*Forschungsberichte aus dem
Institut für Nachrichtentechnik des
Karlsruher Instituts für Technologie*

Tobias Renk

Cooperative Communications: Network Design and Incremental Relaying

Band 24

Druck: Engelhardt und Bauer Druck- und Verlagsgesellschaft mbH
Käppelestraße 10, 76131 Karlsruhe

ISSN: 1433-3821

Forschungsberichte aus dem Institut für Nachrichtentechnik des Karlsruher Instituts für Technologie

---

Herausgeber: Prof. Dr. rer. nat. Friedrich Jondral

Band 1 Marcel Kohl
Simulationsmodelle für die Bewertung von Satellitenübertragungsstrecken im 20/30 GHz Bereich

Band 2 Christoph Delfs
Zeit-Frequenz-Signalanalyse: Lineare und quadratische Verfahren sowie vergleichende Untersuchungen zur Klassifikation von Klaviertönen

Band 3 Gunnar Wetzker
Maximum-Likelihood Akquisition von Direct Sequence Spread-Spectrum Signalen

Band 4 Anne Wiesler
Parametergesteuertes Software Radio für Mobilfunksysteme

Band 5 Karl Lütjen
Systeme und Verfahren für strukturelle Musteranalysen mit Produktionsnetzen

Band 6 Ralf Machauer
Multicode-Detektion im UMTS

Band 7 Gunther M. A. Sessler
Schnell konvergierender Polynomial Expansion Multiuser Detektor mit niedriger Komplexität

Band 8 Henrik Schober
Breitbandige OFDM Funkübertragung bei hohen Teilnehmergeschwindigkeiten

Band 9 Arnd-Ragnar Rhiemeier
Modulares Software Defined Radio

Forschungsberichte aus dem Institut für Nachrichtentechnik des Karlsruher Instituts für Technologie

---

Herausgeber: Prof. Dr. rer. nat. Friedrich Jondral

Band 10 Mustafa Mengüç Öner
Air Interface Identification for Software Radio Systems

Band 11 Fatih Çapar
Dynamische Spektrumverwaltung und elektronische Echtzeitvermarktung von Funkspektren in Hotspotnetzen

Band 12 Ihan Martoyo
Frequency Domain Equalization in CDMA Detection

Band 13 Timo Weiß
OFDM-basiertes Spectrum Pooling

Band 14 Wojciech Kuropatwiński-Kaiser
MIMO-Demonstrator basierend auf GSM-Komponenten

Band 15 Piotr Rykaczewski
Quadraturempfänger für Software Defined Radios: Kompensation von Gleichlauffehlern

Band 16 Michael Eisenacher
Optimierung von Ultra-Wideband-Signalen (UWB)

Band 17 Clemens Klöck
Auction-based Medium Access Control

Band 18 Martin Henkel
Architektur eines DRM-Empfängers und Basisbandalgorithmen zur Frequenzakquisition und Kanalschätzung

Band 19 Stefan Edinger
Mehrträgerverfahren mit dynamisch-adaptiver Modulation zur unterbrechungsfreien Datenübertragung in Störfällen

Forschungsberichte aus dem Institut für Nachrichtentechnik des Karlsruher Instituts für Technologie

---

Herausgeber: Prof. Dr. rer. nat. Friedrich Jondral

Band 20 Volker Blaschke
Multiband Cognitive Radio-Systeme

Band 21 Ulrich Berthold
Dynamic Spectrum Access using OFDM-based Overlay Systems

Band 22 Sinja Brandes
Suppression of Mutual Interference in OFDM-based Overlay Systems

Band 23 Christian Körner
Cognitive Radio – Kanalsegmentierung und Schätzung von Periodizitäten

Band 24 Tobias Renk
Cooperative Communications: Network Design and Incremental Relaying

Vorwort des Herausgebers

In der Mobilkommunikation wird in naher Zukunft mit einem explosionsartigen Anwachsen des Datenaufkommens gerechnet. Damit werden Untersuchungen aller verfügbaren Methoden, die eine bessere Nutzung der bekanntlich stark eingeschränkt verfügbaren Ressource Frequenz versprechen, notwendig. Dazu gehört die Einführung von Multiple Input Multiple Output (MIMO) Strategien genauso wie der Einsatz von Cognitive Radios (CRs) in Overlaysystemen oder die Nutzung der Ultra Wide Band (UWB) Technik. Eine weitere, schon lange bekannte Methode, besteht darin, Relais einzusetzen[1]. Sie ist in letzter Zeit erneut unter dem Namen Kooperative Kommunikation stark in das Interesse der wissenschaftlichen Diskussion gerückt.

Praktische Anwendungen finden Relais heute in der Satellitenkommunikation, im Richtfunk und in Sonderanwendungen. In zellularen Systemen werden kooperative Systeme bisher nicht eingesetzt, da sie hier zu einem stark erhöhten Signalisierungsaufkommen führen. Allerdings zeichnen sich zurzeit Anwendungen ab, die eine tiefer gehende Beschäftigung mit Prinzipien der Kooperativen Kommunikation sinnvoll erscheinen lassen. Sollen sich z.B. in einer Produktionshalle für den Materialtransport genutzte Plattformen vollständig autonom bewegen können, muss jede Plattform über eine Kommunikationskomponente verfügen, die sowohl eine Kommunikation von jeder Plattform zu jeder anderen als auch die Verbindung jeder Plattform mit dem Steuerrechner gestattet. Hier ist Kooperative Kommunikation der von den Plattformen getragenen Endgeräte untereinander gefragt, z.B. wenn eine Plattform von der direkten Verbindung mit dem Steuerrechner abgeschattet ist.

Mit seiner Dissertation *Cooperative Communications: Network Design and Incremental Relaying* liefert Tobias Renk interessante Beiträge zur Diskussion über den Einsatz von Relais. Er hat dabei folgende Beiträge zum Fortschritt von Wissenschaft und Technik geleistet:

[1]Siehe z.B. Arthur C. Clarke: Extra-terrestrial Relays – Can Rocket Stations Give World-wide Radio Coverage? Wireless World, October 1945, p. 305

- Entwicklung und Verifikation des Adaptive Relay Selection Protocols (ARSPs)
- Beschreibung einer Methode zur optimalen Leistungs- und Zeitallokation in Relaisnetzen
- Beschreibung eines neuartigen, bezüglich der Kombinationsstrategie hybriden, Diversitätsempfängers
- Grundlegende Untersuchungen zur theoretischen Beschreibung und zur Performanceabgrenzung des Incremental Relaying

Karlsruhe, im Juli 2010
Friedrich Jondral

Cooperative Communications: Network Design and Incremental Relaying

Zur Erlangung des akademischen Grades eines

DOKTOR-INGENIEURS

der Fakultät für
Elektrotechnik und Informationstechnik
des Karlsruher Instituts für Technologie

genehmigte

DISSERTATION

von

Dipl.-Ing. (BA) Tobias Renk

aus

Kronach

Tag der mündlichen Prüfung: 15. Juli 2010
Hauptreferent: Prof. Dr. rer. nat. Friedrich K. Jondral
Korreferent: Prof. Dr.-Ing. Frank H. P. Fitzek

In rebus multo plus quam ea, quae primo
aspectu praeterientes cognoscere possumus,
est. Artium honestarum est hoc investigare.

To the loving memory of my grandmother.
She always put others first in her life and
showed me that one can have a rich and
fulfilled life by making others happy. How
proud would she be today, how proud!

Acknowledgment

There'll be lots of hard times along the way.

Rude Bones

My supervisor Prof. Dr. rer. nat. Friedrich K. Jondral deserves many thanks. He guided me in many ways, gave me the opportunity to do this thesis and the freedom to pursue my own ideas. His numerous efforts and dedication kept me focused during my time at his lab. Being more than just an inspiration from an academic point of view, he always aims at shaping his students personally and prepare them for whatever will come. I am sure that I will benefit from all this in my future. His mentoring has left a permanent mark!

In addition, I would like to thank Prof. Dr.-Ing. Frank Fitzek for serving on my thesis committee and improving the final version of this work. This dissertation has definitely benefited from his comments and suggestions. Moreover, I thank Prof. Dr.-Ing. Thomas Zwick, Prof. Dr. rer. nat. Wilhelm Storck, and Prof. Dr. rer. nat. Olaf Dössel for serving on my thesis committee.

I am very grateful to Prof. Andrea Goldsmith for giving me the opportunity to come to the Wireless Systems Laboratory at Stanford as a visiting researcher. I very much enjoyed my time at Stanford and made a lot of new friends. This thesis benefited from her great knowledge and I remember having numerous valuable discussions on my research topic. I had a fantastic time at Stanford. And yeah – thanks for the shirt!

During the course of a dissertation, you meet a lot of new people. And sometimes you know that they will be with you for the rest of your life. From an academic perspective, these are certainly Dr.-Ing. Holger Jäkel and Deniz Gündüz. Holger always had a calm way of discussing my research and he brought up a lot of new ideas.

Moreover, I will never forget him as a great sportsmate. He was also proof-reading a draft version of this dissertation and gave very helpful and valuable comments. I met Deniz during my time at Stanford and once visited him at Princeton University. His great knowledge and creativity have been very enriching to me. Chapter 6 has definitely benefited from his contributions.

Of many colleagues, I deeply thank Dipl.-Ing. Maximilian Hauske for proof-reading this thesis and improving it with numerous helpful comments. With Dipl.-Ing. Dennis Burgkhardt and Dipl.-Ing. Stefan Nagel I shared wonderful times at and outside the lab. I also enjoyed having a lot of non-technical discussions with those guys. And remember, you two are the next in line! In addition, thanks also go to my former roommate Dr.-Ing. Clemens Klöck and my former colleague Dr.-Ing. Volker Blaschke. They already got it done. If there is one guy at the lab who is passionate about sports, it is Dipl.-Ing. (FH) Reiner Linnenkohl. I thank him for the time we shared in the gym and the music he shared with me. Excellent taste!

My parents sent me – against my will – to the Gymnasium and this step still influences my life every day. They were right at that time – little did I know then. They showed me to believe in higher education and to accept new challenges with enthusiasm and courage. Even if they could not understand every step I made in the past, they supported me with endless love and pride. My gratitude cannot be expressed in words.

Having an older brother is the best motivation I can ever think of. And the most beneficial thing about this is that he does things first – and you can try to be better. I do not know if he actually likes this confession, but that is just the way it was (and probably still is).

Finally, my most sincere thanks and heartfelt gratitude go to my girlfriend Nina who loves dancing, her dog Alma, and me. Thank you for supporting me and us in a lot of ways. The story we are writing together is more wonderful than I could ever dream of!

Zusammenfassung

Mobile Kommunikation ist längst ein fester Bestandteil unseres täglichen Lebens geworden. Dabei hat die Verwendung mobiler Endgeräte sämtliche Bereiche unseres Lebens ergriffen. Der integrierte Wecker unseres Mobilgerätes weckt uns am Morgen, auf dem Weg zur Arbeit hören wir Musik, die auf unserem Mobilgerät gespeichert ist, wir verwenden das Mobilgerät als Restaurantführer, Kalender, Straßenkarte und vieles mehr. Die Vision der modernen Mobilkommunikation lässt sich demnach klar formulieren: Wir möchten auf drahtlose Dienste zurückgreifen können wann immer wir wollen und wo immer wir sind. Diese Entwicklung führt zu einem rasant ansteigenden Datenvolumen und es ist fraglich, ob heutige zellulare Kommunikationssysteme die notwendigen Datenraten bei entsprechender Übertragungsqualität liefern können. Ein weiterer wichtiger Aspekt in diesem Zusammenhang ist ein Ansteigen der Nutzerdichte in Kommunikationsnetzen. Sicherlich wird der erhöhte Bandbreitebedarf in naher Zukunft dazu führen, dass die derzeitige Belegung spektraler Ressourcen überdacht werden muss, da die strikte Regulierung zu einer ineffizienten Nutzung geführt hat.

Eine Möglichkeit, dem hohen Bedarf an Datenrate und den anspruchsvollen Anforderungen zukünftiger Technologien zu begegnen, ist die kooperative Kommunikation verschiedener Teilnehmer. Dabei ist der Begriff der Kooperation in der Kommunikationstechnik so alt wie die Kommunikation selbst. Man denke beispielsweise an die Rauchzeichen der Indianer oder die Festungsanlagen entlang der Grenzgebiete des Imperium Romanum. In diesem Fall hatten beide Beispiele vorrangig den Zweck, Informationen über eine längere Wegstrecke übertragen zu können – ein weiterer Vorteil der kooperativen Kommunikation. Die vorliegende Arbeit bietet eine umfassende Bearbeitung der Thematik "Kooperative Kommunikation", angefangen von der Auswahl geeigneter Relays über optimale Ressourcenvergabe bis hin zu einer neuartigen hybriden Empfängerstruktur.

