of $\mathcal{I}(\boldsymbol{x})$ corresponding to constraints $\{g_i(\boldsymbol{x})\}_{i=1}^3$, and since at least one of the three constraints must be tight, $\mathcal{I}_1(\boldsymbol{x}) \neq \emptyset$. $\mathcal{I}_2(\boldsymbol{x})$ is a subset of $\mathcal{I}(\boldsymbol{x})$ corresponding to non-negative constraints $\{g_i(\boldsymbol{x})\}_{i=4}^5$, while $\mathcal{I}_3(\boldsymbol{x})$ is a subset of $\mathcal{I}(\boldsymbol{x})$ corresponding to non-negative constraints $\{g_i(\boldsymbol{x})\}_{i=6}^7$.

The intuition into why $\mathcal{I}_2(\boldsymbol{x}) \neq (4,5)$ and $\mathcal{I}_3(\boldsymbol{x}) \neq (6,7)$ are given below:

- due to (3.55), if $p_1 = 0$, then $q_1 > 0$. Hence $g_4(\boldsymbol{x})$ is tight $\implies$ $g_5(\boldsymbol{x})$ is not tight, i.e., $4 \in \mathcal{I}_2(\boldsymbol{x}) \implies 5 \notin \mathcal{I}_2(\boldsymbol{x})$. Similarly, if $q_1 = 0$, then due to (3.55), $p_1 > 0$, and hence $5 \in \mathcal{I}_2(\boldsymbol{x}) \implies 4 \notin \mathcal{I}_2(\boldsymbol{x})$. Moreover, if both $p_1 > 0, q_1 > 0$, then $\mathcal{I}_2(\boldsymbol{x}) = \emptyset$. Therefore, depending on $\boldsymbol{x}$, $\mathcal{I}_2(\boldsymbol{x})$ is either of $= \emptyset, \{4\}$, or $\{5\}\}$, but $\mathcal{I}_2(\boldsymbol{x}) \neq (4,5)$.
- the rationale for $\mathcal{I}_3(\boldsymbol{x}) \subseteq \{6,7\} \setminus (6,7)$ follows from those for $\mathcal{I}_2(\boldsymbol{x})$ explained above, after replacing constraints 4 and 5 on variables p_1 and q_1 for $\mathcal{I}_2(\boldsymbol{x})$ with constraints 6 and 7 on variables p_2 and q_2 for $\mathcal{I}_3(\boldsymbol{x})$ that satisfy (3.56).

Finally, constraint $g_8(\boldsymbol{x})$ is never tight, i.e., $8 \notin \mathcal{I}(\boldsymbol{x})$, since R is always $R > 0$ due the fact that $A, A_1, A_2 > 0$.

It thus follows from (B.70) that 63 combinations of active sets are possible. As an illustrative example, we first consider the following case and show that the MFCQ holds.

Case with Active Set $\mathcal{I}(\boldsymbol{x}) = \{1,2,3\} \cup \{4\} \cup \{6\}$: Clearly, in this case constrains $g_1(\boldsymbol{x}), g_2(\boldsymbol{x})$ and $g_3(\boldsymbol{x})$, and non-negativity constraints $g_4(\boldsymbol{x})$ and $g_6(\boldsymbol{x})$ are tight, along with the equality constraints $h_1(\boldsymbol{x})$ and $h_2(\boldsymbol{x})$.

Consider a direction vector $\boldsymbol{t} = (t_1, t_2, t_3, t_4, t_5)^T$. First, to satisfy $\nabla h_1(\boldsymbol{x})^T \boldsymbol{t} = 0$, and $\nabla h_2(\boldsymbol{x})^T \boldsymbol{t} = 0$, vector $\boldsymbol{t}$ needs to satisfy

$$\begin{aligned} \nabla h_1(\boldsymbol{x})^T \boldsymbol{t} = 0 &\implies t_2 = -\beta/\bar{\beta} t_1, \\ \nabla h_2(\boldsymbol{x})^T \boldsymbol{t} = 0 &\implies t_4 = -\beta/\bar{\beta} t_3. \end{aligned} \tag{B.71}$$

For the MFCQ to hold, in addition to (B.71), $\nabla g_k(\boldsymbol{x})^T \boldsymbol{t} < 0, k = 1,2,3,4,6$, must also hold. For $\nabla g_4(\boldsymbol{x})^T \boldsymbol{t} < 0$ and $\nabla g_6(\boldsymbol{x})^T \boldsymbol{t} < 0$ to hold, $\boldsymbol{t}$ must satisfy

$$\nabla g_4(\boldsymbol{x})^T \boldsymbol{t} < 0 \iff t_1 > 0, \quad \text{and} \quad \nabla g_6(\boldsymbol{x})^T \boldsymbol{t} < 0 \iff t_3 > 0.$$

After substituting for t_2 and t_4 as in (B.71), for $\nabla g_1(\boldsymbol{x})^T \boldsymbol{t} < 0$ to hold, $\boldsymbol{t}$ must satisfy

$$\nabla g_1(\boldsymbol{x})^T \boldsymbol{t} < 0 \iff t_5 + t_3 \underbrace{\frac{\alpha\beta\kappa}{2}\left(\frac{c_{21}^2}{1 + c_{21}^2 q_2} - \frac{d_2^2}{1 + d_2^2 p_2}\right)}_{=:s_1} - t_1 \underbrace{\frac{\alpha\beta\kappa}{2}\frac{d_1^2}{1 + d_1^2 p_1}}_{>0} < 0. \tag{B.72}$$

While $t_1 > 0$ and $t_3 > 0$, t_5 is free. For inequality (B.72) to hold, if $s_1 > 0$ finite, $t_5 < 0$ needs to be taken to be a sufficiently large negative number, while if $s_1 \leq 0$, any $t_5 < 0$ is sufficient.

Similarly, substituting for t_2 and t_4 as in (B.71), for $\nabla g_2(\boldsymbol{x})^T \boldsymbol{t} < 0$ to hold, $\boldsymbol{t}$ must satisfy

$$\nabla g_2(\boldsymbol{x})^T \boldsymbol{t} < 0 \iff t_5 + t_1 \underbrace{\frac{\alpha\beta\kappa}{2}\left(\frac{c_{12}^2}{1 + c_{12}^2 q_1} - \frac{d_1^2}{1 + d_1^2 p_1}\right)}_{=:s_2} - t_3 \underbrace{\frac{\alpha\beta\kappa}{2}\frac{d_2^2}{1 + d_2^2 p_2}}_{>0} < 0. \tag{B.73}$$

As before, $t_1 > 0, t_3 > 0$, and t_5 is free. For inequality (B.73) to hold, if $s_2 > 0$ finite, $t_5 < 0$ needs to be taken to be a sufficiently large negative number, while if $s_2 \leq 0$, any $t_5 < 0$ is sufficient.

Finally, for $\nabla g_3(\boldsymbol{x})^T \boldsymbol{t} < 0$ to hold, $\boldsymbol{t}$ must satisfy

$$\nabla g_3(\boldsymbol{x})^T \boldsymbol{t} < 0 \iff t_5 - t_1 \underbrace{\frac{\alpha\beta\kappa}{2}\frac{d_1^2}{1 + d_1^2 p_1}}_{>0} - t_3 \underbrace{\frac{\alpha\beta\kappa}{2}\frac{d_2^2}{1 + d_2^2 p_2}}_{>0} < 0. \tag{B.74}$$

Since $t_1 > 0, t_3 > 0$, inequality (B.73) holds by taking any $t_5 < 0$.

Therefore, we have found $\boldsymbol{t} = (t_1, t_2, t_3, t_4, t_5)^T$, where $t_1, t_3 > 0$, $t_2 = -\beta/\bar{\beta} t_1, t_4 = -\beta/\bar{\beta} t_3$, and t_5 is the minimum of the two sufficiently large negative numbers found for inequalities (B.72) and (B.73). Therefore, the problem satisfies the MFCQ for the active

set $\mathcal{I}(\boldsymbol{x}) = \{1,2,3\} \cup \{4\} \cup \{6\}$.

Other Active Sets: While the above analysis is for the specific active set $\mathcal{I}(\boldsymbol{x}) = \{1,2,3\} \cup \{4\} \cup \{6\}$, similar analysis holds for all other active sets. Since the analysis for other active sets is tedious and repetitive, we omit their explicit derivation.

The Linear Independence Constraint Qualification

For problem $[\mathcal{P}']$ in (B.62)-(B.64), if $\boldsymbol{x}^*$ is a feasible point, the linear independence constraint qualification (LICQ) [155, Section 2.1.2], [156] is said to hold if

- $\{\nabla h_i(\boldsymbol{x}^*), \;\; i = 1,\ldots,n\}, \quad \text{and} \quad \{\nabla g_i(\boldsymbol{x}^*), \;\; i \in \mathcal{I}(\boldsymbol{x}^*)\}$ are linearly independent.

Problem $[\mathcal{P}1]$ Does Not Satisfy the LICQ in General

We observe that $[\mathcal{P}1]$ *does not satisfy the LICQ in general.* For a feasible point $\boldsymbol{x}$ for problem $[\mathcal{P}1]$, the total number of constraints in an active set $\mathcal{I}(\boldsymbol{x})$ defined in (B.70), may potentially be as large as 5. In addition, there are two equality constraints as in (B.66)-(B.67). Therefore, the total number of gradient vectors to be considered for LICQ can be as large as 7. However, the feasible point $\boldsymbol{x}$ has the dimension of 5, and thus no more than 5 vectors can be linearly independent. Thus $[\mathcal{P}1]$ does not satisfy the LICQ in general.

While LICQ is not guaranteed to hold for $[\mathcal{P}1]$ in general, it may hold for active sets with cardinality 3 or less. However, there exists a large number of such active sets for $[\mathcal{P}1]$, and verifying if LICQ holds for all such sets is tedious and repetitive. Therefore, instead of treating the general case of problem $[\mathcal{P}1]$, we limit the analysis for the case in Section 3.4.2 as this specific case has been studied in detail in this chapter. In particular, we consider the case where problem $[\mathcal{P}1]$ is subject to the assumptions [A1] through [A4] in Section 3.4 and condition $c_{21}^2 > c_{12}^2$ in Section 3.4.2.

As a result, the potential number of active sets for the problem now reduces to only four, given by either $\mathcal{I}(\boldsymbol{x}) = \{1,2,4,6\}$, $\mathcal{I}(\boldsymbol{x}) = \{1,2,4\}$, $\mathcal{I}(\boldsymbol{x}) = \{1,2,3\}$, or $\mathcal{I}(\boldsymbol{x}) = \{1,2\}$, depending on the appropriate feasible $\boldsymbol{x}$, presented as valid LGRs in Table B.1 and detailed in Section B.11.

Active set $\mathcal{I}(\boldsymbol{x}) = \{1, 2, 4, 6\}$: LICQ *does not hold*, and the optimal Lagrange multipliers are *non-unique* [156].
For active set $\mathcal{I}(\boldsymbol{x}) = \{1, 2, 4, 6\}$, when the two equality constraints are also included, for LICQ to hold, 6 vectors $\{\nabla h_1(\boldsymbol{x}), \nabla h_2(\boldsymbol{x}), \nabla g_1(\boldsymbol{x}), \nabla g_2(\boldsymbol{x}), \nabla g_4(\boldsymbol{x}), \nabla g_6(\boldsymbol{x})\}$ need to be linearly independent. Since the feasible point has a dimension of 5, these vectors are *linearly dependent*, and hence *LICQ does not hold*. Therefore, the optimal Lagrange multipliers are non-unique [156].

Active set $\mathcal{I}(\boldsymbol{x}) = \{1, 2, 4\}$: LICQ *holds*, and the optimal Lagrange multipliers are *unique*.
For active set $\mathcal{I}(\boldsymbol{x}) = \{1, 2, 4\}$, when the two equality constraints are included, for LICQ to hold, 5 vectors $\{\nabla h_1(\boldsymbol{x}), \nabla h_2(\boldsymbol{x}), \nabla g_1(\boldsymbol{x}), \nabla g_2(\boldsymbol{x}), \nabla g_4(\boldsymbol{x})\}$ need to be linearly independent. Simple algebraic manipulations show that these vectors are indeed *linearly independent*. Hence, *LICQ holds*, and the optimal Lagrange multipliers are *unique* [156].

LICQ also holds for $\mathcal{I}(\boldsymbol{x}) = \{1, 2, 3\}$, and $\mathcal{I}(\boldsymbol{x}) = \{1, 2\}$ as well, but the details are omitted to avoid repetition.

B.13 The Optimal Power Allocation for the Asymmetric DCLIC

Recall that in Section 3.4.1, problem $[\mathcal{P}1]$ was formulated under assumptions [A2], i.e., $A_1 = A_2$ in (3.50), which provides symmetry in the underlying microwave IC, as well as $c_{12}^2 < c_{21}^2$ in the mm-wave cross-links. We now characterize the optimal power allocation scheme for the asymmetric DCLIC where the symmetry condition of [A2] does not necessarily hold. More specifically, keeping $c_{12}^2 < c_{21}^2$ unchanged, we now formulate problem $[\mathcal{P}1]$ under a more general assumption [A2G]

$$\text{[A2G]:} \quad A_2 < A_1. \tag{B.75}$$

In this case, *seven* LGRs are needed to describe the optimal allocation scheme (IC-OPA). The definition and the expressions of the optimal powers of the LGRs are found by solving the problem $[\mathcal{P}1]$ under [A2G] in a manner similar to that for problem $[\mathcal{P}1]$ under [A2]. Hence, the explicit derivation is omitted, and only the resulting LGRs are discussed.

Of the seven LGRs, four LGRs denoted $\hat{\mathcal{A}}_{\mathsf{d,d}}$, $\hat{\mathcal{A}}_{\mathsf{c,cd}}$, $\hat{\mathcal{A}}_{\mathsf{cd,cd}}$, and $\hat{\mathcal{S}}_{\mathsf{cd,cd}}$ are counterparts of

Table B.2: The optimal link powers for the asymmetric DCLIC in terms of channel parameters $\boldsymbol{a} = (d_1, d_2, c_{12}, c_{21}, \gamma_1, \gamma_2, \eta, P)$, with $\hat{F}(P)$ defined in (B.76), and γ_1, γ_2 and η defined in (B.77).

LGR	**Optimal Power Allocation**			
$\hat{\mathcal{A}}_{\mathsf{d,d}}$	$p_1 = \dfrac{P}{\beta}$,	$q_1 = 0$,	$p_2 = \dfrac{P}{\beta}$,	$q_2 = 0$
$\hat{\mathcal{A}}_{\mathsf{c,cd}}$	$p_1 = 0$,	$q_1 = \dfrac{P}{\bar{\beta}}$,	$p_2 = P - q_2$,	$q_2 = \dfrac{Pc_{12}^2/\bar{\beta} + 1 - \eta}{\eta c_{21}^2}$
$\hat{\mathcal{S}}_{\mathsf{cd,cd}}$	$p_1 = \dfrac{P}{\beta} - \dfrac{\bar{\beta}(\gamma_1 - 1)}{\beta c_{12}^2}$,	$q_1 = \dfrac{\gamma_1 - 1}{c_{12}^2}$,	$p_2 = \dfrac{P}{\beta} - \dfrac{\bar{\beta}(\gamma_2 - 1)}{\beta c_{21}^2}$,	$q_2 = \dfrac{\gamma_2 - 1}{c_{21}^2}$
$\hat{\mathcal{A}}_{\mathsf{cd,cd}}$	$p_1 = \dfrac{P - \bar{\beta} q_1}{\beta}$,	$q_1 = \dfrac{\hat{F}(P)}{c_{12}^2}$,	$p_2 = \dfrac{P - \bar{\beta} q_2}{\beta}$,	$q_2 = \dfrac{1 + c_{12}^2 q_1}{\eta c_{21}^2} - \dfrac{1}{c_{21}^2}$

the four LGRs $\mathcal{A}_{\mathsf{d,d}}, \mathcal{A}_{\mathsf{c,cd}}, \mathcal{A}_{\mathsf{cd,cd}}$, and $\mathcal{S}_{\mathsf{cd,cd}}$ in Table 3.1, in the sense that the mode of link power allocation in an LGR $\hat{\mathcal{A}}_{(.,.)}$ in the former set is similar to its counterpart $\mathcal{A}_{(.,.)}$ in the latter set as detailed below. In addition to these four LGRs, three "new" LGRs $\hat{\mathcal{A}}_{\mathsf{c,d}}, \hat{\mathcal{A}}_{\mathsf{cd,d}}$ and $\hat{\mathcal{S}}_{\mathsf{cd,d}}$ are found: the mode of link power allocation in these LGRs are different than those in any of the LGRs in Table 3.1, and hence designated "new".

In Table B.2, the optimal power allocations for the first four LGRs are presented, where

$$\hat{F}(P) := \frac{1}{2\bar{\beta}(1+\beta)}\left(E_1 + \eta E_2 - \sqrt{(E_1 - \eta E_2)^2 + 4\beta^2 \eta E_1 E_2}\right) - 1, \tag{B.76}$$

with E_1, E_2, defined in (3.59), and γ_1, γ_2 and η are defined as follows

$$\begin{aligned}
\gamma_1 &:= \left(\frac{(1+P_1)(1+P_2)}{1+P_2+a_{21}^2 P_2}\right)^{1/\alpha\bar{\beta}}, \\
\gamma_2 &:= \left(\frac{(1+P_1)(1+P_2)}{1+P_1+a_{12}^2 P_1}\right)^{1/\alpha\bar{\beta}}, \\
\eta &:= \left(\frac{1+P_1+a_{21}^2 P_2}{1+P_2+a_{12}^2 P_1}\right)^{1/\alpha\bar{\beta}}.
\end{aligned} \tag{B.77}$$

We omit the definitions of these LGRs for the sake of conciseness. These LGRs, i.e., $\hat{\mathcal{A}}_{\mathsf{d,d}}, \hat{\mathcal{A}}_{\mathsf{c,cd}}, \hat{\mathcal{A}}_{\mathsf{cd,cd}}$, and $\hat{\mathcal{S}}_{\mathsf{cd,cd}}$, are similar to their counterparts in Table 3.1 in that they follow the same Waterfilling-like (WF-like) and saturation properties: (a) when the power budget P is sufficiently small, the IC-OPA allocates power to only the direct-links (in $\hat{\mathcal{A}}_{\mathsf{d,d}}$), or in both cross-links and one direct-link (in $\hat{\mathcal{A}}_{\mathsf{c,cd}}$) depending on whether the direct- or the cross-links are *stronger* in a sense similar to that for $\mathcal{A}_{\mathsf{d,d}}$ and $\mathcal{A}_{\mathsf{c,cd}}$; (b) as P is increased, the IC-OPA allocates power to all channels (in $\hat{\mathcal{A}}_{\mathsf{cd,cd}}$) as in $\mathcal{A}_{\mathsf{cd,cd}}$; and (c) when P is sufficiently large, both cross-links become saturated (in $\hat{\mathcal{S}}_{\mathsf{cd,cd}}$) as in $\mathcal{S}_{\mathsf{cd,cd}}$.

