Diss. ETH No. 19720

Distributed Forwarding in Multiuser Multihop Wireless Networks

A dissertation submitted to the
ETH ZURICH

for the degree of
Doctor of Sciences

presented by
JÖRG WAGNER
Master of Science, ETH Zurich
born October 5, 1980
citizen of Germany

accepted on the recommendation of
Prof. Dr. Armin Wittneben, examiner
Prof. Dr. Robert Heath, co-examiner

2011

Day of Doctoral Examination: 16 May, 2011

Abstract

This thesis is concerned with physical layer signal processing and fundamental performance limits in wireless multiple-input multiple-output (MIMO) multihop networks. Such networks consist of a stage of source antennas, one or several stages of relay antennas and a stage of destination antennas. Signals traverse all these stages (one at a time) from source to destination antennas on the same physical channel. The schemes proposed in this work aim for transferring performance gains that are well known to be enabled by multiple transmit and receive antennas in point-to-point channels to this more general topology. Particular interest is on

- capacity scaling in the number of antennas per stage,
- spatial multiplexing and diversity gain.

Besides the topology (i) the amount of cooperation among antennas within a stage and (ii) the channel state information (CSI) that is accessible by the individual antennas must be specified in order to fully characterize a network configuration. In the special case that all antennas in the network possess the full CSI of their preceding hop each, and cooperation within stages is unconstrained, the above mentioned transfer is trivial: transmission can be decoupled into independent point-to-point MIMO transmissions, i.e., code-words are decoded and re-encoded in each stage (*decode & forward*). This thesis focuses on two selected network configurations that should not be decoupled in such fashion, since this straightforward approach does not suffice for realizing the desired multiple-antenna gains. Both these network configurations share the property that both cooperation among relay antennas and among source antennas is fully disabled as far as the exchange of information about the transmit signals is concerned.

The first considered configuration allows for full cooperation among antennas in the destination stage and assumes the relay antennas to possess the full CSI of the MIMO channel corresponding to the preceding hop each. The setting is referred to as *multihop multiple access channel* in this work. The simplest approach that comes to mind for coping with the spatial inter-symbol interference at the relay antennas is to have them perform a simple *amplify & forward* operation. However, under the assumption of an equal number of antennas in each stage, prior work has shown that this approach does not suffice for sustaining *linear capacity scaling* in the limit of infinitely many hops. More precisely, the constant of proportionality that relates the number of antennas per stage and sum-capacity tends to zero as the number of hops grows large. This thesis devises two approaches for coping with this issue.

- The first approach drops the amplify & forward strategy and resorts to a *quantization based forwarding* strategy. The presented scheme which applies *Slepian & Wolf compression* on top of the quantization of the received signals at the relay antennas succeeds in enabling linear sum-capacity in the limit of infinitely many hops, as long as the number of antennas in each relay stage grows linearly with the number of source and destination antennas.
- The second approach sticks to the amplify & forward strategy, but allows for increasing the ratio of relay antennas per stage to source and destination antennas per stage with the number of hops. As a key finding, it is proven, that this increase must be at least linear in the number of hops in order to sustain a linear sum-capacity scaling in the number of source and destination antennas in the limit of large numbers of hops.

The second considered configuration disables every cooperation even in the destination stage and assumes dedicated source-destination antenna pairs. The setting is referred to as *multihop interference channel* in this work. In contrast to the above multiple access scenario, the CSI of the full network is assumed to be known to each relay antenna. A decoupling of the network into point-to-point transmissions, e.g. through interference alignment and a decode & forward strategy, cannot yield a spatial multiplexing gain larger than $n/2$ for n source-destination antenna pairs (neglecting the pre-log) due to the lack of antenna cooperation. The contribution of this work in this context is summarized as follows:

- This thesis proposes a distributed zero-forcing scheme that achieves the optimal (given a sufficiently large number of relay antennas) spatial multiplexing gain, which is equal to the number of source-destination antenna pairs. The scheme is based on a *coherent* amplify & forward strategy.
- The optimization problem of maximizing the achievable rate of the weakest source-destination pair with respect to the relay gain coefficients in the coherent

amplify & forward framework is studied. For two hop-networks, the problem is turned into a quasi-convex problem. Such problems can be solved by standard optimization methods.

- An upper-bound on the achievable diversity-multiplexing tradeoff (DMT) curve is provided. The bound is based on the assumption of full relay cooperation within relay stages.
- Coherent amplify & forward relaying schemes are proposed, which – according to numerical evidence – achieve the derived DMT upper-bound, whenever the network topology allows for the full multiplexing gain.

Kurzfassung

In dieser Arbeit werden Verfahren zur Signalverarbeitung in der Bitübertragungsschicht sowie fundamentale Performance-Grenzen für drahtlose Multiple-Input Multiple-Output (MIMO) Multi-Hop-Netzwerke behandelt. Solche Netzwerke bestehen aus einer Gruppe von Quellenantennen, einer Gruppe oder mehreren Gruppen von Relaisantennen und einer Gruppe von Zielantennen. Signale durchlaufen Relaisgruppen von Quellen- zu Zielgruppe durch den selben physikalischen Übertragungskanal. Die Verfahren, die in dieser Arbeit vorgeschlagen werden, zielen darauf ab, Effizienzsteigerungen, die im Zusammenhang von Punkt-zu-Punkt-Übertragungen bereits wohlbekannt sind, auf diese verallgemeinerte Netzwerktopologie zu transferieren. Die Schwerpunkte liegen dabei auf

- der Skalierung der Kapazität mit der Antennenanzahl pro Quellen-, Relais- und Zielantennengruppe,
- räumlichem Multiplexgewinn ("Spatial Multiplexing Gain") und Diversitätsgewinn ("Diversity Gain").

Neben der Netzwerk-Topology müssen (i) der Umfang der Kooperation zwischen Antennen innerhalb einer Gruppe und (ii) die an den Antennen bekannte Information über den Kanalzustand spezifiziert werden, um die Netzwerkkonfiguration vollständig zu charakterisieren. Im Spezialfall, dass alle Antennen im Netzwerk den Kanalzustand ihres vorhergehenden Hops perfekt kennen und Antennen innerhalb der gleichen Gruppe unbeschränkt kooperieren können, ist die oben erwähnte Verallgemeinerung der Effizienzsteigerungen von Punkt-zu-Punkt-Netzwerken auf Multi-Hop-Netzwerke trivial. Die Übertragung kann in diesem Fall in unabhängige Punkt-zu-Punkt-MIMO-Übertragungen entkoppelt werden, d.h. Codewörter können in jeder Antennengruppe decodiert und dann erneut encodiert und weitergeleitet werden ("Decode & Forward"). Diese Arbeit konzentriert sich auf zwei ausgewählte Netzwerkkonfigurationen, in denen diese Entkopplung nicht vorgenommen werden sollte, da ein solch einfacher Ansatz nicht ausreicht, um die gewünschten MIMO-Gewinne zu erzielen. Die beiden untersuchten Netzwerkkonfigurationen teilen die gemeinsame Eigenschaft, dass sowohl

Quellenantennen als auch Relaisantennen keinerlei Informationen über ihre Sende- und Empfangssignale austauschen können.

Die erste betrachtete Konfiguration lässt unbeschränkte Kooperation in der Zielantennengruppe zu und nimmt an, dass Relaisantennen vollständige Information über den Kanalzustand des jeweils vorhergehenden Hops besitzen. Diese Anordnung wird als *Multi-Hop-Vielfachzugriffskanal* bezeichnet. Der naheliegendste Ansatz, mit der auftretenden Intersymbol-Interferenz an den Relaisantennen umzugehen, ist eine einfache Verstärkung und Weiterleitung des Empfangssignals ("Amplify & Forward"). Allerdings hat eine frühere Arbeit gezeigt, dass dieser Ansatz unter Annahme gleich grosser Antennengruppen nicht ausreicht, um eine lineare Skalierung der Kapazität mit der Antennenanzahl im Grenzwert unendlich vieler Hops aufrecht zu erhalten. Genauer gesagt, konvergiert die Proportionalitätskonstante, die die Anzahl Antennen pro Gruppe und die Summen-Kapazität verbindet, für eine wachsende Anzahl von Hops gegen Null. Diese Arbeit schlägt zwei Ansätze vor, dieses Problem zu lösen.

- Der erste Ansatz sieht vor, statt der "Amplify & Forward"-Strategie eine quantisierungsbasierte Methode anzuwenden. Das vorgestellte Verfahren, das zusätzlich zur Quantisierung eine Slepian & Wolf Komprimierung vornimmt, erlaubt eine lineare Skalierung der Summen-Kapazität im Grenzwert unendlich vieler Hops, solange die Anzahl Antennen in jeder Relaisantennengruppe zumindest linear mit der Anzahl Quellen- und Zielantennen wächst.
- Der zweite Ansatz hält an der "Amplify & Forward"-Strategie fest, erlaubt allerdings, das Verhältnis der Anzahl Relaisantennen pro Gruppe zu der Anzahl Quellen- und Zielantennen mit der Anzahl Hops im Netzwerk zu erhöhen. Ein Kernresultat dieser Arbeit ist der Beweis, dass dieses Verhältnis zumindest linear mit der Anzahl Hops anwachsen muss, um eine lineare Skalierung der Summen-Kapazität in der Anzahl Quellen- und Zielantennen auch im Grenzwert unendlich vieler Hops aufrecht zu erhalten.

Die zweite in dieser Arbeit betrachtete Netzwerkkonfiguration unterbindet jegliche Kooperation in der Zielantennengruppe des Netzwerks und nimmt an, dass Quellen- und Zielantennen in Kommunikationspaare gruppiert sind. Diese Anordnung wird als *Multi-Hop-Intereferenzkanal* bezeichnet. Im Gegensatz zum obigen Vielfachzugriff-Szenario wird hier angenommen, dass jede Relaisantenne den Kanalzustand des gesamten Netzwerks kennt. Eine Entkopplung des Netzwerks in Punkt-zu-Punkt-Übertragungen, z.B. durch "Interference Alignment"-Verfahren in Kombination mit einer "Decode& forward"-Strategie, lässt für n Quellen-Ziel-Antennenpaare aufgrund der unterbundenen Antennenkooperation keinen grösseren räumlichen Multiplexgewinn als $n/2$ (Pre-log vernachlässigt) zu.

Die Beiträge dieser Arbeit in diesem Zusammenhang lassen sich wie folgt zusammenfassen:

- Es wird ein Verfahren zur Interferenzunterdrückung vorgeschlagen, dass für eine hinreichend grosse Anzahl von Relaisantennen den optimalen räumlichen Multiplexgewinn erreicht. Dieses Verfahren basiert auf einer kohärenten "Amplify & Forward"-Strategie.
- Es wird das Optimierungsproblem der Maximierung der erreichbaren Rate des schwächsten Quellen-Ziel-Antennenpaares bezüglich der Relais-Verstärkungskoeffizienten untersucht. Für Zweihop-Netzwerke wird dieses Problem in ein quasikonvexes Problem überführt, das durch Standartoptimierungsmethoden gelöst werden kann.
- Es wird eine obere Schranke auf den erreichbaren "Diversity-Multiplexing Tradeoff" (DMT) hergeleitet. Die Schranke basiert auf der Annahme voller Kooperation zwischen Relaisantennen innerhalb der gleichen Relaisantennengruppe.
- Es werden kohärente "Amplify & Forward"-Verfahren vorgeschlagen, die gemäss numerischer Indizien an die hergeleitete obere Schranke für den DMT heranreichen, solange die Netzwerktopologie das Erreichen des vollen Multiplexgewinns erlaubt.

Acknowledgments

There is a group of individuals that has contributed greatly to the results of this thesis in one way or another.

First and foremost, my advisor Prof. Armin Wittneben who maintained an excellent working environment throughout my time with the Communication Technology Laboratory of ETH Zurich and provided me with guidance whenever it was needed. His way of thinking was a great source of inspiration.

Dr. Boris Rankov and Dr. Ying-Chang Liang who got me enthusiastic about the field of Wireless Communications during my MSc studies and thus laid the foundation for this work.

Dr. Marc Kuhn, Heinrich Lücken and Dr. Ingmar Hammerström who were always open to exchanging thoughts on current topics, sharing insights and providing advice. Their perspective on things was food for thought and motivation at the same time.

A variety of BSc and MSc students who did their student projects and theses with me and thus supported my research.

Contents

Abstract i

Kurzfassung v

Acknowledgments ix

1. Introduction 1

1.1. Background & Motivation 1

1.2. Contribution and Outline 8

2. Spatial Multiplexing in Multihop Interference Networks 13

2.1. Introduction & Related Work 13

2.2. Signal Model and Communication Protocol 18

2.3. Spatial Multiplexing Gain 21

2.3.1. Topology Requirements for Distributed Zero-Forcing in Networks with Two Source-Destination Pairs 24

2.3.2. Topology Requirements for Distributed Zero-Forcing in Networks with More Than Two Source-Destination Pairs 26

2.3.3. On the Number of Zero-Forcing Solutions 29

2.3.4. Achievability of Full Spatial Multiplexing Gain 30

2.3.5. Experimental Insights into Distributed Zero-Forcing 32

2.4. Excursus: "Ping-Pong" in Single-Hop Interference Networks 34

2.5. An Upper-Bound on the Achievable Diversity-Multiplexing Tradeoff . . 37

2.6. Max-Min Rate Optimization 44

2.6.1. Two-Hop Networks ($L = 1$) 48

2.6.2. General Multihop Networks 54

2.6.3. Is the DMT Upper-Bound Achievable? 65

2.6.3.1. Two-Hop Networks ($L = 1$) 67

2.6.3.2. Long Networks ($L \geq n$) 67

2.6.3.3. Short Networks ($2 \leq L \leq n-1$) 73
2.7. Concluding Remarks 74

3. Capacity Scaling of "Long" Multihop MIMO Multiple Access Networks 75
3.1. Introduction & Related Work 75
3.2. Signal Model & Communication Protocol 79
3.3. Multihop Relaying Techniques 82
3.3.1. Decode & Forward Relaying 82
3.3.2. Quantization Based Relaying 84
3.3.2.1. Pure Quantization 86
3.3.2.2. Quantization with Slepian & Wolf Compression 87
3.3.3. Amplify & Forward Relaying 88
3.4. Capacity Scaling 90
3.4.1. Decode & Forward Networks 90
3.4.2. Quantize & Forward Networks 91
3.4.2.1. Quantize & Forward without Slepian & Wolf compression 91
3.4.2.2. Quantization with Slepian & Wolf Compression 96
3.4.3. Amplify & Forward Networks 102
3.5. Concluding Remarks 122

4. Conclusions 123
4.1. Results 123
4.2. Open Aspects 124
4.3. Practical Challenges 125

A. Appendix 131
A.1. Proof of Proposition 4 131

B. Appendix 137
B.1. Proofs of Lemmata for Theorems 3 and 4 137
B.2. Proofs of Lemmata for Theorem 5 145
B.3. Proof of Proposition 5 155

Acronyms 159

Glossary 161

Bibliography 165

1. Introduction

1.1. Background & Motivation

In the last decade, technical development in the field of *wireless communications* has experienced a humongous boost. The driver of this development is found in the ever-growing demand for high data rates and the corresponding need for a more efficient use of the radio spectrum, which went along with the beginning of the age of multimedia. From that time on, academic research has provided a plethora of fundamentally new insights, which in turn have triggered a vast number of technical innovations and inventions. As a result, the wireless industry is one of the most dynamic of its kind and constantly undergoes far-reaching changes.

One of the most prominent examples for the rapid evolution of wireless technology is the area of mobile cellular networks (e.g. [1]). In the recent past, Switzerland's leading telecommunication provider Swisscom has recorded a doubling of the bandwidth requirements in its mobile networks every seven months [2]. And forecasts of mobile data requirements predict this massive growth to continue in the next years. A current world-wide forecast of Cisco Systems is shown in Fig. 1.1. The key to meeting these demands is the utilization of recent academic findings and the rapid realization of the corresponding signal processing algorithms in faster and faster very-large-scale integration (VLSI) technologies. The first commercial digital mobile communication system, GSM, was launched in the early 1990s and provided data rates up to 10 kbit/s at that time. Since then, new technologies have led to an explosive growth of supported data rates. The most promising candidate for the next generation of mobile communication systems, LTE Advanced, promises peak data rates up to 1 Gbit/s in the near future. An overview of the evolution of mobile communication systems is provided in Tab. 1.1.

Current wireless research is mainly attracted by two types of networks as far as rising of throughput is concerned. These are *wireless ad-hoc networks* on the one hand and *radio access networks* on the other hand. A wireless ad-hoc network (e.g. [3]) is usually

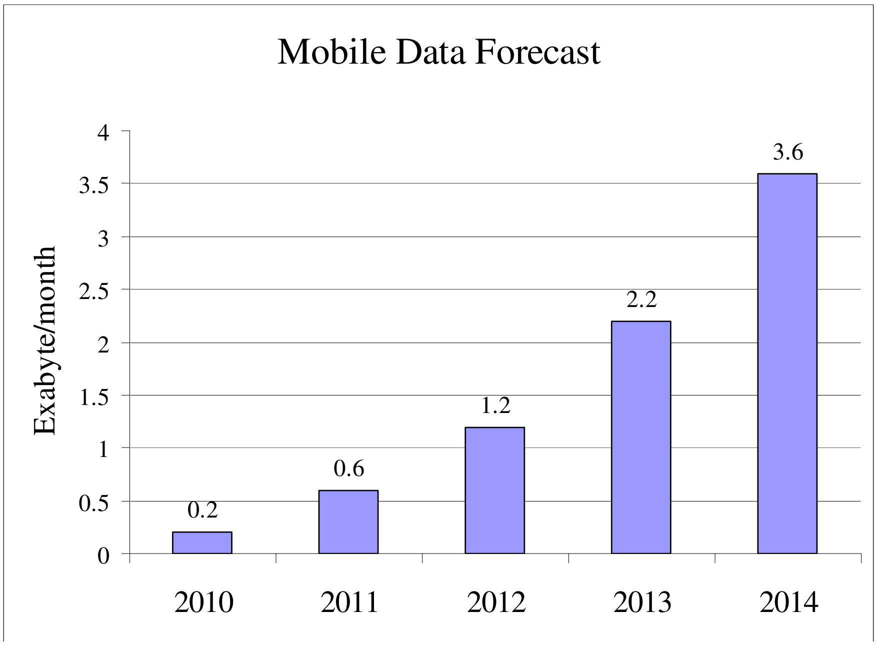


Figure 1.1.: Forecast of mobile data traffic per month. *Source: Cisco.*

Generation	Standards	Year Introduced	Peak Data Rates (Downlink)
2G	GSM	1992	9.6 kbit/s
	GPRS	2001	40 kbit/s
	EDGE	2005	200 kbit/s
3G	UMTS	2004	350 kbit/s
	HSDPA	2006	1.8 Mbit/s
	HSPA+	2009	28 Mbit/s
	LTE	2011	150 Mbit/s
4G	LTE Advanced	?	1 Gbit/s

Table 1.1.: Mobile network generations as introduced and operated by Swisscom.

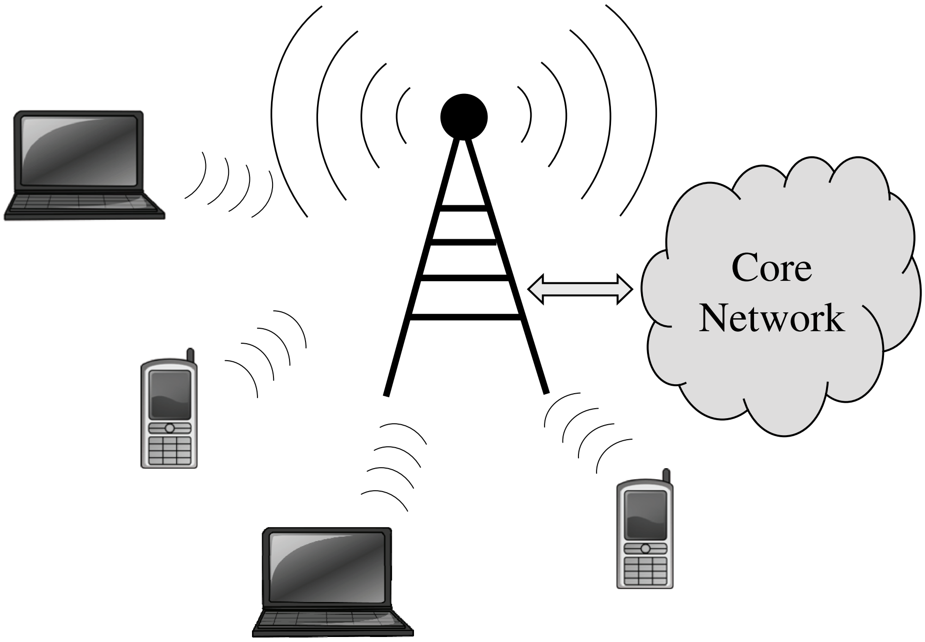


Figure 1.2.: Sketch of a typical present-day radio access network.

defined as a collection of mobile nodes that form a temporary wireless network without the aid of any centralized administration. Ad-hoc networks are typically symmetric in the sense that all nodes have equal rights and thus equal transmission capabilities. A radio access network (e.g. [4]) is the part of a mobile telecommunication system between mobile devices and base stations that connect to a core network. In such networks, base stations serve as central coordination units, and the base-station-to-mobile link (downlink) customarily supports larger data rates than the mobile-to-base-station link (uplink). Radio access networks are thus asymmetric in the sense that nodes have different functionality and that higher data rates are supported in one or the other direction.

Up to the time this writing, ad-hoc networks and radio access networks that have been brought to market have rather simple topologies. Traditional radio access networks are organized in "star topology", that is, all mobile nodes are in radio range of a base station and communicate to no other than this node (cf. Fig. 1.2). Wireless ad-hoc networks, such as WLAN in ad-hoc mode, have a fully connected mesh structure, that is, it is required that all terminals are in radio range of each other and communicate to each other directly (cf. Fig. 1.3).

From a physical layer engineering perspective, the understanding of radio access networks in this basic form is significantly more advanced than that of the basic wireless

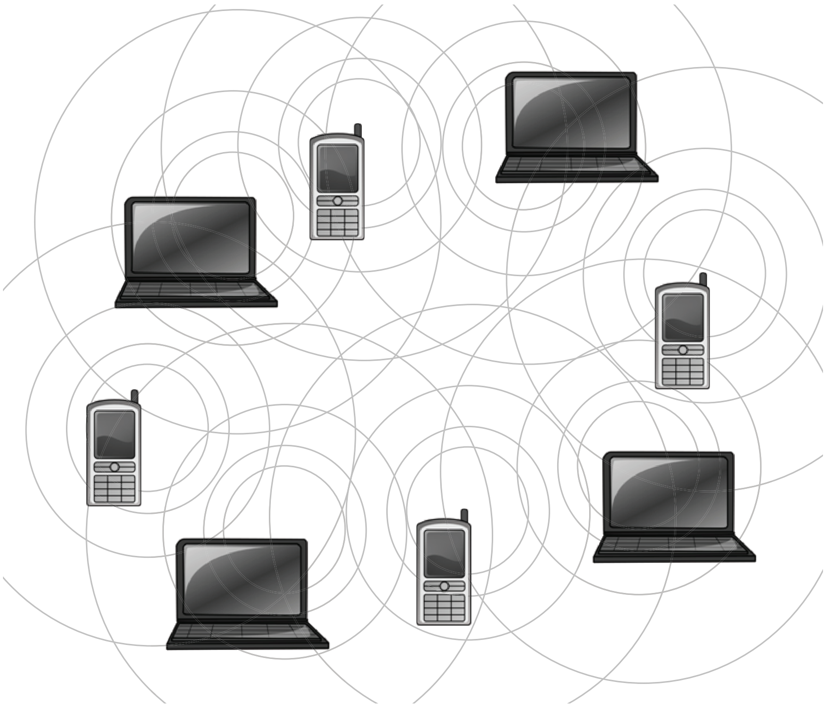


Figure 1.3.: Sketch of a typical present-day wireless ad-hoc network.

ad-hoc network. It is commonly accepted that the key to further leaps in data rates in future technologies is the use of multiple antennas at base stations and even in mobile devices. The revolutionary concept behind this approach is *multiple-input multiple-output (MIMO) communication* ([5–8]). While classical uplink and downlink schemes use the dimensions time or frequency or a combination of both for interference mitigation, the MIMO concept provides a "spatial" dimension for interference equalization. Using this new dimension thus facilitates higher spectral efficiency, which is defined as the number of transmitted bits per second and Hertz. In contrast to time and frequency, this dimension can be used "for free" in the sense that the only limiting factor is the area or volume of the antenna arrays. Luckily, for making the MIMO principle work, it is only essential to have multiple antennas either on the transmit or on the receive side. On the respective other side, the multiple antennas can be distributed, i.e., integrated into different devices. That is, a multi-antenna base station can exploit MIMO gains in up- and downlink, even if mobile devices are single-antenna terminals. This characteristic is important, since a base station is easily equipped with multiple antennas, while space constraints are more severe in mobile terminals. Recently launched radio access networks, such as LTE, WiMAX or WLAN 802.11n, are the first of their kind that realize the advantages of MIMO communication.

The fundamental limits of the basic wireless ad-hoc network of Fig. 1.3 are considerably less explored. In particular, it has remained unclear for a long time, to which extend the spatial dimension can be taken advantage of. A surprising and very recent finding is the following. Under the assumption that there are multiple source-destination node pairs (all with an equal number of antennas) that wish to exchange private messages each, the number of utilizable spatial dimensions is half the total number of transmit (or receive) antennas [9]. This insight is remarkable and implies that the spatial dimension can be gained access to, even if both transmit and receive antennas are distributed and unable to cooperate. The key concept behind this result is *interference alignment* – a novel scheme in communication engineering that currently attracts a lot of research activity. Needless to say, current wireless ad-hoc network, such as WLAN in ad-hoc mode, ZigBee or Bluetooth, do not make use of these spatial dimensions yet.

Future wireless networks are likely to have more complex structures than the two aforementioned basic networks. Recent advances in the area of *cooperative communication* (e.g. [10]) have revealed great potentials of cooperation among nodes in a wireless network. A wireless network is said to be cooperative, if there are nodes that not only act as either a source or a destination, but also as a helper for the communication of other nodes. An example of cooperative communication, which might soon be applied in practice, is cooperation among base stations in a cellular communication system (e.g. [11, 12]). The upcoming LTE-Advanced standard envisions this approach for increasing throughput both in up- and downlink [13]. The idea is to have base stations exchange information about their respective transmit or receive signals. This shall enable *distributed MIMO communication* (aka. *coordinated multi-point transmission*) through *virtual antenna arrays* and thus a significantly improved interference management. Cooperation among base stations is likely to occur over a wired backbone network for this application. However, in more distant future also wireless cooperation among mobile nodes is likely to come into play. In fact, such cooperation is the key to sustaining reasonable per-user throughput in large wireless ad-hoc networks [14, 15].

Multihopping is known as an efficient means for increasing the coverage of a network. In particular, when path loss is high, large wireless ad-hoc networks should make use of this option in order to achieve optimal throughput scaling in the number of nodes [14, 16]. It is highly relevant for future design both of wireless ad-hoc networks and of radio access networks. The lack of available spectrum will force wireless networks to higher and higher carrier frequencies, where unfavorable antenna and propagation properties lead to severe coverage limitations. The upcoming LTE-Advanced standard will be the first to apply relay nodes in a cellular network. Relays shall be operated by the provider and thus be part of the fixed network infrastructure. Multihopping in ad-hoc networks is significantly more challenging than in radio access networks due to the mobility of the relay nodes, and for implementation in commercial networks, there is still a long road to go.

The simplest form of multihopping bases upon the a forwarding technique known as *decode & forward*. In this scheme, a relay node decodes the message of the source node, re-encodes it, and forwards it either to another relay node or to the destination. This technique is probably the first that comes to mind, and, indeed, it is known to be optimal in many situations. However, it shows severe limitations in many aspects of cooperative communication. This thesis starts with one such aspect, namely the issue of *distributed forwarding* in wireless networks, which has its roots in the concept of *cooperative relaying* [17, 18] in the broadest sense. Distributed forwarding refers to the technique of having signals relayed not only via a single path from source to destination, but via several of them. The limitation of the decode & forward strategy is evident in this case: if any relay node on any path incurs a decoding error, the whole transmission is likely to be affected in a negative way. There are forwarding strategies that do not require relays to decode their receive signal. Thereby, they enable a relay node to effectively contribute to the transmission from source to destination, even if its link quality to adjacent nodes is weak. Such schemes are typically referred to as *non-regenerative*. The most prominent members of this family of schemes are *amplify & forward* and *quantize (compress) & forward*. These schemes are often superior to decode & forward, if interfering signals of multiple sources are relayed over multiple paths through a multihop network. An example of a three-hop radio access network with two user nodes and two parallel relay stations per stage is provided in Fig. 1.4. Likewise, Fig. 1.5 shows an example of a three-hop chain in a wireless ad-hoc network. Networks of this kind are in the center of this thesis.

Generally, the whole field of cooperative communication is a cutting-edge topic in wireless communication research. Although the potentials of many important building blocks have become evident in the recent past, it is still far from being fully understood and even further away from showing its full impact in practice. This thesis aims for shedding further light on the "multihop component" in cooperative communication, and focuses on two fundamental properties of it. An overview of the contributions is provided in the following section.

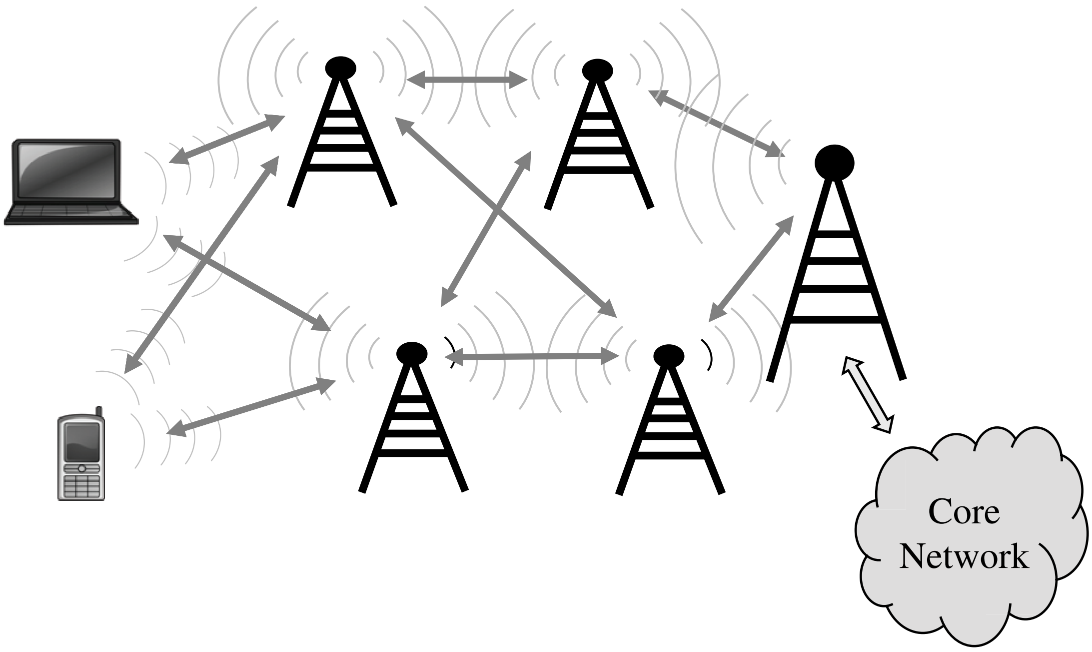


Figure 1.4.: Sketch of a multihop radio access network with infrastructure relays.

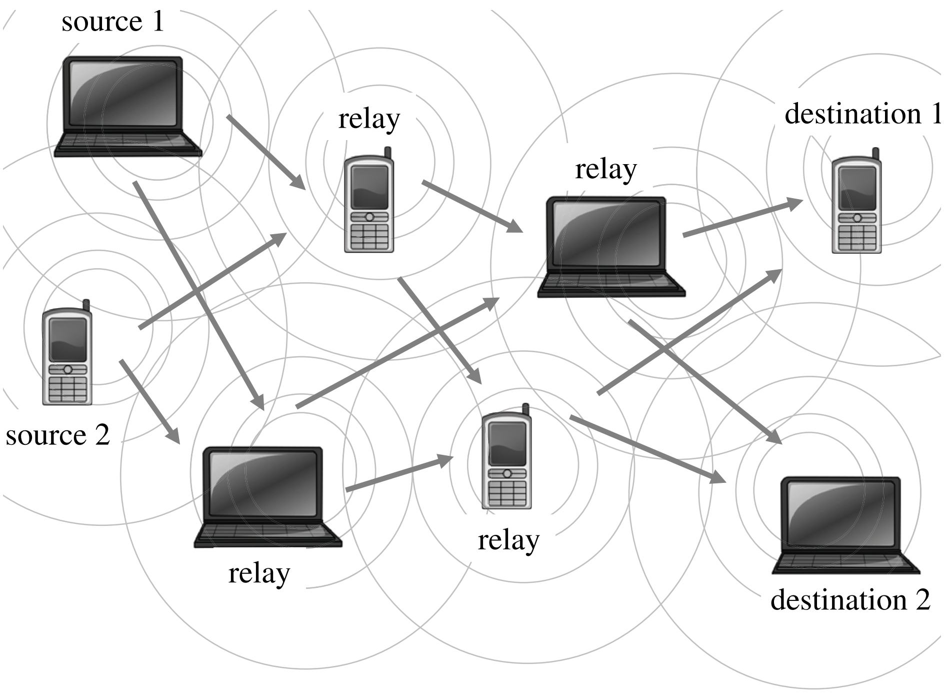


Figure 1.5.: Sketch of a multihop chain in a wireless ad-hoc network.

1.2. Contribution and Outline

The results of this thesis are formulated on the basis of data rates that are achieved in the information theoretic sense of Shannon [19, 20], which nowadays is a standard performance measure for physical layer signal processing schemes. In the notion of Shannon, a source node can transmit *reliably* to a destination node at rate R, if there exists a family of block codes with block-length n and 2^{nR} codewords, such that the average probability of a decoding error at the destination node tends to zero as n tends to infinity. Although the concept of achievable rates was introduced in the 1940s, its tremendous importance did not become evident until 1993. At that time, the advent of iterative decoding [21] enabled the realization of transmission close to the Shannon capacity with reasonable decoding complexity.

In contrast to wired communication channels, the wireless channel is often modeled as a random object. Multipath propagation, shadowing and the mobility of nodes cause variation in a wireless channel over time that are unpredictable without exhaustive geographic side-information. This phenomenon is referred to as *fading*. Fading is classified as either *slow* or *fast* according to the temporal variation of the channel gains and as either *flat* or *selective* according to the spectral variation of the channel gains. This work focuses on slowly and frequency selective fading channels.

As a consequence of the random nature of the wireless channel, also achievable rates are random quantities. A physical layer signal processing scheme is thus fully characterized by the joint statistical distribution of the respective achievable rates. Unfortunately, for most channels there is no analytical access to this distribution. Instead, studies attempt to gain insights into the performance of a scheme by a characterization of the distribution in asymptotic regimes. Also this thesis follows this approach. Specifically, two network structures are considered.

Chapter 2 studies *multihop interference networks*. Such a network consists of a cluster of n source nodes, a cluster of n destination nodes, and L relay node clusters with $n_{\mathcal{R}}^{(l)}$ relay nodes in the lth relay cluster. Source and destination nodes are grouped into n communication pairs. Source nodes simultaneously transmit their signals to the first relay cluster. Likewise, each relay cluster with index $l \in \{1, \ldots, L\}$ simultaneously forwards its signals either to the relay cluster with index $l+1$ or to the destination cluster (if $l = L$). All nodes in the network are assumed to have a single antenna and each relay node is assumed to know the channel state information (CSI) of the full network. This thesis proposes a scheme that guarantees under certain requirements on the network topology with probability one the existence of a set of achievable rates $(R_1, \ldots, R_n)$, R_i the rate corresponding to the ith source-destination pair, that as a function of the per-cluster transmit power P[1]

$$\lim_{P\to\infty} \frac{\sum_{i=1}^{n} R_i}{\log P} = n. \tag{1.1}$$

This sum-rate is achieved through a diagonalization of the network in the sense that all interference among source signals is eliminated (zero-forced) at the destination nodes. The proposed scheme is based on a coherent amplify & forward technique. Schemes that fulfill 1.1 are said to achieve a *spatial multiplexing gain* of n. This result is interesting in the sense that over single-hop interference networks without relay clusters no more than spatial multiplexing gain $n/2$ can be achieved [22].

In a second step, the proposed scheme is refined to design the relay gain coefficients for the amplify & forward operation to minimize the probability of an outage event. An outage event refers to the scenario that not all source-destination pairs are able to reliably communicate at a given target rate simultaneously due to poor channel conditions. This optimization corresponds to maximizing the achievable rate of the weakest source-destination pair with respect to the relay gain coefficients. For two hop-networks, this thesis turns the problem into a quasi-convex problem, which can be solved by standard methods.

A further contribution is the derivation of an upper-bound on the achievable diversity-multiplexing tradeoff (DMT) curve [23] in the case of identically and independently distributed (i.i.d.) circularly symmetric complex Gaussian (CSCG) channel coefficients. In the context of this work, the DMT curve characterizes the joint cumulative distribution function (CDF) of maximally achievable source-destination pair rates, $F_{R_1,\ldots,R_n}(r_1, \ldots, r_n)$, at the point $r_1 = \ldots = r_n = r$ for large per-cluster transmit power P, where r is a function of P. The bound is based on the assumption of full relay cooperation within relay stages. Thereupon, coherent amplify & forward relaying schemes are proposed, which – according to numerical evidence – achieve the derived DMT upper-bound, whenever the network topology facilitates the full multiplexing gain.

Chapter 3 studies *multihop multiple access networks*. Such a network differs form a multihop interference channel in that it allows full cooperation among the nodes

[1]It is assumed here that all clusters can transmit and receive simultaneously, and that clusters receive signals only from the preceding cluster.

in the destination cluster. That is, the n single-antenna source nodes transmit their messages to a single n-antenna destination via the multihop network. Again, the network comprises L relay stages, each of which is assumed to contain $n_{\mathcal{R}}$ single-antenna relay nodes. Schemes that are considered for this network in this thesis require nodes to have at most CSI of their respective preceding hop. The central assumption of this chapter is that all channel coefficients in the network are i.i.d. with zero mean and finite second and fourth moments. This class of channel distributions is quite broad in the sense that there are no further requirements on higher moments of the distribution. The quantity of interest in our setting is the following deterministic limiting value[2] of the random sum of rates that are achieved by the n source nodes:

$$c \triangleq \lim_{n\to\infty} \frac{\sum_{i=1}^{n} R_i}{n}. \tag{1.2}$$

This limit is evaluated for a constant value of the per-cluster transmit power P. It would be desirable to have this quantity nonzero, which is well known to be the case for the respective single-hop multiple access channels [6]. However, the question, whether or not this nice scaling behavior carries over to multihop networks, is nontrivial to answer in the context of distributed forwarding. Research on achievable rate scaling of multihop networks has so far focused on the amplify & forward relaying strategy. Two important prior results on this forwarding technique are the following. If $n_{\mathcal{R}}$ grows linearly with n, the quantity c is nonzero, if the number of hops is finite [24], but tends to zero[3], if L tends to infinity [25]. These results are the starting point for our studies, which aim for finding strategies that sustain a nonzero c, even in the limit of large L. The analysis of the corresponding schemes relies on results from large random matrix theory (e.g. [26]).

Our first approach drops the amplify & forward strategy and resorts to a *quantization based forwarding* strategy. The presented scheme, which applies *Slepian & Wolf compression* on top of the quantization of the received signals at the relay antennas, succeeds in keeping c bounded away from zeros in the limit of infinitely many hops, if $n_{\mathcal{R}}$ grows linearly with n. It is shown that the Slepian & Wolf compression step is not required for linear sum-rate scaling, however, it helps to reduce the required per-cluster transmit power P from exponential to linear in L.

Our second approach sticks to the amplify & forward strategy, but introduces an increasing ratio $n_{\mathcal{R}}/n$ with the number of hops. As a key finding, it is proved, that

[2]Convergence will be shown to be in the almost sure sense.
[3]This result is for noiseless relay nodes.

this increase must be at least linear in the number of hops in order to sustain a linear sum-rate scaling in n and thus a nonzero c in the limit of large numbers of hops.

In principle the results of Chapters 2 and 3 can find application both in future wireless ad-hoc and future radio access networks. Still, the symmetric setup of the multihop interference network in Chapter 2 is closer to wireless ad-hoc networks, while the asymmetric scenario with a multiple-antenna destination node in the multiple access network is closer to the classic scenario of a radio access network.

Chapter 4 finally provides concluding remarks and an outlook on future research on multiuser multihop networks.

2. Spatial Multiplexing in Multihop Interference Networks

2.1. Introduction & Related Work

The control of interference in communication networks that consist of multiple source-destination pairs is a crucial difficulty on the way towards an efficient utilization of the radio spectrum in future wireless ad-hoc networks. Accordingly, a lot of research has been pursued on this topic in recent years. A basic form of a wireless ad-hoc network network leads to the model of the *interference channel*. The interference channel models transmission between n single-antenna source-destination pairs over the same physical channel, where each destination is in radio range of all sources in general. Natural schemes for communication over this channel are time division multiple access (TDMA), frequency division multiple access (FDMA) or code division multiple access (CDMA). These schemes schedule source-destination pairs over orthogonal resources in time and/or frequency. Since the resources need to be shared among n source-destination pairs, the maximally achievable per-source-destination-pair data-rate scales like

$$1/n \log P + o(1/n \log(P)),$$

where P denotes the per-node transmit power. That is, in large networks, source-destination pairs communicate at unacceptable data-rates.

Recent research has discovered two fundamentally different approaches that overcome this severe limitation. Both methods use the spatial dimension as an additional resource and result into per-source-destination-pair data-rates that scale like

$$d \cdot \log P + o(\log(P)),$$

where $d = 1/2$ is independent of n, and the quantity $d \cdot n$ is typically referred to as spatial multiplexing gain or degrees of freedom. This result is remarkable: In the regime of high signal-to-noise ratio (SNR), each source-destination pair can, irrespective of the size of the network, operate at about half the capacity that would result, if there were no interfering nodes. In the following, a brief overview over these schemes is provided.

- *Interference alignment*: Interference alignment is a technique that introduces linear dependence among interference signals at each destination. That is, interference is at each destination concentrated in a subspace of small dimension. The idea of interference alignment goes back to the 1990s [27], where it has been used in the context of source coding. It has been rediscovered in [28] and [29] in the context of the so-called MIMO X channel. This channel is a generalized interference channel, where each source has messages for each destination. Interference alignment has finally been applied to the interference channel in [9] with the result that a spatial multiplexing gain $n/2$ is achieved in a sufficiently time- or frequency-selective channel. That is, the achievable data-rate of each source-destination pair is bounded away from zero, as the network grows large. Compared to the classical scenario of fully cooperating transmit and receive antenna arrays with n elements each, this corresponds to a loss of half the multiplexing gain. A major drawback of the interference alignment scheme of [9] are the considerable requirements on channel selectivity. In particular, the number of required fading realizations that a codeword must undergo, and thus the length of the codewords, grows rapidly with n. However, an interference alignment method has recently been proposed which even achieves the spatial multiplexing gain $n/2$ over a static channel [30].

- *Ergodic interference alignment*: Ergodic interference alignment [31] is a concept that applies to certain ergodic fading processes, in particular, also to i.i.d. (in space and time) Rayleigh fading. The basic idea is the matching of complementary channel realizations whose channel matrices add up to (almost) identity matrices. If the sources apply repetition coding with rate $1/2$ over two complementary channels, destinations can add up their receive signals in order to eliminate all interference. Also this scheme results into a spatial multiplexing gain of $n/2$, where the factor $1/2$ is an immediate consequence of the repetition code. Ergodic interference alignment involves significantly simpler coding than regular interference alignment. However, its restriction to very specific channel models limits its practical applicability. It is important to keep in mind that also regular interference alignment achieves a spatial multiplexing gain of $n/2$ over almost all ergodic channels.

This work goes beyond the basic model of the interference channel, and considers multihop interference networks. We expand the basic interference channel by an arbitrary number of relay stages, L, and assume that the lth relay stage comprises $n_{\mathcal{R}}^{(l)}$ single-antenna relay nodes (see sketch in Fig. 2.1). A straightforward approach for communication over a multihop interference network decouples the network into a cascade of single-hop interference networks. Such a decoupling is achieved, if all messages are decoded in every relay stage (decode & forward). It is then a straightforward conclusion that a spatial multiplexing gain of $\min(n/2, \min_l n_{\mathcal{R}}^{(l)}/2)$ is achievable[1] by the interference alignment technique of [9].

However, this is not the end of the story. Consider the following example of a two-hop network with two source-destination pairs and two relays. The two-hop interference channel can be turned into a concatenation of two X networks. In the first hop, each source transmits two messages to the relays (one to each), and, in the second hop, each relay has one message for each of the destinations. Even for non-selective channels, this approach allows for a spatial multiplexing gain of $4/3$ [33, 34], which is larger than $n/2$. Still, the approach of decoupling the multihop network into a concatenation of X networks through decode & forward relaying is suboptimal in general. The spatial multiplexing gain of an X network with n source nodes and m destination nodes is given by $mn/(m+n-1)$ [33] for sufficiently selective channels. That is, the achievable spatial multiplexing gain is strictly smaller than $\min(m, n)$. For our multihop network, this implies that the spatial multiplexing gain is strictly smaller than $\min(n, \min_l n_{\mathcal{R}}^{(l)})$. We conclude that the approach cannot be optimal in general, since there is the following method for two-hop networks, which does not decouple the network, and allows for the full spatial multiplexing gain of n for a sufficiently large number of relay nodes.

Distributed zero-forcing: Consider a two-hop network with $n_{\mathcal{R}}^{(1)} \geq n^2 - n + 1$ single-antenna relay nodes. For this network, it is known from [35] that all multiuser interference at the destination nodes can be canceled by the relays through a coherent amplify & forward architecture with a specific relay gain allocation. That is, each relay amplifies and phase-rotates its receive signal in a specific way before re-transmission, such that interference terms at the destination nodes add up destructively. Complexity is completely transferred to the relay nodes, and the individual source-destination pairs can perform standard single-input single-output (SISO) coding and decoding. Thus, n data streams are conveyed to the destination nodes over n parallel additive white

[1] At this point we ignore the facts that (i) relays typically cannot transmit and receive at the same time (half-duplex constraint, see [32] for a mitigation technique), and (ii) relay stages might pose interference to each other.

Gaussian noise (AWGN) channels, which results into a spatial multiplexing again of n. This distributed zero-forcing approach does not require the channels to be time or frequency selective and applies for generic channel coefficients.

This result on two-hop networks is the starting point for the contribution of this chapter. In the first instance, we aim for a generalization of distributed zero-forcing from two-hop networks to arbitrary multihop networks. A proof is provided that the generalized distributed zero-forcing scheme achieves the full spatial multiplexing gain 2 in multihop interference networks with $n = 2$, $L \geq 2$, $n_{\mathcal{R}}^{(l)} \geq 2$ for all $l \in \{1, \ldots, L\}$, and non-selective channels for generic channel coefficients. Moreover, the evidence is provided that the considerable constraint on the required number of relay nodes in the two-hop network can be significantly relaxed in "longer" networks. In particular, it is conjectured that $n_{\mathcal{R}}^{(l)} = n$ relay nodes in all stages stage $l \in \{1, \ldots, L\}$ suffice to enable distributed zero-forcing in networks with $L \geq n$, non-selective channels and generic channel coefficients.

There are very recent results on distributed spatial multiplexing in multiuser interference networks with *ergodic* channels [36–38]. These contributions propose a scheme that resembles the ergodic interference alignment approach for single-hop interference networks. It achieves full multiplexing gain n in multihop interference networks with $L = n - 1$ and $n_{\mathcal{R}}^{(l)} \geq n$ for all $l \in \{1, \ldots, L\}$ for a class of ergodic fading processes, in particular, also for i.i.d. (in space and time) Rayleigh fading. The corresponding approach matches (in the $L + 1$ hops) channel realizations whose corresponding channel matrices multiply to (almost) diagonal matrices. If relay stages properly delay and permute their receive signals, this allows for (almost) interference free communication over n parallel AWGN channels.

In a second step, we focus not only on the spatial multiplexing capabilities of the coherent amplify & forward architecture in multihop interference networks, but, consider also optimization of the relay gain coefficients with respect to link reliability. In particular, the relay gain allocation is optimized with respect to the smallest signal-to-interference-plus-noise ratio (SINR) among the source-destination pairs, which minimizes the outage probability. We succeed in turning this novel problem into a quasi-convex problem in the case $L = 1$. There is an efficient interior point method for the resulting problem. For $L > 1$, methods for finding local optima are devised. Among these local optima we identify a significantly smaller subset that contains potential candidates for the global optimum. Finally, two two types of low-complexity relay gain allocation methods are proposed.

For an evaluation of the optimized gain allocations, the standard approach of the DMT framework [23] is followed, which fully captures the relation between data-rate and outage probability in the regime of large SNR. An upper-bound on the achievable DMT curve is provided, which is essentially based on full cooperation among relay nodes within the same stage. Interestingly, numerical evidence suggests that the DMT upper-bound is achievable, whenever distributed zero-forcing is feasible. This would imply that the lack of cooperation among relay nodes within the same stage comes with no penalty in these cases as far as the DMT is concerned. A result of this flavor is known from [39] for two-hop networks with a single source-destination pair and multiple relays. If our observations are correct, the result of [39] can be generalized in the sense that it carries over both to multiple hops and multiple source-destination pairs.

There is, to our knowledge, no other work on the DMT of multihop interference networks. For the single-hop interference channel with two source-destination pairs, the DMT is studied in [40, 41]. There are also a variety of works that are related to the DMT of multihop networks with a single source-destination pair (also with multiple-antenna nodes) [42–45].

It is finally important to note that our concept of coherent amplify & forward relaying assumes all relay nodes to possess the CSI of the full network ("global CSI"). Whether or not the assumption of global CSI at the relay nodes, as it is made in our contribution, is reasonable in real-world networks certainly depends on the coherence time of the channel. Major progress in reducing the CSI dissemination overhead for coherent amplify & forward in two-hop interference networks has recently been made in [46, 47]. These contributions propose to determine interference suppressing relay gain coefficients through a gradient descent with respect to the SINRs of the source-destination pairs. Remarkably, the respective gradients can be computed based on limited feedback from the destination nodes that does not depend on the number of relay nodes in the network. Ongoing work shows that these findings also extend to networks with an arbitrary number of hops [48].

Organization of the chapter: Section 2.2 formally introduces the concept of coherent amplify & forward relaying and mathematically formulates the constraints on the relay gain coefficients that guarantee interference free communication in multihop interference networks. Necessary and sufficient conditions on the network topology that allow for meeting these constraints and the achievable spatial multiplexing gain are subject of Section 2.3. Section 2.4 is a brief excursus to single-hop interference networks. Section 2.5 provides an upper-bound on the achievable DMT curve of coherent amplify

& forward relaying in multihop interference networks. Optimal relay gain allocations with respect to the minimal SINR among the source-destination pairs and the DMT are subject of Section 2.6. Section 2.7 provides concluding remarks.

2.2. Signal Model and Communication Protocol

We consider a network that consists of $L+2$ stages: a source stage $\mathcal{S} = \{\mathrm{S}_1, \ldots, \mathrm{S}_n\}$ and a destination stage $\mathcal{D} = \{\mathrm{D}_1, \ldots, \mathrm{D}_n\}$ with n nodes S_k and D_k each, as well as L relay stages $\mathcal{R}_l = \{\mathrm{R}_1^{(l)}, \ldots, \mathrm{R}_{n_{\mathcal{R}}^{(l)}}^{(l)}\}$, $l = \{1, \ldots, L\}$, with $n_{\mathcal{R}}^{(l)}$ nodes $\mathrm{R}_k^{(l)}$ in the lth relay stage. All nodes are equipped with a single antenna, and nodes within the same stage are assumed not to exchange any information about their transmit or receive signals. We consider the transmission of n SISO codewords — one per source-destination pair $\{\mathrm{S}_k, \mathrm{D}_k\}$ — over the *same physical channel*. Transmissions are divided into $L+1$ time slots and initiated by the source nodes *simultaneously* and *in the same frequency band* in the first time slot. In time slot l, nodes in relay stage $\mathcal{R}_l$ receive interfering signals from relay stage $\mathcal{R}_{l-1}$, if $l > 1$, or the source stage $\mathcal{S}$, if $l = 1$. Each relay $\mathrm{R}_k^{(l)} \in \mathcal{R}_l$ scales the amplitude and rotates the phase of its receive signal. That is, before retransmission in time slot $l+1$, it performs a complex multiplication in equivalent base-band (*amplify & forward*). In time slot $L+1$, the nodes of the destination stage $\mathcal{D}$ receive the signals that are transmitted by the nodes of stage $\mathcal{R}_L$. A graph of the considered network is depicted in Fig. 2.1.

After at most $L+1$ time slots, the source stage can inject new narrow-band signals into the network without interfering transmissions of previous codewords. This interval can be reduced under certain circumstances. Let us, for instance, assume that nodes receive signals only from their adjacent stages. That is, channel coefficients between nodes that are located in nonadjacent stages are zero. Then, the source stage can inject new signals into the network in every third time slot ("reuse factor 1/3") without causing any temporal interference. The corresponding interference situations for reuse 1, 1/2 and 1/3 are sketched in Fig. 2.2. Note that such temporal interference is not an issue in two-hop networks. Thus, depending on whether relay nodes are full- or half-duplex, reuse 1 or 1/2 is feasible. The reuse factor is denoted by p^{-1} in the following.

We assume a slow and frequency flat fading channel model and denote the multiplicative equivalent base-band fading coefficient that corresponds to the transmission

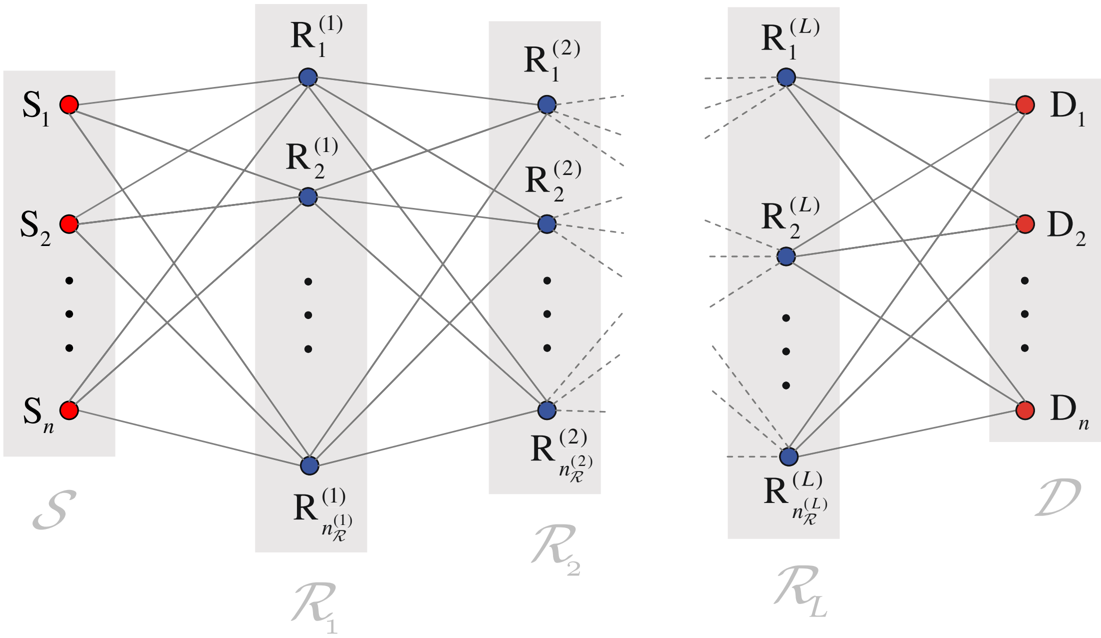


Figure 2.1.: Graph of the considered multihop interference network with source stage $\mathcal{S}$, destination stage $\mathcal{D}$, and relay stages $\mathcal{R}_1, \ldots, \mathcal{R}_L$.

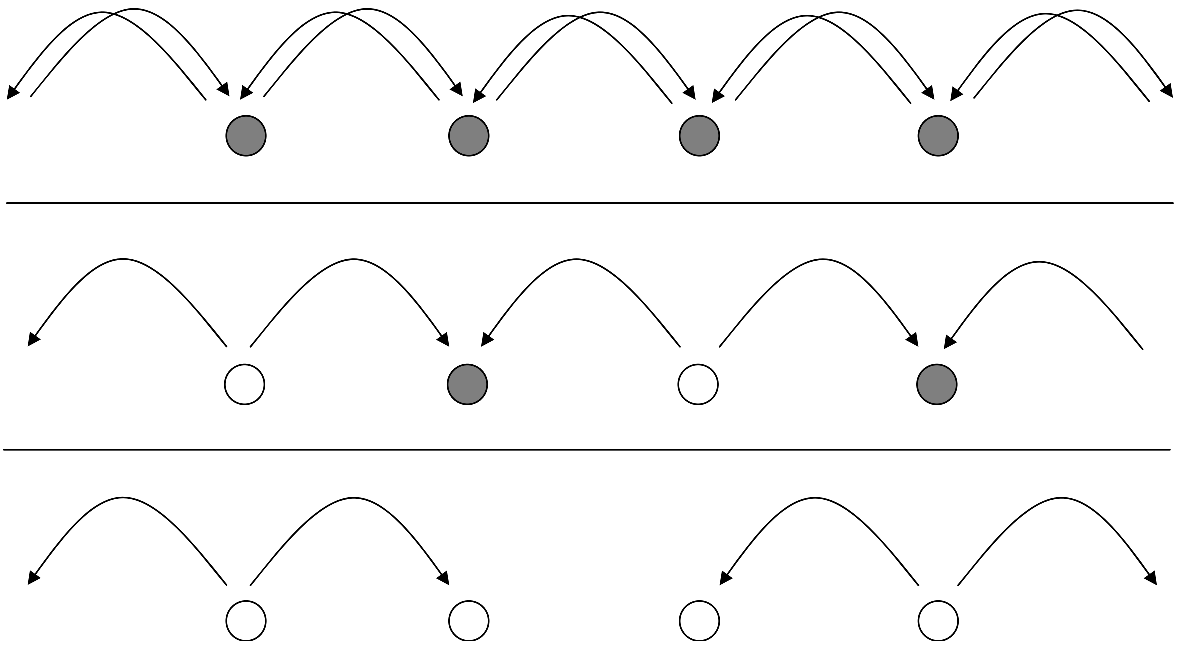


Figure 2.2.: Interference situation in a multihop chain for reuse 1 (top), 1/2 (middle) and 1/3 (bottom). Nodes with interfering receive signals are shaded. Channel coefficients between nodes that are located in nonadjacent stages are zero.

from node X to node Y by $h_{\mathrm{YX}} \in \mathbb{C}$. The symbol-discrete input-output (IO) relations from sources to destinations are fully described by the set of equations

$$\begin{aligned} y_{\mathrm{Y}} =& h_{\mathrm{YX}} \cdot x_{\mathrm{X}} + w_{\mathrm{Y}}, \\ & \qquad (\mathrm{X},\mathrm{Y}) \in (\mathcal{S} \times \mathcal{R}_1) \cup (\mathcal{R}_1 \times \mathcal{R}_2) \cup \ldots \cup (\mathcal{R}_L \times \mathcal{D}), \quad (2.1) \\ x_{\mathrm{Y}} =& g_{\mathrm{X}} \cdot y_{\mathrm{Y}}, \qquad \mathrm{Y} \in \mathcal{R}_1 \cup \ldots \cup \mathcal{R}_L. \quad (2.2) \end{aligned}$$

Here, $x_{\mathrm{X}} \in \mathbb{C}$ and $y_{\mathrm{Y}} \in \mathbb{C}$ denote the equivalent base-band representations of transmit and receive signal, $w_{\mathrm{Y}} \in \mathbb{C}$ AWGN of variance σ^2, and $g_{\mathrm{X}} \in \mathbb{C}$ the relay gain coefficient. The *effective* multiplicative fading coefficient $d_{\mathrm{D}_i\mathrm{S}_j}$ that corresponds to the transmission from source node S_j to destination node D_i is obtained as the superposition of all $\prod_{l=1}^{L} n_{\mathcal{R}}^{(l)}$ paths that connect these nodes in the network graph:

$$d_{\mathrm{D}_i\mathrm{S}_j} = \sum_{(\mathrm{R}_{k_1}^{(1)},\ldots,\mathrm{R}_{k_L}^{(L)}) \in \mathcal{R}_1 \times \ldots \times \mathcal{R}_L} h_{\mathrm{R}_{k_1}^{(1)}\mathrm{S}_j} h_{\mathrm{R}_{k_2}^{(2)}\mathrm{R}_{k_1}^{(1)}} \cdots h_{\mathrm{R}_{k_L}^{(L)}\mathrm{R}_{k_{L-1}}^{(L-1)}} h_{\mathrm{D}_i\mathrm{R}_{k_L}^{(L)}} g_{\mathrm{R}_{k_1}^{(1)}} \cdots g_{\mathrm{R}_{k_L}^{(L)}}. \quad (2.3)$$

Likewise, we obtain for the relay-to-destination and source-to-relay links:

$$d_{\mathrm{D}_i\mathrm{R}_j^{(l)}} = \sum_{(\mathrm{R}_{k_{l+1}}^{(l+1)},\ldots,\mathrm{R}_{k_L}^{(L)}) \in \mathcal{R}_{l+1} \times \ldots \times \mathcal{R}_L} h_{\mathrm{R}_{k_{l+1}}^{(l+1)}\mathrm{R}_j^{(l)}} \cdots h_{\mathrm{D}_i\mathrm{R}_{k_L}^{(L)}} h_{\mathrm{R}_{k_L}^{(L)}\mathrm{R}_{k_{L-1}}^{(L-1)}} g_{\mathrm{R}_{k_{l+1}}^{((l+1))}} \cdots g_{\mathrm{R}_{k_L}^{(L)}}, \quad (2.4)$$

$$d_{\mathrm{R}_i^{(l)}\mathrm{S}_j} = \sum_{(\mathrm{R}_{k_1}^{(1)},\ldots,\mathrm{R}_{k_{l-1}}^{(l-1)}) \in \mathcal{R}_1 \times \ldots \times \mathcal{R}_{l-1}} h_{\mathrm{R}_{k_1}^{(1)}\mathrm{S}_j} h_{\mathrm{R}_{k_2}^{(2)}\mathrm{R}_{k_1}^{(1)}} \cdots h_{\mathrm{R}_i^{(l)}\mathrm{R}_{k_{l-1}}^{(l-1)}} g_{\mathrm{R}_{k_1}^{(1)}} \cdots g_{\mathrm{R}_{k_{l-1}}^{(l-1)}}. \quad (2.5)$$

For notational convenience, we define the following matrices:

$$\mathbf{H}_l = \begin{cases} \left(h_{\mathrm{R}_i^{(1)}\mathrm{S}_j}\right)_{i=1,\ldots,n_{\mathcal{R}}^{(1)};j=1,\ldots,n} & \text{if } l = 1, \\ \left(h_{\mathrm{R}_i^{(l)}\mathrm{R}_j^{(l-1)}}\right)_{i=1,\ldots,n_{\mathcal{R}}^{(l)};j=1,\ldots,n_{\mathcal{R}}^{(l-1)}} & \text{if } 1 < l \leq L, \\ \left(h_{\mathrm{D}_i\mathrm{R}_j^{(L)}}\right)_{i=1,\ldots,n,j=1,\ldots,n_{\mathcal{R}}^{(L)}} & \text{if } l = L+1, \end{cases} \quad (2.6)$$

$$\mathbf{G}_l = \mathrm{Diag}\left(\mathbf{g}_l\right), \text{ with } \mathbf{g}_l = \left(g_{\mathrm{R}_k^{(l)}}\right)_{k=1}^{n_{\mathcal{R}}^{(l)}}, l \in \{1,\ldots,L\}, \quad (2.7)$$

$$\mathbf{D}_l = \begin{cases} \left(d_{\mathrm{D}_i\mathrm{S}_j}\right)_{i=1,\ldots,n;j=1,\ldots,n} = \mathbf{H}_{L+1}\mathbf{G}_L\mathbf{H}_L \cdots \mathbf{G}_1\mathbf{H}_1, \text{ if } l = 0 \\ \left(d_{\mathrm{D}_i\mathrm{R}_k^{(l)}}\right)_{i=1,\ldots,n;k=1,\ldots,n_{\mathcal{R}}^{(l)}} = \mathbf{H}_{L+1}\mathbf{G}_L\mathbf{H}_L \cdots \mathbf{H}_{l+1}\mathbf{G}_l, \text{if } l \in \{1,\ldots,L\}. \end{cases} \quad (2.8)$$

With this notation, the vector of received signals at the destination antennas, $\mathbf{y}_{\mathcal{D}}$, is obtained from the vector of source transmit signals, $\mathbf{x}_{\mathcal{S}}$, through the affine transformation

$$\mathbf{y}_{\mathcal{D}} = \mathbf{D}_0 \cdot \mathbf{x}_{\mathcal{S}} + \mathbf{w}_{\mathcal{D}} + \sum_{l=1}^{L} \mathbf{D}_l \cdot \mathbf{w}_{\mathcal{R}_l}, \tag{2.9}$$

where $\mathbf{w}_{\mathcal{D}}$ and $\mathbf{w}_{\mathcal{R}_l}$, $l \in \{1, \ldots, l\}$, denote the additive noise vectors of the respective stages.

Let P_{X} denote the average transmit power of node X. The assumptions on the power allocation are as follows:

- All source nodes transmit with equal power:

$$P_{\mathrm{S}_i} = \frac{P_{\mathcal{S}}}{n} \text{ for all } \mathrm{S}_i \in \mathcal{S}. \tag{2.10}$$

- The average sum-power of each relay stage $\mathcal{R}_l$, $P_{\mathcal{R}_l}$, $l \in \{1, \ldots, L\}$, is subject to the constraint:

$$P_{\mathcal{R}_l} \triangleq \sum_{k=1}^{n_{\mathcal{R}}^{(l)}} P_{\mathrm{R}_k^{(l)}} \leq \bar{P}_{\mathcal{R}_l}. \tag{2.11}$$

We address achievable spatial multiplexing and diversity gains in the course of this chapter, and thereby take the quantities $\bar{P}_{\mathcal{S}}$ and $\bar{P}_{\mathcal{R}_l}$ to infinity. In this context, we make the essential assumption that $P_{\mathcal{R}_l}$ as a function of $P_{\mathcal{S}}$ fulfills

$$\lim_{P_{\mathcal{S}} \to \infty} \frac{P_{\mathcal{R}_l}}{P_{\mathcal{S}}} = \gamma_l \text{ for all } l \in \{1, \ldots L\}, \tag{2.12}$$

where the γ_l fulfill $0 < \gamma_l < \infty$.

2.3. Spatial Multiplexing Gain

This section is concerned with the spatial multiplexing gain (also referred to as degrees of freedom) that can be achieved in the considered network. We use the standard definition of the spatial multiplexing gain [8]. Let R_i be an achievable rate of the ith source destination pair $\{\mathrm{S}_i, \mathrm{D}_i\}$. A spatial multiplexing gain of r is said to be achievable, if there exists a rate vector $(R_1, \ldots, R_n)$ that fulfills as a function of

$P \triangleq P_{\mathcal{S}} = P_{\mathcal{R}_1} = \ldots = P_{\mathcal{R}_L}$:

$$r \triangleq \lim_{P \to \infty} \frac{\sum_{i=1}^{n} R_i}{\log P}. \tag{2.13}$$

If $n_{\mathcal{R}}^{(l)} \geq n$ for all $l \in \{1, \ldots, L\}$, an upper-bound on the achievable spatial multiplexing gain is given by

$$r \leq n/p, \tag{2.14}$$

where we assume that the source stage injects new signals into the network in every pth time slot, and p is sufficiently large such that consecutive codewords do not interfere with each other. The upper-bound is established by allowing for full cooperation within all stages of the network. The network is then turned into a single-user MIMO network with a single multi-antenna node in each relay stage. Let us denote the sequences of random transmit and receive signal vectors of a stage $\mathcal{X}$ by $\left(\mathsf{x}_{\mathcal{X}}^{(j)}\right)_{j=1}^{N}$ and $\left(\mathsf{y}_{\mathcal{X}}^{(j)}\right)_{j=1}^{N}$, respectively, where N denotes the length of the sequences. Achievable (sum-) rates are then upper-bounded by [49]

$$\sum_{i=1}^{n} R_i < \frac{1}{p} \cdot \lim_{N \to \infty} \frac{1}{N} \cdot I\left(\left(\mathsf{x}_{\mathcal{S}}^{(j)}\right)_{j=1}^{N}; \left(\mathsf{y}_{\mathcal{D}}^{(j)}\right)_{j=1}^{N}\right). \tag{2.15}$$

Due to the Markov chain

$$\left(\mathsf{x}_{\mathcal{S}}^{(j)}\right)_{j=1}^{N} \to \left(\mathsf{y}_{\mathcal{R}_1}^{(j)}\right)_{j=1}^{N} \to \left(\mathsf{x}_{\mathcal{R}_1}^{(j)}\right)_{j=1}^{N} \to \ldots \to \left(\mathsf{y}_{\mathcal{R}_L}^{(j)}\right)_{j=1}^{N} \to \left(\mathsf{x}_{\mathcal{R}_L}^{(j)}\right)_{j=1}^{N} \to \left(\mathsf{y}_{\mathcal{D}}^{(j)}\right)_{j=1}^{N}, \tag{2.16}$$

the data-processing inequality [20] yields the upper-bound

$$\begin{aligned} & I\left(\left(\mathsf{x}_{\mathcal{S}}^{(j)}\right)_{j=1}^{N}; \left(\mathsf{y}_{\mathcal{D}}^{(j)}\right)_{j=1}^{N}\right) \\ & \quad \leq \min\left(I\left(\left(\mathsf{x}_{\mathcal{S}}^{(j)}\right)_{j=1}^{N}; \left(\mathsf{y}_{\mathcal{R}_1}^{(j)}\right)_{j=1}^{N}\right), I\left(\left(\mathsf{x}_{\mathcal{R}_L}^{(j)}\right)_{j=1}^{N}; \left(\mathsf{y}_{\mathcal{D}}^{(j)}\right)_{j=1}^{N}\right), \right. \\ & \qquad \left. \min_{l \in \{1,\ldots,L-1\}} I\left(\left(\mathsf{x}_{\mathcal{R}_l}^{(j)}\right)_{j=1}^{N}; \left(\mathsf{y}_{\mathcal{R}_{l+1}}^{(j)}\right)_{j=1}^{N}\right)\right) \qquad (2.17) \\ & \quad \leq N \cdot \min\left(I\left(\mathsf{x}_{\mathcal{S}}^{(j)}; \mathsf{y}_{\mathcal{R}_1}^{(j)}\right), \min_{l \in \{1,\ldots,L-1\}} I\left(\mathsf{x}_{\mathcal{R}_l}^{(j)}; \mathsf{y}_{\mathcal{R}_{l+1}}^{(j)}\right), I\left(\mathsf{x}_{\mathcal{R}_L}^{(j)}; \mathsf{y}_{\mathcal{D}}^{(j)}\right)\right). \qquad (2.18) \end{aligned}$$

For $\lim_{P\to\infty} P_{\mathcal{R}_l}/P = \gamma_l$, $0 < \gamma_l < \infty$, $l \in \{1,\dots,L\}$, we thus obtain [6]:

$$\lim_{P\to\infty} \frac{I\left(\mathsf{x}_{\mathcal{S}}^{(j)};\mathsf{y}_{\mathcal{D}}^{(j)}\right)}{\log P} \leq \min\{n, n_{\mathcal{R}}^{(1)}, \dots, n_{\mathcal{R}}^{(l)}\}, \tag{2.19}$$

which establishes (2.14).