Einer der wichtigsten Aspekte, wenn man von einer Kooperation unter mobilen Teilnehmern ausgeht, ist die Auswahl der Relays. Ausgehend von dieser Fragestellung wird in dieser Arbeit ein Protokoll entwickelt, welches die Relays auf intelligente Art und Weise auswählt. Grundgedanke hierbei ist, dass sich die Anzahl ausgewählter Relays an der geforderten Leistungsfähigkeit des Systems orientiert. Als Kriterium für die Leistungsfähigkeit wird die Fehlerwahrscheinlichkeit des Endteilnehmers herangezogen. Die Auswahl der Relays basiert auf den Kanaleigenschaften zwischen den Teilnehmern, der verbleibenden Batterieleistung pro Teilnehmer, der Bereitschaft zur Kooperation sowie der Bewegungsrichtung der einzelnen Teilnehmer. Jeder Teilnehmer bestimmt seine jeweiligen Parameter und sendet diese an eine Zentraleinheit. Diese ermittelt unter Beachtung der empfangenen Parameter sowohl die Anzahl als auch die Übertragungsreihenfolge der ausgewählten Relays. Die Leistungsfähigkeit des Protokolls wird an Hand simulativer Untersuchungen evaluiert.

Enorme Performancegewinne lassen sich durch eine optimale Zuteilung der Übertragungsressourcen erzielen. Dabei liegt in dieser Arbeit das Hauptaugenmerk auf einer optimalen Allokation der Ressourcen Leistung und Zeit. Ein Optimierungsalgorithmus basierend auf der sog. Brent-Methode wird vorgestellt, um einen Kompromiss zwischen der Konvergenzgeschwindigkeit und der Zuverlässigkeit verschiedener Algorithmen zu finden. Als Beurteilungskriterien werden die momentane und die verzögerungskritische Kapazität untersucht. Es zeigt sich, dass relaybasierte Netze vor allem dann große Kapazitätsgewinne gegenüber einem einzelnen Kommunikationspaar erzielen, wenn die gesamte zur Verfügung stehende Systemleistung gering ist. Des Weiteren hat die Lage eines Relays einen großen Einfluss auf die Leistungsfähigkeit. Die Untersuchungen ermöglichen es, die optimale Lage eines Relays zu bestimmen. Eine Ausweitung der Ergebnisse auf zellulare Netze, in denen zusätzliche feste Relays installiert werden sollen, ist möglich. Dabei sind die Erkenntnisse besonders bei der infrastrukturellen Planung von Bedeutung, um etwaige Installationskosten zu minimieren und die Performance zu maximieren.

Im Anschluss daran wird eine neuartige hybride Empfängerstruktur vorgestellt. Hierbei steht bei der Entwicklung der Kompromiss zwischen effizienter Implementierung und Leistungsfähigkeit im Mittelpunkt. Aus diesem Grund wird auf Kombinierungsverfahren zurückgegriffen, die keine Schätzung der Kanalkoeffizienten benötigen. Als Kriterium für die Leistungsfähigkeit wird erneut die Fehlerwahrscheinlichkeit des Empfängers herangezogen. Um sich den ständig ändernden Ausbreitungsbedingungen des Kommunikationsmediums anzupassen, schaltet der Empfänger adaptiv zwischen zwei Kombinierungsverfahren um. Eine statische Festlegung der Schaltschwelle ist dabei nicht mehr möglich. Vielmehr muss sich die Schaltschwelle den Ausbreitungsbedingungen dynamisch anpassen. Dies führt zur Entwicklung eines leistungsfähigen Algorithmus, bei dem die Schaltschwelle basierend auf Messungen der Signalleistungen der unterschiedlichen Empfangspfade bestimmt wird. Die Leistungsfähigkeit des hybriden Empfängers wird mathematisch beschrieben und analysiert.

Der Nachteil der ursprünglichen Kooperationsstrategien ist eine ineffiziente Nutzung der Freiheitsgrade des Übertragungskanals. Diese kann gesteigert werden, wenn die Relays nur dann an der Kommunikation teilnehmen, nachdem sie vom Endteilnehmer dazu aufgefordert wurden. Für diese Art von Kooperationsprotokollen hat sich der Begriff Incremental Relaying eingebürgert. Der hierfür notwendige Feedback-Kanal vom Endteilnehmer zu den restlichen Teilnehmern wird in dieser Arbeit zunächst als fehlerfrei angenommen. Es werden die Ausfallkapazitäten unter Verwendung zweier unterschiedlicher Kooperationsprotokolle ermittelt: Einerseits dekodiert das Relay die Nachricht der Quelle vollständig, andererseits verstärkt es lediglich die pulsartig übertragene Quellennachricht. Daran anschließend werden der Feedback-Kanal als symmetrischer Binärkanal modelliert und die Ausfallkapazitäten für die beschriebenen Kooperationsprotokolle hergeleitet und analysiert. Die gefundenen Ergebnisse werden auf Netze erweitert, die aus einer Vielzahl von Relays bestehen, um dem Aspekt einer ansteigenden Nutzerdichte in Kommunikationsnetzen gerecht zu werden. Für den Fall eines unzuverlässigen Feedback-Kanals wird eine Übertragungsstrategie vorgestellt, die auf Grund mathematischer Untersuchungen entwickelt werden konnte.

Abstract

Mobile communications has already become an essential element of our daily life. In doing so, the usage of mobile terminals has practically influenced all parts of it. The integrated alarm bell of our mobile phone wakes us in the morning, on the way to work we listen to music that is stored on our mobile phone, we use our mobile phone as restaurant guide, organizer, street map, and many more. As a consequence, the vision of modern mobile communications can be stated explicitly: We want to have ubiquitous access to wireless services whenever we want and wherever we are. This development leads to a rapidly increasing demand for data rate, and it is questionable if today's cellular communications systems are able to provide the required data rates at sufficiently high quality of service. Another important aspect in this context is an increasing number of users in communications networks. In the near future, the demand for more bandwidth will clearly yield a rethinking of the current allocation of spectral resources, since the strict regulations have caused an inefficient usage of spectrum.

One possibility in order to meet the high demand for data rates and the challenging requirements of future technologies is cooperative communications amongst users. Noteworthy, the concept of cooperation in communications is as old as communications itself. Take smoke signal of Indians, for example, or watchtowers along the borders of the Imperium Romanum. Both cases had the purpose of transmitting information over large distances – which is another advantage of cooperative communications. This thesis provides a comprehensive treatment of "cooperative communications," starting from relay selection over optimal resource allocation to the development of a novel hybrid receiver structure.

One of the most important aspects of cooperation amongst mobile terminals is relay selection. On the basis of this question we develop a protocol in this thesis, which intelligently selects relays. The main idea is that the amount of selected relays is adapted to the required performance of the system. The performance criterion is the

error probability at the destination. Relay selection is based on parameters including channel gains between users, remaining battery power, willingness to cooperate, and direction of movement of users. Each user determines its parameters and transmits them to a central unit. The central unit then identifies the amount of selected relays and the sequence of transmission. The performance of this protocol is evaluated by simulations.

Enormous performance gains can be achieved by optimal resource allocation. In this thesis, the main focus is the optimal allocation of power and transmission time. An optimization algorithm based on Brent's method is presented in order to balance the aspects of speed of convergence and reliability. Performance criteria are the instantaneous and the delay-limited capacity. It is shown that relay-based networks are especially beneficial over a single communications pair for low overall system powers. Moreover, the relay location has a great impact on the system performance and the optimal relay location is determined. An extension of the results to cellular networks, in which additional fixed relays are installed, is possible. In particular, the results are of special importance for infrastructural planning in order to minimize installation costs and to maximize system performance.

Furthermore, a novel hybrid receiver structure is introduced. The trade-off between efficient implementation and performance is one of the main issues. For that reason, combining strategies are used that do not require an estimation of the channel gains. Again, the error probability at the destination is applied as performance criterion. The receiver switches adaptively between two combining strategies in order to adjust to the varying propagation conditions. It is not possible to use a fixed threshold anymore. In contrast to that, the threshold must adapt dynamically to the propagation conditions. This yields the development of an efficient algorithm, where the threshold is determined on the basis of signal-to-noise ratio measurements. The performance of the hybrid receiver structure is described mathematically and properly analyzed.

The disadvantage of the original cooperation strategies is an inefficient use of the degrees of freedom of the channel. This problem can be overcome if the relays do not always participate actively in the communications process, but rather aid communications after having received a request from the destination. These kinds of protocols are called incremental relaying. The feedback channel from the destination to all other users is first considered to be perfect in this thesis. Outage capacities of two different cooperation strategies are determined: First, the relay has to fully decode the source signal, second, the relay only amplifies the bursty source transmission. In addition to that, the feedback channel is considered to be imperfect and is modeled as a binary symmetric channel. Again, outage capacities for the two mentioned cooperation strategies are determined. The results are extended to networks with an arbitrary number of relays in order to deal with the aspect of an increasing number of users in communications networks. In case of unreliable feedback, a transmission strategy is presented based on mathematical investigations.

Contents

1

Introduction

None is so great that he needs no help, and none is so small that he cannot give it.

King Solomon

There is a clear vision of wireless communications for the next decades. People use portable devices that carry multimedia traffic anywhere and at any time. One witness of that vision clearly is the huge success of mobile telephones and the services they offer (e.g., short message service or multimedia messaging service). Due to the increasing number of multimedia applications, there is great demand for higher and higher data rates – especially in the area of wireless communications. And, indeed, wireless communications is the most promising means to break this communications frontier. This aspect was highlighted in a statement by Mischa Schwartz back in 1999 [1]: "I don't like to predict things, because I am always wrong! I can't tell what's going to happen. But clearly, the Internet is driving the show now. That and wireless. [...] I find wireless networking [...] one of the major engineering challenges now."

However, the wireless channel is one of the most challenging propagation channels. In addition to path loss and shadowing, e.g., due to buildings, there are many scattered rays arriving at the receiver because of multipath propagation and each ray suffers different impacts. Destructive and/or constructive superposition of these rays leads to multipath signal fading that causes a great fluctuation of the received signal strength. Therefore, an improvement of the propagation quality is indispensable to meet the requirements for new services and seamless access. An approach to fight multipath propagation efficiently is the creation of diversity. On the one hand, diversity can be created by multiple-input multiple-output (MIMO) systems [2–4]. However, due to size, costs, and hardware limitations it is frequently not possible to incorporate numerous antenna elements into a mobile terminal. One alternative, on the other hand, is cooperation among mobile users. Usage for mobile cellular networks is imaginable, but today plenty of research results indicate that the application area of cooperation – at least for the short term – lies in the field of sensor and ad-hoc networks [5, 6]. The basic idea of cooperation is that many mobile terminals, which are equipped with one antenna each, for instance, pool their resources in order to create a virtual antenna array to exploit the advantages of MIMO techniques [7, 8]. Hence, this kind of cooperation is sometimes called virtual MIMO [9, 10].

Cooperation among mobile users has a lot of advantages. First, coverage area can be increased, which has been the original intention in building up cooperative networks even thousands of years ago (see [11] and Section 2.2 for more information on that topic). Another very important aspect, especially for big cities, is the possibility to have connection to the network even if one is currently located between two skyscrapers. Generally, diversity leads to a reduction of error and outage probability and to energy savings. The last point is of great importance in wireless networks where mobile terminals only possess a limited amount of energy. For instance, consider transmission from a source to a destination, where one relay aids communication. Then, it is not necessary that both source as well as relay transmit with all their available power in order to achieve a certain performance. These energy savings eventually lead to an increased battery lifetime [12].

Though relay networks and user cooperation have received enormous interest recently, they are neither a new phenomenon nor limited to the field of wireless communications. They also have applications in the field of transatlantic cable-laying, computer networking, directional radio, satellite communications as well as sensor, ad-hoc, and mesh networking. Consequently, user cooperation is not only part of academic life, but also has great impact on practical realizations. The Swedish company TerraNet offers free calls over "meshed" mobile phones without a regular network [13]. These "mesh"-phones possess integrated Voice-over-IP-clients, which allow to forward up to seven external phones calls while one's own telephone is in use. In order to achieve an adequate coverage and maintenance in buildings, a frequency band below 1 GHz is used. The project has started in South America in 2008 and in Europe in 2009.

1.1 Motivation

What is now the motivation for relaying and user cooperation? In this section, we show briefly the advantages of relaying and give some simple scenario examples where cooperation clearly is beneficial.