Table B.3: Definition of the "new" LGRs and optimal link powers for the asymmetric DCLIC in terms of parameters $\boldsymbol{a} = (d_1, d_2, c_{12}, c_{21}, \gamma_1, \gamma_2, \eta, P)$, with γ_1, γ_2 and η defined in (B.77), and the threshold powers defined in (B.78).

LGR	**Optimal Power Allocation**	
$\hat{\mathcal{A}}_{\mathsf{c,d}} := \{\boldsymbol{a} : 0 \le P \le \min(\bar{\mathsf{P}}_5, \bar{\mathsf{P}}_6)\}$	$p_1 = 0,$	$q_1 = \dfrac{P}{\bar{\beta}},$
	$p_2 = \dfrac{P}{\beta},$	$q_2 = 0$
$\hat{\mathcal{A}}_{\mathsf{cd,d}} := \{\boldsymbol{a} : \max(\bar{\mathsf{P}}_6, \bar{\mathsf{P}}_7) \le P \le \bar{\mathsf{P}}_8\}$	$p_1 = P + \dfrac{\bar{\beta}}{c_{12}^2} - \dfrac{\bar{\beta}}{d_1^2},$	$q_1 = P + \dfrac{\beta}{d_1^2} - \dfrac{\beta}{c_{12}^2},$
	$p_2 = \dfrac{P}{\beta},$	$q_2 = 0$
$\hat{\mathcal{S}}_{\mathsf{cd,d}} := \{\boldsymbol{a} : \max(\bar{\mathsf{P}}_5, \bar{\mathsf{P}}_8) \le P \le \bar{\mathsf{P}}_9\}$	$p_1 = \dfrac{P}{\beta} - \dfrac{\bar{\beta}(\eta-1)}{\beta c_{12}^2},$	$q_1 = \dfrac{\eta - 1}{c_{12}^2},$
	$p_2 = \dfrac{P}{\beta},$	$q_2 = 0$

Note that, if assumption [A2G] simplifies to [A2], we have $\eta = 1$ and $\gamma_1 = \gamma_2$, and the powers in Table B.2 simplify to those in Table 3.1.

Next, in Table B.3 we present definitions of the three "new" LGRs and the corresponding optimal link powers, where the threshold power are defined as below, with $\varrho[f(P)]$ denoting the non-negative square root of polynomial $f(P)$

$$
\begin{aligned}
\bar{\mathsf{P}}_5 &:= \bar{\beta}\frac{\eta - 1}{c_{12}^2}, \\
\bar{\mathsf{P}}_6 &:= \bar{\beta}\left(\frac{1}{d_1^2} - \frac{1}{c_{12}^2}\right), \\
\bar{\mathsf{P}}_7 &:= \beta\left(\frac{1}{c_{12}^2} - \frac{1}{d_1^2}\right), \\
\bar{\mathsf{P}}_8 &:= \frac{\eta - \bar{\beta}}{c_{12}^2} - \frac{\beta}{d_1^2}, \\
\bar{\mathsf{P}}_9 &:= \varrho\left[\frac{d_2^2}{c_{21}^2(1 + d_2^2 P/\beta)} + \frac{\beta\eta}{P c_{12}^2 + \bar{\beta} + \beta c_{12}^2/d_1^2 - \bar{\beta}\eta}\right] - 1.
\end{aligned}
\tag{B.78}
$$

The power allocation for the new LGRs can be interpreted following the WF-like and saturation properties as follows. In $\hat{\mathcal{A}}_{\mathsf{c,d}}$, we have $\bar{\mathsf{P}}_6 > 0 \iff d_1^2 < c_{12}^2$ (i.e., the cross-link from source S_1 is stronger than the corresponding direct-link), and thus following the WF-like property, for sufficiently small P, the power budget is entirely allocated to the

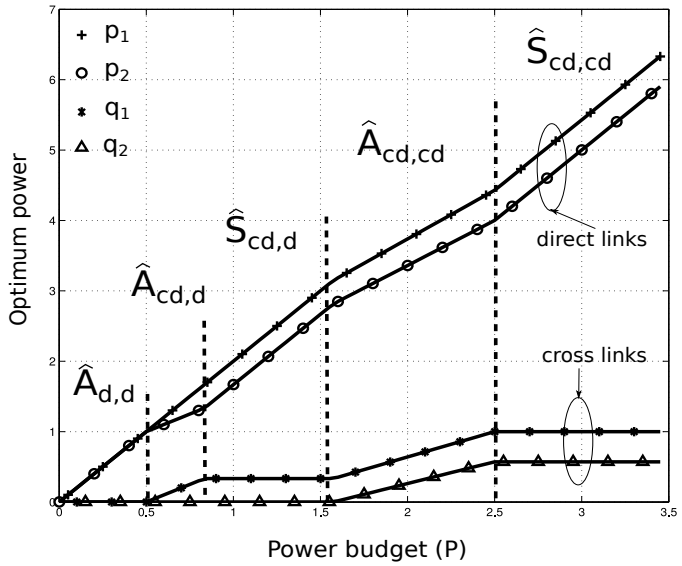


Figure B.1: An example of the optimum power allocation in the asymmetric case. The IC-OPA follows path $\hat{\mathcal{A}}_{\mathsf{d,d}} \to \hat{\mathcal{A}}_{\mathsf{cd,d}} \to \hat{\mathcal{S}}_{\mathsf{cd,d}} \to \hat{\mathcal{A}}_{\mathsf{cd,cd}} \to \hat{\mathcal{S}}_{\mathsf{cd,cd}}$, where $\hat{\mathcal{A}}_{\mathsf{cd,d}}$ and $\hat{\mathcal{S}}_{\mathsf{cd,d}}$ are new.

cross-link, i.e., $q_1 = P/\bar{\beta}, p_1 = 0$. However, under assumptions [A2G] and $c_{12}^2 < c_{21}^2$, it is sub-optimal to allocate power to the other cross-link as such allocation results in $\Sigma_2 < \Sigma_1$, and does not increase $\min\{\Sigma_2, \Sigma_1\}$ as much as possible, resulting in a suboptimal sum-rate. Therefore, for source S_2, it is optimal to allocate P only to the direct-link, i.e., $p_2 = P/\beta, q_2 = 0$ as long as $P < \max(\bar{\mathsf{P}}_5, \bar{\mathsf{P}}_6)$. This results in the maximum increases of the bottleneck rate Σ_2 in this range of P.

As P increases, additional benefits from transmitting only in the cross-link from source S_1 reduces. Hence, when $P > \max\{\bar{\mathsf{P}}_6, \bar{\mathsf{P}}_7\}$, it becomes optimal to allocate a fraction of P to the direct-link as well following the WF-like property, resulting in the allocation in $\hat{\mathcal{A}}_{\mathsf{cd,d}}$, and this allocation remains optimal for $P < \bar{\mathsf{P}}_8$.

Note that in $\hat{\mathcal{A}}_{\mathsf{c,d}}$ and $\hat{\mathcal{A}}_{\mathsf{cd,d}}$, the power budget P is small enough such that the optimal sum-rate is $R = \Sigma_2 < \Sigma_1$. When $P > \max\{\bar{\mathsf{P}}_5, \bar{\mathsf{P}}_8\}$, enough power has been allocated to the cross-link of source S_1 to cause $\Sigma_2 = \Sigma_1$. As P increases further, it becomes optimal to maintain $\Sigma_2 = \Sigma_1$ by keeping this cross-link power fixed at $q_1 = \dfrac{\eta - 1}{c_{12}^2}$, while all additional increments of P are allotted to the direct-link only. This allocation, described by $\hat{\mathcal{S}}_{\mathsf{c,d}}$, remains optimal as long as $P < \bar{\mathsf{P}}_9$, beyond which transmitting in all four links as in $\hat{\mathcal{A}}_{\mathsf{cd,cd}}$ in Table B.2 becomes optimal.

In summary, it is apparent that while the optimal power allocation scheme for the asymmetric DCLIC with [A2G] is more complex (with seven LGRs) than that of the case with [A2] (with only four LGRs), it can nonetheless be satisfactorily explained by the WF-like and saturation properties similar to those for the case with [A2]. Since the power allocation does not reveal nay additional properties, for the clarity of exposition, we omit any further discussion on this case.

Finally, in Figure B.1, we provide an example of the optimum power allocation in the asymmetric case with parameters with $a_{12}^2 = 1.5, a_{21}^2 = 2, P_1 = P_2 = 2, \alpha = 2, \beta = 0.5, d_1^2 = 1, d_2^2 = 0.5, c_{12}^2 = c_{21}^2 = 0.5$. In this example, the optimal power allocation follows path $\hat{\mathcal{A}}_{\mathsf{d,d}} \to \hat{\mathcal{A}}_{\mathsf{cd,d}} \to \hat{\mathcal{S}}_{\mathsf{cd,d}} \to \hat{\mathcal{A}}_{\mathsf{cd,cd}} \to \hat{\mathcal{S}}_{\mathsf{cd,cd}}$, where LGRs $\hat{\mathcal{A}}_{\mathsf{cd,d}}$ and $\hat{\mathcal{S}}_{\mathsf{cd,d}}$ are new.

Appendix C

Appendices for Chapter 4

C.1 Proof of Theorem 4.1

Outer Bound: Assume that source S_k transmits $M_k, k \in \{1,2\}$. Since the destination knows $\boldsymbol{H}_{\mathsf{D}}^n$ and $\bar{\boldsymbol{H}}_{\mathsf{D}}^{n_1}$ where $\boldsymbol{H}_{\mathsf{D},i} := \{H_{m\mathsf{D},i}\}_{m\in\{1,2,\mathsf{R}\}}, i = 1,\ldots,n$, $\bar{\boldsymbol{H}}_{\mathsf{D},\ell} := \{\bar{H}_{m\mathsf{D},\ell}\}_{m\in\{1,2,\mathsf{R}\}}, \ell = 1,\ldots,n_1$, by Fano's inequality we have

$$
\begin{aligned}
nR_1 &= H(M_1) \\
&= I(M_1; Y_{\mathsf{D}}^n, \bar{Y}_{\mathsf{RD}}^{n_1}, \bar{Y}_{1\mathsf{D}}^{n_1}) + H(M_1|Y_{\mathsf{D}}^n, \bar{Y}_{\mathsf{RD}}^{n_1}, \bar{Y}_{1\mathsf{D}}^{n_1}) \\
&\le I(X_1^n, \hat{X}_1^{n_1}, \bar{X}_1^{n_1}; Y_{\mathsf{D}}^n, \bar{Y}_{\mathsf{RD}}^{n_1}, \bar{Y}_{1\mathsf{D}}^{n_1}, \boldsymbol{H}_{\mathsf{D}}^n, \bar{\boldsymbol{H}}_{\mathsf{D}}^{n_1}) + n\epsilon_n \\
&\overset{(a)}{\le} I(X_1^n, \hat{X}_1^{n_1}, \bar{X}_1^{n_1}; Y_{\mathsf{D}}^n, \bar{Y}_{\mathsf{RD}}^{n_1}, \bar{Y}_{1\mathsf{D}}^{n_1}|\boldsymbol{H}_{\mathsf{D}}^n, \bar{\boldsymbol{H}}_{\mathsf{D}}^{n_1}) + n\epsilon_n \\
&\overset{(b)}{\le} I(X_1^n, \hat{X}_1^{n_1}; Y_{\mathsf{D}}^n, \bar{Y}_{\mathsf{RD}}^{n_1}|\boldsymbol{H}_{\mathsf{D}}^n, \bar{H}_{\mathsf{RD}}^{n_1}) \\
&\quad + \sum_{\ell=1}^{n_1} h(\bar{G}_{1\mathsf{D},\ell}^{1/2} e^{j\bar{\Theta}_{1\mathsf{D},\ell}} \bar{X}_{1,\ell} + \bar{Z}_{1\mathsf{D},\ell}|\bar{G}_{1\mathsf{D},\ell}, \bar{\Theta}_{1\mathsf{D},\ell}) - h(\bar{Z}_{1\mathsf{D},\ell}) + n\epsilon_n \\
&\overset{(c)}{\le} I(X_1^n, \hat{X}_1^{n_1}; Y_{\mathsf{D}}^n, \bar{Y}_{\mathsf{RD}}^{n_1}|\boldsymbol{H}_{\mathsf{D}}^n, \bar{H}_{\mathsf{RD}}^{n_1}) + \sum_{l=1}^{n_1} \mathbb{E}[\log(1 + \bar{G}_{1\mathsf{D}}\bar{P}_{1,l})] + n\epsilon_n \\
&\overset{(d)}{\le} I(X_1^n, \hat{X}_1^{n_1}; Y_{\mathsf{D}}^n, \bar{Y}_{\mathsf{RD}}^{n_1}|\boldsymbol{H}_{\mathsf{D}}^n, \bar{H}_{\mathsf{RD}}^{n_1}) + n_1\mathbb{E}[\tilde{\mathsf{C}}(\bar{G}_{1\mathsf{D}}\bar{P}_1)] + n\epsilon_n \qquad \text{(C.1)}
\end{aligned}
$$

where (a) follows since $(X_1^n, \hat{X}_1^{n_1}, \bar{X}_1^{n_1}) \perp\!\!\!\perp (\boldsymbol{H}_{\mathsf{D}}^n, \bar{\boldsymbol{H}}_{\mathsf{D}}^{n_1})$; (b) follows by first expanding (a) into four $I(.;.)$ terms using chain rule where two $I(.;.)$ terms turn out to be zero due to Markov chains (MC) $\bar{X}_1^{n_1} \to (X_1^n, \hat{X}_1^{n_1}, \boldsymbol{H}_{\mathsf{D}}^n, \bar{H}_{\mathsf{RD}}^{n_1}) \to (Y_{\mathsf{D}}^n, \bar{Y}_{\mathsf{RD}}^{n_1})$, and $(X_1^n, \hat{X}_1^{n_1}, \boldsymbol{H}_{\mathsf{D}}^n, \bar{H}_{\mathsf{RD}}^{n_1}, Y_{\mathsf{D}}^n, \bar{Y}_{\mathsf{RD}}^{n_1}) \to (\bar{X}_1^{n_1}, \bar{H}_{1\mathsf{D}}^{n_1}) \to \bar{Y}_{1\mathsf{D}}^{n_1}$; the last two terms follow from the Gaussian model and applying chain rule and unconditioning to one of the remaining $I(.)$ terms; (c) follows from maximizing the

first $h(.)$ term in (b) by using $\bar{X}_{1,\ell} \sim \mathcal{CN}(0, \bar{P}_{1,\ell})$ where $\frac{1}{n_1}\sum_{\ell=1}^{n_1} \bar{P}_{1,\ell} \leq \bar{P}_1$ and expectations are over $\bar{G}_{1\mathsf{D},\ell}$ i.i.d., with $\tilde{\mathsf{C}}(x) = \log(1+x)$; (d) follows by applying the Jensen's inequality.

Bounding R_2 similarly, the following bounds

$$R_k \leq \frac{1}{n} I(X_k^n, \hat{X}_k^{n_1}; Y_{\mathsf{D}}^n, \bar{Y}_{\mathsf{RD}}^{n_1} | \boldsymbol{H}_{\mathsf{D}}^n, \bar{H}_{\mathsf{RD}}^{n_1}) + \frac{n_1}{n}\mathbb{E}[\tilde{\mathsf{C}}(\bar{G}_{k\mathsf{D}}\bar{P}_k)], \quad k \in \{1,2\} \tag{C.2}$$

are found where expectations are over $\bar{G}_{k\mathsf{D}}$. Taking $n \to \infty$ such that $n_1/n \to \alpha$ and $\epsilon_n \to 0$, then gives the bounds in Theorem 4.1, for some empirical probability mass function (pmf) distributed as a

$$\begin{aligned}
&p(x_1^n, \hat{x}_1^{n_1}, \bar{x}_1^{n_1}) p(x_2^n, \hat{x}_2^{n_1}, \bar{x}_2^{n_1}) \prod_{i=1}^{n} p(y_{\mathsf{R},i}, y_{\mathsf{D},i} | x_{1,i}, x_{2,i}, x_{\mathsf{R},i}) \\
&\prod_{i=1}^{n} p(x_{\mathsf{R},i} | y_{\mathsf{R}}^{i-1}, \{\bar{y}_{k\mathsf{R}}^{n_1(i-1)}, h_{k\mathsf{R}}^{i-1}, \bar{h}_{k\mathsf{R}}^{n_1(i-1)}\}_{k=1}^{2}) \prod_{\ell=1}^{n_1} p(\bar{x}_{\mathsf{R},\ell} | y_{\mathsf{R}}^{n(l-1)}, \{\bar{y}_{k\mathsf{R}}^{l-1}, h_{k\mathsf{R}}^{n(l-1)}, \bar{h}_{k\mathsf{R}}^{l-1}\}_{k=1}^{2}) \\
&\prod_{\ell=1}^{n_1} p(\bar{y}_{\mathsf{RD},\ell} | \bar{x}_{\mathsf{R},\ell})\; p(\bar{y}_{1\mathsf{D},\ell} | \bar{x}_{1,\ell})\; p(\bar{y}_{2\mathsf{D},\ell} | \bar{x}_{2,\ell})\; p(\bar{y}_{1\mathsf{R},\ell} | \hat{x}_{1,\ell})\; p(\bar{y}_{2\mathsf{R},\ell} | \hat{x}_{2,\ell}).
\end{aligned} \tag{C.3}$$

Finally, the outer bound to the capacity region of the underlying R-MARC is given by the first $I(.;.)$ term in (C.2).