In the following, we are interested in sets of relay gain coefficients that achieve this upper-bound. Specifically, we show that it suffices to choose the relay gain coefficients in a way, such that each destination node D_i receives no other signals than those of its associated source node S_i. We refer to this technique, which suppresses all spatial interference between the n source-destination pair, as *distributed zero-forcing.* In this context, "distributed" refers to the constraint that beamforming must be performed solely based on a multiplication of the received signal by a complex scalar at each relay. This is in contrast to a matrix multiplication of the receive signal vector of a relay stage, as it is enabled by full relay cooperation. In mathematical terms, the network is zero-forced, if both the following conditions are fulfilled:

$$d_{\mathrm{D}_i\mathrm{S}_j} = 0 \text{ for all } (i,j) \in \{1,\dots,n\}^2 \text{ such that } i \neq j, \tag{2.20}$$

$$d_{\mathrm{D}_i\mathrm{S}_i} \neq 0 \text{ for all } i \in \{1,\dots,n\}. \tag{2.21}$$

or, equivalently,

$$\mathbf{D}_0 - \mathbf{D}_0 \odot \mathbf{I}_n = \mathbf{0}_{n\times n}, \tag{2.22}$$

$$\mathrm{diag}(\mathbf{D}_0) \in (\mathbb{C}\backslash\{0\})^n. \tag{2.23}$$

The first condition ensures that all interference in the network is eliminated, while the second condition ensures that the desired signals are sustained. The $d_{\mathrm{D}_i\mathrm{S}_j}$ are multivariate polynomials in the $g_{\mathrm{R}_k^{(l)}}$. Zero-forcing of the network is thus possible, whenever the polynomials $d_{\mathrm{D}_i\mathrm{S}_j}$, $i \neq j$, have at least a single common root that is not a root of any of the polynomials $d_{\mathrm{D}_i\mathrm{S}_i}$. Note that solving the system (2.20) for the $g_{\mathrm{R}_k^{(l)}}$ requires global CSI.

In the following two subsections, we study the requirements on the network topology for the feasibility of distributed zero-forcing, or, equivalently the solvability of the system of equations (2.20) and inequations (2.21). In this context, we make the essential assumption that the vector of all elements h_{YX}, $(\mathrm{X},\mathrm{Y}) \in (\mathcal{S}\times\mathcal{R}_1)\cup(\mathcal{R}_1\times\mathcal{R}_2)\cup\ldots\cup(\mathcal{R}_L\times\mathcal{D})$ is a realization of a random vector whose joint distribution is continuous and nondegenerate. This assumption is reasonable, if the spacing between all pairs of nodes in the network is large compared to the carrier wavelength. We first consider the special case $n = 2$ in Subsection 2.3.1, where we prove a necessary and sufficient condition on the network topology for the feasibility of distributed zero-forcing. In Subsection 2.3.2, we consider the general case and devise a condition on the network topology that we conjecture to be necessary and sufficient for the feasibility of distributed zero-forcing for arbitrary n. Note that the feasibility of distributed zero-forcing is studied independently of all power constraints in the first instance. In Subsection 2.3.4, it is then shown that the power constraints can be enforced a posteriori through an appropriate scaling of the relay gain coefficients in each stage.

2.3.1. Topology Requirements for Distributed Zero-Forcing in Networks with Two Source-Destination Pairs

We consider the special case of two source-destination pairs ($n = 2$). From [35] it is known that a zero-forcing gain allocation in a two-hop network ($L = 1$) with two source-destination pairs exists with probability one, if and only if[2] $n_{\mathcal{R}}^{(1)} \geq 3$. In this subsection, we show that the corresponding necessary and sufficient condition for the case $L \geq 2$ is given by

$$n_{\mathcal{R}}^{(l)} \geq 2 \qquad \forall\, l \in \{1, \dots, L\}. \tag{2.24}$$

From conditions (2.22) and (2.23), it is clear that the rank of $\mathbf{D}_0$ must be two. Hence, the necessity of condition (2.24) follows from the inequality

$$\mathrm{rk}(\mathbf{D}_0) \leq \min\left(\min_{l\in\{1,\dots,L\}} (\mathrm{rk}\{\mathbf{G}_l\}), \min_{l\in\{1,\dots,L+1\}} (\mathrm{rk}\{\mathbf{H}_l\}) \right) \leq \min_l (\mathrm{rk}\{\mathbf{G}_l\}) \leq \min_l (n_{\mathcal{R}}^{(l)}), \tag{2.25}$$

where we use that the rank of the matrix product $\mathbf{D}_0$ (see (2.8)) is smaller than or equal to the rank of each involved matrix.

In order to prove sufficiency, we can restrict ourselves to the case $n_{\mathcal{R}}^{(l)} = 2$ for all $l \in \{1, \dots, L\}$. Once this case is established, the generalization to cases with $n_{\mathcal{R}}^{(l)} \geq 2$

[2] This condition is only claimed to be necessary in [35]. Under our assumptions on the fading distribution it is also sufficient.

is straightforward: If there are more than two relay nodes in a stage, one can choose an arbitrary subset of two relay nodes and assign zero gain coefficients to all remaining relay nodes in the network. Thus, one obtains a reduced network that fulfills (2.24) with equality. If $L \geq 3$, we set the gain coefficients of both relay nodes in each of the stages $\mathcal{R}_3, \ldots, \mathcal{R}_L$ to one. Likewise, we fix $g_{\mathrm{R}_1^{(1)}} = g_{\mathrm{R}_1^{(2)}} = 1$. With the above substitutions, the effective fading coefficients (2.3) simplify and (2.20) reduces to

$$\begin{aligned} d_{\mathrm{D}_2\mathrm{S}_1} =& h_{\mathrm{R}_1^{(1)}\mathrm{S}_1} h_{\mathrm{R}_1^{(2)}\mathrm{R}_1^{(1)}} \tilde{h}_{\mathrm{D}_2\mathrm{R}_1^{(2)}} + h_{\mathrm{R}_1^{(1)}\mathrm{S}_1} h_{\mathrm{R}_2^{(2)}\mathrm{R}_1^{(1)}} \tilde{h}_{\mathrm{D}_2\mathrm{R}_2^{(2)}} g_{\mathrm{R}_2^{(2)}} \\ &+ h_{\mathrm{R}_2^{(1)}\mathrm{S}_1} h_{\mathrm{R}_1^{(2)}\mathrm{R}_2^{(1)}} \tilde{h}_{\mathrm{D}_2\mathrm{R}_1^{(2)}} g_{\mathrm{R}_2^{(1)}} + h_{\mathrm{R}_2^{(1)}\mathrm{S}_1} h_{\mathrm{R}_2^{(2)}\mathrm{R}_2^{(1)}} \tilde{h}_{\mathrm{D}_2\mathrm{R}_2^{(2)}} g_{\mathrm{R}_2^{(1)}} g_{\mathrm{R}_2^{(2)}} = 0, \qquad (2.26) \\ d_{\mathrm{D}_1\mathrm{S}_2} =& h_{\mathrm{R}_1^{(1)}\mathrm{S}_2} h_{\mathrm{R}_1^{(2)}\mathrm{R}_1^{(1)}} \tilde{h}_{\mathrm{D}_1\mathrm{R}_1^{(2)}} + h_{\mathrm{R}_1^{(1)}\mathrm{S}_2} h_{\mathrm{R}_2^{(2)}\mathrm{R}_1^{(1)}} \tilde{h}_{\mathrm{D}_1\mathrm{R}_2^{(2)}} g_{\mathrm{R}_2^{(2)}} \\ &+ h_{\mathrm{R}_2^{(1)}\mathrm{S}_2} h_{\mathrm{R}_1^{(2)}\mathrm{R}_2^{(1)}} \tilde{h}_{\mathrm{D}_1\mathrm{R}_1^{(2)}} g_{\mathrm{R}_2^{(1)}} + h_{\mathrm{R}_2^{(1)}\mathrm{S}_2} h_{\mathrm{R}_2^{(2)}\mathrm{R}_2^{(1)}} \tilde{h}_{\mathrm{D}_1\mathrm{R}_2^{(2)}} g_{\mathrm{R}_2^{(1)}} g_{\mathrm{R}_2^{(2)}} = 0, \qquad (2.27) \end{aligned}$$

where

$$\tilde{h}_{\mathrm{D}_i\mathrm{R}_j^{(2)}} = \begin{cases} h_{\mathrm{D}_i\mathrm{R}_j^{(2)}}, \text{ if } L = 2, \\ \sum_{(\mathrm{R}_{k_3}^{(3)},\ldots,\mathrm{R}_{k_L}^{(L)}) \in \mathcal{R}_3 \times \ldots \times \mathcal{R}_L} h_{\mathrm{R}_{k_3}^{(3)}\mathrm{R}_j^{(2)}} \cdots h_{\mathrm{D}_i\mathrm{R}_{k_L}^{(L)}}, \text{ if } L \geq 3. \end{cases} \qquad (2.28)$$

The unknowns $g_{\mathrm{R}_2^{(1)}}$ and $g_{\mathrm{R}_2^{(2)}}$ are isolated as follows:

$$0 = a_1 + b_1 g_{\mathrm{R}_2^{(1)}} + c_1 g_{\mathrm{R}_2^{(1)}}^2 \qquad (2.29)$$

$$0 = a_2 + b_2 g_{\mathrm{R}_2^{(2)}} + c_2 g_{\mathrm{R}_2^{(2)}}^2, \qquad (2.30)$$

where a_1, b_1, c_1, a_2, b_2 and c_2 are functions of all fading coefficients h_{YX} and $\tilde{h}_{\mathrm{YX}}$. The equations have the solution(s) $g_{\mathrm{R}_2^{(i)}}^{(1,2)} = (-b_i \pm \sqrt{b_i^2 - 4a_i c_i})/(2c_i), i \in \{1,2\}$, if c_1 and c_2 are non-zero. This is the case, if and only if

$$\left(h_{\mathrm{R}_2^{(1)}\mathrm{S}_1}, h_{\mathrm{R}_2^{(1)}\mathrm{S}_2}, h_{\mathrm{R}_1^{(2)}\mathrm{R}_2^{(1)}}, h_{\mathrm{R}_2^{(2)}\mathrm{R}_1^{(1)}}, h_{\mathrm{R}_2^{(2)}\mathrm{R}_2^{(1)}}, \tilde{h}_{\mathrm{D}_2\mathrm{R}_2^{(2)}}, \tilde{h}_{\mathrm{D}_1\mathrm{R}_2^{(2)}} \right) \in (\mathbb{C}\backslash\{0\})^7 \qquad (2.31)$$

and both the following conditions are fulfilled:

$$h_{\mathrm{R}_1^{(1)}\mathrm{S}_1} h_{\mathrm{R}_2^{(1)}\mathrm{S}_2} \neq h_{\mathrm{R}_2^{(1)}\mathrm{S}_1} h_{\mathrm{R}_1^{(1)}\mathrm{S}_2} \qquad (2.32)$$

$$\tilde{h}_{\mathrm{D}_1\mathrm{R}_1^{(2)}} \tilde{h}_{\mathrm{D}_2\mathrm{R}_2^{(2)}} \neq \tilde{h}_{\mathrm{D}_2\mathrm{R}_1^{(2)}} \tilde{h}_{\mathrm{D}_1\mathrm{R}_2^{(2)}}. \qquad (2.33)$$

These conditions are equivalent to $\det(\mathbf{H}_1) \neq 0$ and $\det(\tilde{\mathbf{H}}_3) \neq 0$. That is, the channel matrices $\mathbf{H}_1$ and $\tilde{\mathbf{H}}_3$ need to have full rank. Since by assumption the channel coefficients are drawn from a nondegenerate continuous distribution, these conditions hold with probability one.

It remains to verify that condition (2.21) is fulfilled for the obtained solutions. Substitution of a solution into (2.21) yields that $d_{\mathrm{D}_1\mathrm{S}_1} \neq 0$ and $d_{\mathrm{D}_2\mathrm{S}_2} \neq 0$, if and only if

$$\left(h_{\mathrm{R}_1^{(1)}\mathrm{S}_1}, h_{\mathrm{R}_1^{(1)}\mathrm{S}_2}, h_{\mathrm{R}_1^{(2)}\mathrm{R}_1^{(1)}}, \tilde{h}_{\mathrm{D}_1\mathrm{R}_1^{(2)}}, \tilde{h}_{\mathrm{D}_2\mathrm{R}_1^{(2)}}\right) \in (\mathbb{C}\backslash\{0\})^5 \tag{2.34}$$

and

$$h_{\mathrm{R}_1^{(2)}\mathrm{R}_1^{(1)}} h_{\mathrm{R}_2^{(2)}\mathrm{R}_2^{(1)}} \neq h_{\mathrm{R}_1^{(2)}\mathrm{R}_2^{(1)}} h_{\mathrm{R}_2^{(2)}\mathrm{R}_1^{(1)}}. \tag{2.35}$$

The first condition interestingly involves all channel coefficients that have not been involved in (2.31). The second condition is equivalent to $\det(\mathbf{H}_2) \neq 0$. That is, the channel matrix $\mathbf{H}_2$ needs to have full rank. Again, these conditions are fulfilled with probability one.

In summary, the network can be zero-forced, if and only if all channel matrices have full rank, and none of the channel coefficients is zero.

We have thus proven, that a multiplexing gain $r = 2/p$ is achievable for $n = 2$ with probability one, if (2.24) is fulfilled.

2.3.2. Topology Requirements for Distributed Zero-Forcing in Networks with More Than Two Source-Destination Pairs

In this subsection, we consider the general case of an arbitrary number of source-destination pairs, n. For two-hop networks ($L = 1$), the system of equations and inequations (2.20) and (2.21) is linear, and therefore well analyzable by standard linear algebra methods [35]. A zero-forcing relay gain allocation in a two-hop network with n source-destination pairs exists with probability one, if and only if $n_{\mathcal{R}}^{(1)} \geq n^2 - n + 1$ relay nodes are available in the relay stage. If $L \geq 2$, the polynomial system of equations and inequations (2.20) and (2.21) becomes multi-linear. Whether or not solutions to such systems do exist, is well understood in the case that the coefficients of the involved polynomials are *generic*:

Definition 1. *A property is said to hold generically for polynomials $p_1, \ldots, p_m$, if there exists a non-zero polynomial in the coefficients of the p_i, such that the property holds* *for all $p_1, \ldots, p_m$ for which the polynomial is non-vanishing. Moreover, coefficients of the p_i for which this polynomial is non-vanishing are said to be generic.*

That is, for all coefficients except for a subset of measure zero, the solvability of a system of polynomial equations and inequations is fully determined by the structure of the involved monomials [50,51]. Unfortunately, the coefficients in (2.20) and (2.21) are not necessarily generic in spite of the randomness of the involved channel coefficients. There are $n^2 \prod_l n_{\mathcal{R}}^{(l)}$ monomials in (2.20) and (2.21). However, the respective coefficients depend on $n \cdot n_{\mathcal{R}}^{(1)} + \sum_{l=1}^{L-1} n_{\mathcal{R}}^{(l)} n_{\mathcal{R}}^{(l+1)} + n_{\mathcal{R}}^{(L)} \cdot n$ fading coefficients only. Thus, they are subject to a certain structure and therefore potentially non-generic. For this reason, it is difficult to prove general conditions on sets $\{n_{\mathcal{R}}^{(1)}, \ldots, n_{\mathcal{R}}^{(L)}\}$ that guarantee the existence of a zero-forcing gain allocation.

In this thesis, we restrict ourselves to stating the following conjecture, which is in line with numerical evidence (see Subsection 2.3.3 and Tab. 2.1).

Conjecture 1. *The subsequent conditions are necessary and sufficient (both conditions together) for the existence of a zero-forcing relay gain allocation in a multihop network with n source-destination pairs and L relay stages with probability one:*

$$\sum_{l=1}^{L} n_{\mathcal{R}}^{(l)} \geq n(n-1) + L, \tag{2.36}$$

$$n_{\mathcal{R}}^{(l)} \geq n \qquad \forall l \in \{1, \ldots, L\}. \tag{2.37}$$

For $L = 1$, condition (2.36) reduces to the well-known necessary and sufficient condition $n_{\mathcal{R}}^{(1)} \geq n(n-1) + 1$ for two-hop networks, while condition (2.37) is redundant [35]. Likewise, for $n = 2$ the condition (2.37) coincides with (2.24), while (2.36) is redundant. In full generality, we can neither prove sufficiency of both conditions nor necessity of the first condition. The necessity of the second condition is provable. From conditions (2.22) and (2.23) it is clear that the rank of $\mathbf{D}_0$ must be n. Hence, the necessity of condition (2.37) follows from inequality (2.25).

In the following, we illuminate a motivation for condition (2.36). Suppose the system (2.20) has a solution $(\mathbf{g}_1, \ldots, \mathbf{g}_L) = (\mathbf{g}_1^*, \ldots, \mathbf{g}_L^*)$ that fulfills (2.21). Then, there must exist infinitely many such solutions that constitute an L-dimensional affine variety: Let $c_1, \ldots, c_L$ be arbitrary non-zero complex scalars. Then, also $(\mathbf{g}_1, \ldots, \mathbf{g}_L) = (c_1\mathbf{g}_1^*, \ldots, c_L\mathbf{g}_L^*)$ fulfills (2.20) and (2.21). This is seen by inspection of (2.3), which reveals

$$d_{\mathrm{D}_i\mathrm{S}_j}\Big|_{(\mathbf{g}_1,\ldots,\mathbf{g}_L)=(c_1\mathbf{g}_1^*,\ldots,c_L\mathbf{g}_L^*)} = c_1\cdots c_L\cdot d_{\mathrm{D}_i\mathrm{S}_j}\Big|_{(\mathbf{g}_1,\ldots,\mathbf{g}_L)=(\mathbf{g}_1^*,\ldots,\mathbf{g}_L^*)}\ \forall\,(\mathrm{S}_j,\mathrm{D}_i)\in\mathcal{S}\times\mathcal{D}. \tag{2.38}$$

The left hand side of this equation is zero/non-zero, if and only if $d_{\mathrm{D}_i\mathrm{S}_j}\Big|_{(\mathbf{g}_1,\ldots,\mathbf{g}_L)=(\mathbf{g}_1^*,\ldots,\mathbf{g}_L^*)}$ is zero/non-zero. Due to property (2.38) the considered system (2.20) and (2.21) is said to be *L-homogeneous* (homogeneous in the L groups of gain coefficients) or also *L-linear* (since the c_l occur to the first power on the right hand side of (2.38)). While such systems have either no or infinitely many solutions in $\mathbb{C}^{n_{\mathcal{R}}^{(1)}}\times\ldots\times\mathbb{C}^{n_{\mathcal{R}}^{(L)}}$, they can have a finite number of solutions in the product of projective spaces[3] $\mathbb{P}^{n_{\mathcal{R}}^{(1)}-1}\times\ldots\times\mathbb{P}^{n_{\mathcal{R}}^{(L)}-1}$ (e.g. [52]).

For the consideration of the solvability of (2.20) and (2.21) we can fix without loss of generality $g_{\mathrm{R}_1^{(l)}}=1$ for all $l\in\{1,\ldots,L\}$. This dehomogenization reduces varieties of solutions to the original system to unique solutions in the dehomogenized problem. From [50], the following is known about the solvability of such dehomogenized system with *generic coefficients*:

- If condition (2.36) holds, there is one or several solutions to the dehomogenized system of equations (2.20) in[4] $(\mathbb{C}\backslash\{0\})^{(n_{\mathcal{R}}^{(1)}-1)}\times\ldots,\times(\mathbb{C}\backslash\{0\})^{(n_{\mathcal{R}}^{(L)}-1)}$. That is, the system is exactly determined.
- The solutions to the dehomogenized system of equations (2.20) cannot violate the dehomogenized system of inequations (2.21): If (2.21) contained equations rather than inequations, the resulting dehomogenized system of n^2 equations would be overdetermined [50]. We conjecture that also our coefficients render this system overdetermined in $(\mathbb{C}\backslash\{0\})^{(n_{\mathcal{R}}^{(1)}-1)}\times\ldots,\times(\mathbb{C}\backslash\{0\})^{(n_{\mathcal{R}}^{(L)}-1)}$, if condition (2.37) holds.
- Clearly, a reduction of the number of unknowns (total number of relay nodes) renders the dehomogenized system (2.20) overdetermined and does not allow for any solution in $(\mathbb{C}\backslash\{0\})^{(n_{\mathcal{R}}^{(1)}-1)}\times\ldots,\times(\mathbb{C}\backslash\{0\})^{(n_{\mathcal{R}}^{(L)}-1)}$. Vice versa, if the number of relay nodes is increased, the system is underdetermined and there are infinitely many solutions.

Thus, for generic coefficients, condition (2.36) would provably be necessary and suffi-

[3] The projective space $\mathbb{P}^k$ is the set of all k dimensional lines in $\mathbb{C}^{k+1}$ that pass through the origin.

[4] For the original system, there is in this case a solution in the product of projective spaces $\mathbb{P}^{n_{\mathcal{R}}^{(1)}-1}\times\ldots\times\mathbb{P}^{n_{\mathcal{R}}^{(L)}-1}$ (allows for zero components) for arbitrary non-zero coefficients [53].

cient for the existence of a solution to (2.20) that does not violate (2.21). According to this, we implicitly assume in our conjecture that this generic property carries over to the specific structure of the coefficients in (2.20) and (2.21), whenever (2.37) holds.

Generally, a necessary and sufficient condition for the existence of a solution is given by the projective weak Nullstellensatz [52]. The problem with this approach is the requirement of a reduced Groebner basis, which is difficult to obtain for general networks.

2.3.3. On the Number of Zero-Forcing Solutions

If condition (2.36) holds with equality, the dehomogenized system (2.20) has as many unknowns as equations. Suppose our conjecture is true, and the dehomogenized system (2.20) is exactly determined in $(\mathbb{C}\backslash\{0\})^{(n_{\mathcal{R}}^{(1)}-1)} \times \ldots, \times (\mathbb{C}\backslash\{0\})^{(n_{\mathcal{R}}^{(L)}-1)}$. Then, there is an upper-bound on the number of solutions. This bound is due to D.N. Bernstein [50], who linked the number of solutions of a system of polynomial equations to the structure of the Newton polytopes of the polynomials. A Newton polytope is defined as follows:

Definition 2. *Consider a polynomial $p(z_1, \ldots, z_m) = \sum_\alpha c_\alpha \prod_{i=1}^m z_i^{\alpha_i}$. The Newton polytope of p, denoted by Δ_p, is the convex hull of the set of exponents α, considered as vectors in $\mathbb{Z}^m$.*

By inspection of (2.3), we realize that all $d_{\mathrm{D}_i\mathrm{S}_j}$ share the same Newton polytope. Such systems are said to be *unmixed* and allow to use a corollary to Bernsteins' general result [54]:

Theorem 1 (Kushnirenko)**.** *If m polynomials $p_1, p_2, \ldots, p_m$ with identical Newton polytope have a finite number of joint zeros in $(\mathbb{C}\backslash 0)^m$, their number is upper-bounded by $m!\mathrm{Volume}(\Delta_p)$. The bound holds with equality for generic coefficients.*

For our problem this upper-bound is evaluated as follows [55, 56]:

$$\#\text{solutions} \leq \frac{(n^2-n)!}{\prod_{l=1}^{L}(n_{\mathcal{R}}^{(l)}-1)!}. \tag{2.39}$$

An important conclusion from this theorem is the following. The number of solutions to a system of polynomial equations is fully determined by the structure of the monomials for generic coefficients. As discussed in the previous section, the coefficients in (2.3) depend on a number of parameters which is significantly smaller than the number

Table 2.1.: Number of zero-forcing gain allocations for various network topologies.

$(n, n_{\mathcal{R}}^{(1)}, \ldots, n_{\mathcal{R}}^{(L)})$	solutions	upper-bound
$(n, n^2 - n + 1)$	1	1
$(2, 2, 2)$	2	2
$(3, 5, 3)$	6	15
$(3, 3, 5)$	6	15
$(3, 4, 4)$	12	20
$(3, 3, 3, 3)$	18	90
$(4, 7, 7)$	≥ 528	924
$(4, 5, 5, 5)$	n/a	34650
$(4, 4, 4, 4, 4)$	n/a	369600

of coefficients. According to the theorem, this lack of genericity can only lead to a reduction of the number of solutions.

In Tab. 2.1, we show for a couple of network topologies, how the actual number of solutions (as numerically identified by means of Newton's method and random initializations) compares to the upper-bound. We observe that both the actual number of solutions and the upper-bound increase rapidly both in the number of relay stages L and the number of source-destination pairs n. The upper-bound holds with equality in the cases $L = 1$ and $(n, L) = (2, 2)$. For larger L, the bound is loose in general. A comparison of the networks $(3, 4, 4)$, $(3, 3, 5)$ and $(3, 5, 3)$ suggests that uniform relay allocations over the stages yield more solutions than asymmetric allocations.

2.3.4. Achievability of Full Spatial Multiplexing Gain

In this subsection, we establish that the maximal spatial multiplexing gain n/p is achievable through distributed zero-forcing with probability one, whenever the network topology allows for an enabling relay gain allocation. The effective IO-relation between the n sources and n destinations under distributed zero-forcing decouples into n parallel AWGN channels:

$$y_{\mathrm{D}_i} = d_{\mathrm{D}_i\mathrm{S}_i} x_{\mathrm{S}_i} + \sum_{l=1}^{L} \sum_{k=1}^{n_{\mathcal{R}}^{(l)}} d_{\mathrm{D}_i\mathrm{R}_k^{(l)}} g_{\mathrm{R}_k^{(l)}} w_{\mathrm{R}_k^{(l)}} + w_{\mathrm{D}_i}, \qquad i = \{1, \ldots, n\}. \tag{2.40}$$

The maximum rate that is achievable over the ith of these parallel channels under the source power constraint is given by

$$\bar{R}_i = \frac{1}{p} \cdot \log \left(1 + \frac{P}{n} \cdot |d_{\mathrm{D}_i\mathrm{S}_i}|^2 \Bigg/ \left(\sigma^2 + \sum_{l=1}^{L} \sum_{k=1}^{n_{\mathcal{R}}^{(l)}} \left| g_{\mathrm{R}_k^{(l)}} \cdot d_{\mathrm{D}_i\mathrm{R}_k^{(l)}} \right|^2 \cdot \sigma^2 \right) \right), \tag{2.41}$$

where the $g_{\mathrm{R}_k^{(l)}}$ are subject to the sum-power constraints (2.11) with $\bar{P}_{\mathcal{R}_l} = P$. We fulfill all power constraints with equality, which yields

$$\mathbf{g}_l = \sqrt{\frac{P}{\tilde{P}}} \cdot \tilde{\mathbf{g}}_l \text{ for all } l \in \{1, \ldots, L\}, \tag{2.42}$$

with

$$\begin{aligned} \tilde{P} =& \frac{P}{n} \cdot \operatorname{Tr}\left[\tilde{\mathbf{G}}_l \mathbf{H}_l \cdots \tilde{\mathbf{G}}_1 \mathbf{H}_1 \mathbf{H}_1^H \tilde{\mathbf{G}}_1^H \cdots \mathbf{H}_l^H \tilde{\mathbf{G}}_l^H \right] \\ &+ \sigma^2 \cdot \operatorname{Tr}\left[\tilde{\mathbf{G}}_l \left(\mathbf{I}_{n_{\mathcal{R}}^{(l)}} + \mathbf{H}_L \left(\tilde{\mathbf{G}}_{l-1} \left(\mathbf{I}_{n_{\mathcal{R}}^{(l-1)}} + \ldots \mathbf{H}_{l'} \tilde{\mathbf{G}}_{l'} \tilde{\mathbf{G}}_{l'}^H \mathbf{H}_{l'}^H \ldots \right) \tilde{\mathbf{G}}_{l-1}^H \right) \mathbf{H}_L^H \right) \tilde{\mathbf{G}}_l^H \right], \end{aligned} \tag{2.43}$$

where $(\tilde{\mathbf{g}}_1, \ldots, \tilde{\mathbf{g}}_L)$ is a solution to the dehomogenized system (2.20) and (2.21) and $\tilde{\mathbf{G}}_l = \mathrm{Diag}(\tilde{\mathbf{g}}_l)$. Due to property (2.38), it is clear that the relay gain allocation (2.42) zero-forces the network.

We conclude from (2.41) that a multiplexing gain n/p is achieved in the network, if for each of the parallel channels

$$\lim_{P\to\infty} |d_{\mathrm{D}_i\mathrm{S}_i}|^2 > 0 \qquad \text{and} \qquad \lim_{P\to\infty} \sum_{l=1}^{L} \sum_{k=1}^{n_{\mathcal{R}}^{(l)}} \left| g_{\mathrm{R}_k^{(l)}} \cdot d_{\mathrm{D}_i\mathrm{R}_k^{(l)}} \right|^2 < \infty. \tag{2.44}$$

Both these conditions hold, if all relay gain coefficients converge to positive constants as $P \to \infty$. For the first relay stage, this follows immediately from (2.42):

$$\lim_{P\to\infty} \mathbf{g}_l = \tilde{\mathbf{g}}_1 \Bigg/ \sqrt{\operatorname{Tr}\left[\frac{1}{n} \cdot \tilde{\mathbf{G}}_1 \mathbf{H}_1 \mathbf{H}_1^H \tilde{\mathbf{G}}_1^H \right]}. \tag{2.45}$$

Now, it follows by induction that also the relay gain coefficients in all other relay stages converge to positive constants. If all relay gain coefficients in the stages $\mathcal{R}_{l'}$, $l' < l$, converge to positive constants, so do the gain coefficients in $\mathcal{R}_l$. From (2.42) we conclude:

$$\lim_{P\to\infty} \mathbf{g}_l = \tilde{\mathbf{g}}_l \Bigg/ \sqrt{\operatorname{Tr}\left[\frac{1}{n}\cdot \tilde{\mathbf{G}}_l\mathbf{H}_l \cdots \tilde{\mathbf{G}}_1\mathbf{H}_1\mathbf{H}_1^H\tilde{\mathbf{G}}_1^H \cdots \mathbf{H}_l^H\tilde{\mathbf{G}}_l^H\right]}. \tag{2.46}$$

This establishes that distributed zero-forcing achieves the full multiplexing gain n/p in the network with probability one, whenever it is feasible with probability one for the network topology.

2.3.5. Experimental Insights into Distributed Zero-Forcing

In this subsection, the sensitivity of the distributed zero-forcing performance to the choice of the zero-forcing solution in a network with a finite number of solutions is studied by means of simulations. Let us denote the maximally achievable rate (2.41) of source-destination pair $\{\mathrm{S_i}, \mathrm{D_i}\}$ under zero-forcing solution j (for $p = 1$) by $\bar{R}_i^{(j)}$. We are particularly interested in the sum-rate $\bar{R}_\Sigma^{(j)} \triangleq \sum_i \bar{R}_i^{(j)}$ and in the rate of the weakest source-destination pair $\bar{R}_{\min}^{(j)} \triangleq \min_i \bar{R}_i^{(j)}$ (referred to as minimum-rate). Realizations of the fading coefficients h_{YX} are generated independently from a CSCG random variable of unit variance. We consider a network with three source-destination pairs and two relay stages containing four relay nodes each. This network exhibits twelve different zero-forcing solutions. We conduct the following experiment: For 1000 channel realizations, we determine all twelve zero-forcing solutions numerically, and thereupon evaluate the corresponding rates $\bar{R}_\Sigma^{(j)}$ and $\bar{R}_{\min}^{(j)}$, $j \in \{1, \ldots, 12\}$. Average transmit powers are set to $P_\mathcal{S} = P_{\mathcal{R}_1} = P_{\mathcal{R}_2} = 1000$, and the noise variance to $\sigma^2 = 1$. We obtain empirical CDFs of the order statistics of $\bar{R}_\Sigma^{(j)}$ and $\bar{R}_{\min}^{(j)}$. The respective plots are shown in Fig. 2.3. The key conclusion to be drawn from these plots is that both for $\bar{R}_\Sigma^{(j)}$ and $\bar{R}_{\min}^{(j)}$ there are tremendous differences in performance. The average rate for the best solution is around five times larger than the one for the worst solution in terms of minimum-rate and still more than twice larger in terms of sum-rate. While the identification of the best zero-forcing solution based on a brute force search is manageable for the case at hand with only twelve solutions, such an approach seems to be hopeless for larger networks, when the number of solutions grows rapidly (cf. Tab. 2.1).

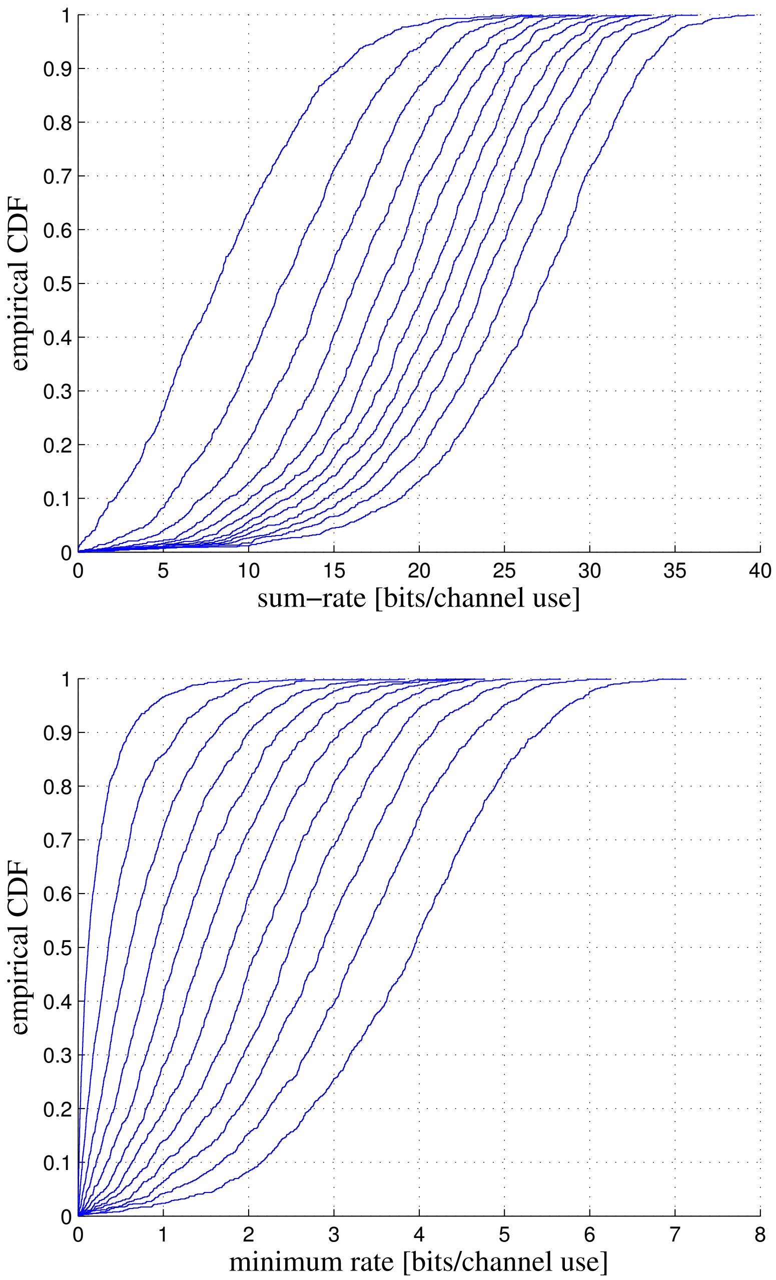


Figure 2.3.: Empirical CDFs of order statistics of sum-rates $(\bar{R}_{\Sigma}^{(1)}, \ldots, \bar{R}_{\Sigma}^{(12)})$ and minimum-rates $(\bar{R}_{\min}^{(1)}, \ldots, \bar{R}_{\min}^{(12)})$ for the three-hop network with $n = 3$, $n_{\mathcal{R}}^{(1)} = 4$, and $n_{\mathcal{R}}^{(2)} = 4$. The pre-log is set to $p^{-1} = 1$, the SNR to $P/\sigma^2 = 1000$.

2.4. Excursus: "Ping-Pong" in Single-Hop Interference Networks

Distributed zero-forcing in multihop interference networks inspires a new approach to communication in single-hop interference networks. A set of n source-destination pairs can communicate in an interference free fashion over the same physical channel, if both sets of nodes transmit their signals forth and back several times. This is illustrated on the left hand side of Fig. 2.4.

Let us define the matrix of channel coefficients $\mathbf{H} = (h_{\mathrm{D}_i\mathrm{S}_j})_{i,j=1,\dots,n}$. If channels are reciprocal, the fading matrix for the backward transmission from destination to source nodes is given by $\mathbf{H}^T$. Conditions (2.36) and (2.37) suggest that a multihop interference network with n relay stages with n relay nodes each can be zero-forced. Both conditions are fulfilled with equality for such a network. Accordingly, we ask the question, whether or not zero-forcing is also feasible in a single-hop interference network with n (even number) source-destination pairs by transmitting $n+1$ times forth and back. This strategy yields an equivalent $(n+1)$-hop network with effective fading matrix (cf. (2.8))

$$\mathbf{D}_0 = \mathbf{H}\mathbf{G}_n\mathbf{H}^T\mathbf{G}_{n-1}\cdots\mathbf{H}\mathbf{G}_2\mathbf{H}^T\mathbf{G}_1\mathbf{H}, \tag{2.47}$$

where all matrices $\mathbf{G}_l$ with odd indexes correspond to amplify & forward operations at the destination nodes and those with even indexes to amplify & forward operations at the source nodes. The $\mathbf{G}_l$, $l \in \{1, \dots, n\}$, need to be chosen to zero-force the equivalent $(n+1)$-hop network. For this particular setting, numerical experiments suggest that zero-forcing is indeed feasible. Also, the numbers of obtained solutions are in line with those obtained in the equivalent multihop networks. This is not evident, because the coefficients in (2.3) exhibit even more structure in this "ping-pong" scenario than in the case of a true multihop network.

The above example is of little practical value. In total, n data streams are transmitted over $n+1$ channel uses. This corresponds to a spatial multiplexing gain of $n/(n+1) < 1$. A simple time-division multiple access scheme outperforms this scheme. Nevertheless, the approach can be rendered useful. In order to achieve spatial multiplexing gains larger than one, the source and/or the destination stage need to be assisted by additional relay nodes. This is illustrated on the right hand side of Fig. 2.4. In doing so, we achieve zero-forcing with a reduced number of forth and back transmis-

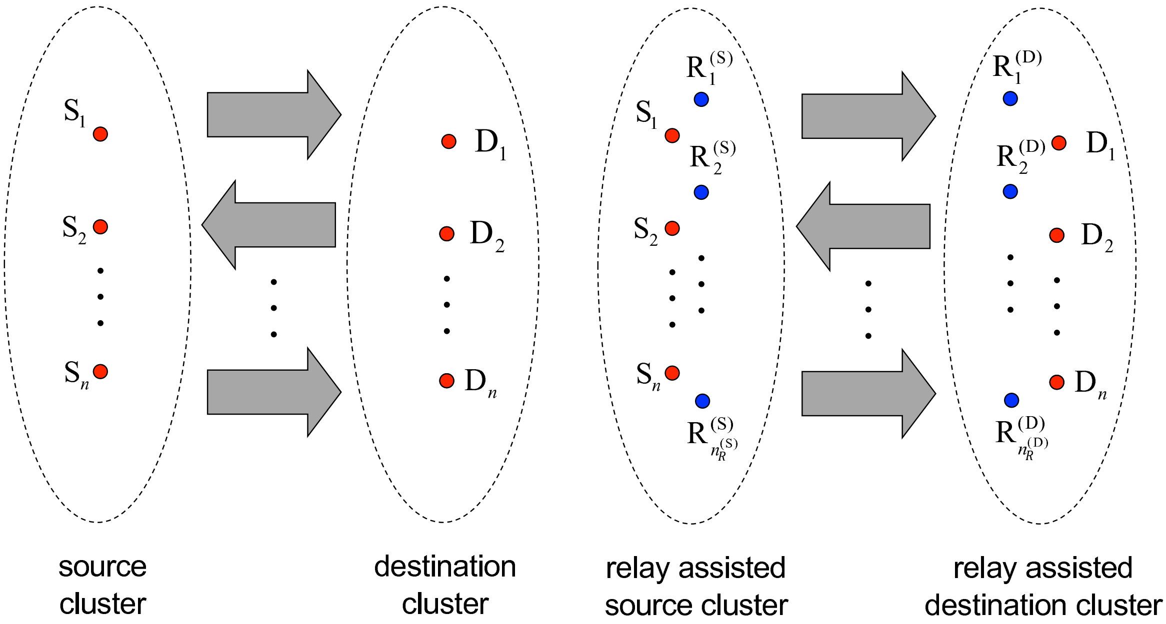


Figure 2.4.: "Ping-pong" in a single-hop network without (left) and with (right) assisting relay nodes.

sions K (an odd number ≥ 3). The respective spatial multiplexing gain is then given by n/K. Consider for example a network of four source-destination pairs, where source and destination stage are assisted by three additional relay nodes each. Indeed, this network can be zero-forced through three forth and back transmissions, which turns the network into an equivalent three-hop network with $n_{\mathcal{R}}^{(1)} = n_{\mathcal{R}}^{(2)} = 7$. Hence, a spatial multiplexing gain of $4/3$ is achieved.

If additional relay nodes are allocated uniformly over source and destination stage ($n_{\mathcal{R}}$ per stage), our approach achieves a spatial multiplexing gain of n/K according to (2.36), if

$$n_{\mathcal{R}} \geq \frac{n^2 - n}{K - 1} + 1 - n. \tag{2.48}$$

Numerical evidence suggests that for symmetric relay allocations the number of solutions to the system of equations (2.20) and inequations (2.21) is sustained despite of the added structure in the coefficients of the polynomials. This observation is food for the following thought:

Consider a single-hop interference network with $n' = n^2$ source-destination pairs. Apply the following time-division multiplexing scheme. The source-destination pairs

$(n, n_{\mathcal{R}}^{(1)}, n_{\mathcal{R}}^{(2)})$	solutions
$(3,4,4)$	12
$(3,5,3)$	0
$(3,3,5)$	0

Table 2.2.: Ping-Pong

are grouped into n disjoint sets $\mathcal{P}_i$*,* $i \in \{1, \ldots, n\}$*, of cardinality n each. There are n periods of* $K = 3$ *time slots each. In period i, the source-destination pairs in* $\mathcal{P}_i$ *are active, while all pairs in* $\bigcup_{j \neq i} \mathcal{P}_j$ *act as additional relay nodes. Thus,* (2.48) *is fulfilled, and in each period i, the multiplexing gain is* $n/3$ *or, equivalently* $\sqrt{n'}/3$*. After n periods, all source-destination pairs have been scheduled for transmission exactly once. Thus, the spatial multiplexing gain per source-destination pair is* $1/(3\sqrt{n'})$*, while the spatial multiplexing gain in the network scales with the square root of the number of source-destination pairs.*

Let us finally consider an asymmetric relay node allocation of additional relay nodes over source and destination stage. For simplicity, we assume a network of three source-destination pairs with two additional relay nodes in the destination stage, and no additional relay nodes in the source stage. Again, condition (2.36) is fulfilled with equality, and also condition (2.37) holds. In this example, however, we observe that no solution is found by our numerical solver. Here, the effect of added structure in the coefficients of the polynomials kicks in. Note that this problem is circumvented, if the forth and back transmissions are performed over different subcarriers in a frequency selective environment. Then, each transmission is associated with a different fading matrix, and the solvability conditions coincide with those for multihop networks.

Remark: A similar "ping-pong" technique with different purpose has been proposed in [57]. The scheme called *time reversal mirroring* bounces signals forth and back between a source and a mirror array for focusing signals in time and/or space. The technique has its origin in acoustics, and has also been applied to electromagnetic waves [58].

2.5. An Upper-Bound on the Achievable Diversity-Multiplexing Tradeoff

So far, we have been interested in the pure spatial multiplexing capabilities of the coherent amplify & forward architecture. In this section, we also shed light on the diversity performance and investigate the achievable DMT under the assumptions that the channels between any two nodes in adjacent stages are quasi-static and the channel matrices have i.i.d. entries $h_{\mathrm{YX}} \sim \mathcal{CN}(0,1)$. We provide an upper-bound that — according to numerical evidence — appears to be achievable, whenever distributed zero-forcing is feasible.

We start out with the definition of the tradeoff between diversity and multiplexing gain according to [23] tailored to our network. Assume that each of the n source-destination pair makes use of the same set of SISO codes and that the average per-stage transmit powers fulfill $P_{\mathcal{S}} = P_{\mathcal{R}_1} = \ldots = P_{\mathcal{R}_L} \triangleq P$. For a specific value of P, each source node chooses the same code, such that the code rate R as a function of P fulfills

$$\lim_{P\to\infty} \frac{R(P)}{\log P} \triangleq \frac{r}{n}. \tag{2.49}$$

The quantity r/n is referred to as the multiplexing gain of the set of codes. The diversity that is achieved by such a set of codes is defined as

$$-\lim_{P\to\infty} \frac{\log \mathsf{P}\left[\bigcup_{i=1}^{n} E_i \mid r\right]}{\log P} \triangleq d(r),$$

where E_i denotes the event of a maximum likelihood decoding error at the ith destination. That is, an error at a single destination node suffices to declare the whole network to be in outage. We refer to the function $d(r)$ as DMT curve. If capacity achieving codebooks are used, a decoding error occurs either due to a suboptimal relay gain allocation or due to an outage of the channel. In this case, the probability of a decoding error as a function of the SINRs at the destination nodes SINR_i, $i \in \{1,\ldots,n\}$, is written as

$$\mathsf{P}\left[\bigcup_{i=1}^{n} E_i \mid r\right] = \mathsf{P}\left[\bigcup_{i=1}^{n} \left\{\frac{1}{p}\log\left(1+\mathsf{SINR}_i\right) < R(P)\right\}\right]. \tag{2.50}$$

Again, the pre-log factor $1/p$ stems from the fact that new signals are injected into the network by the source stage in every pth time slot only.

We construct an upper-bound on the achievable DMT curve of an $L+1$-hop network with $n_{\mathcal{R}}^{(l)} \geq n$ relay nodes in the lth stage, $l \in \{1,\ldots,L\}$. To this end, we develop $L+1$ upper-bounds $\bar{d}_l(r)$, $l \in \{1,\ldots,L+1\}$, on the achievable DMT curve $d(r)$. Each bound $\bar{d}_l(r)$ is obtained through a specific isolation of the l-th hop in the network. Eventually, we combine all these bounds into the bound

$$d(r) \leq \min_l \bar{d}_l(r) \triangleq \bar{d}(r). \tag{2.51}$$

In order to obtain the individual bounds $\bar{d}_l(r)$, $l \in \{1,\ldots,L+1\}$, we apply the following relaxations, which for each value of the multiplexing gain r can only increase the DMT curve of the network $d(r)$:

- We neglect all noise in the network except for the noise that is introduced in the respective receive stage of hop l, which is $\mathcal{R}_l$, if $l \leq L$, and $\mathcal{D}$, if $l = L+1$.
- If $l > 1$, we replace the IO-relation of the subnetwork from source stage $\mathcal{S}$ to relay stage $\mathcal{R}_{l-1}$ by an arbitrary linear map that is defined by the matrix $\mathbf{G}_t^{(l-1)} \in \mathbb{C}^{n_{\mathcal{R}}^{(l-1)} \times n}$ and fulfills the sum-power constraint on $\mathcal{R}_{l-1}$.
- Likewise, if $l \leq L$, we replace the IO-relation of the subnetwork from relay stage $\mathcal{R}_l$ to the destination stage $\mathcal{D}$ by an arbitrary linear map that is defined by the matrix $\mathbf{G}_r^{(l)} \in \mathbb{C}^{n \times n_{\mathcal{R}}^{(l)}}$.

The second and third relaxations yield an upper-bound on the DMT, since we allow for an arbitrary linear processing on the transmit and/or receive side of the hop. That is, neither $\mathbf{G}_r^{(l)}$ needs to follow the structure $\mathbf{H}_{L+1}\mathbf{G}_L \cdots \mathbf{H}_2\mathbf{G}_l$ nor $\mathbf{G}_t^{(l)}$ needs to follow the structure $\mathbf{G}_l\mathbf{H}_l \cdots \mathbf{G}_1\mathbf{H}_1$. Note that in the physical network only the diagonal elements of the $\mathbf{G}_l$ can be varied.

For the evaluation of the resulting upper-bounds, it turns out that three cases have to be distinguished. These are (i) $l = 1$, (ii) $l = L+1$ and (iii) $2 \leq l \leq L$.

Case $l = 1$: This case corresponds to the bound that is obtained through the isolation of the hop between $\mathcal{S}$ and $\mathcal{R}_1$. The IO-relation of the subnetwork from $\mathcal{R}_1$ to $\mathcal{D}$ is replaced by an arbitrary linear map that is defined by the matrix $\mathbf{G}_r^{(1)} \in \mathbb{C}^{n \times n_{\mathcal{R}}^{(1)}}$. The IO-relation of the modified network is then given by

$$\mathbf{y}_{\mathcal{D}} = \mathbf{G}_r^{(1)} \cdot (\mathbf{H}_1\mathbf{x}_{\mathcal{S}} + \mathbf{w}_{\mathcal{R}_1}). \tag{2.52}$$

Thus, a single-hop network is obtained. It corresponds to a MIMO multiple access scenario with a linear receiver, where the n users are constrained to an average power

P/n each. A sketch of this network is depicted in Fig. 2.5 (a).

Proposition 1. *The optimal DMT curve of this network is achieved through receive zero-forcing and given in terms of the function* $x^+ \triangleq \max(0, x)$ *by (taking into account the pre-log* $1/p$*)*

$$\bar{d}_1(r) = \left(n_{\mathcal{R}}^{(1)} - n + 1\right) \cdot \left(1 - \frac{p \cdot r}{n}\right)^+ . \tag{2.53}$$

Proof. The proposition is proved in [59]. □

Case $2 \leq l \leq L$: This case corresponds to bounds that are obtained through the isolation of the hop between any two adjacent relay stages $\mathcal{R}_{l-1}$ and $\mathcal{R}_l$. The IO-relation of the subnetwork from $\mathcal{S}$ to $\mathcal{R}_{l-1}$ is replaced by a linear map that is determined by the matrix $\mathbf{G}_t^{(l-1)} \in \mathbb{C}^{n_{\mathcal{R}}^{(l-1)} \times n}$ and fulfills the sum-power constraint on $\mathcal{R}_{l-1}$. Likewise, the IO-relation of the subnetwork from $\mathcal{R}_l$ to $\mathcal{D}$ is replaced by a linear map that is determined by the matrix $\mathbf{G}_r^{(l)} \in \mathbb{C}^{n \times n_{\mathcal{R}}^{(l)}}$. The IO-relation of this modified network is given by

$$\mathbf{y}_{\mathcal{D}} = \mathbf{G}_r^{(l)} \cdot \left(\mathbf{H}_l \cdot \mathbf{G}_t^{(l-1)} \mathbf{x}_{\mathcal{S}} + \mathbf{w}_{\mathcal{R}_l}\right) . \tag{2.54}$$

Thus, a single-hop channel is obtained. It corresponds to a point-to-point MIMO link with joint linear transmit and receive beamforming and n parallel spatial streams. A sketch of this channel is depicted in Fig. 2.5 (b). Note that due to the fixed number of n spatial streams, the optimal DMT curve from [23] of the general MIMO channel is not achievable.

Proposition 2. *The optimal DMT curve of this channel is achieved by channel diagonalization via singular value decomposition, parallel transmission over the* n *strongest eigenmodes and equalization of the receive SNRs through power loading. It is given by*

$$\bar{d}_l(r) = \left(n_{\mathcal{R}}^{(l)} - n + 1\right) \cdot \left(n_{\mathcal{R}}^{(l-1)} - n + 1\right) \cdot \left(1 - \frac{p \cdot r}{n}\right)^+ . \tag{2.55}$$

Proof. From [60, Theorem 2], we know that the joint linear transmit-receive beamforming that maximizes the minimum SINR among n parallel symbol streams subject to an average transmit-power constraint, decouples the MIMO channel into its eigenmodes and applies an equal receive power policy to the n strongest eigenmodes. That is, the IO-relation (2.54) can be written in terms of the non-zero eigenvalues of $\mathbf{H}_l \mathbf{H}_l^H$,

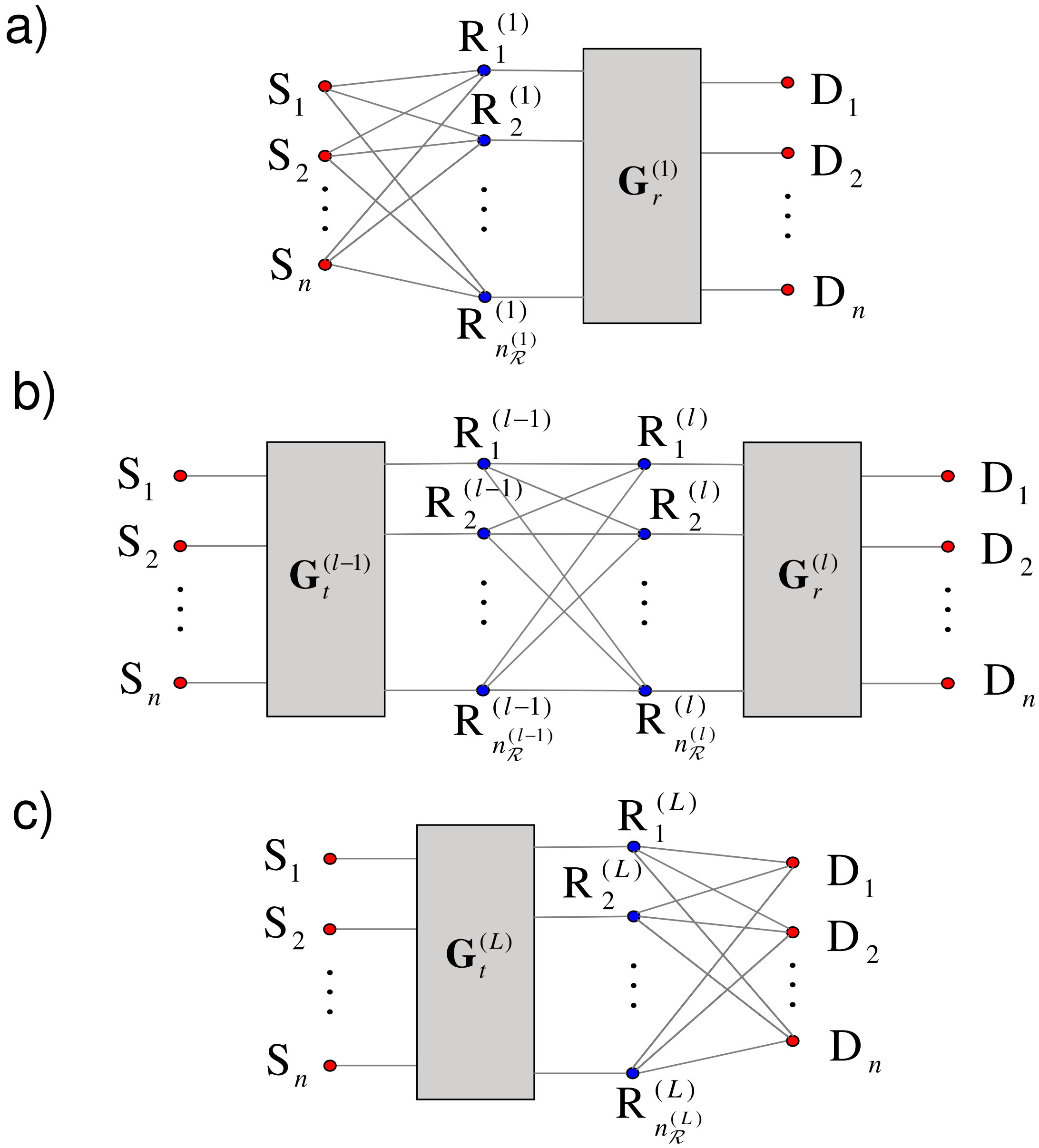


Figure 2.5.: Receive (a), joint transmit/receive (b), transmit (c) beamforming.

$\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_{\min(n_{\mathcal{R}}^{(l-1)}, n_{\mathcal{R}}^{(l)})}$, as

$$y_{\mathrm{D}_i} = \sqrt{\lambda_i} \cdot x_{\mathrm{S}_i} + \tilde{w}_{\mathrm{D}_i}, i \in \{1, \ldots, n\}, \tag{2.56}$$

where the noise vector $\tilde{\mathbf{w}}_{\mathcal{R}_l} = (\tilde{w}_{\mathrm{R}_1^{(l)}}, \ldots, \tilde{w}_{\mathrm{R}^{(l)}_{n_{\mathcal{R}}^{(l)}}})^T$ follows the same distribution as $\mathbf{w}_{\mathcal{R}_l}$, the power allocation fulfills

$$\sum_{i=1}^{n} P_{\mathrm{S}_i} = P \tag{2.57}$$

and

$$\mathsf{SINR}_1 = P_{\mathrm{S}_1} \cdot \lambda_1 = \ldots = \mathsf{SINR}_n = P_{\mathrm{S}_n} \cdot \lambda_n. \tag{2.58}$$

Thus, we have the following inequality for the average transmit power of the weakest subchannel:

$$\frac{P}{n} \leq P_{\mathrm{S}_n} \leq P. \tag{2.59}$$

We obtain the following lower- and upper-bounds on the outage probability (2.50) for a set of codes whose rates fulfill $R(P) = r/n \log P + \mathcal{O}(\log P^{1-\varepsilon})$ for arbitary $\varepsilon > 0$:

$$\begin{aligned} \mathsf{P}\left[\bigcup_{i=1}^{n} E_i \mid r\right] &= \mathsf{P}\left[\frac{1}{p}\log\left(1 + \frac{P_{\mathrm{S}_n}}{\sigma^2}\cdot\lambda_n\right) < \frac{r}{n}\log P + \mathcal{O}(\log P^{1-\varepsilon})\right] &(2.60)\\ &\geq \mathsf{P}\left[\frac{1}{p}\log\left(1 + \frac{P}{\sigma^2}\cdot\lambda_n\right) < \frac{r}{n}\log P + \mathcal{O}(\log P^{1-\varepsilon})\right] &(2.61)\\ &= \mathsf{P}\left[\lambda_n < \sigma^2 \cdot P^{-\left(1-\frac{p\cdot r}{n}\right)} + \mathcal{O}(P^{-\varepsilon})\right] &(2.62)\end{aligned}$$

$$\begin{aligned} \mathsf{P}\left[\bigcup_{i=1}^{n} E_i \mid r\right] &= \mathsf{P}\left[\frac{1}{p}\log\left(1 + \frac{P_{\mathrm{S}_n}}{\sigma^2}\cdot\lambda_n\right) < \frac{r}{n}\log P + \mathcal{O}(\log P^{1-\varepsilon})\right] &(2.63)\\ &\leq \mathsf{P}\left[\frac{1}{p}\log\left(1 + \frac{P}{n\sigma^2}\cdot\lambda_n\right) < \frac{r}{n}\log P + \mathcal{O}(\log P^{1-\varepsilon})\right] &(2.64)\\ &= \mathsf{P}\left[\lambda_n < n\sigma^2 \cdot P^{-\left(1-\frac{p\cdot r}{n}\right)} + \mathcal{O}(P^{-\varepsilon})\right]. &(2.65)\end{aligned}$$

For large P both upper- and lower-bound are of the form

$$\mathsf{P}\left[\lambda_n < c_{\text{up/low}} \cdot P^{-\left(1-\frac{p\cdot r}{n}\right)}\right]. \tag{2.66}$$

If $1-p\cdot r/n<0$, it is obvious, that (2.66) tends to one for large P, and thus $\bar{d}_l(r)=0$. In the following, we study the case $1-p\cdot r/n>0$. For notational convenience, we define $n_{\min}=\min\{n_{\mathcal{R}}^{(l-1)}, n_{\mathcal{R}}^{(l)}\}$. From [23, Eq. (15)][5], we know that for the nonnegative vector $(b_i)_{i=1}^{n}$, $b_i \in \mathbb{R}^+$,

$$\lim_{x\to\infty} -\frac{\mathsf{P}\left[\cap_{i=1}^{n_{\min}}\left\{\lambda_i < x^{-b_i}\right\}\right]}{\log x} \tag{2.67}$$

$$= \sum_{i=1}^{n_{\min}} \left(\left|n_{\mathcal{R}}^{(l-1)} - n_{\mathcal{R}}^{(l)}\right| - 2\cdot(n_{\min}-i+1)-1\right)\cdot b_i. \tag{2.68}$$

We define the sets $\mathcal{B}_1=\{1,\ldots,n-1\}$ and $\mathcal{B}_2=\{n,\ldots,n_{\min}\}$. We can now show that the upper- and lower-bound on the DMT curve of the considered channel are independent of $c_{\text{up/low}}$ and thus coincide:

$$\bar{d}_l(r) = \lim_{P\to\infty} -\frac{\log \mathsf{P}\left[\lambda_n < c_{\text{up/low}}\cdot P^{-(1-p\cdot r/n)}\right]}{\log P} \tag{2.69}$$

$$= \lim_{P\to\infty} -\frac{\log \mathsf{P}\left[\bigcap_{i\in\mathcal{B}_2}\left\{\lambda_i < c_{\text{up/low}}\cdot P^{-(1-p\cdot r/n)}\right\}\right]}{\log P} \tag{2.70}$$

$$= \lim_{P\to\infty} -\frac{\log \mathsf{P}\left[\bigcap_{i\in\mathcal{B}_2}\left\{\lambda_i < c_{\text{up/low}}\cdot P^{-(1-p\cdot r/n)}\right\}\right]}{\log P}$$

$$+ \lim_{P\to\infty} -\frac{\log \mathsf{P}\left[\bigcap_{i\in\mathcal{B}_1}\{\lambda_i<1\}\,\middle|\,\bigcap_{i\in\mathcal{B}_2}\left\{\lambda_i < c_{\text{up/low}}\cdot P^{-(1-p\cdot r/n)}\right\}\right]}{\log P} \tag{2.71}$$

$$= \lim_{P\to\infty} -\frac{\log \mathsf{P}\left[\left\{\bigcap_{i\in\mathcal{B}_1}\{\lambda_i<1\}\right\}\bigcap\left\{\bigcap_{i\in\mathcal{B}_2}\left\{\lambda_i < c_{\text{up/low}}\cdot P^{-(1-p\cdot r/n)}\right\}\right\}\right]}{\log P} \tag{2.72}$$

$$= \lim_{P\to\infty} -\frac{\log \mathsf{P}\left[\left\{\bigcap_{i\in\mathcal{B}_1}\{\lambda_i<1\}\right\}\bigcap\left\{\bigcap_{i\in\mathcal{B}_2}\left\{\lambda_i < P^{-(1-p\cdot r/n)}\right\}\right\}\right]}{\log P} \tag{2.73}$$

$$= \sum_{i=n}^{n_{\min}} \left(\left|n_{\mathcal{R}}^{(l-1)} - n_{\mathcal{R}}^{(l)}\right| + 2\cdot(n_{\min}-i+1)-1\right)\cdot(1-p\cdot r/n) \tag{2.74}$$

$$= (n_{\mathcal{R}}^{(l)}-n+1)\cdot(n_{\mathcal{R}}^{(l-1)}-n+1)\cdot(1-p\cdot r/n)\,. \tag{2.75}$$

Note that the added term in (2.71) is zero, since the probability therein is independent of P. In (2.74), we apply (2.68) with $b_i=(1-p\cdot r/n)$ for all $i\in\mathcal{B}_2$ and $b_i=0$ for all

[5]Note that the eigenvalues are ordered in reversed order in [23]. Therefore, (2.68) is modified accordingly.

$i \in \mathcal{B}_1$. Thus, the proof is complete. □

Case $l = L + 1$: This case corresponds to the bound that is obtained through the isolation of the hop between $\mathcal{R}_L$ and $\mathcal{D}$. The IO-relation of the subnetwork from $\mathcal{S}$ to $\mathcal{R}_L$ is replaced by an arbitrary linear map that is defined by the matrix $\mathbf{G}_t^{(L)} \in \mathbb{C}^{n_{\mathcal{R}}^{(L)} \times n}$ and fulfills the sum-power constraint on $\mathcal{R}_L$. The IO relation of this modified network is then given by

$$\mathbf{y}_{\mathcal{D}} = \mathbf{H}_{L+1} \cdot \mathbf{G}_t^{(L)} \mathbf{x}_{\mathcal{S}} + \mathbf{w}_{\mathcal{D}}. \tag{2.76}$$

Thus, a single-hop network is obtained. It corresponds to a MIMO broadcast scenario with a linear transmitter, where the n spatial streams are constrained to an average sum-power P. A sketch of this network is depicted in Fig. 2.5 (c).

Proposition 3. *The optimal DMT curve of this network is achieved through transmit zero-forcing and given by*

$$\bar{d}_{L+1}(r) = \left(n_{\mathcal{R}}^{(L)} - n + 1\right) \cdot \left(1 - \frac{p \cdot r}{n}\right)^{+}. \tag{2.77}$$

Proof. We know the following from the uplink-downlink duality [61]: Every set of SINRs $(\mathsf{SINR}_1, \ldots, \mathsf{SINR}_n)$ which is achievable in the dual multiple access (uplink) network with linear receiver is achievable in our average sum-power constrained broadcast (downlink) network with linear transmitter. In particular, every set $(\mathsf{SINR}_1, \ldots, \mathsf{SINR}_n)$ that is achievable through receive zero-forcing in a dual uplink network is achievable through transmit zero-forcing in our downlink network. Therefore, the achievability of the DMT curve (2.77) is an immediate consequence of the DMT curve for the dual uplink network with linear receiver [59] (cf. (2.53)) and the uplink-downlink duality.

Note that the [59] constrains each user to transmit with power P/n in the uplink network. Therefore, our downlink network with sum-transmit power P and arbitrary power-allocation could potentially achieve a better DMT curve, and we need a converse to establish the proposition. Again, we consider the dual uplink problem. We upper-bound the optimal DMT curve of a *sum-power constrained* dual uplink network with sum-power P by a per-user-power constrained uplink network with transmit power P per user. Now, we can repeat the original proof of [59] with the modified constraint of an n-fold transmit power. As P is taken to infinity, this modification has no impact on the result. □

In the following two sections, we attempt to come up with relay gain allocations that might achieve this upper-bound. Specifically, we study optimization of the SINR of the weakest source-destination pair in Section 2.6 and distributed zero-forcing with relay set selection in Section 2.6.3.

2.6. Max-Min Rate Optimization

This section is concerned with the following constrained optimization problem in the complex gain coefficients of the relay nodes:

$$\text{maximize} \min_{i\in\{1,\ldots,n\}} \mathsf{SINR}_i \tag{2.78}$$

$$\text{subject to } P_{\mathcal{R}_l} \leq \bar{P}_{\mathcal{R}_l} \text{ for all } l \in \{1,\ldots,L\}, \tag{2.79}$$

where SINR_i denotes the SINR of source-destination pair i. As in the previous sections, the transmit power of each relay stage $\mathcal{R}_l$, $\bar{P}_{\mathcal{R}_l}$, is constrained not to exceed a maximum sum-power $\bar{P}_{\mathcal{R}_l}$. Maximization of the weakest SINR corresponds to maximization of the rate $\bar{R}_{\min}$ that can be achieved by each of the n source-destination pairs. This rate is given by

$$\bar{R}_{\min} = \min_{i\in\{1,\ldots,n\}} \log_2\left(1+\mathsf{SINR}_i\right) = \log_2\left(1+\min_{i\in\{1,\ldots,n\}} \mathsf{SINR}_i\right). \tag{2.80}$$

This objective is motivated by the outage probability (2.50): The choice of an $\bar{R}_{\min}$ maximizing gain allocation ensures that an outage occurs only due to poor channel realizations and not due to suboptimal gain allocations. In this sense, the scheme is optimal with respect to outage probability, and, in particular, with respect to the DMT.

The following proposition states that the above optimization problem can be turned into an unconstrained problem in the complex vectors $\tilde{\mathbf{g}}_l$, $l \in \{1,\ldots,L\}$, which are related to the $\mathbf{g}_l$'s according to

$$\mathbf{g}_l = \sqrt{\frac{\bar{P}_{\mathcal{R}_l}}{\tilde{P}_{\mathcal{R}_l}}} \cdot \tilde{\mathbf{g}}_l. \tag{2.81}$$

Here, $\tilde{P}_{\mathcal{R}_l}$ denotes the transmit power of relay stage $\mathcal{R}_l$ that would result, if $\tilde{\mathbf{g}}_l$ was applied as the gain vector of stage $\mathcal{R}_l$. Thus, the $\tilde{\mathbf{g}}_l$'s inherently lead to $\mathbf{g}_l$'s that fulfill the power constraints with equality.

Proposition 4. *The constrained optimization problem in* $(\mathbf{g}_1, \ldots, \mathbf{g}_L)$ *can be turned into an unconstrained optimization problem in* $(\tilde{\mathbf{g}}_1, \ldots, \tilde{\mathbf{g}}_L)$ *that for any* $l \in \{1, \ldots, L\}$ *can be written as*

$$\text{maximize} \min_{i \in \{1,\ldots,n\}} \frac{\tilde{\mathbf{g}}_l^H \mathbf{A}_{l,i} \tilde{\mathbf{g}}_l}{\tilde{\mathbf{g}}_l^H \mathbf{B}_{l,i} \tilde{\mathbf{g}}_l}, \tag{2.82}$$

where for all $i \in \{1, \ldots, n\}$

$$\mathsf{SINR}_i = \frac{\tilde{\mathbf{g}}_l^H \mathbf{A}_{l,i} \tilde{\mathbf{g}}_l}{\tilde{\mathbf{g}}_l^H \mathbf{B}_{l,i} \tilde{\mathbf{g}}_l}, \tag{2.83}$$

and the matrices $\mathbf{A}_{l,i}$ *and* $\mathbf{B}_{l,i}$ *are functions of all vectors* $\{\tilde{\mathbf{g}}_1, \ldots, \tilde{\mathbf{g}}_{l-1}, \tilde{\mathbf{g}}_{l+1}, \ldots, \tilde{\mathbf{g}}_L\}$, *but not of* $\tilde{\mathbf{g}}_l$. *Moreover,* $\mathbf{A}_{l,i} = \mathbf{A}_{l,i}^H$, $\mathbf{A}_{l,i} \succeq \mathbf{0}$, $\mathbf{B}_{l,i} = \mathbf{B}_{l,i}^H$, $\mathbf{B}_{l,i} \succ \mathbf{0}$ *for all* l *and* i.