Consider a scenario where one is using his cellular telephone but the connection to the base station is weak. Why not using resources of an adjacently located phone in order to improve the quality of the call and prevent an outage? The aid of a nearby

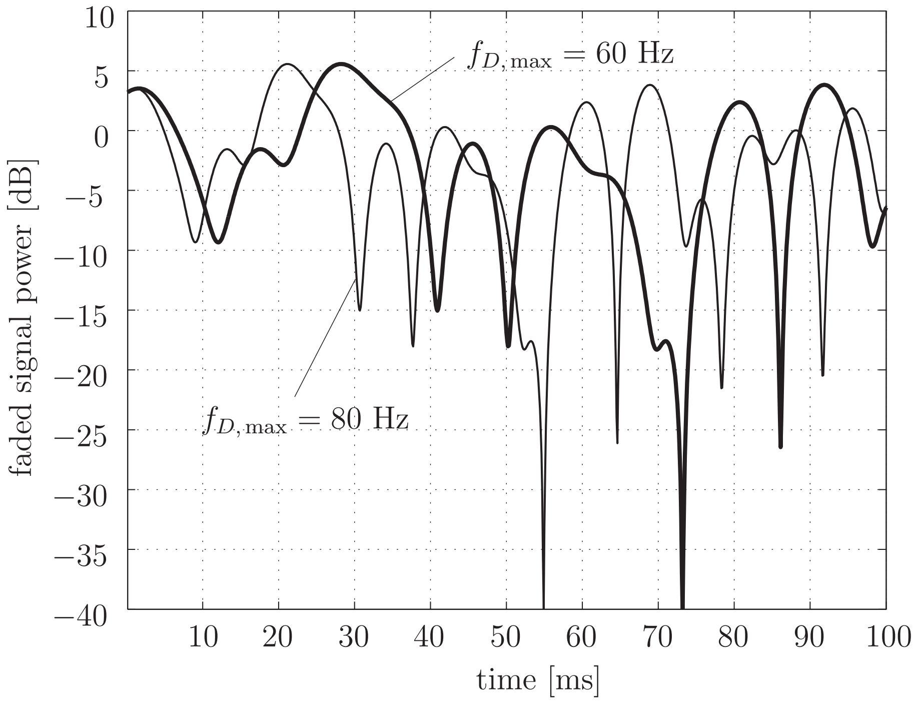

Figure 1.1: Faded signal power [dB] of a signal that is transmitted over two different Rayleigh fading channels with a maximum Doppler shift of $f_{D,\max} = 80$ Hz (thin line) and $f_{D,\max} = 60$ Hz (thick line), respectively. Sampling time equals 100 μs.

mobile surely can increase reliability to a certain extent. In a similar fashion one could profit from cooperation while walking down to a subway station. If the connection to the base station gets weaker and weaker, nearby users with mobile phones can help keeping up the connection. However, why should someone allow another person to "waste" his resources? This is the underlying philosophy of cooperation, which can be found in nature in several different occurrences. For instance, a special kind of vampire bats (desmondus rotundus) shows a highly developed social behavior – they share food among each other. Due to this technique of cooperation, the actual mortality rate is approximately 24 %. Without this kind of cooperation, the mortality rate would be up to 82 % [14]. A hungry bat gains up to 18 hours to starvation by this donation of food, whereas the donator bat loses only 6 hours. This kind of cooperation is known as reciprocal altruism.

The technical background of the underlying idea is illustrated in Fig. 1.1 and Fig. 1.2. Fig. 1.1 shows the faded signal power over time of a signal transmitted over two different Rayleigh[1] fading channels (see Subsection 2.1.1). The average power

[1] John William Strutt, 3rd Baron Rayleigh, ∗ November 12, 1842, † June 30, 1919. English physicist. Earned the Nobel Prize for Physics in 1904 with William Ramsay for the discovery of the element argon.

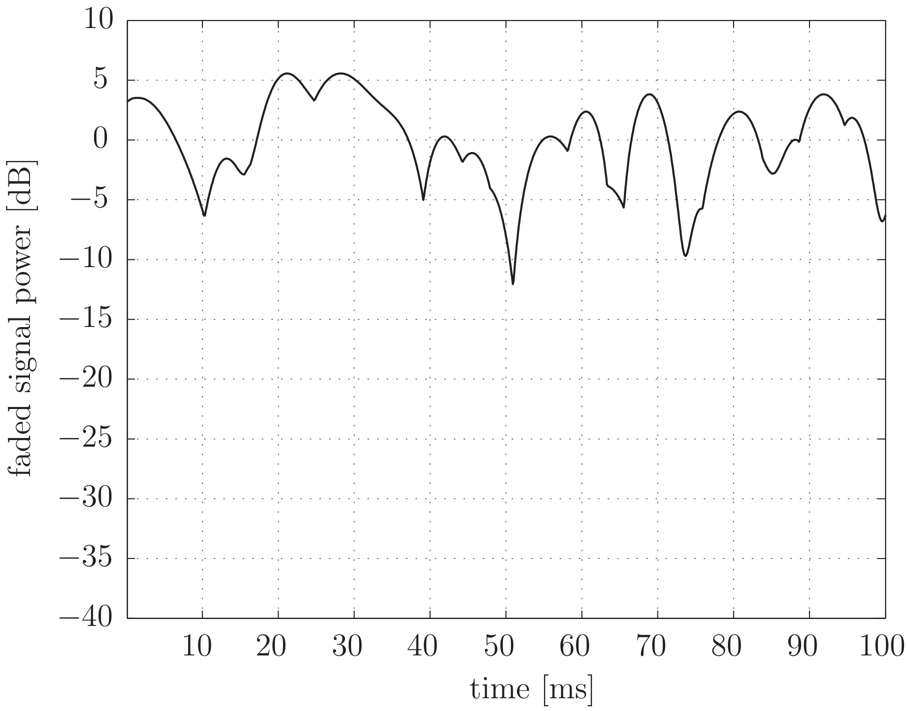

Figure 1.2: Faded signal power [dB] of a signal that is transmitted over the two Rayleigh fading channels shown in Fig. 1.1. Only the best among the two paths is selected to demonstrate the benefit of diversity.

value is normalized to 0 dB. The first channel is characterized by a maximum Doppler frequency shift of $f_{D,\max} = 80$ Hz, the second one by a maximum Doppler frequency shift of $f_{D,\max} = 60$ Hz. We see that the signal suffers from fading and therefore inherits great fluctuations in the signal power. We even can recognize deep fades with an attenuation of down to −40 dB in both cases which makes it practically very hard to detect the signal. Now imagine that one of those channels represents the channel from the source to the destination and the other the channel from the relay to the destination. Further assume that the relay has been able to decode the source signal reliably. If we now transmit the signal first from the source and second from the relay to the destination and the destination only decides for the "better" signal, which means for the signal with higher faded signal power at a certain time instance, we get an overall faded signal power as shown in Fig. 1.2. The deepest fade now is −12 dB which is a great improvement compared to the aforementioned example. Noteworthy, that the use of a relay is not necessary if we can wait long enough so that the channel between source and destination changes significantly. However, this clearly is not practical and alternative solutions must be found.

The idea of relaying and user cooperation is also part of some working groups that try to put relaying into standardization. For instance, the Cooperative Network working group (CoNet) of the Wireless World Research Forum (WWRF) that describes a

Beyond-third-Generation (B3G) vision in [15]. They present architectural principles, research challenges, and candidate approaches and point out that B3G systems will be built over generic Internet Protocol (IP) networking technologies. The networks should be able to self-organize dynamically. Mobility management, multiple access as well as moving networks are mentioned as key components and technologies. The authors emphasize that cooperation demands a cross-layered approach. The layers are divided into application, connectivity, and access layer.

Relaying is also included as amendment to the 802.16 standard that is known as WiMAX (worldwide inter-operability for microwave access) [16–18]. WiMAX has been developed to address the problem of the “last mile” and find a wireless solution for broadband access that can compete with wired networks [19]. In order to achieve higher data rates, WiMAX employs advanced signal processing techniques like orthogonal frequency division multiple access (OFDMA) and MIMO. Nonetheless, higher data rates require a certain signal-to-noise ratio (SNR) that may be difficult to obtain at the cell edges. On the one hand WiMAX must be highly reliable and on the other hand it must provide good coverage to compete with 3G cellular networks and ensure maximum mobility. But these aspects are contradictory. Either one increases the data rate and thus reduces reliability or one increases reliability at the cost of a reduced coverage area. Mostly, this issue is solved by shrinking the cell size and installing additional base stations. But this also means that providers have to pay additional costs for antenna space at the base stations and for the wired backhaul network.

An alternative solution to this is the insertion of (fixed) relays. Those relays only aid communication from a base station to a mobile station and vice versa. In literature, this kind of network is often referred to as multi-hop cellular network [20]. A task force has been formed within the IEEE 802.16 working group to extend the IEEE 802.16e-2005 standard to relay-based multi-hop communications (802.16j).

Within the amendment 802.16j three types of relaying are included, namely transparent relaying, non-transparent relaying, and cooperative relaying. Transparent relaying describes the fact that sometimes higher throughput can be achieved if a mobile station, though it is able to decode control information from the base station, uses several relays. The reason for this is that in this case the relays do not have to transmit control information as well. The term transparent relaying denotes the fact that the mobile station is not aware of those relays. If the mobile station cannot decode control information from the base station, it is necessary that the relays transmit control information. These relays are then called non-transparent relays. The latter type of relaying, cooperative relaying, can be divided into three diversity mechanisms: cooperative source diversity, cooperative transmit diversity, and cooperative hybrid diversity. The first mechanism describes simultaneous transmission of identical signals from relays and base stations, the second one the application of space-time codes, and the latter is a combination of the two aforementioned mechanisms.

The Long Term Evolution (LTE) technology, which is a new air interface for cellular communications systems, utilizes concepts based on MIMO and orthogonal frequency division multiplexing (OFDM) in order to deliver high data rates over small cell sizes [21, 22]. Currently, there are several restrictions that limit the cell size and possible capacity extensions, for instance, path loss attenuation at high carrier frequencies or the high degree of the base station antenna down tilting angle. Hence, one study point of LTE-Advanced[2] is the enhancement of current LTE networks by employing decode-and-forward relays. The Heinrich-Hertz-Institut and Nokia Siemens Networks built up a transceiver test-bed and demonstrated by field trials that relaying indeed has a great impact on the coverage area and the achievable data rates in LTE networks.

There are still a lot of open points and unsolved problems in the field of relaying and user cooperation and a lot of more work is required in order to meet the great challenges that eventually bring us one step closer to the wireless vision mentioned at the beginning of this chapter.

1.2 Background and Related Work

In this section, we give a historical survey on the research in relaying and user cooperation. We state explicitly that this survey is by far not exhaustive. The publications presented in this section are considered to be groundbreaking and deal with the most important aspects of relaying.

The theoretical basis for the analysis of relay networks has been set by van der Meulen in 1968 [23]. In 1971, van der Meulen derived upper and lower bounds for the capacity of the “classical” relay channel [24]. Indeed, the capacity of the relay channel is unknown to date, but the results of van der Meulen could have been improved enormously by Cover and El Gamal in 1979 [25]. This publication is still seen as the most important and influential work with respect to relay networking. Cover and El Gamal considered channels that consisted of one source, one relay, and one destination. An attempt to extend the results of Cover and El Gamal to networks with multiple relays was done, for instance, in [26]. An overview over the state-of-the-art of relaying in the late 1970s was given by van der Meulen in [27]. Other publications that contributed enormously to the understanding of relaying and user cooperation are [28–31]. After that the interest in relaying and user cooperation diminished more and more, though there have been some publications on that subjects. Reasons might have been the high technical challenges in implementing user cooperation in mobile networks.

Yet with the discovery of MIMO systems in [2–4] and space-time coding in [32, 33] in the late 1990s, the interest in relaying and user cooperation rose again. In [34], new information-theoretic results have been given. Here, no direct source-to-

[2]LTE-Advanced is considered as a major enhancement of LTE. It is standardized by the 3rd Generation Partnership Project (3GPP).

destination link has been considered, but there are two relays that aid communication (so-called parallel relay channel). Another very important publication is the two-part paper by Sendonaris et al. [35, 36]. There, user cooperation is employed in the uplink to achieve diversity gains. Especially, part I describes the concept of user cooperation and proposes a cooperation strategy for a code division multiple access (CDMA) system. Part II considers practical implementation issues related to the presented cooperation concept. In [37] and [38] Laneman et al. deal with cooperation strategies over fading channels. There, common network models (see Section 2.2) and cooperation strategies (see Section 2.3) are investigated. The authors give expressions on mutual information and derive closed-form expressions for the outage probabilities in the high SNR regime.