Achievability: We pick integers (n, n_1) and a distribution that factors as (C.3), and then code over t blocks of symbols together. Define $U_k := (X_k^n, \hat{X}_k^{n_1})$ and $\bar{U}_k := \bar{X}_k^{n_1}$ where $U_1 \perp\!\!\!\perp U_2$, and $\bar{U}_k = \bar{X}_k^{n_1} \sim \mathcal{CN}(0, \bar{P}_k)$ i.i.d., $k = 1, 2$. To encode $M_k \in \mathcal{M}_k$, we generate 2^{tnR_k} i.i.d. sequences $u_k^t(M_k)$ and $\bar{u}_k^t(M_k)$, distributed according to $p(u_k^t) = \prod_{i=1}^t p(u_{k,i}) = \prod_{i=1}^t p(x_{k,(i-1)n+1}^{in}, \hat{x}_{k,(i-1)n_1+1}^{in_1})$ and $p(\bar{u}_k^t) = \prod_{i=1}^t p(\bar{u}_{k,i})$, $k = 1, 2$. To communicate M_k, we transmit $u_k^t(M_k)$ and $\bar{u}_k^t(M_k)$ through the underlying RL-MARC and the S_k-D direct-links respectively. The relay assists each (n, n_1) block of symbols, by producing codewords according to the relay-distribution in (C.3), and forwarding them. The destination then decodes M_k from the received signals, $(Y_{\mathsf{D}}^{nt}, \bar{Y}_{\mathsf{RD}}^{n_1 t}, \bar{Y}_{k\mathsf{D}}^{n_1 t})$, using the CSI $(\boldsymbol{H}_{\mathsf{D}}^n, \bar{\boldsymbol{H}}_{\mathsf{D}}^{n_1})$. Applying standard random coding techniques as in [138, Chapter 8.7], the rates

$$R_k < \frac{1}{n} \mathrm{I}(X_k^n, \hat{X}_k^{n_1}; Y_{\mathsf{D}}^n, \bar{Y}_{\mathsf{RD}}^{n_1} | \boldsymbol{H}_{\mathsf{D}}^n, \bar{H}_{\mathsf{RD}}^{n_1}) + \frac{n_1}{n}\mathbb{E}[\tilde{\mathsf{C}}(\bar{G}_{k\mathsf{D}}\bar{P}_k)], \quad k \in \{1,2\} \tag{C.4}$$

are achievable. Finally, an achievable rate pair on the RL-MARC is given by the first term in (C.4), and its capacity of the R-MARC is the closure of the union of sets of all achievable rate pairs on the RL-MARC where the union is over all (n, n_1) and pmfs factoring as in (C.3) with $\bar{y}_{k\mathsf{D},\ell} = \bar{x}_{k,\ell} = \emptyset, k = 1, 2$.

C.2 Proof of Theorem 4.2

The achievable region is obtained by performing block Markov encoding over $B+1$ blocks with i.i.d., circularly symmetric complex Gaussian (CSCG) codewords and backward decoding at destination following a scheme similar to that for the single-band fading MARC as presented in [77]. As such, we present below only the notable details of the scheme.

Encoding: Encoding for block $b \in \{1, \ldots, B+1\}$ proceeds as follows: (i) the block lengths (n, n_1), and the input distributions $p_m(x), \bar{p}_m(\bar{x})$ and $\hat{p}_k(\hat{x}), m \in \{1, 2, \mathsf{R}\}, k \in \{1, 2\}$ are chosen; (ii) the message $M_{k,b} \in \mathcal{M}_k$ from S_k is encoded into codewords $x_k^n(M_{k,b})$ and $\hat{x}_k^{n_1}(M_{k,b})$, generated according to $\prod_{i=1}^{n} p_k(x_{k,i}(M_{k,b}))$ and $\prod_{\ell=1}^{n_1} \hat{p}_k(\hat{x}_{k,\ell}(M_{k,b}))$, $k \in \{1, 2\}$, and transmitted; (iii) assuming that the relay estimated $(M_{1,b-1}, M_{2,b-1})$ in block $b-1$ correctly, they are encoded into codewords $x_{\mathsf{R}}^n(M_{1,b-1}, M_{2,b-1})$ and $\bar{x}_{\mathsf{R}}^{n_1}(M_{1,b-1}, M_{2,b-1})$, generated according to $\prod_{i=1}^{n} p_{\mathsf{R}}(x_{\mathsf{R},i}(M_{1,b-1}, M_{2,b-1}))$ and $\prod_{\ell=1}^{n_1} \bar{p}_{\mathsf{R}}(\bar{x}_{\mathsf{R},\ell}(M_{1,b-1}, M_{2,b-1}))$, and transmitted. The messages $M_{k,0}$ and $M_{k,B+1}$ are known at the destination, $k \in \{1, 2\}$ as in [77, 124].

Decoding at the Relay: Assume that the message pair $(M_{1,b-1}, M_{2,b-1})$ was correctly decoded in block $b-1$. The relay then uses the side information $x_{\mathsf{R}}^n(M_{1,b-1}, M_{2,b-1})$ and $\bar{x}_{\mathsf{R}}^{n_1}(M_{1,b-1}, M_{2,b-1})$ and the CSI at block b, i.e., $\{H_{k\mathsf{R}}^n(b), \bar{H}_{k\mathsf{R}}^{n_1}(b)\}_{k=1}^{2}$, and estimates $(M_{1,b}, M_{2,b})$ from the signals received in block b as in [138, Chapter 14.3.1]. Such decoding yields certain rate constraints on R_1, R_2 and $R_1 + R_2$ which are then maximized by using i.i.d. CSCG codewords $X_m \sim \mathcal{CN}(0, P_m), \hat{X}_k \sim \mathcal{CN}(0, \hat{P}_k), m \in \{1, 2, 3\}, k \in \{1, 2\}$. Finally, the achievable rates are obtained by averaging the resulting rate constraints over i.i.d. squared-magnitudes of fading coefficients $G_{k\mathsf{R}}$ and $\bar{G}_{k\mathsf{R}}$ (since rate constraints are independent of the phases), as given in (4.12)-(4.14).

Decoding at the Destination: (Backward decoding) Assuming that $(M_{1,b+1}, M_{2,b+1})$ were decoded correctly in block $b+1$, the decoder estimates $(M_{1,b}, M_{2,b})$ from the signals received in blocks b and $b+1$ as in [138, Chapter 14.3.1] by using the side information $x_k^n(M_{k,b+1})$ and $\hat{x}_k^{n_1}(M_{k,b+1}), k \in \{1, 2\}$, and CSI in blocks b and $b+1$, given by $\{H_{m\mathsf{D}}^n(\ell), \bar{H}_{\mathsf{RD}}^{n_1}(\ell)\}_{\ell=b}^{b+1}, m \in \{1, 2, \mathsf{R}\}$. The resulting rate constraints are maximized by the same i.i.d. CSCG codewords as for the relay, and achievable rates are obtained by taking expectation over i.i.d. $G_{k\mathsf{D}}$ and $\bar{G}_{\mathsf{RD}}$, as given by (4.15)-(4.17).

C.3 Proof of Theorem 4.3

The outer-bounds are derived by applying the cut-set bounding technique such as those in [138, Chapter 14.10].

For notational convenience, define $\mathsf{U} \subseteq \{1,2\}$ and $\mathsf{U}^c := \{1,2\} \setminus \mathsf{U}$, such that $M_{\mathsf{U}} := \{M_k, k \in \mathsf{U}\}$, and $X_{\mathsf{U}} := \{X_k, k \in \mathsf{U}\}$. In the following, we characterize the upper bound on $\sum_{k\in\mathsf{U}} R_k$ for an arbitrary $\mathsf{U} \subseteq \{1,2\}$, from which the upper bounds on rates R_1, R_2, and $R_1 + R_2$ are obtained by taking $\mathsf{U} = \{1\}, \mathsf{U} = \{2\}$, and $\mathsf{U} = \{1,2\}$, respectively.

Assuming that source S_k transmits the message $M_k, k \in \{1,2\}$, we have

$$
\begin{aligned}
n\sum_{k\in\mathsf{U}} R_k &= H\left(M_{\mathsf{U}}\right) \\
&\overset{(a)}{=} H\left(M_{\mathsf{U}}|M_{\mathsf{U}^c}\right) \\
&\overset{(b)}{=} I\left(M_{\mathsf{U}}; Y_{\mathsf{D}}^n, \bar{Y}_{\mathsf{RD}}^{n_1}, \boldsymbol{H}_{\mathsf{D}}^n, \bar{H}_{\mathsf{RD}}^{n_1}|M_{\mathsf{U}^c}\right) + H\left(M_{\mathsf{U}}|Y_{\mathsf{D}}^n, \bar{Y}_{\mathsf{RD}}^{n_1}, \boldsymbol{H}_{\mathsf{D}}^n, \bar{H}_{\mathsf{RD}}^{n_1}, M_{\mathsf{U}^c}\right) \\
&\overset{(c)}{\leq} I\left(M_{\mathsf{U}}; Y_{\mathsf{D}}^n, \bar{Y}_{\mathsf{RD}}^{n_1}, \boldsymbol{H}_{\mathsf{D}}^n, \bar{H}_{\mathsf{RD}}^{n_1}|M_{\mathsf{U}^c}\right) + n\epsilon_n \\
&\overset{(d)}{=} I\left(M_{\mathsf{U}}; Y_{\mathsf{D}}^n|\boldsymbol{H}_{\mathsf{D}}^n, \bar{H}_{\mathsf{RD}}^{n_1}, M_{\mathsf{U}^c}\right) + I\left(M_{\mathsf{U}}; \bar{Y}_{\mathsf{RD}}^{n_1}|Y_{\mathsf{D}}^n, \boldsymbol{H}_{\mathsf{D}}^n, \bar{H}_{\mathsf{RD}}^{n_1}, M_{\mathsf{U}^c}\right) + n\epsilon_n \\
&\overset{(e)}{=} \sum_{i=1}^{n} h\left(Y_{\mathsf{D},i}|Y_{\mathsf{D}}^{i-1}, \boldsymbol{H}_{\mathsf{D}}^n, \bar{H}_{\mathsf{RD}}^{n_1}, M_{\mathsf{U}^c}, X_{\mathsf{U}^c,i}\right) \\
&\quad - h\left(Y_{\mathsf{D},i}|Y_{\mathsf{D}}^{i-1}, \boldsymbol{H}_{\mathsf{D}}^n, \bar{H}_{\mathsf{RD}}^{n_1}, M_{\mathsf{U}}, M_{\mathsf{U}^c}, X_{\mathsf{U},i}, X_{\mathsf{U}^c,i}\right) \\
&\quad + \sum_{l=1}^{n_1} h\left(\bar{Y}_{\mathsf{RD},l}|\bar{Y}_{\mathsf{RD}}^{l-1}, Y_{\mathsf{D}}^n, \boldsymbol{H}_{\mathsf{D}}^n, \bar{H}_{\mathsf{RD}}^{n_1}, M_{\mathsf{U}^c}\right) \\
&\quad - h\left(\bar{Y}_{\mathsf{RD},l}|\bar{Y}_{\mathsf{RD}}^{l-1}, Y_{\mathsf{D}}^n, \boldsymbol{H}_{\mathsf{D}}^n, \bar{H}_{\mathsf{RD}}^{n_1}, M_{\mathsf{U}^c}, M_{\mathsf{U}}\right) + n\epsilon_n \\
&\overset{(f)}{\leq} \sum_{i=1}^{n} h\left(Y_{\mathsf{D},i}|Y_{\mathsf{D}}^{i-1}, \boldsymbol{H}_{\mathsf{D}}^n, \bar{H}_{\mathsf{RD}}^{n_1}, M_{\mathsf{U}^c}, X_{\mathsf{U}^c,i}\right) \\
&\quad - h\left(Y_{\mathsf{D},i}|Y_{\mathsf{D}}^{i-1}, \boldsymbol{H}_{\mathsf{D}}^n, \bar{H}_{\mathsf{RD}}^{n_1}, M_{\mathsf{U}}, M_{\mathsf{U}^c}, X_{\mathsf{U},i}, X_{\mathsf{U}^c,i}, X_{\mathsf{R},i}\right) \\
&\quad + \sum_{l=1}^{n_1} h\left(\bar{Y}_{\mathsf{RD},l}|\bar{Y}_{\mathsf{RD}}^{l-1}, Y_{\mathsf{D}}^n, \boldsymbol{H}_{\mathsf{D}}^n, \bar{H}_{\mathsf{RD}}^{n_1}, M_{\mathsf{U}^c}\right) \\
&\quad - h\left(\bar{Y}_{\mathsf{RD},l}|\bar{Y}_{\mathsf{RD}}^{l-1}, Y_{\mathsf{D}}^n, \boldsymbol{H}_{\mathsf{D}}^n, \bar{H}_{\mathsf{RD}}^{n_1}, M_{\mathsf{U}^c}, M_{\mathsf{U}}, \bar{X}_{\mathsf{R},l}\right) + n\epsilon_n \\
&\overset{(g)}{\leq} \sum_{i=1}^{n} h\left(Y_{\mathsf{D},i}|X_{\mathsf{U}^c,i}, \boldsymbol{H}_{\mathsf{D},i}\right) - h\left(Y_{\mathsf{D},i}|X_{\mathsf{U},i}, X_{\mathsf{R},i}, X_{\mathsf{U}^c,i}, \boldsymbol{H}_{\mathsf{D},i}\right) \\
&\quad + \sum_{l=1}^{n_1} h\left(\bar{Y}_{\mathsf{RD},l}|\bar{H}_{\mathsf{RD},l}\right) - h\left(\bar{Z}_{\mathsf{RD},l}\right)
\end{aligned}
$$

$$
\begin{aligned}
&\overset{(h)}{=} \sum_{i=1}^{n} h\left(\sum_{k\in\mathsf{U}} G_{k\mathsf{D}}^{1/2} e^{j\Theta_{k\mathsf{D},i}} X_{k,i} + G_{\mathsf{RD}}^{1/2} e^{j\Theta_{\mathsf{RD},i}} X_{\mathsf{R},i} + Z_{\mathsf{D},i} | \{G_{m\mathsf{D},i}, \Theta_{m\mathsf{D},i}\}_{m\in\{1,2,\mathsf{R}\}}\right) \\
&\quad - h\left(Z_{\mathsf{D},i}\right) + \sum_{l=1}^{n_1} h\left(\bar{G}_{\mathsf{RD}}^{1/2} e^{j\bar{\Theta}_{\mathsf{RD},l}} \bar{X}_{\mathsf{R},l} + \bar{Z}_{\mathsf{RD},l} | \bar{G}_{\mathsf{RD},l}, \bar{\Theta}_{\mathsf{RD},l}\right) - h\left(\bar{Z}_{\mathsf{RD},l}\right) \\
&\overset{(i)}{\leq} \sum_{i=1}^{n} \mathbb{E}\left[\log\left(1 + G_{\mathsf{RD}} P_{\mathsf{R},i} + \sum_{k\in\mathsf{U}} \left(G_{k\mathsf{D}} P_{k,i} + 2 G_{k\mathsf{D}}^{1/2} G_{\mathsf{RD}}^{1/2}\, \mathsf{Re}\{e^{j(\Theta_{k\mathsf{D}} - \Theta_{\mathsf{RD}})} \mathbb{E}\left[X_{k,i} X_{\mathsf{R},i}^{*}\right]\}\right)\right)\right] \\
&\quad + \sum_{l=1}^{n_1} \mathbb{E}\left[\log\left(1 + \bar{G}_{\mathsf{RD}} \bar{P}_{\mathsf{R},l}\right)\right] \\
&\overset{(j)}{\leq} \sum_{i=1}^{n} \mathbb{E}\left[\log\left(1 + \sum_{k\in\mathsf{U}} G_{k\mathsf{D}} P_{k,i} + G_{\mathsf{RD}} P_{\mathsf{R},i}\right)\right] + \sum_{l=1}^{n_1} \mathbb{E}\left[\log\left(1 + \bar{G}_{\mathsf{RD}} \bar{P}_{\mathsf{R},l}\right)\right] \\
&\overset{(k)}{\leq} n\mathbb{E}\left[\log\left(1 + \sum\nolimits_{k\in\mathsf{U}} G_{k\mathsf{D}} P_{\mathsf{D}} + G_{\mathsf{RD}} P_{\mathsf{R}}\right)\right] + n_1 \mathbb{E}\left[\log\left(1 + \bar{G}_{\mathsf{RD}} \bar{P}_{\mathsf{R}}\right)\right] \qquad \text{(C.5)}
\end{aligned}
$$