Proof. The proof of Proposition 4 is provided in Appendix A.1.

The problem in the complex vectors $(\tilde{\mathbf{g}}_1, \ldots, \tilde{\mathbf{g}}_L)$ is easily translated into a problem in the real vectors $(\mathbf{z}_1, \ldots, \mathbf{z}_L)$, where $\mathbf{z}_l \triangleq (\Re\{\tilde{\mathbf{g}}_l\}^T \ , \ \Im\{\mathbf{g}_l\}^T)^T$. To this end, the following real and symmetric matrices are introduced:

$$\bar{\mathbf{A}}_{l,i} \triangleq \begin{pmatrix} \Re\{\mathbf{A}_{l,i}\} & -\Im\{\mathbf{A}_{l,i}\} \\ \Im\{\mathbf{A}_{l,i}\} & \Re\{\mathbf{A}_{l,i}\} \end{pmatrix} \quad \text{and} \quad \bar{\mathbf{B}}_{l,i} \triangleq \begin{pmatrix} \Re\{\mathbf{B}_{l,i}\} & -\Im\{\mathbf{B}_{l,i}\} \\ \Im\{\mathbf{B}_{l,i}\} & \Re\{\mathbf{B}_{l,i}\} \end{pmatrix}.$$

With this notation, the SINRs are rewritten for any $l \in \{1, \ldots, L\}$ as

$$\mathsf{SINR}_i = \frac{\tilde{\mathbf{g}}_l^H \mathbf{A}_{l,i} \tilde{\mathbf{g}}_l}{\tilde{\mathbf{g}}_l^H \mathbf{B}_{l,i} \tilde{\mathbf{g}}_l} = \frac{\mathbf{z}_l^T \bar{\mathbf{A}}_{l,i} \mathbf{z}_l}{\mathbf{z}_l^T \bar{\mathbf{B}}_{l,i} \mathbf{z}_l}. \tag{2.84}$$

Assumption: Throughout the remainder of this section, it is assumed that the generalized Rayleigh quotients SINR_i, $i \in \{1, \ldots, n\}$, are expanded in a way such that

$$\mathbf{z}_1^T \bar{\mathbf{B}}_{1,i} \mathbf{z}_1 = \ldots = \mathbf{z}_L^T \bar{\mathbf{B}}_{L,i} \mathbf{z}_L. \tag{2.85}$$

This assumption is obviously without loss of generality.

Our optimization problem falls into the class of maximin (or minimax) problems. A comprehensive theory of minimax optimization is provided in [62]. For general network topologies the problem is a nonlinear programming problem. Necessary conditions for a local maximum of $\min_{i \in \{1,\ldots,L\}} \mathsf{SINR}_i$ are given as follows [62]. Let $\mathcal{A}$ (typically referred to as "active set") be defined as

$$\mathcal{A} \triangleq \left\{ i \in \{1,\ldots,n\} \;\middle|\; \mathsf{SINR}_i = \min_{j\in\{1,\ldots,L\}} \mathsf{SINR}_j \right\}. \tag{2.86}$$

If $\min_{i\in\{1,\ldots,L\}} \mathsf{SINR}_i$ has a local maximum in the point $(\mathbf{z}_1^{(0)},\ldots,\mathbf{z}_L^{(0)})$, there are $\gamma_i \geq 0$, $i \in \{1,\ldots,n\}$, such that

$$\sum_{i\in\mathcal{A}} \gamma_i \cdot \boldsymbol{\nabla}_{\mathbf{z}_1,\ldots,\mathbf{z}_L} \mathsf{SINR}_i \Big|_{(\mathbf{z}_1,\ldots,\mathbf{z}_L)=(\mathbf{z}_1^{(0)},\ldots,\mathbf{z}_L^{(0)})} = \mathbf{0} \qquad \text{and} \qquad \sum_{i\in\mathcal{A}} \gamma_i = 1. \tag{2.87}$$

These conditions can be interpreted geometrically: The zero vector must be contained in the convex hull of the set of gradients $\{\boldsymbol{\nabla}_{\mathbf{z}_1,\ldots,\mathbf{z}_L}\mathsf{SINR}_i \mid i \in \mathcal{A}\}$. They are straightforwardly derived as the Karush-Kuhn-Tucker (KKT) conditions [63,64] for the following constrained optimization problem in $(\mathbf{z}_1,\ldots,\mathbf{z}_L)$:

$$\text{maximize} \quad \lambda \tag{2.88}$$

$$\text{subject to } \mathsf{SINR}_i - \lambda \geq 0 \text{ for all } i \in \{1,\ldots,n\}. \tag{2.89}$$

Problem (2.89) is equivalent to problem (2.82), and the optimal λ fulfills $\lambda = \min_{j\in\{1,\ldots,n\}}\{\mathsf{SINR}_j\}$. This re-formulation is a common trick that allows for the elimination of the non-differentiable minimum-function. The Lagrangian of this problem is given by

$$\mathcal{L}(\lambda,\mathbf{z}_1,\ldots,\mathbf{z}_L,\gamma_1\ldots,\gamma_n) = \lambda + \sum_{i=1}^{n} \gamma_i \cdot (\mathsf{SINR}_i - \lambda)\,. \tag{2.90}$$

Thus, the KKT conditions are summarized as:

$$\sum_{i=1}^{n} \gamma_i \cdot \boldsymbol{\nabla}_{\mathbf{z}_1,\ldots,\mathbf{z}_L} \mathsf{SINR}_i = \mathbf{0}, \tag{2.91}$$

$$\sum_{i=1}^{n} \gamma_i = 1, \tag{2.92}$$

$$\mathsf{SINR}_i - \lambda \geq 0 \text{ for all } i \in \{1,\ldots,n\}, \tag{2.93}$$

$$\gamma_i \geq 0 \text{ for all } i \in \{1,\ldots,n\}, \tag{2.94}$$

$$\gamma_i \cdot (\mathsf{SINR}_i - \lambda) = 0 \text{ for all } i \in \{1,\ldots,n\}. \tag{2.95}$$

Since $\lambda = \min_{j\in\{1,\dots,n\}} \mathsf{SINR}_j$, condition (2.93) is redundant. Moreover, condition (2.95) implies that Lagrange multipliers that correspond to inactive constraints are zero, i.e., $\gamma_i = 0$ for all $i \notin \mathcal{A}$. Hence, the KKT conditions are equivalent to (2.87).

The gradient vectors for condition (2.87) are given for $i \in \{1,\dots,n\}$ by

$$\boldsymbol{\nabla}_{\mathbf{z}_1,\dots,\mathbf{z}_L}\mathsf{SINR}_i = \left(\boldsymbol{\nabla}^T_{\mathbf{z}_1}\mathsf{SINR}_i,\dots,\boldsymbol{\nabla}^T_{\mathbf{z}_L}\mathsf{SINR}_i\right)^T, \tag{2.96}$$

where

$$\boldsymbol{\nabla}_{\mathbf{z}_l}\mathsf{SINR}_i = \frac{2}{\mathbf{z}_l^T\bar{\mathbf{B}}_{l,i}\mathbf{z}_l}\cdot\left(\bar{\mathbf{A}}_{l,i}\mathbf{z}_l - \mathsf{SINR}_i\cdot\bar{\mathbf{B}}_{l,i}\mathbf{z}_l\right). \tag{2.97}$$

Since each multiplicative term $2/\mathbf{z}_l^T\bar{\mathbf{B}}_{l,i}\mathbf{z}_l$ corresponds just to a scaling of the ith gradient, condition (2.87) is equivalent to the existence of positive γ_i, $i \in \{1,\dots,n\}$, such that

$$\sum_{i\in\mathcal{A}}\gamma_i\cdot\left(\bar{\mathbf{A}}_{1,i}\mathbf{z}_1 - \mathsf{SINR}_i\cdot\bar{\mathbf{B}}_{1,i}\mathbf{z}_1\right) = \mathbf{0}\;\forall\, l\in\{1,\dots,L\} \quad \text{and} \quad \sum_{i\in\mathcal{A}}\gamma_i \;= 1. \tag{2.98}$$

Suppose there was a set of positive γ_i', $i \in \{1,\dots,n\}$, that fulfills (2.87). Then, (2.98) is fulfilled by the following positive γ_i, $i \in \{1,\dots,n\}$:

$$\gamma_i = \frac{1}{\sum_{j\in\mathcal{A}}\frac{2}{\mathbf{z}_l^T\bar{\mathbf{B}}_{l,j}\mathbf{z}_l}\cdot\gamma_j'}\cdot\frac{2}{\mathbf{z}_l^T\bar{\mathbf{B}}_{l,i}\mathbf{z}_l}\cdot\gamma_i', \tag{2.99}$$

where it is required, that $\mathbf{z}_1^T\bar{\mathbf{B}}_{1,i}\mathbf{z}_1 = \ldots = \mathbf{z}_L^T\bar{\mathbf{B}}_{L,i}\mathbf{z}_L$, in order to obtain the same set of γ_i, $i \in \mathcal{A}$, for all $l \in \{1,\dots,L\}$. This is guaranteed by assumption (2.85).

Note that only SINR_i that correspond to source-destination pairs in the active set are involved in (2.98). Thus, all of them fulfill $\mathsf{SINR}_i = \min_{j\in\mathcal{A}}\mathsf{SINR}_j = \lambda$ by definition. The conditions (2.98) can thus be rewritten as

$$\sum_{i\in\mathcal{A}}\gamma_i\cdot\left(\bar{\mathbf{A}}_{l,i}\mathbf{z}_l - \lambda\cdot\bar{\mathbf{B}}_{l,i}\mathbf{z}_l\right) = \mathbf{0} \text{ for all } l\in\{1,\dots,L\}, \tag{2.100}$$

$$\sum_{i\in\mathcal{A}}\gamma_i = 1, \tag{2.101}$$

$$\mathsf{SINR}_i = \lambda \text{ for all } i\in\mathcal{A}. \tag{2.102}$$

The focus of the next subsection is on two-hop networks, i.e., the case $L = 1$. In this case, we can turn the optimization problem (2.82) into a quasi-convex problem that can be solved by standard algorithms. Moreover, for the special case $L = 1$ and $n = 2$, a customized optimization method is proposed that allows for a nice geometric interpretation.

2.6.1. Two-Hop Networks ($L = 1$)

If the network is restricted to two hops, the necessary condition for a local maximum reduces to the existence of positive γ_i, $i \in \mathcal{A}$, such that

$$\sum_{i\in\mathcal{A}} \gamma_i \cdot \left(\bar{\mathbf{A}}_{1,i}\mathbf{z}_1 - \lambda \cdot \bar{\mathbf{B}}_{1,i}\mathbf{z}_1\right) = \mathbf{0}, \tag{2.103}$$

$$\sum_{i\in\mathcal{A}} \gamma_i = 1, \tag{2.104}$$

$$\mathsf{SINR}_i = \lambda \text{ for all } i \in \mathcal{A}, \tag{2.105}$$

where (2.103) is a generalized eigenvalue problem in the matrices $\sum_{i\in\mathcal{A}} \gamma_i \cdot \bar{\mathbf{A}}_{1,i}$ and $\sum_{i\in\mathcal{A}} \gamma_i \cdot \bar{\mathbf{B}}_{1,i}$ with eigenvalues $\lambda_j\{\sum_{i\in\mathcal{A}} \gamma_i \cdot \bar{\mathbf{A}}_{1,i}, \sum_{i\in\mathcal{A}} \gamma_i \cdot \bar{\mathbf{B}}_{1,i}\}$, $j \in \{1, \ldots, 2n_{\mathcal{R}}^{(1)}\}$, which are ordered in descending order.

Remark: A given set of γ_i, $i \in \mathcal{A}$, that fulfills (2.103)-(2.105) for one of the eigenvalues of $\sum_{i\in\mathcal{A}} \gamma_i \cdot \bar{\mathbf{A}}_{1,i}$ and $\sum_{i\in\mathcal{A}} \gamma_i \cdot \bar{\mathbf{B}}_{1,i}$, will not fulfill (2.105) for any of the remaining $2 \cdot n_{\mathcal{R}}^{(1)} - 1$ eigenvalues in general.

For two-hop networks, we do not attempt to solve equations (2.100)-(2.102) directly. Instead, we succeed in turning the optimization problem (2.82) into a quasi-convex problem. The key to this is the following chain of identities. For any $\mathbf{z} \in \mathbb{R}^{2\cdot n_{\mathcal{R}}^{(1)}}$, we have

$$\begin{aligned}
\min_{i\in\{1,\ldots,n\}} \mathsf{SINR}_i &= \min_{i\in\{1,\ldots,n\}} \frac{\mathbf{z}_1^T \bar{\mathbf{A}}_{1,i}\mathbf{z}_1}{\mathbf{z}_1^T \bar{\mathbf{B}}_{1,i}\mathbf{z}_1} && (2.106)\\
&= \min_{\substack{(\gamma_1,\ldots,\gamma_n)\in[0,\infty)^n \\ \sum_{i=1}^n \gamma_i > 0}} \frac{\sum_{i=1}^n \gamma_i \cdot \mathbf{z}_1^T \bar{\mathbf{A}}_{1,i}\mathbf{z}_1}{\sum_{i=1}^n \gamma_i \cdot \mathbf{z}_1^T \bar{\mathbf{B}}_{1,i}\mathbf{z}_1} && (2.107)\\
&\le \min_{\substack{(\gamma_1,\ldots,\gamma_n)\in[0,\infty)^n \\ \sum_{i=1}^n \gamma_i > 0}} \max_{\mathbf{z}_1\in\mathbb{R}^{2\cdot n_{\mathcal{R}}^{(1)}}} \frac{\mathbf{z}_1^T \left(\sum_{i=1}^n \gamma_i \cdot \bar{\mathbf{A}}_{1,i}\right)\mathbf{z}_1}{\mathbf{z}_1^T \left(\sum_{i=1}^n \gamma_i \cdot \bar{\mathbf{B}}_{1,i}\right)\mathbf{z}_1} && (2.108)\\
&= \min_{\substack{(\gamma_1,\ldots,\gamma_n)\in[0,\infty)^n \\ \sum_{i=1}^n \gamma_i > 0}} \lambda_1\left\{\bar{\mathbf{A}}_1(\boldsymbol{\gamma}), \bar{\mathbf{B}}_1(\boldsymbol{\gamma})\right\}, && (2.109)
\end{aligned}$$

where $\bar{\mathbf{A}}_1(\boldsymbol{\gamma}) \triangleq \sum_{i=1}^n \gamma_i \cdot \bar{\mathbf{A}}_{1,i}$, $\bar{\mathbf{B}}_1(\boldsymbol{\gamma}) \triangleq \sum_{i=1}^n \gamma_i \cdot \bar{\mathbf{B}}_{1,i}$ and $\boldsymbol{\gamma} \triangleq (\gamma_1, \ldots, \gamma_n)$. In order to understand the equality between (2.106) and (2.107), it is helpful to note the following:

(i) At the minimum of (2.107) all source-destination pairs with index $i, i \notin \mathcal{A}$, are assigned zero weights, i.e., $\gamma_i = 0$.

(ii) If $\gamma_i = 0$ for all $i \notin \mathcal{A}$, (2.107) is obviously equal to (2.106) and independent of the values $\{\gamma_i | i \in \mathcal{A}\}$.

Inequality (2.108) follows, since maximization of the generalized Rayleigh-quotient for every fixed $(\gamma_1, \ldots, \gamma_n)$ with respect to $\mathbf{z}_1$ cannot decrease the minimum value with respect to $(\gamma_1, \ldots, \gamma_n)$. Finally, (2.108) coincides with the largest generalized eigenvalue of the matrices $\bar{\mathbf{A}}_1(\boldsymbol{\gamma})$ and $\bar{\mathbf{B}}_1(\gamma)$, which yields (2.109). The above inequality holds with equality, if and only if $\mathbf{z}_1$ is the principal generalized eigenvector of $\bar{\mathbf{A}}_1(\boldsymbol{\gamma})$ and $\bar{\mathbf{B}}_1(\boldsymbol{\gamma})$. That is, the problem (2.82) is equivalent to the following problem in $\boldsymbol{\gamma}$:

$$\begin{aligned} \text{minimize} \quad & \lambda_1 \left\{ \bar{\mathbf{A}}_1(\boldsymbol{\gamma}), \bar{\mathbf{B}}_1(\boldsymbol{\gamma}) \right\} \\ \text{subject to} \quad & \sum_{i=1}^{n} \gamma_i > 0, \\ & \gamma_i \geq 0 \text{ for all } i \in \{1, \ldots, n\}. \end{aligned} \tag{2.110}$$

The largest generalized eigenvalue $\lambda_1\{\mathbf{A}, \mathbf{B}\}$ is quasi-convex in the symmetric matrices $\mathbf{A} = \mathbf{A}^T$ and $\mathbf{B} = \mathbf{B}^T$ [65]. Thus, problem (2.110) is a quasi-convex problem in standard form [66]. Accordingly, it is has a more desirable structure than the original problem (2.82). An efficient algorithm for this problem has been proposed in [65]. It is based on the barrier method and accepts any problem of the form

$$\begin{aligned} \text{minimize} \quad & \lambda_1 \left\{ \mathbf{M}(\boldsymbol{\gamma}), \mathbf{N}(\boldsymbol{\gamma}) \right\} \\ \text{subject to} \quad & \mathbf{N}(\boldsymbol{\gamma}) \succ \mathbf{0}_{2n_{\mathcal{R}}^{(1)} \times 2n_{\mathcal{R}}^{(1)}}, \\ & \mathbf{P}(\boldsymbol{\gamma}) \succ \mathbf{0}_{(n+2)\times(n+2)}, \end{aligned} \tag{2.111}$$

where the matrices depend linearly on $\boldsymbol{\gamma} \in \mathbb{R}^n$ as follows:

$$\mathbf{M}(\boldsymbol{\gamma}) = \sum_{m=i}^{n} \gamma_i \cdot \mathbf{M}_i, \qquad \mathbf{M}_i = \mathbf{M}_i^T \text{ for all } i \in \{1, \ldots, n\}, \tag{2.112}$$

$$\mathbf{N}(\boldsymbol{\gamma}) = \sum_{m=i}^{n} \gamma_i \cdot \mathbf{N}_i, \qquad \mathbf{N}_i = \mathbf{N}_i^T \text{ for all } i \in \{1, \ldots, n\}, \tag{2.113}$$

$$\mathbf{P}(\boldsymbol{\gamma}) = \mathbf{P}_0 + \sum_{m=i}^{n} \gamma_i \cdot \mathbf{P}_i, \qquad \mathbf{P}_i = \mathbf{P}_i^T \text{ for all } i \in \{1, \ldots, n\}. \tag{2.114}$$

In the sequel, the problem (2.110) is fit into this framework. For reasons that will become evident, the constraint $\sum_{i=1}^{n} \gamma_i > 0$ in (2.110) is replaced by the two constraints

$$\sum_{i=1}^{n} \gamma_i < 1.1 \qquad \text{and} \qquad \sum_{i=1}^{n} \gamma_i > 0.9. \tag{2.115}$$

This modification does not change the optimization problem, since the co-domain of the generalized eigenvalue in (2.110) would be unchanged, even if $\sum_{i=1}^{n} \gamma_i$ was constrained to a fixed positive value. Next, the following matrices are associated with each other:

$$\mathbf{M}(\boldsymbol{\gamma}) = \bar{\mathbf{A}}(\boldsymbol{\gamma}) \qquad \text{and} \qquad \mathbf{M}_i = \bar{\mathbf{A}}_{1,i}, i \in \{1, \ldots, n\}, \tag{2.116}$$

$$\mathbf{N}(\boldsymbol{\gamma}) = \bar{\mathbf{B}}(\boldsymbol{\gamma}) \qquad \text{and} \qquad \mathbf{N}_i = \bar{\mathbf{B}}_{1,i}, i \in \{1, \ldots, n\}. \tag{2.117}$$

The matrices $\mathbf{P}_i$, $i \in \{0, 1, \ldots, n\}$, capture the constraints and are defined as follows:

$$p_{o_{\mu\nu}} = \begin{cases} -1.1, \text{ if } \mu = \nu = n+1, \\ 0.9, \text{ if } \mu = \nu = n+2, \\ 0, \text{ else.} \end{cases} \qquad \text{for } i = 0, \tag{2.118}$$

$$p_{i_{\mu\nu}} = \begin{cases} 1, \text{ if } \mu = \nu = i, \\ -1, \text{ if } \mu = \nu = n+1, \\ 1, \text{ if } \mu = \nu = n+2, \\ 0, \text{ else,} \end{cases} \qquad \text{for } i \in \{1, \ldots, n\}. \tag{2.119}$$

Note that the constraints $\gamma_i \geq 0$ are replaced by $\gamma_i > 0$ here. In the case that $|\mathcal{A}| < n$ at the optimum $\boldsymbol{\gamma}$, the optimum thus corresponds to the infimum of the objective function in the feasibility region.

The algorithm of [65] for the above problem (2.111) is subject to the following assumptions:

1. $\mathbf{N}(\boldsymbol{\gamma})$ is bounded away from singularity on the feasible set, i.e., there is $c > 0$, such that $\mathbf{P}(\boldsymbol{\gamma}) \succ \mathbf{0}_{(n+2)\times(n+2)}$ implies $\mathbf{N}(\boldsymbol{\gamma}) \succeq c \cdot \mathbf{I}_{2n_{\mathcal{R}}^{(1)}}$.
2. The feasible set is bounded, i.e., there is some c such that $\mathbf{P}(\boldsymbol{\gamma}) \succ \mathbf{0}_{(n+2)\times(n+2)}$ implies $\|\boldsymbol{\gamma}\|_2 < c$.
3. There is a pair $(\lambda^{(0)}, \boldsymbol{\gamma}^{(0)})$, such that $\mathbf{N}(\boldsymbol{\gamma}^{(0)}) \succ \mathbf{0}_{2n_{\mathcal{R}}^{(1)}\times 2n_{\mathcal{R}}^{(1)}}$, $\mathbf{P}(\boldsymbol{\gamma}^{(0)}) \succ \mathbf{0}_{(n+2)\times(n+2)}$ and $\lambda^{(0)}\mathbf{N}(\boldsymbol{\gamma}^{(0)}) - \mathbf{M}(\boldsymbol{\gamma}^{(0)}) \succ \mathbf{0}_{2n_{\mathcal{R}}^{(1)}\times 2n_{\mathcal{R}}^{(1)}}$.

For the problem at hand these assumptions are verified as follows:

1. Due to the constraint $\sum_{i=1}^{n} \gamma_i > 0.9$ and the positive-definiteness of all $\bar{\mathbf{B}}_{1,i}$, $i \in \{1, \ldots, n\}$, the following holds for the smallest eigenvalue:

$$\lambda_{2n_{\mathcal{R}}^{(1)}} \left\{ \sum_{i=1}^{n} \gamma_i \bar{\mathbf{B}}_{1,i} \right\} \geq 0.9 \cdot \min_{i \in \{1,\ldots,n\}} \lambda_{2n_{\mathcal{R}}^{(1)}} \left\{ \bar{\mathbf{B}}_{1,i} \right\} > 0.$$

   We conclude that $\det(\bar{\mathbf{B}}_1(\boldsymbol{\gamma}) - c \cdot \mathbf{I}_{2n_{\mathcal{R}}^{(1)}}) \geq 0$ and thus $\bar{\mathbf{B}}_1(\boldsymbol{\gamma}) \succeq c \cdot \mathbf{I}_{2n_{\mathcal{R}}^{(1)}}$ for all $c < 0.9 \cdot \min_{i \in \{1,\ldots,n\}} \lambda_{2n_{\mathcal{R}}^{(1)}} \left\{ \bar{\mathbf{B}}_{1,i} \right\}$.

2. The feasible set is obviously bounded, since $0 < \gamma_i < 1.1$ for all $i \in \{1, \ldots, n\}$. Here, the upper-bounds are due to the constraint $\sum_{i=1}^{n} \gamma_i < 1.1$.

3. By definition, $\mathbf{P}(\boldsymbol{\gamma}^{(0)}) \succ \mathbf{0}_{(n+2)\times(n+2)}$ for every feasible $\boldsymbol{\gamma}^{(0)}$. We know that $\bar{\mathbf{B}}_1(\boldsymbol{\gamma}^{(0)}) \succ \mathbf{0}_{2n_{\mathcal{R}}^{(1)} \times 2n_{\mathcal{R}}^{(1)}}$, and thus that $\lambda_{2n_{\mathcal{R}}^{(1)}}\{\bar{\mathbf{B}}_1(\boldsymbol{\gamma}^{(0)})\} > 0$ for all $\boldsymbol{\gamma}^{(0)}$. Since $\lambda_1 \left\{ \bar{\mathbf{A}}_1(\boldsymbol{\gamma}) \right\}$ is finite for every feasible $\boldsymbol{\gamma}^{(0)}$, any choice $\lambda^{(0)} > \lambda_1\{\bar{\mathbf{A}}_1(\boldsymbol{\gamma}^{(0)})\}/\lambda_{2n_{\mathcal{R}}^{(1)}}\{\bar{\mathbf{B}}_1(\boldsymbol{\gamma}^{(0)})\}$ fulfills $\lambda^{(0)} \cdot \bar{\mathbf{B}}_1(\boldsymbol{\gamma}^{(0)}) - \bar{\mathbf{A}}_1(\boldsymbol{\gamma}^{(0)}) \succ \mathbf{0}_{2n_{\mathcal{R}}^{(1)} \times 2n_{\mathcal{R}}^{(1)}}$.

Next, the special case $n = 2$ is considered. We device a particularly simple and insightful optimization method for this case.

Case $n = 2$: Consider the necessary conditions (2.103) - (2.105). For $n = 2$, the active set $\mathcal{A}$ at the global optimum is either $\{1\}$, $\{2\}$ or $\{1, 2\}$. If $\mathcal{A} = \{1\}$ at the global optimum, the necessary conditions (2.103) - (2.105) imply that the optimum is given by the largest generalized eigenvalue $\lambda_1\{\bar{\mathbf{A}}_{1,1}, \bar{\mathbf{B}}_{1,1}\}$, and the optimum $\mathbf{z}_1$ is given by the corresponding principal eigenvector. We thus have

$$\max_{\mathbf{z}_1 \in \mathbb{R}^{2 \cdot n_{\mathcal{R}}^{(1)}}} \min_{i \in \{1,2\}} \mathsf{SINR}_i = \max_{\mathbf{z}_1 \in \mathbb{R}^{2 \cdot n_{\mathcal{R}}^{(1)}}} \mathsf{SINR}_1 < \mathsf{SINR}_2 \big|_{\mathbf{z}_1 = \operatorname{argmax} \mathsf{SINR}_1}. \tag{2.120}$$

Likewise, if $\mathcal{A} = \{2\}$ at the global optimum, the necessary conditions (2.103) - (2.105) imply that the optimum is given by the largest generalized eigenvalue $\lambda_1\{\bar{\mathbf{A}}_{1,2}, \bar{\mathbf{B}}_{1,2}\}$, and the optimum $\mathbf{z}_1$ is given by the corresponding principal eigenvector. We thus have

$$\max_{\mathbf{z}_1 \in \mathbb{R}^{2 \cdot n_{\mathcal{R}}^{(1)}}} \min_{i \in \{1,2\}} \mathsf{SINR}_i = \max_{\mathbf{z}_1 \in \mathbb{R}^{2 \cdot n_{\mathcal{R}}^{(1)}}} \mathsf{SINR}_2 < \mathsf{SINR}_1 \big|_{\mathbf{z}_1 = \operatorname{argmax} \mathsf{SINR}_2}. \tag{2.121}$$

Lastly, if at the optimum $\mathcal{A} = \{1, 2\}$, the necessary conditions (2.103) - (2.105) simplify to the existence of $\gamma_1 \in [0, 1]$, such that

$$\left((1 - \gamma_1)\bar{\mathbf{A}}_{1,1} + \gamma_1 \bar{\mathbf{A}}_{1,2}\right) \mathbf{z}_1 = \lambda \cdot \left((1 - \gamma_1)\bar{\mathbf{B}}_{1,1} + \gamma_1 \bar{\mathbf{B}}_{1,2}\right) \mathbf{z}_1, \tag{2.122}$$

$$\lambda = \mathsf{SINR}_1 = \mathsf{SINR}_2. \tag{2.123}$$

From (2.106) - (2.109), we know that λ is given by the generalized eigenvalue $\lambda_1\{(1-\gamma_1)\bar{\mathbf{A}}_{1,1} + \gamma_1\bar{\mathbf{A}}_{1,2}, (1-\gamma_1)\bar{\mathbf{B}}_{1,1} + \gamma_1\bar{\mathbf{B}}_{1,2}\}$ and $\mathbf{z}_1$ by the corresponding principal eigenvector $\mathbf{v}_1(\gamma_1)$ in the global optimum. We fix $\mathbf{z}_1 = \mathbf{v}_1(\gamma_1)$ in the following. Then, a $\gamma_1 \in [0,1]$ that fulfills $\mathsf{SINR}_1 - \mathsf{SINR}_2 = 0$ must exist. This follows due to the intermediate value theorem, since

$$\mathsf{SINR}_1|_{\mathbf{z}_1=\mathbf{v}_1(0)} - \mathsf{SINR}_2|_{\mathbf{z}_1=\mathbf{v}_1(0)} = \max_{\mathbf{z}_1 \in \mathbb{R}^{2\cdot n_{\mathcal{R}}^{(1)}}} \mathsf{SINR}_1 - \mathsf{SINR}_2|_{\mathbf{z}_1=\mathbf{v}_1(0)} \geq 0, \tag{2.124}$$

$$\mathsf{SINR}_1|_{\mathbf{z}_1=\mathbf{v}_1(1)} - \mathsf{SINR}_2|_{\mathbf{z}_1=\mathbf{v}_1(1)} = \mathsf{SINR}_1|_{\mathbf{z}_1=\mathbf{v}_1(1)} - \max_{\mathbf{z}_1 \in \mathbb{R}^{2\cdot n_{\mathcal{R}}^{(1)}}} \mathsf{SINR}_2 \leq 0. \tag{2.125}$$

If one of these two inequalities did not hold, the active set at the global optimum could not be given by $\mathcal{A} = \{1,2\}$ (cf. (2.120) and (2.121)). The problem is thus reduced to a root finding problem on the interval $[0,1]$. It remains to show that there exists exactly one root on this interval. Let $\gamma_1^{(0)}$ be a root. There exists a unique root, if for any $\delta > 0$ the inequalities $\mathsf{SINR}_2|_{\mathbf{z}_1=\mathbf{v}_1(\gamma_1^{(0)}+\delta)} > \mathsf{SINR}_1|_{\mathbf{z}_1=\mathbf{v}_1(\gamma_1^{(0)}+\delta)}$ and $\mathsf{SINR}_2|_{\mathbf{z}_1=\mathbf{v}_1(\gamma_1^{(0)}-\delta)} < \mathsf{SINR}_1|_{\mathbf{z}_1=\mathbf{v}_1(\gamma_1^{(0)}-\delta)}$ hold, i.e.,

$$\left.\frac{\mathbf{z}_1^T\bar{\mathbf{A}}_{1,2}\mathbf{z}_1}{\mathbf{z}_1^T\bar{\mathbf{B}}_{1,2}\mathbf{z}_1}\right|_{\mathbf{z}_1=\mathbf{v}_1(\gamma_1^{(0)}+\delta)} > \left.\frac{\mathbf{z}_1^T\bar{\mathbf{A}}_{1,1}\mathbf{z}_1}{\mathbf{z}_1^T\bar{\mathbf{B}}_{1,1}\mathbf{z}_1}\right|_{\mathbf{z}_1=\mathbf{v}_1(\gamma_1^{(0)}+\delta)}, \tag{2.126}$$

$$\left.\frac{\mathbf{z}_1^T\bar{\mathbf{A}}_{1,2}\mathbf{z}_1}{\mathbf{z}_1^T\bar{\mathbf{B}}_{1,2}\mathbf{z}_1}\right|_{\mathbf{z}_1=\mathbf{v}_1(\gamma_1^{(0)}-\delta)} < \left.\frac{\mathbf{z}_1^T\bar{\mathbf{A}}_{1,1}\mathbf{z}_1}{\mathbf{z}_1^T\bar{\mathbf{B}}_{1,1}\mathbf{z}_1}\right|_{\mathbf{z}_1=\mathbf{v}_1(\gamma_1^{(0)}-\delta)}. \tag{2.127}$$

Eq. (2.126) — and (2.127) analogously — is proved through contradiction:

$$\max_{\mathbf{z}_1 \in \mathbb{R}^{2\cdot n_{\mathcal{R}}^{(1)}}} \frac{\mathbf{z}_1^T\left((1-\gamma_1^{(0)}-\delta)\cdot\bar{\mathbf{A}}_{1,1} + (\gamma_1^{(0)}+\delta)\cdot\bar{\mathbf{A}}_{1,2}\right)\mathbf{z}_1}{\mathbf{z}_1^T\left((1-\gamma_1^{(0)}-\delta)\cdot\bar{\mathbf{B}}_{1,1} + (\gamma_1^{(0)}+\delta)\cdot\bar{\mathbf{B}}_{1,2}\right)\mathbf{z}_1} \tag{2.128}$$

$$= \left.\frac{\mathbf{z}_1^T\left((1-\gamma_1^{(0)}-\delta)\cdot\bar{\mathbf{A}}_{1,1} + (\gamma_1^{(0)}+\delta)\cdot\bar{\mathbf{A}}_{1,2}\right)\mathbf{z}_1}{\mathbf{z}_1^T\left((1-\gamma_1^{(0)}-\delta)\cdot\bar{\mathbf{B}}_{1,1} + (\gamma_1^{(0)}+\delta)\cdot\bar{\mathbf{B}}_{1,2}\right)\mathbf{z}_1}\right|_{\mathbf{z}_1=\mathbf{v}_1(\gamma_1^{(0)}+\delta)} \tag{2.129}$$

$$\leq \left.\frac{\mathbf{z}_1^T\left((1-\gamma_1^{(0)})\cdot\bar{\mathbf{A}}_{1,1} + \gamma_1^{(0)}\cdot\bar{\mathbf{A}}_{1,2}\right)\mathbf{z}_1}{\mathbf{z}_1^T\left((1-\gamma_1^{(0)})\cdot\bar{\mathbf{B}}_{1,1} + \gamma_1^{(0)}\cdot\bar{\mathbf{B}}_{1,2}\right)\mathbf{z}_1}\right|_{\mathbf{z}_1=\mathbf{v}_1(\gamma_1^{(0)}+\delta)} \tag{2.130}$$

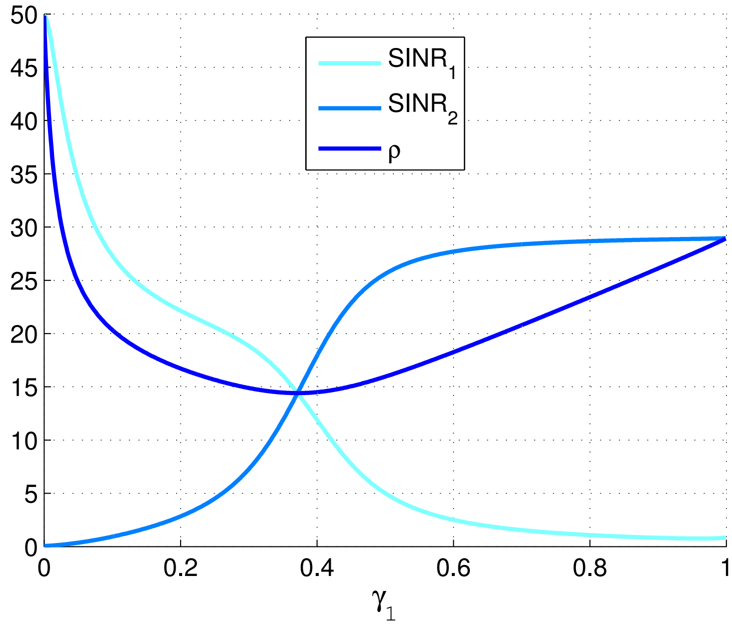


Figure 2.6.: The quantities SINR_1, SINR_2 and ρ are plotted versus $\gamma_1 \in [0,1]$ for a sample channel with $\mathcal{A} = \{1,2\}$ at the optimum.

$$< \left. \frac{\mathbf{z}_1^T \left((1-\gamma_1^{(0)}) \cdot \bar{\mathbf{A}}_{1,1} + \gamma_1^{(0)} \cdot \bar{\mathbf{A}}_{1,2} \right) \mathbf{z}_1}{\mathbf{z}_1^T \left((1-\gamma_1^{(0)}) \cdot \bar{\mathbf{B}}_{1,1} + \gamma_1^{(0)} \cdot \bar{\mathbf{B}}_{1,2} \right) \mathbf{z}_1} \right|_{\mathbf{z}_1 = \mathbf{v}_1(\gamma_1^{(0)})} \tag{2.131}$$

$$= \left. \frac{\mathbf{z}_1^T \left((1-\gamma_1^{(0)} - \delta) \cdot \bar{\mathbf{A}}_{1,1} + (\gamma_1^{(0)} + \delta) \cdot \bar{\mathbf{A}}_{1,2} \right) \mathbf{z}_1}{\mathbf{z}_1^T \left((1-\gamma_1^{(0)} - \delta) \cdot \bar{\mathbf{B}}_{1,1} + (\gamma_1^{(0)} + \delta) \cdot \bar{\mathbf{B}}_{1,2} \right) \mathbf{z}_1} \right|_{\mathbf{z}_1 = \mathbf{v}_1(\gamma_1^{(0)})} . \tag{2.132}$$

Inequality (2.130) holds under the hypothesis

$$\left. \frac{\mathbf{z}_1^T \bar{\mathbf{A}}_{1,2} \mathbf{z}_1}{\mathbf{z}_1^T \bar{\mathbf{B}}_{1,2} \mathbf{z}_1} \right|_{\mathbf{z}_1 = \mathbf{v}_1(\gamma_1^{(0)} + \delta)} \geq \left. \frac{\mathbf{z}_1^T \bar{\mathbf{A}}_{1,1} \mathbf{z}_1}{\mathbf{z}_1^T \bar{\mathbf{B}}_{1,1} \mathbf{z}_1} \right|_{\mathbf{z}_1 = \mathbf{v}_1(\gamma_1^{(0)} + \delta)} , \tag{2.133}$$

which is to be contradicted. Finally, (2.132) follows, since for equal generalized Rayleigh quotients $\mathbf{z}_1^T \bar{\mathbf{A}}_{1,1} \mathbf{z}_1 / \mathbf{z}_1^T \bar{\mathbf{B}}_{1,1} \mathbf{z}_1$ and $\mathbf{z}_1^T \bar{\mathbf{A}}_{1,2} \mathbf{z}_1 / \mathbf{z}_1^T \bar{\mathbf{B}}_{1,2} \mathbf{z}_1$, the expression is independent of δ. A comparison of (2.128) and (2.132) reveals the contradiction.

An illustration of the SINRs SINR_1 and SINR_2 as they evolve for varying $\gamma_1 \in [0,1]$ is provided in Fig. 2.6. Indeed, there is a unique intercept point of the curves for SINR_1 and SINR_2. Moreover, we have plotted the quantity

$$\rho \triangleq \lambda_1 \{ (1-\gamma_1) \bar{\mathbf{A}}_{1,1} + \gamma_1 \bar{\mathbf{A}}_{1,2}, (1-\gamma_1) \bar{\mathbf{B}}_{1,1} + \gamma_1 \bar{\mathbf{B}}_{1,2} \}.$$

This quantity corresponds to the upper-bound on the smaller of the two SINRs in (2.109). As it was predicted, the upper-bound is met at the intercept point of the SINR curves. The gap to this upper-bound can nicely be used for the following stopping criterion in any iterative optimization method. For a given tolerance $\varepsilon > 0$, terminate an algorithm, if

$$\lambda_1 \left\{ \sum_{i\in\mathcal{A}} \gamma_i \cdot \bar{\mathbf{A}}_{1,i}, \sum_{i\in\mathcal{A}} \gamma_i \cdot \bar{\mathbf{B}}_{1,i} \right\} - \min_{i\in\{1,\ldots,n\}} \mathsf{SINR}_i < \varepsilon, \tag{2.134}$$

that is, in our case if

$$\rho - \min\{\mathsf{SINR}_1, \mathsf{SINR}_2\} < \varepsilon. \tag{2.135}$$

In Algorithm 1, we solve the root finding problem in γ_1 by means of the method of bisection [66]. Thus, we obtain a method for the solution of the optimization problem (2.82) that provably converges to the global optimum of the problem. Bisection is well known to converge linearly [66].

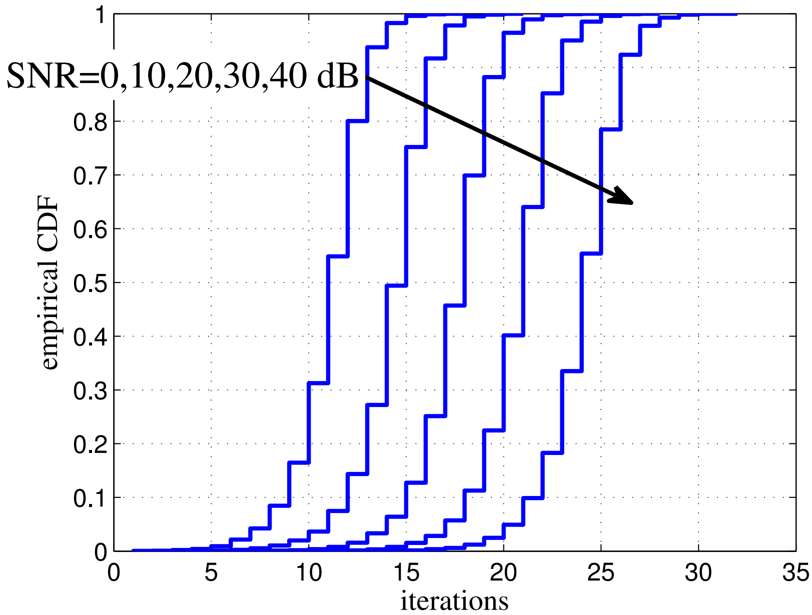
In order to obtain a rough impression of the number of iterations that is typically required by Algorithm 1, we consider Fig. 2.7. Two sets of empirical CDFs of iteration numbers for $\varepsilon = 10^{-3}$ are shown. They are obtained from 10^4 sample channels each. Realizations of the fading coefficients h_{YX} are generated independently from a CSCG random variable of unit variance. Moreover, the per-stage transmit powers are fixed to $P_{\mathcal{S}} = P_{\mathcal{R}_1} = \ldots = P_{\mathcal{R}_L} = P$. In the upper plot, a network with $n_{\mathcal{R}}^{(1)} = 10$ is considered for different $\mathrm{SNR} = P/\sigma^2$. It is observed that the number of iterations increases for increasing SNRs, which is not surprising. In the lower plot, networks are considered for different $n_{\mathcal{R}}^{(1)}$ and fixed SNR $P/\sigma^2 = 10$. It is observed that the number of iterations increases for increasing numbers of relay nodes. Also this effect is expected, since more relay nodes typically result into larger SINRs. Generally, the number of iterations is very moderate. Still, one should keep in mind that each iteration requires the computation of a principal generalized eigenvalue/eigenvector pair.

2.6.2. General Multihop Networks

In this subsection, general networks with an arbitrary number of relay stages, L, are considered. We do not have a method to turn the optimization problem (2.82) into a quasi-convex problem for $L > 1$. Therefore, we focus on iterative method for finding

Algorithm 1 Bisection-based optimization algorithm for special case $L = 1$ and $n = 2$.

```
given l = 0, u = 1,  tolerance ε > 0.
if (2.120) holds, i.e., A = {1} then
  z_1 = v_1(0).
  return z_1.
else if (2.121) holds, i.e., A = {2} then
  z_1 = v_1(1).
  return z_1.
else
  loop
    γ_1 ← (l + u)/2.
    ρ ← λ_1{(1 − γ_1)Ā_{1,1} + γ_1 Ā_{1,2}, (1 − γ_1)B̄_{1,1} + γ_1 B̄_{1,2}} .
    z_1 ← v_1(γ_1).
    Δ ← SINR_1|_{z_1} − SINR_2|_{z_1}.
    if ρ − min{SINR_1, SINR_2} < ε then
      break;
    else if Δ > 0 then
      l ← γ_1;
    else
      u ← γ_1.
    end if
  end loop
  return z_1.
end if
```



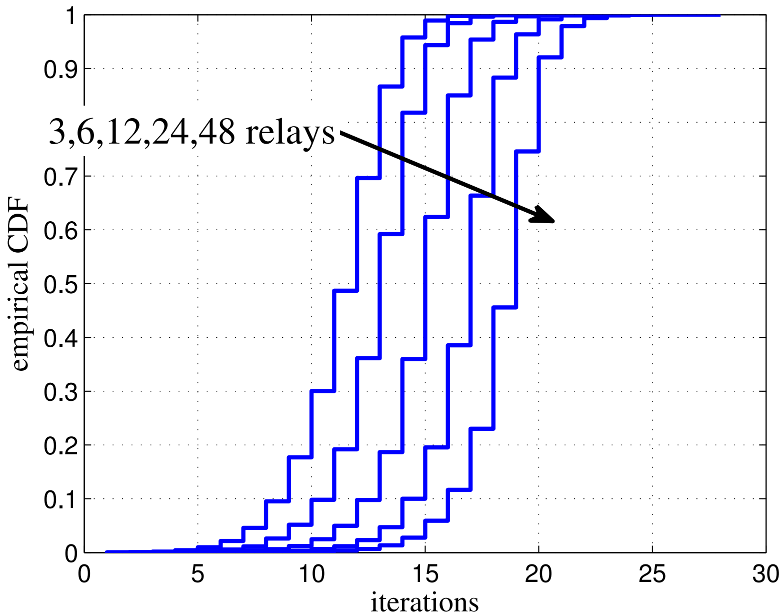


Figure 2.7.: Empirical CDF of iterations that are required by Algorithm 1 for $\varepsilon = 10^{-3}$. 10^4 sample channels are considered each. For each channel, realizations of the fading coefficients h_{YX} are generated independently from a CSCG random variable of unit variance. In the upper plot, a network with $n_{\mathcal{R}}^{(1)} = 10$ is considered for different SNR P/σ^2. In the lower plot, networks with different $n_{\mathcal{R}}^{(1)}$ are considered for a fixed SNR $P/\sigma^2 = 10$.

local optima in the following. An application of the developed method is, for instance, the tracking a global optimum in time-variant channels. As a basis for the applied optimization methods, we compute gradients and Hessians of the n source-destination pair SINRs. To this end, we recall from (2.84) that the ith source-destination pair SINR is a generalized Rayleigh quotient in every $\mathbf{z}_l$, $l \in \{1, \dots, L\}$:

$$\mathsf{SINR}_i = \frac{\mathbf{z}_l^T \bar{\mathbf{A}}_{l,i} \mathbf{z}_l}{\mathbf{z}_l^T \bar{\mathbf{B}}_{l,i} \mathbf{z}_l}. \tag{2.136}$$

The gradient of the ith source-destination pair SINR, $i \in \{1, \dots, n\}$, is thus given by

$$\boldsymbol{\nabla}_{\mathbf{z}_1,\dots,\mathbf{z}_L} \mathsf{SINR}_i = \begin{pmatrix} (\boldsymbol{\nabla}_{\mathbf{z}_1} \mathsf{SINR}_i)^T & \cdots & (\boldsymbol{\nabla}_{\mathbf{z}_L} \mathsf{SINR}_i)^T \end{pmatrix}^T, \tag{2.137}$$

where

$$\boldsymbol{\nabla}_{\mathbf{z}_l} \mathsf{SINR}_i = \frac{2}{\mathbf{z}_l^T \bar{\mathbf{B}}_{l,i} \mathbf{z}_l} \cdot \left(\bar{\mathbf{A}}_{l,i} \mathbf{z}_l - \mathsf{SINR}_i \cdot \bar{\mathbf{B}}_{l,i} \mathbf{z}_l \right). \tag{2.138}$$

The Hessians are constructed from the matrices $\boldsymbol{\nabla}^2_{\mathbf{z}_{l_1}\mathbf{z}_{l_2}} \mathsf{SINR}_i$, $(l_1, l_2) \in \{1, \dots, L\}^2$, as the block matrix

$$\boldsymbol{\nabla}^2_{\mathbf{z}} \mathsf{SINR}_i = \left(\boldsymbol{\nabla}^2_{\mathbf{z}_{l_1}\mathbf{z}_{l_2}} \mathsf{SINR}_i \right)_{l_1=1,\dots,L; l_2=1,\dots,L}, \tag{2.139}$$

where $\boldsymbol{\nabla}^2_{\mathbf{z}_{l_1}\mathbf{z}_{l_2}} \mathsf{SINR}_i$ is the matrix in the l_1th "block row" and l_2th "block column". The expression for the $\boldsymbol{\nabla}^2_{\mathbf{z}_{l_1}\mathbf{z}_{l_2}} \mathsf{SINR}_i$ on the block-diagonal, i.e., for $l_1 = l_2 = l$, is readily obtained as the Hessian of the generalized Rayleigh quotient:

$$\boldsymbol{\nabla}^2_{\mathbf{z}_l\mathbf{z}_l} \mathsf{SINR}_i = \frac{2}{\mathbf{z}_l^T \bar{\mathbf{B}}_{l,i} \mathbf{z}_l} \left(\bar{\mathbf{A}}_{l,i} - \mathsf{SINR}_i \cdot \bar{\mathbf{B}}_{l,i} - (\boldsymbol{\nabla}_{\mathbf{z}_l} \mathsf{SINR}_i)\, \mathbf{z}_l^T \bar{\mathbf{B}}_{l,i} - \bar{\mathbf{B}}_{l,i} \mathbf{z}_l \boldsymbol{\nabla}^T_{\mathbf{z}_l} \mathsf{SINR}_i \right). \tag{2.140}$$

The case $l_1 \neq l_2$ is more tricky. We start out by rewriting the gradient $\boldsymbol{\nabla}_{\mathbf{z}_{l_1}} \mathsf{SINR}_i$ as

$$\begin{aligned} \boldsymbol{\nabla}_{\mathbf{z}_{l_1}} \mathsf{SINR}_i &= \frac{2}{\mathbf{z}_{l_1}^T \bar{\mathbf{B}}_{l_1,i} \mathbf{z}_{l_1}} \cdot \left(\begin{pmatrix} \mathbf{e}_1^T \bar{\mathbf{A}}_{l_1,i} \mathbf{z}_{l_1} \\ \mathbf{e}_2^T \bar{\mathbf{A}}_{l_1,i} \mathbf{z}_{l_1} \\ \vdots \\ \mathbf{e}_{2n_{\mathcal{R}}^{(l_1)}}^T \bar{\mathbf{A}}_{l_1,i} \mathbf{z}_{l_1} \end{pmatrix} - \mathsf{SINR}_i \cdot \begin{pmatrix} \mathbf{e}_1^T \bar{\mathbf{B}}_{l_1,i} \mathbf{z}_{l_1} \\ \mathbf{e}_2^T \bar{\mathbf{B}}_{l_1,i} \mathbf{z}_{l_1} \\ \vdots \\ \mathbf{e}_{2n_{\mathcal{R}}^{(l_1)}}^T \bar{\mathbf{B}}_{l_1,i} \mathbf{z}_{l_1} \end{pmatrix} \right) \\ &= \frac{2}{\mathbf{z}_{l_2}^T \bar{\mathbf{B}}_{l_2,i} \mathbf{z}_{l_2}} \times \end{aligned} \tag{2.141}$$

$$\left(\begin{pmatrix} \mathbf{z}_{l_2}^T \left(\bar{\mathbf{A}}_{l_2,i}\big|_{\tilde{\mathbf{g}}_{l_1}^H=\mathbf{e}_1^T}\right)\mathbf{z}_{l_2} \\ \mathbf{z}_{l_2}^T \left(\bar{\mathbf{A}}_{l_2,i}\big|_{\tilde{\mathbf{g}}_{l_1}^H=\mathbf{e}_2^T}\right)\mathbf{z}_{l_2} \\ \vdots \\ \mathbf{z}_{l_2}^T \left(\bar{\mathbf{A}}_{l_2,i}\big|_{\tilde{\mathbf{g}}_{l_1}^H=\mathbf{e}^T_{n_{\mathcal{R}}^{(l_1)}}}\right)\mathbf{z}_{l_2} \\ \mathbf{z}_{l_2}^T \left(\bar{\mathbf{A}}_{l_2,i}\big|_{\tilde{\mathbf{g}}_{l_1}^H=\imath\cdot\mathbf{e}_1^T}\right)\mathbf{z}_{l_2} \\ \mathbf{z}_{l_2}^T \left(\bar{\mathbf{A}}_{l_2,i}\big|_{\tilde{\mathbf{g}}_{l_1}^H=\imath\cdot\mathbf{e}_2^T}\right)\mathbf{z}_{l_2} \\ \vdots \\ \mathbf{z}_{l_2}^T \left(\bar{\mathbf{A}}_{l_2,i}\big|_{\tilde{\mathbf{g}}_{l_1}^H=\imath\cdot\mathbf{e}^T_{n_{\mathcal{R}}^{(l_1)}}}\right)\mathbf{z}_{l_2} \end{pmatrix} - \mathsf{SINR}_i \cdot \begin{pmatrix} \mathbf{z}_{l_2}^T \left(\bar{\mathbf{B}}_{l_2,i}\big|_{\tilde{\mathbf{g}}_{l_1}^H=\mathbf{e}_1^T}\right)\mathbf{z}_{l_2} \\ \mathbf{z}_{l_2}^T \left(\bar{\mathbf{B}}_{l_2,i}\big|_{\tilde{\mathbf{g}}_{l_1}^H=\mathbf{e}_2^T}\right)\mathbf{z}_{l_2} \\ \vdots \\ \mathbf{z}_{l_2}^T \left(\bar{\mathbf{B}}_{l_2,i}\big|_{\tilde{\mathbf{g}}_{l_1}^H=\mathbf{e}^T_{n_{\mathcal{R}}^{(l_1)}}}\right)\mathbf{z}_{l_2} \\ \mathbf{z}_{l_2}^T \left(\bar{\mathbf{B}}_{l_2,i}\big|_{\tilde{\mathbf{g}}_{l_1}^H=\imath\cdot\mathbf{e}_1^T}\right)\mathbf{z}_{l_2} \\ \mathbf{z}_{l_2}^T \left(\bar{\mathbf{B}}_{l_2,i}\big|_{\tilde{\mathbf{g}}_{l_1}^H=\imath\cdot\mathbf{e}_2^T}\right)\mathbf{z}_{l_2} \\ \vdots \\ \mathbf{z}_{l_2}^T \left(\bar{\mathbf{B}}_{l_2,i}\big|_{\tilde{\mathbf{g}}_{l_1}^H=\imath\cdot\mathbf{e}^T_{n_{\mathcal{R}}^{(l_1)}}}\right)\mathbf{z}_{l_2} \end{pmatrix}\right). \tag{2.142}$$

The first equality follows directly from (2.138). The expression (2.142) is due Proposition 4. It implies that both the nominator and the denominator are generalized Rayleigh quotients in all $\tilde{\mathbf{g}}_l$, $l \in \{1, \ldots, L\}$, and thus also in all $\mathbf{z}_l$, $l \in \{1, \ldots, L\}$. The $\tilde{\mathbf{g}}_{l_1}^H$ in (2.142) explicitly refer to variables that appear in complex conjugated form. That is, all $\tilde{\mathbf{g}}_{l_1}$ that are not complex conjugated are *not* substituted. Note that assumption (2.85) is required for the second equality to hold. Let the kth row of $\boldsymbol{\nabla}^2_{\mathbf{z}_{l_1}\mathbf{z}_{l_2}}\mathsf{SINR}_i$ be denoted by $\left(\boldsymbol{\nabla}^2_{\mathbf{z}_{l_1}\mathbf{z}_{l_2}}\mathsf{SINR}_i\right)_k$. It is obtained from (2.142) as

$$\begin{aligned}\left(\boldsymbol{\nabla}^2_{\mathbf{z}_{l_1}\mathbf{z}_{l_2}}\mathsf{SINR}_i\right)_k = \frac{2}{\mathbf{z}_{l_2}^T\tilde{\mathbf{B}}_{l_2,i}\mathbf{z}_{l_2}} \Big(& 2\cdot \mathbf{z}_{l_2}^T\tilde{\mathbf{A}}_{l_2,l_1,i} - 2\cdot\mathsf{SINR}_i\cdot\mathbf{z}_{l_2}^T\tilde{\mathbf{B}}_{l_2,l_1,i} \\ & -\mathbf{z}_{l_2}^T\tilde{\mathbf{B}}_{l_2,l_1,i}\mathbf{z}_{l_2}\boldsymbol{\nabla}^T_{\mathbf{z}_{l_2}}\mathsf{SINR}_i + 2\cdot\left(\boldsymbol{\nabla}_{\mathbf{z}_{l_1}}\mathsf{SINR}_i\right)_k\cdot\mathbf{z}_{l_2}^T\tilde{\mathbf{B}}_{l_2,i}\Big),\end{aligned} \tag{2.143}$$

where $\left(\boldsymbol{\nabla}_{\mathbf{z}_{l_1}}\mathsf{SINR}_i\right)_k$ denotes the kth element of $\boldsymbol{\nabla}_{\mathbf{z}_{l_1}}\mathsf{SINR}_i$ and

$$\tilde{\mathbf{A}}_{l_2,l_1,i} = \frac{1}{2}\cdot\left(\begin{pmatrix} \bar{\mathbf{A}}_{l_2}\big|_{\tilde{\mathbf{g}}_{l_1}^H=\mathbf{e}_1^T} \\ \vdots \\ \bar{\mathbf{A}}_{l_2}\big|_{\tilde{\mathbf{g}}_{l_1}^H=\mathbf{e}_{n_{\mathcal{R}}^{(l_1)}}^T} \\ \bar{\mathbf{A}}_{l_2}\big|_{\tilde{\mathbf{g}}_{l_1}^H=\imath\cdot\mathbf{e}_1^T} \\ \vdots \\ \bar{\mathbf{A}}_{l_2}\big|_{\tilde{\mathbf{g}}_{l_1}^H=\imath\cdot\mathbf{e}_{n_{\mathcal{R}}^{(l_1)}}^T} \end{pmatrix} + \begin{pmatrix} \bar{\mathbf{A}}_{l_2}\big|_{\tilde{\mathbf{g}}_{l_1}^H=\mathbf{e}_1^T} \\ \vdots \\ \bar{\mathbf{A}}_{l_2}\big|_{\tilde{\mathbf{g}}_{l_1}^H=\mathbf{e}_{n_{\mathcal{R}}^{(l_1)}}^T} \\ \bar{\mathbf{A}}_{l_2}\big|_{\tilde{\mathbf{g}}_{l_1}^H=\imath\cdot\mathbf{e}_1^T} \\ \vdots \\ \bar{\mathbf{A}}_{l_2}\big|_{\tilde{\mathbf{g}}_{l_1}^H=\imath\cdot\mathbf{e}_{n_{\mathcal{R}}^{(l_1)}}^T} \end{pmatrix}^T\right), \tag{2.144}$$

$$\tilde{\mathbf{B}}_{l_2,l_1,i} = \frac{1}{2}\cdot\left(\begin{pmatrix} \bar{\mathbf{B}}_{l_2}\big|_{\tilde{\mathbf{g}}_{l_1}^H=\mathbf{e}_1^T} \\ \vdots \\ \bar{\mathbf{B}}_{l_2}\big|_{\tilde{\mathbf{g}}_{l_1}^H=\mathbf{e}_{n_{\mathcal{R}}^{(l_1)}}^T} \\ \bar{\mathbf{B}}_{l_2}\big|_{\tilde{\mathbf{g}}_{l_1}^H=\imath\cdot\mathbf{e}_1^T} \\ \vdots \\ \bar{\mathbf{B}}_{l_2}\big|_{\tilde{\mathbf{g}}_{l_1}^H=\imath\cdot\mathbf{e}_{n_{\mathcal{R}}^{(l_1)}}^T} \end{pmatrix} + \begin{pmatrix} \bar{\mathbf{B}}_{l_2}\big|_{\tilde{\mathbf{g}}_{l_1}^H=\mathbf{e}_1^T} \\ \vdots \\ \bar{\mathbf{B}}_{l_2}\big|_{\tilde{\mathbf{g}}_{l_1}^H=\mathbf{e}_{n_{\mathcal{R}}^{(l_1)}}^T} \\ \bar{\mathbf{B}}_{l_2}\big|_{\tilde{\mathbf{g}}_{l_1}^H=\imath\cdot\mathbf{e}_1^T} \\ \vdots \\ \bar{\mathbf{B}}_{l_2}\big|_{\tilde{\mathbf{g}}_{l_1}^H=\imath\cdot\mathbf{e}_{n_{\mathcal{R}}^{(l_1)}}^T} \end{pmatrix}^T\right). \tag{2.145}$$

Here, the fact is used that a quadratic form $\mathbf{x}^T\mathbf{K}\mathbf{x}$ solely depends on the even symmetric part $1/2(\mathbf{K}+\mathbf{K}^T)$ of the matrix $\mathbf{K}$.

Below, two optimization schemes are applied to the problem (2.82). The first method has been introduced in [62] and requires only first order information, i.e., the gradients of the source-destination pair SINRs. The second method requires second order information, i.e., also the corresponding Hessians, and applies Newton's method [66].

Optimization based on first order information: Since the minimum function is not differentiable, the classic method of steepest descent cannot be applied to the problem (2.82). An alternative optimization method based on first order information has been proposed in [62]. It is a generalization of the method of steepest descent and provably converges to a local maximum. As the method of steepest descent, the two main building blocks of the algorithm are the computation of a search direction and a line search along this direction. The identification of the search direction requires the concept of an ε-active set. This set is defined in the style of (2.86) as

$$\mathcal{A}_\varepsilon^{(k)} \triangleq \left\{ i \in \{1,\dots,n\} \;\middle|\; \mathsf{SINR}_i\big|_{\mathbf{z}=\mathbf{z}^{(k)}} - \min_{j\in\{1,\dots,n\}} \mathsf{SINR}_j\big|_{\mathbf{z}=\mathbf{z}^{(k)}} < \varepsilon \right\}, \tag{2.146}$$

where $\mathbf{z}^{(k)}$ denotes the value of $\mathbf{z} \triangleq (\mathbf{z}_1^T,\dots,\mathbf{z}_L^T)^T$ in iteration k. Let $(\gamma_i^{(0)})_{i\in\mathcal{A}_\varepsilon}$ be the minimizer of the optimization problem:

$$\begin{aligned} \text{minimize} \quad & \left\| \sum_{i\in\mathcal{A}_\varepsilon} \gamma_i \cdot \boldsymbol{\nabla}_{\mathbf{z}} \mathsf{SINR}_i\big|_{\mathbf{z}=\mathbf{z}^{(k)}} \right\|^2, \\ \text{subject to} \quad & \sum_{i\in\mathcal{A}_\varepsilon^{(k)}} \gamma_i = 1, \\ & \gamma_i \geq 0 \text{ for all } i \in \mathcal{A}_\varepsilon^{(k)}. \end{aligned} \tag{2.147}$$

The search direction is then given by

$$\boldsymbol{\Delta}^{(k)} = \frac{\sum_{i\in\mathcal{A}_\varepsilon^{(k)}} \gamma_i^{(0)} \cdot \boldsymbol{\nabla}_{\mathbf{z}} \mathsf{SINR}_i\big|_{\mathbf{z}=\mathbf{z}^{(k)}}}{\left\| \sum_{i\in\mathcal{A}_\varepsilon^{(k)}} \gamma_i^{(0)} \cdot \boldsymbol{\nabla}_{\mathbf{z}} \mathsf{SINR}_i\big|_{\mathbf{z}=\mathbf{z}^{(k)}} \right\|}. \tag{2.148}$$

The problem (2.147) is a quadratic programming problem and efficiently solved by standard methods [66]. Geometrically, the search direction corresponds to the point on the surface of the convex hull of all $\boldsymbol{\nabla}_{\mathbf{z}}\mathsf{SINR}_i$, $i \in \mathcal{A}_\varepsilon^{(k)}$, that is closest to the origin in Eukledian distance. At this point, all inner products $\boldsymbol{\Delta}^T \cdot \boldsymbol{\nabla}_{\mathbf{z}}\mathsf{SINR}_i$, $i \in \mathcal{A}_\varepsilon^{(k)}$, are equal. Intuitively, all source-destination pairs that are in the ε-active set can thus expect the same "gain" for small step-sizes. Moreover, the search-direction is non-zero, unless $\mathbf{z}^{(k)}$ fulfills the necessary conditions for a local optimum (2.87).

The conventional update equation

$$\mathbf{z}^{(k+1)} = \mathbf{z}^{(k)} + \mu_0^{(k)} \cdot \boldsymbol{\Delta}^{(k)}, \tag{2.149}$$

with corresponding line-search

$$\mu_0^{(k)} = \operatorname*{argmax}_{\mu>0} \min_{i\in\{1,\dots,n\}} \mathsf{SINR}_i \Big|_{\mathbf{z}=\mathbf{z}^{(k)}+\mu\cdot\boldsymbol{\Delta}^{(k)}} \tag{2.150}$$

has a problem in the case at hand. An important assumption in the convergence proof of the considered algorithm is the boundedness of the set

$$\mathcal{M}(\mathbf{z}^{(k)}) \triangleq \left\{ \mathbf{z} \in \mathbb{R}^{2 \cdot \sum_{l=1}^{L} n_{\mathcal{R}}^{(l)}} \middle| \max_{\mathbf{z}} \min_{i \in \{1,\ldots,n\}} \mathsf{SINR}_i > \mathsf{SINR}_i\big|_{\mathbf{z}=\mathbf{z}^{(k)}} \right\}. \tag{2.151}$$

This condition would ensure that the maximizing μ in (2.150) can be attained. However, since all SINR_i are invariant under scaling of $\mathbf{z}$, $\mathcal{M}(\mathbf{z}^{(k)})$ cannot be bounded. This issue is avoided by the update-equation

$$\mathbf{z}^{(k+1)} = \frac{(1-\mu_0^{(k)}) \cdot \mathbf{z}^{(k)} + \mu_0^{(k)} \cdot \boldsymbol{\Delta}^{(k)}}{\|(1-\mu_0^{(k)}) \cdot \mathbf{z}^{(k)} + \mu_0^{(k)} \cdot \boldsymbol{\Delta}^{(k)}\|} \tag{2.152}$$

with corresponding line-search

$$\mu_0^{(k)} = \operatorname*{argmax}_{0<\mu\leq 1} \min_{i \in \{1,\ldots,n\}} \mathsf{SINR}_i \Bigg|_{\mathbf{z}=\frac{(1-\mu)\cdot\mathbf{z}^{(k)}+\mu\cdot\boldsymbol{\Delta}^{(k)}}{\|(1-\mu)\cdot\mathbf{z}^{(k)}+\mu\cdot\boldsymbol{\Delta}^{(k)}\|}}. \tag{2.153}$$

Here, the search interval is bounded, and

$$\max_{\mu>0} \min_{i \in \{1,\ldots,n\}} \mathsf{SINR}_i \Bigg|_{\mathbf{z}=\mathbf{z}^{(k)}+\mu\cdot\boldsymbol{\Delta}^{(k)}} = \max_{0<\mu\leq 1} \min_{i \in \{1,\ldots,n\}} \mathsf{SINR}_i \Bigg|_{\mathbf{z}=\frac{(1-\mu)\cdot\mathbf{z}^{(k)}+\mu\cdot\boldsymbol{\Delta}^{(k)}}{\|(1-\mu)\mathbf{z}^{(k)}+\mu\cdot\boldsymbol{\Delta}^{(k)}\|}}. \tag{2.154}$$

The method as proposed in [62] is summarized in Algorithm 2. Unfortunately, we do not have an upper-bound on the gap between current minimum SINR and the true optimum, and thus no reliable stopping criterion for the case $L > 1$.

Algorithm 2 Optimization algorithm based on first order information [62]

given $\mathbf{z}^{(0)}$, $k = 0$, $\varepsilon_0 > 0$, $\rho_0 > 0$.
loop
  determine $\mathcal{A}_\varepsilon^{(k)}$.
  for all $i \in \mathcal{A}_\varepsilon^{(k)}$ compute gradients $\boldsymbol{\nabla}_{\mathbf{z}}\mathsf{SINR}_i\big|_{\mathbf{z}=\mathbf{z}^{(k)}}$.
  determine search direction $\boldsymbol{\Delta}^{(k)}$.
  obtain $\mathbf{z}^{(k+1)}$ through line search.
  if $\min_{i\in\mathcal{A}_\varepsilon^{(k)}} \boldsymbol{\Delta}^{(k)} \cdot \boldsymbol{\nabla}_{\mathbf{z}}^T\mathsf{SINR}_i\big|_{\mathbf{z}=\mathbf{z}^{(k)}} < \rho_k$ **then**
    $\varepsilon_k \leftarrow \varepsilon_0/2^k$;
    $\rho_k \leftarrow \rho_0/2^k$.
  end if
  $k \leftarrow k+1$.
end loop

Optimization based on second order information: All of the above optimization algorithms are well suited for obtaining a rough estimate of the locally optimal $\mathbf{z}$. However, none of them shows good convergence properties in the immediate vicinity of the optimum, since they converge only linearly in the best case. Therefore, we introduce here Newton's method, which converges quadratically. This method is particularly desirable as a complement to Algorithm 2, since fast convergence near the optimum is of special importance, when no proper stopping criterion is known. It should be emphasized that Newton's method converges only locally in general [66]. Therefore, it requires a good initialization, which is assumed to be achieved through K iterations in Algorithm 2 in the following.

We propose a method which in first the instance assumes that $\mathcal{A}_\varepsilon^{(K)}$, as obtained by Algorithm 2, is the true active set $\mathcal{A}$ at the local maximum. Under this assumption, we can consider the following equality constrained optimization problem in $(\mathbf{z}_1, \ldots, \mathbf{z}_L)$:

$$\begin{aligned} &\text{maximize } \mathsf{SINR}_j, \\ &\text{subject to } \mathsf{SINR}_i = \mathsf{SINR}_j \text{ for all } i \in \mathcal{A}_\varepsilon^{(K)} \backslash \{j\}, \end{aligned} \tag{2.155}$$

where $j \in \mathcal{A}_\varepsilon^{(K)}$. The local optimum is then found by solving the following system of $\left(\sum_{l=1}^{L} n_{\mathcal{R}}^{(l)}\right) + |\mathcal{A}_\varepsilon^{(K)}| - 1$ equations in just as many unknowns (cf. (2.100)-(2.102)):

$$\boldsymbol{\nabla}_{\mathbf{z}} \mathsf{SINR}_j + \sum_{i \in \mathcal{A}_\varepsilon^{(K)} \backslash \{j\}} \tilde{\gamma}_i \cdot \boldsymbol{\nabla}_{\mathbf{z}} \mathsf{SINR}_i = 0, \tag{2.156}$$

$$\mathsf{SINR}_i - \mathsf{SINR}_j = 0 \text{ for all } i \in \mathcal{A}_\varepsilon^{(K)} \backslash \{j\}. \tag{2.157}$$

This system is obtained by taking derivatives in the Lagrangian of (2.155), where the $\tilde{\gamma}_i$ are the respective Lagrangian multipliers. A comparison of (2.156) and (2.87) reveals that the $\tilde{\gamma}_i$ and γ_i, $i \in \mathcal{A}_\varepsilon^{(k)} \backslash \{j\}$, are related as follows:

$$\tilde{\gamma}_i = \frac{1}{\gamma_j} \cdot \gamma_i. \tag{2.158}$$

A solution of the system is thus a local optimum (for a sufficiently good initialization), if $\tilde{\gamma}_i \geq 0$ for all $i \in \mathcal{A}_\varepsilon^{(K)}$. If any of the involved γ_i is negative, the assumption that $\mathcal{A}_\varepsilon^{(K)}$ coincides with the active set at the true local optimum is wrong. In this case, either the initialization has to be refined or the assumed active set needs to be reduced.