Though, generally, high values of SNR are of minor interest for wireless communications, since we are not able to employ such high transmit powers due to regulation and health concerns, this concept led to a new performance metric called diversity order. The diversity order describes the slope of the outage probability curves in the high SNR regime and thus gives information about the differential outage behavior of cooperative networks. Moreover, the use of space-time block coding (STBC) in a two-phase relay network has been investigated, where the destination does not know a priori which relays participate in the communication process. The authors demonstrate that full spatial diversity in the number of cooperating terminals can be achieved and that these schemes are preferable over repetition based schemes for higher spectral efficiencies. A very good overview over information-theoretic aspects of relay networking is given in [39], where different coding strategies are presented and analyzed. The authors concentrate on two important cooperation strategies, namely decode-and-forward (see Subsection 2.3.2), where the relay decodes its received signal and reencodes it before transmission, and compress-and-forward (see Subsection 2.3.3), where the relay's transmit signal is a compressed and quantized version of its receive signal. For more information, we refer the reader to Section 2.3.

New results with respect to power control have been presented by Høst-Madsen and Zhang in [40]. The authors study upper bounds and lower bounds on the outage capacity and the ergodic capacity considering practical constraints at the relay node as well as the synchronization between the source and the relay. It is shown that power allocation has a significant impact on the performance of the network. A theoretical analysis of multi-hop relay networks is presented in [41, 42]. In these papers, the authors use the expression multi-hop for both, multi-hop networks that do not create diversity at the destination as well as multi-route networks that achieve diversity gains at the destination. They compare decode-and-forward to amplify-and-forward (see Subsection 2.3.1) and state that amplify-and-forward achieves better results with respect to error probability even though noise is propagated (which is not the case for decode-and-forward). This is especially true for multi-hop networks without diversity, since there the weakest link limits performance.

Apart from these more theoretical investigations, Bletsas et al. dealt mainly with the basics of implementing relay networks practically [43–45]. The authors state that cooperative diversity is, by nature, a cross-layer approach and requires consideration of the physical, link, and routing layers together. This will be done in Chapter 3 as well, where we present a novel relay selection protocol that selects the required number of relays depending on the target bit error rate at the destination. Bletsas et al. provided an implementation of cooperative diversity antenna arrays using commodity hardware. In their scheme, no channel state information (CSI) (see Subsection 2.1.3) is required at the source, which also means that no rate adaptation or beamforming is possible in their scheme. The basic idea is that there is only one relay that aids communication. Relay selection is based on instantaneous channel states between source, relay, and destination and is not based on the averaged channel qualities.

Recently, the idea of network information theory has gathered enormous interest among researchers. An information-theoretic investigation becomes involved with a growing number of relays. An approach how to deal with this issue is attempted in [46, 47]. In [8] capacity scaling laws for MIMO networks are presented, combining both relaying as well as MIMO techniques. It is demonstrated that the network capacity in a setup with one source, one destination, and an arbitrary number of K relays scales as $C = (M/2)\log_2(K) + O(1)$, where M denotes the number of antennas at the source and the destination, respectively, and has to be fixed. Furthermore, this result is only valid for $K \to \infty$ assuming perfect CSI at the destination and the relays and no CSI at the source.

The great advances in the understanding of relaying and the technological progresses in communications and signal processing made user cooperation a promising candidate for wireless ad-hoc, sensor, and mesh networks. One important issue in that context is energy consumption at each terminal in the network. Since we are considering mobile terminals, energy is a limited resource and proper allocation algorithms are indispensable. These aspects are partially covered in [48–50]. Indeed, [48] gives an excellent overview over new applications that are possible with ad-hoc networks, but also points out significant design challenges. The authors highlight the importance of energy constraints and emphasize – once again – that cross-layer designs are required in order to meet the emerging application requirements and technical challenges. Moreover, link design issues like coding, power control, and adaptive resource allocation are discussed. The authors stress that link design is particularly challenging for wireless networks due to effects caused by multipath fading and delay spread. The scarcity of spectrum leads to the aspect of medium access control, especially for large networks. This treats issues related to channelization, random access, and scheduling. Discovering neighboring nodes is essential for wireless networks. Therefore, the authors also deal with network design issues, e.g., routing and scalability. Routing is particularly challenging, since the exchange of routing data already consumes energy. This “loss of energy” should be compensated by the gain through cooperation. In addition, the question must be addressed if perfect knowledge of the network topology is necessary, as gathering perfect knowledge might lead to severe delay issues. With respect to scalability the main focus is on self-organization, distributed routing, mobility management, and security. Finally, the importance of the application layer in a cross-layer design is highlighted. Delivering a guaranteed quality of service (QoS) is unrealistic in a wireless environment due to the mobility of users and time-varying characteristics of the channels. It is, therefore, constituted that applications have to adapt to the offered QoS. For instance, utilization of a rate-delay trade-off curve is discussed. This means that the application layer decides on which point of that curve to work. Another trade-off is, e.g., one that takes energy vs. lifetime into account. In [51] the maximum lifetime routing in wireless sensor networks is discussed. In this paper, the aspect of fairness is pointed out which is very important for the design and the performance of cooperative networks.

1.3 A Note on Information Theory

This section describes the basic means we apply for the analysis of the investigated networks. So why to choose means of information theory? The answer to that question is pretty simple. Information theory provides knowledge about the ultimate data compression and the ultimate transmission rate of a communications system. Obviously, the first is given by the differential entropy h (for continuous random variables) and the latter by the channel capacity C. But information theory is not only a proper means in communications theory. Moreover, it has a great impact on computer science, economics, and mathematics (statistics and probability theory). The following paragraphs are mainly due to [52].

The differential entropy of a continuous random variable X with probability density function $f_X(x)$ and support set $\mathcal{X}$ expressed in bits is given by

$$h(X) = -\int_{\mathcal{X}} f_X(x) \log_2 f_X(x) \, \mathrm{d}x. \tag{1.1}$$

The support set $\mathcal{X}$ of a random variable is the set where $f_X(x) > 0$. Originally, the term entropy was introduced by Ludwig Boltzmann[3] to provide an expression on the second law of thermodynamics[4]. It thus describes the uncertainty in the random variable X. Clearly, entropy can also be interpreted as the expected value of $-\log_2 f_X(X)$. Hence,

$$h(X) = -\mathbb{E}(\log_2 f_X(X)). \tag{1.2}$$

[3] Ludwig Eduard Boltzmann, ∗ February 20, 1844, † September 5, 1906. Austrian physicist. Had the equation $S = k \ln W$ inscribed on his gravestone.

[4] The second law of thermodynamics states that the entropy of an isolated system will tend to increase over time.

The mutual information between two continuous random variables X and Y describes the amount of information X contains about Y (and vice versa). It is defined as

$$I(X;Y) = h(X) - h(X|Y) = h(Y) - h(Y|X). \tag{1.3}$$

This definition provides a descriptive interpretation by means of communications. Associate X with the transmitted signal and Y with the received signal of a communications system. If the uncertainty over X remains after having observed Y, i.e., $h(X|Y) = h(X)$, mutual information becomes 0. This means that the receiver is unable to decide which possible realization of X has been transmitted. On the other hand, if all uncertainty over X is resolved after having observed Y, i.e., $h(X|Y) = 0$, mutual information becomes the entropy of X.[5]

Channel capacity is defined as

$$C = \max_{f_X(x)} I(X;Y), \tag{1.4}$$

where the maximum is taken over all possible input distributions $f_X(x)$. We can now derive channel capacity for the additive white Gaussian noise (AWGN) case, where the receive signal Y is given by $X+Z$. Z denotes noise and possesses Gaussian character. Therefore, (1.4) is maximized for (complex-valued) Gaussian X and becomes

$$C = \log_2\left(1 + \frac{P}{\tilde{N}}\right) \quad \text{bits per transmission}, \tag{1.5}$$

where P is the average power of X and $\tilde{N}$ is the average power of Z. A band-limited signal with bandwidth $2B$ $(-B,B)$ and duration T can be represented by (approximately) $2BT$ samples. Channel capacity in that case can be shown to be

$$C = B\log_2\left(1 + \frac{P}{N_0 B}\right), \tag{1.6}$$

where N_0 is the one-sided noise power spectral density given in W/Hz. We use the normalized capacity (spectral efficiency) in this thesis, which is

$$\mathcal{C} = \frac{C}{B} = \log_2\left(1 + \frac{P}{N_0 B}\right) \quad \text{[bit/s/Hz]}. \tag{1.7}$$

This channel capacity describes the highest transmission rate with which an error-free transmission is possible (if the codeword length tends to ∞).

If we consider transmission over wireless channels, the transmitted signal will be affected by fading. We omit a closer description of fading and refer the interested reader to Subsection 2.1.1 and standard works on wireless communications such as [53–55]. Let h denote the channel gain of a wireless link between X and Y. If the

[5] In the following, we will not make use of the differential entropy $h(\cdot)$ anymore, but rather use h as the channel gain of a wireless link.

channel gain is chosen randomly, but fixed for at least the length of one codeword, the instantaneous channel capacity becomes

$$\mathcal{C} = \log_2\left(1 + |h|^2 \frac{P}{\tilde{N}}\right). \tag{1.8}$$

However, if the channel gain h varies over the transmission of one codeword (but all moments remain the same from codeword to codeword), we get the so-called ergodic capacity

$$\mathcal{C} = \mathbb{E}_h\left(\log_2\left(1 + |h|^2 \frac{P}{\tilde{N}}\right)\right). \tag{1.9}$$

The underlying meaning of ergodic capacity is, that the transmission time is so long that it reveals the ergodic character of the fading process [56]. Ergodic capacity is often also called Shannon capacity, e.g., [57], or throughput capacity [58]. The above equation refers to the case when channel state information is available at the receiver (cf. Subsection 2.1.3). If channel state information is also available at the transmitter, e.g., through a separate and reliable feedback channel, the transmitter can adapt its parameters to the channel conditions (at least at a certain extent). Then, the ergodic capacity becomes [59]

$$\mathcal{C} = \max_{P(\cdot)} \mathbb{E}_h\left(\log_2\left(1 + |h|^2 \frac{P(h)}{\tilde{N}}\right)\right), \tag{1.10}$$

where $P(\cdot)$ denotes the power allocation function, which is, for instance, subject to the average power constraint $\mathbb{E}(P(h)) \leq P$. The ergodic capacity can be achieved by waterfilling.

In non-ergodic fading environments, the ergodic capacity is not a useful measure anymore as it is often zero. The reason for this is that the channel gain h can be close or even equal to zero, e.g., when we consider Rayleigh fading, and we cannot guarantee reliable communications with a fixed predefined (nonzero) rate. In this case, the notion capacity-vs.-outage has been introduced in [60] and formulated in a more general way in [61]. Here, we allow a certain outage probability while transmitting over the channel, which is defined as

$$p_{\text{out}} := \Pr\left(\log_2\left(1 + |h|^2 \frac{P}{\tilde{N}}\right) < R\right). \tag{1.11}$$

In the above equation R denotes the target rate in bit/s/Hz. Then, the ϵ-outage capacity $\mathcal{C}_\epsilon$ is the highest rate R such that outage probability p_{out} satisfies $p_{\text{out}} = \Pr(I < \mathcal{C}_\epsilon) \leq \epsilon$ with $0 \leq \epsilon \leq 1$. For a given ϵ, we have

$$\mathcal{C}_\epsilon := \sup\{R : p_{\text{out}}(R, \mathsf{SNR}) \leq \epsilon\} \tag{1.12}$$

with $\mathsf{SNR} = P/\tilde{N}$. A special case of the ϵ-outage capacity is the delay-limited (or zero-outage) capacity [57, 62] which refers to an outage probability of zero. In single-user channels delay-limited capacity is associated with channel inversion which requires channel state information at the transmitter [56].

In this dissertation, we will make extensive use of channel gains that are fixed for at least the length of one codeword (block fading). The value $|h|^2 P/\tilde{N}$ is often referred to as instantaneous SNR. Then the average SNR is given by $\overline{\mathsf{SNR}} = \mathbb{E}(|h|^2)P/\tilde{N}$.