where (a) follows since messages are independent, i.e., $M_{\mathsf{U}} \perp\!\!\!\perp M_{\mathsf{U}^c}$; (b) follows since the channel-state information for the channels to the destination, i.e., $\boldsymbol{H}_{\mathsf{D}}^{n}$ and $\bar{H}_{\mathsf{RD}}^{n_1}$ where $\boldsymbol{H}_{\mathsf{D},i} := \{H_{m\mathsf{D},i}\}_{m\in\{1,2,\mathsf{R}\}}, i = 1,\ldots,n$, is known at the destination, along with the received signals $Y_{\mathsf{D}}^{n}, \bar{Y}_{\mathsf{RD}}^{n_1}$; (c) follows from the Fano's inequality; (d) follows from the chain rule and noticing the fact that $I(M_{\mathsf{U}}; \boldsymbol{H}_{\mathsf{D}}^{n}, \bar{H}_{\mathsf{RD}}^{n_1} | M_{\mathsf{U}^c}) = 0$; (e) follows by expressing the mutual information $I(.;.)$ terms as the difference of two differential entropies $h(.)$, and then applying chain rule, while conditioning with $X_{\mathsf{U},i}(M_{\mathsf{U}})$ and $X_{\mathsf{U}^c,i}(M_{\mathsf{U}^c})$ (deterministic functions of M_{U} and M_{U^c}) do not alter entropy; (f) follows since conditioning the negative $h(.)$ terms with $X_{\mathsf{R},i}$ and $\bar{X}_{\mathsf{R},l}$ does not decrease entropy; in (g), the two positive $h(.)$ terms follow from unconditioning, while the two negative $h(.)$ terms follow from the two Markov chains $(Y_{\mathsf{D}}^{i-1}, \boldsymbol{H}_{\mathsf{D}}^{n\setminus i}, \bar{H}_{\mathsf{RD}}^{n_1}, M_{\mathsf{U}}, M_{\mathsf{U}^c}) \to (X_{\mathsf{U},i}, X_{\mathsf{U}^c,i}, X_{\mathsf{R},i}, \boldsymbol{H}_{\mathsf{D},i}) \to Y_{\mathsf{D},i}$ and $(\bar{Y}_{\mathsf{RD}}^{l-1}, Y_{\mathsf{D}}^{n}, \boldsymbol{H}_{\mathsf{D}}^{n}, \bar{H}_{\mathsf{RD}}^{n_1\setminus l}, M_{\mathsf{U}}, M_{\mathsf{U}^c}) \to (\bar{X}_{\mathsf{R},l}, \bar{H}_{\mathsf{RD},l}) \to \bar{Y}_{\mathsf{RD},l}$, respectively, imposed by the memoryless system model, where a vector $F^{m\setminus j} := \{F_i\}_{i=1}^{m} \setminus F_j$; (h) follows directly from the fading Gaussian model; (i) follows by maximizing the first positive $h(.)$ term of (h) by choosing $X_{k,i} \sim \mathcal{CN}(0, P_{k,i})$ [77], with $P_{k,i} := \mathbb{E}[|X_{k,i}|^2], k \in \{1,2,\mathsf{R}\}$, where $\mathbb{E}[X_{k,i} X_{\mathsf{R},i}^{*}]$ denotes the cross-correlation between $X_{k,i}$ and $X_{\mathsf{R},i}$ with the expectation taken over column i of the codebook, and $\mathsf{Re}(.)$ denotes the real part; the other positive $h(.)$ term in (h) is similarly maximized by $\hat{X}_{k,l} \sim \mathcal{CN}(0, \hat{P}_{k,l})$; the outer expectation for the first term is over i.i.d. fading magnitudes and phases, while the same for the second term is over the i.i.d. fading magnitudes only; (j) follows from [77, Theorem 8]: since in the first term of (i), $\tilde{\Theta} := \Theta_{k\mathsf{D}} - \Theta_{\mathsf{RD}} \sim \mathcal{U}[0, 2\pi)$, each summand of the form $\mathbb{E}_{\tilde{\Theta},G,B} \log(1 + G + 2G^{1/2}B^{1/2}\mathsf{Re}\{e^{j\tilde{\Theta}}\rho\})$ can be upper bounded as $\mathbb{E}_{\tilde{\Theta},G,B} \log(1 + G + 2G^{1/2}B^{1/2}\mathsf{Re}\{e^{j\tilde{\Theta}}\rho\}) \leq \mathbb{E}_{A} \log(1 + G)$ when

$\tilde{\Theta} \sim \mathcal{U}[0, 2\pi)$, $\rho \in \mathbb{C}$; and (k) follows from applying Jensen's inequality as in steps (c)-(d) of (C.1).

Finally, dividing both sides of (C.5) by n, as $n, n_1 \to \infty$, we have

$$R_{\mathsf{U}} \leq \mathbb{E}\left[\tilde{\mathsf{C}}\left(\sum\nolimits_{k\in\mathsf{U}} G_{k\mathsf{D}}P_{\mathsf{D}} + G_{\mathsf{RD}}P_{\mathsf{R}}\right)\right] + \alpha\mathbb{E}\left[\tilde{\mathsf{C}}\left(\bar{G}_{\mathsf{RD}}\bar{P}_{\mathsf{R}}\right)\right], \quad \mathsf{U} \subseteq \{1,2\}, \tag{C.6}$$

from which individual bounds on R_1, R_2, and $R_1 + R_2$ are obtained by choosing $\mathsf{U} = \{1\}$, $\mathsf{U} = \{2\}$, and $\mathsf{U} = \{1,2\}$, respectively.

C.4 Solution of Problem $[\mathcal{P}2]$ in Section 4.4.1

Convexity of Problem: We first show that $[\mathcal{P}2]$ is convex. We denote a feasible point by $\mathbf{x} := (p_1, q_1, p_2, q_2, R) \in \mathbb{R}^5_+$, and use the equivalent objective, *minimize* $-R$. Note that the objective is linear, and the equality constraints in (4.36) are affine. Moreover, the constraint in (4.34) is convex as its Hessian is a positive semidefinite matrix with

$$\frac{\alpha\kappa}{2}\left(\frac{d_1^2}{(1+d_1p_1)^2}, \frac{r_1^2}{(1+r_1q_1)^2}, \frac{d_2^2}{(1+d_2p_2)^2}, \frac{r_2^2}{(1+r_2q_2)^2}, 0\right)$$

on its leading diagonal. Similarly, constraint (4.35) can also be shown to be convex. In addition, (4.36) and (4.37), and $\mathbf{x} \succeq \mathbf{0}$ imply that the feasible set is compact for a given $P > 0$. Hence, $[\mathcal{P}2]$ is a convex optimization problem over a compact set.

The KKT Conditions: Problem $[\mathcal{P}2]$ satisfies Slater's condition [132, Chapter 5.2.3]: it can be illustrated by considering the point $\tilde{\mathbf{x}} := \left(P-\epsilon,\ \epsilon,\ P-\epsilon,\ \epsilon,\ \sigma_{\mathsf{R}}\right)$, which is strictly feasible for sufficiently small $\epsilon > 0$. Moreover, $[\mathcal{P}2]$ also satisfies the MFCQ as illustrated in Section 4.4.1. Therefore, $[\mathcal{P}2]$ can be solved using KKT conditions [132, Chapter 5.5.3].

The Lagrangian function for $[\mathcal{P}2]$ is given by

$$\mathfrak{L} = -R + \lambda_1(R - \Sigma_{\mathsf{R}}) + \lambda_2(R - \Sigma_{\mathsf{D}}) + \textstyle\sum_{k=1}^{2} \mu_k(p_k + q_k - P) - \rho_1 p_1 - \rho_2 q_1 - \rho_3 p_2 - \rho_4 q_2 - \rho_5 R,$$

where $\{\lambda_k\}_{k=1}^2, \{\mu_k\}_{k=1}^2$ and $\{\rho_i\}_{i=1}^5$ are Lagrange multipliers corresponding to constraints (4.34)-(4.35), (4.36), and $(p_1, q_1, p_2, q_2, R) \succeq \mathbf{0}$ respectively, with Σ_{R} and Σ_{D} in (4.30)–(4.31). With slight abuse of notation, we denote the optimal primal variables by (p_1, q_1, p_2, q_2, R), and the optimal Lagrange multipliers (OLM) by $(\lambda_1, \lambda_2, \rho_1, \rho_2, \rho_3, \rho_4)$ and (μ_1, μ_2), and they satisfy the following KKT conditions

$$\lambda_1 + \lambda_2 = 1, \tag{C.7}$$

$$\rho_1 = \mu_1 - \frac{\alpha}{2}\frac{d_1}{1+d_1p_1}, \quad \rho_2 = \mu_1 - \frac{\alpha}{2}\frac{\lambda_1 r_1}{1+r_1q_1}, \quad \rho_3 = \mu_2 - \frac{\alpha}{2}\frac{d_2}{1+d_2p_2}, \quad \rho_4 = \mu_2 - \frac{\alpha}{2}\frac{\lambda_1 r_2}{1+r_2q_2}, \tag{C.8}$$

$$p_1 + q_1 = P, \quad p_2 + q_2 = P, \tag{C.9}$$

$$R - \Sigma_{\mathsf{R}} \leq 0, \quad R - \Sigma_{\mathsf{D}} \leq 0, \quad \lambda_1(R - \Sigma_{\mathsf{R}}) = 0, \quad \lambda_2(R - \Sigma_{\mathsf{D}}) = 0, \tag{C.10}$$

$$\rho_1 p_1 = 0, \quad \rho_2 q_1 = 0, \quad \rho_3 p_2 = 0, \quad \rho_4 q_2 = 0, \tag{C.11}$$

$$(p_1, q_1, p_2, q_2, R) \succeq \mathbf{0}, \quad (\lambda_1, \lambda_2, \rho_1, \rho_2, \rho_3, \rho_4) \succeq \mathbf{0}, \quad (\mu_1, \mu_2) \neq \mathbf{0}. \tag{C.12}$$

with $\rho_5 = 0$ since $R \geq \min(\sigma_{\mathsf{D}}, \sigma_{\mathsf{R}}) > 0$.

Partitioning the Set of OLMs: We now partition the set of all $(\boldsymbol{\rho}, \boldsymbol{\lambda})$-tuples where $\boldsymbol{\rho} := (\rho_1, \rho_2, \rho_3, \rho_4) \succeq \mathbf{0}$ and $\boldsymbol{\lambda} := (\lambda_1, \lambda_2) \succeq \mathbf{0}$, into 18 subsets. First, the set of (ρ_1, ρ_2)-tuples is partitioned into 3 subsets, $\mathcal{I}_1 := \{(\rho_1, \rho_2) : \rho_1 > 0, \rho_2 = 0\}$, $\mathcal{I}_2 := \{(\rho_1, \rho_2) : \rho_1 = 0, \rho_2 > 0\}$, and $\mathcal{I}_3 := \{(\rho_1, \rho_2) : \rho_1 = 0, \rho_2 = 0\}$, since subset $\mathcal{I}_4 := \{(\rho_1, \rho_2) : \rho_1 > 0, \rho_2 > 0\}$ violates (C.9) by requiring $p_1 = q_1 = 0$. The set of (ρ_3, ρ_4)-tuples is similarly partitioned into 3 subsets $\mathcal{J}_k, k \in \{1, 2, 3\}$.

Finally, the set of $\boldsymbol{\lambda}$-tuples is partitioned into 2 subsets $\mathcal{L}_1 := \{\boldsymbol{\lambda} : \lambda_1 = 1, \lambda_2 = 0\}$ and $\mathcal{L}_2 := \{\boldsymbol{\lambda} : \lambda_1 > 0, \lambda_2 > 0\}$, since subset $\mathcal{L}_3 := \{\boldsymbol{\lambda} : \lambda_1 = 0, \lambda_2 = 1\}$ violates the assumption $\gamma > 1$ in the MARC-OPA by requiring $\Sigma_{\mathsf{D}} < \Sigma_{\mathsf{R}}$, and $\mathcal{L}_4 := \{\boldsymbol{\lambda} : \lambda_1 = 0, \lambda_2 = 0\}$ violates (4.34)–(4.35) by requiring $R < \min(\Sigma_{\mathsf{D}}, \Sigma_{\mathsf{R}})$. Thus, the set of $(\boldsymbol{\rho}, \boldsymbol{\lambda})$-tuples are now partitioned into 18 subsets $\mathcal{I}_k \cap \mathcal{J}_l \cap \mathcal{L}_m, k, l \in \{1, 2, 3\}, m \in \{1, 2\}$.

Note that a $(\boldsymbol{\rho}, \boldsymbol{\lambda})$-tuple now satisfies the KKT conditions as well as the condition of the subset to which it belongs. When all conditions on $(\boldsymbol{\rho}, \boldsymbol{\lambda})$ are expressed in terms of $(P, r_1, r_2, d_1, d_2, \gamma)$, each subset leads to an LGR as presented in Table C.1. However, only 14 LGRs are valid, since 3 are subsumed into an existing LGR ($\tilde{\mathcal{A}}_{(.,.)} \subseteq \mathcal{A}_{(.,.)}$), and $\tilde{\mathcal{A}}_{\mathsf{d},\mathsf{d}}$ is invalid as it violates the assumption $\gamma > 1$.

Power Allocation in LGRs: Next, we express the conditions on $(\boldsymbol{\rho}, \boldsymbol{\lambda})$ in each LGR in terms of P and threshold powers in Table 4.3. We also derive the expression of optimal powers in this process.

LGR $\mathcal{A}_{\mathsf{r},\mathsf{r}}$: Here, $\boldsymbol{\rho} \in \mathcal{I}_1 \cap \mathcal{J}_1$ and $\boldsymbol{\lambda} \in \mathcal{L}_1$. For $\boldsymbol{\rho} \in \mathcal{I}_1 \cap \mathcal{J}_1$, we have $\rho_1 > 0, \rho_2 = 0, \rho_3 > 0, \rho_4 = 0$, which require $p_1 = 0, q_1 = P, p_2 = 0, q_2 = P$ from (C.9), (C.11)-(C.12). Now, $\boldsymbol{\lambda} \in \mathcal{L}_1$ requires $\Sigma_{\mathsf{R}} < \Sigma_{\mathsf{D}}$ that results in $P < \mathsf{P}_{\mathsf{r},\mathsf{r}}$ from (C.10). The conditions for $\rho_1 > 0, \rho_3 > 0$ are derived by substituting $\rho_2 = \rho_4 = 0$ in (C.8) and eliminating (μ_1, μ_2).

Table C.1: Set of $(\boldsymbol{\rho}, \boldsymbol{\lambda})$-tuples are partitioned into 18 subsets and the LGR corresponding to each subset is provided.

	$\boldsymbol{\lambda} \in \mathcal{L}_1$	$\boldsymbol{\lambda} \in \mathcal{L}_2$
$\boldsymbol{\rho} \in \mathcal{I}_1 \cap \mathcal{J}_1$	$\mathcal{A}_{\mathsf{r,r}}$	$\tilde{\mathcal{A}}_{\mathsf{r,r}} \subseteq \mathcal{A}_{\mathsf{r,r}}$
$\boldsymbol{\rho} \in \mathcal{I}_2 \cap \mathcal{J}_1$	$\mathcal{A}_{\mathsf{d,r}}$	$\tilde{\mathcal{A}}_{\mathsf{d,r}} \subseteq \mathcal{A}_{\mathsf{d,r}}$
$\boldsymbol{\rho} \in \mathcal{I}_3 \cap \mathcal{J}_1$	$\mathcal{A}_{\mathsf{rd,r}}$	$\mathcal{S}_{\mathsf{rd,r}}$
$\boldsymbol{\rho} \in \mathcal{I}_1 \cap \mathcal{J}_2$	$\mathcal{A}_{\mathsf{r,d}}$	$\tilde{\mathcal{A}}_{\mathsf{r,d}} \subseteq \mathcal{A}_{\mathsf{r,d}}$
$\boldsymbol{\rho} \in \mathcal{I}_2 \cap \mathcal{J}_2$	$\mathcal{A}_{\mathsf{d,d}}$	$\tilde{\mathcal{A}}_{\mathsf{d,d}}$ is invalid
$\boldsymbol{\rho} \in \mathcal{I}_3 \cap \mathcal{J}_2$	$\mathcal{A}_{\mathsf{rd,d}}$	$\mathcal{S}_{\mathsf{rd,d}}$
$\boldsymbol{\rho} \in \mathcal{I}_1 \cap \mathcal{J}_3$	$\mathcal{A}_{\mathsf{r,rd}}$	$\mathcal{S}_{\mathsf{r,rd}}$
$\boldsymbol{\rho} \in \mathcal{I}_2 \cap \mathcal{J}_3$	$\mathcal{A}_{\mathsf{d,rd}}$	$\mathcal{S}_{\mathsf{d,rd}}$
$\boldsymbol{\rho} \in \mathcal{I}_3 \cap \mathcal{J}_3$	$\mathcal{A}_{\mathsf{rd,rd}}$	$\mathcal{S}_{\mathsf{rd,rd}}$

Hence, the conditions for $\mathcal{A}_{\mathsf{r,r}}$ are given by

$$P \leq \mathsf{P}'_{\mathsf{d,d}} = d_1^{-1} - r_1^{-1}, \quad P \leq \hat{\mathsf{P}}'_{\mathsf{d,d}} = d_2^{-1} - r_2^{-1}, \quad P < \mathsf{P}_{\mathsf{r,r}}. \tag{C.13}$$

The conditions of the counterpart $\tilde{\mathcal{A}}_{\mathsf{r,r}}$ (with $\boldsymbol{\lambda} \in \mathcal{L}_2$ instead of $\boldsymbol{\lambda} \in \mathcal{L}_1$) is valid only for a set of measure zero at $P = \mathsf{P}_{\mathsf{r,r}}$ but the optimum powers are the same as in $\mathcal{A}_{\mathsf{r,r}}$, thus it is subsumed in $\mathcal{A}_{\mathsf{r,r}}$.