Let us define the vector $\tilde{\boldsymbol{\gamma}}^{(k)} \triangleq (\tilde{\gamma}_i^{(k)})_{i \in \mathcal{A}_\varepsilon^{(K)} \backslash \{j\}}$, where the $\tilde{\gamma}_i^{(k)}$ correspond to the values of the $\tilde{\gamma}_i$ in iteration k. The Newton update-equation is then given by

$$\mathbf{z}^{(k+1)} = \mathbf{z}^{(k)} + \boldsymbol{\Delta}_{\mathbf{z}}^{(k)}, \tag{2.159}$$

$$\tilde{\boldsymbol{\gamma}}^{(k+1)} = \tilde{\boldsymbol{\gamma}}^{(k)} + \boldsymbol{\Delta}_{\tilde{\boldsymbol{\gamma}}}^{(k)}, \tag{2.160}$$

where the update vectors $\boldsymbol{\Delta}_{\tilde{\mathbf{z}}}^{(k)}$ and $\boldsymbol{\Delta}_{\tilde{\boldsymbol{\gamma}}}^{(k)}$ in iteration k are given by the solution to the linear equation system

$$\underbrace{\begin{pmatrix} \mathbf{A}_{11}^{(k)} & \mathbf{A}_{12}^{(k)} \\ \mathbf{A}_{21}^{(k)} & \mathbf{A}_{22}^{(k)} \end{pmatrix}}_{\mathbf{A}^{(k)}} \cdot \underbrace{\begin{pmatrix} \boldsymbol{\Delta}_{\tilde{\mathbf{z}}}^{(k)} \\ \boldsymbol{\Delta}_{\tilde{\boldsymbol{\gamma}}}^{(k)} \end{pmatrix}}_{\boldsymbol{\Delta}^{(k)}} = - \underbrace{\begin{pmatrix} \mathbf{b}_1^{(k)} \\ \mathbf{b}_2^{(k)} \end{pmatrix}}_{\mathbf{b}^{(k)}}, \tag{2.161}$$

with

$$\mathbf{A}_{11}^{(k)} = \boldsymbol{\nabla}_{\mathbf{z}}^2 \mathsf{SINR}_j\big|_{\mathbf{z}=\mathbf{z}^{(k)}} + \sum_{i \in \mathcal{A}_\varepsilon^{(K)} \setminus \{j\}} \tilde{\gamma}_i^{(k)} \cdot \boldsymbol{\nabla}_{\mathbf{z}}^2 \mathsf{SINR}_i\big|_{\mathbf{z}=\mathbf{z}^{(k)}}, \tag{2.162}$$

$$\mathbf{A}_{12}^{(k)} = \left(\left(\boldsymbol{\nabla}_{\mathbf{z}}^T \mathsf{SINR}_i\big|_{\mathbf{z}=\mathbf{z}^{(k)}} \right)_{i \in \mathcal{A}_\varepsilon^{(K)} \setminus \{j\}} \right)^T, \tag{2.163}$$

$$\mathbf{A}_{21}^{(k)} = \left(\boldsymbol{\nabla}^T \mathsf{SINR}_i\big|_{\mathbf{z}=\mathbf{z}^{(k)}} - \boldsymbol{\nabla}^T \mathsf{SINR}_j\big|_{\mathbf{z}=\mathbf{z}^{(k)}} \right)_{i \in \mathcal{A}_\varepsilon^{(K)} \setminus \{j\}}, \tag{2.164}$$

$$\mathbf{A}_{22}^{(k)} = \mathbf{0}_{(|\mathcal{A}_\varepsilon^{(K)}|-1) \times (|\mathcal{A}_\varepsilon^{(K)}|-1)} \tag{2.165}$$

and

$$\mathbf{b}_1^{(k)} = \boldsymbol{\nabla}_{\mathbf{z}} \mathsf{SINR}_j\big|_{\mathbf{z}=\mathbf{z}^{(k)}} + \sum_{i \in \mathcal{A}_\varepsilon^{(K)} \setminus \{j\}} \tilde{\gamma}_i^{(k)} \cdot \boldsymbol{\nabla}_{\mathbf{z}} \mathsf{SINR}_i\big|_{\mathbf{z}=\mathbf{z}^{(k)}}, \tag{2.166}$$

$$\mathbf{b}_2^{(k)} = \left(\mathsf{SINR}_i\big|_{\mathbf{z}=\mathbf{z}^{(k)}} - \mathsf{SINR}_j\big|_{\mathbf{z}=\mathbf{z}^{(k)}} \right)_{i \in \mathcal{A}_\varepsilon^{(K)} \setminus \{j\}}. \tag{2.167}$$

A problem of the method outlined so far, is the fact that the matrix $\mathbf{A}^{(k)}$ in (2.161) is singular in the optimum. Thus, an important assumption that is required in standard proofs for quadratic convergence of Newton's method is violated. This problem is circumvented by a dehomogenization of the problem as discussed in Section 2.3.2. That is, the first and the $(n_{\mathcal{R}}^{(l)}+1)$st element of each $\mathbf{z}_l$, $l \in \{1, \ldots, L\}$, i.e., the real and imaginary part of the first element of $\tilde{\mathbf{g}}_l$, need to be fixed to arbitrary values that are not both zero. If the vector of remaining unknowns is denoted by $\tilde{\mathbf{z}}$, the corresponding update-equation is obtained from (2.159) and (2.161) by deleting all lines and columns in $\mathbf{A}^{(k)}$ and elements in $\boldsymbol{\Delta}^{(k)}$ and $\mathbf{b}^{(k)}$ that correspond to derivatives with respect to any of the $2L$ eliminated variables. The insight that the resulting matrix in (2.165) is full rank in the optimum point for the modified update-equation stems from numerical experiments and is not proved in this thesis.

An algorithm that uses Newton's method for the polishing of an estimate of a local optimum, which is obtained by Algorithm 2, is outlined in Algorithm 3. The basic idea is to run this algorithm after a fixed number K of iterations in Algorithm 2. Newton's method will

- either confirm the choice of the active set and polish the optimum,
- or contradict the assumption of the active set, since either
  - one or more Lagrange multipliers are negative
  - or $\exists j \notin \mathcal{A}_\varepsilon^{(K)}$, such that $\mathsf{SINR}_j < \min_{i \in \mathcal{A}_\varepsilon^{(K)}} \mathsf{SINR}_i$,
- or it will not converge.

In the latter two cases, Algorithm 2 is run for another K iterations, until eventually Newton's method converges and confirms the assumed active set.

Algorithm 3 Optimization algorithm based on second order information

given $K \in \mathbb{N}^+$.
run Algorithm 2 for K iterations with random initialization $\mathbf{z}^{(0)}$.
$\mathbf{z}^{(0)} \leftarrow \mathbf{z}^{(K)}$.
loop
  run Newton's method under assumption of active set $\mathcal{A}_\varepsilon^{(K)}$.
  if Newton's method converged to $\mathbf{z} = \mathbf{z}^*$ and $\tilde{\gamma}_i = \tilde{\gamma}_i^*$
      and $\min_{i \in \mathcal{A}_\varepsilon^{(K)}} \tilde{\gamma}_i^* \geq 0$ *and* $\min_{j \notin \mathcal{A}_\varepsilon^{(K)}} \mathsf{SINR}_j\big|_{\mathbf{z}=\mathbf{z}^*} \geq \min_{i \in \mathcal{A}_\varepsilon^{(K)}} \mathsf{SINR}_i\big|_{\mathbf{z}=\mathbf{z}^*}$ **then**
    return $\mathbf{z}^*$;
  else
    run Algorithm 2 with initialization $\mathbf{z}^{(0)}$;
    $\mathbf{z}^{(0)} \leftarrow \mathbf{z}^{(K)}$.
  end if
end loop

Note that this algorithm could potentially converge to a local minimum or a saddle point, since conditions (2.156) and (2.157) are only necessary and not sufficient for a local maximum. Both events are unlikely to occur, since very poor initializations through Algorithm 2 are required. In the following, we identify the class of stationary points that we are actually interested in. If we choose an arbitrary $l' \in \{1, \dots, L\}$ and fix all $\mathbf{z}_l$, $l \in \{1, \dots, L\} \backslash \{l'\}$, we obtain an equivalent two-hop network with a single relay stage $\mathcal{R}_{l'}$. We know from Subsection 2.6.1 that $\mathbf{z}_{l'}$ is suboptimal, if (2.106)-(2.109) does not hold with equality. Therefore, the global maximum must fulfill

$$\lambda_1 \left\{ \sum_{i \in \mathcal{A}} \gamma_i \cdot \bar{\mathbf{A}}_{l,i}, \sum_{i \in \mathcal{A}} \gamma_i \cdot \bar{\mathbf{B}}_{l,i} \right\} = \min_{i \in \{1,\ldots,n\}} \mathsf{SINR}_i \text{ for all } l \in \{1, \ldots, L\}. \tag{2.168}$$

Whether or not an identified stationary point does fulfill this condition, is straightforward to check. There are, in general, multiple of such stationary points, and computer simulations suggest that their number depends for a given channel realization on the ratio P/σ^2. For an illustration consider the plots of Fig. 2.8. We show for a three-hop network with $n = n_{\mathcal{R}}^{(1)} = n_{\mathcal{R}}^{(2)} = 2$, fixed channel and various random initializations the evolution of the minimum SINR with the iterations in Algorithm 2. The two plots correspond to the cases $P/\sigma^2 = 50$ (top) and $P/\sigma^2 = 100$ (bottom), and we show only curves whose corresponding stationary point fulfills (2.168). We observe (also based on larger sets of sample curves than shown here) that the number of stationary points in this network is either one or two for this (and also every other generic) channel. Generally, observations from computer experiments suggests the following about stationary points that fulfill (2.168) in generic channels:

(i) Their number is one for all network topologies in the regime of sufficiently small P/σ^2, and increases up to a certain limit for increasing P/σ^2.

(ii) This limit on the number of solutions in the regime of large P/σ^2 grows rapidly for n and L.

The second point might be related to the fact that also the number of zero-forcing solutions (see Tab. 2.1) grows rapidly with the network size.

2.6.3. Is the DMT Upper-Bound Achievable?

The upper-bound on the DMT curve of Section 2.5 is very "optimistic" in fact. Not only does it ignore most of the noise introduced in the network, it also assumes full cooperation among relay nodes that are within the same stage. Accordingly, it is an obvious question, whether this upper-bound is meaningful in any sense. Although we are not able to answer this question in a rigorous way, we provide some evidence that the upper-bound might be achievable, whenever distributed zero-forcing is feasible.

For our outage simulation results, realizations of the fading coefficients h_{YX} are generated independently from a CSCG random variable of unit variance. Moreover, the per-stage transmit powers are fixed to $P_{\mathcal{S}} = P_{\mathcal{R}_1} = \ldots = P_{\mathcal{R}_L} = P$ and the noise

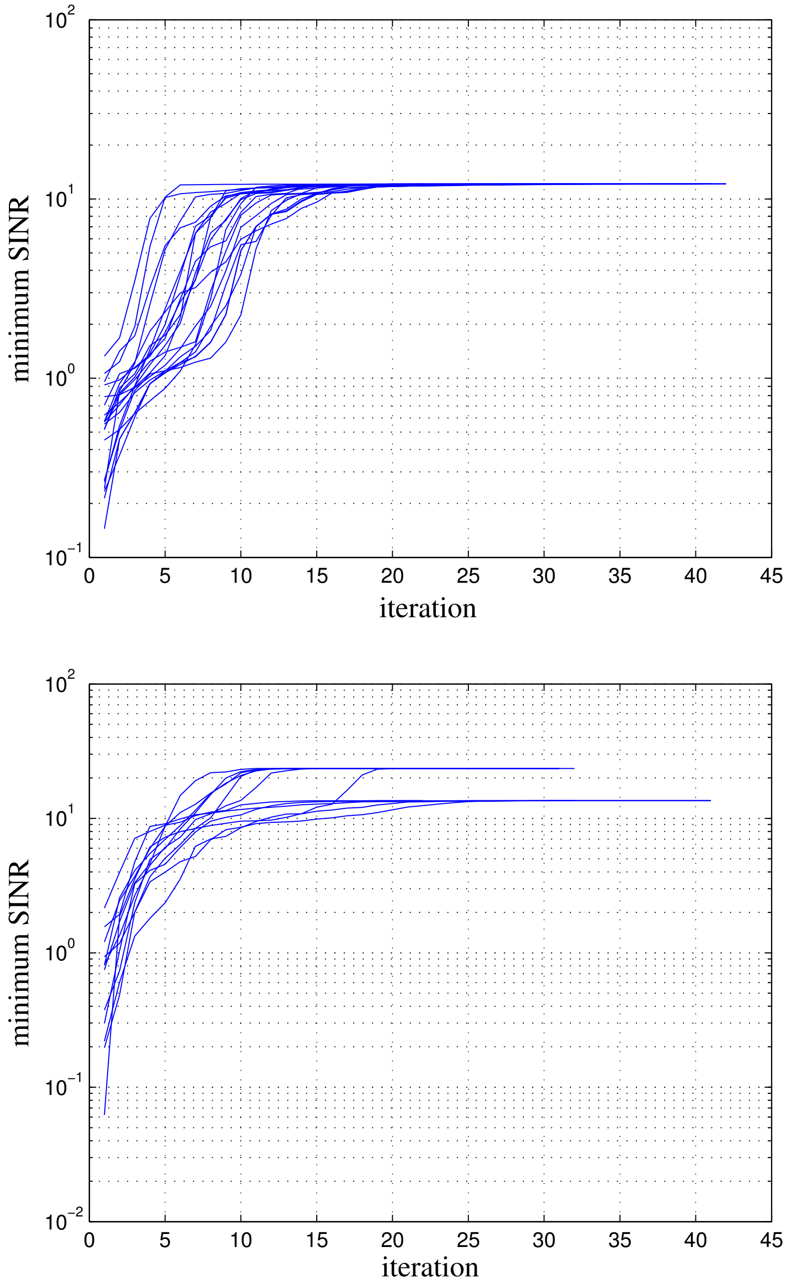


Figure 2.8.: Evolution of the minimum SINR with the iterations in Algorithm 2 for a three-hop network with $n = n_{\mathcal{R}}^{(1)} = n_{\mathcal{R}}^{(2)} = 2$, fixed channel and various random initializations. The two plots correspond to the cases $P/\sigma^2 = 50$ (top) and $P/\sigma^2 = 100$ (bottom), and only curves whose corresponding stationary point fulfills (2.168) are shown.

variance is set to $\sigma^2 = 1$. Moreover, we assume a constant code rate $R = 1$ bit per channel use, which corresponds to a multiplexing gain $r = 0$.

2.6.3.1. Two-Hop Networks ($L = 1$)

In this case, we can apply the methods of Subsection 2.6.1 in order to globally maximize the SINR of the weakest source-destination pair. Thus, we can simulate the maximal achievable diversity. Results are shown in Fig. 2.9. Remarkably, numerical evidence suggests that the diversity upper-bound (2.51), which evaluates to $d(0) \leq n_{\mathcal{R}}^{(1)} - n + 1$, is achieved with equality, whenever distributed zero-forcing is feasible. The upper plot shows performance results for two-hop networks with two source-destination pairs for various numbers of relay nodes. In the case $n_{\mathcal{R}}^{(1)} = 2$, it is not surprising that outage probability does not tend to zero as the SNR is increased. This is due to the fact, that the two relay nodes are not able to completely eliminate interference in the network. Therefore, the error floor is a result of interference limitation. For three relay nodes, the observed diversity is $d(0) = 2$, which coincides with the upper-bound (2.51). The same observation applies for larger relay numbers. This means, that each additional relay provides a diversity increment of one order. The plot on the right-hand side of Fig. 2.9 shows the corresponding curves for networks with two, three and four source-destination pairs and the minimum number of relay nodes that is required for distributed zero-forcing, i.e., topologies with $n_{\mathcal{R}}^{(1)} = n \cdot (n - 1) + 1$. Also for these networks the upper-bound (2.51) appears to be met according to numerical evidence.

2.6.3.2. Long Networks ($L \geq n$)

For networks with $L \geq 2$, we do not have methods to find the global maximum (except for an exhaustive search over all local maxima). For this reason, we consider less complex relay gain allocation strategies to achieve the maximal diversity.

We consider a network with n source-destination pairs and at least n relay stages with $n_{\mathcal{R}}^{(l)} \geq n$ for all $l \in \{1, \dots, L\}$. According to conditions (2.36) and (2.37), the existence of a zero-forcing gain allocation in such a network is guaranteed with probability one. In particular, it suffices to use only n relay nodes in each stage for distributed zero-forcing. This motivates the concept of *relay set selection*. The method tests different relay sets with respect to the maximal rate that is achievable for the weakest source-destination pair and schedules the best set for data transmission.

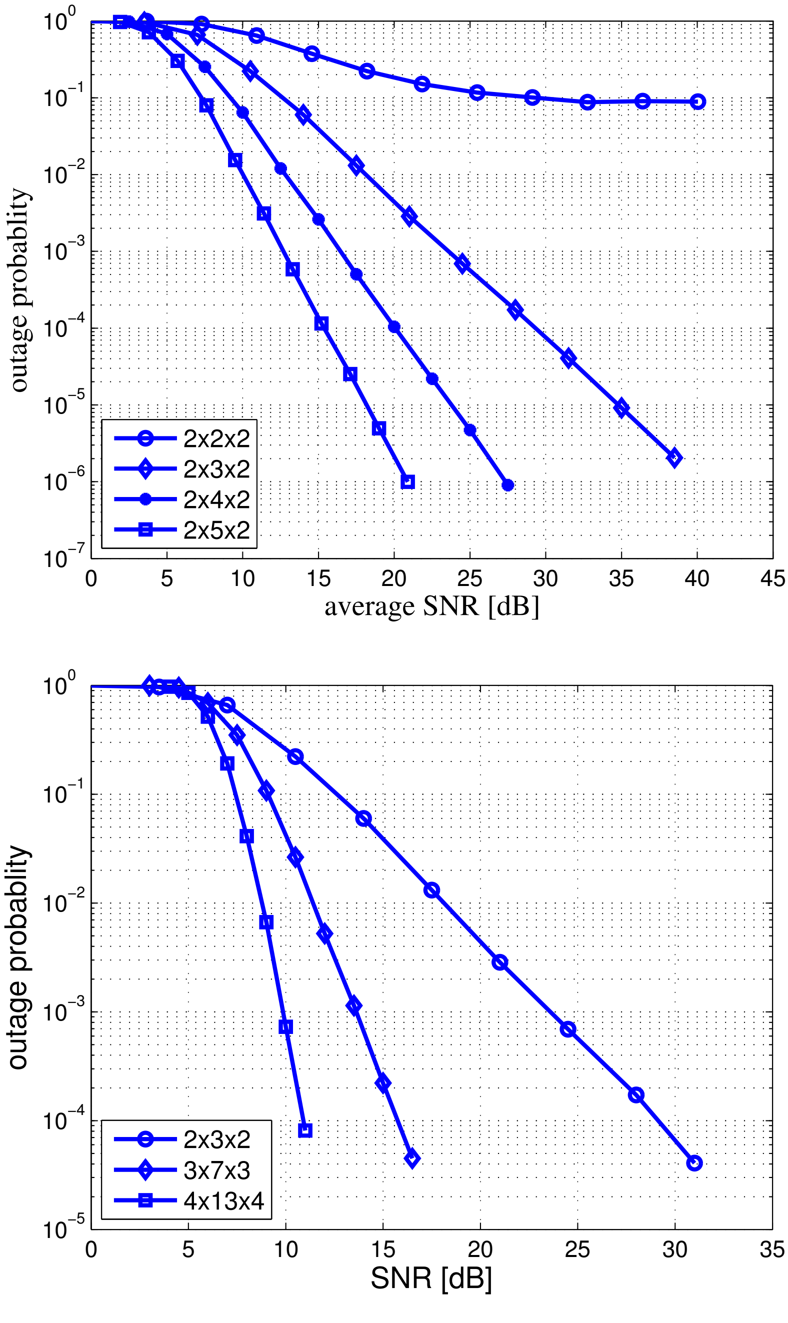


Figure 2.9.: Outage probability vs. SNR for a constant code rate $R = 1$ and various two-hop networks. The pre-log is set to $p^{-1} = 1$.

The key issue with such a selection algorithm is the assembling of the relay sets for the individual test cycles. The number of relay sets that contain n relay nodes of each stage is given by

$$\prod_{l=1}^{L} \binom{n_{\mathcal{R}}^{(l)}}{n}.$$

Therefore, an exhaustive search over all these sets is feasible only for very small numbers of excess relay nodes. It would be desirable, if tests of only $\bar{d}(0)$ sets were sufficient for achieving a diversity order $\bar{d}(0)$.

Algorithm 4 Relay Set Selection Algorithm for $L \geq n$

given $\bar{d}(0)$, n, $n_{\mathcal{R}}^{(1)}, \ldots, n_{\mathcal{R}}^{(L)}$, $n_{\mathcal{R}}^{(L+1)} \triangleq n$, $\mathsf{SNR}^* = 0$, $\mathbf{k} = \mathbf{0}_{L\times 1}$.
for $m = 1$ to $\bar{d}(0)$ **do**
  for $l = 1, 3, 5, ...$ **do**
    $k_l \leftarrow \Big((m-1) \bmod (n_{\mathcal{R}}^{(l)} - n + 1)\Big) + 1$.
    $\mathcal{R}_l^{(m)} \leftarrow \Big\{\mathrm{R}_{k_l}^{(l)}, \ldots, \mathrm{R}_{k_l+n-1}^{(l)}\Big\}$.
  end for
  for $l = 2, 4, 6, ...$ **do**
    $k_l \leftarrow \Big\lceil m / \min(n_{\mathcal{R}}^{(l-1)} - n + 1, n_{\mathcal{R}}^{(l+1)} - n + 1)\Big\rceil$.
    $\mathcal{R}_l^{(m)} \leftarrow \Big\{\mathrm{R}_{k_l}^{(l)}, \ldots, \mathrm{R}_{k_l+n-1}^{(l)}\Big\}$.
  end for
  Compute an arbitrary zero-forcing gain allocation for $\mathcal{R}_1^{(m)}, \ldots, \mathcal{R}_L^{(m)}$.
  for $j = 1$ to n **do**
    $\mathsf{SNR}_j \leftarrow$ SNR at D_j for current gain allocation.
  end for
  if $\min_j \mathsf{SNR}_j \geq \mathsf{SNR}^*$ **then**
    $\mathsf{SNR}^* \leftarrow \min_j \mathsf{SNR}_j$.
    $\mathcal{R}^* \leftarrow \Big\{\mathcal{R}_1^{(m)}, \ldots, \mathcal{R}_L^{(m)}\Big\}$.
  end if
end for
return $\mathcal{R}^*$, SNR^*.

With this objective in mind, we propose the following heuristic relay set selection algorithm. The relay set for test cycle m is given by $\cup_{l=1}^{L} \mathcal{R}_l^{(m)}$, where $\mathcal{R}_l^{(m)} \subseteq \mathcal{R}_l$ and $|\mathcal{R}_l^{(m)}| = n$ for all $l \in \{1, \ldots, L\}$. We distinguish relay stages with odd and even index l. For stages with odd indexes, we choose

$$\mathcal{R}_l^{(m)} = \left\{\mathrm{R}^{(l)}_{(m-1)\left(\mathrm{mod}(n_{\mathcal{R}}^{(l)}-n+1)\right)+1}, \mathrm{R}^{(l)}_{(m-1)\left(\mathrm{mod}(n_{\mathcal{R}}^{(l)}-n+1)\right)+2}, \ldots, \mathrm{R}^{(l)}_{(m-1)\left(\mathrm{mod}(n_{\mathcal{R}}^{(l)}-n+1)\right)+n}\right\}.$$

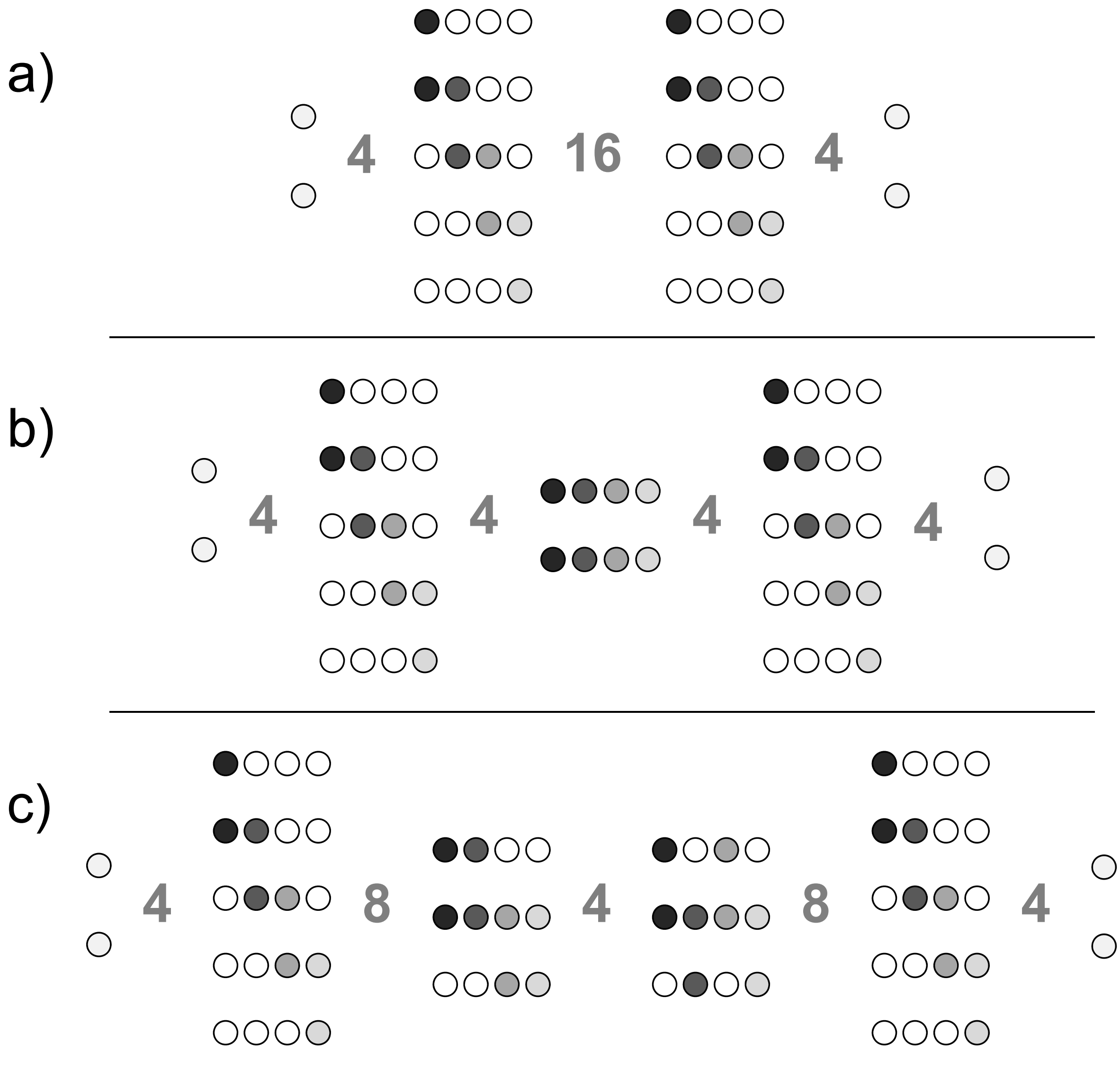


Figure 2.10.: Relay set selection algorithm applied to three different sample topologies with diversity order four: 2×5×5×2 (top), 2×5×2×5×2 (middle), 2×5×3×3×5×2 (bottom). Relay nodes as selected in cycles 1 to 4 are marked in the respective column within a relay stage each. Bounds $\bar{d}_l(0)$ are indicated between relay stages.

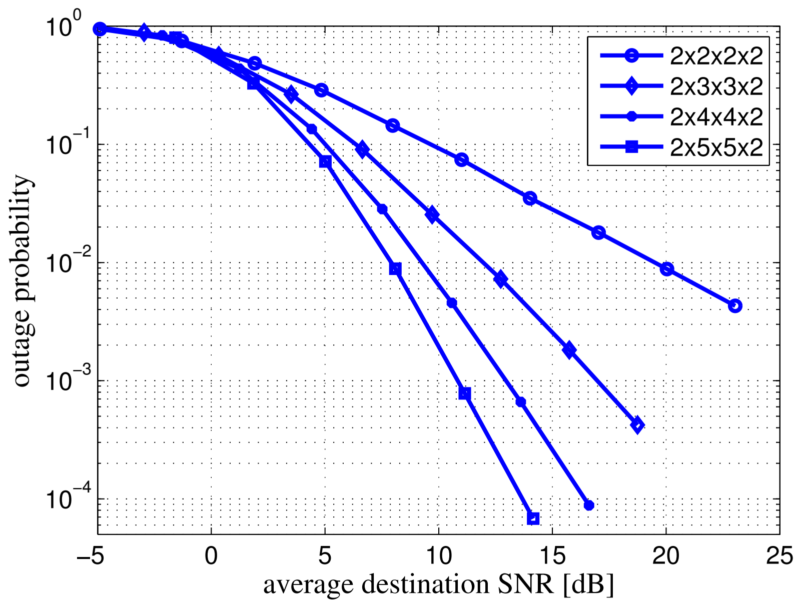


Figure 2.11.: Outage probability vs. SNR for a constant code rate $R = 1$ and various three-hop networks. The pre-log is set to $p^{-1} = 1$.

For stages with even indexes, we define $m_l \triangleq \min(n_{\mathcal{R}}^{(l-1)} - n + 1, n_{\mathcal{R}}^{(l+1)} - n + 1)$ and choose

$$\mathcal{R}_l^{(m)} = \left\{ \mathrm{R}^{(l)}_{\lceil m/m_l \rceil}, \mathrm{R}^{(l)}_{\lceil m/m_l \rceil+1}, \ldots, \mathrm{R}^{(l)}_{\lceil m/m_l \rceil+n-1} \right\},$$

where $n_{\mathcal{R}}^{(L+1)} \triangleq n$. Since $\bar{d}(0)/m_l \leq n_{\mathcal{R}}^{(l)} - n + 1$, we change the candidate set $n_{\mathcal{R}}^{(l)} - n + 1$ times at most in relay stages with even index. This relay set selection algorithm ensures by construction that, in each cycle m, there is at least a single pair of transmit and receive antennas in each hop that has not been jointly tested in any prior cycle.

In every test cycle, the zero-forcing gain allocation is obtained by assigning equal gain coefficients to the n tested relay nodes in each of the stages $\mathcal{R}_1, \ldots, \mathcal{R}_{L-n}$ and by using the gain coefficients of the n relay nodes in each of the stages $\mathcal{R}_{L-n+1}, \ldots, \mathcal{R}_L$ to solve the resulting system of equations (2.20) and inequations (2.21). We randomly pick any out of the finitely many solutions in the product of projective spaces as potential relay gain allocation and test it, subject to the sum-power constraints, with respect to the maximum rate that is achievable by the weakest source-destination pair.

A pseudo-code formulation of the whole relay set selection method is provided in Algorithm 4. Moreover, we provide illustrations for three sample networks in Fig. 2.10.

In the following, we apply the algorithm to several network topologies and numerically evaluate the corresponding diversity performance. We start out with a three-hop

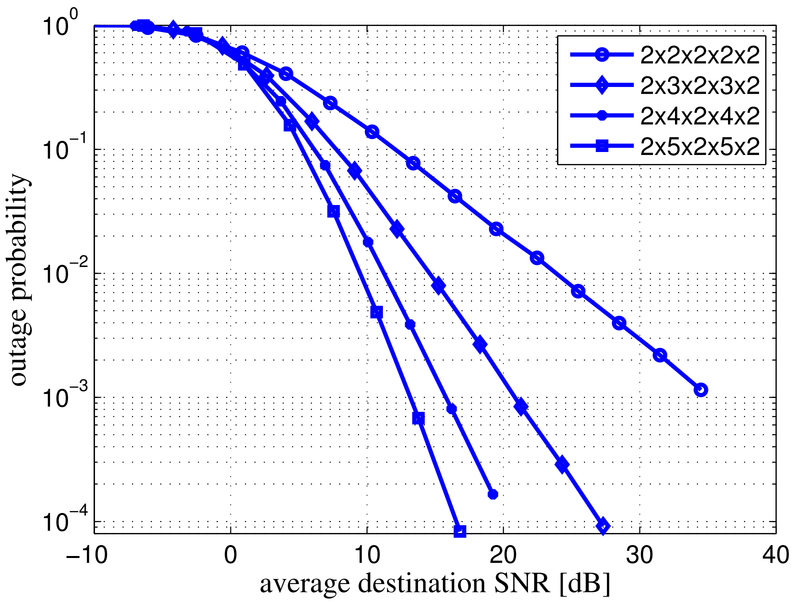


Figure 2.12.: Outage probability vs. SNR for a constant code rate $R = 1$ and various four-hop networks. The pre-log is set to $p^{-1} = 1$.

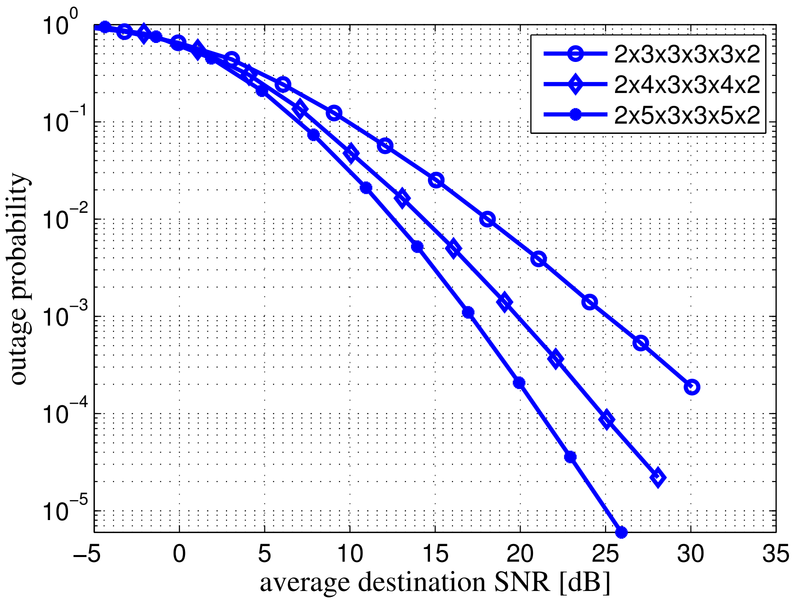


Figure 2.13.: Outage probability vs. SNR for a constant code rate $R = 1$ and various five-hop networks. The pre-log is set to $p^{-1} = 1$.

network of topology $2 \times n_{\mathcal{R}} \times n_{\mathcal{R}} \times 2$, $n_{\mathcal{R}} \in \{2, 3, 4, 5\}$. The upper-bound on diversity at $r = 0$ is $\bar{d}(0) = n_{\mathcal{R}} - 1$. For the case $n_{\mathcal{R}} = 5$, the relay sets that are tested in the four test cycles of the relay set selection algorithm are depicted in Fig. 2.10a). The outage probability versus SNR curves are shown in Fig. 2.11. We observe, that the asymptotic slope of the curve in log-log scale indeed tends to $-(n_{\mathcal{R}} - 1)$, which corresponds to the upper-bound.

We repeat the experiment for four-hop networks of topology $2 \times n_{\mathcal{R}} \times 2 \times n_{\mathcal{R}} \times 2$ for $n_{\mathcal{R}} \in \{2, 3, 4, 5\}$. The upper-bound on diversity at $r = 0$ is $\bar{d}(0) = n_{\mathcal{R}} - 1$. For the case $n_{\mathcal{R}} = 5$, the relay sets that are tested in the four test cycles of the relay set selection algorithm are depicted in Fig. 2.10b). The corresponding performance plots are shown in Fig. 2.12. Again, the upper-bound is $n_{\mathcal{R}} - 1$, and the slopes of the outage probability curves suggest that this upper-bound is met.

As a last example, we consider five-hop networks of topology $2 \times n_{\mathcal{R}} \times 3 \times 3 \times n_{\mathcal{R}} \times 2$ for $n_{\mathcal{R}} \in \{3, 4, 5\}$. The upper-bound on diversity at $r = 0$ is $\bar{d}(0) = n_{\mathcal{R}} - 1$. For the case $n_{\mathcal{R}} = 5$ the relay sets that are tested in the four test cycles of the relay set selection algorithm are depicted in Fig. 2.10c). The corresponding plots are shown in Fig. 2.13. Although there is a clear trend of an increasing slope for increasing $n_{\mathcal{R}}$, a clear identification of the diversity order is not possible. In order to observe the full slope $-(n_{\mathcal{R}} - 1)$, (most likely) larger values of SNR need to be considered in this case.

2.6.3.3. Short Networks ($2 \leq L \leq n - 1$)

In the previous subsection, we have assumed that the network consists of n relay stages at least. For such networks, it was possible to identify $\bar{d}(0)$ relay sets that in each hop differ at least in a single relay node (either on transmit or receive side). This is not feasible, when the network has less than n relay stages. Consider the extreme case of a $3 \times 4 \times 4 \times 3$-network. The DMT upper-bound for this network evaluates to $\bar{d}(0) = 2$. Since all relay nodes in both relay stages are indispensable for zero-forcing the network, relay set selection is not an option for meeting the upper-bound. However, we can resort to another means for inducing diversity, namely the multiple solutions to the system of equations (2.20) and inequations (2.21). Our approach is thus to test any two (out of the twelve, cf. Tab. 2.1) different solutions with respect to the maximal achievable rate of the weakest of the three source-destination pairs, and to choose the best of these solutions for data transmission. A respective numerical experiment shows that the outage probability versus SNR curve in log-log scale (not shown) exhibits slope -2 for large values of SNR.

We could imagine in fact (and all our numerical performance results support this conjecture) that the above discussed means for inducing diversity, namely relay set selection and selection from finite sets of solutions to (2.20) (or combinations of both) suffice for meeting the upper-bound on the DMT curve, whenever $L \geq 2$ (there is only a single solution in projective space for $L = 1$) and distributed zero-forcing is feasible.

2.7. Concluding Remarks

We have generalized the concept of distributed zero-forcing in interference networks from two-hop networks to networks with an arbitrary number of hops. Interestingly, our conjecture on the topology requirements for the feasibility of distributed zero-forcing suggests that the required number of relay nodes per stage decreases significantly with increasing numbers of hops, and is given by n for sufficiently long networks. The second major insight, which is conveyed by this work, is the fact that the lack of cooperation among relay nodes within the same stage appears not to result into a loss of performance in terms of the DMT that is achievable by coherent amplify & forward schemes.

3. Capacity Scaling of "Long" Multihop MIMO Multiple Access Networks

3.1. Introduction & Related Work

Consider wireless transmission from n transmit antennas to an n-antenna receive terminal over a random static and frequency-flat channel. Assume that each transmit antenna transmits with power P/n, where P corresponds to the sum-transmit power. If the channel coefficients between all pairs of transmit and receive antennas are i.i.d. random variables with zero mean and nonzero variance, (sum-)capacity scales linearly in n almost surely. This result does not depend on whether transmit antennas can cooperate or not [6], since the capacity of an $n \times n$ point-to-point MIMO channel with white transmit covariance matrix, coincides with the sum-capacity of an MIMO multiple access channel with n single-antenna transmit terminals and an n-antenna receive terminal [67].

Envision the scenario that the transmit antennas are shadowed from the receive terminal. Wireless connectivity can then be sustained through the installation of properly positioned intermediate nodes that relay the source signals to the destination via multiple hops (see Fig. 3.1). We refer to this modified network as multihop MIMO multiple access network. If the number of antennas in each relay stage, $n_{\mathcal{R}}$, grows linearly with n, then also the sum-capacity of the network scales linearly in n for any fixed number of relay stages, L. This result holds even for non-cooperative relay stages[1] [24]. This establishes the generalization of the result that transmit antenna cooperation is not

[1]The term non-cooperative relay stage refers in this work to the scenario that joint processing of receive signals of antennas within a stage is disabled, i.e., each relay antenna is associated with a single-antenna node. Likewise, a cooperative relay stage can be viewed as a single multi-antenna node.

crucial in multi-antenna single-hop networks, *neither source nor relay antenna cooperation is crucial for linear sum-capacity scaling in multi-antenna multihop networks.*

The above statement says nothing about the asymptotic constant of proportionality c_L^{XF} that fulfills almost surely

$$c_L^{\mathrm{XF}} = \lim_{n \to \infty} \frac{R_L^{\mathrm{XF}}}{n}, \tag{3.1}$$

where R_L^{XF} denotes the supremum of the set of sum-rates that are achievable through a certain relaying scheme "XF" in an $L+1$-hop network, except for the fact that it is strictly positive for every L. In particular, it does not preclude the scenario that $c_L^{\mathrm{XF}} \to 0$ as $L \to \infty$.

Accordingly, the central question of this work is, how c_L^{XF} evolves[1] for increasing L under various relaying strategies. The provided answers reveal fundamental differences in this asymptotic behavior with respect to L not only between cooperative and non-cooperative relaying strategies, but also between different non-cooperative strategies. Specifically, the following four relaying strategies are investigated:

- *Decode & forward* (DF): For this strategy, full cooperation among antennas within relay stages is assumed. Relay stages can then decode message(s) from their preceding stage efficiently, and re-encode them prior to the forwarding. Thus, messages are regenerated in every stage of the network and propagate from stage to stage until they reach the destination. This strategy is optimal in multihop networks according to the cut-set bound (or the data-processing inequality) [20].
- *Pure quantize & forward* (QF): This strategy can be executed in a completely decentralized fashion, i.e., without any relay antenna cooperation. The receive signal of each relay antenna is quantized. The index of the quantization codeword is then encoded and forwarded. For decoding, the destination successively recovers the quantized relay receive signals of each stage, until it can decode the source messages based on the quantized receive signals of the first relay stage.
- *Quantize & forward with Slepian & Wolf compression* (CF): The above quantize & forward strategy QF can be enhanced through Slepian & Wolf compression [68] in each relay stage. However, this extra step requires additional dissemination of CSI.

[1] In this work, the focus is not on the pre-log factor that is obviously incurred, if the source stage does not inject new signals into the network in every time slot. Accordingly, it is not taken into account in R_L^{XF} and set to one.

- *Amplify & forward*: (AF) This strategy operates in a completely decentralized fashion, too. The receive signal of each relay antenna is amplified by a constant gain factor prior to transmission to the succeeding relay or destination stage. Thus, end-to-end transmission from source to destination stage occurs over an equivalent single-hop channel. The strategy is particularly appealing due to its simplicity and low complexity.

In order to ensure a fair comparison, the sequence $(c_L^{\mathrm{XF}})_{L=0}^{\infty}$ must be considered together with a second sequence $(P_L)_{L=0}^{\infty}$, whose Lth element corresponds to the per-stage transmit power that is allocated to the source antenna stage and each of the relay stages in an $L+1$-hop network.

The results of this work are outlined as follows (see also Tab. 3.1):

- Decode & forward: For $n_{\mathcal{R}} = n$, c_L^{DF} is constant with respect to L, if also P_L is constant with respect to L.
- Quantize & forward with Slepian & Wolf compression: For $n_{\mathcal{R}} = n$, there exists a linearly increasing sequence $(P_L)_{L=0}^{\infty}$, such that c_L^{CF} is constant with respect to L. Moreover, for every $P_L \propto L$, c_L converges to a positive constant as $L \to \infty$. This is the best among the non-cooperative strategies.
- Amplify & forward: The convergence $c_L^{\mathrm{AF}} \to 0$ as $L \to \infty$ is avoided for $P_L \propto L$, if and only if the ratio $\beta = n_{\mathcal{R}}/n$ grows at least linearly with L.
- Pure quantize & forward: For $n_{\mathcal{R}} = n$, there exists an exponentially increasing sequence $(P_L)_{L=0}^{\infty}$, such that c_L^{CF} is constant with respect to L. Moreover, for every $P_L \propto \exp L$, c_L converges to a positive constant as $L \to \infty$.

Remarks: It is not surprising, that a constant P_L does not suffice for a constant c_L^{XF}, if a non-cooperative strategy is applied. This is an immediate consequence of the inherent noise accumulation, which needs to be compensated by an increased transmit power. Moreover, the obtained results do not allow for a fair comparison between amplify & forward and pure quantize & forward. No statement is made about c_L^{AF} for $n_{\mathcal{R}} = n$ and an exponential growth of P_L. Vice versa, no statement is made on how c_L^{QF} can benefit from more than n relay nodes per stage.

Related Work: Multihop MIMO multiple access networks currently receive a lot of attention in the wireless research community. The work [24] establishes linear sum-capacity scaling in n (the number of antennas per stage) of amplify & forward multihop MIMO multiple access networks for finite numbers of hops. This work is a generalization of a result on two-hop MIMO multiple access networks in [69, 70]. The work [25]

	P_L	c_L
DF ($n_{\mathcal{R}} = n$)	$P_L = \text{const}$	$\lim_{L\to\infty} c_L^{\text{DF}} = \text{const} > 0$
CF ($n_{\mathcal{R}} = n$)	$P_L \propto L$	$\lim_{L\to\infty} c_L^{\text{CF}} = \text{const} > 0$
QF ($n_{\mathcal{R}} = n$)	$\log P_L \propto L$	$\lim_{L\to\infty} c_L^{\text{QF}} = \text{const} > 0$
AF ($n_{\mathcal{R}}/n = \beta$)	$P_L \propto L$	$\lim_{L\to\infty} c_L^{\text{AF}} = \begin{cases} \text{const} > 0, \text{if} L \in \mathcal{O}(\beta) \\ 0, \text{if} L \in \Omega(\beta^{1+\varepsilon}), \varepsilon > 0 \end{cases}$

Table 3.1.: Summary of results.

considers a system that corresponds to an amplify & forward multihop MIMO multiple access network with noiseless relays. For this network, it is shown that the asymptotic constant of proportionality between sum-capacity and n tends to zero as the number of hops grows large. The work [71] studies the same setting for a finite number of hops with correlated fading and cooperative relay antennas that not only amplify, but also linearly combine the receive signals within a stage. In contrast to the above works, the work [72] studies multihop MIMO multiple access networks for finite n in terms of the achievable spatial multiplexing gain (degrees of freedom) in the limit of infinitely many hops. Moreover, the works [73] (decode & forward) and [42–44] (amplify & forward) study the DMT of multihop MIMO networks under various assumptions on the level of cooperation in the source and relay stages.

Organization of the chapter: Section 3.2 introduces the signal model and the applied multihop communication protocol. Section 3.3 presents in detail the relaying strategies that are considered. Section 3.4 constitutes the core of this work and provides four theorems that characterize the capacity scaling of each relaying strategy. Finally, Section 3.5 provides concluding remarks.

In the following, the standard $\mathcal{O}(\cdot), \Omega(\cdot), \Theta(\cdot)$ notation is used for the characterization of the asymptotic behavior of some function $f(\cdot)$ according to

$$\begin{aligned}
&f(n) \in \mathcal{O}(g(n)), \text{ if } \exists M, n_0 > 0 : M|g(n)| > |f(n)|, \forall n \geq n_0, \\
&f(n) \in \Omega(g(n)), \text{ if } \exists M, n_0 > 0 : M|g(n)| < |f(n)|, \forall n > n_0, \\
&f(n) \in \Theta(g(n)), \text{ if } f(n) \in \mathcal{O}(g(n)) \text{ and } f(n) \in \Omega(g(n)).
\end{aligned}$$

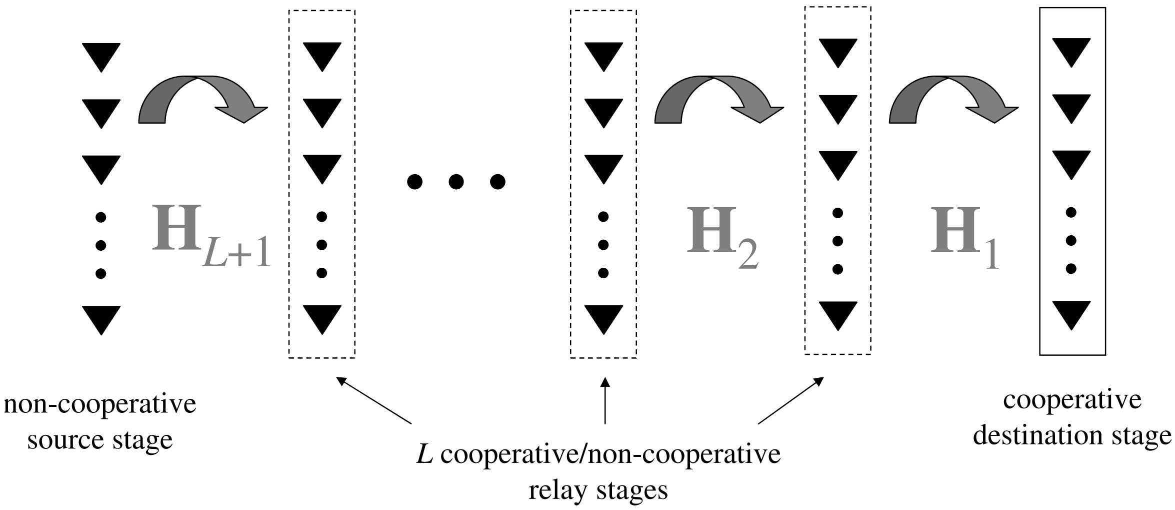


Figure 3.1.: $n_{\mathcal{S}}$ non-cooperating source antennas transmit to a destination terminal with $n_{\mathcal{D}}$ antennas via L stages of $n_{\mathcal{R}}$ relay antennas. Relay antennas are cooperative or non-cooperative depending on the relaying strategy.

3.2. Signal Model & Communication Protocol

A stage of $n_{\mathcal{S}}$ single-antenna source nodes $\mathcal{S}$ aims to transmit data to a destination node that has access to the receive signals of a cluster $\mathcal{D}$ of $n_{\mathcal{D}}$ antennas. Communication is assisted by L relay stages of $n_{\mathcal{R}}$ antennas each (see Fig. 3.1). Two cases are distinguished:

- *Fully cooperative relay stages:* All relay antennas in a stage are connected to a central node. This central node has access to the receive signals of all antennas in the stage, and, based on this knowledge, determines the transmit signals of the relay antennas.
- *Non-cooperative relay stages:* Each relay antenna corresponds to a single node that has to determine its transmit signal solely based on the knowledge of its own receive signal.

The relay stages are labeled by $\mathcal{R}_l$, $l \in \{1, \ldots, L\}$. Moreover, the kth antenna in source, relay and destination stage is labeled by S_k, $\mathrm{R}_k^{(l)}$ and D_k, respectively. The network employs a multihop protocol. More precisely, source signals traverse all relay stages in *descending*[2] order of indexes, i.e., $\mathcal{R}_L, \mathcal{R}_{L-1}, \ldots, \mathcal{R}_1$, before they are received by the destination stage. Transmission is divided into $L+1$ time slots, one dedicated to each hop, of N symbol durations each. That is, the transmissions of the stages are

[2]Note that the order is reversed in comparison to Chapter 2 for notational convenience.

orthogonal in time. Specifically,

- stage $\mathcal{S}$ transmits to stage $\mathcal{R}_L$ in time slot $j = 1$,
- stage $\mathcal{R}_{L-j+2}$ transmits to stage $\mathcal{R}_{L-j+1}$ in time slot j, where $j \in \{2, \dots, L\}$,
- stage $\mathcal{R}_1$ transmits to stage $\mathcal{D}$ in time slot $j = L + 1$.

Channels between any two antennas are quasi-static and frequency-flat over the bandwidth of interest. The channel coefficient that corresponds to the link between transmit node A and receive node B is denoted by h_{BA}. With this notation, the channel matrices are written as

$$\mathbf{H}_l = \begin{cases} \left(h_{\mathrm{R}_k^{(L)}\mathrm{S}_{k'}}\right)_{k=1,\dots,n_{\mathcal{R}},k'=1,\dots,n_{\mathcal{S}}}, & \text{if } l = L+1, \\ \left(h_{\mathrm{R}_k^{(l)}\mathrm{R}_{k'}^{(l-1)}}\right)_{k=1,\dots,n_{\mathcal{R}},k'=1,\dots,n_{\mathrm{R}}}, & \text{if } l \in \{2,\dots,L\}, \\ \left(h_{\mathrm{D}_k\mathrm{R}_{k'}^{(1)}}\right)_{k=1,\dots,n_{\mathcal{D}},k'=1,\dots,n_{\mathcal{R}}}, & \text{if } l = 1. \end{cases} \tag{3.2}$$

Furthermore, the sequence of signals that is transmitted by antenna A in its dedicated transmit time slot is denoted by $(x_{\mathrm{A}}^{(i)})_{i=1}^N$. The vector of transmit signals of antennas in a stage $\mathcal{A} = \{\mathrm{A}_1, \dots, \mathrm{A}_{|\mathcal{A}|}\}$ is denoted by

$$\mathbf{x}_{\mathcal{A}}^{(i)} = \left(x_{\mathrm{A}_1}^{(i)}, \dots, x_{\mathrm{A}_{|\mathcal{A}|}}^{(i)}\right)^T. \tag{3.3}$$

Likewise, the sequences of receive and additive noise signals, as they are observed by antenna B in its dedicated receive time slot, are denoted by $(y_{\mathrm{B}}^{(i)})_{i=1}^N$ and $(w_{\mathrm{B}}^{(i)})_{i=1}^N$, respectively. The receive signal and noise vectors of a stage $\mathcal{B} = \{\mathrm{B}_1, \dots, \mathrm{B}_{|\mathcal{B}|}\}$ are denoted by

$$\mathbf{y}_{\mathcal{B}}^{(i)} = \left(y_{\mathrm{B}_1}^{(i)}, \dots, y_{\mathrm{B}_{|\mathcal{B}|}}^{(i)}\right)^T, \tag{3.4}$$

$$\mathbf{w}_{\mathcal{B}}^{(i)} = \left(w_{\mathrm{B}_1}^{(i)}, \dots, w_{\mathrm{B}_{|\mathcal{B}|}}^{(i)}\right)^T. \tag{3.5}$$

The transmission in time slot j is thus described by the IO relation

$$\mathbf{y}_{\mathcal{R}_L}^{(i)} = \mathbf{H}_{L+1}\mathbf{x}_{\mathcal{S}}^{(i)} + \mathbf{w}_{\mathcal{R}_L}^{(i)}, \text{ if } j = 1, \tag{3.6}$$

$$\mathbf{y}_{\mathcal{R}_{L-j+1}}^{(i)} = \mathbf{H}_{L-j+2}\mathbf{x}_{\mathcal{R}_{L-j+2}}^{(i)} + \mathbf{w}_{\mathcal{R}_{L-j+1}}^{(i)}, \text{ if } j \in \{2, \dots, L\}, \tag{3.7}$$

$$\mathbf{y}_{\mathcal{D}}^{(i)} = \mathbf{H}_1\mathbf{x}_{\mathcal{R}_1}^{(i)} + \mathbf{w}_{\mathcal{D}}^{(i)}, \text{ if } j = L+1. \tag{3.8}$$

Note that transmit antennas within the same stage are assumed to be symbol-synchronized.

The elements of the vectors $\mathbf{w}_{\mathcal{R}_l}^{(i)}$, $l \in \{1, \ldots, L\}$, and $\mathbf{w}_{\mathcal{D}}^{(i)}$ represent the thermal noise that is introduced at the receiver front-ends. They are i.i.d. (both in space and time) CSCG random variables of variance σ_{w}^2. The elements of the channel matrices $\mathbf{H}_l$, $l \in \{1, \ldots, L+1\}$, are i.i.d. random variables that follow an arbitrary distribution with zero-mean, unit variance and bounded fourth moment.

The transmit signals of all antennas in source and relay stages are subject to an average power constraint[3]. Specifically, source and relay signals must fulfill almost surely

$$P_{\mathrm{S}_k} \triangleq \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} \left| x_{\mathrm{S}_k}^{(i)} \right|^2 \leq P_L / n_{\mathcal{S}}, \text{ for all } k \in \{1, \ldots, n_{\mathcal{S}}\}, \tag{3.9}$$

$$P_{\mathrm{R}_k^{(l)}} \triangleq \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} \left| x_{\mathrm{R}_k^{(l)}}^{(i)} \right|^2 \leq P_L / n_{\mathcal{R}}, \text{ for all } l \in \{1, \ldots, L\}, k \in \{1, \ldots, n_{\mathcal{R}}\}, \tag{3.10}$$

where P_L corresponds to the sum-power that is transmitted by each stage. The parameter P_L will be treated as a sequence in L.

It remains to specify, how each relay stage determines its transmit signals from its receive signals, i.e., a map

$$g_{\mathcal{R}_l} : \mathbb{C}^{N \times n_{\mathcal{R}}} \longrightarrow \mathcal{C}_{\mathcal{R}_l} : \left(\mathbf{y}_{\mathcal{R}_l}^{(i)}\right)_{i=1}^{N} \longrightarrow \left(\mathbf{x}_{\mathcal{R}_l}^{(i)}\right)_{i=1}^{N}, \tag{3.11}$$

where $\mathcal{C}_{\mathcal{R}_l}$ denotes the set of transmit signal vector sequences of relay stage $\mathcal{R}_l$. In the case of non-cooperative relay stages, the transmit signal sequence of each antenna must solely depend on the corresponding receive signal sequence of the antenna. That is, the map $g_{\mathcal{R}_l}$ decouples into the following maps:

$$g_{\mathrm{R}_k^{(l)}} : \mathbb{C}^{N} \longrightarrow \mathcal{C}_{\mathrm{R}_k^{(l)}} : \left(y_{\mathrm{R}_k^{(l)}}^{(i)}\right)_{i=1}^{N} \longrightarrow \left(x_{\mathrm{R}_k^{(l)}}^{(i)}\right)_{i=1}^{N}, k \in \{1, \ldots, n_{\mathcal{R}}\}, \tag{3.12}$$

where $\mathcal{C}_{\mathrm{R}_k^{(l)}}$ denotes the set of transmit signals of relay $\mathrm{R}_k^{(l)}$. Three fundamentally different relaying techniques are investigated: (i) decode & forward, (ii) quantize & forward (with an optional Slepian & Wolf compression) and (iii) amplify & forward. While

[3] This power constraint will be slightly relaxed in the case of amplify & forward relaying (see Section 3.3.3).

	Relay Stages	Destination Node
DF	preceding hop channel matrix	preceding hop channel matrix
CF	receive signal powers	all channel matrices
QF	receive signal powers	all channel matrices
AF	receive signal powers	all channel matrices

Table 3.2.: Receive Channel State Information

	Source Nodes	Relay Stages
DF	channel code rates	channel code rates
CF	channel code rates	channel code and compression rates
QF	channel code rates	channel code rates
AF	channel code rates	no requirements

Table 3.3.: Rate Feedback

the decode & forward strategy is optimal in terms of sum-capacity, it requires fully cooperative relay stages. The other forwarding strategies are known to be suboptimal, but do not require cooperative relay stages. For all relaying strategies, transmitting nodes are assumed not to possess any transmit CSI. The amount of receive CSI in the relay stages and at the destination node is specified in Table 3.2 for each scheme. Moreover, transmitting nodes require certain rate feedback, which is provided through perfect feedback links from the destination. This rate feedback is summarized in Table 3.3 for each scheme. The three relaying techniques and corresponding achievable rates are revisited in the context of the considered multihop network in the next section.

3.3. Multihop Relaying Techniques

In the following, each source node S_k chooses randomly according to a uniform distribution a message m_{S_k} out of the message set $\mathcal{M}_{\mathrm{S}_k}$ with $2^{NR_k^{(0)}}$ messages for transmission. The channel codebook of node S_k has rate $R_k^{(0)}$ and is denoted by $\mathcal{C}_{\mathrm{S}_k}$. Furthermore, for each node S_k an encoding function is defined:

$$f_{\mathrm{S}_k} : \mathcal{M}_{\mathrm{S}_k} \longrightarrow \mathcal{C}_{\mathrm{S}_k} : m_{\mathrm{S}_k} \longrightarrow \left(x_{\mathrm{S}_k}^{(i)}\right)_{i=1}^{N}. \tag{3.13}$$

3.3.1. Decode & Forward Relaying

Although decode & forward relaying can be applied to networks with arbitrary $n_{\mathcal{S}}$, $n_{\mathcal{R}}$ and $n_{\mathcal{D}}$, attention is restricted to the case $n_{\mathcal{S}} = n_{\mathcal{R}} = n_{\mathcal{D}} \triangleq n$ in this work. Decode &

forward relaying refers to the technique of decoding and re-encoding messages in each relay stage. For this approach, fully cooperative relay stages are assumed. That is, the antennas in each relay stage are connected to a central node. In order to enable coherent decoding in all relay stages and at the destination node, each relay stage as well as the destination node is assumed to possess perfect CSI of its preceding hop. The maps (3.11) can be split into two parts each, i.e., $g_{\mathcal{R}_l} = \hat{g}_{\mathcal{R}_l} \circ \tilde{g}_{\mathcal{R}_l}$. First, the $n_{\mathcal{S}}$ messages — one corresponding to each source node — as transmitted by the preceding stage are decoded based on the observed sequence $\left(\mathbf{y}_{\mathcal{R}_l}^{(i)}\right)_{i=1}^{N}$ and the knowledge of the channel matrix $\mathbf{H}_{l+1}$. The corresponding decoding function is

$$\tilde{g}_{\mathcal{R}_l} : \mathbb{C}^{N \times n_{\mathcal{R}}} \longrightarrow \mathcal{M}_{\mathrm{S}_1} \times \ldots \times \mathcal{M}_{\mathrm{S}_{n_{\mathcal{S}}}} : \left(\mathbf{y}_{\mathcal{R}_l}^{(i)}\right)_{i=1}^{N} \longrightarrow \left(\hat{m}_{\mathrm{S}_1}^{(l)}, \ldots, \hat{m}_{\mathrm{S}_{n_{\mathcal{S}}}}^{(l)}\right). \tag{3.14}$$

Then, the decoded messages are re-encoded:

$$\hat{g}_{\mathcal{R}_l} : \mathcal{M}_{\mathrm{S}_1} \times \ldots \times \mathcal{M}_{\mathrm{S}_{n_{\mathcal{S}}}} \longrightarrow \mathcal{C}_{\mathcal{R}_l} : \left(\hat{m}_{\mathrm{S}_1}^{(l)}, \ldots, \hat{m}_{\mathrm{S}_{n_{\mathcal{S}}}}^{(l)}\right) \longrightarrow \left(\mathbf{x}_{\mathcal{R}_l}^{(i)}\right)_{i=1}^{N}, \tag{3.15}$$

where $\mathcal{C}_{\mathcal{R}_l}$ denotes the channel codebook of relay stage $\mathcal{R}_l$. The rate of this channel codebook is given by $R_l = \sum_{k=1}^{n} R_k^{(0)} \triangleq R_{\mathcal{S}}^{\Sigma}$.

A sum-rate $R_{\mathcal{S}}^{\Sigma}$ is achievable with the decode & forward strategy, if it is supported by each individual hop. For the first hop between the source stage $\mathcal{S}$ and the fully cooperative relay stage $\mathcal{R}_L$, transmission occurs over an equivalent uplink channel with n single-antenna transmit terminals and an n-antenna receive terminal. A sum-rate $R_{\mathcal{S}}^{\Sigma}$ is achievable over the first hop channel [67], if

$$R_{\mathcal{S}}^{\Sigma} < \log\det\left(\mathbf{I}_n + \frac{P_L}{n \cdot \sigma_{\mathrm{w}}^2} \mathbf{H}_{L+1} \mathbf{H}_{L+1}^H\right). \tag{3.16}$$

Note that this work focuses on the achievable sum-rate $R_{\mathcal{S}}^{\Sigma}$ and does not care about the individual rates $R_k^{(0)}$. For all following hops, transmission occurs over equivalent point-to-point channels, since both transmit and receive stage are fully cooperative. For the following hops, relay stage $\mathcal{R}_l$, $l \in \{1, \ldots, L\}$, can transmit reliably to its succeeding stage at a rate R_l, if [6]

$$R_l < \log\det\left(\mathbf{I}_n + \frac{P_L}{n \cdot \sigma_{\mathrm{w}}^2} \mathbf{H}_l \mathbf{H}_l^H\right). \tag{3.17}$$

Here, a spatially white transmit covariance matrix is used due to the absence of transmit CSI. In order to achieve $R_{\mathcal{S}}^{\Sigma}$ over the end-to-end channel, (3.16) and (3.17) need to be fulfilled for $R_{\mathcal{S}}^{\Sigma} = R_L = \ldots = R_1$. This corresponds to the condition

$$R_{\mathcal{S}}^{\Sigma} < \min_{l \in \{1,\ldots,L+1\}} \log\det\left(\mathbf{I}_n + \frac{P_L}{n \cdot \sigma_{\mathrm{w}}^2}\mathbf{H}_l\mathbf{H}_l^H\right) \triangleq R_L^{\mathrm{DF}}. \tag{3.18}$$

3.3.2. Quantization Based Relaying

Although quantization based relaying can be applied to networks with arbitrary $n_{\mathcal{S}}$, $n_{\mathcal{R}}$ and $n_{\mathcal{D}}$, attention is restricted to the case $n_{\mathcal{S}} = n_{\mathcal{R}} = n_{\mathcal{D}} \triangleq n$ in this work. In this non-cooperative relaying scheme, the maps $g_{\mathrm{R}_k^{(l)}}$, $k \in \{1,\ldots,n\}$, as defined in (3.12), are implemented in two steps, i.e., $g_{\mathrm{R}_k^{(l)}} = \hat{g}_{\mathrm{R}_k^{(l)}} \circ \tilde{g}_{\mathrm{R}_k^{(l)}}$. In a first step, each relay $\mathrm{R}_k^{(l)}$ quantizes its receive sequence, i.e., approximates its receive signal by the sequence $\left(\hat{y}_{\mathrm{R}_k^{(l)}}^{(i)}\right)_{i=1}^{N} \in \tilde{\mathcal{C}}_{\mathrm{R}_k^{(l)}}$, where $\tilde{\mathcal{C}}_{\mathrm{R}_k^{(l)}}$ denotes the quantization codebook of the relay. The sequence of additive quantization noise is defined as

$$\left(q_{\mathrm{R}_k^{(l)}}^{(i)}\right)_{i=1}^{N} = \left(\hat{y}_{\mathrm{R}_k^{(l)}}^{(i)} - y_{\mathrm{R}_k^{(l)}}^{(i)}\right)_{i=1}^{N}, \tag{3.19}$$

and the quantization noise vector of stage $\mathcal{R}_l$ is given by

$$\mathbf{q}_{\mathcal{R}_l}^{(i)} = \left(q_{\mathrm{R}_1^{(l)}}^{(i)}, \ldots, q_{\mathrm{R}_n^{(l)}}^{(i)}\right)^T. \tag{3.20}$$

The corresponding map is

$$\tilde{g}_{\mathrm{R}_k^{(l)}} : \mathbb{C}^N \longrightarrow \tilde{\mathcal{C}}_{\mathrm{R}_k^{(l)}} : \left(y_{\mathrm{R}_k^{(l)}}^{(i)}\right)_{i=1}^{N} \longrightarrow \left(\hat{y}_{\mathrm{R}_k^{(l)}}^{(i)}\right)_{i=1}^{N}, \tag{3.21}$$

where the quantization codebook $\tilde{\mathcal{C}}_{\mathrm{R}_k^{(l)}}$ has rate $\tilde{R}_k^{(l)}$.

In a second step, the index of the quantized sequence is encoded by the channel coder, i.e., mapped onto a codeword of the channel codebook $\mathcal{C}_{\mathrm{R}_k^{(l)}}$ whose rate is larger than or equal to $\tilde{R}_k^{(l)}$:

$$\hat{g}_{\mathrm{R}_k^{(l)}} : \tilde{\mathcal{C}}_{\mathrm{R}_k^{(l)}} \longrightarrow \mathcal{C}_{\mathrm{R}_k^{(l)}} : \left(\hat{y}_{\mathrm{R}_k^{(l)}}^{(i)}\right)_{i=1}^{N} \longrightarrow \left(x_{\mathrm{R}_k^{(l)}}^{(i)}\right)_{i=1}^{N}. \tag{3.22}$$

The decoding in the destination stage is then performed in a successive fashion. In a first step, the messages sent by the relay stage $\mathcal{R}_1$ are decoded based on the sequence of receive vectors $(\mathbf{y}_{\mathcal{D}}^{(i)})_{i=1}^{N}$ and the knowledge of $\mathbf{H}_1$. Note that the transmission from

$\mathcal{R}_1$ to $\mathcal{D}$ occurs over an uplink channel with n single-antenna transmit terminals and a receive terminal with n antennas. Since the functions $g_{\mathrm{R}_k^{(1)}}$ are not injective, and thus not invertible, it is impossible to obtain a perfect reconstruction of the sequence of receive vectors $(\mathbf{y}_{\mathcal{R}_1}^{(i)})_{i=1}^N$. This ambiguity is accounted for by the sequence of additive quantization noise vectors $(\mathbf{q}_{\mathcal{R}_1}^{(i)})_{i=1}^N$. The ith element in the sequence of quantized receive vectors $\hat{\mathbf{y}}_{\mathcal{R}_1}^{(i)}$ is then written in terms of the ith element in the corresponding sequence of transmit vectors $\mathbf{x}_{\mathcal{R}_2}^{(i)}$ as

$$\hat{\mathbf{y}}_{\mathcal{R}_1}^{(i)} = \mathbf{H}_2 \mathbf{x}_{\mathcal{R}_2}^{(i)} + \mathbf{w}_{\mathcal{R}_1}^{(i)} + \mathbf{q}_{\mathcal{R}_1}^{(i)}. \tag{3.23}$$

With the knowledge of the sequence $(\hat{\mathbf{y}}_{\mathcal{R}_1}^{(i)})_{i=1}^N$ and $\mathbf{H}_2$, the destination proceeds with decoding the messages of the nodes in $\mathcal{R}_2$. These messages can be considered as being transmitted over a virtual uplink with n single-antenna transmit terminals and a receive terminal with n antennas.