1.4 Outline of the Dissertation

This dissertation continues as follows. In Chapter 2 we give an overview over cooperative networks. Especially, we review some important characteristics of the wireless channel and describe general aspects of cooperative network models. Moreover, cooperation strategies are presented and discussed in detail. Chapter 3 treats the issue of relay selection. We present an adaptive relay selection protocol, where the number of selected relays is variable and depends on the target bit error rate at the destination. The issue of optimal resource allocation is investigated in Chapter 4. The optimization problem is solved by applying an algorithm that is based on Brent's method. Results are given for the instantaneous channel capacity as well as the delay-limited capacity. Chapter 5 presents a combining receiver for a dual-diversity wireless network. The receiver selects adaptively between two combining techniques based on a signal-to-noise ratio criterion. Our main theoretical results are contained in Chapter 6, where we deal with incremental relaying. We derive the ϵ-outage capacities of various cooperation strategies for the case of perfect and imperfect feedback. The results are extended to networks with an arbitrary number of relays. Finally, Chapter 7 concludes our findings and points out some areas for future research.

2

Cooperative Networks

> It is through cooperation, rather than conflict, that your greatest successes will be derived.
>
> *Ralph Charell*

2.1 Wireless Channel

There are salient effects of the wireless channel that strongly affect the performance of communications and influence design decisions. One of the main challenges for future mobile communications is to deal with these impairments and with the unpredictability of the wireless channel in a proper way, so that the demand for higher data rates in multimedia applications can be met. Although it may sound reasonable to face the degradation of the receive signal amplitude just by transmitting with higher powers, this may not be a proper means in practice due to practical limitations and regulatory restrictions [57]. Regulation in that sense leads to average power or peak power constraints and orthogonality constraints (to reduce interference caused by other users). Another challenge is the usage of an inherently scarce resource – frequency. On the one hand the frequency range suitable for mobile communications is restricted, on the other hand access to frequencies is also regulated by governmental bodies.

Generally, uncorrelated (or independent) channels can be generated in three physical domains, namely time, frequency, and space. All three possibilities lead to diversity at the receiver and are, thus, called time diversity, frequency diversity, and spatial diversity.[1] Diversity is a powerful technique to compensate for fading in a wireless channel [55]. The principle of diversity is very simple. Assume there are multiple paths from a transmitter to a receiver. Then, it is very likely that at least one path does not undergo a deep fade and the average SNR at the receiver can be increased. Time diversity means that a signal is transmitted repeatedly at different time instants which are greater than the coherence time (see Subsection 2.1.1) of the channel. Drawback is that time diversity can result in large system delays. When frequency diversity is employed, a signal is transmitted on several carrier frequencies. The basic idea is that the difference between those carrier frequencies exceeds the coherence bandwidth of the channel (see Subsection 2.1.1). Frequency diversity clearly leads to a waste of bandwidth [63]. Spatial diversity means that multiple receive antennas are placed in a way that they see different signal paths. This diversity technique has gained much interest recently and is one of the underlying ideas of cooperation amongst mobile terminals.

The most severe impairments to wireless communications are caused by path loss, shadowing and fading. These issues will be discussed in Subsection 2.1.1. In general, in wireless communications signals are emitted from an antenna around a carrier frequency. One reason for this is that antennas can emit only power at a certain frequency that is determined by the size and structure of the antenna. Another reason is the purpose of separating multiple users to mitigate interference. From an analytical perspective, it is more convenient to treat those signals as equivalent discrete-time baseband signals, since this allows to model the wireless channel as a (time-varying) linear filter [57].

2.1.1 Path Loss, Shadowing and Fading

In the following paragraphs, we give an overview over characteristics of the wireless channel that degrade the performance of a communication system. Profound descriptions of the wireless channel and its characteristics can be found in [53–56, 64–66].

In a wireless scenario, the receive signal can be modeled as the superposition of distorted versions of the transmit signal. According to [55, 67], there are three mechanisms that influence signal propagation in a wireless channel: reflection, diffraction, and scattering.[2] Signals over several paths are affected by different attenuation factors and delays and superpose either constructively or destructively at the receiver.

Path loss, also known as large-scale fading, leads to an attenuation of the receive signal amplitude due to propagation over large distances [67]. The path loss, i.e., the ratio of receive power P_d and transmit power P_s, where the subscripts stand for

[1] Another diversity technique is antenna polarization which is mostly used at base stations [55] to reduce costs. The disadvantage is that there are only two diversity branches. However, it is advantageous that several antenna elements can be co-located.

[2] Another mechanism that is often mentioned is refraction by different atmospheric layers.

Table 2.1: Typical path loss exponents (cf. [53]).

Urban macrocells	3.7-6.5
Urban microcells	2.7-3.5
Office building (same floor)	1.6-3.5
Office building (multiple floors)	2-6
Store	1.8-2.2
Factory	1.6-3.3
Home	3

destination and source, respectively, is given by the Friis[3] transmission equation [68]

$$\frac{P_\mathrm{d}}{P_\mathrm{s}} = G_\mathrm{s} G_\mathrm{d} \left(\frac{\lambda}{4\pi d_\mathrm{sd}} \right)^{\alpha}, \tag{2.1}$$

where G_s and G_d are the antenna gains of the source and the destination antenna, respectively, λ is the wavelength, d_sd is the distance between the source and the destination, and α denotes the path loss exponent typically between 2 (free space) and 4. Tab. 2.1 shows typical values of the path loss exponent for different environments.

The signal degradation caused by shadowing is mainly due to the blocking of the transmitted signal by obstacles between the source and the destination. These random variations of the signal depend on the physical and electrical properties of the blocking objects. For this case, the ratio of transmit and receive power $\psi = P_\mathrm{s}/P_\mathrm{d}$ is modeled as a log-normally distributed random variable with the probability density function [53]

$$f_\Psi(\psi) = \begin{cases} \frac{\xi}{\sqrt{2\pi}\sigma_\mathrm{dB}\psi} \exp\left(-\frac{(\psi_\mathrm{dB} - \mu_\mathrm{dB})^2}{2\sigma_\mathrm{dB}^2} \right) & : \quad \psi > 0 \\ 0 & : \quad \text{otherwise} \end{cases}, \tag{2.2}$$

where $\xi = 10/\ln 10 \approx 4.3429$, μ_dB is the mean value of $\psi_\mathrm{dB} = 10\log_{10}\psi$ in dB, and σ_dB is the standard deviation of ψ_dB in dB. Path loss and shadowing can be combined in order to consider both effects simultaneously. We get

$$\frac{P_\mathrm{d}}{P_\mathrm{s}} \text{ in dB } = 10\log_{10}(G_\mathrm{s}G_\mathrm{d}) + 10\alpha\log_{10}\left(\frac{\lambda}{4\pi d_\mathrm{sd}} \right) - \psi_\mathrm{dB}. \tag{2.3}$$

In a realistic scenario, the signal at the destination is a superposition of a number of different versions of the transmit signal that have experienced signal attenuations and propagation delays through different paths. This combination generates random fluctuations of the received power. Let $y(t)$ denote the discrete-time receive signal

[3]Harald T. Friis, ∗ 1893, † 1976. Danish-American radio engineer. Pioneering work in the areas of radio propagation, radio astronomy, and radar.

and $x(t)$ denote the discrete-time transmit signal. The multipath fading channel can then be modeled as

$$y(t) = \sum_{i=1}^{L} h_i(t)x(t - \tau_i(t)) + n(t), \tag{2.4}$$

where L is the number of paths, $\tau_i(t)$ is the time delay of the i-th path, and $n(t)$ is additive white Gaussian noise. The delay spread T_D of the channel is

$$T_D = \max_{i,j \in \{1,\ldots,L\}} \tau_i - \tau_j. \tag{2.5}$$

The coherence bandwidth is approximated by the inverse of the delay spread

$$B_c \approx \frac{1}{T_D}. \tag{2.6}$$

It describes the impact on the transmit signal in the frequency domain. If the signal bandwidth B_x is greater than the coherence bandwidth, the channel is frequency selective (frequency selective fading). In this case, the frequency components of the transmitted signal are affected differently by the channel and undergo independent attenuations. If the signal bandwidth is less than the coherence bandwidth, the channel is frequency non-selective (flat fading) [54].

We next describe the fading characteristics in the time domain. Assume that the distance between the source and the destination varies over time. As a consequence, the receive signal is shifted by the so-called Doppler[4] frequency shift, which can be expressed as[5]

$$f_D = f_0 \frac{v}{c} \cos\theta, \tag{2.7}$$

where f_0 is the carrier frequency of the transmitted signal, v is the speed of the moving destination, c is the speed of light in free space, and θ is the angle between the direction of propagation of the electro-magnetic wave and the direction of movement. The maximum difference in Doppler shifts is called the Doppler spread and describes the spectral broadening of the signal. The coherence time can be approximated by the inverse of the Doppler spread

$$T_c \approx \frac{1}{B_D}. \tag{2.8}$$

If the duration of the transmitted signal T_x is less than the coherence time, the channel does not vary noticeably and the different versions of the signal are affected by the same distortion (slow fading). If the signal duration is greater than the coherence time, the distortion becomes relevant and the attenuations become independent (fast fading).

Fig. 2.1 summarizes the described channel characteristics with respect to the signal's transmission time T_x and the signal's bandwidth B_x.

---

[4] Christian Andreas Doppler, ∗ November 29, 1803, † March 17, 1853. Austrian mathematician and physicist.

[5] For the sake of simplicity, we only consider one receive path here.

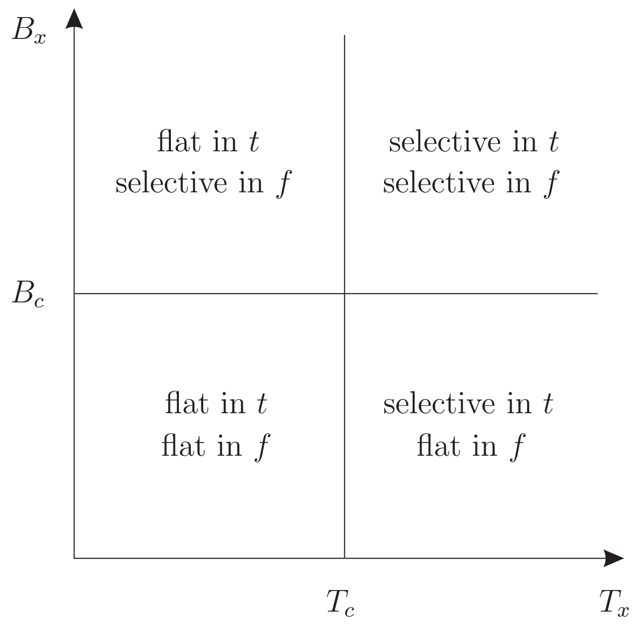

Figure 2.1: Classification of wireless channels (cf. [69]).

If the channel contains many scatterers, i.e., the number of paths L is large, the central limit theorem can be applied. This means that the channel gains follow a Gaussian distribution. Therefore, the channel gains[6] h are modeled as zero-mean[7], independent, circular-symmetric complex-valued random variables with variances σ^2. This means that inphase and quadrature phase components have variance $\sigma^2/2$ each. Then, the magnitude $|h| = \sqrt{h_I^2 + h_Q^2}$ follows a Rayleigh distribution given by

$$f_{|H|}(|h|) = \begin{cases} \frac{2|h|}{\sigma^2} \exp\left(-\frac{|h|^2}{\sigma^2}\right) & : \quad |h| \geq 0 \\ 0 & : \quad \text{otherwise} \end{cases}, \tag{2.9}$$

where $\sigma^2 = \mathbb{E}(|h|^2)$. The phase is uniformly distributed in $[0,2\pi)$. Furthermore, we can recall that the random variable $|h|^2$ is exponentially distributed with mean σ^2. Hence,

$$f_{|H|^2}(|h|^2) = \begin{cases} \frac{1}{\sigma^2} \exp\left(\frac{-|h|^2}{\sigma^2}\right) & : \quad |h|^2 \geq 0 \\ 0 & : \quad \text{otherwise} \end{cases}. \tag{2.10}$$

Fig. 2.2 shows the Rayleigh and the exponential distribution for $\sigma^2 = 1$.

[6] We omit the dependence on i and t for the sake of presentation in the following.

[7] Zero-mean in this case refers to the fact that there is no line-of-sight (NLOS) between source and destination. If there is line-of-sight (LOS), the wireless channel is modeled as a Rician fading channel, named after Stephen O. Rice (∗ November 29, 1907, † November 18, 1986).