LGR $\mathcal{A}_{\mathsf{d,rd}}$ and $\mathcal{S}_{\mathsf{d,rd}}$: In $\mathcal{A}_{\mathsf{d,rd}}$, $\boldsymbol{\rho} \in \mathcal{I}_2 \cap \mathcal{J}_3$ and $\boldsymbol{\lambda} \in \mathcal{L}_1$. For $\boldsymbol{\rho} \in \mathcal{I}_2 \cap \mathcal{J}_3$, we have $\rho_1 = 0, \rho_2 > 0, \rho_3 = 0, \rho_4 = 0$, which require $p_1 = P, q_1 = 0, p_2 \geq 0, q_2 \geq 0$ from (C.9), (C.11)-(C.12). First, by substituting $\rho_3 = \rho_4 = 0, \lambda_1 = 1$ in (C.8), we obtain $p_2 = 0.5(P + r_2^{-1} - d_2^{-1})$ and $q_2 = 0.5(P + d_2^{-1} - r_2^{-1})$, and conditions $(p_2, q_2) \succeq \mathbf{0}$ require $P \geq \hat{\mathsf{P}}'_{\mathsf{d,d}}, P \geq \hat{\mathsf{P}}_{\mathsf{d,d}}$. The condition for $\rho_2 > 0$, found by substituting $\rho_1 = 0, \lambda_1 = 1$ in (C.8), requires $P \leq \mathsf{P}_{\mathsf{d,d}} = r_1^{-1} - d_1^{-1}$. Finally, $\boldsymbol{\lambda} \in \mathcal{L}_1$ requires $\Sigma_{\mathsf{R}} < \Sigma_{\mathsf{D}}$, i.e., $P < \mathsf{P}_{\mathsf{d,rd}} = (2\gamma - 1)r_2^{-1} - d_2^{-1}$. Thus, the conditions for $\mathcal{A}_{\mathsf{d,rd}}$ are

$$\min(\mathsf{P}_{\mathsf{d,d}}, \mathsf{P}_{\mathsf{d,rd}}) \geq P \geq \max(\hat{\mathsf{P}}_{\mathsf{d,d}}, \hat{\mathsf{P}}'_{\mathsf{d,d}}).$$

In $\mathcal{S}_{\mathsf{d,rd}}$, $\boldsymbol{\rho} \in \mathcal{I}_2 \cap \mathcal{J}_3$, which still requires $p_1 = P, q_1 = 0, p_2 \geq 0, q_2 \geq 0$. However, now $\boldsymbol{\lambda} \in \mathcal{L}_2$, i.e., $(\lambda_1, \lambda_2) \succeq \mathbf{0}$, which requires $\Sigma_{\mathsf{R}} = \Sigma_{\mathsf{D}}$, resulting in $q_2 = (\gamma - 1)r_2^{-1}$, and $p_2 = P - (\gamma - 1)r_2^{-1}$. Due to $\gamma > 1$, we have $q_2 > 0$, but $p_2 > 0$ additionally requires $P > \mathsf{P}_{\mathsf{d,r}} = (\gamma - 1)r_2^{-1}$. Since $\lambda_1 + \lambda_2 = 1$ in (C.7), $(\lambda_1, \lambda_2) \succeq \mathbf{0}$ is equivalent to $1 > \lambda_1 > 0$.

Solving for λ_1 by substituting (p_2, q_2) above and $\rho_3 = \rho_4 = 0$ in (C.8), the condition $1 > \lambda_1 > 0$ requires $P > \mathsf{P}_{\mathsf{d,rd}}$. The condition for $\rho_2 > 0$, found by substituting $\rho_1 = 0$ in (C.8), requires $P > \bar{\mathsf{P}}_{\mathsf{d,rd}}$ if $\boldsymbol{r} \in \mathcal{R}_{S2}$, and $P < \bar{\mathsf{P}}_{\mathsf{d,rd}}$ otherwise. Therefore, the conditions of $\mathcal{S}_{\mathsf{d,rd}}$ are

$$P \geq \max(\mathsf{P}_{\mathsf{d,r}}, \mathsf{P}_{\mathsf{d,rd}}, \bar{\mathsf{P}}_{\mathsf{d,rd}}), \text{ if } \boldsymbol{r} \in \mathcal{R}_{S2}, \text{ and } \max(\mathsf{P}_{\mathsf{d,r}}, \mathsf{P}_{\mathsf{d,rd}}) \leq P < \bar{\mathsf{P}}_{\mathsf{d,rd}}, \text{otherwise.}$$

LGR $\mathcal{A}_{\mathsf{rd,r}}$ and $\mathcal{S}_{\mathsf{rd,r}}$: In $\mathcal{A}_{\mathsf{rd,r}}$, $\boldsymbol{\rho} \in \mathcal{I}_3 \cap \mathcal{J}_1$ and $\boldsymbol{\lambda} \in \mathcal{L}_1$. For $\boldsymbol{\rho} \in \mathcal{I}_3 \cap \mathcal{J}_1$, we have $\rho_1 = 0, \rho_2 = 0, \rho_3 > 0, \rho_4 = 0$, which require $p_1 \geq 0, q_1 \geq 0, p_2 = 0, q_2 = P$ from from (C.9), (C.11)-(C.12). First, by substituting $\rho_1 = \rho_2 = 0, \lambda_1 = 1$ in (C.8) we find $p_1 = 0.5(P + r_1^{-1} - d_1^{-1})$ and $q_1 = 0.5(P + d_1^{-1} - r_1^{-1})$, and $(p_1, q_1) \succeq \mathbf{0}$ require $P \geq \mathsf{P}'_{\mathsf{d,d}}$ and $P \geq \mathsf{P}_{\mathsf{d,d}}$. The condition for $\rho_3 > 0$, found by substituting $\rho_4 = 0, \lambda_1 = 1$ in (C.8), requires $P \leq \hat{\mathsf{P}}'_{\mathsf{d,d}}$. Also, $\boldsymbol{\lambda} \in \mathcal{L}_1$ (i.e., $\Sigma_{\mathsf{R}} < \Sigma_{\mathsf{D}}$) requires $P < \mathsf{P}_{\mathsf{rd,r}}$. Thus, the conditions for $\mathcal{A}_{\mathsf{rd,r}}$ are given by

$$P \geq \max(\mathsf{P}_{\mathsf{d,d}}, \mathsf{P}'_{\mathsf{d,d}}), \quad P \leq \min(\hat{\mathsf{P}}'_{\mathsf{d,d}}, \mathsf{P}_{\mathsf{rd,r}}).$$

In $\mathcal{S}_{\mathsf{rd,r}}$, $\boldsymbol{\rho} \in \mathcal{I}_3 \cap \mathcal{J}_1$ still requires $p_1 \geq 0, q_1 \geq 0, p_2 = 0, q_2 = P$, but $\boldsymbol{\lambda} \in \mathcal{L}_2$ now requires $\Sigma_{\mathsf{R}} = \Sigma_{\mathsf{D}}$, from which we solve for λ_1. Then, using λ_1 and $\rho_1 = \rho_2 = 0$ in (C.8), we find $p_1 = P - r_1^{-1}(\gamma/(1 + Pr_2) - 1)$ and $q_1 = r_1^{-1}(\gamma/(1 + Pr_2) - 1)$, and $(p_1, q_1) \succeq \mathbf{0}$ require $\mathsf{P}_{\mathsf{d,r}} > P > \mathsf{P}_{\mathsf{r,r}}$. Conditions (C.7) and $\boldsymbol{\lambda} \in \mathcal{L}_2$ simplify to $1 > \lambda_1 > 0$ which requires $P > \mathsf{P}_{\mathsf{rd,r}}$, while the condition for $\rho_3 > 0$ requires $P < \bar{\mathsf{P}}_{\mathsf{rd,r}}$. Thus, conditions for $\mathcal{S}_{\mathsf{rd,r}}$ are

$$\min(\mathsf{P}_{\mathsf{d,r}}, \bar{\mathsf{P}}_{\mathsf{rd,r}}) \geq P \geq \max(\mathsf{P}_{\mathsf{r,r}}, \mathsf{P}_{\mathsf{rd,r}}).$$

LGR $\mathcal{S}_{\mathsf{rd,rd}}$: Here, $\boldsymbol{\rho} \in \mathcal{I}_3 \cap \mathcal{J}_3$, i.e., $\boldsymbol{\rho} = \mathbf{0}$, and $\boldsymbol{\lambda} \in \mathcal{L}_2$: this require $\Sigma_{\mathsf{R}} = \Sigma_{\mathsf{D}}$, from which we solve for λ_1. Conditions (C.7) and $\boldsymbol{\lambda} \in \mathcal{L}_2$ simplify to $1 > \lambda_1 > 0$ which requires $P > \mathsf{P}_{\mathsf{rd,rd}}$. Using the expression of λ_1 and $\boldsymbol{\rho} = \mathbf{0}$ in (C.8), we find q_2 and q_1 as in the last and third to last rows of Table 4.1. From (C.9) we have $p_k = P - q_k$, and $p_k > 0, k = 1, 2$, requires $P > \max(\bar{\mathsf{P}}_{\mathsf{r,rd}}, \bar{\mathsf{P}}_{\mathsf{rd,r}})$. Finally, depending on the relay-link gains, condition $q_k > 0, k = 1, 2$, simplify to either of the following three conditions

$$\max(\bar{\mathsf{P}}_{\mathsf{rd,d}}, \bar{\mathsf{P}}_{\mathsf{d,rd}}) < P, \text{ for } \boldsymbol{r} \in \mathcal{R}_2 \cup \mathcal{R}_1,$$
$$\bar{\mathsf{P}}_{\mathsf{rd,d}} < P < \bar{\mathsf{P}}_{\mathsf{d,rd}}, \text{ for } \boldsymbol{r} \in \mathcal{R}_{S2},$$
$$\bar{\mathsf{P}}_{\mathsf{d,rd}} < P < \bar{\mathsf{P}}_{\mathsf{rd,d}}, \text{ for } \boldsymbol{r} \in \mathcal{R}_{S1},$$

as in Table 4.1.

The optimal powers and conditions for $\mathcal{A}_{\mathsf{rd,d}}, \mathcal{S}_{\mathsf{rd,d}}, \mathcal{A}_{\mathsf{r,rd}}$ and $\mathcal{S}_{\mathsf{r,rd}}$ are derived from $\mathcal{A}_{\mathsf{d,rd}}, \mathcal{S}_{\mathsf{d,rd}}, \mathcal{A}_{\mathsf{rd,r}}$ and $\mathcal{S}_{\mathsf{rd,r}}$ by exchanging the roles of the direct-links and relay-links, while those for $\mathcal{A}_{\mathsf{d,d}}$, $\mathcal{A}_{\mathsf{d,r}}$, $\mathcal{A}_{\mathsf{r,d}}$ and $\mathcal{A}_{\mathsf{rd,rd}}$ are derived through similar tedious algebraic manipu-

lations. The details are omitted here.

C.5 MF Constraint Qualification for Problem $[\mathcal{P}2]$

The Mangasarian-Fromovitz Constraint Qualification

Consider an optimization problem $[\mathcal{P}']$ over variable $\boldsymbol{x} \in \mathbb{R}^v$ as follows

$$[\mathcal{P}'] \quad \text{minimize} \quad f(\boldsymbol{x}) \tag{C.14}$$

$$\text{subject to:} \quad g_i(\boldsymbol{x}) \leq 0, \quad i = 1, \ldots, m, \tag{C.15}$$

$$h_i(\boldsymbol{x}) = 0, \quad i = 1, \ldots, n, \tag{C.16}$$

where $f, g_i, h_i : \mathbb{R}^v \to \mathbb{R}$ are assumed to be continuously differentiable functions. For a feasible point $\boldsymbol{x}^*$ of problem $[\mathcal{P}']$, the active set is defined as

$$\mathcal{I}(\boldsymbol{x}^*) := \{i \in \{1, \ldots, m\} : g_i(\boldsymbol{x}^*) = 0\}. \tag{C.17}$$

The Mangasarian-Fromovitz constraint qualification conditions [155, Section 2.1.2], [154, Section 4.3] holds at $\boldsymbol{x}^*$ if

- $\nabla h_i(\boldsymbol{x}^*)$ are linearly independent for $i = 1, \ldots, n$, and
- there exists a direction vector $\boldsymbol{t} \in \mathbb{R}^v$ such that
  1. $\nabla h_i(\boldsymbol{x}^*)^T \boldsymbol{t} = 0, \quad i = 1, \ldots, n$, and
  2. $\nabla g_i(\boldsymbol{x}^*)^T \boldsymbol{t} < 0, \quad i \in \mathcal{I}(\boldsymbol{x}^*)$.

Constraints and Gradients in Problem $[\mathcal{P}2]$

Note that in problem $[\mathcal{P}2]$ defined in (4.33)-(4.38), all equality and inequality constraints are *differentiable*. Given a set of channel parameters $(d_1, d_2, r_1, r_2, \gamma, P)$, consider a *feasible point* $\boldsymbol{x} = (p_1, q_1, p_2, q_2, R)$ for problem $[\mathcal{P}2]$.

For convenience, we denote the inequality constraints (4.34) and (4.35) by

$$g_1(\boldsymbol{x}) = R - \sigma_{\mathsf{R}} - \alpha \log(1 + r_1 q_1) - \alpha \log(1 + r_2 q_2) - \alpha \log(1 + d_1 p_1) - \alpha \log(1 + d_2 p_2),$$

$$g_2(\boldsymbol{x}) = R - \sigma_{\mathsf{D}} - \alpha \log(1 + d_1 p_1) - \alpha \log(1 + d_2 p_2),$$

where $\sigma_{\mathsf{R}}, \sigma_{\mathsf{D}}, \alpha, r_1, r_2, d_1, d_2$ are non-negative constants that do not depend on the feasible point $\boldsymbol{x}$. The gradients of these functions, $\nabla g_1(\boldsymbol{x})$ and $\nabla g_2(\boldsymbol{x})$, are given by

$$\nabla g_1(\boldsymbol{x}) = \left(-\frac{\alpha\kappa d_1}{1+d_1p_1},\ -\frac{\alpha\kappa r_1}{1+r_1q_1},\ -\frac{\alpha\kappa d_2}{1+d_2p_2},\ -\frac{\alpha\kappa r_2}{1+r_2q_2},\ 1 \right)^T,$$
$$\nabla g_2(\boldsymbol{x}) = \left(-\frac{\alpha\kappa d_1}{1+d_1p_1},\ 0,\ -\frac{\alpha\kappa d_2}{1+d_2p_2},\ 0,\ 1 \right)^T,$$

where $\mathbf{a}^T$ denotes the transpose of vector $\mathbf{a}$, and $\kappa = \log_2 e > 0$.

Next, the non-negativity constraints in (4.38) and their gradients are given as follows

$$\begin{aligned} g_3(\boldsymbol{x}) &= -p_1, & \nabla g_3(\boldsymbol{x}) &= (-1,0,0,0,0)^T, \\ g_4(\boldsymbol{x}) &= -q_1, & \nabla g_4(\boldsymbol{x}) &= (0,-1,0,0,0)^T, \\ g_5(\boldsymbol{x}) &= -p_2, & \nabla g_5(\boldsymbol{x}) &= (0,0,-1,0,0)^T, \\ g_6(\boldsymbol{x}) &= -q_2, & \nabla g_6(\boldsymbol{x}) &= (0,0,0,-1,0)^T, \\ g_7(\boldsymbol{x}) &= -R, & \nabla g_7(\boldsymbol{x}) &= (0,0,0,0,-1)^T. \end{aligned}$$

Finally, the equality constraints (4.36) and (4.37) are denoted as

$$h_1(\boldsymbol{x}) = p_1 + q_1 - P, \tag{C.18}$$
$$h_2(\boldsymbol{x}) = p_2 + q_2 - P, \tag{C.19}$$

and their gradients are given as follows

$$\nabla h_1(\boldsymbol{x}) = (1,1,0,0,0)^T \tag{C.20}$$
$$\nabla h_2(\boldsymbol{x}) = (0,0,1,1,0)^T. \tag{C.21}$$

[$\mathcal{P}2$] Satisfies the MFCQ

First, note that $\nabla h_1(\boldsymbol{x})$ and $\nabla h_2(\boldsymbol{x})$ in (C.20)-(C.21) are linearly independent. Next, for notational convenience, the active set $\mathcal{I}(\boldsymbol{x})$ at $\boldsymbol{x}$ is denoted as

$$\begin{aligned} &\mathcal{I}(\boldsymbol{x}) = \mathcal{I}_1(\boldsymbol{x}) \cup \mathcal{I}_2(\boldsymbol{x}) \cup \mathcal{I}_3(\boldsymbol{x}), \quad \text{where} \\ &\mathcal{I}_1(\boldsymbol{x}) \subseteq \{1,2\} \setminus \emptyset, \quad \mathcal{I}_2(\boldsymbol{x}) \subseteq \{3,4\} \setminus (3,4), \quad \mathcal{I}_3(\boldsymbol{x}) \subseteq \{5,6\} \setminus (5,6). \end{aligned} \tag{C.22}$$

More specifically, $\mathcal{I}_1(\boldsymbol{x})$ is a subset of $\mathcal{I}(\boldsymbol{x})$ corresponding to constraints $\{g_i(\boldsymbol{x})\}_{i=1}^2$, and since at least one of the two constraints must be tight, $\mathcal{I}_1(\boldsymbol{x}) \neq \emptyset$. $\mathcal{I}_2(\boldsymbol{x})$ is a subset of $\mathcal{I}(\boldsymbol{x})$ corresponding to non-negative constraints $\{g_i(\boldsymbol{x})\}_{i=3}^4$, while $\mathcal{I}_3(\boldsymbol{x})$ is a subset of $\mathcal{I}(\boldsymbol{x})$ corresponding to non-negative constraints $\{g_i(\boldsymbol{x})\}_{i=5}^6$.

The intuition into why $\mathcal{I}_2(\boldsymbol{x}) \neq (3,4)$, and $\mathcal{I}_3(\boldsymbol{x}) \neq (5,6)$ are given below:

- due to (4.36), if $p_1 = 0$, then $q_1 > 0$. Hence $g_3(\boldsymbol{x})$ is tight $\implies$ $g_4(\boldsymbol{x})$ is not tight, i.e., $3 \in \mathcal{I}_2(\boldsymbol{x}) \implies 4 \notin \mathcal{I}_2(\boldsymbol{x})$. Similarly, if $q_1 = 0$, then $p_1 > 0$, and hence $4 \in \mathcal{I}_2(\boldsymbol{x}) \implies 3 \notin \mathcal{I}_2(\boldsymbol{x})$. Moreover, if both $p_1 > 0, q_1 > 0$, then $\mathcal{I}_2(\boldsymbol{x}) = \emptyset$. Therefore, $\mathcal{I}_2(\boldsymbol{x}) \subseteq \{3,4\} \setminus (3,4)$.
- the rationale for $\mathcal{I}_3(\boldsymbol{x}) \subseteq \{5,6\} \setminus (5,6)$ follows from those for $\mathcal{I}_2(\boldsymbol{x})$ explained above, after replacing constraints 3 and 4 on variables p_1 and q_1 for $\mathcal{I}_2(\boldsymbol{x})$ with constraints 5 and 6 on variables p_2 and q_2 for $\mathcal{I}_3(\boldsymbol{x})$ that satisfy (4.37).

Finally, constraint $g_7(\boldsymbol{x})$ is never tight, i.e., $7 \notin \mathcal{I}(\boldsymbol{x})$ as R is always $R > 0$ due to the fact that $\sigma_{\mathsf{R}}, \sigma_{\mathsf{D}} > 0$.