Proceeding this way iteratively allows for tracing back through the relay chain stage by stage based on the sequences of quantized receive vectors and the knowledge of the respective channel matrix. In the lth iteration, the decoder obtains the sequence of quantized receive vectors $(\hat{\mathbf{y}}_{\mathcal{R}_l}^{(i)})_{i=1}^N$ by decoding the messages of $\mathcal{R}_l$. These messages are transmitted over a virtual uplink channel with n single-antenna transmit terminals and a receive terminal with n antennas. The effective IO relation between $\mathcal{R}_{l+1}$ and $\mathcal{D}$ can thus be written as

$$\hat{\mathbf{y}}_{\mathcal{R}_l}^{(i)} = \mathbf{H}_{l+1} \mathbf{x}_{\mathcal{R}_{l+1}}^{(i)} + \mathbf{w}_{\mathcal{R}_l}^{(i)} + \mathbf{q}_{\mathcal{R}_l}^{(i)}, \tag{3.24}$$

where $\mathbf{q}_{\mathcal{R}_l}^{(i)}$ is the ith element in the respective sequence of quantization noise vectors. In the $(L+1)$-st iteration the decoder finally arrives at the source stage, whose messages are transmitted over a virtual uplink channel with n single-antenna transmit terminals and a receive terminal with n antennas. They are decoded based on the effective IO relation

$$\hat{\mathbf{y}}_{\mathcal{R}_L}^{(i)} = \mathbf{H}_{L+1} \mathbf{x}_{\mathcal{S}}^{(i)} + \mathbf{w}_{\mathcal{R}_L}^{(i)} + \mathbf{q}_{\mathcal{R}_L}^{(i)} \tag{3.25}$$

and the knowledge of $\mathbf{H}_{L+1}$ and $\hat{\mathbf{y}}_{\mathcal{R}_L}^{(i)}$.

The channel and quantization codebooks that are used later on render the quantization noise vectors $\mathbf{q}_{\mathcal{R}_l}^{(i)}$, $l \in \{1, \ldots, L\}$, i.i.d. (both in space and time) CSCG with variance σ_l^2. Thus, for the end-to-end channel from $\mathcal{S}$ to $\mathcal{D}$, a sum-rate rate

$R_{\mathcal{S}}^{\Sigma} = \sum_{k=1}^{n} R_k^{(0)}$, where $R_k^{(0)}$ denotes the rate of the channel codebook of S_k, is achievable, if it fulfills

$$R_{\mathcal{S}}^{\Sigma} < \log\det\left(\mathbf{I}_n + \frac{P_L}{n\cdot(\sigma_{\mathrm{w}}^2 + \sigma_L^2)} \cdot \mathbf{H}_{L+1}\mathbf{H}_{L+1}^H\right) \triangleq R_L^{\mathrm{QF/CF}}. \tag{3.26}$$

The rate of the channel codebook of relay node $\mathrm{R}_k^{(l)}$ is denoted by $R_k^{(l)}$ in the following. In this work, rates of channel codebooks are required to coincide within each stage $\mathcal{R}_l$, i.e., $R_1^{(l)} = \ldots = R_n^{(l)} \triangleq R_l$. The rate R_l will in each case determine the quantization noise variance σ_l^2. For given quantization noise variances σ_{l-1}^2, $l \in \{1, \ldots, L\}$, where $\sigma_0^2 = 0$, the rate R_l is achievable, if [6, 67]

$$|\tilde{\mathcal{R}}_l| R_l < \log\det\left(\mathbf{I}_n + \frac{P_L}{n\cdot\left(\sigma_{\mathrm{w}}^2 + \sigma_{l-1}^2\right)} \cdot (\mathbf{H}_l)_{\tilde{\mathcal{R}}_l} (\mathbf{H}_l)_{\tilde{\mathcal{R}}_l}^H\right) \text{ for all } \tilde{\mathcal{R}}_l \subseteq \mathcal{R}_l. \tag{3.27}$$

Here, $(\mathbf{H}_l)_{\tilde{\mathcal{R}}_l}$ denotes the $n \times |\tilde{\mathcal{R}}_l|$ matrix, that collects all columns of $\mathbf{H}_l$ that correspond to the nodes that are contained in $\tilde{\mathcal{R}}_l$. Moreover, the supremum of the set of rates R_l that fulfill (3.27) is denoted by $\overline{R}_l$ in the sequel.

In order to achieve these rates, the source and relay nodes need to generate their channel codebooks by choosing the entries $x_{\mathrm{S}_k}^{(i)}$ and $x_{\mathrm{R}_k^{(l)}}^{(i)}$ as independent realizations of CSCG random variable $\mathsf{x}_{\mathcal{S}}$ and $\mathsf{x}_{\mathcal{R}}$, respectively. The variances of these random variable are chosen to fulfill the average power constraints (3.9) and (3.10) with equality, i.e., $\mathsf{E}_{\mathsf{x}_{\mathcal{S}}}[|\mathsf{x}_{\mathcal{S}}|^2] = P_L/n$ and $\mathsf{E}_{\mathsf{x}_{\mathcal{R}}}[|\mathsf{x}_{\mathcal{R}}|^2] = P_L/n$.

In the following, the quantization codebooks of the relay nodes and the resulting quantization noise variances are specified. In this context, "pure quantization" and "quantization with Slepian & Wolf compression" is distinguished.

3.3.2.1. Pure Quantization

Pure quantization refers to a receive signal quantization method that does not exploit the statistical correlation among receive signals of relay nodes within the same stage. Each relay node $\mathrm{R}_k^{(l)}$ generates a quantization codebook $\tilde{\mathcal{C}}_k^{(l)}$ of rate $\tilde{R}_k^{(l)}$. The elements of the quantization noise vector $\mathbf{q}_{\mathcal{R}_l}^{(i)}$ are rendered i.i.d. (both in space and time) CSCG with variance σ_l^2 by choosing the entries of the quantization codebook as statistically independent realizations of a CSCG random variable $\hat{\mathsf{y}}$ of variance $Q_k^{(l)} + \sigma_{\mathrm{w}}^2 + \sigma_l^2$. Here, $Q_k^{(l)}$ denotes the average power of the desired part of the receive signal at node $\mathrm{R}_k^{(l)}$. A codeword is declared to be the quantization of the observed sequence, if both are jointly typical. In order to ensure that a quantization codeword for the observed sequence is found based on a joint typicality check with probability one as $N \to \infty$, the mutual information between original and quantized observation of node $\mathrm{R}_k^{(l)}$ must fulfill

$$\tilde{R}_k^{(l)} > I\left(\mathrm{y}_{\mathrm{R}_k^{(l)}}^{(i)}; \hat{\mathrm{y}}_{\mathrm{R}_k^{(l)}}^{(i)}\right) = \log\left(1 + \frac{Q_k^{(l)} + \sigma_\mathrm{w}^2}{\sigma_l^2}\right), \tag{3.28}$$

or equivalently

$$\sigma_l^2 > \frac{Q_k^{(l)} + \sigma_\mathrm{w}^2}{2^{\tilde{R}_k^{(l)}} - 1}. \tag{3.29}$$

This inequality must be fulfilled for all nodes in $\mathcal{R}_l$. Moreover, the rates of the quantization codebooks $\tilde{R}_k^{(l)}$ must be smaller than or equal to the rate of the channel codebook, R_l. For a given $R_l < \overline{R}_l$ a quantization noise variance σ_l^2 is achievable, if it fulfills

$$\sigma_l^2 > \max_{\mathrm{R}_k^{(l)} \in \mathcal{R}_l} \frac{Q_k^{(l)} + \sigma_\mathrm{w}^2}{2^{R_l} - 1}. \tag{3.30}$$

3.3.2.2. Quantization with Slepian & Wolf Compression

A more efficient relaying method performs a Slepian & Wolf compression of the receive signal upon quantization. Thus, the correlation in the receive signals of relay nodes within the same stage can be efficiently exploited. As in the case of pure quantization, the elements of the quantization noise vector $\mathbf{q}_{\mathcal{R}_l}^{(i)}$ are rendered i.i.d. (both in space and time) CSCG with variance σ_l^2. The rate of the compressed quantization codebook of relay node $\mathrm{R}_k^{(l)}$ is denoted by $\tilde{R}_k^{(l)}$ in the sequel. Rates of quantization codebooks are chosen to coincide within each stage $\mathcal{R}_l$, i.e., $\tilde{R}_1^{(l)} = \ldots = \tilde{R}_n^{(l)} \triangleq \tilde{R}_l$. The compression problem at hand has been studied in [74] in the context of a two-hop setup with orthogonal second hop. The quantized observation vector can be reconstructed with probability one as $N \to \infty$, if

$$|\tilde{\mathcal{R}}_l|\tilde{R}_l > I\left(\left(\mathrm{y}_{\mathrm{R}_k^{(l)}}^{(i)}\right)_{\mathrm{R}_k^{(l)} \in \tilde{\mathcal{R}}_l}; \left(\hat{\mathrm{y}}_{\mathrm{R}_k^{(l)}}^{(i)}\right)_{\mathrm{R}_k^{(l)} \in \tilde{\mathcal{R}}_l} \middle| \left(\hat{\mathrm{y}}_{\mathrm{R}_k^{(l)}}^{(i)}\right)_{\mathrm{R}_k^{(l)} \in \tilde{\mathcal{R}}_l^C}\right) \quad \text{for all } \tilde{\mathcal{R}}_l \subseteq \mathcal{R}_l. \tag{3.31}$$

For a given quantization noise variance, the conditional mutual information expression can be written as (refer to the proof of Lemma 3 in Appendix B.1 for a derivation)

$$I\left(\left(\mathrm{y}_{\mathrm{R}_k^{(l)}}^{(i)}\right)_{\mathrm{R}_k^{(l)}\in\tilde{\mathcal{R}}_l};\left(\hat{\mathrm{y}}_{\mathrm{R}_k^{(l)}}^{(i)}\right)_{\mathrm{R}_k^{(l)}\in\tilde{\mathcal{R}}_l}\middle|\left(\hat{\mathrm{y}}_{\mathrm{R}_k^{(l)}}^{(i)}\right)_{\mathrm{R}_k^{(l)}\in\tilde{\mathcal{R}}_l^C}\right) \tag{3.32}$$

$$\begin{aligned} &= \log\det\left(\mathbf{I}_n + \frac{P_L}{(\sigma_\mathrm{w}^2+\sigma_l^2)\cdot n}\mathbf{H}_{l+1}\mathbf{H}_{l+1}^H\right) \\ &\quad - \log\det\left(\mathbf{I}_n + \frac{P_L}{(\sigma_\mathrm{w}^2+\sigma_l^2)\cdot n}\left((\mathbf{H}_{l+1})_{\tilde{\mathcal{R}}_l^C}\right)^H(\mathbf{H}_{l+1})_{\tilde{\mathcal{R}}_l^C}\right) \\ &\quad + \left|\tilde{\mathcal{R}}_l\right|\log\left(1+\frac{\sigma_\mathrm{w}^2}{\sigma_l^2}\right). \end{aligned} \tag{3.33}$$

Here, $(\mathbf{H}_l)_{\tilde{\mathcal{R}}_l^C}$ denotes the $|\tilde{\mathcal{R}}_l^C|\times n$ matrix that collects all rows of $\mathbf{H}_l$ which correspond to the nodes contained in $\tilde{\mathcal{R}}_l^C$. Again, the rate of the quantization codebook of each node is required to be smaller than or equal to the rate of the channel codebook of the node, i.e., $\tilde{R}_l \leq R_l$. For a given $R_l < \overline{R}_l$, a quantization noise variance σ_l^2 is achievable, if it fulfills (3.31) for

$$\tilde{R}_l = R_l. \tag{3.34}$$

3.3.3. Amplify & Forward Relaying

Amplify & forward refers to the technique of deriving a transmit signal from the receive signal through simple amplification. This work focuses on the case that each relay within stage $\mathcal{R}_l$ applies the same gain factor $\sqrt{\alpha/n_\mathcal{R}}$, $\alpha > 0$ and sufficiently small, such that (3.10) is fulfilled. That is, $g_{\mathrm{R}_k^{(l)}}$ as defined in (3.12) is given by

$$g_{\mathrm{R}_k^{(l)}}:\mathbb{C}^N \longrightarrow \mathbb{C}^N:\left(y_{\mathrm{R}_k^{(l)}}^{(i)}\right)_{i=1}^N \longrightarrow \left(\sqrt{\alpha/n_\mathcal{R}}\cdot y_{\mathrm{R}_k^{(l)}}^{(i)}\right)_{i=1}^N. \tag{3.35}$$

Equivalently, the relay receive and transmit vectors are related as follows:

$$\mathbf{x}_{\mathcal{R}_l}^{(i)} = \sqrt{\frac{\alpha}{n_\mathcal{R}}}\mathbf{y}_{\mathcal{R}_l}^{(i)} = \begin{cases} \sqrt{\frac{\alpha}{n_\mathcal{R}}}\left(\mathbf{H}_{l+1}\mathbf{x}_{\mathcal{R}_{l+1}}^{(i)}+\mathbf{w}_{\mathcal{R}_l}\right), & \text{if } l\in\{1,\dots,L-1\}, \\ \sqrt{\frac{\alpha}{n_\mathcal{R}}}\left(\mathbf{H}_{l+1}\mathbf{x}_{\mathcal{S}}^{(i)}+\mathbf{w}_{\mathcal{R}_1}\right), & \text{if } l = L, \end{cases} \tag{3.36}$$

where the second equality follows from (3.6) and (3.7), respectively. The effective IO relation from source to destination stage is then obtained through recursive derivation of $\mathbf{x}^{(i)}_{\mathcal{R}_{l+1}}$ from $\mathbf{x}^{(i)}_{\mathcal{R}_l}$ according to (3.36), and given by

$$\mathbf{y}^{(i)}_{\mathcal{D}} = \frac{\alpha^{L/2}}{n_{\mathcal{R}}^{L/2}}\mathbf{H}_1 \cdots \mathbf{H}_{L+1}\mathbf{x}^{(i)}_{\mathcal{S}} + \mathbf{w}^{(i)}_{\mathcal{D}} + \sum_{l=1}^{L} \frac{\alpha^{l/2}}{n_{\mathcal{R}}^{l/2}}\mathbf{H}_1 \cdots \mathbf{H}_l \mathbf{w}^{(i)}_{\mathcal{R}_l}. \tag{3.37}$$

This IO relation represents an $n_{\mathcal{S}}$-user uplink channel with $n_{\mathcal{D}}$ receive antennas. The sum-capacity for a fixed channel realization is achieved by a Gaussian codebook, i.e., the $x^{(i)}_{\mathrm{S}_{(k)}}$ are i.i.d. CSCG with variance $P_L/n_{\mathcal{S}}$. Under this input-distribution a sum-rate $R^{\Sigma}_{\mathcal{S}}$ is achievable, if it fulfills [6]

$$R^{\Sigma}_{\mathcal{S}} < \log\det\left(\mathbf{I}_{n_{\mathcal{D}}} + \mathbf{R}_s\mathbf{R}_n^{-1}\right) \triangleq R^{\mathrm{AF}}_L, \tag{3.38}$$

where $\mathbf{R}_s$ and $\mathbf{R}_n$ denote the covariance matrices of desired receive signal and accumulated noise at the destination, respectively:

$$\mathbf{R}_s = \frac{P_L\alpha^L}{n_{\mathcal{S}}n_{\mathcal{R}}^L}\mathbf{H}_1 \cdots \mathbf{H}_{L+1}\mathbf{H}^H_{L+1} \cdots \mathbf{H}^H_1, \tag{3.39}$$

$$\mathbf{R}_n = \sigma^2_{\mathrm{w}} \cdot \left(\mathbf{I}_{n_{\mathcal{D}}} + \sum_{l=1}^{L} \frac{\alpha^l}{n_{\mathcal{R}}^l}\mathbf{H}_1 \cdots \mathbf{H}_l\mathbf{H}^H_l \cdots \mathbf{H}^H_1\right). \tag{3.40}$$

In the sequel, α is chosen as $\alpha = P_L/(P_L + \sigma^2_{\mathrm{w}})$. This choice does not fulfill the power constraints (3.10) exactly, but in the following sense:

Proposition 5. *Let $\mathbf{H}_2, \ldots, \mathbf{H}_L \in \mathbb{C}^{n_{\mathcal{R}} \times n_{\mathcal{R}}}$ and $\mathbf{H}_{L+1} \in \mathbb{C}^{n_{\mathcal{R}} \times n_{\mathcal{S}}}$ be statistically independent random matrices, each with i.i.d. elements of zero-mean, unit-variance and bounded fourth moment, and fix the ratios $n_{\mathcal{R}}/n_{\mathcal{D}}$ and $n_{\mathcal{R}}/n_{\mathcal{S}}$. If $\alpha = P_L/(P_L + \sigma^2_{\mathrm{w}})$, then*

- *the sum-transmit power of each stage $\mathcal{R}_l$ fulfills*

$$\lim_{n_{\mathcal{R}} \to \infty} \sum_{k=1}^{n_{\mathcal{R}}} P_{\mathrm{R}^{(l)}_k} = P_L \text{ almost surely}, \tag{3.41}$$

- *for every fixed $\gamma < 1$ and every stage $\mathcal{R}_l$*

$$\mathsf{Pr}\left[\forall\varepsilon > 0\ \exists n_0 : \ \frac{1}{n_{\mathcal{R}}}\sum_{k=1}^{n_{\mathcal{R}}} 1\left\{\left|n_{\mathcal{R}}P_{\mathrm{R}^{(l)}_k} - P_L\right| < \varepsilon\ \forall n_{\mathcal{R}} \geq n_0\right\} > \gamma\right] = 1. \tag{3.42}$$

Remark: Note that this proposition does not imply for a given l that

$$\lim_{n_{\mathcal{R}}\to\infty} \max_{k\in\{1,\ldots,n_{\mathcal{R}}\}} \left| n_{\mathcal{R}} \cdot P_{\mathrm{R}_k^{(l)}} - P_L \right| = 0 \text{ almost surely.} \tag{3.43}$$

The proof of the proposition is provided in Appendix B.3. Since it relies on notation and concepts introduced in Section 3.4.3, it is recommended to return to this proposition and its proof later.

3.4. Capacity Scaling

This section establishes for each of the introduced relaying schemes the capacity scaling result as outlined above. Four theorems, one for each relaying scheme, are provided and proved. Decode & forward is examined in Subsection 3.4.1, quantize & forward without and with Slepian & Wolf compression in Subsections 3.4.2.1 and 3.4.2.2, and amplify & forward in Subsection 3.3.3.

3.4.1. Decode & Forward Networks

The following theorem characterizes the scaling of the supremum of achievable sum-rates in decode & forward multihop MIMO networks:

Theorem 2. *Let $\mathbf{H}_1, \ldots, \mathbf{H}_{L+1} \in \mathbb{C}^{n\times n}$ be statistically independent random channel matrices, each with i.i.d. elements of zero-mean, unit-variance and bounded fourth moment. Then, the supremum of the set of sum-rates that are achievable by the decode & forward strategy, R_L^{DF}, fulfills for all $L \in \mathbb{N}_0$*

$$\lim_{n\to\infty} \frac{1}{n} R_L^{\mathrm{DF}} = \psi\left(\frac{P_L}{\sigma_{\mathrm{w}}^2}\right) \triangleq c_L^{\mathrm{DF}} \textit{ almost surely,} \tag{3.44}$$

where

$$\psi(x) \triangleq 2\log\left(1 + x - \frac{1}{4}\left(\sqrt{4x+1} - 1\right)^2\right) - \frac{\log e}{4x}\left(\sqrt{4x+1} - 1\right)^2. \tag{3.45}$$

Note that the constant of proportionality that asymptotically relates R_L^{DF} and n is constant with respect to L, if P_L is constant with respect to L.

Proof of Theorem 2. Under the assumptions of the theorem the following holds for every $l \in \{1,\dots,L\}$ according to [75, 76] (see also [26, p. 10]):

$$\lim_{n\to\infty} \frac{1}{n} \cdot \log\det\left(\mathbf{I}_n + \frac{P_L}{n\cdot\sigma_{\mathrm{w}}^2}\mathbf{H}_l\mathbf{H}_l^H\right) = \psi\left(\frac{P_L}{\sigma_{\mathrm{w}}^2}\right) \text{ almost surely.} \tag{3.46}$$

Thus, there exists almost surely for every $\varepsilon > 0$ and arbitrary L an $n_0(L)$, such that for all $n \geq n_0(L)$

$$\max_{l\in\{1,\dots,L+1\}} \left| \frac{1}{n} \cdot \log\det\left(\mathbf{I}_n + \frac{P_L}{n\cdot\sigma_{\mathrm{w}}^2}\mathbf{H}_l\mathbf{H}_l^H\right) - \psi\left(\frac{P_L}{\sigma_{\mathrm{w}}^2}\right)\right| < \varepsilon. \tag{3.47}$$

For the supremum of the set of achievable sum-rates (3.18), this implies

$$\lim_{n\to\infty} \frac{1}{n} R_L^{\mathrm{DF}} = \lim_{n\to\infty} \min_{l\in\{1,\dots,L+1\}} \frac{1}{n} \log\det\left(\mathbf{I}_n + \frac{P_L}{n\cdot\sigma_{\mathrm{w}}^2}\mathbf{H}_l\mathbf{H}_l^H\right) = \psi\left(\frac{P_L}{\sigma_{\mathrm{w}}^2}\right) \text{ almost surely.} \tag{3.48}$$

This establishes the theorem. □

3.4.2. Quantize & Forward Networks

3.4.2.1. Quantize & Forward without Slepian & Wolf compression

The following theorem characterizes the scaling of the supremum of achievable sum-rates in quantize & forward multihop MIMO multiple access networks that do not apply Slepian & Wolf compression:

Theorem 3. *Let $\mathbf{H}_1,\dots,\mathbf{H}_{L+1} \in \mathbb{C}^{n\times n}$ be statistically independent random channel matrices, each with i.i.d. elements of zero-mean, unit-variance and bounded fourth moment. Then, the supremum of the set of sum-rates that are achievable by the quantize & forward strategy, R_L^{QF}, fulfills for all $L \in \mathbb{N}_0$*

$$\lim_{n\to\infty} \frac{1}{n} R_L^{\mathrm{QF}} = \psi\left(\mathrm{snr}_{\mathcal{D}}\right) \triangleq c_L^{\mathrm{QF}} \text{ almost surely,} \tag{3.49}$$

where $\mathrm{snr}_{\mathcal{D}} = P_L/(\sigma_L^2 + \sigma_{\mathrm{w}}^2)$, and $\psi(\cdot)$ is defined in (3.45). Moreover, the per-stage transmit power P_L that is required for rendering $\mathrm{snr}_{\mathcal{D}}$ constant with respect to L grows exponentially with L.

Note that the constant of proportionality that asymptotically relates R_L^{QF} and n is constant with respect to L, if the SNR at the destination stage, $\mathrm{snr}_{\mathcal{D}}$, is constant with respect to L.

Proof of Theorem 3. In order to prove that the supremum of the set of achievable end-to-end sum-rates scales linearly in n in the limit $L \to \infty$, it is sufficient to show that there exists a sequence $(P_L)_{L=0}^{\infty}$, such that the SNR at the destination antennas, given by $\mathrm{snr}_{\mathcal{D}} = P_L/(\sigma_{\mathrm{w}}^2 + \sigma_L^2)$, is kept constant and bounded away from zero as $L \to \infty$. Then, the supremum of the set of achievable sum-rates as given by (3.26) fulfills [75,76]

$$\lim_{n\to\infty} \frac{1}{n} \cdot R_L^{\mathrm{QF}} = \psi(\mathrm{snr}_{\mathcal{D}}) \text{ almost surely,} \tag{3.50}$$

independently of L. Since $\mathrm{snr}_{\mathcal{D}} > 0$, this limit is strictly positive.

First, a result on the MIMO uplink channel with n single-antenna sources and an n antenna destination is stated. Under the assumptions on the fading distributions of the theorem, all sources can, in the limit of large n, simultaneously achieve a fraction $1/n$ of the supremum of achievable sum-rates. More precisely, the following lemma, whose proof is given in Appendix B.1, holds:

Lemma 1. *Consider an uplink channel with n single-antenna transmit terminals with transmit power P/n each, and an n-antenna receive terminal with spatial noise covariance matrix $\sigma^2\mathbf{I}_n$. The elements of the channel matrix $\mathbf{H} \in \mathbb{C}^{n\times n}$ are distributed according to the assumptions of Theorem 3. Let*

$$\xi = \psi\left(\frac{P}{\sigma^2}\right). \tag{3.51}$$

Then, there exists almost surely for every rate $R < \xi$ an n_0, such that for all $n \geq n_0$ the rate tuple $(R^{(1)}, R^{(2)}, \ldots, R^{(n)}) = (R, R, \ldots, R)$, where $R^{(i)}$ denotes the rate of the ith transmit terminal, is achievable.

This lemma implies that in the limit of large n the set of achievable R_l is fully determined by the constraint in (3.27) that corresponds to the set $\tilde{\mathcal{R}}_l = \mathcal{R}_l$, since [75,76]

$$\lim_{n\to\infty} \frac{1}{n} \log\det\left(\mathbf{I}_n + \frac{P_L}{n(\sigma_{l-1}^2 + \sigma_{\mathrm{w}}^2)}\mathbf{H}_l\mathbf{H}_l^H\right) = \xi_l \text{ almost surely,} \tag{3.52}$$

where

$$\xi_l \triangleq \psi\left(\frac{P_L}{\sigma_{\mathrm{w}}^2 + \sigma_{l-1}^2}\right). \tag{3.53}$$

Thus, there is for every fixed L and given sequence $(\sigma_l^2)_{l=0}^{L-1}$ almost surely an $n_0(L)$, such that for all $n \geq n_0(L)$ all rates R_l, $l \in \{1, \ldots, L\}$, are achievable simultaneously, if $R_l < \xi_l$ for all l.

The next step, relies on the fact that as $n \to \infty$, the receive power of all receive antennas in the various relay stages converges almost surely to $P_L + \sigma_{\mathrm{w}}^2$. To make this precise, the following lemma, whose proof is given in Appendix B.1, is stated:

Lemma 2. *Let* $\mathbf{H} \in \mathbb{C}^{n \times n}$ *be a random matrix whose elements are distributed according to the assumptions of Theorem 3 and* $P > 0$. *Denote by* $\mathbf{h}_k^T$ *the* k*th row of* $\mathbf{H}$. *Then,*

$$\lim_{n\to\infty} \max_{k \in \{1,\ldots,n\}} \left| \frac{P}{n} \cdot \|\mathbf{h}_k^T\|^2 - P \right| = 0 \text{ almost surely.} \tag{3.54}$$

Thus, the $Q_k^{(l)}$ in (3.30) fulfill for every every fixed L

$$\lim_{n\to\infty} \max_{l \in \{1,\ldots,L\}} \max_{k \in \{1,\ldots,n\}} \left| Q_k^{(l)} - P_L \right| = 0 \text{ almost surely.} \tag{3.55}$$

From the conclusions of Lemmata 1 and 2, we infer that there is for every fixed L almost surely an $n_0(L)$, such that for all $n \geq n_0(L)$ the sequence of quantization noise variances $(\sigma_l^2)_{l=1}^{L}$ is achievable according to (3.30), if for all $l \in \{1, \ldots, L\}$

$$\sigma_l^2 > \frac{P_L + \sigma_{\mathrm{w}}^2}{2^{\xi_l} - 1} \triangleq \left(\sigma_l^{(\mathrm{inf})}\right)^2, \tag{3.56}$$

where the $\left(\sigma_l^{(\mathrm{inf})}\right)^2$, $l \in \{1, \ldots, L\}$, denote the infima of achievable noise variances. Substitution of the asymptotic suprema of achievable rates (3.53) into (3.56) yields a first order difference equation in σ_l^2 with $\sigma_0^2 = 0$. This difference equation is the starting point for the proof that any SNR value, $\mathrm{snr}_{\mathcal{D}}$, can be sustained at the destination antennas for increasing L by increasing P_L appropriately.

We apply the inequality[4]

$$2^{\xi_l} - 1 > \frac{P_L}{e \cdot (\sigma_{l-1}^2 + \sigma_{\mathrm{w}}^2)}, \tag{3.57}$$

which holds for $\xi_l > 0$, to upper-bound $\left(\sigma_l^{(\mathrm{inf})}\right)^2$ in (3.56) as follows:

$$\left(\sigma_l^{(\mathrm{inf})}\right)^2 = \frac{P_L + \sigma_{\mathrm{w}}^2}{2^{\xi_l} - 1} < \left(1 + \frac{\sigma_{\mathrm{w}}^2}{P_L}\right) \cdot e \cdot (\sigma_{l-1}^2 + \sigma_{\mathrm{w}}^2). \tag{3.58}$$

Hence, we conclude that the sequence of quantization noise variances $(\sigma_l)_{l=0}^L$ that is characterized by the following difference equation is achievable almost surely in the limit $n \to \infty$:

$$\sigma_l^2 = \left(1 + \frac{\sigma_{\mathrm{w}}^2}{P_L}\right) \cdot e \cdot (\sigma_{l-1}^2 + \sigma_{\mathrm{w}}^2) \quad \text{with} \quad \sigma_0^2 = 0. \tag{3.59}$$

The solution to this first order difference equation is given by

$$\sigma_l^2 = \sigma_{\mathrm{w}}^2 \cdot \left(1 + \frac{\sigma_{\mathrm{w}}^2}{P_L}\right) \cdot e \cdot \frac{1 - \left(\left(1 + \frac{\sigma_{\mathrm{w}}^2}{P_L}\right) \cdot e\right)^l}{1 - \left(1 + \frac{\sigma_{\mathrm{w}}^2}{P_L}\right) \cdot e}. \tag{3.60}$$

Thus, for sustaining a certain SNR, $\mathrm{snr}_{\mathcal{D}} = P_L/(\sigma_L^2 + \sigma_{\mathrm{w}}^2)$, P_L and L need to be coupled as follows:

$$P_L = \mathrm{snr}_{\mathcal{D}} \cdot \sigma_{\mathrm{w}}^2 \cdot \left(\left(1 + \frac{\sigma_{\mathrm{w}}^2}{P_L}\right) \cdot e \cdot \frac{1 - \left(\left(1 + \frac{\sigma_{\mathrm{w}}^2}{P_L}\right) \cdot e\right)^L}{1 - \left(1 + \frac{\sigma_{\mathrm{w}}^2}{P_L}\right) \cdot e} + 1\right). \tag{3.61}$$

This implicit equation corresponds to an exponential growth of P_L with L, since

$$\lim_{L \to \infty} \frac{P_L}{e^L} = \mathrm{snr}_{\mathcal{D}} \cdot \sigma_{\mathrm{w}}^2 \cdot \frac{e}{e - 1}. \tag{3.62}$$

In order to show that an exponential growth of P_L with L is necessary for sustaining a

[4]This inequality follows, since both sides evaluate to zero for $x \triangleq \frac{P_L}{\sigma_{l-1}^2 + \sigma_{\mathrm{w}}^2} = 0$, and the slope of the left hand side is strictly larger than the slope of the right hand side for all $x > 0$:

$$\frac{\partial}{\partial x}(2^{\xi_l} - 1) = \exp\left(-\frac{(-1 + \sqrt{1 + 4x})^2}{4x}\right) > e^{-1} = \frac{\partial}{\partial x}(e^{-1}x).$$

constant destination SNR, a lower-bound on $\left(\sigma_l^{(\mathrm{inf})}\right)^2$, as defined in (3.56), is developed. There exists for every destination SNR value $\mathrm{snr}_{\mathcal{D}} = P_L/(\sigma_L^2 + \sigma_{\mathrm{w}}^2)$ a $c > 1$, such that for arbitrarily large L

$$2^{\xi_l} - 1 \leq \frac{1}{c} \cdot \rho_l \text{ for all } l \in \{1, \ldots, L\}, \tag{3.63}$$

where $\rho_l \triangleq P_L/(\sigma_{l-1}^2 + \sigma_{\mathrm{w}}^2)$. In order to prove this, c is chosen as[5]

$$c = \min_{l \in \{1,\ldots,L\}} \frac{\rho_l}{2^{\xi_l} - 1} = \frac{\rho_L}{2^{\xi_L} - 1} = \frac{\mathrm{snr}_{\mathcal{D}}}{2^{\xi_L} - 1}. \tag{3.66}$$

This c is larger than one for positive $\mathrm{snr}_{\mathcal{D}}$, since c takes on values on the interval $(1, e)$:

$$\lim_{\mathrm{snr}_{\mathcal{D}} \to 0} c = 1, \tag{3.67}$$

$$\lim_{\mathrm{snr}_{\mathcal{D}} \to \infty} c = e, \tag{3.68}$$

$$\frac{\partial c}{\partial \mathrm{snr}_{\mathcal{D}}} > 0 \text{ for all } \mathrm{snr}_{\mathcal{D}} > 0. \tag{3.69}$$

Hence, we conclude that

$$\left(\sigma_l^{(\mathrm{inf})}\right)^2 = \frac{P_L + \sigma_{\mathrm{w}}^2}{2^{\xi_l} - 1} \geq \frac{P_L}{\frac{1}{c} \cdot \rho_l} = c \cdot (\sigma_{l-1}^2 + \sigma_{\mathrm{w}}^2) \text{ for all } l \in \{1, \ldots, L\}. \tag{3.70}$$

With $\left(\sigma_0^{(\mathrm{inf})}\right)^2 = 0$, this yields the following lower-bound on $\left(\sigma_L^{(\mathrm{inf})}\right)^2$:

$$\left(\sigma_L^{(\mathrm{inf})}\right)^2 > \sigma_{\mathrm{w}}^2 \cdot c \cdot \frac{1 - c^L}{1 - c}, \tag{3.71}$$

[5]Here, $l = L$ is the minimizer, since $\rho_l > \rho_L$ for every $l < L$, and $\rho_l/(2^{\xi_l} - 1)$ is monotonically increasing in ρ_l:

$$\frac{\partial}{\partial \rho_l} \frac{\rho_l}{2^{\xi_l} - 1} = \frac{2e^{\frac{(-1+\sqrt{1+4\rho_l})^2}{4\rho_l}} \left(1 + 4\rho_l - \sqrt{1+4\rho_l}\left(2e^{\frac{(\sqrt{1+4\rho_l}-1)^2}{4\rho_l}} - 1\right)\right)}{\sqrt{1+4\rho_l}\left(1 - 2e^{\frac{(-1+\sqrt{1+4\rho_l})^2}{4\rho_l}} + 2\rho_l + \sqrt{1+4\rho_l}\right)^2} > 0. \tag{3.64}$$

The positiveness of the derivative follows, since $e^x < (1-x)^{-1}$ for $x < 1$, and thus

$$2e^{\frac{(\sqrt{1+4\rho_l}-1)^2}{4\rho_l}} - 1 < \frac{2}{1 - \frac{(\sqrt{1+4\rho_l}-1)^2}{4\rho_l}} - 1 = \sqrt{1+4\rho_l}. \tag{3.65}$$

That is, the power P_L that is required for sustaining a constant SNR, $\mathrm{snr}_{\mathcal{D}}$, is lower-bounded according to

$$P_L = \mathrm{snr}_{\mathcal{D}} \cdot \left(\sigma_L^2 + \sigma_{\mathrm{w}}^2\right) > \mathrm{snr}_{\mathcal{D}} \cdot \sigma_{\mathrm{w}}^2 \cdot \left(c \cdot \frac{1-c^L}{1-c} + 1\right). \tag{3.72}$$

This lower-bound implies an exponential growth of the required P_L with L. □

3.4.2.2. Quantization with Slepian & Wolf Compression

The following theorem characterizes the scaling of the supremum of achievable sum-rates in quantize & forward multihop MIMO multiple access networks that apply Slepian & Wolf compression:

Theorem 4. *Let* $\mathbf{H}_1, \ldots, \mathbf{H}_{L+1} \in \mathbb{C}^{n\times n}$ *be statistically independent random channel matrices, each with i.i.d. elements of zero-mean, unit-variance and bounded fourth moment. Then, the supremum of the set of sum-rates that are achievable by the quantize & forward strategy with additional Slepian & Wolf compression,* R_L^{CF}*, fulfills for all* $L \in \mathbb{N}_0$

$$\lim_{n\to\infty} \frac{1}{n} R_L^{\mathrm{CF}} = \psi\left(\mathrm{snr}_{\mathcal{D}}\right) \triangleq c_L^{\mathrm{CF}} \text{ almost surely,} \tag{3.73}$$

where $\mathrm{snr}_{\mathcal{D}} = P_L/(\sigma_L^2 + \sigma_{\mathrm{w}}^2)$*, and* $\psi(\cdot)$ *is defined in* (3.45)*. Moreover, the per-stage transmit power* P_L *that is required for rendering* $\mathrm{snr}_{\mathcal{D}}$ *constant with respect to* L *grows linearly with* L*.*

That is, the constant of proportionality that asymptotically relates R_L^{CF} and n is constant with respect to L, if the SNR at the destination stage, $\mathrm{snr}_{\mathcal{D}}$, is constant with respect to L.

Proof of Theorem 4. The proof is along the lines of the proof of Theorem 3. In order to prove that the supremum of the set of achievable end-to-end sum-rates scales linearly in $\min\{n_{\mathcal{S}}, n_{\mathcal{R}}, n_{\mathcal{D}}\}$ in the limit $L \to \infty$, it is sufficient to show that there is for every L a P_L, such that the SNR at the destination antennas, given by $\mathrm{snr}_{\mathcal{D}} = P_L/(\sigma_{\mathrm{w}}^2 + \sigma_L^2)$, is kept constant and nonzero. Then, the supremum of the set of achievable sum-rates, as defined in (3.26), fulfills [75, 76]:

$$\lim_{n\to\infty} \frac{1}{n} \cdot R_L^{\mathrm{CF}} = \psi\left(\mathrm{snr}_{\mathcal{D}}\right) \text{ almost surely,} \tag{3.74}$$

independently of L. This limit is strictly positive for positive $\mathrm{snr}_{\mathcal{D}} > 0$.

Again, Lemma 1 implies that for every fixed L and given sequence $(\sigma_l^2)_{l=0}^{L-1}$ there exists almost surely an $n_0(L)$, such that for all $n \geq n_0(L)$ all rates R_l, $l \in \{1, \ldots, L\}$, are achievable simultaneously, if $R_l < \xi_l$ for all l.

Next, the quantization noise variances are evaluated. The following lemma serves as a starting point:

Lemma 3. *Let* y *be a Gaussian random vector with zero-mean and covariance matrix* $\mathbf{K}_{\mathsf{y}} = \sigma_{\mathrm{w}}^2 \mathbf{I}_n + \frac{P}{n}\mathbf{H}\mathbf{H}^H$*, where* $\mathbf{H} \in \mathbb{C}^{n\times n}$ *is distributed according to the assumptions of Theorem 4 and* $P > 0$*. Let* $\hat{\mathsf{y}}$ *be the quantization of* y*, which is obtained as* $\hat{\mathsf{y}} = \mathsf{y} + \mathsf{z}$*, where the quantization noise vector* z *is a Gaussian random vector with zero-mean and covariance matrix* $\mathbf{K}_{\mathsf{z}} = \sigma_{\mathrm{q}}^2 \mathbf{I}_n$*. Let*

$$\zeta = \psi\left(\frac{P}{\sigma_{\mathrm{w}}^2 + \sigma_{\mathrm{q}}^2}\right) + \log\left(1 + \frac{\sigma_{\mathrm{w}}^2}{\sigma_{\mathrm{q}}^2}\right), \tag{3.75}$$

where $\psi(\cdot)$ *is defined in* (3.45)*. Then, there exists for every tuple of rates of compressed quantization codebooks* $(R^{(1)}, R^{(2)}, \ldots, R^{(n)}) = (R, R, \ldots, R)$ *with* $R > \zeta$ *almost surely an* n_0*, such that for all* $n \geq n_0$ *the quantization noise variance* σ_{q}^2 *is achievable in the sense of* (3.31)*.*

This lemma implies that, in the limit of large n, the set of achievable σ_l^2 is fully determined by the constraint in (3.31) that corresponds to the set $\tilde{\mathcal{R}}_l = \mathcal{R}_l$, since [75,76]

$$\lim_{n\to\infty} \frac{1}{n} \log\det\left(\left(1 + \frac{\sigma_{\mathrm{w}}^2}{\sigma_l^2}\right)\mathbf{I}_n + \frac{P_L}{n(\sigma_l^2 + \sigma_{\mathrm{w}}^2)}\mathbf{H}_{l+1}\mathbf{H}_{l+1}^H\right) = \zeta_l \text{ almost surely}, \tag{3.76}$$

where

$$\zeta_l = \psi\left(\frac{P_L}{\sigma_{\mathrm{w}}^2 + \sigma_l^2}\right) + \log\left(1 + \frac{\sigma_{\mathrm{w}}^2}{\sigma_l^2}\right). \tag{3.77}$$

Thus, there is for every fixed L and given sequence $(\tilde{R}_l)_{l=1}^{L}$, where $\tilde{R}_l > \zeta_l$ for all l, almost surely an $n_0(L)$, such that for all $n \geq n_0(L)$ all quantization noise variances σ_l^2, $l \in \{1, \ldots, L\}$, that fulfill (3.77) are achievable simultaneously.

The conclusions of Lemmata 1 & 3 allow to infer that there exists for every fixed L almost surely an $n_0(L)$, such that for all $n \geq n_0(L)$ the sequence of quantization noise variances $(\sigma_l^2)_{l=1}^L$ is achievable according to (3.34), if for all $l \in \{1, \ldots, L\}$

$$\zeta_l < \tilde{R}_l = R_l < \xi_l. \tag{3.78}$$

For a given σ_{l-1}^2 the infimum of the achievable quantization noise variances in stage $\mathcal{R}_l$ is denoted by $\left(\sigma_l^{(\text{inf})}\right)^2$. It is convenient to define

$$\Delta_l^{(\text{inf})} \triangleq \begin{cases} \left(\sigma_l^{(\text{inf})}\right)^2 - \sigma_{l-1}^2, & \text{if } l > 1, \\ \left(\sigma_l^{(\text{inf})}\right)^2, & \text{if } l = 1. \end{cases} \tag{3.79}$$

The quantities $\left(\sigma_l^{(\text{inf})}\right)^2$, or, equivalently $\Delta_l^{(\text{inf})}$, $l \in \{1, \ldots, L\}$, are determined by the equations $\zeta_l = \xi_l$, $l \in \{1, \ldots, L\}$:

$$\psi\left(\frac{P_L}{\sigma_{l-1}^2 + \sigma_\text{w}^2}\right) = \psi\left(\frac{P_L}{\sigma_{l-1}^2 + \Delta_l^{(\text{inf})} + \sigma_\text{w}^2}\right) + \log\left(1 + \frac{\sigma_\text{w}^2}{\sigma_{l-1}^2 + \Delta_l^{(\text{inf})}}\right). \tag{3.80}$$

Theorem 4 is established by showing that there exist constants $c > 0$ and $d > 1$, such that $\sigma_l^2 = (l+1)c$ and $P_L = Ld$ fulfill (3.78) for all $l \in \{1, \ldots, L\}$, where L can grow arbitrarily large. To this end, $\Delta_l^{(\text{inf})}$ is upper-bounded by reducing the transmit power of relay stage $\mathcal{R}_l$ from $P_L = Ld$ to $P_L = ld$. This yields the upper-bound $\Delta_l^{(\text{inf})} < \overline{\Delta}_l^{(\text{inf})}$, since

(i) $\Delta_l^{(\text{inf})}$ can only increase for a fixed σ_{l-1}^2, when the transmit power *both* of $\mathcal{R}_l$ and $\mathcal{R}_{l+1}$ are reduced from Ld to $(l+1)d$ in a first step[6].

[6]Proof: The derivative $\partial\Delta_l^{(\text{inf})}/\partial P_L$ is obtained according to the implicit function theorem:

$$\begin{aligned}\frac{\partial\Delta_l^{(\text{inf})}}{\partial P_L} = \Big[&-2\left(4P_L + \sigma_\text{w}^2 + \Delta_l^{(\text{inf})} + \sigma_\text{w}^2\beta - \Delta_l^{(\text{inf})}\beta - \sigma_{l-1}^2\beta + \sigma_{l-1}^2\right) \\ &\times\left(\sigma_{l-1}^4(\beta-\alpha) + \sigma_\text{w}^2(\sigma_\text{w}^2 + \Delta_l^{(\text{inf})})(\beta-\alpha) + \sigma_{l-1}^2(4P_L + 2\sigma_\text{w}^2 + \Delta_l^{(\text{inf})})(\beta-\alpha)\right. \\ &\left.\left.+P_L\left(4\sigma_\text{w}^2(\beta-\alpha) + 4\Delta_l^{(\text{inf})}\beta\right)\right)\right] \Big/ \\ &\Big[P_L(\sigma_{l-1}^2 + \sigma_\text{w}^2)\alpha(1+\alpha)(\sigma_{l-1}^2 + \Delta_l^{(\text{inf})}) \\ &\times(\sigma_{l-1}^2 + \sigma_\text{w}^2 + \Delta_l^{(\text{inf})})(\sigma_{l-1}^2 + 4P_L + \sigma_\text{w}^2 + \Delta_l^{(\text{inf})})(1+\beta)^2/(\log e)^2\Big],\end{aligned}$$

where $\alpha = \sqrt{1 + \frac{4P_L}{\sigma_\text{w}^2 + \sigma_{l-1}^2}}$ and $\beta = \sqrt{1\frac{4P_L}{\sigma_\text{w}^2 + \sigma_{l-1}^2 + \Delta_l^{(\text{inf})}}}$. This derivative is nonpositive for positive P_L, $\Delta_l^{(\text{inf})}$, σ_w^2, and σ_{l-1}^2. This is seen as follows. The denominator is obviously positive. The numerator

(ii) the upper-bound on $\Delta_l^{(\mathrm{inf})}$ from (i) can again only increase for a fixed a transmit power $(l+1)d$ of stage $\mathcal{R}_{l+1}$ and fixed σ_{l-1}^2, when the transmit power of $\mathcal{R}_l$ is reduced from $(l+1)d$ to ld in a second step.

Eqs. (3.80) are rewritten in a first step as

$$\frac{\log e}{\frac{4ld}{lc+\overline{\Delta}_l^{(\mathrm{inf})}+\sigma_\mathrm{w}^2}}\left(\sqrt{\frac{4ld}{lc+\overline{\Delta}_l^{(\mathrm{inf})}+\sigma_\mathrm{w}^2}+1}-1\right)^2+\phi(\overline{\Delta}_l^{(\mathrm{inf})})=0,\quad l\in\{1,\dots,L\},\tag{3.81}$$

where

$$\phi(\overline{\Delta}_l^{(\mathrm{inf})})=-2\log\left(1+\frac{ld}{lc+\overline{\Delta}_l^{(\mathrm{inf})}+\sigma_\mathrm{w}^2}-\frac{1}{4}\left(\sqrt{\frac{4ld}{lc+\overline{\Delta}_l^{(\mathrm{inf})}+\sigma_\mathrm{w}^2}+1}-1\right)^2\right)$$
$$-\log\left(1+\frac{\sigma_\mathrm{w}^2}{lc+\overline{\Delta}_l^{(\mathrm{inf})}}\right)+\xi_l\Bigg|_{P_L=ld,\sigma_{l-1}^2=lc},\tag{3.82}$$

and in a second step as

$$\overline{\Delta}_l^{(\mathrm{inf})}=-\frac{ld\log e}{\phi(\overline{\Delta}_l^{(\mathrm{inf})})}-\sigma_\mathrm{w}^2-cl-2ld-\frac{ld\phi(\overline{\Delta}_l^{(\mathrm{inf})})}{\log e},\quad l\in\{1,\dots,L\}.\tag{3.83}$$

We take limits on both sides (Bernoulli l'Hospital) and obtain:

$$\lim_{l\to\infty}\overline{\Delta}_l^{(\mathrm{inf})}=\frac{2d\left(c+4d+c\sqrt{1+\frac{4d}{c}}\right)\lim_{l\to\infty}\overline{\Delta}_l^{(\mathrm{inf})}-\sigma_\mathrm{w}^2(c+4d)\left(c+2d+c\sqrt{1+\frac{4d}{c}}\right)}{2cd\sqrt{1+\frac{4d}{c}}}.\tag{3.84}$$

This equation is solved for $\lim_{l\to\infty}\overline{\Delta}_l^{(\mathrm{inf})}$ as follows:

$$\lim_{l\to\infty}\overline{\Delta}_l^{(\mathrm{inf})}=\sigma_\mathrm{w}^2\cdot\left(1+\frac{1+\sqrt{1+\frac{4d}{c}}}{2\frac{d}{c}}\right).\tag{3.85}$$

has zeros only at $\Delta_l^{(\mathrm{inf})}=0$, $\Delta_l^{(\mathrm{inf})}=-\sigma_\mathrm{w}^2-\sigma_{l-1}^2-4P_L$ and $\Delta_l^{(\mathrm{inf})}=-\sigma_{l-1}^2-\frac{\sigma_\mathrm{w}^2P_L}{P_L+\sigma_\mathrm{w}^2}$, which implies that P_L, $\Delta_l^{(\mathrm{inf})}$, σ_w^2, and σ_{l-1}^2 cannot be positive simultaneously at a zero, and, therefore, the numerator has the same sign for all positive tuples $(P_L,\Delta_l^{(\mathrm{inf})},\sigma_\mathrm{w}^2,\sigma_{l-1}^2)$. It thus remains to show that the numerator is negative for an arbitrary choice of the positive tuple $(P_L,\Delta_l^{(\mathrm{inf})},\sigma_\mathrm{w}^2,\sigma_{l-1}^2)$, e.g., for $P_L=6$, $\sigma_\mathrm{w}^2=1$, $\sigma_{l-1}^2=2$, $\Delta_l^{(\mathrm{inf})}=5$, it evaluates to -5760.

Thus, there is for every $\varepsilon > 0$ an l_0, such that for all $l \geq l_0$

$$\left| \overline{\Delta}_l^{(\mathrm{inf})} - \sigma_{\mathrm{w}}^2 \cdot \left(1 + \frac{1 + \sqrt{1 + \frac{4d}{c}}}{2\frac{d}{c}} \right) \right| < \varepsilon. \tag{3.86}$$

In order to ensure that

$$\begin{aligned} \sigma_l^2 - \sigma_{l_0-1}^2 &= (l - l_0 + 1)c & (3.87)\\ &\geq (l - l_0 + 1) \max_{l \in \{l_0, \ldots, l\}} \left\{ \overline{\Delta}_l^{(\mathrm{inf})} \right\} & (3.88)\\ &\geq \sum_{k=l_0}^{l} \overline{\Delta}_k^{(\mathrm{inf})} > \sum_{k=l_0}^{l} \Delta_k^{(\mathrm{inf})}, & (3.89) \end{aligned}$$

for all $l \in \{l_0, \ldots, L\}$, we fix

$$c = \sigma_{\mathrm{w}}^2 \cdot \left(1 + \frac{1 + \sqrt{1 + \frac{4d}{c}}}{2\frac{d}{c}} \right) + \varepsilon, \tag{3.90}$$

which for small ε leads to the coupling ($d > 1$ by assumption)

$$c \approx \frac{\sigma_{\mathrm{w}}^2 \cdot \sqrt{d}}{\sqrt{d} - \sigma_{\mathrm{w}}}. \tag{3.91}$$

Thus, we have shown that

$$\Delta_l^{(\mathrm{inf})} < \sigma_{\mathrm{w}}^2 \cdot \left(1 + \frac{1 + \sqrt{1 + \frac{4d}{c}}}{2\frac{d}{c}} \right) + \varepsilon \text{ for all } l > l_0. \tag{3.92}$$

It remains to show that there exists an L_0, such that for all $L > L_0$

$$\Delta_l^{(\mathrm{inf})} < \sigma_{\mathrm{w}}^2 \cdot \left(1 + \frac{1 + \sqrt{1 + \frac{4d}{c}}}{2\frac{d}{c}} \right) + \varepsilon \text{ for all } l \leq l_0. \tag{3.93}$$

To this end, we substitute $P_L = dL$ and $\sigma_{l-1}^2 = cl$ in (3.81) and take L to infinity. This results for $1 \leq l \leq l_0$ into the first order difference equation (with $\sigma_0^2 = 0$)

$$\lim_{L \to \infty} \left(\sigma_l^{(\mathrm{inf})} \right)^2 = \sigma_{l-1}^2 + \sigma_{\mathrm{w}}^2, \tag{3.94}$$

or, equivalently,

$$\lim_{L\to\infty} \Delta_l^{(\mathrm{inf})} = \sigma_{\mathrm{w}}^2. \tag{3.95}$$

Thus, we conclude that there is for every $\varepsilon > 0$ an L_0, such that for all $L \geq L_0$ there exists almost surely an $n_0(L)$, such that for all $n \geq n_0(L)$

$$\sigma_L^2 = L \cdot c = L \cdot \left(\sigma_{\mathrm{w}}^2 \cdot \left(1 + \frac{1 + \sqrt{1 + \frac{4d}{c}}}{2^{\frac{d}{c}}} \right) + \varepsilon \right) > \sum_{l=1}^{L} \Delta_l^{(\mathrm{inf})} \tag{3.96}$$

is achievable. In order to sustain a certain value of $\mathrm{snr}_{\mathcal{D}}$ in the large n limit, it is thus sufficient that

$$P_L = \mathrm{snr}_{\mathcal{D}} \cdot (\sigma_L^2 + \sigma_{\mathrm{w}}^2) = \mathrm{snr}_{\mathcal{D}} \cdot \sigma_{\mathrm{w}}^2 \cdot \left(1 + L \cdot \left(1 + \frac{1 + \sqrt{1 + \frac{4d}{c}}}{2^{\frac{d}{c}}} + \varepsilon \right) \right). \tag{3.97}$$

Since $d/c \to \mathrm{snr}_{\mathcal{D}}$ as $L \to \infty$, this leads to the following asymptotically linear coupling between P_L and L:

$$\lim_{L\to\infty} \frac{P_L}{L} = \sigma_{\mathrm{w}}^2 \cdot \left((1 + \varepsilon) \cdot \mathrm{snr}_{\mathcal{D}} + \frac{1}{2} + \frac{\sqrt{1 + 4\mathrm{snr}_{\mathcal{D}}}}{2} \right). \tag{3.98}$$

For the converse, i.e., the proof that any P_L that scales slower than linearly with L cannot sustain a constant SNR at the destination stage, (3.80) serves as the starting point, again:

$$\psi\left(\frac{P_L}{\sigma_{l-1}^2 + \sigma_{\mathrm{w}}^2} \right) = \psi\left(\frac{P_L}{\sigma_{l-1}^2 + \Delta_l^{(\mathrm{inf})} + \sigma_{\mathrm{w}}^2} \right) + \log\left(1 + \frac{\sigma_{\mathrm{w}}^2}{\sigma_{l-1}^2 + \Delta_l^{(\mathrm{inf})}} \right). \tag{3.99}$$

The inequality[7]

$$\log\left(1 + \frac{P_L}{\sigma_{l-1}^2 + \sigma_{\mathrm{w}}^2} \right) - \log\left(1 + \frac{P_L}{\sigma_{l-1}^2 + \Delta_l^{(\mathrm{inf})} + \sigma_{\mathrm{w}}^2} \right)$$

[7]Proof: Rewrite the inequality as

$$\begin{aligned} \log\left(1 + \frac{P_L}{\sigma_{l-1}^2 + \sigma_{\mathrm{w}}^2} \right) - \psi\left(\frac{P_L}{\sigma_{l-1}^2 + \sigma_{\mathrm{w}}^2} \right) \\ \geq \log\left(1 + \frac{P_L}{\sigma_{l-1}^2 + \Delta_l^{(\mathrm{inf})} + \sigma_{\mathrm{w}}^2} \right) - \psi\left(\frac{P_L}{\sigma_{l-1}^2 + \Delta_l^{(\mathrm{inf})} + \sigma_{\mathrm{w}}^2} \right). \end{aligned} \tag{3.100}$$

$$\geq \psi\left(\frac{P_L}{\sigma_{l-1}^2+\sigma_{\mathrm{w}}^2}\right)-\psi\left(\frac{P_L}{\sigma_{l-1}^2+\Delta_l^{(\mathrm{inf})}+\sigma_{\mathrm{w}}^2}\right). \tag{3.102}$$

implies the inequality

$$\log\left(1+\frac{P_L}{\sigma_{l-1}^2+\sigma_{\mathrm{w}}^2}\right)\geq \log\left(1+\frac{P_L}{\sigma_{l-1}^2+\Delta_l^{(\mathrm{inf})}+\sigma_{\mathrm{w}}^2}\right)+\log\left(1+\frac{\sigma_{\mathrm{w}}^2}{\sigma_{l-1}^2+\Delta_l^{(\mathrm{inf})}}\right), \tag{3.103}$$

which leads to the following lower-bounds on $\Delta_l^{(\mathrm{inf})}$ and $\left(\sigma_L^{(\mathrm{inf})}\right)^2$:

$$\Delta_l^{(\mathrm{inf})} \geq \sigma_{\mathrm{w}}^2, \tag{3.104}$$

$$\left(\sigma_L^{(\mathrm{inf})}\right)^2 \geq L\cdot\sigma_{\mathrm{w}}^2. \tag{3.105}$$

Thus, the power required to sustain a constant SNR, $\mathrm{snr}_{\mathcal{D}}$, must fulfill

$$P_L=\mathrm{snr}_{\mathcal{D}}\cdot(\sigma_L^2+\sigma_{\mathrm{w}}^2)\geq \mathrm{snr}_{\mathcal{D}}\cdot(L+1)\cdot\sigma_{\mathrm{w}}^2. \tag{3.106}$$

3.4.3. Amplify & Forward Networks

The following theorem characterizes the scaling of the supremum of achievable sum-rates in amplify & forward multihop MIMO multiple access networks.

Theorem 5. *Let* $\mathbf{H}_1\in\mathbb{C}^{n_{\mathcal{D}}\times n_{\mathcal{R}}}$, $\mathbf{H}_2,\ldots,\mathbf{H}_L\in\mathbb{C}^{n_{\mathcal{R}}\times n_{\mathcal{R}}}$ *and* $\mathbf{H}_{L+1}\in\mathbb{C}^{n_{\mathcal{R}}\times n_{\mathcal{S}}}$ *be statistically independent random matrices, each with i.i.d. elements of zero-mean, unit-variance and bounded fourth moment, and fix* $\frac{n_{\mathcal{S}}}{n_{\mathcal{D}}}\triangleq\beta_{\mathcal{S}}$ *and* $\frac{n_{\mathcal{R}}}{n_{\mathcal{D}}}\triangleq\beta_{\mathcal{R}}$. *Let* $\mathrm{snr}_{\mathcal{D}}$ *(signal-to-noise ratio at destination stage) be a positive constant. Fix*

$$P_L=\sigma_{\mathrm{w}}^2\cdot\left(\sqrt[L+1]{\frac{\mathrm{snr}_{\mathcal{D}}+1}{\mathrm{snr}_{\mathcal{D}}}}-1\right)^{-1} \quad\text{and}\quad \alpha=\frac{P_L}{P_L+\sigma_{\mathrm{w}}^2}. \tag{3.107}$$

This inequality holds, since $\Delta_l^{(\mathrm{inf})}\geq 0$ and $\log(1+x)+\psi(x)$ is monotonously increasing in x:

$$\frac{\partial\left(\log(1+x)-\psi(x)\right)}{\partial x}=\frac{\log e\sqrt{1+4x}\left((1+2x)-\sqrt{1+4x}\right)}{x(1+x)\sqrt{1+4x}\left(1+\sqrt{1+4x}\right)}>0 \text{ for all } x>0. \tag{3.101}$$

Let $(c_L^{\mathrm{AF}})_{L=0}^{\infty}$ *be the sequence, such that* $c_L^{\mathrm{AF}} = \lim_{n_{\mathcal{D}}\to\infty} \frac{R_L^{\mathrm{AF}}}{n_{\mathcal{D}}}$ *almost surely. Then,*

$$\lim_{L\to\infty} c_L^{\mathrm{AF}} = \begin{cases} \beta_{\mathcal{S}} \log\left(1 + \mathrm{snr}_{\mathcal{D}} - \frac{1}{4}\chi(\mathrm{snr}_{\mathcal{D}}, \beta_{\mathcal{S}})\right) & \\ + \log\left(1 + \mathrm{snr}_{\mathcal{D}}\beta_{\mathcal{S}} - \frac{1}{4}\chi(\mathrm{snr}_{\mathcal{D}}, \beta_{\mathcal{S}})\right) & \\ - \frac{\log e}{4\mathrm{snr}_{\mathcal{D}}}\chi(\mathrm{snr}_{\mathcal{D}}, \beta_{\mathcal{S}}) \triangleq c_0, & \textit{if } \beta_{\mathcal{R}} \in \Omega\left(L^{1+\varepsilon}\right), \\ 0, & \textit{if } \beta_{\mathcal{R}} \in \mathcal{O}\left(L^{1-\varepsilon}\right), \end{cases} \tag{3.108}$$

where $\chi(x, z) = \left(\sqrt{x(1+\sqrt{z})^2+1} - \sqrt{x(1-\sqrt{z})^2+1}\right)^2$. *Moreover, if* $\beta_{\mathcal{R}} \in \Theta(L)$,

$$\limsup_{L\to\infty} c_L^{\mathrm{AF}} \leq c_0 \qquad \textit{and} \qquad \liminf_{L\to\infty} c_L^{\mathrm{AF}} > 0. \tag{3.109}$$

Remark 1: The per-stage transmit power P_L is chosen such that the destination SNR, $\mathrm{snr}_{\mathcal{D}}$, is constant with respect to L. It scales linearly with L:

$$\lim_{L\to\infty} \frac{P_L}{L} = \frac{\sigma_{\mathrm{w}}^2 \log e}{\log\left(1 + \frac{1}{\mathrm{snr}_{\mathcal{D}}}\right)}. \tag{3.110}$$

Remark 2: Theorem 5 allows to conclude the following:

- R_L^{AF} scales linearly in $\min\{n_{\mathcal{S}}, n_{\mathcal{D}}\}$, if and only if the ratio $\beta_{\mathcal{R}}$ scales at least linearly with L.
- The asymptotic sum-capacity of a single-hop multiple access channel with $n_{\mathcal{S}}$ single-antenna sources and an $n_{\mathcal{D}}$-antenna destination (cf. [6, 76]) is approached for faster than linear growth of $\beta_{\mathcal{R}}$ with L.

For the proof of the theorem the following notation is introduced. The empirical eigenvalue distribution (EED) of some Hermitian matrix $\mathbf{A} \in \mathbb{C}^{n\times n}$ is defined in terms of the indicator function $1\{\cdot\}$ as $F_{\mathbf{A}}^{(\gamma_1,\ldots,\gamma_K)}(x) \triangleq \frac{1}{n}\sum_{i=1}^{n} 1\{\lambda_i\{\mathbf{A}\} < x\}$. The superscripts $\gamma_1, \ldots, \gamma_K$ indicate parameters the EED depends on. Whenever one of these parameters is taken to infinity, the respective superscript is dropped. The expression for R_L^{AF} in (3.38) can be expressed in terms of the EED of $\mathbf{R}_s\mathbf{R}_n^{-1}$ as follows:

$$R_L^{\mathrm{AF}} = n_{\mathcal{D}} \cdot \int_0^{\infty} \log(1+x) \cdot dF_{\mathbf{R}_s\mathbf{R}_n^{-1}}^{(n_{\mathcal{D}},\beta_{\mathcal{R}},L,\beta_{\mathcal{S}})}(x). \tag{3.111}$$

Moreover, the Stieltjes transform of some EED $F(\cdot)$ is defined as

$$G(s) \triangleq \int_{-\infty}^{\infty} \frac{1}{s+x} \cdot dF(x). \tag{3.112}$$

This definition is adopted from [25] here, while it is generally more common to define the Stieltjes transform with a minus sign in the denominator. The EED is uniquely determined by its Stieltjes transform. Moreover, the proof of Theorem 5 relies on the property that $F_{\mathbf{A}}^{(\gamma)}(x)$ converges to $F_{\mathbf{A}}(x)$ pointwise with respect to γ, if and only if $G_{\mathbf{A}}^{(\gamma)}(s)$ converges to $G_{\mathbf{A}}(s)$ pointwise [77, Corollary 1].

The Marčenko-Pastur law [75,78] is a fundamental result in large random matrix theory. Let $\mathbf{X} \in \mathbb{C}^{k_0 \times k_1}$ be a random matrix whose entries are i.i.d. zero-mean distributed and of unit-variance. If both $k_0 \to \infty$ and $k_1 \to \infty$, but $\beta = k_1/k_0$ is kept finite, then the EED $F_{\frac{1}{k_1}\mathbf{X}\mathbf{X}^H}^{(k_1\beta)}(x)$ converges uniformly and almost surely to an asymptotic EED $F_{\frac{1}{k_1}\mathbf{X}\mathbf{X}^H}^{(\beta)}(x)$ whose Stieltjes transform is given by

$$G_{\mathrm{MP}}^{(\beta)}(s) = \frac{\beta^{-1} - 1 - s \pm \sqrt{s^2 + 2\left(\beta^{-1}+1\right)s + \left(\beta^{-1}-1\right)^2}}{2s\beta^{-1}}. \tag{3.113}$$

Moreover, if, in addition, the elements of $\mathbf{X}$ have finite fourth moments, the maximum eigenvalue of $\frac{1}{k_1}\mathbf{X}\mathbf{X}^H$ converges almost surely to [79]

$$\lim_{k_1 \to \infty} \lambda_{\max}\left\{\frac{1}{k_1}\mathbf{X}\mathbf{X}^H\right\} = \left(1 + \frac{1}{\sqrt{\beta}}\right)^2. \tag{3.114}$$

Two transform pairs that appear frequently in the course of this section are the following:

$$F_{\mathrm{MP}}^{(\beta)}(x) \circ\!\!-\!\!\bullet\ G_{\mathrm{MP}}^{(\beta)}(s), \tag{3.115}$$

$$\sigma(x - x_0) \circ\!\!-\!\!\bullet\ \frac{1}{s + x_0}. \tag{3.116}$$

The following Propositions 6 & 7

(i) state that the EEDs of $\mathbf{R}_s$ and $\mathbf{R}_n^{-1}$, $F_{\mathbf{R}_s}^{(n_\mathcal{D},\beta_\mathcal{R},L,\beta_\mathcal{S})}(x)$ and $F_{\mathbf{R}_n^{-1}}^{(n_\mathcal{D},\beta_\mathcal{R},L)}(x)$, convergence in the limit $n_\mathcal{D} \to \infty$ uniformly and almost surely to asymptotic EEDs $F_{\mathbf{R}_s}^{(\beta_\mathcal{R},L,\beta_\mathcal{S})}(x)$ and $F_{\mathbf{R}_n^{-1}}^{(\beta_\mathcal{R},L)}(x)$.

(ii) characterize the asymptotic EEDs $F_{\mathbf{R}_s}^{(\beta_\mathcal{R},L,\beta_\mathcal{S})}(x)$ and $F_{\mathbf{R}_n^{-1}}^{(\beta_\mathcal{R},L)}(x)$ in the limit

$L \to \infty$ for the different couplings between $\beta_{\mathcal{R}}$ and L.

Rather than $\mathbf{R}_s$ and $\mathbf{R}_n$, the following normalized matrices are considered:

$$\tilde{\mathbf{R}}_s = P_L^{-1}\alpha^{-L} \cdot \mathbf{R}_s \quad \text{and} \quad \tilde{\mathbf{R}}_n = \sigma_{\mathrm{w}}^{-2} \cdot \frac{1-\alpha}{1-\alpha^{L+1}} \cdot \mathbf{R}_n. \tag{3.117}$$

Note that due to the particular choice of α and P_L in the theorem, the relation $\mathbf{R}_s\mathbf{R}_n^{-1} = \mathrm{snr}_{\mathcal{D}} \cdot \tilde{\mathbf{R}}_s\tilde{\mathbf{R}}_n^{-1}$ holds.

Proposition 6. *Given the assumptions of Theorem 5, the EED of $\tilde{\mathbf{R}}_s$ converges uniformly as $n_{\mathcal{D}} \to \infty$. That is, there is an asymptotic EED $F_{\tilde{\mathbf{R}}_s}^{(\beta_{\mathcal{R}},L,\beta_{\mathcal{S}})}(x)$, such that*

$$\lim_{n_{\mathcal{D}}\to\infty} \sup_x \left| F_{\tilde{\mathbf{R}}_s}^{(n_{\mathcal{D}},\beta_{\mathcal{R}},L,\beta_{\mathcal{S}})}(x) - F_{\tilde{\mathbf{R}}_s}^{(\beta_{\mathcal{R}},L,\beta_{\mathcal{S}})}(x) \right| = 0 \text{ almost surely.} \tag{3.118}$$

Moreover, the following statements hold:

- *If $\beta_{\mathcal{R}} \in \Omega(^{1+\varepsilon})$, the asymptotic EED of $\tilde{\mathbf{R}}_s$ converges in the limit $L \to \infty$ pointwise to*

$$F_{\tilde{\mathbf{R}}_s}^{(\beta_{\mathcal{S}})}(x) = F_{\mathrm{MP}}^{(\beta_{\mathcal{S}})}(x) \tag{3.119}$$

- *If $\beta_{\mathcal{R}} \in \Theta(L)$, there exists an $a > 0$, such that the asymptotic EED of $\tilde{\mathbf{R}}_s$ fulfills*

$$\liminf_{L\to\infty} 1 - F_{\tilde{\mathbf{R}}_s}^{(\beta_{\mathcal{S}},\beta_{\mathcal{R}},L)}(a) > 0. \tag{3.120}$$

- *If $\beta_{\mathcal{R}} \in \mathcal{O}(L^{1-\varepsilon})$, the asymptotic EED of $\tilde{\mathbf{R}}_s$ converges in the limit $L \to \infty$ pointwise to*

$$F_{\tilde{\mathbf{R}}_s}^{(\beta_{\mathcal{S}})}(x) = \sigma(x). \tag{3.121}$$

Proof of Proposition 6. The uniform and almost sure convergence of $F_{\tilde{\mathbf{R}}_s}^{(n_{\mathcal{D}},\beta_{\mathcal{R}},L,\beta_{\mathcal{S}})}(x)$ to $F_{\tilde{\mathbf{R}}_s}^{(\beta_{\mathcal{R}},L,\beta_{\mathcal{S}})}(x)$ as $n_{\mathcal{D}} \to \infty$ and the respective implicit equation for the Stieltjes transform $G_{\tilde{\mathbf{R}}_s}^{(\beta_{\mathcal{R}},L,\beta_{\mathcal{S}})}(s)$ follows from a result in [25]. It is obtained by means of the S-transform [80, 81], and repeated in the following lemma:

Lemma 4. *Let $\mathbf{M}_1 \in \mathbb{C}^{k_0 \times k_1}, \ldots, \mathbf{M}_N \in \mathbb{C}^{k_{N-1} \times k_N}$ be statistically independent random matrices that fulfill the conditions for the Marčenko-Pastur law and have elements of* *unit-variance. Define* $\beta_n = \frac{k_n}{k_0}$*. Then, the EED of*

$$\mathbf{A} \triangleq 1/\left(k_1 \cdots k_N\right) \mathbf{M}_1 \cdots \mathbf{M}_N \mathbf{M}_N^H \cdots \mathbf{M}_1^H \tag{3.122}$$

converges uniformly and almost surely as $k_0 \to \infty$*: there exists an asymptotic* $F_{\mathbf{A}}^{(\beta_1,\ldots,\beta_N)}(x)$*, such that*

$$\lim_{k_0\to\infty} \sup_x \left| F_{\mathbf{A}}^{(k_0,\beta_1,\ldots,\beta_N)}(x) - F_{\mathbf{A}}^{(\beta_1,\ldots,\beta_N)}(x) \right| = 0 \text{ almost surely.} \tag{3.123}$$

Moreover, the Stieltjes transform of this asymptotic EED fulfills the implicit equation

$$\frac{G_{\mathbf{A}}^{(\beta_1,\ldots,\beta_N)}(s)}{\beta_N} \prod_{n=0}^{N-1} \frac{sG_{\mathbf{A}}^{(\beta_1,\ldots,\beta_N)}(s) - 1 + \beta_{n+1}}{\beta_n} + sG_{\mathbf{A}}^{(\beta_1,\ldots,\beta_N)}(s) = 1. \tag{3.124}$$

Note that $\mathbf{A}$ is normalized with respect to k_N here, while it is normalized with respect to k_0 in [25]. The Stieltjes transform $\tilde{G}(s)$ therein relates to $G(s)$ as $G(s) = \beta_N \tilde{G}(\beta_N s)$.