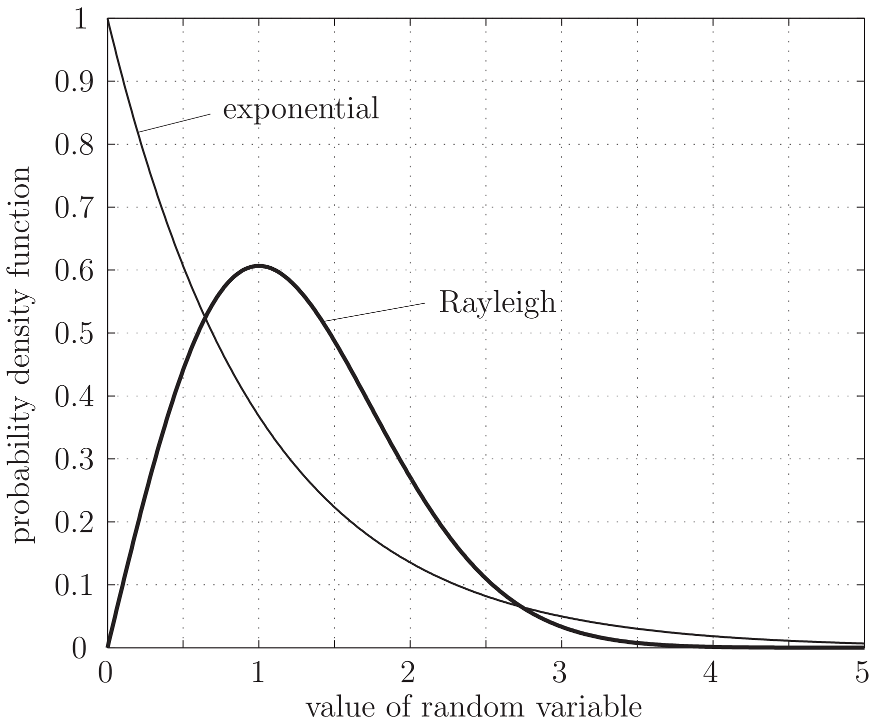

Figure 2.2: Probability density function of a Rayleigh and exponential random variable for $\sigma^2 = 1$, respectively. For the Rayleigh distribution the random variable is $|h|$, and for the exponential distribution the random variable is $|h|^2$.

2.1.2 Full-Duplex vs. Half-Duplex

In practice, it is quite difficult to let a mobile node transmit and receive at the same time (full-duplex), since the transmit signal power is usually much higher than the receive signal power. Typically, the difference between transmit and receive signal power is 100 . . . 150 dB [57]. Let y_r be the receive signal of a relay node, h_sr the channel gain between a source and a relay node, x_s the source transmit signal, and n_r additive white Gaussian noise at the relay. The half-duplex constraint is then modeled as

$$y_\mathrm{r} = \begin{cases} h_\mathrm{sr} x_\mathrm{s} + n_\mathrm{r} & : \quad x_\mathrm{r} = 0 \\ 0 & : \quad \text{otherwise} \end{cases}, \tag{2.11}$$

where x_r is the transmit signal of the relay. In [70] it is stated, that some devices with good echo cancelation can operate in full-duplex mode. However, this requires additional efforts. In the course of this dissertation, we will deal with half-duplex mobile nodes.

2.1.3 Channel State Information

Channel state information (CSI) is basically the knowledge of the instantaneous channel gains [71]. Consider a point-to-point transmission from a source to a destination, then CSI is $h_{\mathrm{sd}} = |h_{\mathrm{sd}}|e^{j\varphi_{\mathrm{sd}}}$. CSI can be divided into three categories, namely full, partial, and statistical CSI. We talk about full CSI at a transceiver, when the transceiver knows both the absolute value and the phase of the channel gain. Hence,

$$h_{\mathrm{sd}}^{\mathrm{full}} = h_{\mathrm{sd}}.$$

When partial CSI is available at a transceiver, the transceiver only knows the absolute value of the channel gain and we have

$$h_{\mathrm{sd}}^{\mathrm{partial}} = |h_{\mathrm{sd}}|.$$

The “weakest” notion of CSI is statistical CSI. Here, a transceiver has knowledge about the correlation properties of the channel [72], but no knowledge about a realization at a certain time instant.

There are several questions that arise with the notion of CSI. How can CSI be obtained? What advantage do we get if we know CSI? Where is CSI available: At the transmitter or the receiver? In the following paragraphs, we give brief answers to these questions.

CSI at the receiver (CSIR) can be obtained through (periodic) pilot sequences. Due to these pilot sequences the receiver can estimate the channel gains that affect its receive signal. If the receiver has a good estimate, it can adjust its parameters to improve the decision-making process and, thus, reliability of communication. CSI at the transmitter (CSIT), in contrast, can either be obtained through a separate feedback channel from the receiver to the transmitter or through training sequences in case of bidirectional traffic. If CSI is available at the transmitter, it can adjust its transmit power or rate to the channel conditions and improve reliability. Clearly, there is a trade-off. On the one hand, obtaining CSI requires additional costs (especially CSIT), i.e., additional resources (overhead). However, on the other hand, knowledge of CSI is necessary in order to perform resource allocation.

In [73] outage minimization and optimal power control for the fading relay channel is investigated for the case of full CSIT and CSIR. Dynamic resource allocation depending on channel states is considered in [74]. The authors assume partial CSIT and full CSIR and show that for the relay channel, in contrast to the single source-single destination channel, a non-zero delay-limited capacity is achievable in a Rayleigh fading environment. We will make the same assumptions on CSI in Chapter 4.

Tab. 2.2 summarizes combinations of fading characteristics and availability of CSIT and their corresponding channel capacity expressions.

Table 2.2: Combinations of fading characteristics and CSIT and corresponding channel capacity expressions.

slow fading, CSIT	delay-limited capacity
slow fading, no CSIT	ϵ-outage capacity
fast fading, no CSIT	ergodic capacity

2.2 Network Models

2.2.1 Multi-Hop Networks

The principle of multi-hop communications is as old as communications itself [11]. In ancient times "relay" stations have already been installed to transmit messages over large distances. The Greek playwright Aeschylus[8] described how the news of the Greek victory over Troy in 1184 BC was transmitted through a chain of fire signals over a distance of 550 km. The Romans protected the borders of their Imperium Romanum – among other things – by building watchtowers in such a way that there has been line-of-sight from one to another. In case of an attack, a message could have been sent to armed forces. One example is the Limes that ranged from the Rhine to the Danube. Indian smoke signals and telegraph poles during the Napoleonic[9] wars are more examples for "historic" multi-hop networks with the purpose of increasing coverage range and rate of transmission.

Figure 2.3: Principle of a multi-hop network with one source S, K relays R_k, $k = 1, \ldots, K$, and one destination D.

Fig. 2.3 shows the principle of multi-hop networks. A source S sends a message to a relay R_1. The relay sends the received message to the next relay R_2. The way how the relay treats the receive signal and creates a "new" transmit signal is discussed in Section 2.3. The message forwarding continues until relay R_K sends its message to the destination D. Since source and relays access the channel at different time instants, there is no interference and no conflicts with respect to medium access. An important aspect is that in multi-hop networks there is no direct link between source and destination. The maximal capacity of such a network is limited by the "weakest" link (see [75] for the proof). The major disadvantage is that an outage of

[8] Aeschylus, ∗ c. 525 BC/524 BC, † c. 456 BC/455 BC. Ancient Greek playwright. Recognized as the founder of tragedy. Famous works are *The Persians*, *The Oresteia*, and *Prometheus Bound*.

[9] Napoleon Bonaparte, ∗ August 15, 1769, † May 5, 1821. Military and political leader of France in the early 19th century. Shaped the politics in Europe at his time. Buried at Les Invalides in Paris.

an intermediate link leads to an overall system outage. The error performance of multi-hop networks with half-duplex relays was investigated in [76]. Multi-hop networks can be considered as series connection from a system-theoretic perspective.

2.2.2 Multi-Route Networks

In contrast to multi-hop networks, there is a direct link between source and destination in multi-route networks. Hence, the relays are not necessarily required for information transfer, but serve in order to increase the performance of communications.

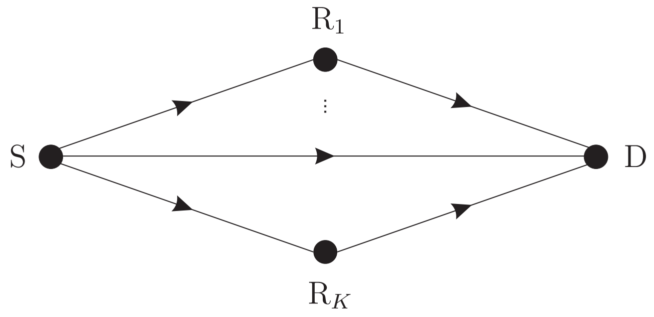

Figure 2.4: Principle of a multi-route network with one source S, K relays R_k, $k = 1,\dots,K$, and one destination D.

The principle of a multi-route network is shown in Fig. 2.4. Source S sends its message simultaneously to all relay nodes R_k, $k = 1,\dots,K$, and the destination D. After that the relays send their receive signals to the destination and the source remains silent. In order to mitigate interference between the relay messages, the relays either transmit in orthogonal time slots or apply space-time block coding (STBC) [38]. The destination now receives several signals and increases reliability by employing suitable combining strategies, e.g., selection combining, equal gain combining, or maximal ratio combining [77]. A comprehensive overview over combining strategies and a suitable receiver structure for the case of two receive signals (source signal included) is given in Chapter 5. Of course, there are other possibilities how transmission can be managed and further improved. For instance, the source can also transmit while the relays transmit. The usage of several relays either leads to a large delay if they transmit in orthogonal time slots or requires a rather complex control of medium access, which can be very challenging in practical networks.

The quality of transmission is not limited by the “weakest” link as it is the case for multi-hop networks. Consequently, outage events of single links can be compensated. In [42] the authors refer to multi-route networks as multi-hop diversity networks in contrast to the “usual” multi-hop networks. From a system-theoretic perspective, multi-route networks can be considered as parallel connections.

2.2.3 Adaptive Multi-Route Networks

The main disadvantage of multi-route networks is that performance of cooperation is limited by the source-to-relay links. For instance, if the relay always has to transmit the source message, we cannot ensure that the relay's transmit signal is error-free. This aspect further limits the performance of multi-route networks. In order to overcome this drawback, adaptive protocols have been proposed [37]. Adaptive multi-route networks can be regarded as a special case of multi-route networks. This is pointed out by the dashed lines in Fig. 2.5.

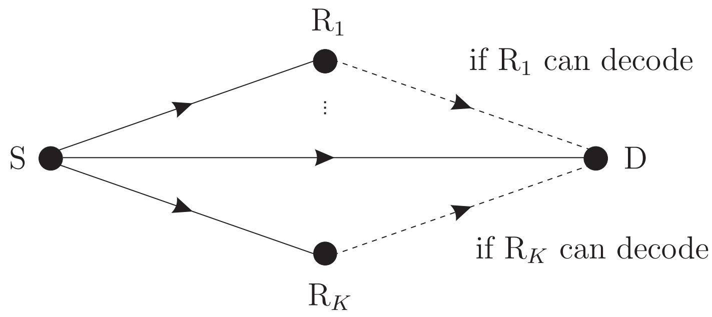

Figure 2.5: Principle of an adaptive multi-route network with one source S, K relays R_k, $k = 1, \ldots, K$, and one destination D.

In [78], two different adaptive protocols have been mentioned for a network that consists of one source, one relay, and one destination. The first one is called simple adaptive decode-and-forward (more information on cooperation strategies is given in Section 2.3). Here, the relay transmits if it has been able to decode the source message. If not, both source as well as relay remain silent in the second time slot. For the other protocol, called complex decode-and-forward, the relay also only transmits if it has been able to decode the source message. If not, the source sends its message again to the destination in the second time slot which increases the receive SNR at the destination. The advantage compared to the simple adaptive decode-and-forward protocol is that there is also a transmission in the second time slot and it does not remain idle.

2.2.4 Incremental Relaying Networks

Drawback of the already mentioned networks is a rather inefficient use of the degrees of freedom of the channel. This is essentially due to the fact that the relays retransmit the source message most of the time (depending on their receive SNR which is determined by the target rate of transmission) even if it is not necessary. A much better approach is that a relay only retransmits if it has received a request for retransmission from the destination. Such a network has been introduced in [37] as relaying with feedback, in [12] as requested relaying, and in [38] as incremental relaying. In the course of this dissertation, we refer to such an approach as incremental relaying. The main advantage is a much better use of the degrees of freedom of the channel. However, incremental relaying requires higher signaling efforts and leads to a higher system complexity. Fig. 2.6 illustrates the principle of an incremental relaying network. The feedback from the destination to the relays is shown by the dashed lines.