It thus follows from (C.22) that 27 combinations of active sets are potentially possible. As an illustrative example, we first consider the following case and show that the MFCQ holds.

Case with Active Set $\mathcal{I}(\boldsymbol{x}) = \{1,2\} \cup \{3\} \cup \{5\}$: Clearly, in this case constrains $g_1(\boldsymbol{x})$ and $g_2(\boldsymbol{x})$, and non-negativity constraints $g_3(\boldsymbol{x})$ and $g_5(\boldsymbol{x})$ are tight, along with the equality constraints $h_1(\boldsymbol{x})$ and $h_2(\boldsymbol{x})$.

Consider a direction vector $\boldsymbol{t} = (t_1, t_2, t_3, t_4, t_5)^T$. First, to satisfy $\nabla h_1(\boldsymbol{x})^T \boldsymbol{t} = 0$, and $\nabla h_2(\boldsymbol{x})^T \boldsymbol{t} = 0$, vector $\boldsymbol{t}$ needs to satisfy

$$\begin{aligned} \nabla h_1(\boldsymbol{x})^T \boldsymbol{t} = 0 &\implies t_2 = -t_1, \\ \nabla h_2(\boldsymbol{x})^T \boldsymbol{t} = 0 &\implies t_4 = -t_3. \end{aligned} \tag{C.23}$$

For the MFCQ to be satisfied, in addition to (C.23), $\nabla g_k(\boldsymbol{x})^T \boldsymbol{t} < 0, k = 1,2,3,5$, must hold as well. For $\nabla g_3(\boldsymbol{x})^T \boldsymbol{t} < 0$ and $\nabla g_5(\boldsymbol{x})^T \boldsymbol{t} < 0$ to hold, $\boldsymbol{t}$ must satisfy

$$\nabla g_3(\boldsymbol{x})^T \boldsymbol{t} < 0 \iff t_1 > 0, \quad \text{and} \quad \nabla g_5(\boldsymbol{x})^T \boldsymbol{t} < 0 \iff t_3 > 0.$$

Then, after substituting for t_2 and t_4 as in (C.23), for $\nabla g_1(\boldsymbol{x})^T \boldsymbol{t} < 0$ to hold, $\boldsymbol{t}$ must satisfy

$$\nabla g_1(\boldsymbol{x})^T \boldsymbol{t} < 0 \iff t_5 + t_1 \underbrace{\left(\frac{\alpha\kappa r_1}{1 + r_1 q_1} - \frac{\alpha\kappa d_1}{1 + d_1 p_1} \right)}_{=:s_1} + t_3 \underbrace{\left(\frac{\alpha\kappa r_2}{1 + r_2 q_2} - \frac{\alpha\kappa d_2}{1 + d_2 p_2} \right)}_{=:s_2} < 0. \tag{C.24}$$

While $t_1 > 0$ and $t_3 > 0$, t_5 is free. Hence, so long as (s_1, s_2) are finite, irrespective of whether (s_1, s_2) satisfy either (a) $s_1 > 0, s_2 > 0$, (b) $s_1 > 0, s_2 < 0$, (c) $s_1 < 0, s_2 > 0$, or (d) $s_1 < 0, s_2 < 0$, inequality (C.24) can be satisfied by taking t_5 to be a sufficiently large negative number.

Similarly, substituting for t_2 and t_4 in (C.23), for $\nabla g_2(\boldsymbol{x})^T \boldsymbol{t} < 0$ to hold, $\boldsymbol{t}$ must satisfy

$$\nabla g_2(\boldsymbol{x})^T \boldsymbol{t} < 0 \iff t_5 - t_1 \underbrace{\frac{\alpha\kappa d_1}{1 + d_1 p_1}}_{>0} - t_3 \underbrace{\frac{\alpha\kappa d_2}{1 + d_2 p_2}}_{>0} < 0. \tag{C.25}$$

As before, $t_1 > 0, t_3 > 0$, and t_5 is free. Hence, taking any $t_5 < 0$ is sufficient for inequality (C.25) to hold.

Therefore, we have found $\boldsymbol{t} = (t_1, t_2, t_3, t_4, t_5)^T$, where $t_1, t_3 > 0$, $t_2 = -t_1, t_4 = -t_3$, and t_5 is the minimum of the two sufficiently large negative numbers found for inequalities (C.24) and (C.25). Therefore, the problem satisfies the MFCQ for the active set $\mathcal{I}(\boldsymbol{x}) = \{1, 2\} \cup \{3\} \cup \{5\}$.

Other Active Sets: While the above analysis is for the specific active set $\mathcal{I}(\boldsymbol{x}) = \{1, 2\} \cup \{3\} \cup \{5\}$, similar analysis holds for all other active sets. Since the analysis for other active sets is tedious and repetitive, we omit their explicit derivation.

[$\mathcal{P}1$] Does Not Satisfy the LICQ in General

For problem [$\mathcal{P}'$] in (C.14)-(C.16), if $\boldsymbol{x}^*$ is a feasible point, the linear independence constraint qualification (LICQ) [155, Section 2.1.2], [156] is said to hold if

- $\{\nabla h_i(\boldsymbol{x}^*), \ \ i = 1, \ldots, n\}$, and $\{\nabla g_i(\boldsymbol{x}^*), \ \ i \in \mathcal{I}(\boldsymbol{x}^*)\}$ are linearly independent.

We observe that [$\mathcal{P}2$] *does not satisfy the LICQ in general.*

To verify, consider a feasible point $\boldsymbol{x}$ for problem [$\mathcal{P}2$], for which the active set has a *cardinality of* 4, i.e., the active set is either $\mathcal{I}(\boldsymbol{x}) = \{1, 2\} \cup \{3\} \cup \{5\}$, $\mathcal{I}(\boldsymbol{x}) = \{1, 2\} \cup \{4\} \cup \{5\}$, $\mathcal{I}(\boldsymbol{x}) = \{1, 2\} \cup \{3\} \cup \{6\}$, or $\mathcal{I}(\boldsymbol{x}) = \{1, 2\} \cup \{4\} \cup \{6\}$. For these active sets, when the two *equality constraints* in (C.18)-(C.19) are also included, the *total number* of gradient vectors to be considered becomes 6. Since feasible point $\boldsymbol{x}$ has dimension 5, any 6 vectors *cannot* be linearly independent. Hence for these active sets [$\mathcal{P}2$] *does not satisfy the LICQ*, and thus the optimal Lagrange multipliers are *non-unique* [156].

For all other active sets, the *cardinality is at most* 3. Hence, including the two equality constraints in (C.18)-(C.19), the number of gradient vectors to be considered for linear independence is at most 5. For each such active set, we convert the problem of *verifying linear independence* into that of *determining the rank* of a matrix associated with a system

of linear equations. Then applying standard row reduction technique, we observe that the vectors for all such active sets are linearly independent. As such, for these active sets, problem [$\mathcal{P}2$] satisfies the LICQ, and thus the optimal Lagrange multipliers are *unique* [156]. The details are omitted.

Appendix D

Appendices for Chapter 5

D.1 Proof of Theorem 5.1

Fix a distribution $p \in \mathcal{P}$, and let the time-sharing variable Q be available to all sources and destinations. Suppose that source S_k wants to communicate its message M_k at rate R_k where the common message is $M_{k\mathsf{c}}$ with rate $R_{k0}, k \in \{1,2,3,4\}$. The encoding proceeds as follows. First, a time-sharing sequence q^n is generated according to $q^n \sim p_{Q^n}(q^n) = \prod_{i=1}^n p_Q(q_i)$. Source S_k then generates $2^{nR_{k0}}$ codewords u_k^n according to $u_k^n \sim p_{U_k^n|Q^n}(u_k^n|q^n) = \prod_{i=1}^n p_{U_k|Q}(u_{k,i}|q_i)$, and indexes them by the common message $M_{k\mathsf{c}} \in \{1,\ldots,2^{nR_{k0}}\}$. To employ superposition encoding, we take each $u_k^n(M_{k\mathsf{c}})$ as a cloud center, and generate $2^{n(R_k-R_{k0})}$ codewords x_k^n according to $x_k^n \sim p_{X_k^n|U_k^n,Q^n}(x_k^n|u_k^n(M_{k\mathsf{c}}),q^n) = \prod_{i=1}^n p_{X_k|U_k,Q}(x_{k,i}|u_{k,i}(M_{k\mathsf{c}}),q_i)$, and index them by common-private message pair $(M_{k\mathsf{c}}, M_{k\mathsf{p}}) \in \{1,\ldots,2^{nR_{k0}}\} \times \{1,\ldots,2^{n(R_k-R_{k0})}\}$. After the codebook for all sources are generated, they are revealed to the destinations and kept fixed. To communicate message M_k, source S_k sends $x_k^n(M_{k\mathsf{c}}, M_{k\mathsf{p}})$.

The codewords from all sources are superimposed at both the destinations. Destination D_1 employs joint typical decoding [7] to find the pair $(\hat{M}_1, \hat{M}_2) = ((M_{1\mathsf{c}}, M_{1\mathsf{p}}), (M_{2\mathsf{c}}, M_{2\mathsf{p}}))$ such that $(q^n, (x_k^n(M_{k\mathsf{c}}, M_{k\mathsf{p}}), u_k^n(M_{k\mathsf{c}}))_{k=1}^2, (u_\ell^n(M_{\ell\mathsf{c}}))_{\ell=3}^4) \in \mathcal{A}_\epsilon^{(n)}$, for some $(M_{3\mathsf{c}}, M_{4\mathsf{c}})$. The error events and the corresponding rate constraints are then derived by adapting the techniques of [7] for 2-user IC to the 2-user MAIC, and the resulting rate region is given by $\mathcal{R}_1(p)$ in Theorem 5.1. The decoding for D_2 proceeds similarly with the roles of S_1 and S_2 swapped with those of S_3 and S_4. The decoding procedure and notations are quite standard and similar to the two-user IC case detailed in [97, Chapter 6], and hence the

details are omitted for brevity.

D.2 Proof of Theorem 5.2

We bound the sum-rate as follows

$$\begin{aligned}
2nC_{\mathsf{S}} &= 2n(R_1+R_2)+2n(R_3+R_4)\\
&= n(R_1+R_2)+n(R_1+R_2)+n(R_3+R_4)+n(R_3+R_4)\\
&\overset{(a)}{\le} I(X_1^n,X_2^n;Y_1^n)+I(X_1^n,X_2^n;Y_1^n)+I(X_3^n,X_4^n;Y_2^n)+I(X_3^n,X_4^n;Y_2^n)+n\epsilon_n\\
&\overset{(b)}{\le} I(X_1^n,X_2^n;Y_1^n)+I(X_1^n,X_2^n;Y_1^n,S_1^n)+I(X_3^n,X_4^n;Y_2^n)+I(X_3^n,X_4^n;Y_2^n,U_2^n)+n\epsilon_n
\end{aligned} \tag{D.1}$$

where (a) follows from Fano's inequality, and (b) follows by providing S_1^n to one of the terms for D_1 and U_2^n to one of the terms for D_2.

We now expand the terms in (D.1) as follows:

$$I(X_1^n,X_2^n;Y_1^n) = h(Y_1^n) - \underbrace{h(h_3(X_3^n+X_4^n)+Z_1^n)}_{=:\alpha_1},$$

$$\begin{aligned}
I(X_1^n,X_2^n;Y_1^n,S_1^n) &= I(X_1^n,X_2^n;Y_1^n|S_1^n)+I(X_1^n,X_2^n;S_1^n)\\
&= h(Y_1^n|S_1^n)+h(S_1^n)-h(Y_1^n|S_1^n,X_1^n,X_2^n)-h(S_1^n|X_1^n,X_2^n)\\
&= h(X_1^n+X_2^n+h_3(X_3^n+X_4^n)+Z_1^n|h_1(X_1^n+X_2^n)+N_1^n)\\
&\quad + h(h_1(X_1^n+X_2^n)+N_1^n)\\
&\quad - h(X_1^n+X_2^n+h_3(X_3^n+X_4^n)+Z_1^n|X_1^n,X_2^n,N_1^n)-h(N_1^n)\\
&= \underbrace{h(h_3(X_3^n+X_4^n)+V_1^n|S_1^n)}_{=:\alpha_2}+\underbrace{h(h_1(X_1^n+X_2^n)+N_1^n)}_{=:\alpha_3}\\
&\quad - \underbrace{h(h_3(X_3^n+X_4^n)+Z_1^n|N_1^n)}_{=:\alpha_4} - nh(N_1),
\end{aligned}$$

$$I(X_3^n,X_4^n;Y_2^n) = h(Y_2^n) - \underbrace{h(h_1(X_1^n+X_2^n)+Z_2^n)}_{=:\alpha_5},$$

$$\begin{aligned}
I(X_3^n,X_4^n;Y_2^n,U_2^n) &= I(X_3^n,X_4^n;U_2^n)+I(X_3^n,X_4^n;Y_2^n|U_2^n)\\
&\overset{(a)}{=} h(X_3^n+X_4^n+h_1(X_1^n+X_2^n)+Z_2^n|h_1(X_1^n+X_2^n)+W_2^n)\\
&\quad - h(X_3^n+X_4^n+h_1(X_1^n+X_2^n)+Z_2^n|X_3^n,X_4^n,h_1(X_1^n+X_2^n)+W_2^n)\\
&= h(X_3^n+X_4^n+Z_2^n-W_2^n|U_2^n)-h(Z_2^n-W_2^n|U_2^n)
\end{aligned}$$

$$\overset{(b)}{=} \underbrace{h(X_3^n + X_4^n + V_2^n|U_2^n)}_{=:\alpha_6} + \underbrace{h(h_1(X_1^n + X_2^n) + W_2^n)}_{=:\alpha_7}$$
$$- \underbrace{h(h_1(X_1^n + X_2^n) + W_2^n|Z_2^n - W_2^n)}_{=:\alpha_8} - nh(Z_2 - W_2), \tag{D.2}$$

where (a) follows since $(X_3^n, X_4^n) \perp\!\!\!\perp U_2^n$, and thus $I(X_3^n, X_4^n; U_2^n) = 0$; and in (b) $V_1^n \sim \mathcal{N}(0, \sigma^2_{Z_1 - N_1/h_1})$, i.i.d., and $V_2^n \sim \mathcal{N}(0, \sigma^2_{Z_2 - W_2})$, i.i.d. are used for notational convenience.

We pair the negative and positive multi-letter $h(.)$ terms as follows $\alpha_2 - \alpha_1, \alpha_6 - \alpha_4, \alpha_3 - \alpha_8$ and $\alpha_7 - \alpha_5$, and single-letterize them by applying the WAN technique [144]. For example, $\alpha_2 - \alpha_1$ is bounded as

$$\begin{aligned}
\alpha_2 - \alpha_1 &= h(h_3(X_3^n + X_4^n) + V_1^n|S_1^n) - h(h_3(X_3^n + X_4^n) + Z_1^n) \\
&\overset{(a)}{\leq} h(h_3(X_3^n + X_4^n) + V_1^n|S_1^n) - h(h_3(X_3^n + X_4^n) + Z_1^n|S_1^n) \\
&\overset{(b)}{\leq} nh(h_3(X_3^* + X_4^*) + V_1|S_1^*) - nh(h_3(X_3^* + X_4^*) + V_1 + \tilde{V}_1|S_1^*)
\end{aligned}$$

where (a) follows by conditioning the negative $h(.)$ term by S_1^n, which does not reduce entropy, and (b) follows from the worst additive noise (WAN) result in [144, Lemma 2]: if $\sigma^2_{Z_1 - N_1/h_1} \leq \sigma^2_{Z_1} = 1$, (a) is maximized by choosing $X_k^* \sim \mathcal{N}(0, P_k)$, i.i.d., resulting in (b), where $S_1^* := h_1(X_1^* + X_2^*) + N_1$, and $\tilde{V}_1 \sim \mathcal{N}(0, 1 - \sigma^2_{Z_1 - N_1/h_1})$ and independent of all other variables. The resulting term is given by $1/2 \log(\mu_1)$ in (5.34) and the condition as [A1] in (5.31).

Similarly, $\alpha_6 - \alpha_4$ is bounded as

$$\begin{aligned}
\alpha_6 - \alpha_4 =& h(X_3^n + X_4^n + V_2^n|U_2^n) - h(h_3(X_3^n + X_4^n) + Z_1^n|N_1^n) \\
\leq^{(a)}& h(X_3^n + X_4^n + V_2^n|U_2^n) - h(h_3(X_3^n + X_4^n) + Z_1^n|N_1^n, U_2^n) \\
\leq^{(b)}& nh(X_3^* + X_4^* + V_2|U_2^*) - nh(X_3^* + X_4^* + V_2 + \tilde{V}_2|U_2^*) - n \log |h_3|
\end{aligned}$$

where (a) follows from conditioning, and (b) follows from the WAN result in [144, Lemma 2]: under $\sigma^2_{Z_2 - W_2} \leq \sigma^2_{Z_1|N_1}/h_3^2$, (a) is maximized by taking $X_k^* \sim \mathcal{N}(0, P_k)$ i.i.d., which results in (b), where $U_2^* := h_1(X_1^* + X_2^*) + W_2$, and $\tilde{V}_2 \sim \mathcal{N}(0, \sigma^2_{Z_1|N_1}/h_3^2 - \sigma^2_{Z_2 - W_2})$ and independent of all other variables. The resulting term is given by $1/2 \log(\mu_3) - 1/2 \log(h_3^2)$ in (5.34), and the condition is given as [A2] in (5.31).

Similarly, $\alpha_3 - \alpha_8$ and $\alpha_7 - \alpha_5$ are single letterized if $\sigma^2_{N_1} \leq \sigma^2_{W_2|Z_2 - W_2}$ and $\sigma^2_{W_2} \leq \sigma^2_{Z_2} = 1$ hold respectively, given in (5.33). The resulting terms are given as $1/2 \log(\mu_2)$ and $1/2 \log(\mu_4)$ in (5.34). Moreover, the expression in (5.34) remain unchanged, if terms $\alpha_3, \alpha_8, \alpha_7, \alpha_5$ are alternatively combined as $\alpha_3 - \alpha_5$ and $\alpha_7 - \alpha_8$, which are single letterized

under conditions $\sigma_{N_1}^2 \leq 1$ and $\sigma_{W_2}^2 \leq \sigma_{W_2|Z_2-W_2}^2$, given in (5.32).