For the setting considered in this work, N is identified with $L+1$, $\mathbf{M}_n$ with $\mathbf{H}_l$, k_N with $n_{\mathcal{S}}$, and k_n, with $n_{\mathcal{R}}$ for all $n < N$. This yields for our setting the implicit equation

$$\Psi_{\tilde{\mathbf{R}}_s} \cdot \frac{G_{\tilde{\mathbf{R}}_s}^{(\beta_{\mathcal{R}},L,\beta_{\mathcal{S}})}(s)}{\beta_{\mathcal{S}}} \cdot \frac{sG_{\tilde{\mathbf{R}}_s}^{(\beta_{\mathcal{R}},L,\beta_{\mathcal{S}})}(s) - 1 + \beta_{\mathcal{S}}}{\beta_{\mathcal{R}}} \cdot \left(sG_{\tilde{\mathbf{R}}_s}^{(\beta_{\mathcal{R}},L,\beta_{\mathcal{S}})}(s) - 1 + \beta_{\mathcal{R}} \right) + sG_{\tilde{\mathbf{R}}_s}^{(\beta_{\mathcal{R}},L,\beta_{\mathcal{S}})}(s) = 1. \tag{3.125}$$

where

$$\Psi_{\tilde{\mathbf{R}}_s} = \left(\frac{sG_{\tilde{\mathbf{R}}_s}^{(\beta_{\mathcal{R}},L,\beta_{\mathcal{S}})}(s) - 1}{\beta_{\mathcal{R}}} + 1 \right)^{L-1}. \tag{3.126}$$

For the asymptotic analysis $L \to \infty$, the following lemma, whose proof is provided in Appendix B.2, is used:

Lemma 5. *Let* $\varepsilon > 0$ *and* g *be some function* $g : \mathbb{N} \longrightarrow \mathbb{Q}^+ : \kappa \longrightarrow g(\kappa)$*.*

- *Then, for all* $c \in \mathbb{C}$ *and* $g(\kappa) \in \Omega(\kappa^{1+\varepsilon})$

$$\lim_{\kappa\to\infty} \left(\frac{c}{g(\kappa)} + 1 \right)^{\kappa} = 1. \tag{3.127}$$

- *Then, for all negative c and $g(\kappa) \in \Theta(\kappa)$ there exist finite constants M_1 and M_2, $M_2 \geq M_1 > 0$, such that*

$$\liminf_{\kappa\to\infty} \left(\frac{c}{g(\kappa)}+1\right)^{\kappa} \geq e^{cM_2}, \tag{3.128}$$

$$\limsup_{\kappa\to\infty} \left(\frac{c}{g(\kappa)}+1\right)^{\kappa} \leq e^{cM_1}. \tag{3.129}$$

- *Then, for all negative c and $g(\kappa) \in \mathcal{O}(\kappa^{1-\varepsilon})$,*

$$\lim_{\kappa\to\infty} g(\kappa) \cdot \left(\frac{c}{g(\kappa)}+1\right)^{\kappa} = 0. \tag{3.130}$$

Lemma 5 is applied to $\Psi_{\tilde{\mathbf{R}}_s}$, where L is identified with κ and $\beta_{\mathcal{R}}$ with $g(\kappa)$. In the limit $L \to \infty$, the implicit equation (3.125) simplifies as follows:

1. If $\beta_{\mathcal{R}} \in \Omega(L^{1+\varepsilon})$, then $\Psi_{\tilde{\mathbf{R}}_s} \to 1$, and thus

$$\beta_{\mathcal{S}}^{-1} s {G_{\tilde{\mathbf{R}}_s}^{(\beta_{\mathcal{S}})}}^2(s) + \left(s+1-\beta_{\mathcal{S}}^{-1}\right) G_{\tilde{\mathbf{R}}_s}^{(\beta_{\mathcal{S}})}(s) = 1. \tag{3.131}$$

   The solution to this equation is the Stieltjes transform of the Marčenko-Pastur law with parameter $\beta_{\mathcal{S}}$.

2. If $\beta_{\mathcal{R}} \in \mathcal{O}(\beta_{\mathcal{R}}^{1-\varepsilon})$, then $\Psi_{\tilde{\mathbf{R}}_s} \to 0$ for $s > 0$. Lemma 5 applies, since $0 \leq sG_{\tilde{\mathbf{R}}_s}^{(\beta_{\mathcal{R}},L,\beta_{\mathcal{S}})}(s) < 1$ for $s > 0$. Note that the Stieltjes transform is positive for positive s, and thus the left hand side of (3.125) would be larger than one, if $sG_{\tilde{\mathbf{R}}_s}^{(\beta_{\mathcal{R}},L,\beta_{\mathcal{S}})}(s) > 1$, cf. proof of [25, Theorem 4]). Thus,

$$G_{\tilde{\mathbf{R}}_s}^{(\beta_{\mathcal{S}})}(s) = \frac{1}{s}. \tag{3.132}$$

   Since the Stieltjes transform is an analytic function on its full domain, (3.132) holds for all s. The corresponding asymptotic EED is $F_{\tilde{\mathbf{R}}_s}^{(\beta_{\mathcal{S}})}(x) = \sigma(x)$.

3. If $\beta_{\mathcal{R}} \in \Theta(L)$, then there exist according to Lemma 5 M_1 and M_2, $0 < M_1 \leq M_2$, such that for all $s > 0$ (again $0 \leq sG_{\tilde{\mathbf{R}}_s}^{(\beta_{\mathcal{S}},\beta_{\mathcal{R}},L)}(s) < 1$):

$$\liminf_{L\to\infty} \Psi_{\tilde{\mathbf{R}}_s}(\beta_{\mathcal{R}}, L) \geq e^{M_2\left(sG_{\tilde{\mathbf{R}}_s}^{(\beta_{\mathcal{S}},\beta_{\mathcal{R}},L)}(s)-1\right)} > 0, \tag{3.133}$$

$$\limsup_{L\to\infty} \Psi_{\tilde{\mathbf{R}}_s}(\beta_{\mathcal{R}}, L) \leq e^{M_1\left(sG_{\tilde{\mathbf{R}}_s}^{(\beta_{\mathcal{S}},\beta_{\mathcal{R}},L)}(s)-1\right)} < 1. \tag{3.134}$$

Moreover,

$$\lim_{L\to\infty}\left|G^{(\beta_{\mathcal{S}},\beta_{\mathcal{R}},L)}_{\tilde{\mathbf{R}}_s}(s)-\frac{1}{s+\Psi_{\tilde{\mathbf{R}}_s}(\beta_{\mathcal{R}},L)\cdot\left(\beta_{\mathcal{S}}^{-1}\left(sG^{(\beta_{\mathcal{S}},\beta_{\mathcal{R}},L)}_{\tilde{\mathbf{R}}_s}(s)-1\right)+1\right)}\right|=0. \tag{3.135}$$

Since $\Psi_{\tilde{\mathbf{R}}_s}(\beta_{\mathcal{R}},L)>0$, the denominator is equal to s, if and only if $sG^{(\beta_{\mathcal{S}},\beta_{\mathcal{R}},L)}_{\tilde{\mathbf{R}}_s}(s)-1=-\beta_{\mathcal{S}}$. This choice, however, leads to the contradiction

$$G^{(\beta_{\mathcal{S}},\beta_{\mathcal{R}},L)}_{\tilde{\mathbf{R}}_s}(s)=(-\beta_{\mathcal{S}}+1)s^{-1}\neq s^{-1}. \tag{3.136}$$

This proves that for every positive s

$$\liminf_{L\to\infty}\left|G^{(\beta_{\mathcal{S}},\beta_{\mathcal{R}},L)}_{\tilde{\mathbf{R}}_s}(s)-\frac{1}{s}\right|>0. \tag{3.137}$$

Since the Stieltjes transform is an analytic function on its full domain, (3.137) holds for all s. For the asymptotic EED, this implies that there is a positive x for which

$$\liminf_{L\to\infty}\left|F^{(\beta_{\mathcal{S}},L,\beta_{\mathcal{R}})}_{\tilde{\mathbf{R}}_s}(x)-\sigma(x)\right|=\liminf_{\beta_{\mathcal{R}}\to\infty}1-F^{(\beta_{\mathcal{S}},L,\beta_{\mathcal{R}})}_{\tilde{\mathbf{R}}_s}(x)>0. \tag{3.138}$$

□

Proposition 7. *Given the assumptions of Theorem 5, the EED of* $\tilde{\mathbf{R}}_n^{-1}$ *converges uniformly as* $n_{\mathcal{D}}\to\infty$*: there is an asymptotic EED* $F^{(\beta_{\mathcal{R}},L)}_{\tilde{\mathbf{R}}_n^{-1}}(x)$*, such that*

$$\lim_{n_{\mathcal{D}}\to\infty}\sup_x\left|F^{(n_{\mathcal{D}},\beta_{\mathcal{R}},L)}_{\tilde{\mathbf{R}}_n^{-1}}(x)-F^{(\beta_{\mathcal{R}},L)}_{\tilde{\mathbf{R}}_n^{-1}}(x)\right|=0 \text{ almost surely.} \tag{3.139}$$

Moreover, the following statements hold:

- *For* $\beta_{\mathcal{R}}\in\Omega(L^{1+\varepsilon})$ *and* $L\to\infty$*, the asymptotic EED of* $\tilde{\mathbf{R}}_n^{-1}$*,* $F^{(L,\beta_{\mathcal{R}})}_{\tilde{\mathbf{R}}_n^{-1}}(x)$*, converges pointwise to*

$$F_{\tilde{\mathbf{R}}_n^{-1}}(x)=\sigma(x-1). \tag{3.140}$$

- *For any pair $(L, \beta_{\mathcal{R}})$, the following inequality holds for $x < 1$:*

$$1 - F_{\tilde{\mathbf{R}}_n^{-1}}^{(L,\beta_{\mathcal{R}})}(x) > x. \tag{3.141}$$

Proof of Proposition 7. The matrix $\mathbf{R}_{n,l}$ is defined as $\mathbf{R}_{n,l} \triangleq \frac{\alpha^l}{n_{\mathcal{R}}^l}\mathbf{H}_1 \cdots \mathbf{H}_l \mathbf{H}_l^H \cdots \mathbf{H}_1^H$, where $l \in \{1, \ldots, L\}$. The EED of $\mathbf{R}_{n,l}$ converges uniformly and almost surely to an asymptotic EED. This and the corresponding implicit equation for the Stieltjes transform of the asymptotic EED of $\mathbf{R}_{n,l}$ follows again from Lemma 4, where N is identified with l, $\mathbf{M}_n$ with $\mathbf{H}_l$ and k_n with $n_{\mathcal{R}}$. Thus, we obtain the following equation for the Stieltjes transform:

$$G_{\mathbf{R}_{n,l}}^{(\beta_{\mathcal{R}},l)}(s) \cdot \left(\Psi_{\mathbf{R}_{n,l}} + s\right) = 1, \tag{3.142}$$

with

$$\Psi_{\mathbf{R}_{n,l}} = \left(\frac{sG_{\mathbf{R}_{n,l}}^{(\beta_{\mathcal{R}},l)}(s) - 1}{\beta_{\mathcal{R}}} + 1\right)^l. \tag{3.143}$$

Again, we assume $s > 0$. Once more, the obtained limiting Stieltjes transform generalizes to its full domain, since it is an analytic function. Since $0 < sG_{\mathbf{R}_{n,l}}^{(\beta_{\mathcal{R}},l)}(s) < 1$, the following statements apply:

- For every fixed $\beta_{\mathcal{R}}$ and L, the following inequality holds for all $l \in \{1, \ldots, L\}$ and $s > 0$:

$$\left|\Psi_{\mathbf{R}_{n,l}} - 1\right| < \left|\left(-\frac{1}{\beta_{\mathcal{R}}} + 1\right)^L - 1\right|. \tag{3.144}$$

- If $\beta_{\mathcal{R}} \in \Omega\left(L^{1+\varepsilon}\right)$, then

$$\lim_{L\to\infty}\left(-\frac{1}{\beta_{\mathcal{R}}} + 1\right)^L = 1. \tag{3.145}$$

This implies that the L terms *converge uniformly* as L and $\beta_{\mathcal{R}}$ tend to infinity:

$$\lim_{L\to\infty}\max_{l\in\{1,\ldots,L\}}\left|\Psi_{\mathbf{R}_{n,l}} - 1\right| = 0, \tag{3.146}$$

which in turn yields

$$\lim_{L\to\infty} \max_{l\in\{1,\dots,L\}} \left| G^{(\beta_{\mathcal{R}},l)}_{\mathbf{R}_{n,l}}(s) - \frac{1}{s+1} \right| = 0. \tag{3.147}$$

In a next step, it is concluded that also $F_{\tilde{\mathbf{R}}_n}(x) = \sigma(x-1)$, if $\beta_{\mathcal{R}} \in \Omega\left(\beta_{\mathcal{R}}^{1+\varepsilon}\right)$. To this end, the following Lemmata 6 and 7, whose proofs are provided in Appendix B.2, are stated. They allow to consider the nuclear norm of the difference between $\tilde{\mathbf{R}}_n$ and the identity matrix.

Lemma 6. *Let $\mathbf{A}^{(n)}(\gamma) \in \mathbb{C}^{n\times n}$ be a sequence of positive semidefinite random matrices with parameter γ whose EED converges uniformly to the asymptotic EED $F^{(\gamma)}_{\mathbf{A}}(x)$ almost surely. Assume for all γ i) $\lim_{n\to\infty} \frac{1}{n}\mathrm{Tr}[\mathbf{A}(\gamma)] = 1$ and ii) $\lambda^{(\gamma)}_{\max} \triangleq \lim_{n\to\infty} \lambda_{\max}\{\mathbf{A}(\gamma)\} < \infty$ almost surely. Then, the following types of convergence are equivalent[8]:*

1. $\lim_{\gamma\to\infty} d(\gamma) = 0$, *where $d(\gamma)$ fulfills for every γ:*

$$\lim_{n\to\infty} \frac{1}{n} \|\mathbf{I}_n - \mathbf{A}(\gamma)\|_* = d(\gamma) \text{ almost surely.}$$

2. $\lim_{\gamma\to\infty} |F^{(\gamma)}_{\mathbf{A}}(x) - \sigma(x-1)| = 0$ *for all x.*
3. $\lim_{\gamma\to\infty} \int_0^\infty |F^{(\gamma)}_{\mathbf{A}}(x) - \sigma(x-1)| \cdot dx = 0.$

Lemma 7. *$\tilde{\mathbf{R}}_n$ fulfills the assumptions of Lemma 6 when the parameter γ is identified with L.*

As a consequence of that, there exists for every $\varepsilon > 0$ an L_0, such that there exists for all $L \geq L_0$ almost surely an $n^{(0)}_{\mathcal{D}}(L_0)$, such that for all $n_{\mathcal{D}} \geq n^{(0)}_{\mathcal{D}}(L_0)$

$$\frac{1}{n_{\mathcal{D}}} \left\| \mathbf{I}_{n_{\mathcal{D}}} - \tilde{\mathbf{R}}_n \right\|_* = \frac{1}{n_{\mathcal{D}}} \left\| \mathbf{I}_{n_{\mathcal{D}}} - \frac{1-\alpha}{1-\alpha^{L+1}} \cdot \mathbf{R}_n \right\|_* \tag{3.148}$$

$$= \frac{1}{n_{\mathcal{D}}} \left\| \frac{1-\alpha}{1-\alpha^{L+1}} \sum_{l=0}^{L} \left(\alpha^l \cdot \mathbf{I}_{n_{\mathcal{D}}} - \mathbf{R}_{n,l} \right) \right\|_* \tag{3.149}$$

$$= \frac{1}{n_{\mathcal{D}}} \left\| \frac{1-\alpha}{1-\alpha^{L+1}} \sum_{l=0}^{L} \alpha^l \left(\mathbf{I}_{n_{\mathcal{D}}} - \frac{1}{n^l_{\mathcal{R}}} \mathbf{H}_1 \cdots \mathbf{H}_l \mathbf{H}_l^H \cdots \mathbf{H}_1^H \right) \right\|_* \tag{3.150}$$

[8] $\|\mathbf{M}\|_* = \sum_i \sigma_i\{\mathbf{M}\}$ denotes the nuclear norm.

$$\leq \frac{1-\alpha}{1-\alpha^{L+1}} \sum_{l=0}^{L} \frac{\alpha^l}{n_{\mathcal{D}}} \left\| \mathbf{I}_{n_{\mathcal{D}}} - \frac{1}{n_{\mathcal{R}}^l} \mathbf{H}_1 \cdots \mathbf{H}_l \mathbf{H}_l^H \cdots \mathbf{H}_1^H \right\|_* \tag{3.151}$$

$$= \frac{1-\alpha}{1-\alpha^{L+1}} \sum_{l=0}^{L} \alpha^l \int_0^\infty \left| F_{\mathbf{R}_{n,l}}^{(n_{\mathcal{D}},\beta_{\mathcal{R}},l)}(x) - \sigma(x-1) \right| dx \tag{3.152}$$

$$\leq \frac{1-\alpha}{1-\alpha^{L+1}} \sum_{l=0}^{L} \alpha^l \Bigg(\int_0^\infty \left| F_{\mathbf{R}_{n,l}}^{(\beta_{\mathcal{R}},l)}(x) - \sigma(x-1) \right| dx + \int_0^\infty \left| F_{\mathbf{R}_{n,l}}^{(n_{\mathcal{D}},\beta_{\mathcal{R}},l)}(x) - F_{\mathbf{R}_{n,l}}^{(\beta_{\mathcal{R}},l)}(x) \right| dx \Bigg) \tag{3.153}$$

$$\leq \max_{l\in\{1,\ldots,L\}} \left\{ \int_0^\infty \left| F_{\mathbf{R}_{n,l}}^{(\beta_{\mathcal{R}},l)}(x) - \sigma(x-1) \right| dx \right\} + \max_{l\in\{1,\ldots,L\}} \left\{ \int_0^\infty \left| F_{\mathbf{R}_{n,l}}^{(n_{\mathcal{D}},\beta_{\mathcal{R}},l)}(x) - F_{\mathbf{R}_{n,l}}^{(\beta_{\mathcal{R}},l)}(x) \right| dx \right\} \tag{3.154}$$

$$< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \tag{3.155}$$

Expression (3.151) is obtained by applying the triangle inequality and using the homogeneity of the nuclear norm. Equality between (3.151) and (3.152) is established by the following chain of identities for a positive semidefinite matrix $\mathbf{A} \in \mathbb{C}^{n\times n}$ (σ_i and λ_i denote ith singular- and eigenvalue, respectively):

$$\frac{1}{n}\|\mathbf{I}_n - \mathbf{A}(\gamma)\|_* = \frac{1}{n}\sum_{i=1}^{n} \sigma_i\{\mathbf{I}_n - \mathbf{A}(\gamma)\}$$

$$= \frac{1}{n}\sum_{i=1}^{n} |\lambda_i\{\mathbf{I}_n - \mathbf{A}(\gamma)\}| = \frac{1}{n}\sum_{i=1}^{n} |1 - \lambda_i\{\mathbf{A}(\gamma)\}| \tag{3.156}$$

$$= \frac{1}{n} \sum_{i:\lambda_i\{\mathbf{A}(\gamma)\}\leq 1} (1 - \lambda_i\{\mathbf{A}(\gamma)\}) + \frac{1}{n} \sum_{i:\lambda_i\{\mathbf{A}(\gamma)\}>1} (\lambda_i\{\mathbf{A}(\gamma)\} - 1) \tag{3.157}$$

$$= \int_0^1 |F_{\mathbf{A}}^{(n,\gamma)}(x)| dx + \int_1^\infty |F_{\mathbf{A}}^{(n,\gamma)}(x) - 1| dx \tag{3.158}$$

$$= \int_0^\infty |F_{\mathbf{A}}^{(n,\gamma)}(x) - \sigma(x-1)| dx. \tag{3.159}$$

Eq. (3.153) follows by adding and substracting $F_{\mathbf{R}_{n,l}}^{(\beta_{\mathcal{R}},l)}(x)$ and repeated application of the triangle inequality. In (3.154), the individual integrals are upper-bounded by the largest ones. In the final step, both terms in (3.154) are upper-bounded by $\varepsilon/2$. For the first term, one can fix an L_0, such that this upper-bound holds for all $L \geq L_0$ by (3.147) and Lemma 6. For fixed L (and thus $\beta_{\mathcal{R}}$), one can finally choose an $n_{\mathcal{D}}^{(0)}(L)$ large enough, such that for all $n_{\mathcal{D}} \geq n_{\mathcal{D}}^{(0)}(L)$ the second term is smaller than $\varepsilon/2$. Note that $\lim_{n_{\mathcal{D}}\to\infty} \lambda_{\max}\{\tilde{\mathbf{R}}_n\} < \infty$ for fixed L almost surely according to Lemma 7. Therefore, the limit can be taken inside the integral due to the uniform convergence of $F_{\tilde{\mathbf{R}}_n}^{(n_{\mathcal{D}},\beta_{\mathcal{R}},L)}$ to $F_{\tilde{\mathbf{R}}_n}^{(\beta_{\mathcal{R}},L)}$). We have thus shown that $F_{\mathbf{R}_n}^{(\beta_{\mathcal{R}},L)}(x)$ converges pointwise to $\sigma(x-1)$ as $L \to \infty$. Since the eigenvalues of the inverse of $\tilde{\mathbf{R}}_n$ are the inverse eigenvalues of $\tilde{\mathbf{R}}_n$, i.e., $\lambda_k\{\tilde{\mathbf{R}}_n^{-1}\} = \lambda_k^{-1}\{\tilde{\mathbf{R}}_n\}$, one can conclude that also $F_{\tilde{\mathbf{R}}_n^{-1}}^{(\beta_{\mathcal{R}},L)}(x)$ converges pointwise to $\sigma(x-1)$.

The last part of the proposition is established as follows. For every pair $(\beta_{\mathcal{R}}, L)$, we have

$$\lim_{n_{\mathcal{D}}\to\infty} \frac{1}{n_{\mathcal{D}}} \sum_{i=1}^{n_{\mathcal{D}}} \lambda_i \left\{\tilde{\mathbf{R}}_n\right\} = \lim_{n_{\mathcal{D}}\to\infty} \frac{1}{n_{\mathcal{D}}} \sum_{i=1}^{n_{\mathcal{D}}} \lambda_i \left\{\frac{1-\alpha}{1-\alpha^{L+1}} \mathbf{R}_n\right\} \tag{3.160}$$

$$= \lim_{n_{\mathcal{D}}\to\infty} \frac{1}{n_{\mathcal{D}}} \mathrm{Tr}\left[\frac{1-\alpha}{1-\alpha^{L+1}} \cdot \mathbf{R}_n\right] \tag{3.161}$$

$$= \frac{1-\alpha}{1-\alpha^{L+1}} \sum_{l=0}^{L} \alpha^l \lim_{n_{\mathcal{D}}\to\infty} \frac{1}{n_{\mathcal{D}}} \mathrm{Tr}\left[\frac{1}{n_{\mathcal{R}}^l} \mathbf{H}_1 \cdots \mathbf{H}_l \mathbf{H}_l^H \cdots \mathbf{H}_1^H\right] \tag{3.162}$$

$$= \frac{1-\alpha}{1-\alpha^{L+1}} \sum_{l=0}^{L} \alpha^l = 1. \tag{3.163}$$

The evaluation of the limit follows from the following lemma, which is proved in Appendix B.2.

Lemma 8. *Let the matrices* $\mathbf{H}_2, \mathbf{H}_3, \ldots, \mathbf{H}_{L+1}$ *be as in Theorem 5. Moreover, let* $\mathbf{H}_1$ *be an arbitrary* $n_{\mathcal{D}} \times n_{\mathcal{R}}$ *random matrix that fulfills* $\lim_{n_{\mathcal{R}}\to\infty} n_{\mathcal{D}}^{-1} \mathrm{Tr}\left[n_{\mathcal{R}}^{-1} \mathbf{H}_1^H \mathbf{H}_1\right] = 1$ *for every fixed ratio* $n_{\mathcal{R}}/n_{\mathcal{D}}$ *almost surely. Define* $\mathbf{A}_l \triangleq \frac{1}{n_{\mathcal{S}} n_{\mathcal{R}}^L} \mathbf{H}_1 \mathbf{H}_2 \cdots \mathbf{H}_{l+1} \mathbf{H}_{l+1}^H \cdots \mathbf{H}_2^H \mathbf{H}_1^H$. *Then, for every pair of fixed ratios* $n_{\mathcal{R}}/n_{\mathcal{S}}$ *and* $n_{\mathcal{D}}/n_{\mathcal{S}}$ *and any* $l \in \{1, \ldots, L\}$

$$\lim_{n_{\mathcal{D}}\to\infty} n_{\mathcal{D}}^{-1} \mathrm{Tr}\left[\mathbf{A}_l\right] = 1 \text{ almost surely.} \tag{3.164}$$

Eq. (3.163) implies for $x > 1$ that

$$1 - F_{\tilde{\mathbf{R}}_n}^{(\beta_{\mathcal{R}},L)}(x) < \frac{1}{x}, \tag{3.165}$$

or, eqivalently, for $x < 1$ that

$$1 - F^{(\beta_{\mathcal{R}},L)}_{\tilde{\mathbf{R}}_n^{-1}}(x) < x, \tag{3.166}$$

which proves the claim. □

With Propositions 6 & 7 ready to hand, the theorem can be proved.

Proof of Theorem 5. We consider each of the three cases separately in the following.

Case $\beta_{\mathcal{R}} \in \mathcal{O}\left(L^{1-\varepsilon}\right)$: It has to be shown that $\lim_{L\to\infty} c_L^{\mathrm{AF}} = 0$. Since $n_{\mathcal{D}}^{-1} R_L^{\mathrm{AF}}$ is known to converge to a nonrandom constant for every L [24] almost surely as $n_{\mathcal{D}} \to \infty$, this is implied by $\lim_{L\to\infty} \lim_{n_{\mathcal{D}}\to\infty} n_{\mathcal{D}}^{-1} \mathsf{E}\left[R_L^{\mathrm{AF}}\right] = 0$.

First, the following lemma is stated:

Lemma 9. *The supremum of the set of ergodically achievable sum-rates* $\mathsf{E}\left[R_L^{\mathrm{AF}}\right]$ *is for all* $L_0 \in \{1, \ldots, L\}$ *upper-bounded according to*

$$\mathsf{E}\left[R_L^{\mathrm{AF}}\right] \leq \mathsf{E}\left[\log\det\left(\mathbf{I}_{n_{\mathcal{D}}} + \frac{P_L s_1 \cdots s_{L_0} \alpha^L}{\sigma_{\mathrm{w}}^2 n_{\mathcal{S}} n_{\mathcal{R}}^{L-L_0} \sum_{l=0}^{L_0} s_1 \cdots s_l \alpha^l} \tilde{\mathbf{V}}_{L_0}^H \mathbf{H}_{L_0+1} \cdots \mathbf{H}_{L+1} \mathbf{H}_{L+1}^H \cdots \mathbf{H}_{L_0+1}^H \tilde{\mathbf{V}}_{L_0}\right)\right], \tag{3.167}$$

where $\tilde{\mathbf{V}}_{L_0}^H$ *is an* $n_{\mathcal{D}} \times n_{\mathcal{R}}$ *matrix with orthonormal rows that is obtained through the following sequence of singular value decompositions:*

$$\tilde{\mathbf{H}}_1 = \mathbf{H}_1 = \mathbf{U}_1 \mathbf{S}_1 \mathbf{V}_1^H = \mathbf{U}_1 \tilde{\mathbf{S}}_1 \tilde{\mathbf{V}}_1^H \tag{3.168}$$
$$\tilde{\mathbf{H}}_2 = \tilde{\mathbf{V}}_1^H \mathbf{H}_2 = \mathbf{U}_2 \mathbf{S}_2 \mathbf{V}_2^H = \mathbf{U}_2 \tilde{\mathbf{S}}_2 \tilde{\mathbf{V}}_2^H \tag{3.169}$$
$$\tilde{\mathbf{H}}_3 = \tilde{\mathbf{V}}_2^H \mathbf{H}_3 = \mathbf{U}_3 \mathbf{S}_3 \mathbf{V}_3^H = \mathbf{U}_3 \tilde{\mathbf{S}}_3 \tilde{\mathbf{V}}_3^H \tag{3.170}$$
$$\vdots$$

The matrices $\tilde{\mathbf{S}}_k$ *and* $\tilde{\mathbf{V}}_k$ *correspond to the first* $n_{\mathcal{D}}$ *columns of* $\mathbf{S}_k$ *and* $\mathbf{V}_k$, *respectively. Furthermore,* $s_k \triangleq n_{\mathcal{R}}^{-1} \mathrm{Tr}\left[\tilde{\mathbf{H}}_k \tilde{\mathbf{H}}_k^H\right]$ *for* $k \in \{1, \ldots, L_0\}$ *and* $\mathsf{E}[s_k] = n_{\mathcal{D}}$ *for all* k.

Remark: This upper-bound corresponds to the capacity of an equivalent network with

- noiseless relay stages $\mathcal{R}_{L_0+1}, \ldots, \mathcal{R}_L$,

- the relay stage $\mathcal{R}_{L_0}$ replaced by an equivalent destination stage with $n_{\mathcal{D}}$ antennas, whose preceding hop has a channel matrix $\tilde{\mathbf{V}}_{L_0}^H \mathbf{H}_{L_0+1}$,

- white noise of power $\sigma_{\mathrm{w}}^2 \cdot \sum_{l=0}^{L_0} \alpha^l$ at the equivalent destination stage (corresponds to the average receive power at the destination stage as it would have been introduced by the removed relay stages $\mathcal{R}_{L_0}, \ldots, \mathcal{R}_1$).

This lemma is applied for $L_0 = \lfloor L/2 \rfloor$ to obtain

$$\begin{aligned} \mathsf{E}\left[\frac{R_L^{\mathrm{AF}}}{n_{\mathcal{D}}}\right] &\leq \frac{1}{n_{\mathcal{D}}} \mathsf{E}\left[\log\det\left(\mathbf{I}_{n_{\mathcal{D}}} + \rho_L \cdot \mathbf{A}\right)\right] \\ &= \frac{1}{n_{\mathcal{D}}} \mathsf{E}\left[\sum_{i:\lambda_i\{\mathbf{A}\}\leq\delta_1} \log\left(1+\rho_L \lambda_i\{\mathbf{A}\}\right)\right] \\ &\quad + \frac{1}{n_{\mathcal{D}}} \mathsf{E}\left[\sum_{i:\lambda_i\{\mathbf{A}\}>\delta_1} \log\left(1+\rho_L \lambda_i\{\mathbf{A}\}\right)\right], \end{aligned} \tag{3.171}$$

where

$$\mathbf{A} \triangleq \frac{1}{n_{\mathcal{S}} n_{\mathcal{R}}^{\lceil L/2 \rceil}} \tilde{\mathbf{V}}_{\lfloor L/2 \rfloor}^H \mathbf{H}_{\lfloor L/2 \rfloor+1} \cdots \mathbf{H}_{L+1} \mathbf{H}_{L+1}^H \cdots \mathbf{H}_{\lfloor L/2 \rfloor+1}^H \tilde{\mathbf{V}}_{\lfloor L/2 \rfloor} \tag{3.172}$$

and

$$\rho_L \triangleq \frac{P_L s_1 \cdots s_{\lfloor L/2 \rfloor} \alpha^L}{\sigma_{\mathrm{w}}^2 \cdot \sum_{l=0}^{\lfloor L/2 \rfloor} s_1 \cdots s_l \alpha^l}, \tag{3.173}$$

with $\tilde{\mathbf{V}}_k$ and s_k constructed as in Lemma 9.

Fix $\varepsilon > 0$ arbitrarily small, and fix $\delta_1 > 0$ sufficiently small, such that

$$\log(1 + 2 \cdot \mathrm{snr}_{\mathcal{D}} \cdot \delta_1) < \frac{\varepsilon}{2}. \tag{3.174}$$

Next, fix $\delta_2 > 0$ sufficiently small, such that

$$\delta_2 \cdot \log\left(1 + 2 \cdot \frac{\mathrm{snr}_{\mathcal{D}}}{\delta_2}\right) < \frac{\varepsilon}{2}. \tag{3.175}$$

The two sums in (3.171) are considered individually. The first sum fulfills independently of L

$$\lim_{n_{\mathcal{D}}\to\infty} \frac{1}{n_{\mathcal{D}}} \sum_{i:\lambda_i\{\mathbf{A}\}\leq\delta_1} \log\left(1+\rho_L\lambda_i\left\{\mathbf{A}\right\}\right) \leq \log\left(1+2\cdot \mathrm{snr}_{\mathcal{D}}\cdot\delta_1\right) < \frac{\varepsilon}{2} \text{ almost surely.} \tag{3.176}$$

This bound is obtained in two steps:

- Since the sum comprises no more than $n_{\mathcal{D}}$ terms that are upper-bounded by $\log\left(1+\rho_L\cdot\delta_1\right)$ each, one obtains

$$\frac{1}{n_{\mathcal{D}}} \sum_{i:\lambda_i\{\mathbf{A}\}\leq\delta_1} \log\left(1+\rho_L\lambda_i\left\{\mathbf{A}\right\}\right) \leq \log\left(1+\rho_L\cdot\delta_1\right). \tag{3.177}$$

- Since $\lim_{n_{\mathcal{D}}\to\infty} s_k = 1$ almost surely for all k, one obtains

$$\lim_{n_{\mathcal{D}}\to\infty} \rho_L = \frac{P_L\alpha^L}{\sigma_{\mathrm{w}}^2\cdot\sum_{l=0}^{\lfloor L/2\rfloor}\alpha^l} = \frac{P_L\alpha^L}{\sigma_{\mathrm{w}}^2\cdot\frac{1-\alpha^{\lfloor L/2\rfloor+1}}{1-\alpha}} \tag{3.178}$$

$$= \mathrm{snr}_{\mathcal{D}}\cdot\frac{1-\alpha^{L+1}}{1-\alpha^{\lfloor L/2\rfloor+1}} \leq \mathrm{snr}_{\mathcal{D}}\cdot\frac{1-\alpha^L}{1-\sqrt{\alpha^L}} < 2\cdot\mathrm{snr}_{\mathcal{D}} \text{ almost surely.} \tag{3.179}$$

The last inequality follows, since $0\leq\alpha^L = \left(P_L/(\sigma_{\mathrm{w}}^2+P_L)\right)^L < 1$.

For the second sum, there exists for every $\varepsilon>0$ an L_0, such that for all $L\geq L_0$

$$\lim_{n_{\mathcal{D}}\to\infty} \frac{1}{n_{\mathcal{D}}} \sum_{i:\lambda_i\{\mathbf{A}\}>\delta_1} \log\left(1+\rho_L\lambda_i\left\{\mathbf{A}\right\}\right) \leq \delta_2\cdot\log\left(1+2\cdot\frac{\mathrm{snr}_{\mathcal{D}}}{\delta_2}\right) \leq \frac{\varepsilon}{2} \text{ almost surely.} \tag{3.180}$$

For establishing this upper-bound, the following lemma is used:

Lemma 10. *Let* $\mathbf{H}_2,\ldots,\mathbf{H}_{L+1}$ *be as in Theorem 5. Moreover, define an* $n_{\mathcal{D}}\times n_{\mathcal{R}}$ *random matrix* $\mathbf{X}$ *whose elements follow a distribution independent of the elements of* $\mathbf{H}_2,\cdots,\mathbf{H}_{L+1}$ *and must fulfill* $\frac{1}{n_{\mathcal{S}}}\mathbf{X}\mathbf{X}^H = \mathbf{I}_{n_{\mathcal{D}}}$. *Then, the EEDs of* $\mathbf{A}\triangleq\frac{1}{n_{\mathcal{R}}^L n_{\mathcal{D}}}\mathbf{X}\mathbf{H}_2\cdots\mathbf{H}_{L+1}\mathbf{H}_{L+1}^H\cdots\mathbf{H}_2^H\mathbf{X}^H$ *and* $\mathbf{B}\triangleq\frac{1}{n_{\mathcal{R}}^{L-1}n_{\mathcal{D}}}\tilde{\mathbf{H}}_2\mathbf{H}_3\mathbf{H}_4\cdots\mathbf{H}_{L+1}\mathbf{H}_{L+1}^H\cdots\mathbf{H}_4^H\mathbf{H}_3^H\tilde{\mathbf{H}}_2^H$ *converge, as* $n_{\mathcal{D}}\to\infty$ *and* $\beta_{\mathcal{S}}$ *and* $\beta_{\mathcal{R}}$ *are fixed, uniformly and almost surely to the same asymptotic EED, where* $\tilde{\mathbf{H}}_2$ *is the matrix that contains the first* $n_{\mathcal{D}}$ *rows of* $\mathbf{H}_2$.

The bound (3.180) is obtained in four steps:

- Due to the concavity of the log-function, Jensen's inequality can be invoked to establish:

$$\frac{n}{n_{\mathcal{D}}}\frac{1}{n}\sum_{i=1}^{n}\log\left(1+\rho_L\lambda_i\{\mathbf{A}\}\right) \leq \frac{n}{n_{\mathcal{D}}}\log\left(1+\rho_L\frac{1}{n}\sum_{i=1}^{n}\lambda_i\{\mathbf{A}\}\right) \tag{3.181}$$

$$\leq \frac{n}{n_{\mathcal{D}}}\log\left(1+\rho_L\frac{1}{n}\sum_{i=1}^{n_{\mathcal{D}}}\lambda_i\{\mathbf{A}\}\right) = \frac{n}{n_{\mathcal{D}}}\log\left(1+\rho_L\frac{n_{\mathcal{D}}}{n}\frac{1}{n_{\mathcal{D}}}\mathrm{Tr}[\mathbf{A}]\right), \tag{3.182}$$

  where we choose n as the maximal i that fulfills $\lambda_i\{\mathbf{A}\} > \delta_1$.

- From Lemma 10, we know that Proposition 6 applies also, if $\tilde{\mathbf{R}}_s$ is replaced by the random matrix $\mathbf{A}$. Note that $\beta_{\mathcal{R}} \in \mathcal{O}\left(L^{1-\varepsilon}\right)$, if and only if also $\lfloor L/2 \rfloor \in \mathcal{O}\left(L^{1-\varepsilon}\right)$. Thus, there exists by Proposition 6 an L_0, such that there exists for all $L \geq L_0$ almost surely an $n_{\mathcal{D}}^{(0)}(L_0)$, such that for all $n_{\mathcal{D}} \geq n_{\mathcal{D}}^{(0)}(L_0)$ the fraction of eigenvalues of $\mathbf{A}$ larger than δ_1 fulfills

$$\frac{n}{n_{\mathcal{D}}} = 1 - F_{\mathbf{A}}^{(n_{\mathcal{D}},\beta_{\mathcal{R}},L)}(\delta_1) < \delta_2. \tag{3.183}$$

- Moreover, Lemma 8 implies $\lim_{n_{\mathcal{D}}\to\infty} n_{\mathcal{D}}^{-1}\mathrm{Tr}\left[\mathbf{A}\right] = 1$ almost surely.
- Finally, we use again that $\lim_{n_{\mathcal{D}}\to\infty} \rho_L < 2 \cdot \mathrm{snr}_{\mathcal{D}}$ almost surely.

Thus, by combining (3.176) and (3.180), we have shown that there exists an L_0, such that there exists for all $L \geq L_0$ almost surely an $n_{\mathcal{D}}^{(0)}(L)$, such that for all $n_{\mathcal{D}} \geq n_{\mathcal{D}}^{(0)}(L_0)$ (3.171) fulfills

$$\mathsf{E}\left[\frac{R_L^{\mathrm{AF}}}{n_{\mathcal{D}}}\right] < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \tag{3.184}$$

Case $\beta_{\mathcal{R}} \in \Omega\left(L^{1+\varepsilon}\right)$: Define the Hermitian matrix

$$\mathbf{\Theta} \triangleq \tilde{\mathbf{R}}_n - \mathbf{I}_{n_{\mathcal{D}}}, \tag{3.185}$$

and rewrite it in terms of its eigenvalue decomposition:

$$\mathbf{\Theta} = \sum_{i=1}^{n_{\mathcal{D}}} \lambda_i\{\Theta\}\mathbf{u}_i\mathbf{u}_i^H = \underbrace{\sum_{i:\lambda_i\{\Theta\}>0} \lambda_i\{\Theta\}\mathbf{u}_i\mathbf{u}_i^H}_{\mathbf{\Theta}^+} + \underbrace{\sum_{i:\lambda_i\{\Theta\}\leq 0} \lambda_i\{\Theta\}\mathbf{u}_i\mathbf{u}_i^H}_{\mathbf{\Theta}^-}, \tag{3.186}$$

where $\mathbf{u}_i$ denotes the eigenvector that corresponds to the ith eigenvalue. Let $\mathbf{A}$ be an arbitrary positive semidefinite matrix with an asymptotic EED that depends on L and $\beta_{\mathcal{R}}$ and fulfills $\lim_{L\to\infty} F_{\mathbf{A}}^{(L,\beta_{\mathcal{R}})}(x) = F_{\mathbf{A}}(x)$. We show that the asymptotic EEDs $F_{\mathbf{A}+\boldsymbol{\Theta}}^{(L,\beta_{\mathcal{R}})}(x)$ and $F_{\mathbf{A}}^{(L,\beta_{\mathcal{R}})}(x)$ coincide in the limit $L \to \infty$:

$$\int_0^\infty \left| F_{\mathbf{A}+\boldsymbol{\Theta}}^{(L,\beta_{\mathcal{R}},n_{\mathcal{D}})}(x) - F_{\mathbf{A}}^{(L,\beta_{\mathcal{R}},n_{\mathcal{D}})}(x) \right| \cdot dx = \frac{1}{n_{\mathcal{D}}} \sum_{i=1}^{n_{\mathcal{D}}} \left| \lambda_i\{\mathbf{A}+\boldsymbol{\Theta}\} - \lambda_i\{\mathbf{A}\} \right| \tag{3.187}$$

$$\begin{aligned}
&= \frac{1}{n_{\mathcal{D}}} \sum_{i:\lambda_i\{\mathbf{A}+\boldsymbol{\Theta}\}>\lambda_i\{\mathbf{A}\}} \left(\lambda_i\{\mathbf{A}+\boldsymbol{\Theta}\} - \lambda_i\{\mathbf{A}\}\right) \\
&\qquad + \frac{1}{n_{\mathcal{D}}} \sum_{j:\lambda_j\{\mathbf{A}+\boldsymbol{\Theta}\}\leq\lambda_j\{\mathbf{A}\}} \left(\lambda_j\{\mathbf{A}\} - \lambda_j\{\mathbf{A}+\boldsymbol{\Theta}\}\right) && (3.188) \\
&\leq \frac{1}{n_{\mathcal{D}}} \sum_{i:\lambda_i\{\mathbf{A}+\boldsymbol{\Theta}\}>\lambda_i\{\mathbf{A}\}} \left(\lambda_i\{\mathbf{A}+\boldsymbol{\Theta}^+\} - \lambda_i\{\mathbf{A}\}\right) \\
&\qquad + \frac{1}{n_{\mathcal{D}}} \sum_{j:\lambda_j\{\mathbf{A}+\boldsymbol{\Theta}\}\leq\lambda_j\{\mathbf{A}\}} \left(\lambda_j\{\mathbf{A}\} - \lambda_j\{\mathbf{A}+\boldsymbol{\Theta}^-\}\right) && (3.189) \\
&\leq \frac{1}{n_{\mathcal{D}}} \sum_{i=1}^{n_{\mathcal{D}}} \left(\lambda_i\{\mathbf{A}+\boldsymbol{\Theta}^+\} - \lambda_i\{\mathbf{A}\}\right) \\
&\qquad + \frac{1}{n_{\mathcal{D}}} \sum_{j=1}^{n_{\mathcal{D}}} \left(\lambda_j\{\mathbf{A}\} - \lambda_j\{\mathbf{A}+\boldsymbol{\Theta}^-\}\right) && (3.190) \\
&= \frac{1}{n_{\mathcal{D}}} \sum_{i=1}^{n_{\mathcal{D}}} \left(\lambda_i\{\mathbf{A}+\boldsymbol{\Theta}^+\} - \lambda_i\{\mathbf{A}+\boldsymbol{\Theta}^-\}\right) && (3.191) \\
&= \frac{1}{n_{\mathcal{D}}} \mathrm{Tr}[\mathbf{A}+\boldsymbol{\Theta}^+ - \mathbf{A} - \boldsymbol{\Theta}^-] && (3.192) \\
&= \frac{1}{n_{\mathcal{D}}} \mathrm{Tr}[\boldsymbol{\Theta}^+ - \boldsymbol{\Theta}^-] = \frac{1}{n_{\mathcal{D}}} \sum_{i=1}^{n_{\mathcal{D}}} |\lambda_i\{\boldsymbol{\Theta}\}| = \frac{1}{n_{\mathcal{D}}} \|\boldsymbol{\Theta}\|_*. && (3.193)
\end{aligned}$$

The first inequality follows, since removing the negative/positive definite part of $\boldsymbol{\Theta}$ can only increase/decrease each individual eigenvalue of $\mathbf{A}+\boldsymbol{\Theta}$. The second inequality follows, since all added terms are nonnegative. With Proposition 7 and the equivalence of 1. and 2. in Lemma 6 we can thus conclude that there exists for every $\varepsilon > 0$ an L_0, such that for all $L \geq L_0$

$$\lim_{n_{\mathcal{D}}\to\infty} \int_0^\infty \left| F_{\mathbf{A}+\boldsymbol{\Theta}}^{(L,\beta_{\mathcal{R}},n_{\mathcal{D}})}(x) - F_{\mathbf{A}}^{(L,\beta_{\mathcal{R}},n_{\mathcal{D}})}(x) \right| \cdot dx$$

$$\leq \lim_{n_{\mathcal{D}}\to\infty} \frac{1}{n_{\mathcal{D}}} \|\mathbf{I}_{n_{\mathcal{D}}} - \tilde{\mathbf{R}}_n\|_* < \varepsilon \text{ almost surely,} \tag{3.194}$$

which implies that $F_{\mathbf{A}+\mathbf{\Theta}}(x)$ converges pointwise to $F_{\mathbf{A}}(x)$.

In the following, we use the identity

$$R_L^{\mathrm{AF}} = \log\det(\mathbf{I}_{n_{\mathcal{D}}} + \mathrm{snr}_{\mathcal{D}} \cdot \tilde{\mathbf{R}}_n^{-1}\tilde{\mathbf{R}}_s) = \log\det(\tilde{\mathbf{R}}_n + \mathrm{snr}_{\mathcal{D}} \cdot \tilde{\mathbf{R}}_s) - \log\det(\tilde{\mathbf{R}}_n). \tag{3.195}$$

We consider the first term on the right hand side and identify $\mathbf{A}$ with $\mathbf{I}_{n_{\mathcal{D}}} + \mathrm{snr}_{\mathcal{D}} \cdot \mathbf{R}_s$. Thus, we obtain

$$\frac{1}{n_{\mathcal{D}}} \log\det\left(\mathbf{I}_{n_{\mathcal{D}}} + \mathbf{\Theta} + \mathrm{snr}_{\mathcal{D}} \cdot \tilde{\mathbf{R}}_s\right) = \frac{1}{n_{\mathcal{D}}} \sum_{i=1}^{n_{\mathcal{D}}} \log(1 + \lambda_i\{\mathbf{\Theta} + \mathrm{snr}_{\mathcal{D}} \cdot \tilde{\mathbf{R}}_s\}) \tag{3.196}$$

$$= \int_0^{\left(1+\sqrt{\beta_{\mathcal{S}}^{-1}}\right)^2} \log(1+x) \cdot dF_{\mathbf{\Theta}+\mathrm{snr}_{\mathcal{D}}\cdot\tilde{\mathbf{R}}_s}^{(n_{\mathcal{D}},\beta_{\mathcal{S}},L,\beta_{\mathcal{R}})}(x)$$

$$+ \frac{1}{n_{\mathcal{D}}} \sum_{i:\lambda_i\{\mathbf{\Theta}+\tilde{\mathbf{R}}_s\}>(1+\sqrt{\beta_{\mathcal{S}}^{-1}})^2} \log(1 + \lambda_i\{\mathbf{\Theta} + \mathrm{snr}_{\mathcal{D}} \cdot \tilde{\mathbf{R}}_s\}), \tag{3.197}$$

where the integration interval corresponds to the support of $\partial F_{\mathrm{MP}}^{(\beta_{\mathcal{S}})}(x)/\partial x$. We drop the sum, use the fact $F_{\mathbf{\Theta}+\tilde{\mathbf{R}}_s}^{(\beta_{\mathcal{S}})}(x)$ converges pointwise to $F_{\tilde{\mathbf{R}}_s}^{(\beta_{\mathcal{S}})}(x)$ as $L \to \infty$, and thus obtain with Proposition 6 that there exists for every $\varepsilon > 0$ an L_0, such that for all $L \geq L_0$ almost surely

$$\lim_{n_{\mathcal{D}}\to\infty} \frac{1}{n_{\mathcal{D}}} \log\det\left(\mathbf{I}_{n_{\mathcal{D}}} + \mathbf{\Theta} + \mathrm{snr}_{\mathcal{D}} \cdot \tilde{\mathbf{R}}_s\right) > \int_0^{\left(1+\sqrt{\beta_{\mathcal{S}}^{-1}}\right)^2} \log(1 + \mathrm{snr}_{\mathcal{D}} \cdot x) \cdot dF_{\mathrm{MP}}^{(\beta_{\mathcal{S}})}(x) - \frac{\varepsilon}{2}. \tag{3.198}$$

Next, we investigate the second term in (3.195). Application of Jensen's inequality yields

$$-\frac{1}{n_{\mathcal{D}}} \log\det(\mathbf{I}_{n_{\mathcal{D}}} + \mathbf{\Theta}) = -\frac{1}{n_{\mathcal{D}}} \sum_{i=1}^{n_{\mathcal{D}}} \log(1 + \lambda_i\{\mathbf{\Theta}\}) \tag{3.199}$$

$$\geq -\log\left(1 + \frac{1}{n_{\mathcal{D}}} \sum_i \lambda_i\{\mathbf{\Theta}\}\right) \tag{3.200}$$

$$= -\log\left(1 + \frac{1}{n_{\mathcal{D}}}\mathrm{Tr}[\boldsymbol{\Theta}]\right). \tag{3.201}$$

Since $|\mathrm{Tr}[\boldsymbol{\Theta}]| \leq \|\boldsymbol{\Theta}\|_*$, we obtain with (3.194) that there exists for every $\varepsilon > 0$ an L_0, such that for all $L \geq L_0$

$$-\lim_{n_{\mathcal{D}}\to\infty} \frac{1}{n_{\mathcal{D}}} \log\det\left(\mathbf{I}_{n_{\mathcal{D}}} + \boldsymbol{\Theta}\right) > -\frac{\varepsilon}{2} \text{ almost surely.} \tag{3.202}$$

We have shown that there exists for every $\varepsilon > 0$ an L_0, such that for all $L \geq L_0$

$$\begin{aligned}&\lim_{n_{\mathcal{D}}\to\infty} \frac{1}{n_{\mathcal{D}}}\left(\log\det\left(\mathbf{I}_{n_{\mathcal{D}}} + \boldsymbol{\Theta} + \mathrm{snr}_{\mathcal{D}}\cdot\tilde{\mathbf{R}}_s\right) - \log\det\left(\mathbf{I}_{n_{\mathcal{D}}} + \boldsymbol{\Theta}\right)\right)\\ &> \int_0^{\left(1+\sqrt{\beta_{\mathcal{S}}^{-1}}\right)^2} \log\left(1 + \mathrm{snr}_{\mathcal{D}}\cdot x\right)\cdot dF_{\mathrm{MP}}^{(\beta_{\mathcal{S}})}(x) - \varepsilon \text{ almost surely.}\end{aligned} \tag{3.203}$$

The integral evaluates to [76]

$$c_0 = \beta_{\mathcal{S}}\log\left(1 + \mathrm{snr}_{\mathcal{D}} - \frac{1}{4}\chi(\mathrm{snr}_{\mathcal{D}}, \beta_{\mathcal{S}})\right) \tag{3.204}$$

$$+\log\left(1 + \mathrm{snr}_{\mathcal{D}}\beta_{\mathcal{S}} - \frac{1}{4}\chi(\mathrm{snr}_{\mathcal{D}}, \beta_{\mathcal{S}})\right) - \frac{\log e}{4\mathrm{snr}_{\mathcal{D}}}\chi(\mathrm{snr}_{\mathcal{D}}, \beta_{\mathcal{S}}), \tag{3.205}$$

which is the sum-capacity of a single-hop multiple access channel with $n_{\mathcal{S}}$ single-antenna sources and an $n_{\mathcal{D}}$-antenna destination.

It remains to prove that the single-hop sum-capacity cannot be exceeded. To this end, we can consider the supremum of ergodically achievable sum-rates $\mathsf{E}[R_L^{\mathrm{AF}}]$ first, since $c_L^{\mathrm{AF}} = \lim_{n_{\mathcal{D}}\to\infty} \mathsf{E}[n_{\mathcal{D}}^{-1} R_L^{\mathrm{AF}}]$. We apply Lemma 9 for $L_0 = L$ to obtain

$$\frac{1}{n_{\mathcal{D}}}\mathsf{E}\left[R_L^{\mathrm{AF}}\right] \leq \frac{1}{n_{\mathcal{D}}}\mathsf{E}\left[\log\det\left(\mathbf{I}_{n_{\mathcal{D}}} + \rho_L\cdot\mathbf{A}\right)\right], \tag{3.206}$$

where

$$\mathbf{A} = \frac{1}{n_{\mathcal{S}} n_{\mathcal{R}}}\tilde{\mathbf{V}}_L^H \mathbf{H}_{L+1}\mathbf{H}_{L+1}^H\tilde{\mathbf{V}}_L \tag{3.207}$$

and

$$\rho_L = \frac{P_L s_1 \cdots s_L \alpha^L}{\sigma_{\mathrm{w}}^2 \cdot \sum_{l=0}^{L} s_1\cdots s_l \alpha^l}, \tag{3.208}$$

with $\tilde{\mathbf{V}}_k$ and s_k constructed as in Lemma 9.

Since $\lim_{n_{\mathcal{D}}\to\infty} s_k = 1$ almost surely for all k, we obtain

$$\lim_{n_{\mathcal{D}}\to\infty} \rho_L = \mathrm{snr}_{\mathcal{D}} \text{ almost surely.} \tag{3.209}$$

Moreover, according to Lemma 10, the asymptotic EED of $\mathbf{A}$ coincides with the asymptotic EED of $\frac{1}{n_{\mathcal{S}}}\mathbf{H}\mathbf{H}^H$, where $\mathbf{H}$ contains the first $n_{\mathcal{D}}$ rows of $\mathbf{H}_{L+1}$. Thus, we can write

$$\lim_{n_{\mathcal{D}}\to\infty} \frac{1}{n_{\mathcal{D}}} \sum_{i=1}^{n_{\mathcal{D}}} \log\left(1 + \mathrm{snr}_{\mathcal{D}} \lambda_i\{\mathbf{A}\}\right) \tag{3.210}$$

$$= \lim_{n_{\mathcal{D}}\to\infty} \int_0^{\left(1+\sqrt{\beta_{\mathcal{S}}^{-1}}\right)^2} \log(1 + \mathrm{snr}_{\mathcal{D}} \cdot x) dF_{\mathbf{A}}^{(\beta_{\mathcal{S}}, n_{\mathcal{D}})}(x)$$

$$+ \lim_{n_{\mathcal{D}}\to\infty} \frac{1}{n_{\mathcal{D}}} \sum_{i:\lambda_i\{\mathbf{A}\} > \left(1+\sqrt{\beta_{\mathcal{S}}^{-1}}\right)^2} \log\left(1 + \mathrm{snr}_{\mathcal{D}} \cdot \lambda_i\{\mathbf{A}\}\right) \tag{3.211}$$

$$= \int_0^{\left(1+\sqrt{\beta_{\mathcal{S}}^{-1}}\right)^2} \log(1 + \mathrm{snr}_{\mathcal{D}} \cdot x) dF_{\mathrm{MP}}(x) \text{ almost surely,} \tag{3.212}$$

which is the sum-capacity of a single-hop multiple access channel. Here, we have taken the limit inside the definite integral according to the bounded convergence theorem. The second term in (3.211) evaluates to zero due to the concavity of the log-function and Jensen's inequality: we choose n as the maxmimal i, such that $\lambda_i\{\mathbf{A}\} > (1 + \sqrt{\beta_{\mathcal{S}}^{-1}})^2$ and write

$$\frac{n}{n_{\mathcal{D}}} \frac{1}{n} \sum_{i=1}^{n} \log\left(1 + \mathrm{snr}_{\mathcal{D}} \lambda_i\{\mathbf{A}\}\right) \leq \frac{n}{n_{\mathcal{D}}} \log\left(1 + \mathrm{snr}_{\mathcal{D}} \frac{1}{n} \sum_{i=1}^{n} \lambda_i\{\mathbf{A}\}\right) \tag{3.213}$$

$$\leq \frac{n}{n_{\mathcal{D}}} \log\left(1 + \mathrm{snr}_{\mathcal{D}} \frac{1}{n} \sum_{i=1}^{n_{\mathcal{D}}} \lambda_i\{\mathbf{A}\}\right) = \frac{n}{n_{\mathcal{D}}} \log\left(1 + \mathrm{snr}_{\mathcal{D}} \frac{n_{\mathcal{D}}}{n} \frac{1}{n_{\mathcal{D}}} \mathrm{Tr}[\mathbf{A}]\right). \tag{3.214}$$

From Lemma 8, we know that $\lim_{n_{\mathcal{D}}\to\infty} n_{\mathcal{D}}^{-1} \mathrm{Tr}\left[\mathbf{A}\right] = 1$ almost surely. Since $n/n_{\mathcal{D}} \to 0$ as $n_{\mathcal{D}} \to \infty$ almost surely — note that $\partial F_{\mathrm{MP}}^{(\beta_{\mathcal{S}})}(x)/\partial x$ is not supported for $x > (1+\sqrt{\beta_{\mathcal{S}}^{-1}})^2$ — the sum converges to zero almost surely. We have thus shown that an achievable sum-rate cannot exceed the sum-capacity of a single-hop multiple access channel.

Case $\beta_{\mathcal{R}} \in \Theta\left(L\right)$: We start out with proving that $\liminf_{L\to\infty} c_L^{\mathrm{AF}} > 0$, or equivalently, that there is a $\delta > 0$ and an L_0, such that for all $L > L_0$

$$c_L^{\mathrm{AF}} > \delta \text{ for every pair } (\beta_{\mathcal{R}}, L). \tag{3.215}$$

Due to the positive semi-definiteness of $\tilde{\mathbf{R}}_s$ and $\tilde{\mathbf{R}}_n^{-1}$ the following lower-bound on the eigenvalues of the product of both matrices holds [82, Theorem 5]:

$$\lambda_k\{\tilde{\mathbf{R}}_s\tilde{\mathbf{R}}_n^{-1}\} \geq \lambda_i\{\tilde{\mathbf{R}}_s\}\lambda_{k+n_{\mathcal{D}}-i}\{\tilde{\mathbf{R}}_n^{-1}\} \text{ for all } i \in \{k, \dots, n_{\mathcal{D}}\}. \tag{3.216}$$

We can fix $i = \lceil \alpha_1 n_{\mathcal{D}} \rceil$, where $\alpha_1 > 0$ is sufficiently small, such that for a sufficiently small $a_1 > 0$

$$\liminf_{L\to\infty} 1 - F_{\tilde{\mathbf{R}}_s}^{(\beta_{\mathcal{S}},\beta_{\mathcal{R}},L)}(a_1) > \alpha_1. \tag{3.217}$$

This choice ensures that there is an L_0, such that there is for all $L > L_0$ almost surely an $n_{\mathcal{D}}^{(0)}(L)$, such that for all $n_{\mathcal{D}} > n_{\mathcal{D}}^{(0)}(L)$, we have $\lambda_{\lceil \alpha_1 n_{\mathcal{D}} \rceil}\{\tilde{\mathbf{R}}_s\} > a_1$. The existence of such a pair (a_1, α_1) is guaranteed through Proposition 6. Next, we fix $k = \lceil \alpha_2 n_{\mathcal{D}} \rceil$, such that $0 < \alpha_2 < \alpha_1$ and $a_2 = 1 - \alpha_1 + \alpha_2$. We can write according to Proposition 7 that

$$1 - F_{\tilde{\mathbf{R}}_n^{-1}}^{(\beta_{\mathcal{R}},L)}(a_2) > a_2 = 1 - \alpha_1 + \alpha_2. \tag{3.218}$$

Thus, there is an L_0, such that there is for all $L > L_0$ almost surely an $n_{\mathcal{D}}^{(0)}(L)$, such that for all $n_{\mathcal{D}} > n_{\mathcal{D}}^{(0)}(L)$, we have $\lambda_{\lceil \alpha_2 n_{\mathcal{D}} \rceil + n_{\mathcal{D}} - \lceil \alpha_1 n_{\mathcal{D}} \rceil}\{\tilde{\mathbf{R}}_n^{-1}\} > a_2$. Substitution into (3.216) yields that there is an L_0, such that there is for all $L > L_0$ almost surely an $n_{\mathcal{D}}^{(0)}(L)$, such that for all $n_{\mathcal{D}} > n_{\mathcal{D}}^{(0)}(L)$

$$\lambda_{\lceil \alpha_2 n \rceil}\{\tilde{\mathbf{R}}_s\tilde{\mathbf{R}}_n^{-1}\} \geq \lambda_{\lceil \alpha_1 n_{\mathcal{D}} \rceil}\{\tilde{\mathbf{R}}_s\}\lambda_{\lceil \alpha_2 n_{\mathcal{D}} \rceil + n_{\mathcal{D}} - \lceil \alpha_1 n_{\mathcal{D}} \rceil}\{\tilde{\mathbf{R}}_n^{-1}\} \geq a_1 \cdot a_2 > 0. \tag{3.219}$$

Note that all $\lambda_j\{\tilde{\mathbf{R}}_s\tilde{\mathbf{R}}_n^{-1}\}$, $j < \lceil \alpha_2 n_{\mathcal{D}} \rceil$, are larger than $\lambda_j\{\tilde{\mathbf{R}}_s\tilde{\mathbf{R}}_n^{-1}\}$. Since these eigenvalues constitute a nonzero fraction ($\alpha_2 > 0$) of the $n_{\mathcal{D}}$ eigenvalues, we conclude

$$\liminf_{L\to\infty} c_L^{\mathrm{AF}} \geq \alpha_2 \log(1 + \mathrm{snr}_{\mathcal{D}} \cdot a_1 \cdot a_2) > 0. \tag{3.220}$$

The upper-bound on $\limsup_{L\to\infty} c_L^{\mathrm{AF}}$ follows immediately from the upper-bound for the case $\beta_{\mathcal{R}} \in \Omega(L^{1+\varepsilon})$, where we did not invoke any assumptions on the scaling of $\beta_{\mathcal{R}}$ with L. □

3.5. Concluding Remarks

In this work, it was shown that the choice of the relaying scheme is crucial for achievable capacity scaling in long multihop networks. While there is unsurprisingly an inherent performance gap between decode & forward and the investigated non-regenerative schemes due to noise accumulation, the key contribution of this work is the identification of the fundamental differences among the performances of non-regenerative relaying schemes in the regime of long multihop networks.

4. Conclusions

This last chapter provides some concluding remarks, which shall help to classify both the theoretical results and highlight practical implications of this thesis. Moreover, it states open problems both on the theoretical and on the practical side.

4.1. Results

Multihop communication is a means without alternative, when it comes to establishing wireless connectivity between nodes whose radio link is shadowed by obstacles. Multihopping in wireless networks has received a lot of attention in the research community in the context of routing in the past (see e.g. [83, 84] for an overview). These studies implicitly assume each relay node to decode its receive signals before the forwarding to the next nodes in the multihop chain. From a physical layer perspective, this approach is suboptimal in general. Performance is particularly diminished, if signals of multiple users interfere with each other in a way, such that signals cannot be spatially resolved through multiple antennas at a relay node. Such a relay node can then become the bottleneck for the multihop transmission. This thesis has shown on the basis of two exemplary scenarios, that cooperative distributed forwarding is an efficient means for avoiding such bottleneck effects. In this spirit, the thesis contributes further examples for the fundamental importance of physical layer cooperation in wireless networks.

Interestingly, the flavors of the results on distributed forwarding in the two example networks are quite different:

- In the multihop interference network, distributed forwarding *enables* a gain that is not achievable in a single-hop interference network: the multiplexing gain in a single-hop interference network is limited to $n/2$, while it can be n (ignoring the pre-log) in multihop interference network.
- In the multihop multiple access network, distributed forwarding is a tool that *sustains* gains that are well known from single-hop multiple access networks: linear

capacity scaling in the number of transmit/receive antennas is well known to be achievable in single-hop multiple access networks. Proper distributed forwarding guarantees this property in arbitrarily long multihop networks.

In summary, this thesis provides the insight that distributed forwarding in multiuser networks is a challenging and nontrivial problem for system designers. Significant MIMO gains are achievable. However, they are crucially dependent on the choice of the applied forwarding scheme. It has been shown, that slightly different network configurations can require fundamentally different approaches.

4.2. Open Aspects

The presented results of this thesis can be seen as two further building blocks towards an understanding of the fundamental limits of wireless networks. However, it is evident that very specific channel models and topologies are assumed. In the remainder of this section we comment on the limitations of the scope of the analysis, and the key generalizations that are desirable in order to cover the full class of generic multihop wireless networks.

Channel Model

Both in the DMT considerations on the multihop interference network and the capacity scaling considerations on the multihop multiple access network, a central assumption is an equal average path loss among all transmit and receive nodes of two adjacent stages. This assumption is very popular in the research community and usually made for analytical tractability. It has all rights to exist, since it paves the way towards important insights on the functioning and properties of a scheme. However, the assumption of an equal average path loss becomes questionable as the "height" of a relay stage (distance between $\mathrm{R}_1^{(l)}$ and $\mathrm{R}_{n_{\mathcal{R}}^{(l)}}^{(l)}$) is in the range of the distances to the adjacent stages. It would thus be desirable to generalize the results of this thesis to more general channel models that also cover heterogeneous average path loss distributions. Our expectation is that the fundamental effects and properties of the studied schemes carry over to many more general channel models.

It is important to note that different average path losses in different hops are not a severe issue in the context of our work, as long as the average path losses within all hops are equal. All results can be generalized in a straightforward fashion to hold true in this generalized scenario.

Network Topology

The considered network topologies are rather specific. In particular, the assumption that source nodes (in the multihop interference network also the destination nodes) are in proximate vicinity of each other is quite restricting. A natural next step in research of wireless multihop networks is thus the transition to more general network structures. In this context, key components are spectrally efficient cooperative physical layer signal processing schemes for the following generalized multihop networks:

- *Unclustered relay stages:* Relay nodes are not clustered into stages. In particular, the sets of relay nodes that assist the transmission of a certain source node in a given time slot need not be identical for the transmissions of different source nodes. See Fig. 4.1 for an illustration.
- *Variable number of hops:* Sources communicate over different numbers of hops to their destination(s). In this setting, nodes potentially can have two functions: they could act both as a source node and as a relay node for other nodes whose transmission "passes by". See Fig. 4.2 for an illustration.
- *Bidirectional traffic:* The terminal nodes are allowed to mutually exchange information. That is, nodes act as sources and destinations simultaneously. Spectrally efficient schemes might allow for simultaneous transmission of forward- and backward-messages over the network[1]. See Fig. 4.3 for an illustration.

All signal processing principles that are used in this thesis can still be applied in these generalized network structures with more or less modifications. We believe that also many of the effects and properties of the schemes that are discovered in this thesis carry over to them, and can give important insights into the design of respective physical layer cooperation schemes.

4.3. Practical Challenges

From a practical perspective, there are still many open problems to solve until physical layer cooperation schemes can be realized to their full effect. Generally, these issues are less severe for networks with infrastructure relay nodes than for networks with mobile relay nodes. In the following we dwell on those challenges that strike us as being the most relevant ones in the context of distributed forwarding.

[1]For two-hop interference networks this scenario has been studied [85]. However, generalizations of the multiplexing gain optimal scheme of this work to more hops are nontrivial.

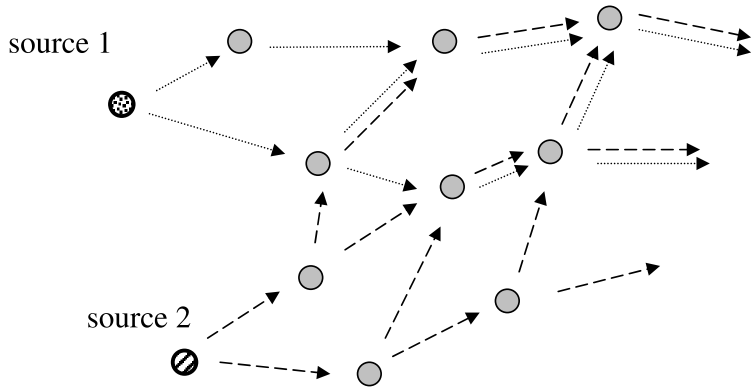


Figure 4.1.: An example of a multiuser multihop network without clustered relay stages. Source nodes use different sets of relay nodes for their transmission. Relay nodes may also serve as helper for different source nodes in different time slots.

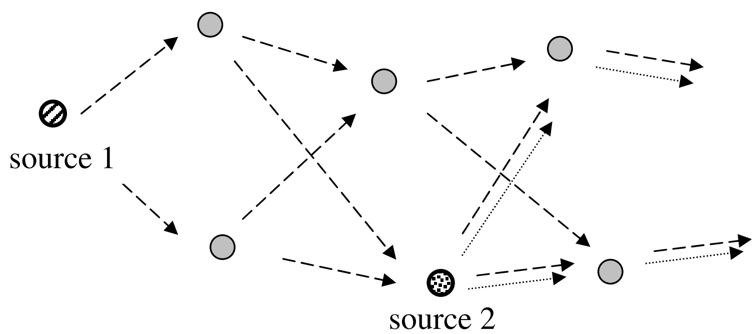


Figure 4.2.: An example of a multiuser multihop network with different number of hops for different transmissions. Source 2 serves both as a source and as a relay node in this example.

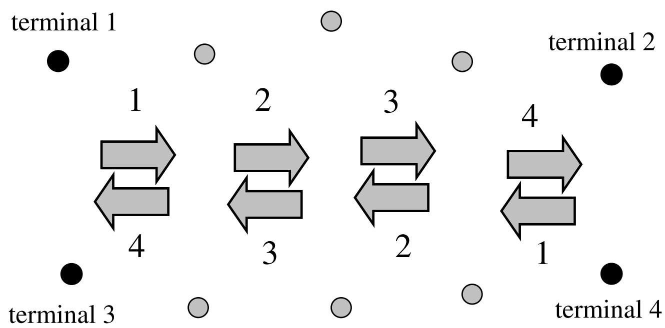


Figure 4.3.: An example of a multihop network with bidirectional traffic. In time slot 1 all terminal nodes initiate a transmission. Signals arrive at the respective terminal nodes at the other end of the network in time slot 4. In time slots 2 and 3, forward and backward transmission interfere with each other.