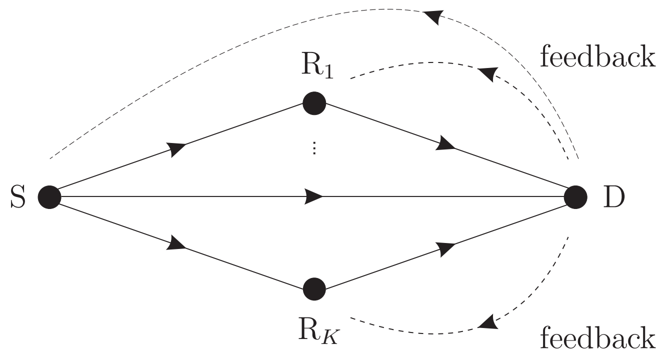

Figure 2.6: Principle of an incremental relaying network with one source S, K relays R_k, $k = 1, \ldots, K$, and one destination D.

A theoretical analysis of incremental relaying gets rather involved due to the variability of the transmission rate. Assume that the target rate of transmission is R. Hence, the source starts transmitting its message with rate R. If the destination has been able to decode the source message properly, there is no need for any relay transmission and the transmission rate matches the target rate. However, if relay transmission is required – assume there is only one relay – the overall transmission rate will become $R/2$ (if the rate of the feedback can be neglected). This is done in [37] for amplify-and-forward (see Subsection 2.3.1) in the high SNR regime where the feedback consists of only one bit that indicates success or failure of source transmission. The authors introduce an average (long-term) transmission rate $\bar{R}$ that depends on the SNR and the target rate. It has been shown in [45] that the average rate $\bar{R}$ is approximately the target rate R for large values of SNR. However, the major drawback of this approach is that the average rate $\bar{R}$ does not occur in the network at a certain time instant. Moreover, the authors examine the high SNR regime. However, if the source is allowed to transmit with a power that tends to ∞, the need for the relay transmission goes to zero and the consideration of cooperation becomes obsolete. It is therefore much more realistic to consider the low SNR regime. This was done in [79, 80] (cf. Chapter 6).

2.3 Cooperation Strategies

In this section, we discuss the three basic cooperation strategies: amplify-and-forward, decode-and-forward, and compress-and-forward. Another cooperation strategy was introduced in [81, 82] and is called coded cooperation. Here, cooperation is integrated into channel coding, which means that a codeword is divided into several blocks and sent over independent paths, i.e., by the source and several relays.

2.3.1 Amplify-and-Forward

Amplify-and-forward (AF) is the "traditional" relay strategy, which means that a signal is only amplified without any further processing in order to be able to transmit over a larger distance. Examples are analog transatlantic cables that connect Europe with North America. Due to the large distance analog repeaters (relays) have been used to mitigate power losses. Since the signal is only amplified and not further processed, AF relays are often called non-regenerative relays (mostly in satellite communications).

Drawback is that the relay amplifies a noisy version of the source signal and, hence, noise is also amplified. However, as both signals (one from the source and the other from the relay) have been transmitted through different (independent) paths and consequently have suffered different (independent) attenuations, diversity gains can be achieved at the destination. Fig. 2.7 schematically illustrates the functionality of AF in a multi-route network with one relay.

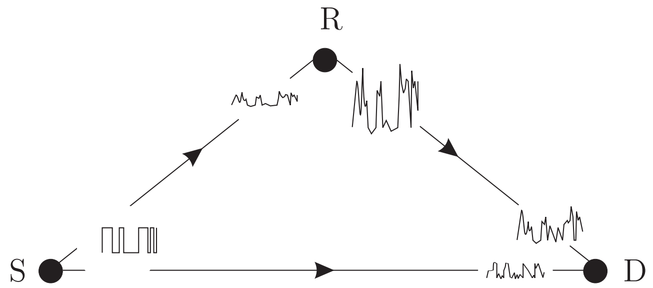

Figure 2.7: The functionality of amplify-and-forward (AF).

Let us assume that transmission takes place in two different phases. For that purpose, we divide a transmission block of duration T into two time slots of duration $T/2$ each. During the first time slot the source broadcasts its signal x_s so that the destination and the relay can receive it, respectively. The received signals at the relay and the destination are

$$y_\mathrm{r}(t) = h_\mathrm{sr}x_\mathrm{s}(t) + n_\mathrm{r}(t) \tag{2.12}$$

$$y_\mathrm{d}(t) = h_\mathrm{sd}x_\mathrm{s}(t) + n_\mathrm{d}(t), \tag{2.13}$$

where $t \in [0,T/2)$. After the source transmission, the relay amplifies its receive signal and forwards it to the destination. Let $x_\mathrm{r}(t)$ denote the relay transmit signal. Then, the receive signal at the destination after the second time slot is given by

$$y_\mathrm{d}(t) = h_\mathrm{rd}x_\mathrm{r}(t) + n_\mathrm{d}(t), \quad t \in [T/2,T). \tag{2.14}$$

In particular, we have $x_{\mathrm{r}}(t) = a y_{\mathrm{r}}(t - T/2)$, where a is the amplification factor. In order to guarantee an average power constraint $\mathbb{E}(|x(t)|^2) \leq P$, the relay must use an amplification factor of [37]

$$a = \sqrt{\frac{P}{|h_{\mathrm{sr}}|^2 P + \tilde{N}}}, \tag{2.15}$$

where $\tilde{N}$ is the average power of additive white Gaussian noise represented by $n_{\mathrm{r}}(t)$.

2.3.2 Decode-and-Forward

Decode-and-forward means that the relay decodes the source signal and reencodes it before transmitting it to the destination. In that manner, the relay sends an estimated version of the source signal to the destination. The destination then can combine the source and the relay signal to improve the decision processing. The main drawback of DF is that an erroneous estimation of the source signal at the relay will probably lead to a wrong decision at the destination, i.e., decoding errors are propagated. The main advantage compared to AF is that due to the decoding procedure at the relay, noisy signal parts will be removed and, hence, there is no noise enhancement.

DF is used in adaptive multi-route networks, where the relay only transmits its message to the destination if it has been able to decode the source signal. Since the relay sends a newly “refreshed” version of the source signal, DF relays are also called regenerative relays. Fig. 2.8 schematically shows the functionality of DF in a multi-route network with one relay.

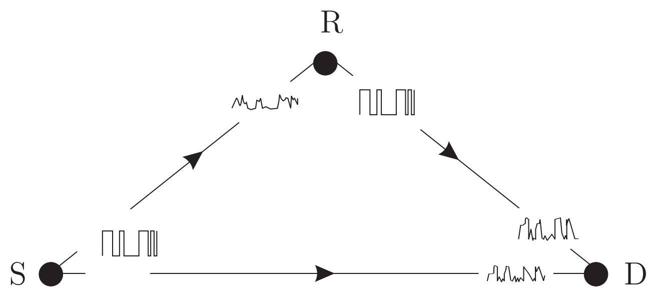

Figure 2.8: The functionality of decode-and-forward (DF).

Assume again that transmission takes place in two different phases (similar to the AF example in Subsection 2.3.1). In the first time slot the source broadcasts its message to the relay and the destination. The receive signals are given by (2.12) and (2.13). Now, the relay decodes and reencodes its receive signal, i.e., the relay's transmit signal is an estimated version of the source's transmit signal. The receive signal at the destination after the second time slot is therefore given by

$$y_{\mathrm{d}}(t) = h_{\mathrm{rd}} \hat{x}_{\mathrm{s}}(t) + n_{\mathrm{d}}(t), \quad t \in [T/2, T), \tag{2.16}$$

where $x_{\mathrm{r}}(t) = \hat{x}_{\mathrm{s}}(t)$.

2.3.3 Compress-and-Forward

The basic idea of compress-and-forward (CF) is closely related to multi-antenna reception [39]. In literature there are several expressions for CF, e.g., estimate-and-forward [25], observe-and-forward [57], or quantize-and-forward [83]. A motivating example for this cooperation strategy goes back to van der Meulen [24]. Fig. 2.9 illustrates the functionality of CF in a multi-route network with one relay.

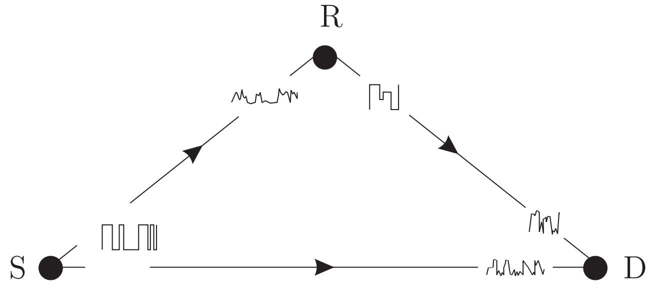

Figure 2.9: The functionality of compress-and-forward (CF).

In the first time slot, as it is the case for AF and DF, the source broadcasts its message. In the second time slot, the relay sends a compressed and quantized version of its receive signal, i.e., of the corrupted source signal, to the destination. The destination then decodes by combining this signal with its own receive signal from the source [39]. Particularly, the relay performs some sort of source coding so that side information can be exploited at the destination. An example is given in [84], where a network with one relay that performs Wyner-Ziv source coding [30] is considered.

2.4 A Note on Synchronization

Throughout this dissertation, we assume that all terminals in a network are synchronized. However, this is a rather strong assumption. The issue of synchronization is a challenging task in practice – especially for distributed relaying networks – and is of major interest for network designers. This problem becomes even more severe, if we consider several stages of relaying. Therefore, and for the sake of completeness, this section briefly reviews some synchronization approaches mainly discussed in [85].

One approach is natural synchronization. If we assume that all terminals that belong to the same hopping stage need approximately the same time for signal processing (decoding, reencoding, retransmission), then relative delay times caused by different path lengths are acceptable if they are less than the symbol duration.

Another approach exploits characteristics of a specific transmission scheme. In particular, the extended cyclic prefix approach treats cooperative networks that employ OFDM. The main purpose of the cyclic prefix, which is simply a repetition of the end of an OFDM symbol at the beginning, is the mitigation of intersymbol interference

(ISI). If the cyclic prefix is longer than the power delay profile of the channel and the expected asynchronism, then ISI is mitigated.

Another approach is the use of robust asynchronous space-time codes. In [86], the problem of asynchronism in a network with two cooperating nodes is treated by using a linear prediction-based channel estimation technique. An asynchronous space-time coded protocol is compared to a synchronous protocol with respect to diversity-multiplexing trade-off in [87]. The author demonstrates that the asynchronous scheme achieves the same diversity order as the synchronous one. Drawback of asynchronous space-time codes is, however, that the issue of synchronization is mostly treated at the expense of spectral efficiency.

3

Relay Selection

Thank you for your cooperation
and vice versa.

Eugene Ormandy

Cooperation depends on several variables such as channel quality, node characteristics, and resource availability to name a few. Almost immediately questions arise that deal with inherent characteristics of cooperation. For instance, is cooperation always useful? Which nodes should act as relays and which nodes should not? Why should some nodes act as relays, whereas others just remain silent (and thus save their own resources)? How is a cooperative network organized and how do the nodes interact?

In this chapter, we present an adaptive relay selection protocol (ARSP) that helps to improve the network performance by intelligently selecting relays. The major aim of the protocol is to select a set of relays that improve the network performance in form of the bit error rate (BER) at the destination. Hence, the selected set might consist of, e.g., one relay, all relays, or even no relay. Whether a node is eventually contained in the set of selected relays or not depends on its suitability for cooperation. The suitability is evaluated by several parameters such as channel quality between source and relay candidate, channel quality between relay candidate and destination, remaining battery power, traffic load, direction of movement, and the willingness of the relay candidate to participate in the cooperative process.

We first give an overview over existing relay selection approaches. After that, we present the ARSP in detail and give some summarizing examples that help to understand the functionality. For the simulations, we used parameters based on IEEE 802.11 and show that the ARSP is particularly suitable for ad-hoc networking [88].