D.3 Solution of Problem (5.40) in Theorem 5.3

Replacing $\sigma_{N_1}^2$ by x for notational convenience, Problem (5.40) is expressed as

$$\text{minimize} \quad -\log x(4h^2 - x) \tag{D.3}$$

$$\text{subject to} \quad 4h^2(1-h^2) \leq x \leq \min(1, 4h^2). \tag{D.4}$$

Simple inspection reveals the objective function to be convex, while the constraints are simple, hence this a convex optimization problem which can be solved by using the approach of KKT conditions [132].

For simpler treatment, we consider two mutually exclusive cases: $h^2 \geq 1/4$, and $h^2 < 1/4$. When $h^2 \geq 1/4$, the constraint in (5.40) simplifies to $x \leq 1$, while for $h^2 < 1/4$, the same constraint simplifies to $x \leq 4h^2$.

First, consider $h^2 \geq 1/4$, in which case the Lagrangian function for Problem (D.3) is

$$\mathfrak{L}_1 = -\log x(4h^2 - x) + \lambda_1 \left(4h^2(1-h^2) - x\right) + \lambda_2(x-1) \tag{D.5}$$

where $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$ are Lagrange multipliers associated with the lower and upper bounds in (D.3). The KKT condition are as follows

$$-\frac{1}{x} + \frac{1}{4h^2 - x} - \lambda_1 + \lambda_2 = 0 \tag{D.6}$$

$$\lambda_1 \left(4h^2(1-h^2) - x\right) = 0, \quad \lambda_1 \geq 0, \quad 4h^2(1-h^2) \leq x \tag{D.7}$$

$$\lambda_2(x-1) = 0, \quad \lambda_2 \geq 0, \quad x \leq 1. \tag{D.8}$$

We consider the 4 cases of (λ_1, λ_2), and solve for the optimal point x^* along with the interval of h^2 under which the optimal solution is valid: (i) for $\lambda_1 > 0, \lambda_2 > 0$, we have $x^* = 1$, if $h^2 = 1/2$; (ii) for $\lambda_1 = 0, \lambda_2 > 0$, we have $x^* = 1$, if $h^2 > 1/2$; (iii) for $\lambda_1 > 0, \lambda_2 = 0$, we have $x^* = 4h^2(1-h^2)$, if $1/4 \leq h^2 < 1/2$; (iv) $\lambda_1 = 0, \lambda_2 = 0$ is invalid.

Considering $h^2 < 1/4$, through similar KKT condition analysis we have the following: (i) for $\lambda_1 > 0, \lambda_2 > 0$, we have $x^* = 1$, if $h^2 = 1/2$; (ii) for $\lambda_1 = 0, \lambda_2 > 0$ is invalid; (iii) for $\lambda_1 > 0, \lambda_2 = 0$, we have $x^* = 4h^2(1-h^2)$, if $0 \leq h^2 < 1/4$; (iv) $\lambda_1 = 0, \lambda_2 = 0$ is invalid.

Accounting for the two cases, in summary, we have

$$x^* = 4h^2(1-h^2), \quad \text{if } h^2 < 1/2,$$

$$x^* = 1, \quad \text{if } h^2 \geq 1/2.$$

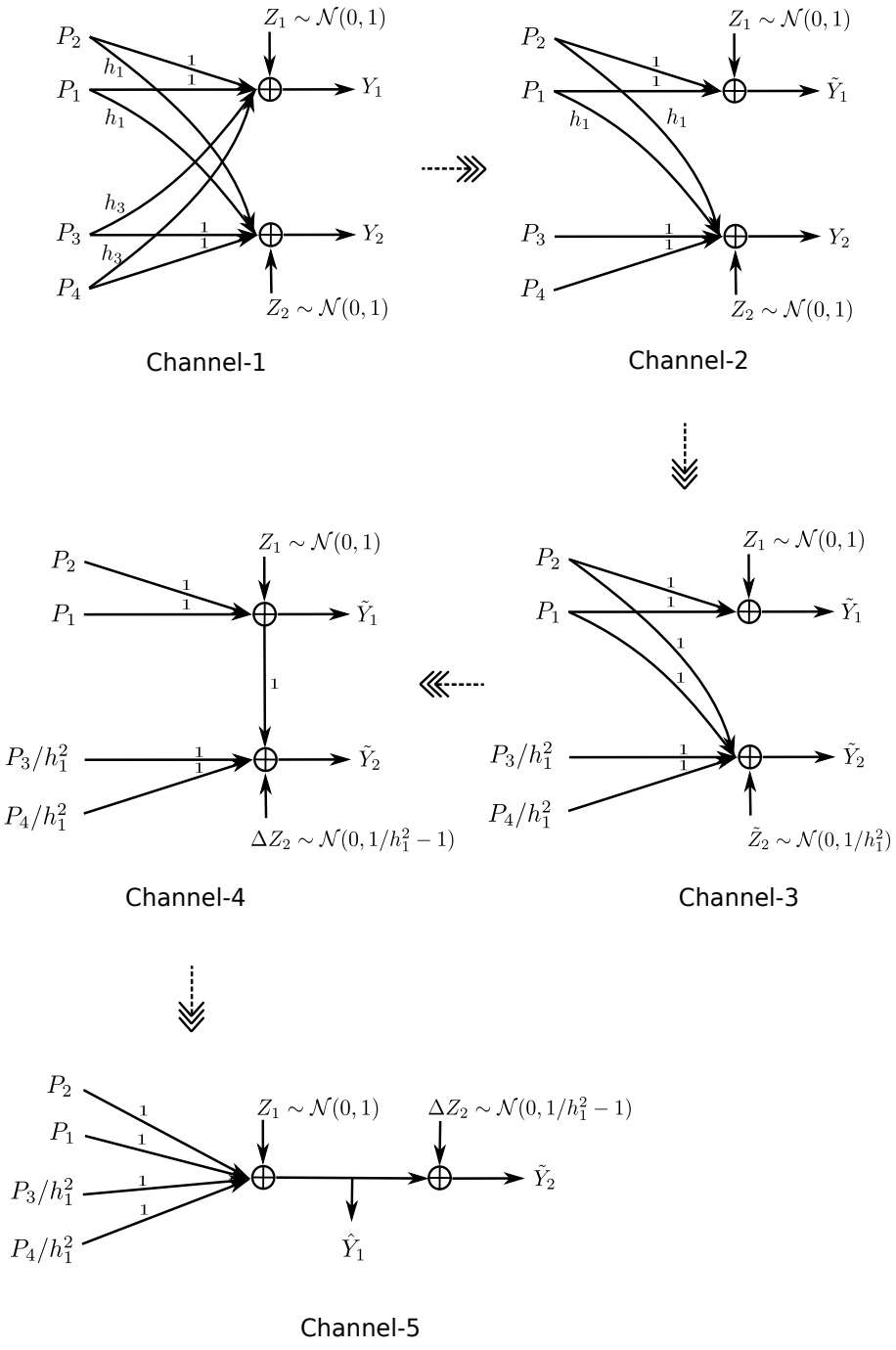


Figure D.1: Intermediate channels for the proof of Theorem 5.4.

D.4 Proof of Theorem 5.4

We prove this result by transforming the GMAIC into a degraded broadcast channel via a few intermediate channels in a step-by-step manner. For intuition, in Figure D.1 we depict this transformation process and the resulting intermediate channels, denoted as

Channel-1 through Channel-5. Finally, we show that the sum-capacity of the GMAIC is upper bounded by a sum-rate upper bound on the broadcast channel, which then provides a sum-rate upper bound for the former.

Consider the MAIC where both destinations are impaired by interference, depicted by Channel-1 in Figure D.1. First, the interfering links from sources S_3 and S_4 to destination D_1 is removed, resulting in a MAIC with partial interference, depicted as Channel-2 in Figure D.1. Equivalently, Channel-2 is found from Channel-1 by providing the interfering signals (X_3^n, X_4^n) to D_1 so that the effect of interference is completely canceled out. Since the capacity of this genie-aided channel (i.e., Channel-2) is always as large as that of the original channel (i.e., Channel-1), the sum-capacity of Channel-1 is upper bounded by that of Channel-2. Note that for Channel-2 the output at D_1 is now given by

$$\tilde{Y}_1^n = (X_1^n + X_2^n) + Z_1^n. \tag{D.9}$$

Next, Channel-3 is constructed from Channel-2 by absorbing the effect of the cross-channel gains h_1 from sources S_1 and S_2 to destination D_2 into the transmit powers of sources S_3 and S_4 and noise variance at D_2, as depicted in Figure D.1. Note that this transformation on Channel-2 does not change its capacity region. To see that, first note that in both channels the signals received at D_1 are the same as given in (D.9). Next, scaling the output at D_2 of Channel-2 by $1/h_1$, we obtain

$$Y_2^n/h_1 = (X_1^n + X_2^n) + (X_3^n + X_4^n)/h_1 + Z_2^n/h_1, \tag{D.10}$$

which is statistically the same as the output at D_2 of Channel-3, given by

$$\tilde{Y}_2^n = (X_1^n + X_2^n) + (\tilde{X}_3^n + \tilde{X}_4^n) + \tilde{Z}_2^n, \tag{D.11}$$

where the transmit power constraints are $\sum_{i=1}^n \mathbb{E}[\tilde{X}_{3,i}^2] \leq P_3/h_1^2$ and $\sum_{i=1}^n \mathbb{E}[\tilde{X}_{4,i}^2] \leq P_4/h_1^2$, while the noise is $\tilde{Z}_2^n \sim \mathcal{N}(0, 1/h_1^2)$, i.i.d.

Since scaling by a constant does not change mutual information, the capacity region of Channel-2 and Channel-3, and in particular, the sum-capacities, are the same.

Note that Channel-3 can be equivalently represented as Channel-4, depicted in Figure D.1. It can be observed that in Channel-3, output $\tilde{Y}_2^n$ can be regarded as a degraded version of output $\tilde{Y}_1^n$. To see this, note that $h_1^2 < 1$, i.e., $1/h_1^2 - 1 > 0$, and thus $\tilde{Z}_2^n$ in (D.11) can be decomposed as $\tilde{Z}_2^n = Z_1^n + \Delta Z_2^n$, where $Z_1^n \sim \mathcal{N}(0,1) \perp\!\!\!\perp \Delta Z_2^n \sim \mathcal{N}(0, 1/h_1^2 - 1)$. As such, substituting $\tilde{Z}_2^n$ in (D.11), one can observe that $\tilde{Y}_2^n$ is a degraded version of

$\tilde{Y}_1^n = (X_1^n + X_2^n) + Z_1^n$, i.e.,

$$\tilde{Y}_2^n = \tilde{Y}_1^n + (\tilde{X}_3^n + \tilde{X}_4^n) + \Delta Z_2^n. \tag{D.12}$$

Since the two received signals are statistically the same in Channel-3 and Channel-4, they posses the same sum-capacity.

Finally, we transform Channel-4 into Channel-5 in Figure D.1, a degraded broadcast channel, and show that the capacity region of these two channels are the same. First, the channel output at destination D_2 for Channel-4, i.e., $\tilde{Y}_2^n = (X_1^n + X_2^n) + (\tilde{X}_3^n + \tilde{X}_4^n) + Z_2^n + \Delta Z_2^n$, is statistically the same as the channel output at D_2 for Channel-5. In contrast, the channel output at destination D_1 for Channel-4 is

$$\tilde{Y}_1^n = (X_1^n + X_2^n) + Z_1^n, \tag{D.13}$$

while the same for Channel-5 is

$$\hat{Y}_1^n = (X_1^n + X_2^n) + (\tilde{X}_3^n + \tilde{X}_4^n) + Z_1^n. \tag{D.14}$$

Therefore, $\tilde{Y}_1^n$ in Channel-4 is *less noisy* than $\hat{Y}_1^n$ in Channel-5. Then, following [157, Figure 6] it can be shown that the capacity region of Channel-5 is at most as large as that of Channel-4.

We now show the other way, i.e., the capacity region of Channel-4 is at most as large as that of Channel-5, or in other words, an achievable rate tuple for Channel-4 is also achievable in Channel-5. Suppose that a rate tuple (R_1, R_2, R_3, R_4) is achievable in Channel-4, and hence from Fano's inequality we have

$$H(W_1, W_2|\tilde{Y}_1^n) \leq n\epsilon_{1,n}, \quad H(W_3, W_4|\tilde{Y}_2^n) \leq n\epsilon_{2,n}, \tag{D.15}$$

where $\epsilon_{1,n}, \epsilon_{2,n} \to 0$, as $n \to \infty$. We then need to show that the same rate tuple is achievable in Channel-5, i.e.,

$$H(W_1, W_2|\hat{Y}_1^n) \leq n\delta_{1,n}, \quad H(W_3, W_4|\hat{Y}_2^n) \leq n\delta_{2,n}, \tag{D.16}$$

where $\delta_{1,n}, \delta_{2,n} \to 0$, as $n \to \infty$.

Since $\hat{Y}_2^n$ in Channel-5 and $\tilde{Y}_2^n$ in Channel-4 are statistically the same, we only need to show that the implication for $H(W_1, W_2|\hat{Y}_1^n) \leq n\delta_{1,n}$. Consider the following term that can be expanded in two ways as follows

$$\begin{aligned} H(W_1, W_2, W_3, W_4|\hat{Y}_1^n) &= H(W_3, W_4|\hat{Y}_1^n) + H(W_1, W_2|W_3, W_4, \hat{Y}_1^n) \\ &= H(W_1, W_2|\hat{Y}_1^n) + H(W_3, W_4|W_1, W_2, \hat{Y}_1^n) \end{aligned}$$

We then have

$$\begin{aligned}
H(W_1,W_2|\hat{Y}_1^n) &= H(W_3,W_4|\hat{Y}_1^n) + H(W_1,W_2|W_3,W_4,\hat{Y}_1^n) - H(W_3,W_4|W_1,W_2,\hat{Y}_1^n) \\
&\overset{(a)}{\leq} H(W_3,W_4|\hat{Y}_1^n) + H(W_1,W_2|W_3,W_4,\hat{Y}_1^n) \\
&\overset{(b)}{\leq} n\epsilon_{2,n} + H(W_1,W_2|W_3,W_4,\hat{Y}_1^n) \\
&\overset{(c)}{=} n\epsilon_{2,n} + H(W_1,W_2|W_3,W_4,\tilde{X}_3^n,\tilde{X}_4^n,\hat{Y}_1^n) \\
&= n\epsilon_{2,n} + H(W_1,W_2|W_3,W_4,\tilde{X}_3^n,\tilde{X}_4^n,\hat{Y}_1^n - (\tilde{X}_3^n+\tilde{X}_4^n)) \\
&\overset{(d)}{=} n\epsilon_{2,n} + H(W_1,W_2|W_3,W_4,\tilde{X}_3^n,\tilde{X}_4^n,\tilde{Y}_1^n) \\
&\overset{(e)}{\leq} n\epsilon_{2,n} + H(W_1,W_2|\tilde{Y}_1^n) \\
&\overset{(f)}{\leq} n\epsilon_{2,n} + n\epsilon_{1,n}
\end{aligned} \tag{D.17}$$

where (a) follows since $H(W_3,W_4|W_1,W_2,\hat{Y}_1^n) \geq 0$; (b) follows by first noticing that $\hat{Y}_1^n$ is less noisy than $\hat{Y}_2^n$, i.e., $H(W_3,W_4|\hat{Y}_1^n) \leq H(W_3,W_4|\hat{Y}_2^n)$, and then from the fact in (D.15) that $H(W_3,W_4|\tilde{Y}_2^n) \leq n\epsilon_{2,n}$; (c) follows since for a given code, $\tilde{X}_3^n$ and $\tilde{X}_4^n$ are deterministic functions of W_3 and W_4; (d) follows since $\hat{Y}_1^n - (\tilde{X}_3^n + \tilde{X}_4^n))$ is statistically the same as $\tilde{Y}_1^n$; (e) follows from unconditioning; and (f) follows from (D.15).

Since $\epsilon_{2,n} + \epsilon_{1,n} \to 0$, as $n \to \infty$, messages (W_1, W_2) can be decoded from $\tilde{Y}_1^n$ in Channel-4. Therefore, Channel-4 has the same capacity region, and in particular, the same sum-capacity as that of Channel-5. Note that Channel-5 is a degraded Gaussian broadcast channel where the less noisy receiver has noise variance $N_1 = 1$ and power constraints P_1 and P_2, while the more noisy receiver has noise variance $N_2 = 1/h_1^2$ and power constraints P_3/h_1^2 and P_4/h_1^2. Standard converse technique for the degraded broadcast channel [139, Prop. 3] shows that the rates of sources S_1 and S_2, and the rates of sources S_3 and S_4 satisfy

$$R_1 + R_2 \leq \frac{1}{2}\log\left(1+P_1'\right), \quad R_3 + R_4 \leq \frac{1}{2}\log\left(1+\frac{P_2'}{P_1'+1/h_1^2}\right) \tag{D.18}$$

for some $P_1', P_2' > 0$ that satisfy $P_1' + P_2' = P_1 + P_2 + (P_3+P_4)/h_1^2$.

D.5 Proof of Theorem 5.5

Similar to that fo the IC in [139, Theorem 2], we provide the following "general" genie-signal

$$\tilde{Y}_2^n = d_1(X_3^n + X_4^n) + d_2(X_1^n + X_2^n) + d_3 Z_2^n + d_4 \tilde{Z}_2^{\,n} \tag{D.19}$$

to destination D_2 where $d_1, d_2, d_3, d_4 \in \mathbb{R}$, and $Z_2^n \sim \mathcal{N}(0,1)$, i.i.d., is defined in (5.3) while $\tilde{Z_2}^n \sim \mathcal{N}(0,1)$, i.i.d., is independent of all other variables.