Coordination

A main challenge is the enormously increased need for coordination among wireless nodes in multihop networks that apply distributed forwarding. Conventional forwarding requires no more than a single path in the network graph that connects a source and a destination node. The identification of such a path might be complex enough in a wireless mobile ad-hoc network, but is still significantly less demanding than the requirements for distributed forwarding. Here, a set of sources, a set of destinations and sets of relay stages as well as a path through the relay stages need to be identified. Moreover, transmissions of nodes within the same stage need to be synchronized at least with respect to the symbol timing. Without any topology knowledge, this is a humongous challenge in a practical system.

Also medium access control (MAC) in the data link layer becomes significantly more complex. Conventionally, a relay is only interested in the transmission of a single user at a given time. It will decode and retransmit the signals of this user and treat all other signals as interference. Thus, classical MAC approaches apply and an efficient decentralized handling of the issue through standard schemes such as carrier sense multiple access (CSMA) is possible. The necessary coordination for distributed forwarding is significantly more demanding. In the multiuser scenarios of this thesis, it is obligatory to have a sufficient number of relay nodes to simultaneously support a set of multiple transmissions. This requires that all relay nodes within a network stage obtain medium access simultaneously. Although the design of cooperative MAC schemes is a hot research topic in the meanwhile (e.g. [86–91]), there are to our knowledge currently no MAC schemes proposed that would allow for an efficient operation of a distributed forwarding multihop chain in a mobile wireless ad-hoc network in the spirit of this thesis.

Channel State Information

Some of the schemes that are considered in this thesis require relay nodes to have CSI of channels between other nodes in the network. This CSI is not accessible for them through training (estimation based on received pilot sequences). Thus, the respective channel coefficients need to be locally estimated and then disseminated over the network. This constitutes an overhead that might well be manageable in small and static networks, but might become prohibitive in large and mobile networks. This overhead can either be handled in-band or, more conveniently, through a secondary network that connects the respective relay nodes in the network. Such a secondary network can either be a wired backbone network (for fixed infrastructure relays) or a wireless network (e.g. WLAN). The CSI overhead problem is particularly severe in the coherent amplify & forward approach of Chapter 2, which requires global CSI for the computation of the relay gain coefficients. Recent research has aimed for reducing this CSI overhead in such multihop interference networks [46, 48, 92]. Progress has been made to an extent that the required CSI dissemination overhead has been rendered independent of the number of relay nodes per stage. Still the overhead remains significant.

Phase Noise and Synchronization

In a multihop network that operates a distributed forwarding scheme, the oscillators of the multiple relay nodes that simultaneously forward signals are subject to statistically independent phase noise. Even if the propagation channel is perfectly static, this imperfection renders the communication channels time-variant in a way such that elements of the effective MIMO channel matrices are affected in different ways. This is not an issue in a point-to-point MIMO channel, since antennas are connected to a single oscillator on transmit and receive side each. Thus, all elements in the MIMO channel matrix undergo the same phase shift, which can be compensated at the receiver and thus has no impact on performance. The stability requirements on oscillators in cooperative distributed forwarding are thus particularly stringent.

A fundamental question is, whether or not the carrier phase of relay nodes within a stage need to be synchronized for a scheme to function properly in a multihop network. For the proposed schemes for the multihop multiple access network, this issue is uncritical, and synchronization not required. For the coherent amplify & forward scheme, however, this question is nontrivial to answer, since relay nodes need to compute and apply specific phase shifts in order to mitigate spatial interference at the destination nodes. The problem has been studied in [93] for two-hop interference networks. The findings therein carry over to arbitrary multihop networks in a one-to-one fashion. The key result is the following: Phase synchronization is not required, if and only if all channel coefficients are measured in the same direction. Typically, this would be the forward direction, since this approach does not rely on channel reciprocity. This means, however, the training sequence of a relay stage cannot be used both at the preceding and the succeeding stage for channel estimation, but only at one of them (typically the succeeding one). There are further situations that favor channel estimation in opposite directions. Another example is a relay stage with lots of nodes that is surrounded by two stages with only few nodes. Then, it is spectrally more efficient, to have the surrounding stages transmit the pilot sequences and have the middle stage estimate the respective channel coefficients. Carrier phase synchronization schemes are provided in [94, 95] for this purpose.

A. Appendix

A.1. Proof of Proposition 4

For any $l \in \{1, \dots, L\}$, the SINR of source-destination pair i can be written as

$$\mathsf{SINR}_i = \frac{\mathbf{g}_l^H \mathbf{M}_\mathrm{s}^{(l,i)} \mathbf{g}_l}{\mathbf{g}_l^H \mathbf{M}_\mathrm{i}^{(l,i)} \mathbf{g}_l + \mathbf{g}_l^H \mathbf{M}_{\mathrm{n},1}^{(l,i)} \mathbf{g}_l + m_{\mathrm{n},2}^{(l,i)}}, \tag{A.1}$$

where

- $\mathbf{g}_l^H \mathbf{M}_\mathrm{s}^{(l,i)} \mathbf{g}_l$ corresponds to the power of the desired part of the received signal, and

$$\mathbf{M}_\mathrm{s}^{(l,i)} = \frac{P_\mathcal{S}}{n} \cdot \mathrm{diag}\left(\mathbf{h}_{\mathrm{D}_i \mathcal{R}_l}^*\right) \mathbf{h}_{\mathcal{R}_l \mathrm{S}_i}^* \mathbf{h}_{\mathcal{R}_l \mathrm{S}_i}^T \mathrm{diag}\left(\mathbf{h}_{\mathrm{D}_i \mathcal{R}_l}\right), \tag{A.2}$$

  with

$$\mathbf{h}_{\mathcal{R}_l \mathrm{S}_i} = \mathbf{H}_l \mathbf{G}_{l-1} \cdots \mathbf{H}_2 \mathbf{G}_1 \mathbf{h}_{\mathcal{R}_1 \mathrm{S}_i}, \tag{A.3}$$

$$\mathbf{h}_{\mathrm{D}_i \mathcal{R}_l}^T = \mathbf{h}_{\mathrm{D}_i \mathcal{R}_L}^T \mathbf{G}_L \mathbf{H}_L \cdots \mathbf{G}_{l+1} \mathbf{H}_{l+1}, \tag{A.4}$$

  and $\mathbf{h}_{\mathcal{R}_1 \mathrm{S}_i}$ the ith column of $\mathbf{H}_1$, $\mathbf{h}_{\mathrm{D}_i \mathcal{R}_L}^T$ the ith row of $\mathbf{H}_{L+1}$,

- $\mathbf{g}_l^H \mathbf{M}_\mathrm{i}^{(l,i)} \mathbf{g}_l$ corresponds to the power of the interference part of the received signal, and

$$\begin{aligned} \mathbf{M}_\mathrm{i}^{(l,i)} &= \frac{P_\mathcal{S}}{n} \cdot \sum_{j \neq i} \mathrm{diag}\left(\mathbf{h}_{\mathrm{D}_i \mathcal{R}_l}^*\right) \mathbf{h}_{\mathcal{R}_l \mathrm{S}_j}^* \mathbf{h}_{\mathcal{R}_l \mathrm{S}_j}^T \mathrm{diag}\left(\mathbf{h}_{\mathrm{D}_i \mathcal{R}_l}\right) \\ &= \frac{P_\mathcal{S}}{n} \cdot \mathrm{diag}\left(\mathbf{h}_{\mathrm{D}_i \mathcal{R}_l}^*\right) \mathbf{H}_{\mathcal{R}_l \mathcal{S} \setminus \{\mathrm{S}_i\}}^* \mathbf{H}_{\mathcal{R}_l \mathcal{S} \setminus \{\mathrm{S}_i\}}^T \mathrm{diag}\left(\mathbf{h}_{\mathrm{D}_i \mathcal{R}_l}\right), \end{aligned} \tag{A.5}$$

  with

$$\mathbf{H}_{\mathcal{R}_l \mathcal{S} \setminus \{\mathrm{S}_i\}} = \left(\mathbf{h}_{\mathcal{R}_l \mathrm{S}_1}, \dots, \mathbf{h}_{\mathcal{R}_l \mathrm{S}_{i-1}}, \mathbf{h}_{\mathcal{R}_l \mathrm{S}_{i+1}}, \dots, \mathbf{h}_{\mathcal{R}_l \mathrm{S}_n}\right), \tag{A.6}$$

- $\mathbf{g}_l^H \mathbf{M}_{\text{n},1}^{(l,i)} \mathbf{g}_l$ corresponds to the power of the noise in the received signal that has been introduced in relay stages $\mathcal{R}_k$, $k \in \{1, \ldots, l\}$, and

$$\mathbf{M}_{\text{n},1}^{(l,i)} = \sigma^2 \cdot \operatorname{diag}\left(\mathbf{h}_{\text{D}_i\mathcal{R}_l}^*\right) \left(\mathbf{I}_{n_\mathcal{R}^{(l)}} + \sum_{m=1}^{l-1} \mathbf{H}_{\mathcal{R}_l\mathcal{R}_m}^* \mathbf{G}_m^* \mathbf{G}_m^T \mathbf{H}_{\mathcal{R}_l\mathcal{R}_m}^T\right) \operatorname{diag}\left(\mathbf{h}_{\text{D}_i\mathcal{R}_l}\right), \tag{A.7}$$

with

$$\mathbf{H}_{\mathcal{R}_l\mathcal{R}_m} = \mathbf{H}_l \mathbf{G}_{l-1} \cdots \mathbf{G}_{m+1} \mathbf{H}_{m+1}, \tag{A.8}$$

- the power of the noise in the received signal that has been introduced in relay stages $\mathcal{R}_k$, $k \in \{l+1, \ldots, L\}$, and at the destination D_i is given by

$$m_{\text{n},2}^{(l,i)} = \sigma^2 \cdot \left(1 + \sum_{m=l+1}^{L} \mathbf{g}_m^H \operatorname{diag}\left(\mathbf{h}_{\text{D}_i\mathcal{R}_m}^*\right) \operatorname{diag}\left(\mathbf{h}_{\text{D}_i\mathcal{R}_m}\right) \mathbf{g}_m\right). \tag{A.9}$$

Moreover, for every $l' \geq l$ the transmit power of relay stage $\mathcal{R}_{l'}$ (for $l' < l$ the transmit power of stage $\mathcal{R}_{l'}$ is independent of $\mathbf{g}_l$) can be written as

$$P_{n_\mathcal{R}^{(l')}} = \mathbf{g}_l^H \mathbf{C}^{(l',l)} \mathbf{g}_l + c^{(l',l)}, \tag{A.10}$$

where

$$\mathbf{C}^{(l',l)} = \begin{cases} \sum_{k=1}^{n_\mathcal{R}^{(l')}} \left|g_{\text{R}_k^{(l')}}\right|^2 \Bigg(\frac{P_\mathcal{S}}{n} \operatorname{diag}\left(\mathbf{h}_{\text{R}_k^{(l')}\mathcal{R}_l}^*\right) \mathbf{H}_{\mathcal{R}_l\mathcal{S}}^* \mathbf{H}_{\mathcal{R}_l\mathcal{S}}^T \operatorname{diag}\left(\mathbf{h}_{\text{R}_k^{(l')}\mathcal{R}_l}\right) \\ \quad + \sigma^2 \operatorname{diag}\left(\mathbf{h}_{\text{R}_k^{(l')}\mathcal{R}_l}^*\right) \operatorname{diag}\left(\mathbf{h}_{\text{R}_k^{(l')}\mathcal{R}_l}\right) \\ \quad + \sigma^2 \sum_{m=1}^{l-1} \operatorname{diag}\left(\mathbf{h}_{\text{R}_k^{(l')}\mathcal{R}_l}^*\right) \mathbf{H}_{\mathcal{R}_l\mathcal{R}_m}^* \mathbf{G}_m^* \mathbf{G}_m^T \mathbf{H}_{\mathcal{R}_l\mathcal{R}_m}^T \operatorname{diag}\left(\mathbf{h}_{\text{R}_k^{(l')}\mathcal{R}_l}\right)\Bigg), & \text{if } l' > l, \\ \left(\frac{P_\mathcal{S}}{n} \mathbf{H}_{\mathcal{R}_l\mathcal{S}}^* \mathbf{H}_{\mathcal{R}_l\mathcal{S}}^T + \sigma^2 \left(\mathbf{I}_{n_\mathcal{R}^{(l)}} + \sum_{m=1}^{l-1} \mathbf{H}_{\mathcal{R}_l\mathcal{R}_m}^* \mathbf{G}_m^* \mathbf{G}_m^T \mathbf{H}_{\mathcal{R}_l\mathcal{R}_m}^T\right)\right) \odot \mathbf{I}_{n_\mathcal{R}^{(l)}}, & \text{if } l' = l, \end{cases} \tag{A.11}$$

$$c^{(l',l)} = \begin{cases} \sigma^2 \cdot \mathrm{Tr}\left(\mathbf{G}_{l'}\left(\mathbf{I}_{n_{\mathcal{R}}^{(l')}} + \sum_{m=l+1}^{l'-1} \mathbf{H}_{\mathcal{R}_{l'}\mathcal{R}_m}\mathbf{G}_m\mathbf{G}_m^H\mathbf{H}_{\mathcal{R}_{l'}\mathcal{R}_m}^H\right)\mathbf{G}_{l'}^H\right), \text{ if } l' > l, \\ 0, \text{ if } l' = l, \end{cases} \tag{A.12}$$

with

$$\mathbf{h}^T_{\mathrm{R}_k^{(l')}\mathcal{R}_l} = \mathbf{h}^T_{\mathrm{R}_k^{(l')}\mathcal{R}_{l'-1}}\mathbf{G}_{l'-1}\cdots\mathbf{G}_{l+1}\mathbf{H}_{l+1}, \tag{A.13}$$

$$\mathbf{H}_{\mathcal{R}_l\mathcal{S}} = \mathbf{H}_l\mathbf{G}_{l-1}\cdots\mathbf{G}_1\mathbf{H}_1, \tag{A.14}$$

and $\mathbf{h}^T_{\mathrm{R}_k^{(l')}\mathcal{R}_{l'-1}}$ the kth row of $\mathbf{H}_{l'}$.

In the following, we assume an arbitrary $l \in \{1, \ldots, L\}$, and substitute (2.81) into (A.1) for all $l' \in \{1, \ldots, L\}$. That is, the gain vectors of all relay stages are chosen to fulfill the respective power constraint with equality each. The optimality of this choice will be proved later on. After these substitutions, the SINR of source-destination pair i is turned into a generalized Rayleigh quotient in $\tilde{\mathbf{g}}_l$.

Substitutions are performed iteratively. Iteration p is allocated to the substitution of the gain vector of stage $\mathcal{R}_{L-p+1}$, $\mathbf{g}_{L-p+1}$. Before each substitution, the SINR of source-destination pair i is written in terms of the parameters $\mathbf{A}^{(l,i,p)}$, $\mathbf{B}_m^{(l,i,p)}$, $b_m^{(l,i,p)}$ and $r^{(p)}$, where the additive interference and noise terms are arranged according to the gain vectors they depend on (indicated in brackets below):

$$\mathsf{SINR}_i = \frac{\mathbf{g}_l^H\mathbf{A}^{(l,i,p)}(\{\mathbf{g}_1, \ldots, \mathbf{g}_{L-p+1}\}\backslash\{\mathbf{g}_l\})\mathbf{g}_l}{\sum_{m=1}^{l}\mathbf{g}_l^H\mathbf{B}_m^{(l,i,p)}(\{\mathbf{g}_m, \ldots, \mathbf{g}_{L-p+1}\}\backslash\{\mathbf{g}_l\})\mathbf{g}_l + \sum_{m=l+1}^{L-p+1} b_m^{(l,i,p)}(\{\mathbf{g}_m, \ldots, \mathbf{g}_{L-p+1}\}) + r^{(p)}}. \tag{A.15}$$

In the first iteration $p = 1$, the parameters are immediately identified by inspection of (A.1) and given by

$$\mathbf{A}^{(l,i,p)} = \mathbf{M}_{\mathrm{s}}^{(l,i)}, \tag{A.16}$$

$$\mathbf{B}_m^{(l,i,p)} = \begin{cases} \mathbf{M}_{\mathrm{i}}^{(l,i)} \cdot \delta[m-1] + \sigma^2 \cdot \mathrm{diag}\left(\mathbf{h}^*_{\mathrm{D}_i\mathcal{R}_l}\right)\mathbf{H}^*_{\mathcal{R}_l\mathcal{R}_m}\mathbf{G}_m^*\mathbf{G}_m\mathbf{H}^T_{\mathcal{R}_l\mathcal{R}_m}\mathrm{diag}\left(\mathbf{h}_{\mathrm{D}_i\mathcal{R}_l}\right), \\ \qquad \text{if } m \in \{1, \ldots, l-1\}, \\ \mathbf{M}_{\mathrm{i}}^{(l,i)} \cdot \delta[m-1] + \sigma^2 \cdot \mathrm{diag}\left(\mathbf{h}^*_{\mathrm{D}_i\mathcal{R}_l}\right)\mathrm{diag}\left(\mathbf{h}_{\mathrm{D}_i\mathcal{R}_l}\right), \text{ if } m = l, \end{cases} \tag{A.17}$$

$$b_m^{(l,i,p)} = \sigma^2 \cdot \mathbf{g}_m^H \mathrm{diag}\left(\mathbf{h}_{\mathrm{D}_i\mathcal{R}_m}^*\right) \mathrm{diag}\left(\mathbf{h}_{\mathrm{D}_i\mathcal{R}_m}\right) \mathbf{g}_m, \text{ for all } m \in \{l+1, \ldots, L\}, \tag{A.18}$$

$$r^{(p)} = \sigma^2. \tag{A.19}$$

The transmit power of stage $\mathcal{R}_{L-p+1}$, $\tilde{P}_{\mathcal{R}_{L-p+1}}$, in (2.81) that would result, if $\tilde{\mathbf{g}}_{L-p+1}$ was used as gain vector of stage $\mathcal{R}_{L-p+1}$, can be written as (A.10)

$$\tilde{P}_{\mathcal{R}_{L-p+1}} = \begin{cases} \sum_{m=1}^{l} \mathbf{g}_l^H \mathbf{C}_m^{(L-p+1,l)}(\{\mathbf{g}_m, \ldots, \mathbf{g}_{L-p+1}\} \backslash \{\mathbf{g}_l\})\big|_{\mathbf{g}_{L-p+1}=\tilde{\mathbf{g}}_{L-p+1}} \mathbf{g}_l \\ + \sum_{m=l+1}^{L-p+1} c_m^{(L-p+1,l)}(\{\mathbf{g}_m, \ldots, \mathbf{g}_{L-p+1}\})\big|_{\mathbf{g}_{L-p+1}=\tilde{\mathbf{g}}_{L-p+1}}, \text{ if } L-p+1 > l, \\ \sum_{m=1}^{l} \tilde{\mathbf{g}}_l^H \mathbf{C}_m^{(L-p+1,l)}(\{\mathbf{g}_m, \ldots, \mathbf{g}_{l-1}\}) \tilde{\mathbf{g}}_l, \text{ if } L-p+1 = l, \end{cases} \tag{A.20}$$

where $\mathbf{C}_m^{(L-p+1,l)}(\{\mathbf{g}_m, \ldots, \mathbf{g}_{L-p+1}\} \backslash \{\mathbf{g}_l\})$ and $c_m^{(L-p+1,l)}(\{\mathbf{g}_m, \ldots, \mathbf{g}_{L-p+1}\})$ correspond to the terms of (A.11) and (A.12) that depend on all these vectors each.

With the substitution $\mathbf{g}_{L-p+1} = \sqrt{\bar{P}_{\mathcal{R}_{L-p+1}}/\tilde{P}_{\mathcal{R}_{L-p+1}}} \cdot \tilde{\mathbf{g}}_{L-p+1}$, the SINR of source-destination pair i (after multiplication of numerator and denominator by $\tilde{P}_{\mathcal{R}_{L-p+1}}/\bar{P}_{\mathcal{R}_{L-p+1}}$) is given by

$$\mathsf{SINR}_i = \begin{cases} \left.\dfrac{\mathbf{g}_l^H \mathbf{A}^{(l,i,p)}(\{\mathbf{g}_1, \ldots, \mathbf{g}_{L-p+1}\} \backslash \{\mathbf{g}_l\}) \mathbf{g}_l}{\mathbf{g}_l^H \left(\sum_{m=1}^{l} \mathbf{B}_m^{(l,i,p)}(\{\mathbf{g}_m, \ldots, \mathbf{g}_{L-p+1}\} \backslash \{\mathbf{g}_l\})\right) \mathbf{g}_l + \sum_{m=l+1}^{L-p+1} b_m^{(l,i,p)}(\{\mathbf{g}_m, \ldots, \mathbf{g}_{L-p+1}\}) + q^{(p)}}\right|_{\mathbf{g}_{L-p+1}=\tilde{\mathbf{g}}_{L-p+1}}, \\ \qquad \text{, if } L-p+1 > l, \\ \dfrac{\tilde{\mathbf{g}}_l^H \mathbf{A}^{(l,i,p)}(\{\mathbf{g}_1, \ldots, \mathbf{g}_{l-1}\}) \tilde{\mathbf{g}}_l}{\tilde{\mathbf{g}}_l^H \left(\sum_{m=1}^{l} \mathbf{B}_m^{(l,i,p)}(\{\mathbf{g}_m, \ldots, \mathbf{g}_{l-1}\})\right) \tilde{\mathbf{g}}_l + q^{(p)}}, \text{ if } L-p+1 = l, \end{cases} \tag{A.21}$$

where

$$q^{(p)} = r^{(p)} \cdot \frac{\tilde{P}_{n_{\mathcal{R}}^{(L-p+1)}}}{\bar{P}_{\mathcal{R}_{L-p+1}}}. \tag{A.22}$$

The substitution yields this nice expression, since the numerator and of all terms of the denominator except for $r^{(p)}$ in (A.15) depend quadratically on $\mathbf{g}_{L-p+1}$. Therefore, the expansion of the fraction does not affect any of the terms except for $r^{(p)}$. The transmit power of each relay stage $\mathcal{R}_k$, $k < L-p+1$, and thus also each corresponding upcoming substitution, is independent of $\tilde{\mathbf{g}}_{L-p+1}$. For this reason, the dependence of the matrices $\mathbf{A}^{(l,i,p)}$ and $\mathbf{B}_m^{(l,i,p)}$ on this vector is dropped for notational convenience subsequently.

The terms in $q^{(p)}$ are distributed over the respective terms in the denominator of (A.21) according to the gain vectors they depend on. If $L-p+1>l$, the SINR of source-destination pair i is written analogously to (A.15) as

$$\mathsf{SINR}_i = \frac{\mathbf{g}_l^H \mathbf{A}^{(l,i,p+1)}(\{\mathbf{g}_1,\dots,\mathbf{g}_{L-p}\}\backslash\{\mathbf{g}_l\})\mathbf{g}_l}{\sum_{m=1}^{l} \mathbf{g}_l^H \mathbf{B}_m^{(l,i,p+1)}(\{\mathbf{g}_m,\dots,\mathbf{g}_{L-p}\}\backslash\{\mathbf{g}_l\})\mathbf{g}_l + \sum_{m=l+1}^{L-p} b_m^{(l,i,p+1)}(\{\mathbf{g}_m,\dots,\mathbf{g}_{L-p}\}) + r^{(p+1)}}, \tag{A.23}$$

where

$$\mathbf{A}^{(l,i,p+1)} = \mathbf{A}^{(l,i,p)}\Big|_{\mathbf{g}_{L-p+1}=\tilde{\mathbf{g}}_{L-p+1}}, \tag{A.24}$$

$$\mathbf{B}_m^{(l,i,p+1)} = \mathbf{B}_m^{(l,i,p)}\Big|_{\mathbf{g}_{L-p+1}=\tilde{\mathbf{g}}_{L-p+1}} + \frac{r^{(p)}}{\bar{P}_{\mathcal{R}_{L-p+1}}} \cdot \mathbf{C}_m^{(L-p+1,l)}\Big|_{\mathbf{g}_{L-p+1}=\tilde{\mathbf{g}}_{L-p+1}} \quad \text{for all } m \in \{1,\dots,l\}, \tag{A.25}$$

$$b_m^{(l,i,p+1)} = b_m^{(l,i,p)}\Big|_{\mathbf{g}_{L-p+1}=\tilde{\mathbf{g}}_{L-p+1}} + \frac{r^{(p)}}{\bar{P}_{\mathcal{R}_{L-p+1}}} \cdot c_m^{(L-p+1,l)}\Big|_{\mathbf{g}_{L-p+1}=\tilde{\mathbf{g}}_{L-p+1}} \quad \text{for all } m \in \{l+1,\dots,L-p\}, \tag{A.26}$$

$$r^{(p+1)} = \frac{r^{(p)}}{\bar{P}_{\mathcal{R}_{L-p+1}}} \cdot c_{L-p+1}^{(L-p+1,l)}\Big|_{\mathbf{g}_{L-p+1}=\tilde{\mathbf{g}}_{L-p+1}} + b_{L-p+1}^{(l,i,p)}\Big|_{\mathbf{g}_{L-p+1}=\tilde{\mathbf{g}}_{L-p+1}}. \tag{A.27}$$

Note that $r^{(p+1)}$ is independent of all $\mathbf{g}_{l'}$ that are not yet substituted. We are thus back to an expression of the form (A.15), and (A.23) serves as the starting point for the next ($p+1$st) iteration. If $L-p+1=l$, the term $q^{(p)}$ in (A.21) is a quadratic form in $\tilde{\mathbf{g}}_l$, such that the SINR of source-destination pair i can be written as

$$\mathsf{SINR}_i = \frac{\tilde{\mathbf{g}}_l^H \mathbf{A}^{(l,i,p)}(\{\mathbf{g}_1,\dots,\mathbf{g}_{l-1}\})\tilde{\mathbf{g}}_l}{\tilde{\mathbf{g}}_l^H \mathbf{B}^{(l,i,p)}(\{\mathbf{g}_1,\dots,\mathbf{g}_{l-1}\})\tilde{\mathbf{g}}_l}, \tag{A.28}$$

where

$$\mathbf{B}^{(l,i,p)} = \sum_{m=1}^{l}\left(\mathbf{B}_m^{(l,i,L-l+1)} + \frac{r^{(p)}}{\bar{P}_{\mathcal{R}_{L-p+1}}} \cdot \mathbf{C}_l^{(L-p+1,l)}\right). \tag{A.29}$$

The remaining substitutions for $l' < l$ are independent of $\tilde{\mathbf{g}}_l$. Thus, the SINR of source-destination pair i can finally be written as

$$\mathsf{SINR}_i = \frac{\tilde{\mathbf{g}}_l^H \mathbf{A}_{l,i} \tilde{\mathbf{g}}_l}{\tilde{\mathbf{g}}_l^H \mathbf{B}_{l,i} \tilde{\mathbf{g}}_l}, \tag{A.30}$$

where $\mathbf{A}_{l,i}$ and $\mathbf{B}_{l,i}$ are functions of $\tilde{\mathbf{g}}_1, \ldots, \tilde{\mathbf{g}}_{l-1}, \tilde{\mathbf{g}}_{l+1}, \ldots, \tilde{\mathbf{g}}_L$, but not of $\tilde{\mathbf{g}}_l$. They are given by

$$\mathbf{A}_{l,i} = \left(\left(\mathbf{A}^{(l,i,p)} \Big|_{\mathbf{g}_{l-1} = \sqrt{\frac{\bar{P}_{\mathcal{R}_{l-1}}}{\tilde{P}_{\mathcal{R}_{l-1}}}} \cdot \tilde{\mathbf{g}}_{l-1}} \right) \Bigg|_{\mathbf{g}_{l-2} = \sqrt{\frac{\bar{P}_{\mathcal{R}_{l-2}}}{\tilde{P}_{\mathcal{R}_{l-2}}}} \cdot \tilde{\mathbf{g}}_{l-2}} \cdots \right) \Bigg|_{\mathbf{g}_1 = \sqrt{\frac{\bar{P}_{\mathcal{R}_1}}{\tilde{P}_{\mathcal{R}_1}}} \cdot \tilde{\mathbf{g}}_1}, \tag{A.31}$$

$$\mathbf{B}_{l,i} = \left(\left(\mathbf{B}^{(l,i,p)} \Big|_{\mathbf{g}_{l-1} = \sqrt{\frac{\bar{P}_{\mathcal{R}_{l-1}}}{\tilde{P}_{\mathcal{R}_{l-1}}}} \cdot \tilde{\mathbf{g}}_{l-1}} \right) \Bigg|_{\mathbf{g}_{l-2} = \sqrt{\frac{\bar{P}_{\mathcal{R}_{l-2}}}{\tilde{P}_{\mathcal{R}_{l-2}}}} \cdot \tilde{\mathbf{g}}_{l-2}} \cdots \right) \Bigg|_{\mathbf{g}_1 = \sqrt{\frac{\bar{P}_{\mathcal{R}_1}}{\tilde{P}_{\mathcal{R}_1}}} \cdot \tilde{\mathbf{g}}_1}. \tag{A.32}$$

It remains to prove that it is optimal to fulfill the sum-power constraints on the relay stages with equality each. To this end, the SINR before the substitution in iteration $p = l - L + 1$ is studied. It is given by (cf. (A.23))

$$\mathsf{SINR}_i = \frac{\mathbf{g}_l^H \mathbf{A}^{(l,i,L-l+1)}(\{\mathbf{g}_1, \ldots, \mathbf{g}_{l-1}\}) \mathbf{g}_l}{\sum_{m=1}^{l} \mathbf{g}_l^H \mathbf{B}_m^{(l,i,L-l+1)}(\{\mathbf{g}_m, \ldots, \mathbf{g}_{l-1}\}) \mathbf{g}_l + r^{(L-l+1)}}, \tag{A.33}$$

where $r^{(L-l+1)}$ is independent of $\mathbf{g}_l$. Since the matrices $\mathbf{A}^{(l,i,L-l+1)}(\{\mathbf{g}_1, \ldots, \mathbf{g}_{l-1}\})$ and $\mathbf{B}_m^{(l,i,L-l+1)}(\{\mathbf{g}_m, \ldots, \mathbf{g}_{l-1}\})$ are positive semidefinite, SINR_i is for every $\mathbf{g}_l = \alpha \cdot \hat{\mathbf{g}}_l$, monotonically increasing in α. That is, it is optimal to allocate the maximum allowable transmit power to relay stage $\mathcal{R}_l$, and the substitution (2.81) is without loss of optimality. If this argument is invoked successively for $l = L, L-1, \ldots, 1$ optimality is established for all relay stages. $\square$

B. Appendix

B.1. Proofs of Lemmata for Theorems 3 and 4

Proof of Lemma 1. Consider the sum-capacity achieving minimum mean square error (MMSE) successive interference cancellation (SIC) receiver structure [67]. We modify this receiver as follows: rather than applying interference cancellation after each decoded codeword, we decode codewords of multiple transmit terminals simultaneously based on a given MMSE equalizer output signal. We group the transmit terminals into the sets $\mathcal{M}_m$, $m \in \{1, \ldots, M\}$. Codewords of transmit terminals within the same set $\mathcal{M}_m$ are decoded simultaneously based on the same MMSE equalizer output signal each, i.e., there are $M-1$ interference cancellation steps in total. Since n is not an integer multiple of M in general, we choose the cardinality of these sets as

$$|\mathcal{M}_m| = \begin{cases} \lceil n/M \rceil, & \text{if } 1 \leq m \leq n \bmod M \\ \lfloor n/M \rfloor, & \text{else.} \end{cases} \tag{B.1}$$

We introduce the one-to-one map $l : \{1, \ldots, M\} \longrightarrow \{1, \ldots, M\} : m \longrightarrow l_m$, where l_m corresponds to the position of set $\mathcal{M}_m$ in the decoding order. E.g., $l_3 = 4$ implies that set $\mathcal{M}_3$ is decoded based on the MMSE output of the fourth "decoding phase", when the transmit signals of three other sets are already canceled.

Let us denote the number of mutually interfering streams in the decoding phase of set $\mathcal{M}_m$ by n_m and consider the ratio

$$\frac{n_m}{n} = \frac{\sum_{k=l_m}^{M} \left|\mathcal{M}_{(l^{-1})_k}\right|}{n}. \tag{B.2}$$

Here, we denote by $(l^{-1})_k$ the inverse function of l_m. This ratio is upper- and lower-bounded as follows:

$$\frac{n_m}{n} \leq \frac{(M - l_m + 1) \cdot \lceil \frac{n}{M} \rceil}{n} < \frac{(M - l_m + 1) \cdot \left(\frac{n}{M} + 1\right)}{n}, \tag{B.3}$$

$$\frac{n_m}{n} \geq \frac{(M - l_m + 1) \cdot \lfloor \frac{n}{M} \rfloor}{n} > \frac{(M - l_m + 1) \cdot \left(\frac{n}{M} - 1\right)}{n}. \tag{B.4}$$

Since both bounds converge to the same value as n grows large, we conclude

$$\lim_{n \to \infty} \frac{n_m}{n} = \frac{M - l_m + 1}{M} \triangleq \beta_{l_m}. \tag{B.5}$$

In the following, we assume that transmission is divided into M time slots of N symbol durations each. Each transmit terminal transmits a sequence of M codewords of length N, one in each time slot. We specify the decoding order for the codewords of the set $\mathcal{M}_m$ that are transmitted in the jth time slot through the function l : $\{1, \ldots, M\}^2 \longrightarrow \{1, \ldots, M\} : (m, j) \longrightarrow l_{m,j}$, where

$$l_{m,j} = ((M - m + j) \bmod M) + 1. \tag{B.6}$$

This function is one-to-one in j for each fixed m and illustrated in Table B.1. The codewords of the transmit terminals in set $\mathcal{M}_m$ for the jth time slot are selected from a common code whose rate $R(l_{m,j})$ is fully determined by the value of $l_{m,j}$, i.e., the respective position in the decoding order for this time slot. The average code rate over the M codebooks of the transmit terminals in set $\mathcal{M}_m$ is given by

$$R^{(m)} = \frac{1}{M} \sum_{j=1}^{M} R(l_{m,j}). \tag{B.7}$$

Since (i) the codomain of $l_{m,j}$ with respect to j is the same for all m, and (ii) for each fixed m the function $l_{m,j}$ is one to one in j, this average rate $R^{(m)}$ is the *same* for all $\mathcal{M}_m$, i.e.,

$$R^{(1)} = R^{(2)} = \ldots = R^{(M)}. \tag{B.8}$$

In the following, we use a result [96, Lemma 3.1] on the t signal-to-interference-plus-noise ratios (SINRs), $(\mathsf{SINR}_k)_{k=1}^{t}$, at the output of an MMSE receiver in an $r \times t$ MIMO channel with channel coefficients as assumed in this lemma. If $t/r \to \beta$ as $n \to \infty$, then there exists a constant $\mathsf{SINR}^{(\infty)}(\beta)$, such that almost surely

$$\lim_{n \to \infty} \max_{k \in \{1, \ldots, t\}} \left| \mathsf{SINR}_k - \mathsf{SINR}^{(\infty)}(\beta) \right| = 0. \tag{B.9}$$

We apply this result, for each of the M time slots, to each of the M decoding phases

set	time slot 1	time slot 2	time slot 3	...	time slot $M-1$	time slot M
$\mathcal{M}_1$	1	2	3	...	$M-1$	M
$\mathcal{M}_2$	M	1	2	...	$M-2$	$M-1$
$\mathcal{M}_3$	$M-1$	M	1	...	$M-3$	$M-2$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
$\mathcal{M}_{M-1}$	3	4	5	...	1	2
$\mathcal{M}_M$	2	3	4	...	M	1

Table B.1.: Illustration of $l_{m,j}$: position of set $\mathcal{M}_m$ in the decoding order of the jth decoding phase.

in the corresponding modified MMSE-SIC receiver. Specifically, since M is finite, we conclude that almost surely

$$\lim_{n\to\infty} \max_{(j,m)\in\{1,\ldots,M\}^2} \max_{k\in\mathcal{M}_m} \left|\mathsf{SINR}_{k,j} - \mathsf{SINR}^{(\infty)}\left(\beta_{l_{m,j}}\right)\right| = 0, \tag{B.10}$$

where $\mathsf{SINR}_{k,j}$ denotes the SINR for the signal of transmit terminal k in time slot j, and $\beta_{l_{m,j}}$ is defined analogously to (B.5). Thus, in the limit of large n, $R(l_{m,j})$ is achievable for all transmit terminals in $\mathcal{M}_m$ in time slot j almost surely, if

$$R(l_{m,j}) < \log\left(1 + \mathsf{SINR}^{(\infty)}\left(\beta_{l_{m,j}}\right)\right). \tag{B.11}$$

The average (over time slots) of the suprema of achievable rates for the transmit terminals in set $\mathcal{M}_m$ is given by

$$R_m = \frac{1}{M}\sum_{j=1}^{M} \log\left(1 + \min_{k\in\mathcal{M}_m} \mathsf{SINR}_{k,j}\right). \tag{B.12}$$

According to (B.10), there exists for every $\varepsilon > 0$ almost surely an n_0, such that for all $n \geq n_0$

$$\max_{m\in\{1,\ldots,M\}} \left|R_m - \tilde{\xi}\right| < \frac{\varepsilon}{2}, \tag{B.13}$$

where due to the choice of $l_{m,j}$

$$\tilde{\xi} = \frac{1}{M}\sum_{j=1}^{M} \log\left(1 + \mathsf{SINR}^{(\infty)}\left(\beta_{l_{1,j}}\right)\right) = \ldots = \frac{1}{M}\sum_{j=1}^{M} \log\left(1 + \mathsf{SINR}^{(\infty)}\left(\beta_{l_{M,j}}\right)\right). \tag{B.14}$$

Next, we show that there is for each $\varepsilon > 0$ an M_0, such that for all $M \geq M_0$

$$\left|\xi - \tilde{\xi}\right| < \frac{\varepsilon}{2}, \tag{B.15}$$

where ξ fulfills [75]

$$\lim_{n\to\infty} \frac{1}{n} \log\det\left(\mathbf{I}_n + \frac{P}{n\cdot\sigma^2}\mathbf{H}\mathbf{H}^H\right) = \xi \text{ almost surely.} \tag{B.16}$$

Let us fix j and denote by $\overline{\mathsf{SINR}}_{k,j}$ the SINR of the kth transmit terminal in time slot j, when interference of each codeword is canceled individually, once it is decoded. That is, we consider the SINRs as seen in the capacity achieving structure. Thus, the following relation holds [67]:

$$\log\det\left(\mathbf{I}_n + \frac{P}{n\cdot\sigma^2}\mathbf{H}\mathbf{H}^H\right) = \sum_{k=1}^{n} \log(1 + \overline{\mathsf{SINR}}_{k,j}). \tag{B.17}$$

We fix the decoding order in this case, such that the codewords of the transmit terminals in set $\mathcal{M}_m$ are decoded in the decoding phases

$$\left\{\sum_{i<l_{m,j}} \left|\mathcal{M}_{\left(l_j^{-1}\right)_i}\right| + 1, \ldots, \sum_{i\leq l_{m,j}} \left|\mathcal{M}_{\left(l_j^{-1}\right)_i}\right|\right\}. \tag{B.18}$$

Here, we denote by $\left(l_j^{-1}\right)_i$ the inverse function of $l_{m,j}$ with respect to m. Since the SINR of each transmit terminal signal becomes the larger the more interference signals are canceled, we have $\overline{\mathsf{SINR}}_{k',j} \leq \mathsf{SINR}_{k,j}$ for all k, k', such that $k \in \mathcal{M}_m$ and $k' \in \mathcal{M}_{m'}$ and $l_{m,j} > l_{m',j}$. Thus, we conclude

$$\sum_{k=1}^{n} \log\left(1 + \mathsf{SINR}_{k,j}\right) = \sum_{i=1}^{M} \sum_{k\in\mathcal{M}_{\left(l_j^{-1}\right)_i}} \log\left(1 + \mathsf{SINR}_{k,j}\right) \tag{B.19}$$

$$\geq \sum_{i=2}^{M} \sum_{k\in\mathcal{M}_{\left(l_j^{-1}\right)_i}} \log\left(1 + \mathsf{SINR}_{k,j}\right) \tag{B.20}$$

$$\geq \sum_{k=1}^{n-\left|\mathcal{M}_{\left(l_j^{-1}\right)_M}\right|} \log\left(1 + \overline{\mathsf{SINR}}_{k,j}\right) \tag{B.21}$$

$$\geq \log\det\left(\mathbf{I}_n + \frac{P}{n\cdot\sigma^2}\mathbf{H}\mathbf{H}^H\right) - \left|\mathcal{M}_{\left(l_j^{-1}\right)_M}\right| \max_{k\in\mathcal{M}_{\left(l_j^{-1}\right)_M}} \log\left(1+\overline{\mathsf{SINR}}_{k,j}\right). \tag{B.22}$$

Here, the last inequality follows from (B.17). Taking the limit with respect to n in (B.19), normalized by n^{-1}, yields almost surely

$$\lim_{n\to\infty}\frac{1}{n}\sum_{j=1}^{M}\sum_{k\in\mathcal{M}_{\left(l_j^{-1}\right)_j}}\log\left(1+\mathsf{SINR}_{k,j}\right) = \sum_{m=1}^{M}\left(\lim_{n\to\infty}\frac{\left|\mathcal{M}_{l_{m,j}}\right|}{n}\right)\log\left(1+\mathsf{SINR}^{(\infty)}\left(\beta_{l_{m,j}}\right)\right) \tag{B.23}$$

$$= \frac{1}{M}\sum_{m=1}^{M}\log\left(1+\mathsf{SINR}^{(\infty)}\left(\beta_{l_{m,j}}\right)\right) = \tilde{\xi}. \tag{B.24}$$

Likewise, we obtain for (B.22) almost surely

$$\lim_{n\to\infty}\frac{1}{n}\log\det\left(\mathbf{I}_n + \frac{P}{n\cdot\sigma^2}\mathbf{H}\mathbf{H}^H\right) - \frac{\left|\mathcal{M}_{\left(l_i^{-1}\right)_M}\right|}{n}\max_{k\in\mathcal{M}_{\left(l_i^{-1}\right)_M}}\log\left(1+\overline{\mathsf{SINR}}_{k,j}\right) \tag{B.25}$$

$$= \xi - \frac{1}{M}\log\left(1+\mathsf{SINR}^{(\infty)}(0)\right). \tag{B.26}$$

Since $\mathsf{SINR}^{(\infty)}(0)$ is finite almost surely according to [96, Lemma 3.1], there exists for every $\varepsilon > 0$ an M_0, such that for all $M \geq M_0$ we have $\frac{1}{M}\log\left(1+\mathsf{SINR}^{(\infty)}(0)\right) < \varepsilon/2$. Thus, we eventually conclude

$$\xi - \frac{\varepsilon}{2} < \tilde{\xi} < \xi, \tag{B.27}$$

which establishes (B.15). The second inequality follows, since ξ is the normalized asymptotic sum-capacity of the channel.

We finally combine (B.13) and (B.15) in order to conclude that there exists for every $\varepsilon > 0$ almost surely an n_0, such that for all $n \geq n_0$

$$\max_{m\in\{1,\ldots,M\}}\left|R_m - \xi\right| \leq \left|R_m - \tilde{\xi}\right| + \left|\xi - \tilde{\xi}\right| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} < \varepsilon, \tag{B.28}$$

where we used the triangle inequality. Thus, the rate tuple $(R^{(1)}, R^{(2)}, \ldots, R^{(n)}) = (R, R, \ldots, R)$ is achievable almost surely in the limit $n \to \infty$, if $R < \xi$. □

Proof of Lemma 2. We decompose the matrix $\mathbf{H} \in \mathbb{C}^{n\times n}$ into the matrices $\tilde{\mathbf{H}}_m \in \mathbb{C}^{n_m\times n}$, $m \in \{1,\dots,M\}$, where

$$n_m = \begin{cases} \lceil n/M \rceil, & \text{if } 1 \le m \le n \mod M \\ \lfloor n/M \rfloor, & \text{else.} \end{cases} \tag{B.29}$$

The matrices $\tilde{\mathbf{H}}_m$ contain disjoint sections of $\mathbf{H}$, such that

$$\mathbf{H} = \left(\tilde{\mathbf{H}}_1^T \dots \tilde{\mathbf{H}}_M^T\right)^T. \tag{B.30}$$

With this notation, we obtain the following bounds on $\|\mathbf{h}_k^T\|^2$:

$$\frac{1}{n}\lambda_{\min}\left\{\tilde{\mathbf{H}}_m\tilde{\mathbf{H}}_m^H\right\} \le \min_{k:\lceil k/M\rceil = m} \frac{1}{n}\|\mathbf{h}_k^T\|^2 \le \max_{k:\lceil k/M\rceil = m} \frac{1}{n}\|\mathbf{h}_k^T\|^2 \le \frac{1}{n}\lambda_{\max}\left\{\tilde{\mathbf{H}}_m\tilde{\mathbf{H}}_m^H\right\}. \tag{B.31}$$

Here, we used that the maximum/minimum Eucledian norm of any row of a matrix is upper-/lower-bounded by the maximum/minimum singular value of the matrix.

From [79] it is known that almost surely

$$\lim_{n\to\infty} \lambda_{\max}\frac{1}{n}\left\{\tilde{\mathbf{H}}_m\tilde{\mathbf{H}}_m^H\right\} = \lim_{n\to\infty}(1+\sqrt{n_m/n})^2 = (1+\sqrt{1/M})^2, \tag{B.32}$$

$$\lim_{n\to\infty} \lambda_{\min}\frac{1}{n}\left\{\tilde{\mathbf{H}}_m\tilde{\mathbf{H}}_m^H\right\} = \lim_{n\to\infty}(1-\sqrt{n_m/n})^2 = (1-\sqrt{1/M})^2. \tag{B.33}$$

Since M is finite, we can conclude that also almost surely

$$\lim_{n\to\infty} \max_{m\in\{1,\dots,M\}} \frac{1}{n}\lambda_{\max}\left\{\tilde{\mathbf{H}}_m\tilde{\mathbf{H}}_m^H\right\} = (1+\sqrt{1/M})^2, \tag{B.34}$$

$$\lim_{n\to\infty} \min_{m\in\{1,\dots,M\}} \frac{1}{n}\lambda_{\min}\left\{\tilde{\mathbf{H}}_m\tilde{\mathbf{H}}_m^H\right\} = (1-\sqrt{1/M})^2. \tag{B.35}$$

In the following, we choose M for a given $\varepsilon > 0$ sufficiently large, such that

$$P(1+\sqrt{1/M})^2 < P + \frac{\varepsilon}{2} \tag{B.36}$$

and

$$P(1-\sqrt{1/M})^2 > P - \frac{\varepsilon}{2}. \tag{B.37}$$

Then, there exists for every $\varepsilon > 0$ almost surely an n_0, such that for all $n \geq n_0$

$$\begin{aligned} \max_{k\in\{1,\dots,n\}} \frac{P}{n}\|\mathbf{h}_k^T\|^2 - P &\leq \max_{m\in\{1,\dots,M\}} \frac{P}{n}\lambda_{\max}\{\mathbf{H}_m\mathbf{H}_m^H\} - P \\ &< P(1+\sqrt{1/M})^2 + \frac{\varepsilon}{2} - P < \frac{\varepsilon}{2}. \end{aligned} \tag{B.38}$$

The first inequality follows from (B.31), the second one from (B.34), and the third one from (B.36). Likewise, there exists for every $\varepsilon > 0$ almost surely an n_0, such that for all $n \geq n_0$.

$$\min_{k\in\{1,\dots,n\}} \frac{P}{n}\|\mathbf{h}_k^T\|^2 - P \geq \min_{m\in\{1,\dots,M\}} \frac{P}{n}\lambda_{\min}\{\mathbf{H}_m\mathbf{H}_m^H\} - P \tag{B.39}$$

$$> P(1-\sqrt{1/M})^2 - \frac{\varepsilon}{2} - P > -\frac{\varepsilon}{2}. \tag{B.40}$$

Again, the first inequality follows from (B.31), the second one from (B.35), and the third one from (B.37). The combination of the bounds (B.38) and (B.40) yields, that there also exists for every $\varepsilon > 0$ almost surely an n_0, such that for all $n \geq n_0$

$$\max_{k\in\{1,\dots,n\}} \left|\frac{P}{n}\|\mathbf{h}_k^T\|^2 - P\right| < \varepsilon, \tag{B.41}$$

which establishes the Lemma. □

Proof of Lemma 3. We rewrite the considered conditional mutual information as follows:

$$I(\mathsf{y}_{\mathcal{S}};\hat{\mathsf{y}}_{\mathcal{S}}|\hat{\mathsf{y}}_{\mathcal{S}^C}) = h(\hat{\mathsf{y}}_{\mathcal{S}}|\hat{\mathsf{y}}_{\mathcal{S}^C}) - h(\hat{\mathsf{y}}_{\mathcal{S}}|\mathsf{y}_{\mathcal{S}},\hat{\mathsf{y}}_{\mathcal{S}^C}) \tag{B.42}$$

$$= h(\hat{\mathsf{y}}_{\mathcal{S}}|\hat{\mathsf{y}}_{\mathcal{S}^C}) - h(\hat{\mathsf{y}}_{\mathcal{S}}|\mathsf{y}_{\mathcal{S}}) \tag{B.43}$$

$$= h(\hat{\mathsf{y}}_{\mathcal{S}},\hat{\mathsf{y}}_{\mathcal{S}^C}) - h(\hat{\mathsf{y}}_{\mathcal{S}^C}) - h(\hat{\mathsf{y}}_{\mathcal{S}}|\mathsf{y}_{\mathcal{S}}) \tag{B.44}$$

$$= h(\hat{\mathsf{y}}) - h(\hat{\mathsf{y}}_{\mathcal{S}^C}) - h(\hat{\mathsf{y}}_{\mathcal{S}}|\mathsf{y}_{\mathcal{S}}). \tag{B.45}$$

Here, we used the fact that $\hat{\mathsf{y}}_{\mathcal{S}}$ is independent of $\hat{\mathsf{y}}_{\mathcal{S}^C}$ when conditioned on $\mathsf{y}_{\mathcal{S}}$ in order to obtain (B.43). The three entropy expressions are evaluated as follows:

$$h(\hat{\mathsf{y}}) = n\log\left(2\pi e(\sigma_{\mathrm{w}}^2+\sigma_{\mathrm{q}}^2)\right) + \log\det\left(\mathbf{I}_n + \frac{P}{(\sigma_{\mathrm{w}}^2+\sigma_{\mathrm{q}}^2)\cdot n}\mathbf{H}\mathbf{H}^H\right), \tag{B.46}$$

$$h(\hat{\mathsf{y}}_{\mathcal{S}^C}) = \left|\mathcal{S}^C\right|\log\left(2\pi e(\sigma_{\mathrm{w}}^2+\sigma_{\mathrm{q}}^2)\right) + \log\det\left(\mathbf{I}_{|\mathcal{S}^C|} + \frac{P}{(\sigma_{\mathrm{w}}^2+\sigma_{\mathrm{q}}^2)\cdot n}\left(\mathbf{H}_{\mathcal{S}^C}\right)^H\mathbf{H}_{\mathcal{S}^C}\right) \tag{B.47}$$

$$= \left|\mathcal{S}^C\right| \log\left(2\pi e(\sigma_\mathrm{w}^2 + \sigma_\mathrm{q}^2)\right) + \log\det\left(\mathbf{I}_n + \frac{P}{(\sigma_\mathrm{w}^2 + \sigma_\mathrm{q}^2)\cdot n}\mathbf{H}_{\mathcal{S}^C}\left(\mathbf{H}_{\mathcal{S}^C}\right)^H\right), \tag{B.48}$$

$$h(\hat{\mathsf{y}}_\mathcal{S}|\mathsf{y}_\mathcal{S}) = h(\mathsf{z}) = |\mathcal{S}|\log(2\pi e\sigma_\mathrm{q}^2). \tag{B.49}$$

Here, $\mathbf{H}_{\mathcal{S}^C}$ denotes the matrix that contains the $|\mathcal{S}^C|$ columns of $\mathbf{H}$ whose indexes are contained in $\mathcal{S}^C$. Thus, we obtain according to (B.45)

$$\begin{aligned} I(\mathsf{y}_\mathcal{S};\hat{\mathsf{y}}_\mathcal{S}|\hat{\mathsf{y}}_{\mathcal{S}^C}) =& \log\det\left(\mathbf{I}_n + \frac{P}{(\sigma_\mathrm{w}^2 + \sigma_\mathrm{q}^2)\cdot n}\mathbf{H}\mathbf{H}^H\right) \\ &- \log\det\left(\mathbf{I}_n + \frac{P}{(\sigma_\mathrm{w}^2 + \sigma_\mathrm{q}^2)\cdot n}\mathbf{H}_{\mathcal{S}^C}\left(\mathbf{H}_{\mathcal{S}^C}\right)^H\right) + |\mathcal{S}|\log(1 + \frac{\sigma_\mathrm{w}^2}{\sigma_\mathrm{q}^2}). \end{aligned} \tag{B.50}$$

Next, we use the following corollary that follows from Lemma 1:

Corollary 1. *Let the elements of the random matrix* $\mathbf{H} \in \mathbb{C}^{n\times n}$ *be distributed according to the assumptions of Theorem 3. Let* P *and* σ^2 *be positive constants. Then, there exists almost surely an* n_0*, such that for all* $n \geq n_0$*:*

$$\begin{aligned} &\frac{|\mathcal{S}|}{n}\left(\frac{1}{n}\log\det\left(\mathbf{I}_n + \frac{P}{n\cdot\sigma^2}\mathbf{H}\mathbf{H}^H\right)\right) \\ &\qquad \leq \frac{1}{n}\log\det\left(\mathbf{I}_n + \frac{P}{n\cdot\sigma^2}\mathbf{H}_\mathcal{S}\left(\mathbf{H}_\mathcal{S}\right)^H\right) \\ &\qquad\qquad \textit{for all } \mathcal{S} \subseteq \{1,\ldots,n\} \end{aligned} \tag{B.51}$$

where $\mathbf{H}_\mathcal{S}$ *denotes the matrix that contains the* $|\mathcal{S}|$ *columns of* $\mathbf{H}$ *whose indexes are contained in* $\mathcal{S}$*.*

The proof of this corollary is provided subsequent to this proof. According to Corollary 1, there exists almost surely an n_0, such that for all $n \geq n_0$

$$\begin{aligned} &\frac{1}{n}\log\det\left(\mathbf{I}_n + \frac{P}{(\sigma_\mathrm{w}^2 + \sigma_\mathrm{q}^2)\cdot n}\mathbf{H}_{\mathcal{S}^C}\left(\mathbf{H}_{\mathcal{S}^C}\right)^H\right) \\ &\qquad \geq \frac{\left|\mathcal{S}^C\right|}{n}\left(\frac{1}{n}\log\det\left(\mathbf{I}_n + \frac{P}{(\sigma_\mathrm{w}^2 + \sigma_\mathrm{q}^2)\cdot n}\mathbf{H}\mathbf{H}^H\right)\right) \\ &\qquad\qquad \text{for all } \mathcal{S} \subseteq \{1,\ldots,n\}. \end{aligned} \tag{B.52}$$

We apply this result to (B.50), and conclude that there is almost surely an n_0, such that for all $n \geq n_0$

$$\frac{1}{n} I(\mathsf{y}; \hat{\mathsf{y}}_{\mathcal{S}} | \hat{\mathsf{y}}_{\mathcal{S}^C}) \leq \frac{1}{n} \cdot \left(1 - \frac{\left|\mathcal{S}^C\right|}{n}\right) \cdot \log\det\left(\mathbf{I}_n + \frac{P}{(\sigma_{\mathrm{w}}^2 + \sigma_{\mathrm{q}}^2) \cdot n} \mathbf{H}\mathbf{H}^H\right) + \frac{|\mathcal{S}|}{n} \log\left(1 + \frac{\sigma_{\mathrm{w}}^2}{\sigma_{\mathrm{q}}^2}\right) \tag{B.53}$$

$$= \frac{|\mathcal{S}|}{n} \cdot \frac{1}{n} \log\det\left(\left(1 + \frac{\sigma_{\mathrm{w}}^2}{\sigma_{\mathrm{q}}^2}\right) \mathbf{I}_n + \frac{P}{n} \mathbf{H}\mathbf{H}^H\right) \text{ for all } \mathcal{S} \subseteq \{1, \ldots, n\}, \tag{B.54}$$

where (B.54) converges to $\frac{|\mathcal{S}|}{n}\zeta$ almost surely as $n \to \infty$, i.e., there exists almost surely for every tuple of rates of compressed quantization codebooks $(R^{(1)}, R^{(2)}, \ldots, R^{(n)}) = (R, R, \ldots, R)$ with $R > \zeta$ an n_0, such that for all $n \geq n_0$ the quantization noise variance σ_{q}^2 is achievable in the sense of (3.31). □

Proof of Corollary 1. Lemma 1 implies that there is for all $R < \xi$ almost surely an n_0, such that for all $n \geq n_0$

$$|\mathcal{S}| R < \log\det\left(\mathbf{I}_n + \frac{P}{n \cdot \sigma^2} \cdot \mathbf{H}_{\mathcal{S}} (\mathbf{H}_{\mathcal{S}})^H\right) \text{ for all } \mathcal{S} \subseteq \{1, \ldots, n\}. \tag{B.55}$$

Since almost surely

$$\xi = \lim_{n \to \infty} \frac{1}{n} \log\det\left(\mathbf{I}_n + \frac{P}{n \cdot \sigma^2} \mathbf{H}\mathbf{H}^H\right), \tag{B.56}$$

this is equivalent to (B.51). □

B.2. Proofs of Lemmata for Theorem 5

Proof of Lemma 5. The lemma follows from the fact that the limit can be taken inside a continuous function, which allows us to write

$$\lim_{\kappa \to \infty} \left(\frac{c}{g(\kappa)} + 1\right)^{\kappa} = \exp\left(\lim_{\kappa \to \infty} \operatorname{Log}\left(\frac{c}{g(\kappa)} + 1\right)^{\kappa}\right) \tag{B.57}$$

$$= \exp\left(\lim_{\kappa \to \infty} \kappa \log\left|\frac{c}{g(\kappa)} + 1\right| + \lim_{\kappa \to \infty} \imath \arg\left\{\frac{c}{g(\kappa)} + 1\right\}\right) \tag{B.58}$$

$$= \exp\left(\lim_{\kappa \to \infty} \kappa \log\left|\frac{c}{g(\kappa)} + 1\right|\right) \tag{B.59}$$

$$= \exp\left(\lim_{\kappa\to\infty} \Re\left\{\kappa \mathrm{Log}\left(\frac{c}{g(\kappa)}+1\right)\right\}\right). \tag{B.60}$$

From the rule of Bernoulli-l'Hospital, we know that

$$\lim_{\kappa\to\infty} \kappa \cdot \mathrm{Log}\left(\frac{c}{M\kappa^\gamma}+1\right) = \lim_{\kappa\to\infty} \frac{c\gamma\kappa}{c+M\kappa^\gamma}, \tag{B.61}$$

where γ and M are positive constants.

If $g(\kappa) \in \Omega(\kappa^{1+\varepsilon})$, there exists by definition some $M > 0$, such that the absolute value of the argument of the exponential function in (B.60) can be upper-bounded according to

$$\lim_{\kappa\to\infty} \left|\kappa \cdot \mathrm{Log}\left(\frac{c}{g(\kappa)}+1\right)\right| \le \lim_{\kappa\to\infty} \left|\kappa \cdot \mathrm{Log}\left(\frac{c}{M\kappa^{1+\varepsilon}}+1\right)\right|. \tag{B.62}$$

Evaluating (B.61) for $\gamma = 1+\varepsilon$ renders this upper-bound zero, which establishes that also the left hand side of (B.62) becomes zero and (B.60) evaluates to one in this case.

For the proof of the other two cases (note that c is real and negative now), we write analogously to (B.60)

$$\lim_{\kappa\to\infty} \left(\frac{c}{g(\kappa)}+1\right)^\kappa = \exp\left(\lim_{\kappa\to\infty} \kappa \log\left(\frac{c}{g(\kappa)}+1\right)\right). \tag{B.63}$$

If $g(\kappa) \in \Theta(\kappa)$, there exist constants M_1 and M_2, $M_1 \ge M_2 > 0$, and κ_0, such that for all $\kappa \ge \kappa_0$, the exponent in (B.63) is sandwiched between

$$\frac{c}{M_1} = \lim_{\kappa\to\infty} \kappa \cdot \log\left(\frac{c}{M_1\kappa}+1\right) \tag{B.64}$$

$$\ge \kappa \cdot \log\left(\frac{c}{g(\kappa)}+1\right) \tag{B.65}$$

$$\ge \lim_{\kappa\to\infty} \kappa \cdot \log\left(\frac{c}{M_2\kappa}+1\right) = \frac{c}{M_2}, \tag{B.66}$$

where the limit is obtained by evaluating the right hand side of (B.61) for $\gamma = 1$. This establishes the second case.

Since $g(\kappa) \in \mathcal{O}(\kappa^{1-\varepsilon})$, there exists some $M > 0$, such that the argument of the exponential function in (B.63) can be upper-bounded according to

$$\lim_{\kappa\to\infty} \kappa \cdot \log\left(\frac{c}{g(\kappa)}+1\right) \le \lim_{\kappa\to\infty} \kappa \cdot \log\left(\frac{c}{M\kappa^{1-\varepsilon}}+1\right). \tag{B.67}$$

The limit (B.61) does not exist for $\gamma = 1 - \varepsilon$, which implies that both sides of (B.67) evaluate to minus infinity, and (B.60) to zero. $\square$

Proof of Lemma 6. We establish the following chain of identities:

$$\begin{aligned}
&\lim_{n\to\infty} \frac{1}{n} \left\| \mathbf{I}_n - \mathbf{A}^{(n)}(\gamma) \right\|_* \\
&\quad = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \sigma_i\{\mathbf{I}_n - \mathbf{A}^{(n)}(\gamma)\} && \text{(B.68)}\\
&\quad = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \left|\lambda_i\{\mathbf{I}_n - \mathbf{A}^{(n)}(\gamma)\}\right| && \text{(B.69)}\\
&\quad = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \left|1 - \lambda_i\{\mathbf{A}^{(n)}(\gamma)\}\right| && \text{(B.70)}\\
&\quad = \lim_{n\to\infty} \frac{1}{n} \left[\sum_{i:\lambda_i\{\mathbf{A}^{(n)}(\gamma)\}\le 1} \left(1 - \lambda_i\{\mathbf{A}^{(n)}(\gamma)\}\right) + \sum_{i:\lambda_i\{\mathbf{A}^{(n)}(\gamma)\}>1} \left(\lambda_i\{\mathbf{A}^{(\gamma)}\} - 1\right) \right] && \text{(B.71)}\\
&\quad = \lim_{n\to\infty} \left[\int_0^1 \left| F_{\mathbf{A}}^{(n,\gamma)}(x) \right| \cdot dx + \int_1^\infty \left| F_{\mathbf{A}}^{(n,\gamma)}(x) - 1 \right| \cdot dx \right] && \text{(B.72)}\\
&\quad = \lim_{n\to\infty} \int_0^\infty \left| F_{\mathbf{A}}^{(n,\gamma)}(x) - \sigma(x-1) \right| \cdot dx && \text{(B.73)}\\
&\quad = \int_0^\infty \left| F_{\mathbf{A}}^{(\gamma)}(x) - \sigma(x-1) \right| \cdot dx \text{ almost surely.} && \text{(B.74)}
\end{aligned}$$

In (B.69), we write the nuclear norm of the matrix $\mathbf{I}_n - \mathbf{A}$ in terms of its singular values σ_i. Since the matrix $\mathbf{I}_n - \mathbf{A}^{(\gamma)}$ is normal, i.e., $(\mathbf{I}_n - \mathbf{A}^{(\gamma)})^H(\mathbf{I}_n - \mathbf{A}^{(\gamma)}) = (\mathbf{I}_n - \mathbf{A}^{(\gamma)})(\mathbf{I}_n - \mathbf{A}^{(\gamma)})^H$, its singular values coincide with the absolute values of its eigenvalues. In (B.71), we arrange the terms, such that they can be related to the EED of $\mathbf{A}^{(\gamma)}$. In (B.74), we take the limit inside the integral. This is justified, since the maximum eigenvalue of $\mathbf{A}^{(\gamma)}$ converges to some bounded constant by assumption. Thus, the integration is over the compact interval $[0, \lambda_{\max}^{(\gamma)}]$, where we integrate over a uniformly convergent sequence (in n) of functions. This establishes the equivalence between 1) and 3). It remains to establish that

$$\lim_{\gamma\to\infty} \int_0^{\lambda_{\max}^{(\gamma)}} |F_{\mathbf{A}}^{(\gamma)}(x) - \sigma(x-1)| \cdot dx = 0$$

$$\iff \lim_{\gamma\to\infty} |F_{\mathbf{A}}^{(\gamma)}(x) - \sigma(x-1)| = 0 \qquad \text{for all } x. \tag{B.75}$$

For the forward part, consider $\epsilon(x) \triangleq |F_{\mathbf{A}}^{(\gamma)}(x) - \sigma(x-1)|$ for $x \in [0,1)$, i.e., $\epsilon(x) = |F_{\mathbf{A}}^{(\gamma)}(x)|$. Fix any $\Delta \in [-1,0)$. Since $\epsilon(x)$ is monotonically increasing on the interval of interest, we can write

$$\int_{1+\Delta}^{1} |F_{\mathbf{A}}^{(\gamma)}(x) - \sigma(x-1)| \cdot dx > |\Delta| \cdot \epsilon(1+\Delta). \tag{B.76}$$

Thus, if $\epsilon(1+\Delta)$ does not tend to zero for all Δ, the integral cannot tend to zero. The same reasoning can be applied for the interval $\Delta \in [1, \lambda_{\max}^{(\gamma)}]$.

For the backward part, we break the integration in (B.74) into two parts. The first integral is from zero to some constant d, $1 < d < \lambda_{\max}^{(\gamma)}$. $F_{\mathbf{A}}^{(\gamma)}(x)$ is a sequence of Riemann integrable functions that is uniformly bounded and pointwise convergent. By the bounded convergence theorem for the Riemann integral (e.g. [97]) we can take the limit inside the integral. Thus, the limit of this first integral is zero. The second part of the integral is from d to $\lambda_{\max}^{(\gamma)}$. Here, the limit cannot be taken inside the integral in general (note that $\lim_{\gamma\to\infty} \lambda_{\max}^{(\gamma)}$ might be unbounded). However, we can write

$$\lim_{\gamma\to\infty} \int_{d}^{\lambda_{\max}^{(\gamma)}} 1 - F_{\mathbf{A}}^{(\gamma)}(x) \cdot dx = \lim_{\gamma\to\infty} \int_{0}^{\lambda_{\max}^{(\gamma)}} 1 - F_{\mathbf{A}}^{(\gamma)}(x) \cdot dx - \lim_{\gamma\to\infty} \int_{0}^{d} 1 - F_{\mathbf{A}}^{(\gamma)}(x) \cdot dx = 0. \tag{B.77}$$

The second term on the right hand side of (B.77) evaluates to one, since the limit can be taken inside the integral. Again, this is justified by the bounded convergence theorem (e.g. [97]). The first integral on the right hand side evaluates to one for every γ, since otherwise there would be a contradiction between the following two statements:

$$\lim_{n\to\infty} \int_{0}^{\lambda_1\{\mathbf{A}(\gamma)\}} 1 - F_{\mathbf{A}}^{(n,\gamma)}(x) \cdot dx = 1 \text{ almost surely for each } \gamma, \tag{B.78}$$

$$\lim_{n\to\infty} \int_{0}^{\lambda_1\{\mathbf{A}(\gamma)\}} 1 - F_{\mathbf{A}}^{(n,\gamma)}(x) \cdot dx = \int_{0}^{\lambda_{\max}^{(\gamma)}} 1 - F_{\mathbf{A}}^{(\gamma)}(x) \cdot dx \text{ almost surely for each } \gamma. \tag{B.79}$$

The first statement corresponds to the assumption $\lim_{n\to\infty} n^{-1}\mathrm{Tr}[\mathbf{A}(\gamma)] = 1$ almost surely for each γ, the second statement follows by taking the limit inside the integral, which is fine for every fixed value of $\lambda_{\max}^{(\gamma)}$. □

Proof of Lemma 7 We go through each of the assumptions required for Lemma 6:

1. To show that the EED of $\tilde{\mathbf{R}}_n$ converges uniformly to a nonrandom distribution irrespective of the values of L and $\beta_{\mathcal{R}}$, we use a result from [98, 99] (also [26, Theorem 2.39]):

   Let $\mathbf{Y} = \frac{1}{k}\mathbf{X}\mathbf{T}\mathbf{X}^H$, where

   - $\mathbf{T}$ is an arbitrary nonnegative definite $k \times k$ random matrix whose EED for every fixed ratio k/n converges uniformly to a nonrandom distribution almost surely as $n \to \infty$,
   - $\mathbf{X}$ is an $n \times k$ random matrix, independent of $\mathbf{T}$, with i.i.d. elements of unit variance.

   Then, the EED of $\mathbf{Y}$ converges uniformly to a nonrandom distribution function almost surely as $n \to \infty$. If $\mathbf{X}$ is identified with $\mathbf{H}_{l+1}$ and $\mathbf{T}$ with

$$\mathbf{I}_{n_{\mathcal{D}}} + \frac{\alpha}{n_{\mathcal{R}}}\mathbf{H}_l\left(\mathbf{I}_{n_{\mathcal{D}}} + \ldots + \frac{\alpha}{n_{\mathcal{R}}}\mathbf{H}_2\left(\mathbf{I}_{n_{\mathcal{D}}} + \frac{\alpha}{n_{\mathcal{R}}}\mathbf{H}_1\mathbf{H}_1^H\right)\mathbf{H}_2^H\cdots\right)\mathbf{H}_l^H,$$

   uniform and almost sure convergence of the EED of $\tilde{\mathbf{R}}_n$ follows by induction.

2. For the trace condition, we obtain for arbitrary L and $\beta_{\mathcal{R}}$ by application of Lemma 8 (see p. 112):

$$\begin{aligned}
&\lim_{n_{\mathcal{D}}\to\infty}\frac{1}{n_{\mathcal{D}}}\mathrm{Tr}\left[\frac{1-\alpha}{1-\alpha^{L+1}}\cdot\mathbf{R}_n\right] \\
&\quad= \frac{1-\alpha}{1-\alpha^{L+1}}\sum_{l=0}^{L}\alpha^l\lim_{n_{\mathcal{D}}\to\infty}\frac{1}{n_{\mathcal{D}}}\mathrm{Tr}\left[\frac{1}{n_{\mathcal{R}}^l}\mathbf{H}_1\cdots\mathbf{H}_l\mathbf{H}_l^H\cdots\mathbf{H}_1^H\right] && \text{(B.80)}\\
&\quad= \frac{1-\alpha}{1-\alpha^{L+1}}\sum_{l=0}^{L}\alpha^l = 1 \text{ almost surely.} && \text{(B.81)}
\end{aligned}$$

   Note that $\beta_{\mathcal{R}}$ and L are fixed here, and thus the sum is finite.