Since for an achievable scheme, destination D_2 is able to decode (X_3^n, X_4^n) from its received signal Y_2^n, after receiving the genie-signal $\tilde{Y}_2^n$, D_2 creates an approximation of Y_1^n, i.e., the signal received at destination D_1, as follows

$$\begin{aligned}
\hat{Y}_1^n &= \alpha Y_2^n + \beta \tilde{Y}_2^n + (h_3 - (\alpha + \beta d_1))\, X_3^n + (h_3 - (\alpha + \beta d_1))\, X_4^n \\
&= \alpha\,(X_3^n + X_4^n + h_1 X_1^n + h_1 X_2^n + Z_2^n) + \beta\left(d_1(X_3^n + X_4^n) + d_2(X_1^n + X_2^n) + d_3 Z_2^n + d_4 \tilde{Z_2}^n\right) \\
&\quad + (h_3 - (\alpha + \beta d_1))\, X_3^n + (h_3 - (\alpha + \beta d_1))\, X_4^n \\
&= h_3 X_3^n + h_3 X_4^n + X_1^n(\alpha h_1 + \beta d_2) + X_2^n(\alpha h_1 + \beta d_2) + V^n, \qquad \text{(D.20)}
\end{aligned}$$

where $\alpha, \beta \in \mathbb{R}$, $V^n := Z_2^n(\alpha + \beta d_3) + \beta d_4 \tilde{Z_2}^n \sim \mathcal{N}(0, \sigma_V^2)$, i.i.d., with $\sigma_V^2 = (\alpha + \beta d_3)^2 + \beta^2 d_4^2$.

Now, taking α such that $\alpha h_1 + \beta d_2 = 1$ in (D.20), results in

$$\begin{aligned}
&\hat{Y}_1^n = h_3 X_3^n + h_3 X_4^n + X_1^n + X_2^n + V^n, \qquad \text{with} \\
&\sigma_V^2 = \left(h_1^{-1} + \beta(d_3 - d_2/h_1)\right)^2 + \beta^2 d_4^2. \qquad \text{(D.21)}
\end{aligned}$$

Adjusting parameters (β, d_2, d_3, d_4) such that $\sigma_V^2 \leq 1$ in (D.21) holds, the noise variance in $\hat{Y}_1^n$ is less than that in Y_1^n, i.e., $\hat{Y}_1^n$ is less noisy than Y_1^n. In other words, given (W_3, W_4) and $(Y_2^n, \tilde{Y}_2^n)$, $\hat{Y}_1^n$ can be constructed as in (D.20), which is less noisy than Y_1^n, i.e.,

$$H(W_1, W_2 | Y_2^n, \tilde{Y}_2^n, W_3, W_4) \leq H(W_1, W_2 | Y_1^n, W_3, W_4). \qquad \text{(D.22)}$$

We now upper bound the sum-rate as follows

$$\begin{aligned}
nC_{\mathsf{S}} &= n(R_1 + R_2) + n(R_3 + R_4) \\
&\overset{(a)}{\leq} I(W_1, W_2; Y_1^n) + I(W_3, W_4; Y_2^n) + n\epsilon_n \\
&\leq I(W_1, W_2; Y_1^n, W_3, W_4) + I(W_3, W_4; Y_2^n) + n\epsilon_n \\
&\overset{(b)}{=} I(W_1, W_2; Y_1^n | W_3, W_4) + I(W_3, W_4; Y_2^n) + n\epsilon_n \\
&= H(W_1, W_2) - H(W_1, W_2 | Y_1^n, W_3, W_4) + I(W_3, W_4; Y_2^n) + n\epsilon_n \\
&\overset{(c)}{\leq} H(W_1, W_2) - H(W_1, W_2 | Y_2^n, \tilde{Y}_2^n, W_3, W_4) + I(W_3, W_4; Y_2^n) + n\epsilon_n \\
&= I(W_1, W_2; Y_2^n, \tilde{Y}_2^n | W_3, W_4) + I(W_3, W_4; Y_2^n) + n\epsilon_n \\
&\overset{(d)}{\leq} I(X_1^n, X_2^n; Y_2^n, \tilde{Y}_2^n | X_3^n, X_4^n) + I(X_3^n, X_4^n; Y_2^n, \tilde{Y}_2^n) + n\epsilon_n \\
&= h(Y_2^n, \tilde{Y}_2^n) - h(Y_2^n, \tilde{Y}_2^n | X_3^n, X_4^n) + h(Y_2^n, \tilde{Y}_2^n | X_3^n, X_4^n) - h(Y_2^n, \tilde{Y}_2^n | X_1^n, X_2^n, X_3^n, X_4^n) \\
&= h(Y_2^n, \tilde{Y}_2^n) - h(Z_2^n, V^n)
\end{aligned}$$

$$
\begin{aligned}
&\overset{(e)}{\leq} nh(Y_{2G}, \tilde{Y}_{2G}) - nh(Z_2, V) \\
&= \frac{n}{2} \log \big[\big(1 + P_3 + P_4 + h_1^2(P_1+P_2)\big) \big(d_3^2 + d_4^2 + d_1^2(P_3+P_4) + d_2^2(P_1+P_2)\big) \\
&\qquad - \big(d_1(P_3+P_4) + h_1 d_2 (P_1+P_2) + d_3\big)^2 \big] - \frac{n}{2} \log d_4^2. \qquad \text{(D.23)}
\end{aligned}
$$

where (a) follows from Fano's inequality; (b) follows since $(W_1, W_2) \perp\!\!\!\perp (W_3, W_4)$; (c) follows from (D.22); (d) follows from providing $\tilde{Y}_2^n$ to the term for D_2; (e) follows by maximizing the $h(.)$ terms by choosing $X_k^n \sim \mathcal{N}(0, P_k), k \in \{1,2,3,4\}$ as in [131, Lemma 1], and the resulting received signals are given by Y_{2G} and $\tilde{Y}_{2G}$.

Hence, the sum-rate upper bound is given by the expression found by dividing the term on the r.h.s. of (D.23) by n, which we denote by $\Sigma(d_1, d_2, d_3, d_4)$. The tightest upper bound can then obtained by $C_{\mathsf{S}} \leq \min_{d_1,d_2,d_3,d_4} \Sigma(d_1, d_2, d_3, d_4)$ under constraint $\sigma_V^2 \leq 1$ in (D.21).

Note that having strict inequality in the constraint (D.21), i.e., $\sigma_V^2 < 1$ makes $\hat{Y}_1^n$ strict less noisy than Y_1^n, which results in strict inequality in step-(c) of (D.23). Hence, the tightest bound is obtained at equality $\sigma_V^2 = 1$, which leads to a quadratic equation in β. For β to be real, the discriminant of that equation must be non-negative, i.e.,

$$
(d_2 - h_1 d_3)^2 - d_4^2(1 - h_1^2) \geq 0. \qquad \text{(D.24)}
$$

Now, we provide a step-by-step procedure to minimize $\Sigma(d_1, d_2, d_3, d_4)$ in (D.23) under constraint (D.24). Instead of optimizing $\Sigma(d_1, d_2, d_3, d_4)$ jointly over (d_1, d_2, d_3, d_4), we first optimize only over d_1. We observe that $\Sigma(d_1, d_2, d_3, d_4)$ is convex over d_1, hence this minimizer is obtained as $d_1 = d_1^* = \frac{d_3 + d_2 h_1 (P_1+P_2)}{1 + h_1^2 (P_1+P_2)}$. Thus, the following upper bound is obtained

$$
\min_{d_1,d_2,d_3,d_4} \Sigma(d_1, d_2, d_3, d_4) \leq \min_{d_2,d_3,d_4} \min_{d_1} \Sigma(d_1, d_2, d_3, d_4) = \min_{d_2,d_3,d_4} \Sigma_1(d_2, d_3, d_4),
$$

where $\Sigma_1(d_2, d_3, d_4) = \min_{d_1} \Sigma(d_1, d_2, d_3, d_4) = \Sigma(d_1, d_2, d_3, d_4)|_{d_1 = d_1^*}$

The problem now simplifies to obtaining $\min_{d_2,d_3,d_4} \Sigma_1(d_2, d_3, d_4)$ under constraint (D.24), which is unchanged as d_1 is not involved in the constraint. As such, we take $d_3^* = -d_2 h_1 (P_1 + P_2)$, which results in the modified constraint $d_2^2(1 + h_1^3(P_1+P_2)) \geq d_4^2(1 - h_1^2)$. Finally, we take $d_4^* = 1$, and take the constraint with equality to obtain $d_2^* = \frac{(1-h_1^2)^{1/2}}{1 + h_1^2(P_1+P_2)}$. Hence, the following upper bound is obtained

$$
\min_{d_2,d_3,d_4} \Sigma_1(d_2, d_3, d_4) \leq \Sigma_1(d_2, d_3, d_4)|_{d_2 = d_2^*, d_3 = d_3^*, d_4 = d_4^*},
$$

and the final expression of this bound is given in Theorem 5.5.

Note that taking $d_3^* = -d_2 h_1 (P_1 + P_2)$ results in $d_1^* = 0$, which is justified as follows.

In (D.19), d_1 represents the weight of intended signals (X_3^n, X_4^n) for destination D_2. Since for an achievable scheme, (X_3^n, X_4^n) is assumed to be decodable at D_2, there is no need to additionally boost (X_3^n, X_4^n) by having $d_1^* \neq 0$, and thus $d_1^* = 0$ is chosen.

D.6 Proof of Theorem 5.6

We take $\ell = 2, k = 3$, and bound the sum-rate by $\mathsf{U}(X_3, X_2)$. Since all other bounds can be similarly obtained, we only provide the proof for $\mathsf{U}(X_3, X_2)$ here.

Note that $\mathsf{U}(X_3, X_2)$ consists of the minimum of two terms, the first of which is bounded as follows. From Fano's inequality, we have

$$
\begin{aligned}
&2n(R_1 + R_2 + R_3 + R_4) \\
&\leq I(X_1^n, X_2^n; Y_1^n) + I(X_1^n, X_2^n; Y_1^n) + I(X_3^n, X_4^n; Y_2^n) + I(X_3^n, X_4^n; Y_2^n) + n\epsilon_n \\
&\overset{(a)}{=} I(X_1^n, X_2^n; Y_1^n) + I(X_2^n; Y_1^n) + I(X_1^n; Y_1^n | X_2^n) \\
&+ I(X_3^n, X_4^n; Y_2^n) + I(X_3^n; Y_2^n) + I(X_4^n; Y_2^n | X_3^n) + n\epsilon_n \\
&\overset{(b)}{\leq} I(X_1^n, X_2^n; Y_1^n, X_3^n) + I(X_2^n; Y_1^n, X_1^n, X_4^n) + I(X_1^n; Y_1^n, X_3^n, X_4^n | X_2^n) \\
&+ I(X_3^n, X_4^n; Y_2^n, X_2^n) + I(X_3^n; Y_2^n, X_1^n, X_2^n, X_4^n) + I(X_4^n; Y_2^n, X_1^n, X_2^n | X_3^n) + n\epsilon_n \\
&\overset{(c)}{=} I(X_1^n, X_2^n; Y_1^n | X_3^n) + I(X_2^n; Y_1^n | X_1^n, X_4^n) + I(X_1^n; Y_1^n | X_3^n, X_4^n, X_2^n) \\
&+ I(X_3^n, X_4^n; Y_2^n | X_2^n) + I(X_3^n; Y_2^n | X_1^n, X_2^n, X_4^n) + I(X_4^n; Y_2^n | X_1^n, X_2^n, X_3^n) + n\epsilon_n \\
&= I(X_1^n, X_2^n; Y_1^n | X_3^n) + I(X_4^n; Y_2^n | X_1^n, X_2^n, X_3^n) \\
&+ I(X_3^n, X_4^n; Y_2^n | X_2^n) + I(X_1^n; Y_1^n | X_3^n, X_4^n, X_2^n) \\
&+ I(X_2^n; Y_1^n | X_1^n, X_4^n) + I(X_3^n; Y_2^n | X_1^n, X_2^n, X_4^n) + n\epsilon_n \\
&\overset{(d)}{=} \underbrace{h(Y_1^n | X_3^n) - h(Z_2^n)}_{d_1} + \underbrace{h(X_4^n + Z_2^n) - h(h_4 X_4^n + Z_1^n)}_{d_2} \\
&+ \underbrace{h(Y_2^n | X_2^n) - h(Z_1^n)}_{d_3} + \underbrace{h(X_1^n + Z_1^n) - h(h_1 X_2^n + Z_2^n)}_{d_4} \\
&+ \underbrace{h(Y_1^n | X_1^n, X_4^n) - h(Z_2^n)}_{d_5} + \underbrace{h(X_3^n + Z_2^n) - h(h_3 X_3^n + Z_1^n)}_{d_6} + n\epsilon_n \\
&\overset{(e)}{\leq} nh(Y_1^* | X_3^*) - nh(Z_2^*) + nh(X_4^* + Z_2^*) - nh(h_4 X_4^* + Z_1^*) \\
&+ nh(Y_2^* | X_2^*) - nh(Z_1^*) + nh(X_1^* + Z_1^*) - nh(h_1 X_2^* + Z_2^*) \\
&+ nh(Y_1^* | X_1^*, X_4^*) - nh(Z_2^*) + nh(X_3^* + Z_2^*) - nh(h_3 X_3^* + Z_1^*) + n\epsilon_n
\end{aligned}
$$

$$\stackrel{(f)}{=} nh(Y_1^*|X_3^*) - nh(Y_1^*|X_1^*, X_2^*, X_3^*, X_4^*) + nh(Y_2^*|X_1^*, X_2^*, X_3^*) - nh(Y_1^*|X_1^*, X_2^*, X_3^*)$$
$$+ nh(Y_2^*|X_2^*) - nh(Y_2^*|X_1^*, X_2^*, X_3^*, X_4^*) + nh(Y_1^*|X_4^*, X_2^*, X_3^*) - nh(Y_2^*|X_4^*, X_2^*, X_3^*)$$
$$+ nh(Y_1^*|X_1^*, X_4^*) - nh(Y_1^*|X_1^*, X_2^*, X_3^*, X_4^*) + nh(Y_2^*|X_1^*, X_2^*, X_4^*) - nh(Y_1^*|X_1^*, X_2^*, X_4^*)$$
$$+ n\epsilon_n \tag{D.25}$$

where (a) follows from the chain rule; (b) follows from providing the first term in (a) with X_3^n, second term with (X_1^n, X_4^n), third term with (X_3^n, X_4^n), and in a complementary manner, providing the fourth term with X_2^n, the fifth term with (X_1^n, X_2^n, X_4^n), and the last term in (X_1^n, X_2^n); (c) follows since any $X_k^n \perp\!\!\!\perp X_j^n, k \neq j \in \{1,2,3,4\}$; (d) follows from expressing mutual information in terms of differential entropies and rearranging; (e) follows by upper bounding d_1, d_3 and d_5 by using $X_k^n \sim \mathcal{N}(0, P_k), k \in \{1,2,3,4\}$ as in [131, Lemma 1], whereas terms d_2, d_4 and d_6 are also upper bounded by using $X_k^n \sim \mathcal{N}(0, P_k), k \in \{1,2,3,4\}$ and applying the worst additive noise result [130]; (f) follows from re-writing the terms in (e) in the form that, when combined, gives the expressions in Theorem 5.6.

This proves the first term inside the minimum of $\mathsf{U}(X_3, X_2)$. To prove the second term, we change the information provided in step (b) of (D.25): instead of (X_1^n, X_4^n), we provide (X_1^n, X_3^n, X_4^n) to the second term of (b), and instead of (X_1^n, X_2^n, X_4^n), we provide (X_1^n, X_4^n) to the fifth term of (b). The resulting terms are then obtained similarly to the above case.

D.7 Proof of Theorem 5.7

We denote the received signal at D_1 in the MAC-IC-MAC by $\tilde{Y}_1^n$, which is statistically the same as the received signal at D_1 in the GMAIC conditioned on X^4, i.e., $Y_1^n - h_4 X_4^n$. Considering a code for the GMAIC as in Def. 5.2, we now show that rate R_1 for the GMAIC can be upper bounded by rate $\tilde{R}_1$ for the MAC-IC-MAC.

For the GMAIC, we have

$$\begin{aligned} nR_1 &= H(M_1) \\ &\stackrel{(a)}{=} H(M_1|M_4) \\ &\stackrel{(b)}{=} I(M_1; Y_1^n|M_4, X_4^n) + H(M_1|Y_1^n, M_4, X_4^n) \\ &\stackrel{(c)}{=} I(M_1; Y_1^n|M_4, X_4^n) + H(M_1|\tilde{Y}_1^n, M_4, X_4^n) \\ &\stackrel{(d)}{\leq} I(M_1; Y_1^n|M_4, X_4^n) + H(M_1|\tilde{Y}_1^n) \end{aligned}$$

$$
\begin{aligned}
&= H(M_1) - H(M_1|Y_1^n, M_4, X_4^n) + H(M_1|\tilde{Y}_1^n) \\
&\overset{(e)}{=} H(M_1) - H(M_1|\tilde{Y}_1^n) + H(M_1|\tilde{Y}_1^n) \\
&= I(M_1; \tilde{Y}_1^n) + H(M_1|\tilde{Y}_1^n) \\
&\overset{(f)}{=} n\tilde{R}_1
\end{aligned} \tag{D.26}
$$

where (a) follows since $M_1 \perp\!\!\!\perp M_4$; (b) follows since X_4^n is deterministic function of M_4; (c) follows from observing that $\tilde{Y}_1^n$ has the same distribution as $Y_1^n|X_4^n$ in the second term; (d) follows from unconditioning; (e) follows from the Markov chain $M_1 \multimap \tilde{Y}_1^n \multimap (M_4, X_4^n)$; (f) follows since the rate of message M_1 in the MAC-IC-MAC is $n\tilde{R}_1 = I(M_1; \tilde{Y}_1^n) + H(M_1|\tilde{Y}_1^n)$.

Similarly, other rates R_2, R_2, R_4 for the GMAIC can be upper bounded by their counterparts in the MAC-IC-MAC, $\tilde{R}_2, \tilde{R}_3, \tilde{R}_4$. Thus, the capacity region of the GMAIC is contained within the capacity region of the MAC-IC-MAC. Finally, [145, Theorem 7] states that the capacity region of the MAC-IC-MAC is outer bounded by the region in (5.45), which hence outer bounds the capacity region of the GMAIC.