3. For the condition on the maximum eigenvalue, we use the triangle inequality and the submultiplicativity of the spectral norm:

$$\lim_{n_{\mathcal{D}}\to\infty}\lambda_{\max}\left\{\frac{1-\alpha}{1-\alpha^{L+1}}\mathbf{R}_n\right\}$$

$$= \frac{1-\alpha}{1-\alpha^{L+1}} \lim_{n_{\mathcal{D}}\to\infty} \lambda_{\max}\left\{\sum_{l=0}^{L} \alpha^{l} \cdot \frac{1}{n_{\mathcal{R}}^{l}} \mathbf{H}_1 \cdots \mathbf{H}_l \mathbf{H}_l^H \cdots \mathbf{H}_1^H\right\} \tag{B.82}$$

$$\leq \frac{1-\alpha}{1-\alpha^{L+1}} \sum_{l=0}^{L} \alpha^{l} \cdot \lim_{n_{\mathcal{D}}\to\infty} \lambda_{\max}\left\{\frac{1}{n_{\mathcal{R}}^{l}} \mathbf{H}_1 \cdots \mathbf{H}_l \mathbf{H}_l^H \cdots \mathbf{H}_1^H\right\} \tag{B.83}$$

$$= \frac{1-\alpha}{1-\alpha^{L+1}} \sum_{l=0}^{L} \alpha^{l} \cdot \lim_{n_{\mathcal{D}}\to\infty} \lambda_{\max}\left\{\frac{1}{n_{\mathcal{R}}^{l}} \mathbf{H}_2 \cdots \mathbf{H}_l \mathbf{H}_l^H \cdots \mathbf{H}_1^H \mathbf{H}_1\right\} \tag{B.84}$$

$$\leq \frac{1-\alpha}{1-\alpha^{L+1}} \sum_{l=0}^{L} \alpha^{l} \times \lim_{n_{\mathcal{D}}\to\infty} \lambda_{\max}\left\{\frac{1}{n_{\mathcal{R}}^{l-1}} \mathbf{H}_2 \cdots \mathbf{H}_l \mathbf{H}_l^H \cdots \mathbf{H}_2^H\right\} \lim_{n_{\mathcal{D}}\to\infty} \lambda_{\max}\left\{\frac{1}{n_{\mathcal{R}}} \mathbf{H}_1 \mathbf{H}_1^H\right\}, \tag{B.85}$$

where we used that the nonzero eigenvalues of the matrix products in (B.83) and (B.84) coincide. Repeated application of steps (B.84) and (B.85) finally yields

$$\lim_{n_{\mathcal{D}}\to\infty} \lambda_{\max}\left\{\frac{1-\alpha}{1-\alpha^{L+1}} \mathbf{R}_n\right\}$$

$$\leq \frac{1-\alpha}{1-\alpha^{L+1}} \sum_{l=0}^{L} \alpha^{l} \prod_{l'=1}^{l} \lim_{n_{\mathcal{D}}\to\infty} \lambda_{\max}\left\{\frac{1}{n_{\mathcal{R}}} \mathbf{H}_{l'} \mathbf{H}_{l'}^H\right\} \tag{B.86}$$

$$= \frac{1-\alpha}{1-\alpha^{L+1}} \sum_{l=0}^{L} \alpha^{l} \cdot 4^{l-1} \left(1 + \sqrt{\frac{n_{\mathcal{D}}}{n_{\mathcal{R}}}}\right)^2 < \infty \text{ almost surely,} \tag{B.87}$$

where each of the multipliers in (B.86) is bounded according to (3.114). The ratio $n_{\mathcal{D}}/n_{\mathcal{R}}$ and L are fixed, such that also the resulting product is finite. □

Proof of Lemma 8. The proof relies on a result of [98,99] (also [26, Theorem 2.39]): Let $\mathbf{Y} = \frac{1}{k}\mathbf{X}\mathbf{T}\mathbf{X}^H$ be a random matrix, where

- $\mathbf{T}$ is an arbitrary nonnegative definite $k \times k$ random matrix whose EED converges uniformly to a nonrandom distribution almost surely as $k \to \infty$ and satisfies $\lim_{k\to\infty} k^{-1}\mathrm{Tr}\,[\mathbf{T}] = T$ almost surely,
- $\mathbf{X}$ is an $n \times k$ random matrix, independent of $\mathbf{T}$, with i.i.d. elements of unit variance.

Then, for every fixed ratio k/n

$$\lim_{n\to\infty} \frac{1}{n}\mathrm{Tr}\,[\mathbf{Y}] = \lim_{n\to\infty} \frac{1}{k}\mathrm{Tr}\,[\mathbf{T}] = T \text{ almost surely.} \tag{B.88}$$

Since the trace of a product is invariant under cyclic permutation of the factors, we can write

$$n_{\mathcal{D}}^{-1}\mathrm{Tr}\left[\frac{1}{n_{\mathcal{R}}^2}\mathbf{H}_1\mathbf{H}_2\mathbf{H}_2^H\mathbf{H}_1^H\right] = n_{\mathcal{D}}^{-1}\mathrm{Tr}\left[\frac{1}{n_{\mathcal{R}}}\mathbf{H}_2^H\left(\frac{1}{n_{\mathcal{R}}}\mathbf{H}_1^H\mathbf{H}_1\right)\mathbf{H}_2\right]. \tag{B.89}$$

If $n_{\mathcal{R}}^{-1}\mathbf{H}_1^H\mathbf{H}_1$ is identified with $\mathbf{T}$, where $n = k = n_{\mathcal{R}}$ and $T = n_{\mathcal{D}}/n_{\mathcal{R}}$, we obtain

$$\lim_{n_{\mathcal{D}}\to\infty} n_{\mathcal{D}}^{-1}\mathrm{Tr}\left[\frac{1}{n_{\mathcal{R}}}\mathbf{H}_2^H\left(\frac{1}{n_{\mathcal{R}}}\mathbf{H}_1^H\mathbf{H}_1\right)\mathbf{H}_2\right] = \lim_{n_{\mathcal{D}}\to\infty} n_{\mathcal{D}}^{-1}\mathrm{Tr}\left[\frac{1}{n_{\mathcal{R}}}\mathbf{H}_1^H\mathbf{H}_1\right] = 1 \text{ almost surely.} \tag{B.90}$$

Repeating the argument for

$$n_{\mathcal{D}}^{-1}\mathrm{Tr}\left[\frac{1}{n_{\mathcal{R}}^3}\mathbf{H}_1\mathbf{H}_2\mathbf{H}_3\mathbf{H}_3^H\mathbf{H}_2^H\mathbf{H}_1^H\right] = n_{\mathcal{D}}^{-1}\mathrm{Tr}\left[\frac{1}{n_{\mathcal{R}}}\mathbf{H}_3^H\left(\frac{1}{n_{\mathcal{R}}^2}\mathbf{H}_2^H\mathbf{H}_1^H\mathbf{H}_1\mathbf{H}_2\right)\mathbf{H}_3\right] \tag{B.91}$$

with $\mathbf{T} = n_{\mathcal{R}}^{-2}\mathbf{H}_2^H\mathbf{H}_1^H\mathbf{H}_1\mathbf{H}_2$ yields

$$\lim_{n_{\mathcal{D}}\to\infty} n_{\mathcal{D}}^{-1}\mathrm{Tr}\left[\frac{1}{n_{\mathcal{R}}}\mathbf{H}_3^H\left(\frac{1}{n_{\mathcal{R}}^2}\mathbf{H}_2^H\mathbf{H}_1^H\mathbf{H}_1\mathbf{H}_2\right)\mathbf{H}_3\right]$$

$$= \lim_{n_{\mathcal{D}}\to\infty} n_{\mathcal{D}}^{-1}\mathrm{Tr}\left[\frac{1}{n_{\mathcal{R}}^2}\mathbf{H}_2^H\mathbf{H}_1^H\mathbf{H}_1\mathbf{H}_2\right] \tag{B.92}$$

$$= \lim_{n_{\mathcal{D}}\to\infty} n_{\mathcal{D}}^{-1}\mathrm{Tr}\left[\frac{1}{n_{\mathcal{R}}}\mathbf{H}_1^H\mathbf{H}_1\right] = 1 \text{ almost surely.} \tag{B.93}$$

Finally, we obtain with $\mathbf{T} = n_{\mathcal{R}}^{-L}\mathbf{H}_L^H\cdots\mathbf{H}_2^H\mathbf{H}_1^H\mathbf{H}_1\mathbf{H}_2\cdots\mathbf{H}_L$

$$\lim_{n_{\mathcal{D}}\to\infty} n_{\mathcal{D}}^{-1}\mathrm{Tr}\left[\frac{1}{n_{\mathcal{S}}}\mathbf{H}_{L+1}^H\left(\frac{1}{n_{\mathcal{R}}^L}\mathbf{H}_L^H\cdots\mathbf{H}_2^H\mathbf{H}_1^H\mathbf{H}_1\mathbf{H}_2\cdots\mathbf{H}_L\right)\mathbf{H}_{L+1}\right] = 1 \text{ almost surely,} \tag{B.94}$$

which completes the proof. □

Proof of Lemma 9. We define the matrices

$$\mathbf{A}_l \triangleq \left(\frac{\alpha}{n_{\mathcal{R}}}\right)^{L-l+2}\mathbf{H}_l\mathbf{H}_{l+1}\cdots\mathbf{H}_L\mathbf{H}_{L+1}\mathbf{H}_{L+1}^H\mathbf{H}_L^H\cdots\mathbf{H}_{l+1}^H\mathbf{H}_l^H \tag{B.95}$$

and

$$\mathbf{B}_l \triangleq \mathbf{I}_{n_\mathcal{R}} + \frac{\alpha}{n_\mathcal{R}}\mathbf{H}_l \left(\mathbf{I}n_\mathcal{R} + \ldots + \frac{\alpha}{n_\mathcal{R}}\mathbf{H}_{L_0-1}\left(\mathbf{I}_{n_\mathcal{R}} + \frac{\alpha}{n_\mathcal{R}}\mathbf{H}_{L_0}\mathbf{H}_{L_0}^H\right)\mathbf{H}_{L_0-1}^H \cdots\right)\mathbf{H}_l^H. \tag{B.96}$$

Note that $\sigma_\mathrm{w}^2 \cdot (\mathbf{I}_{n_\mathcal{D}} + \mathbf{H}_1\mathbf{B}_2\mathbf{H}_1^H)$ corresponds to the noise covariance matrix at the destination antennas under the assumption of noiseless relay stages $\mathcal{R}_L$ to $\mathcal{R}_{L-L_0+1}$. Therefore, we obtain the following upper-bound:

$$\begin{aligned}
\mathsf{E}\left[R_L^{\mathrm{AF}}\right] &\leq \mathsf{E}\left[\log\det\left(\mathbf{I}_{\mathbf{n}_\mathcal{D}} + \frac{P_L}{\sigma_\mathrm{w}^2 \cdot n_\mathcal{S}}\mathbf{H}_1\mathbf{A}_2\mathbf{H}_1^H \cdot \left(\mathbf{I}_{n_\mathcal{D}} + \mathbf{H}_1\mathbf{B}_2\mathbf{H}_1^H\right)^{-1}\right)\right] \\
&= \mathsf{E}\left[\log\det\left(\mathbf{I}_{n_\mathcal{D}} + \frac{P_L}{\sigma_\mathrm{w}^2 \cdot n_\mathcal{S}}\mathbf{U}_1\tilde{\mathbf{S}}_1\tilde{\mathbf{V}}_1^H\mathbf{A}_2\tilde{\mathbf{V}}_1\tilde{\mathbf{S}}_1\mathbf{U}_1^H\right.\right. \\
&\qquad\qquad \left.\left.\times\left(\mathbf{I}_{n_\mathcal{D}} + \mathbf{U}_1\tilde{\mathbf{S}}_1\tilde{\mathbf{V}}_1^H\mathbf{B}_2\tilde{\mathbf{V}}_1\tilde{\mathbf{S}}_1\mathbf{U}_1^H\right)^{-1}\right)\right] &\text{(B.97)} \\
&= \mathsf{E}\left[\log\det\left(\mathbf{I}_{n_\mathcal{D}} + \frac{P_L}{\sigma_\mathrm{w}^2 \cdot n_\mathcal{S}}\tilde{\mathbf{V}}_1^H\mathbf{A}_2\tilde{\mathbf{V}}_1\tilde{\mathbf{S}}_1^2 \cdot \left(\mathbf{I}_{n_\mathcal{D}} + \tilde{\mathbf{V}}_1^H\mathbf{B}_2\tilde{\mathbf{V}}_1\tilde{\mathbf{S}}_1^2\right)^{-1}\right)\right] &\text{(B.98)} \\
&\leq \mathsf{E}\left[\log\det\left(\mathbf{I}_{n_\mathcal{D}} + \frac{P_L}{\sigma_\mathrm{w}^2 \cdot n_\mathcal{S}}\tilde{\mathbf{V}}_1^H\mathbf{A}_2\tilde{\mathbf{V}}_1\frac{\mathrm{Tr}[\tilde{\mathbf{S}}_1^2]}{n_\mathcal{R}} \cdot \left(\mathbf{I}_{n_\mathcal{D}} + \tilde{\mathbf{V}}_1^H\mathbf{B}_2\tilde{\mathbf{V}}_1\frac{\mathrm{Tr}[\tilde{\mathbf{S}}_1^2]}{n_\mathcal{R}}\right)^{-1}\right)\right] &\text{(B.99)} \\
&= \mathsf{E}\left[\log\det\left(\mathbf{I}_{n_\mathcal{D}} + \frac{s_1\alpha P_L}{\sigma_\mathrm{w}^2 \cdot n_\mathcal{S}}\tilde{\mathbf{H}}_2\mathbf{A}_3\tilde{\mathbf{H}}_2^H \cdot \left((1+s_1\alpha)\mathbf{I}_{n_\mathcal{D}} + s_1\alpha\tilde{\mathbf{H}}_2\mathbf{B}_3\tilde{\mathbf{H}}_2^H\right)^{-1}\right)\right]. &\text{(B.100)}
\end{aligned}$$

In (B.97), we use the singular value decomposition $\mathbf{H}_1 = \mathbf{U}_1\mathbf{S}_1\mathbf{V}_1^H = \mathbf{U}_1\tilde{\mathbf{S}}_1\tilde{\mathbf{V}}_1^H$, where $\tilde{\mathbf{S}}_1$ is an $n_\mathcal{D} \times n_\mathcal{D}$ diagonal matrix with all nonzero singular values on its diagonal and $\tilde{\mathbf{V}}_1^H$ an $n_\mathcal{D} \times n_\mathcal{R}$ matrix derived from $\tilde{\mathbf{V}}_1$ by deleting the rows that correspond to the zero singular values.

Eq. (B.98) follows from the following chain of identities, where we define $\mathbf{X} \triangleq \mathbf{U}_1\tilde{\mathbf{S}}_1$:

$$\mathbf{X}\tilde{\mathbf{A}}_2\mathbf{X}^H\left(\mathbf{I}_{n_\mathcal{D}} + \mathbf{X}\tilde{\mathbf{B}}_2\mathbf{X}^H\right)^{-1} = \mathbf{X}\tilde{\mathbf{A}}_2\mathbf{X}^H\mathbf{X}\mathbf{X}^{-1}\left(\mathbf{I}_{n_\mathcal{D}} + \mathbf{X}\tilde{\mathbf{B}}_2\mathbf{X}^H\right)^{-1} \tag{B.101}$$

$$= \mathbf{X}\tilde{\mathbf{A}}_2\mathbf{X}^H\mathbf{X}\left(\mathbf{X} + \mathbf{X}\tilde{\mathbf{B}}_2\mathbf{X}^H\mathbf{X}\right)^{-1} = \mathbf{X}\tilde{\mathbf{A}}_2\mathbf{X}^H\mathbf{X}\left(\mathbf{I}_{n_\mathcal{D}} + \tilde{\mathbf{B}}_2\mathbf{X}^H\mathbf{X}\right)^{-1}\mathbf{X}^{-1} \tag{B.102}$$

and the fact that

$$\lambda_i \left\{ \mathbf{X}\tilde{\mathbf{A}}_2\mathbf{X}^H\mathbf{X}\left(\mathbf{I}_{n_{\mathcal{D}}} + \tilde{\mathbf{B}}_2\mathbf{X}^H\mathbf{X}\right)^{-1}\mathbf{X}^{-1} \right\} = \lambda_i \left\{ \tilde{\mathbf{A}}_2\mathbf{X}^H\mathbf{X}\left(\mathbf{I}_{n_{\mathcal{D}}} + \tilde{\mathbf{B}}_2\mathbf{X}^H\mathbf{X}\right)^{-1}\mathbf{X}^{-1}\mathbf{X} \right\}. \tag{B.103}$$

Eq. (B.99) follows from the fact that $\log\det\left(\mathbf{I}_{n_{\mathcal{D}}} + \tilde{\mathbf{A}}_2\tilde{\mathbf{S}}_1(\mathbf{I}_{n_{\mathcal{D}}} + \tilde{\mathbf{B}}_2\tilde{\mathbf{S}}_1)^{-1}\right)$ is concave on the set of positive definite matrices $\tilde{\mathbf{S}}_1$ according to Lemma 11 on p. 153, and, thus, for a given trace maximized by the respective matrix proportional to the identity [6]. In (B.100), we introduce the matrix $\tilde{\mathbf{H}}_2 = \tilde{\mathbf{V}}_1^H\mathbf{H}_2$, which fulfills

$$\mathsf{E}\left[\mathrm{Tr}\left[\tilde{\mathbf{H}}_2\tilde{\mathbf{H}}_2^H\right]\right] = \frac{1}{n_{\mathcal{R}}}\mathsf{E}\left[\mathrm{Tr}\left[\mathbf{H}_2\mathbf{H}_2^H\right]\right] = \mathsf{E}\left[\mathrm{Tr}\left[\mathbf{H}_1\mathbf{H}_1^H\right]\right].$$

We finally obtain by induction:

$$\mathsf{E}\left[R_L^{\mathrm{AF}}\right] \le \mathsf{E}\Bigg[\log\det\Bigg(\mathbf{I}_{n_{\mathcal{D}}} + \frac{s_1 s_2 \alpha P_L}{\sigma_{\mathrm{w}}^2 \cdot n_{\mathcal{S}}}\tilde{\mathbf{V}}_2^H\mathbf{A}_3\tilde{\mathbf{V}}_2 \times\left((1+s_1\alpha)\cdot\mathbf{I}_{n_{\mathcal{D}}} + s_1 s_2\alpha^2\tilde{\mathbf{V}}_2^H\mathbf{B}_3\tilde{\mathbf{V}}_2\right)^{-1}\Bigg)\Bigg] \tag{B.104}$$

$$= \mathsf{E}\Bigg[\log\det\Bigg(\mathbf{I}_{n_{\mathcal{D}}} + \frac{s_1 s_2 \alpha^2 P_L}{\sigma_{\mathrm{w}}^2 \cdot n_{\mathcal{S}}}\tilde{\mathbf{H}}_3\mathbf{A}_4\tilde{\mathbf{H}}_3^H \times\left((1+s_1\alpha+s_1 s_2\alpha^2)\cdot\mathbf{I}_{n_{\mathcal{D}}} + s_1 s_2\alpha^2\tilde{\mathbf{H}}_3\mathbf{B}_4\tilde{\mathbf{H}}_3^H\right)^{-1}\Bigg)\Bigg] \tag{B.105}$$

$$\vdots$$

$$\le \mathsf{E}\Bigg[\log\det\Bigg(\mathbf{I}_{n_{\mathcal{D}}} + \frac{s_1\cdots s_{L_0}\alpha^{L_0-1}P_L}{\sigma_{\mathrm{w}}^2 \cdot n_{\mathcal{S}}}\tilde{\mathbf{V}}_{L_0}^H\mathbf{A}_{L_0+1}\tilde{\mathbf{V}}_{L_0} \times\left(\sum_{l=0}^{L_0} s_1\cdots s_l\alpha^l\cdot\mathbf{I}_{n_{\mathcal{D}}}\right)^{-1}\Bigg)\Bigg] \tag{B.106}$$

$$= \mathsf{E}\left[\log\det\left(\mathbf{I}_{n_{\mathcal{D}}} + \frac{P_L}{\sigma_{\mathrm{w}}^2\cdot n_{\mathcal{S}} n_{\mathcal{R}}^{L-L_0}}\frac{s_1\cdots s_{L_0}\alpha^{L_0-1}}{\sum_{l=0}^{L_0} s_1\cdots s_l\alpha^l}\tilde{\mathbf{V}}_{L_0}^H\mathbf{A}_{L_0+1}\tilde{\mathbf{V}}_{L_0}\right)\right] \tag{B.107}$$

$$= \mathsf{E}\Bigg[\log\det\Bigg(\mathbf{I}_{n_{\mathcal{D}}} + \frac{P_L}{\sigma_{\mathrm{w}}^2\cdot n_{\mathcal{S}} n_{\mathcal{R}}^{L-L_0}}\frac{s_1\cdots s_{L_0}\alpha^{L}}{\sum_{l=0}^{L_0} s_1\cdots s_l\alpha^l} \times\tilde{\mathbf{V}}_{L_0}^H\mathbf{H}_{L_0+1}\cdots\mathbf{H}_{L+1}\mathbf{H}_{L+1}^H\cdots\mathbf{H}_{L_0+1}^H\tilde{\mathbf{V}}_{L_0}\Bigg)\Bigg]. \tag{B.108}$$

□

Lemma 11. *Let* $\mathbf{A}$ *and* $\mathbf{B}$ *be positive definite matrices. Then, the function*

$$f(\mathbf{Q}) = \log\det\left(\mathbf{I} + \mathbf{A}\mathbf{Q}(\mathbf{I} + \mathbf{B}\mathbf{Q})^{-1}\right) \tag{B.109}$$

is concave on the set of positive semidefinite matrices $\mathbf{Q}$.

Proof of Lemma 11. Define $\mathbf{Q} = \mathbf{Q}_a + t \cdot \mathbf{Q}_b$, where $\mathbf{Q}_a$ is positive semidefinite, $\mathbf{Q}_b$ is Hermitian, and $t \in \mathbb{C}$. We proof concavity of $f(\cdot)$ in $\mathbf{Q}$ by proving that $f(\mathbf{Q}_a + t \cdot \mathbf{Q}_b)$ is convex in t for all t such that $\mathbf{Q}$ is positive semidefinite. We compute the first and second derivative of $f(\cdot)$ with respect to t as

$$\frac{\partial}{\partial t}\log\det\left(\mathbf{I} + \mathbf{A}(\mathbf{Q}_a + t\mathbf{Q}_b)(\mathbf{I} + \mathbf{B}(\mathbf{Q}_a + t\mathbf{Q}_b))^{-1}\right) \tag{B.110}$$

$$= \operatorname{Tr}\left[\left[(\mathbf{A}+\mathbf{B})(\mathbf{Q}_a + t\mathbf{Q}_b) + \mathbf{I}\right]^{-1}(\mathbf{A}+\mathbf{B})\mathbf{Q}_b\right] - \operatorname{Tr}\left[\left[\mathbf{B}(\mathbf{Q}_a + t\mathbf{Q}_b) + \mathbf{I}\right]^{-1}\mathbf{B}\mathbf{Q}_b\right] \tag{B.111}$$

and

$$\frac{\partial^2}{\partial t^2}\log\det\left(\mathbf{I} + \mathbf{A}(\mathbf{Q}_a + t\mathbf{Q}_b)(\mathbf{I} + \mathbf{B}(\mathbf{Q}_a + t\mathbf{Q}_b))^{-1}\right) \tag{B.112}$$

$$= -\operatorname{Tr}\left[\left(\left[(\mathbf{A}+\mathbf{B})(\mathbf{Q}_a + t\mathbf{Q}_b) + \mathbf{I}\right]^{-1}(\mathbf{A}+\mathbf{B})\mathbf{Q}_b\right)^2\right] + \operatorname{Tr}\left[\left(\left[\mathbf{B}(\mathbf{Q}_a + t\mathbf{Q}_b) + \mathbf{I}\right]^{-1}\mathbf{B}\mathbf{Q}_b\right)^2\right] \tag{B.113}$$

$$= -\operatorname{Tr}\left[\left(\left[\mathbf{Q}_a + t\mathbf{Q}_b + (\mathbf{A}+\mathbf{B})^{-1}\right]^{-1}\mathbf{Q}_b\right)^2\right] + \operatorname{Tr}\left[\left(\left[\mathbf{Q}_a + t\mathbf{Q}_b + \mathbf{B}^{-1}\right]^{-1}\mathbf{Q}_b\right)^2\right] \leq 0. \tag{B.114}$$

This expression is nonpositive due to [100, Lemma 2.3], which states that for a positive definite matrix $\mathbf{Z}$, a positive semidefinite matrix $\mathbf{W}$, and a Hermitian matrix $\mathbf{X}$

$$\operatorname{Tr}\left[\left(\mathbf{Z}^{-1}\mathbf{X}\right)^2\right] \geq \operatorname{Tr}\left[\left((\mathbf{Z}+\mathbf{W})^{-1}\mathbf{X}\right)^2\right]. \tag{B.115}$$

In (B.114), we identify $\mathbf{X}$ with $\mathbf{Q}_b$, $\mathbf{Z}$ with $\mathbf{Q}_a + t\mathbf{Q}_b + (\mathbf{A}+\mathbf{B})^{-1}$ and $\mathbf{W}$ with $\mathbf{B}^{-1} - (\mathbf{A}+\mathbf{B})^{-1}$. □

Proof of Lemma 10. We prove the statement by showing that the S-transforms of the asymptotic EEDs of $\mathbf{A}$ and $\mathbf{B}$ coincide. The S-transform of $\mathbf{B}$ is given by [25, Theorem 2]

$$S_{\mathbf{B}}(z) = \begin{cases} \frac{\frac{n_{\mathcal{D}}}{n_{\mathcal{S}}}}{\left(1+\frac{n_{\mathcal{D}}}{n_{\mathcal{S}}}z\right)}, \text{ if } L=1, \\ \left(\frac{1}{1+\frac{n_{\mathcal{D}}}{n_{\mathcal{R}}}z}\right)^{L-2} \frac{\frac{n_{\mathcal{D}}}{n_{\mathcal{S}}}}{\left(1+\frac{n_{\mathcal{D}}}{n_{\mathcal{S}}}z\right)\left(\frac{n_{\mathcal{D}}}{n_{\mathcal{R}}}z+1\right)}, \text{ if } L \geq 2. \end{cases} \tag{B.116}$$

We derive the S-transform of $\mathbf{A}$ in the following. First, we obtain the S-transform of $\tilde{\mathbf{A}} \triangleq \frac{1}{n_{\mathcal{R}}^L}\mathbf{H}_2\mathbf{H}_3\cdots\mathbf{H}_{L+1}\mathbf{H}_{L+1}^H\cdots\mathbf{H}_3^H\mathbf{H}_2^H$ from [25, Theorem 2] as

$$S_{\tilde{\mathbf{A}}}(z) = \left(\frac{1}{1+z}\right)^{L-1} \frac{1}{\frac{n_{\mathcal{S}}}{n_{\mathcal{R}}}+z}. \tag{B.117}$$

Using that the S-transform of $\frac{1}{n_{\mathcal{D}}}\tilde{\mathbf{H}}_1^H\tilde{\mathbf{H}}_1$ is given by $S_{\frac{1}{n_{\mathcal{D}}}\tilde{\mathbf{H}}_1^H\tilde{\mathbf{H}}_1}(z) = \frac{z+1}{1+n_{\mathcal{R}}/n_{\mathcal{D}}z}$ (e.g. [26]) we obtain the S-transform of $\tilde{\tilde{\mathbf{A}}} \triangleq \frac{1}{n_{\mathcal{R}}^L n_{\mathcal{D}}}\mathbf{H}_2\mathbf{H}_3\cdots\mathbf{H}_{L+1}\mathbf{H}_{L+1}^H\cdots\mathbf{H}_3^H\mathbf{H}_2^H\tilde{\mathbf{H}}_1^H\tilde{\mathbf{H}}_1$ as [25, Theorem 1]

$$S_{\tilde{\tilde{\mathbf{A}}}}(z) = S_{\frac{1}{n_{\mathcal{D}}}\tilde{\mathbf{H}}_1^H\tilde{\mathbf{H}}_1}(z)S_{\tilde{\mathbf{A}}}(z) = \frac{z+1}{1+\frac{n_{\mathcal{R}}}{n_{\mathcal{D}}}z}\left(\frac{1}{1+z}\right)^{L-1}\frac{1}{\frac{n_{\mathcal{S}}}{n_{\mathcal{R}}}+z}. \tag{B.118}$$

Finally, we obtain the S-transform of the asymptotic EED of $\mathbf{A}$ as [25, Eq. (15)] as

$$S_{\mathbf{A}}(z) = \frac{z+1}{z+\frac{n_{\mathcal{R}}}{n_{\mathcal{D}}}}S_{\tilde{\tilde{\mathbf{A}}}}\left(\frac{n_{\mathcal{D}}}{n_{\mathcal{R}}}z\right) = \begin{cases} \frac{\frac{n_{\mathcal{D}}}{n_{\mathcal{S}}}}{1+\frac{n_{\mathcal{D}}}{n_{\mathcal{S}}}z}, \text{ if } L=1, \\ \left(\frac{1}{1+\frac{n_{\mathcal{D}}}{n_{\mathcal{R}}}z}\right)^{L-2} \frac{\frac{n_{\mathcal{D}}}{n_{\mathcal{S}}}}{\left(1+\frac{n_{\mathcal{D}}}{n_{\mathcal{S}}}z\right)\left(\frac{n_{\mathcal{D}}}{n_{\mathcal{R}}}z+1\right)}, \text{ if } L \geq 2, \end{cases} \tag{B.119}$$

which coincides with (B.116). □

B.3. Proof of Proposition 5

The sum-transmit power of stage $\mathcal{R}_l$ is given by

$$\begin{aligned} \sum_{k=1}^{n_{\mathcal{R}}} P_{\mathrm{R}_k^{(l)}} =& \frac{P_L}{n_{\mathcal{S}}}\frac{\alpha^{L-l+1}}{n_{\mathcal{R}}^{L-l+1}}\mathrm{Tr}\left[\mathbf{H}_{l+1}\cdots\mathbf{H}_{L+1}\mathbf{H}_{L+1}^H\cdots\mathbf{H}_{l+1}^H\right] \\ &+ \sigma_{\mathrm{w}}^2\cdot\frac{\alpha}{n_{\mathcal{R}}}\cdot\left(1+\sum_{l'=l}^{L-1}\frac{\alpha^{l'-l}}{n_{\mathcal{R}}^{l'-l}}\mathrm{Tr}\left[\mathbf{H}_{l+1}\cdots\mathbf{H}_{l'+1}\mathbf{H}_{l'+1}^H\cdots\mathbf{H}_{l+1}^H\right]\right). \end{aligned} \tag{B.120}$$

Each of the traces converges to one almost surely. This follows by repeated application of the following result from [98, 99]: Let $\mathbf{Y} = \frac{1}{m}\mathbf{X}\mathbf{R}\mathbf{X}^H$ be a random matrix, where

- $\mathbf{R}$ is an arbitrary nonnegative definite $m \times m$ random matrix whose EED converges uniformly to a nonrandom distribution almost surely as $m \to \infty$ and satisfies $\lim_{m\to\infty} m^{-1}\mathrm{Tr}\left[\mathbf{R}\right] = 1$ almost surely,
- $\mathbf{X}$ is an $p\times m$ random matrix, independent of $\mathbf{R}$, with i.i.d. elements of zero-mean and unit variance.

Then, for every fixed ratio m/p, $\lim_{p\to\infty} \frac{1}{p}\mathrm{Tr}\left[\mathbf{Y}\right] = 1$ almost surely.

Due to the convergence of the traces, the following holds:

$$\lim_{n_\mathcal{R}\to\infty} \sum_{k=1}^{n_\mathcal{R}} P_{\mathrm{R}_k^{(l)}} = P_L\alpha^{L-l+1} + \sigma_\mathrm{w}^2 \cdot \sum_{l'=l}^{L-1} \alpha^{L-l} \tag{B.121}$$

$$= P_L\alpha^{L-l+1} + \sigma_\mathrm{w}^2 \cdot \alpha\frac{1-\alpha^{L-l+1}}{1-\alpha} \text{ almost surely,} \tag{B.122}$$

where the right hand side fulfills for $\alpha = P_L/(P_L + \sigma_\mathrm{w}^2)$

$$P_L\alpha^{L-l+1} + \sigma_\mathrm{w}^2 \cdot \alpha\frac{1-\alpha^{L-l+1}}{1-\alpha} = P_L \text{ for all } l \in \{1,\ldots,L\}. \tag{B.123}$$

For the second part, the transmit power of relay $\mathrm{R}_k^{(l)}$ is written as

$$P_{\mathrm{R}_k^{(l)}} = \frac{\alpha}{n_\mathcal{R}} \mathbf{h}_{l+1}^{(k)}{}^H \mathbf{T}_{l+1}\mathbf{h}_{l+1}^{(k)}, \tag{B.124}$$

where $\mathbf{h}_k^{(l)}$ denotes the kth column of $\mathbf{H}_l$ and

$$\mathbf{T}_{l+1} = \begin{cases} \frac{P_L}{n_\mathcal{S}}\mathbf{I}_{n_\mathcal{R}}, \text{ if } l = L, \\ \frac{P_L}{n_\mathcal{S}}\frac{\alpha^{L-l}}{n_\mathcal{R}^{L-l}}\mathbf{H}_{l+2}\cdots\mathbf{H}_{L+1}\mathbf{H}_{L+1}^H\cdots\mathbf{H}_{l+2}^H \\ \qquad +\sigma_\mathrm{w}^2\cdot\left(\mathbf{I}_{n_\mathcal{R}} + \sum_{l'=l}^{L-2}\frac{\alpha^{l'-l+1}}{n_\mathcal{R}^{l'-l+1}}\mathbf{H}_{l+2}\cdots\mathbf{H}_{l'+1}\mathbf{H}_{l'+1}^H\cdots\mathbf{H}_{l+2}^H\right), \\ \qquad \text{if } l \in \{1,\ldots,L-1\}. \end{cases}$$

First, a result from [101, Theorem 1] yields, that for every $k \in \{1,\ldots,n_\mathcal{R}\}$

$$\lim_{n_\mathcal{R}\to\infty} P_{\mathrm{R}_k^{(l)}} = \lim_{n_\mathcal{R}\to\infty} \frac{\alpha}{n_\mathcal{R}}\mathsf{E}\left[\mathbf{h}_{l+1}^{(k)}{}^H\mathbf{T}_{l+1}\mathbf{h}_{l+1}^{(k)}\right] = P_L\alpha^{L-l+1} + \sigma_\mathrm{w}^2 \cdot \sum_{l'=l}^{L-1}\alpha^{L-l'+1} \tag{B.125}$$

$$= P_L\alpha^{L-l+1} + \sigma_{\mathrm{w}}^2 \cdot \alpha \frac{1-\alpha^{L-l+1}}{1-\alpha} \text{ almost surely,} \tag{B.126}$$

where the right hand side fulfills

$$P_L\alpha^{L-l+1} + \sigma_{\mathrm{w}}^2 \cdot \alpha \frac{1-\alpha^{L-l+1}}{1-\alpha} = P_L \text{ for all } l \in \{1,\dots,L\} \tag{B.127}$$

for $\alpha = P_L/(P_L+\sigma_{\mathrm{w}}^2)$. Ref. [101, Theorem 1] requires the $\mathbf{T}_{l+1}$ to be positive semi-definite (obvious) and to fulfill almost surely $\lim_{n_{\mathcal{R}}\to\infty} \frac{1}{n_{\mathcal{R}}}\mathrm{Tr}[\mathbf{T}_{l+1}] < \infty$ (follows by the almost sure convergence of the traces in $\mathbf{T}_{l+1}$).

Thus, there exists for every relay $\mathrm{R}_k^{(l)}$ an $n_0^{(k)}$, such that for all $n_{\mathcal{R}} \geq n_0^{(k)}$

$$\left| n_{\mathcal{R}} P_{\mathrm{R}_k^{(l)}} - P_L \right| < \varepsilon. \tag{B.128}$$

Next, consider for each relay $\mathrm{R}_k^{(l)}$ the smallest such $n_0^{(k)}$ as a random variable. The distribution of these i.i.d. random variables fulfills for all $k \in \{1,\dots,n_{\mathcal{R}}\}$

$$\lim_{n_{\mathcal{R}}\to\infty} \mathsf{Pr}\left[n_0^{(k)} < n_{\mathcal{R}} \right] = 1. \tag{B.129}$$

Now, fix for γ arbitrarily close to one, $n_{\mathcal{R}}^{(0)}$ sufficiently large, such that $\mathsf{Pr}\left[n_0^{(k)} < n_{\mathcal{R}}^{(0)}\right] > \gamma$. Then, for all $n_{\mathcal{R}} > n_{\mathcal{R}}^{(0)}$

$$\mathsf{Pr}\left[\frac{1}{n_{\mathcal{R}}} \sum_{k=1}^{n_{\mathcal{R}}} 1\left\{ n_0^{(k)} < n_{\mathcal{R}} \right\} > \gamma \right] > \mathsf{Pr}\left[\frac{1}{n_{\mathcal{R}}} \sum_{k=1}^{n_{\mathcal{R}}} 1\left\{ n_0^{(k)} < n_{\mathcal{R}}^{(0)} \right\} > \gamma \right]. \tag{B.130}$$

Consider now the inequality for all $n_{\mathcal{R}} > n_{\mathcal{R}}^{(0)}$

$$\frac{1}{n_{\mathcal{R}}} \sum_{k=1}^{n_{\mathcal{R}}} 1\left\{ n_0^{(k)} < n_{\mathcal{R}} \right\} > \frac{1}{n_{\mathcal{R}}} \sum_{k=1}^{n_{\mathcal{R}}} 1\left\{ n_0^{(k)} < n_{\mathcal{R}}^{(0)} \right\}, \tag{B.131}$$

where the lower-bound fulfills due to (B.130) and the strong law of large numbers

$$\lim_{n_{\mathcal{R}}\to\infty} \frac{1}{n_{\mathcal{R}}} \sum_{k=1}^{n_{\mathcal{R}}} 1\left\{ n_0^{(k)} < n_{\mathcal{R}}^{(0)} \right\} > \gamma \text{ almost surely.} \tag{B.132}$$

This establishes the proposition. □

Acronyms

AF	amplify & forward.
AWGN	additive white Gaussian noise.
CDF	cumulative distribution function.
CDMA	code division multiple access.
CF	compress & forward.
CSCG	circularly symmetric complex Gaussian.
CSI	channel state information.
CSMA	carrier sense multiple access.
DCF	distributed coordination function.
DF	decode & forward.
DMT	diversity multiplexing tradeoff.
EDGE	Enhanced Data Rates for Global Evolution.
EED	empirical eigenvalue distribution.
FDMA	frequency division multiple access.
GPRS	Gleneral Packet Radio Service.
GSM	Global System for Mobile Communications.
HSDPA	High Speed Downlink Packet Access.
HSPA	High Speed Packet Access.

i.i.d.	identically and independently distributed.
IO	input-output.
KKT	Karush-Kuhn-Tucker.
LTE	Long Term Evolution.
MAC	medium access control.
MIMO	multiple-input multiple-output.
MMSE	minimum mean square error.
OFDM	orthorgonal frequency division multiplexing.
PDF	probability density function.
PMF	probability mass function.
QF	quantize & forward.
SIC	successive interference cancellation.
SINR	signal-to-interference-plus-noise ratio.
SISO	single-input single-output.
SNR	signal-to-noise ratio.
TDMA	time division multiple access.
UMTS	Universal Mobile Telecommunications System.
VLSI	very-large-scale integration.
WLAN	wireless local area network.

Notation

$\mathbf{a}$	vector $\mathbf{a}$.
a, A	scalars a and A.
$(a)^N$	sequence $(a_1, \ldots, a_N)$.
$[a, b)$	interval from a (including a) to b (excluding b).
$\mathbf{A}$	matrix $\mathbf{A}$.
$(a_{ij})_{i=1,\ldots m, j=1,\ldots n}$	$m \times n$ matrix $\mathbf{A}$.
$(\mathbf{a}_1, \ldots, \mathbf{a}_n)$	$m \times n$ matrix with vector $\mathbf{a}_i$ in column i.
$\mathcal{A}$	set $\mathcal{A}$.
a	random variable a.
a	random vector a.
A	random matrix A.
$\mathrm{Diag}(\mathbf{a})$	diagonal matrix with the vector $\mathbf{a}$ on its diagonal.
$\mathrm{diag}(\mathbf{A})$	vector that contains the diagonal of matrix $\mathbf{A}$.
$\\|\mathbf{a}\\|_p$	p-norm of vector $\mathbf{a}$.
$\mathrm{rk}\{\mathbf{A}\}$	rank of matrix $\mathbf{A}$.
$\\|\mathbf{A}\\|_*$	nuclear norm of matrix $\mathbf{A}$.
$\\|\mathbf{A}\\|_p$	induced p-norm of matrix $\mathbf{A}$.
$\mathrm{Tr}[\mathbf{A}]$	trace of matrix $\mathbf{A}$.
$\det(\mathbf{A})$	determinant of matrix $\mathbf{A}$.
$\lambda_i\{\mathbf{A}\}$	ith ordered (decreasing) eigenvalue of matrix $\mathbf{A}$.
$\lambda_i\{\mathbf{A}, \mathbf{B}\}$	ith ordered (decreasing) generalized eigenvalue of matrices $\mathbf{A}$ and $\mathbf{B}$.
$\sigma_i\{\mathbf{A}\}$	ith ordered (decreasing) singular value of matrix $\mathbf{A}$.
$\mathbf{A}^T$	transpose of matrix $\mathbf{A}$.

$\mathbf{A}^H$	Hermitian transpose of matrix $\mathbf{A}$.
$\mathbf{A}^*$	element-wise complex conjugate of matrix $\mathbf{A}$.
$\Re\{\mathbf{A}\}$	real part of matrix $\mathbf{A}$.
$\Im\{\mathbf{A}\}$	imaginary part of matrix $\mathbf{A}$.
$\mathbf{A} \prec \mathbf{B}$	$\mathbf{B} - \mathbf{A}$ is positive definite.
$\mathbf{A} \preceq \mathbf{B}$	$\mathbf{B} - \mathbf{A}$ is positive semi-definite.
$\mathbf{A} \succ \mathbf{B}$	$\mathbf{A} - \mathbf{B}$ is positive definite.
$\mathbf{A} \succeq \mathbf{B}$	$\mathbf{A} - \mathbf{B}$ is positive semi-definite.
$\mathbf{A} \odot \mathbf{B}$	element-wise (Hadamard) product of matrices $\mathbf{A}$ and $\mathbf{B}$.
$\mathbf{I}_n$	$n \times n$ identity matrix.
$\mathbf{e}_i$	ith unit vector.
$\mathbf{0}_{m \times n}$	$m \times n$ all zero matrix.
π	Archimedes' constant 3.14....
e	Euler's constant 2.71....
$\imath$	imaginary unit $\sqrt{-1}$.
$\mathcal{A}^+$	the set of positive elements in the set $\mathcal{A}$.
$\mathcal{A} \times \mathcal{B}$	Cartesian product of sets $\mathcal{A}$ and $\mathcal{B}$.
$\mathcal{A}^n$	nth Cartesian power of set $\mathcal{A}$.
$\mathcal{A}^C$	complement of set $\mathcal{A}$.
$\mathcal{A} \cup \mathcal{B}$	union of sets $\mathcal{A}$ and $\mathcal{B}$.
$\lvert\mathcal{A}\rvert$	cardinality of set $\mathcal{A}$.
$\mathcal{A} \cap \mathcal{B}$	intersection of sets $\mathcal{A}$ and $\mathcal{B}$.
$\mathcal{A} \backslash \mathcal{B}$	the set of elements in $\mathcal{A}$ that are not contained in $\mathcal{B}$.
$\mathrm{a} \in \mathcal{A}$	a is an element of $\mathcal{A}$.
$\mathrm{a} \notin \mathcal{A}$	a is not an element of $\mathcal{A}$.
$\mathbb{N}$	the set of natural numbers.
$\mathbb{Q}$	the set of rational numbers.
$\mathbb{R}$	the set of real numbers.
$\mathbb{C}$	the set of complex numbers.

$\mathbb{P}^n$	projective space corresponding to $\mathbb{C}^{n+1}$.
$\exists$	exists.
$\forall$	for all.
$f_{\mathsf{x}}(\cdot)$	joint probability density function of the random vector x.
$F_{\mathsf{x}}(\cdot)$	joint cumulative distribution function of the random vector x.
$\mathrm{P}[\mathrm{A}]$	probability of event A.
$\mathrm{P}[\mathrm{A}\vert\mathrm{B}]$	probability of event A given even B.
$\mathrm{E}_{\mathsf{x}}[\cdot]$	expectation with respect to the random vector x.
$\mathsf{x} \sim \mathcal{CN}(\boldsymbol{\mu}, \mathbf{K})$	random vector x is circularly symmetric complex Gaussian distributed with mean vector $\boldsymbol{\mu}$ and covariance matrix $\mathbf{K}$.
$h(\mathsf{x})$	differential entropy of random vector x.
$h(\mathsf{x}\vert\mathsf{y})$	conditional differential entropy of random vector x given the random vector y.
$I(\mathsf{x};\mathsf{y})$	mutual information between random vectors x and y.
$I(\mathsf{x};\mathsf{y}\vert\mathsf{z})$	conditional mutual information between random vectors x and y given the random vector z.
$\mathsf{x} \to \mathsf{y} \to \mathsf{z}$	the random vectors x, y and z form a Markov chain (in this order).
$f : \mathcal{A} \to \mathcal{B} : x \to y$	function $y = f(x)$ with domain $\mathcal{A}$ and co-domain $\mathcal{B}$.
$f \circ g$	composition of function f and g.
max	maximum of a function or set of values.
min	minimum of a function or set of values.
argmax	maximizing argument of a function.
argmin	minimizing argument of a function.
$x \propto y$	x is proportional to y.
$(\cdot)^+$	$\max(0, \cdot)$.
$1\{a < b\}$	indicator function, 1, if $a < b$ is true, 0 else.
$\mathrm{Log}(\cdot)$	complex logarithm to the base e.
$\log(\cdot)$	real logarithm to the base 2 unless stated explicitly.
$\delta(\cdot)$	Dirac delta function.
$\delta[\cdot]$	Kronecker delta function.

$\sigma(\cdot)$	Heaviside step function.
$\lfloor \cdot \rfloor$	floor function.
$\lceil \cdot \rceil$	ceiling function.
$a \bmod b$	a modulo b.
$\mathcal{O}, o, \Omega, \omega, \Theta$	Landau symbols.
$\partial/\partial a$	partial derivative with respect to a.
$\boldsymbol{\nabla}_{\mathbf{a}}$	gradient vector with respect to vector argument $\mathbf{a}$.
$\boldsymbol{\nabla}^2_{\mathbf{a}}$	Hessian matrix with respect to vector argument $\mathbf{a}$.
$\mathcal{L}$	Lagrangian function.
$F^{(n)}_{\mathsf{A}}(\cdot)$	empirical eigenvalue distribution of the Hermitian random matrix A.
$G^{(n)}_{\mathsf{A}}(\cdot)$	Stieltjes transform of empirical eigenvalue distribution of the Hermitian random matrix A.
$S^{(n)}_{\mathsf{A}}(\cdot)$	S-transform of empirical eigenvalue distribution of the Hermitian random matrix A.
S_k	kth source node.
D_k	kth destination node.
$\mathrm{R}^{(l)}_k$	kth relay node in stage l.
$\mathcal{S}$	set of source nodes.
$\mathcal{D}$	set of destination nodes.
$\mathcal{R}^{(l)}$	set of relay nodes in stage l.
L	number of hops.
$n_{\mathcal{S}}$	number of source nodes.
$n_{\mathcal{D}}$	number of destination nodes.
$n^{(l)}_{\mathcal{R}}$	number of relay nodes in $\mathcal{R}^{(l)}$.
n	number of sources and destinations when $n_{\mathcal{S}} = n_{\mathcal{D}}$.
$n_{\mathcal{R}}$	number of relay nodes per relay stage when $n^{(1)}_{\mathcal{R}} = \ldots = n^{(L)}_{\mathcal{R}}$.

Bibliography

[1] A. Mishra, *Fundamentals of cellular network planning and optimization.* John Wiley and Sons, LTd, 2004.

[2] *Annual report.* Switzerland: Swisscom AG, 2009.

[3] C. Perkins, *Ad Hoc Networking.* Addison Wesley Professional, Dec. 2000.

[4] M. Clark, *Wireless Access Networks: Fixed Wireless Access and WLL Networks-Design and Operation.* John Wiley and Sons, Oct. 2001.

[5] D. Gesbert and J. Akhtar, "Breaking the barriers of Shannon's capacity: an overview of MIMO wireless systems," 2002.

[6] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, Nov. 1999.

[7] G. J. Foschini and M. J. Gans, "On Limits of Wireless Communications in a Fading Environment When Using Multiple Antennas," *Wireless Personal Communications*, vol. 6, no. 3, pp. 311–335, Mar. 1998.

[8] D. Tse and P. Viswanath, *Fundamentals of wireless communication.* Cambridge University Press, 2005.

[9] V. Cadambe and S. Jafar, "Interference alignment and degrees of freedom of the K-user interference channel," *IEEE Transactions on Information Theory*, vol. 54, no. 8, pp. 3425–3441, Aug. 2008.

[10] G. Kramer, I. Maric, and R. Yates, *Cooperative communications.* now Publishers Inc., 2006.

[11] O. Simeone, O. Somekh, H. V. Poor, and S. Shamai, "Local base station cooperation via finite-capacity links for the uplink of linear cellular networks," *IEEE Transactions on Information Theory*, vol. 55, no. 1, pp. 190–204, 2009.

[12] ——, "Downlink multicell processing with limited-backhaul capacity," *EURASIP Journal on Advances in Signal Processing — Multiuser MIMO Transmission with Limited Feedback, Cooperation, and Coordination*, 2009.

[13] "Further advancements for EUTRA physical layer aspects," 3GPP, Tech. Rep. TR-36.814, V9.0.0, Mar. 2010.

[14] A. Özgür, O. Lévêque, and D. Tse, "Hierarchical cooperation achieves optimal capacity scaling in ad hoc networks," *IEEE Transactions on Information Theory*, vol. 53, no. 10, pp. 3549–3572, 2007.

[15] P. Gupta and P. R. Kumar, "Towards an information theory of large networks: an achievable rate region," *IEEE Transactions on Information Theory*, vol. 49, no. 8, pp. 1877–1894, Aug. 2003.

[16] ——, "The capacity of wireless networks," *IEEE Transactions on Information Theory*, vol. 46, no. 2, pp. 388–404, Mar. 2000.

[17] N. Laneman and G. Wornell, "Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2415–2425, Oct. 2003.

[18] N. Laneman, D. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Transactions on Information Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[19] C. E. Shannon, "A mathematical theory of communication," *Bell System Techn. J.*, vol. 27, pp. 379–423 and 623–656, July and Oct. 1948.

[20] T. Cover and J. A. Thomas, *Elements of Information Theory.* John Wiley & Sons, 1991.

[21] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo-codes," in *Proc. International Conference on Communications*, 1993, pp. 1064–1070.

[22] A. Host-Madsen and A. Nosratinia, "The multiplexing gain of wireless networks," in *Proc. IEEE International Symposium on Information Theory*, Adelaide, SA, September 2005, pp. 2065–2069.

[23] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: A fundamental trade-off in multiple-antenna channels," *IEEE Transactions on Information Theory*, vol. 49, pp. 1073–1096, 2003.

[24] S. Yeh and O. Lévêque, "Asymptotic capacity of multi-level amplify-and-forward relay networks," in *Proc. IEEE International Symposium on Information Theory*, Nice, France, June 2007.

[25] R. R. Müller, "On the asymptotic eigenvalue distribution of concatenated vector-valued fading channels," *IEEE Transactions on Information Theory*, vol. 48, no. 7, pp. 2086–2091, July 2002.

[26] A. M. Tulino and S. Verdú, *Random matrix theory and wireless communications*. Now Publishers Inc., 2004.

[27] Y. Birk and T. Kol, "Informed-source coding-on-demand (ISCOD) over broadcast channels," in *IEEE Infocomm*, vol. 3, San Francisco, Apr. 1998, pp. 1257 –1264.

[28] "Signaling over MIMO multi-base systems — combination of multi-access and broadcast schemes," in *Proc. IEEE International Symposium on Information Theory*, June 2006.

[29] "Degrees of freedom on the MIMO X channel — optimality of the MMK scheme tech," Tech. Rep., Sep. 2006, available on arXiv:cs.IT/0607099v2.

[30] A. Motahari, M. Gharan, S.O. Maddah-Ali, and A. Khandani, "Forming pseudo-MIMO by embedding infinite rational dimensions along a single real line: Removing barriers in achieving the DOFs of single antenna systems," Tech. Rep., 2009, coRR, vol. abs/0908.2282.

[31] B. Nazer, M. Gastpar, S. Jafar, and S. Vishwanath, "Ergodic interference alignment," in *Proc. IEEE International Symposium on Information Theory*, Seoul, Korea, June/July 2009.

[32] B. Rankov and A. Wittneben, "Spectral efficient signaling for half-duplex relay channels," in *Asilomar Conference on Signals, Systems, and Computers 2005*, Nov. 2005.

[33] V. Cadambe and S. 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.

[34] V. R. Cadambe, S. A. Jafar, and C. Wang, "Interference alignment with asymmetric complex signaling: settling the host-madsen-nosratinia conjecture," *IEEE Transactions on Information Theory*, vol. 56, pp. 4552–4565, September 2010.

[35] A. Wittneben and B. Rankov, "Distributed antenna systems and linear relaying for Gigabit MIMO wireless," in *Proc. IEEE Vehicular Technology Conference*, Los Angeles, LA, Sept. 2004, pp. 3624–3630.

[36] S.-W. Jeon, S.-Y. Chung, and S. Jafar, "Degrees of freedom of multi-source relay networks," in *Proc. Allerton Conference on Communication, Control, and Computing*, Sept. 2009.

[37] S. Jeon and S.-Y. Chung, "Capacity of a class of multi-source relay networks," *IEEE Transactions on Information Theory*, under review for publication — preprint available on arxiv.org.

[38] S.-W. Jeon and S.-Y. Chung, "Capacity of a class of multi-source relay networks," in *in in Information Theory and Applications Workshop*, San Diego, CA, Feb. 2009.

[39] C. Akçaba, P. Kuppinger, and H. Bölcskei, "Distributed transmit diversity in relay networks," in *IEEE Information Theory Workshop*, Jul. 2007, pp. 233–237.

[40] E. Akuiyibo and O. Lévêque, "Diversity-multiplexing tradeoff for the slow fading interference channel," pp. 140–Ű143, Mar. 2008.

[41] C. Akcaba and H. Bölcskei, "Diversity-multiplexing tradeoff in fading interference channels," *IEEE Transactions on Information Theory*, May 2009.

[42] S. Yang and J.-C. Belfiore, "Diversity of MIMO multihop relay channels," submitted to IEEE Transactions on Information Theory.

[43] R. Vaze and R. W. Heath, "To code in space and time or not in multihop relay channels," *Signal Processing, IEEE Transactions on*, vol. 57, no. 7, pp. 2736–2747, July 2009.

[44] ——, "Maximizing reliability in multi-hop wireless networks," in *Proc. IEEE International Symposium on Information Theory*, Toronto, ON, July 2008.

[45] K. Sreeram, S. Birenjith, and P. V. Kumar, "DMT of multi-hop cooperative networks-part II: Half-duplex networks with full-duplex performance," 2008. [Online]. Available: http://www.citebase.org/abstract?id=oai:arXiv.org:0808.0235

[46] C. Esli, J. Wagner, and A. Wittneben, "Distributed gradient based gain allocation for coherent multiuser AF relaying networks," in *Proc. International Conference on Communications*, June 2009.

[47] ——, "Distributed gradient methods for relay gain allocation in coherent AF relaying networks," under review for publication in IEEE Transactions on Wireless Communications.

[48] R. Rolny, J. Wagner, and W. A., "Distributed gain allocation in non-regenerative multiuser multihop MIMO networks," in *Proc. Asilomar Conf. Signals, Systems and Computers*, Pacific Grove, CA, Nov. 2010.

[49] S. Verdu and T. S. Han, "A general formula for channel capacity," *IEEE Transactions on Information Theory*, July 1994.

[50] D. N. Bernstein, "The number of roots of a system of equations," *Functional Analysis and its Applications*, vol. 9, pp. 1–4, 1975.

[51] B. Sturmfels, *Solving systems of polynomial equations.* Oxford University Press, 2002.

[52] D. Cox, J. Little, and O. D., *Ideals, Varieties, and Algorithms*, 3rd ed. Springer, 2007.

[53] I. Gelfand, M. Kapranov, and A. Zelevinsky, *Discriminants, Resultants, and Multidimensional Determinants.* Birkhäuser, 1994.

[54] A. G. Kushnirenko, "Newton polytopes and the bezout theorem," *Functional Analysis and its Applications*, vol. 10, pp. 233–235, 1976.

[55] A. McLennan, "The expected number of real roots of a multihomogeneous system of polynomial equations," 2001.

[56] ——, "The maximal number of real roots of in a multihomogeneous system of polynomial equations," *Beiträge zur Algebra und Geometrie*, vol. 50, no. 2, pp. 343–350, 1999.

[57] M. Fink, "Time reversal of ultrasonic fields — part i: Basic principles," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 39, no. 5, Sept. 1992.

[58] C. Oestges, A. Kim, G. Papanicolaou, and A. Paulraj, "Characterization of space-time focusing in time-reversed random fields," *IEEE Trans. on Ant. and Prop.*, vol. 53, no. 1, pp. 283–293, Jan. 2005.

[59] K. Kumar, G. Caire, and A. Moustakas, "The diversity-multiplexing tradeoff of linear MIMO receivers," in *Proc. IEEE Information Theory Workshop*, Sep. 2007, pp. 487–492.

[60] H. Sampath, P. Stoica, and A. Paulraj, "Generalized linear precoder and decoder design for MIMO channels using the weighted MMSE criterion," *IEEE Trans. Commun*, vol. 49, pp. 2198–2206, 2001.

[61] P. Viswanath and D. Tse, "Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality," *IEEE Transactions on Information Theory*, vol. 49, no. 8, pp. 1912–1921, Aug. 2003.

[62] V. Dem'yanov and V. Malozemov, "On the theory of non-linear minimax problems," *Russ. Math. Surv.*, pp. 26–57, 1971.

[63] W. Karush, *Minima of Functions of Several Variables with Inequalities as Side Constraints.* M.Sc. Dissertation. Dept. of Mathematics, Univ. of Chicago, Chicago, Illinois., 1939.

[64] H. W. Kuhn and A. W. Tucker, "Nonlinear programming," in *2nd Berkeley Symposium*, Berkeley: University of California Press, 1951, pp. 481–492.

[65] S. Boyd and L. E. Ghaoui, "Method of centers for minimizing generalized eigenvalues," in *Linear Algebra Appl*, 1993, pp. 63–111.

[66] S. Boyd and L. Vandenberghe, *Convex Optimization.* Cambridge University Press, 2004.

[67] M. Varanasi and T. Guess, "Optimum decision feedback multiuser equalization with successive decoding achieves the total capacity of the Gaussian multiple-access channel," in *Proc. of Asilomar Conference on Signals, Systems and Computers*, vol. 2, Nov 1997, pp. 1405–1409.

[68] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources." *IEEE Transactions on Information Theory*, pp. 471–480, 1973.

[69] V. I. Morgenshtern and H. Bölcskei, "Crystallization in large wireless networks," *IEEE Transactions on Information Theory, Special Issue on Relaying and Cooperation in Communication Networks*, vol. 53, no. 10, pp. 3319–3349, Oct. 2007.

[70] ——, "Random matrix analysis of large relay networks," in *Proc. Allerton Conference on Communication, Control, and Computing*, Monticello, IL, Sept. 2006.

[71] N. Fawaz, K. Zarifi, M. Debbah, and D. Gesbert, "Asymptotic capacity and optimal precoding in MIMO multi-hop relay networks," *CoRR*, vol. abs/0906.0037, 2009.

[72] S. Borade, L. Zheng, and R. Gallager, "Amplify and forward in wireless relay networks: Rate, diversity and network size," *IEEE Transactions on Information Theory*, vol. 53, no. 10, pp. 3302–3318, Oct. 2007.

[73] D. Gündüz, M. Khojastepour, A. Goldsmith, and H. Poor, "Multi-hop mimo relay networks: diversity-multiplexing trade-off analysis," *IEEE Transactions on Wireless Communications*, vol. 9, no. 5, pp. 1738–1747, Bay 2010.

[74] A. Sanderovich, S. Shamai (Shitz), Y. Steinberg, and M. Peleg, "Decentralized receiver in a MIMO system," in *Proc. IEEE International Symposium on Information Theory*, Seattle, WA, July 2006, pp. 6–10.

[75] V. A. Marcenko and L. A. Pastur, "Distributions of eigenvalues for some sets of random matrices," *Math. USSR-Sbornik*, vol. 1, pp. 457–483, 1967.

[76] S. Verdu and S. Shamai (Shitz), "Spectral efficiency of CDMA with random spreading," *IEEE Transactions on Information Theory*, vol. 48, pp. 3117–3128, Dec. 2002.

[77] J. Geronimo and T. Hill, "Necessary and sufficient condition that the limit of stieltjes transforms is a stieltjes transform," *Journal on Approximation Theory*, vol. 121, no. 1, Mar. 2003.

[78] Z. D. Bai, "Methodologies in spectral analysis of large dimensional random matrices," *Statistica Sinica*, vol. 9, pp. 611–677, 1999.

[79] Z. Bai and J. Silverstein, "No eigenvalues outside the support of the limiting spectral distribution of large dimensional sample covariance matrices," *IEEE Transactions on Information Theory*, vol. 26, pp. 316–345, 1998.

[80] D. Voiculescu, "Multiplication of certain non-commuting random variables," *Journal of Operator Theory*, vol. 18, pp. 223–235, 1987.

[81] R. R. Müller, "Random matrices, free probability and the replica method," in *Proc. of European Signal Processing Conference (Eusipco)*, Vienna, Austria, Sept. 2004.

[82] B. Wang and F. Zhang, "Some inequalities for the eigenvalues of the product of positive semidefinite hermitian matrices," *Linear Algebra Appl.*, vol. 160, pp. 113–118, 1992.

[83] D. Lang, *Routing protocols for mobile ad hoc networks.* VDM, 2008.

[84] C. Toh, *Ad Hoc Mobile Wireless Networks — Protocols and Systems.* Prentice-Hall, 2002.

[85] B. Rankov and A. Wittneben, "Spectral efficient protocols for half-duplex fading relay channels," in *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 2, 2007, pp. 379–389.

[86] A. Ettefagh, M. Kuhn, C. Esli, and A. Wittneben, "Performance of a cluster-based MAC protocol in multiuser MIMO wireless LANs," in *Proc. International ITG Workshop on Smart Antennas*, Bremen, Germany, Feb. 2010.

[87] R. Lin and A. Petropulu, "A new wireless network medium access protocol based on cooperation," *IEEE Transactions on Signal Processing*, vol. 53, no. 12, pp. 4675–4684, Dec. 2005.

[88] ——, "A new wireless network medium access protocol based on cooperation," *IEEE Transactions on Signal Processing*, vol. 53, no. 12, pp. 4675–4684, Dec. 2005.

[89] T. Tang, M. Park, J. Heath, R.W., and S. Nettles, "A joint MIMO-OFDM transceiver and MAC design for mobile ad hoc networking," in *International Workshop on Wireless Ad-Hoc Networks*, May-June 2004, pp. 315–319.

[90] D. Wang and U. Tureli, "Joint MIMO-OFDM and MAC design for broadband multihop ad hoc networks," *EURASIP Journal on Wireless Communications and Networking*, pp. 1–9, 2006.

[91] P. Liu, Z. Tao, S. Narayanan, T. Korakis, and S. S. Panwar, "Coopmac: A cooperative MAC for wireless LANs," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 2, pp. 340–354, Feb. 2007.

[92] P. Fertl, A. Hottinen, and G. Matz, "Perturbation-based distributed beamforming for wireless relay networks," in *Proc. IEEE Global Telecommunications Conference*, New Orleans, LA, Dec. 2008, pp. 1–5.

[93] S. Berger and A. Wittneben, "When do non-regenerative two-hop relaying networks require a global phase reference?" in *Proc. IEEE Global Communications Conference*, Honolulu, Hawaii, Dec. 2009.

[94] Y.-S. Tu and G. Pottie, "Coherent cooperative transmission from multiple adjacent antennas to a distant stationary antenna through AWGN channels," in *Proc. IEEE Vehicular Technology Conference*, Birmingham, Al, May 2002, pp. 130–134.

[95] S. Berger and A. Wittneben, "Carrier phase synchronization of multiple distributed nodes in a wireless network," in *Proc. IEEE Workshop on Signal Processing Advances for Wireless Communications*, Helsinki, Finland, June 2007.

[96] Y.-C. Liang, G. Pan, and Z. Bai, "Asymptotic performance of MMSE receivers for large systems using random matrix theory," *IEEE Transactions on Information Theory*, vol. 53, no. 11, pp. 4173–4190, Nov. 2007.

[97] R. Gordon, "A convergence theorem for the Riemann integral," *Mathematics Magazine*, vol. 73, no. 2, April 2000.

[98] Q. Y. Yin, "Limiting spectral distribution for a class of random matrices," *J. of Multivariate Analysis*, vol. 29, pp. 50–68, 1986.

[99] J. Silverstein, "Strong convergence of the empirical distribution of eigenvalues of large dimensional random matrices," *Journal of Multivariate Analysis*, vol. 55, pp. 331–339, Nov. 1995.

[100] E.-V. Belmega, L. Samson, and M. Debbah, "A trace inequality for positive definite matrices," *Journal of inequalities in pure and applied mathematics*, vol. 10, p. 4pp, 2009.

[101] D. E. Varberg, "Almost sure convergence of quadratic forms in independent random variable," *The Annals of Mathematical Statistics*, vol. 39, no. 5, pp. 1502–1506, 1968.

Curriculum Vitae

Name:	Jörg Wagner
Date of Birth:	October 5, 1980
Place of Birth:	Bad Mergentheim, Germany

Education

11/2006-01/2011 **ETH Zurich, Switzerland**
PhD studies at the Communication Technology Laboratory, Department of Information Technology and Electrical Engineering.

10/2001-10/2006 **ETH Zurich, Switzerland**
Bachelor and Master studies in Information Technology and Electrical Engineering. Degree: Master of Science.

Experience

11/2006-01/2011 **ETH Zurich, Switzerland**
Research assistant at the Communication Technology Laboratory headed by Prof. Dr. Armin Wittneben.

- Teaching and presentation experience
- Project work in industry projects with
  - Swisscom (Schweiz) AG
  - Nokia Research Center Helsinki
- Supervision of Master theses and student projects
- Reviewer of international conference and journal submissions

09/2005-03/2006 **Agency for Science, Technology & Research, Singapore**
Traineeship in Institute for Infocomm Research under supervision of Dr. Ying-Chang Liang. Research work on opportunistic beamforming techniques for cellular downlink transmissions.

Award

06/2006 **Fritz-Kutter Award 2006**
For thesis "Analysis of Dynamics in Mobile Ad-hoc Networks".

Publications

- **Book Chapter**
  - *Distributed Antenna Systems and Linear Relaying for Rank-Deficient MIMO Systems*; B. Rankov, J. Wagner, and A. Wittneben, Chapter in Distributed Antenna Systems: Open Architecture for Future Wireless Communications, Auerbach Publications, CRC Press, 2006, (invited paper).
- **Journal Papers**
  - *On Capacity Scaling of Multi-Antenna Multi-Hop Networks: The Significance of the Relaying Strategy in the "Long Network Limit"*; J. Wagner and A. Wittneben, under review for publication in IEEE Transactions on Information Theory.
  - *Large N Analysis of Amplify-and-Forward MIMO Relay Channels with Correlated Rayleigh Fading*; J. Wagner, B. Rankov, and A. Wittneben, IEEE Transactions on Information Theory, Vol. 54, No. 12, pp. 5735-5746, 2008.
  - *An Empirical Analysis of the Impact of Mobility on Link Failures in an 802.11 Ad Hoc Network*; V. Lenders, J. Wagner, S. Heimlicher, and B. Plattner, IEEE Wireless Communications Magazine, Dec. 2008.
  - *On the Balance of Multiuser Diversity and Spatial Multiplexing Gain in Random Beamforming*; J. Wagner, Y.-C. Liang, and R. Zhang, IEEE Transactions on Wireless Communications, Vol. 7, No. 7, pp. 2512-2525, July 2008.
- **Conference, Symposium, and Workshop Papers**
  - *Distributed Gain Matrix Optimization in Non-Regenerative MIMO Relay Networks*; R. Rolny, J. Wagner, C. Esli, and A. Wittneben, Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, USA, Nov. 2009.
  - *On the Diversity-Multiplexing Tradeoff of Multiuser Ampfliy & Forward Multihop Networks*; J. Wagner and A. Wittneben, Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, Nov. 2009.
  - *Distributed Gradient Based Gain Allocation for Coherent Multiuser AF Relaying*; Networks C. Esli, J. Wagner, and A. Wittneben, IEEE International Conference on Communications (ICC 2009) , June 2009.
  - *On Capacity Scaling of (Long) MIMO Amplify-and-Forward Multi-Hop Networks*; J. Wagner and A. Wittneben, IEEE Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA, Oct. 2008.
  - *Multihop-Enabled Orthogonalization in Distributed MIMO Networks*; J. Wagner, M. Kuhn, and A. Wittneben, IEEE Personal, Indoor and Mobile Radio Conference (PIMRC), Cannes, France, Sept. 2008, (invited paper).

- *Capacity Scaling of (Long) Non-Regenerative MIMO Multi-Hop Channels*; J. Wagner and A. Wittneben, IEEE International Symposium on Information Theory, Toronto, ON, Canada, July 2008.
- *Cooperative Processing for the WLAN Uplink*; M. Kuhn, J. Wagner, and A. Wittneben, IEEE Wireless Communications and Networking Conference, WCNC 2008, Las Vegas, USA, Mar. 2008.
- *On the Asymptotic Capacity of the Rayleigh Fading Amplify-and-Forward MIMO Relay Channel*; J. Wagner, B. Rankov, and A. Wittneben, IEEE International Symposium on Information Theory, Nice, France, June 2007.
- *Random Beamforming with Systematic Beam Selection*; J. Wagner, Y.-C. Liang, and R. Zhang, The 17th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, Helsinki, Finland, Sept. 2006.
- *Measurements from an 802.11b Mobile Ad Hoc Network*; V. Lenders, J. Wagner, and M. May, IEEE WoWMoM Workshop on Advanced Experimental Activities on Wireless Networks and Systems, Niagara-Falls/Buffalo, New York, USA, June 2006.
- *Analyzing the Impact of Mobility in Ad Hoc Networks*; V. Lenders, J. Wagner, and M. May, ACM/Sigmobile Workshop on Multi-hop Ad Hoc Networks: from theory to reality, Florence, Italy, May 2006.