3.1 Relay Selection Approaches

In the following, we briefly describe important relay selection approaches found in literature. Especially, we outline the applied system models, the selection metrics, and their specific characteristics.[1]

Bletsas et al. proposed a relay selection scheme in a slow-fading Rayleigh scenario with one source, one destination, and K relay candidates [45, 89–91]. The scheme differentiates between three phases, namely distributed relay selection, transmission from the source to the relay, and transmission from the relay to the destination. Relay selection is done by means of opportunistic relaying, which means that the relay which provides the best end-to-end path between source and destination is selected. The protocol is distributed in a way that each relay candidate measures the quality of its channel to the source and to the destination and then acts in accordance to its measurement outcome. The channel measurements are instantaneous and are possible due to the `RTS/CTS` (request-to-send/clear-to-send) handshaking, which also provides knowledge of CSI at the relays. Channel reciprocity must be assumed, which means that the channel gain from a node i to a node j, h_{ij}, is the same as the channel gain from node j to i, h_{ji}. The protocol does not exploit knowledge about CSI to perform beamforming or rate adaptation. Once relay k has measured the channel gains h_{sr_k} and $h_{\mathrm{r}_k\mathrm{d}}$, it calculates the overall channel quality q_k which is a measure for the suitability of the relay. There are two ways of calculating the channel quality. The first one is the minimum criterion

$$\mathrm{q}_k = \min\left\{|h_{\mathrm{sr}_k}|^2, |h_{\mathrm{r}_k\mathrm{d}}|^2\right\} \tag{3.1}$$

and the second one is the harmonic mean criterion[2]

$$\mathrm{q}_k = \frac{|h_{\mathrm{sr}_k}|^2 |h_{\mathrm{r}_k\mathrm{d}}|^2}{|h_{\mathrm{sr}_k}|^2 + |h_{\mathrm{r}_k\mathrm{d}}|^2}. \tag{3.2}$$

For the minimum criterion, the channel quality q_k is determined by the worst of the channel gains h_{sr_k}, $h_{\mathrm{r}_k\mathrm{d}}$. In contrast to that, the harmonic mean criterion is an average of the two channel gains. The idea is in both cases to maximize q_k. After calculating q_k, relay k sets up its timer in accordance to $\vartheta_k = 1/\mathrm{q}_k$. We see that the duration of the timer is inversely proportional to q_k. Therefore, the timer of the relay with the best channel quality q_k expires before the timers of the other nodes. After the timer has expired, the relay transmits a flag in order to show its presence for cooperation. Now, the other relays start a back-off procedure and the relay that has previously sent the flag transmits toward the destination. The relay can either use DF or AF. It is mentioned in [45], that both cooperation strategies achieve the same performance

[1] For consistency and reasons of readability we change the variables of the different works so that they match the nomenclature of this dissertation.

[2] Note that a factor of 2 is omitted in the numerator. This, however, does not effect the decision-making process.

under opportunistic relaying. This protocol does not need any information about the network topology. The measurements of the channel gains are done locally at each relay and the overhead in selecting the best relay is minimal. Nevertheless, this protocols requires significant modifications to almost all layers.

In [92], relay selection is based on the path loss (PL). The relay with the smallest path loss is chosen. This means that for each relay the path loss between source and relay and relay and destination is evaluated. The largest value of these determines the channel quality. Eventually, the relay with the smallest path loss is selected. That is

$$\text{selected relay} = \arg\min\max\{\mathsf{PL}_{\mathrm{sr}_k}, \mathsf{PL}_{\mathrm{r}_k\mathrm{d}}\}. \tag{3.3}$$

Relay selection ends here. However, since the model uses TDMA, the time slot in which the relay can transmit must also be selected. This selection process is based on the carrier-to-interference ratio, CIR_k. There are three propositions for the channel selection: smart channel selection, semi-smart channel selection, and random channel selection. Smart channel selection favors the channel which maximizes CIR_k when the channel is reusable. Semi-smart channel selection is similar to smart channel selection, but without checking the channel reusability. Lastly, random channel selection – as the name already states – randomly selects the channel.

Chu et al. analyze a system with half-duplex relays where coded cooperation is used [93]. It is assumed that CSI is known at all nodes, SNR is known at the destination, and CSI of the relay-to-destination channels is available at the relays. The source transmits its message in the first phase by means of repeat-accumulate (RA) codes[3]. The selected relay then demodulates and reencodes the source message before retransmission in the second phase. Noteworthy, that rate adaptation is possible by puncturing. The relay selection is based on three different approaches:

1. *Optimal relay selection:* The relay that minimizes the BER is selected. The approach is based on density evolution, which is an iterative procedure in order to obtain the probability density function. This is impractical for low complexity networks.

2. *Maximum mutual information:* The relay with the largest mutual information is selected. This requires knowledge of the SNR.

3. *Max-min source-to-relay-to-destination channel:* The relay with the largest minimum SNR between the source-to-relay and relay-to-destination path is selected, i.e., selected relay $= \arg\max\min\{\mathsf{SNR}_{\mathrm{sr}_k}, \mathsf{SNR}_{\mathrm{r}_k\mathrm{d}}\}$. Drawback of this approach is poor diversity. It inherently assumes that the BER is limited by the worst of the source-to-relay and relay-to-destination channels, which yields poor performance.

[3]Repeat-accumulate (RA) codes are codes where bits are repeated several times and then interleaved.

Relay selection approaches described by Nosratinia et al. in [94] are based on non-altruistic cooperation, i.e., each node itself has data to transmit. Furthermore, it is assumed that cooperation is not reciprocal. The networks consists of multiple nodes, where M nodes have own data to transmit. Each user is assigned to an orthogonal multiple access channel, which can either be in the frequency, time, or code domain (FDMA, TDMA, or CDMA). Cooperating nodes perform DF or selection DF[4]. Two selection approaches are presented, namely distributed partner selection and centralized partner selection.

1. *Distributed partner selection:* The nodes decide individually whom to assist. Each node can help n different nodes, where CSIR is available at the nodes, but no CSIT. Distributed partner selection is further divided into three schemes: random selection, fixed priority selection, and receive SNR selection. Random selection means that a node randomly selects the nodes to assist. For the fixed priority selection, selection is based on a priority list. A node tries to assist n other nodes starting with the one with highest priority, i.e., the first one in the list. Receive SNR selection means that a node measures the receive SNR and decides to cooperate with the node whose transmission is most likely to succeed.

2. *Centralized partner selection:* There are two different approaches. In the first one, the "best" relay is selected based on path losses or channel gains. One partner for each node is selected randomly. Then, outage probability is calculated and appropriate changes are made to eventually minimize outage probability. In the second approach coded cooperation is applied. A node transmits a fraction of N_1 bits of its message (which consists of $N_1 + N_2$ bits). In a second frame the remaining N_2 bits are transmitted by another node. Coded cooperation and selection DF achieve the same order of diversity under this scheme [94].

Hwuang et al. propose a relay selection algorithm that reduces the number of channel estimations in comparison to opportunistic relaying (see [45, 89–91]) from $2K$ estimations to K in a network that comprises of K nodes [95]. In spite of being suboptimal, the main advantage of this approach is that it reduces complexity and power consumption. For the selection process, the SNR between the source and each relay candidate is compared to a threshold SNR $\mathsf{SNR}_{\mathrm{th}}$. If it exceeds the threshold, then the SNR values between the relays and the destination are compared. Hence,

$$\text{selected relay} = \arg\max\{\mathsf{SNR}_{\mathrm{r}_k\mathrm{d}}\} \quad \text{if} \quad \mathsf{SNR}_{\mathrm{sr}_k} > \mathsf{SNR}_{\mathrm{th}}. \tag{3.4}$$

$$\text{selected relay} = \arg\max\min\{\mathsf{SNR}_{\mathrm{sr}_k}, \mathsf{SNR}_{\mathrm{r}_k\mathrm{d}}\}. \tag{3.5}$$

In [96] the authors propose a distributed weighted cooperative routing algorithm for a multi-hop environment. It uses the destination sequenced distance vector (DSDV)

[4]Selection DF means that a relay only transmits if it has been able to decode its receive signal correctly. Thus, it is nothing but DF in an AMR network (see Subsection 2.2.3).

routing protocol. This is a table-driven proactive protocol in which each node establishes and maintains a routing table that points to the next hop. Relay selection is performed by means of a weighted metric that comprises of the remaining energy of the relays and CSI, that is

$$w = \Omega(1 - e^{-\mathrm{q}_k}) + (1 - \Omega)\frac{E_{\mathrm{r},k}}{E_k} f(E_{\mathrm{r},k}), \tag{3.6}$$

where Ω is a weighting coefficient, q_k is the channel quality given in (3.1) and (3.2), $E_{\mathrm{r},k}$ is the remaining energy of relay k, and E_k is the initial energy of relay k. The function $f(E_{\mathrm{r},k})$ can be expressed as

$$f(E_{\mathrm{r},k}) = \begin{cases} 0 & E_{\mathrm{r},k} < E_k/2 \\ 1 & E_{\mathrm{r},k} > E_k/2 \end{cases}. \tag{3.7}$$

As for the opportunistic relaying approach, a timer is set up whose duration is inversely proportional to the channel quality.

Li et al. describe a relay selection scheme with a hybrid relaying protocol in [97]. The two-hop relay network consists of one source, K relay candidates, and one destination. Quasi-static fading is considered where the channel coefficients are constant over one frame length and change independently from one frame to another. The protocol classifies the relay candidates into two groups – one DF and one AF group. The DF group consists of nodes that have been able to decode the source message. The rest of the relay candidates is part of the AF group. Eventually, the destination measures the received SNR for each relay candidate and selects the most suited one whether it is in the DF or the AF group.

Del Coso and Ibars propose a relay selection and power allocation algorithm in order to maximize the transmission rate [98]. The relay candidates perform either DF or partial decoding. Partial decoding is more appropriate when the source has the ability to adapt the amount of information transmitted by relays with respect to the channel conditions. Relay selection is done by maximizing mutual information. Power allocation among the relays is shown to be optimal beamforming which results in waterfilling. For this algorithm a trade-off between power and cooperation can be seen that has already been mentioned in [99]: The higher the power allocated to the source, the more relays become part of the decoding set; however, they have less power to transmit.

In [100] Oechtering and Boche investigate a bidirectional communication environment. The bidirectional communication is performed in two phases, a multiple access (MAC) and a broadcast (BC) phase. During the MAC phase, two end nodes transmit to a relay node which decodes the two signals. During the BC phase, the relay broadcasts a composition of those signals. The relay that achieves the largest weighted rate sum for any bidirectional rate pair is finally selected.

In [101] the authors study optimal power allocation in a sense that the total energy consumption for two cooperating nodes is minimized. Such a strategy is particularly important in wireless sensor networks. Three matching algorithms are presented for partner selection. The first one is a maximum weighted algorithm where the energy gain between two cooperating nodes is maximized. Drawback of this approach is, however, that CSI of all interuser channels has to be available. Energy gain is defined as the ratio of the sum energy spent by users without cooperation to the sum of energy spent by users with cooperation. The second algorithm is a worst-link-first maximal gain algorithm. This means that the total energy gain is maximized by taking into account that the partner with worst channel quality and highest energy consumption has the priority in order to select its partner. The third algorithm is a worst-link-first matching algorithm. Here, the aim is to minimize the maximum energy consumption.

Souryal and Moayeri propose a channel adaptive relaying scheme, where a fixed number of relays is opportunistically selected in accordance to channel measurements [102]. The authors deal with several metrics. The first two take the position of relay candidates into account, that is

$$\text{selected relay} = \arg\min_{k\in\mathcal{N}} \{d_{\mathrm{r}_k\mathrm{d}}\} \tag{3.8}$$

$$= \arg\max_{k\in\mathcal{N}} \{d_{\mathrm{sd}} - d_{\mathrm{r}_k\mathrm{d}}\}, \tag{3.9}$$

where $\mathcal{N}$ is the set of neighboring nodes, $d_{\mathrm{r}_k\mathrm{d}}$ is the distance between the k-th relay candidate and the destination, and d_{sd} is the distance between the source and the destination. In order to account for the expected progress, the packet success probability to the k-th relay candidate $\mathrm{Pr}_{\mathrm{sr}_k}$ can be considered and we have

$$\text{selected relay} = \arg\max_{k\in\mathcal{N}}\{(d_{\mathrm{sd}} - d_{\mathrm{r}_k\mathrm{d}})\mathrm{Pr}_{\mathrm{sr}_k}\}. \tag{3.10}$$

The main drawback of these metrics is that they require information of the candidates' position. Whenever this is not possible, the distance term can be removed, which yields

$$\text{selected relay} = \arg\max_{k\in\mathcal{N}}\{\mathrm{Pr}_{\mathrm{sr}_k}(\mathsf{SNR}_{\mathrm{sr}_k})\} \tag{3.11}$$

$$= \arg\max_{k\in\mathcal{N}}\{\mathsf{SNR}_{\mathrm{sr}_k}\}. \tag{3.12}$$

Gómez-Vilardebó and Pérez-Neira present an iterative relay selection algorithm which optimizes the capacity per unit energy $C(\mathbf{E})/E$ with a total energy constraint [103]. That is

$$\eta = \max_{E_k, \sum E_k = E} \frac{C(\mathbf{E})}{E}, \tag{3.13}$$

where E_k is the average energy associated to user k, E is the total energy, and $\mathbf{E}$ is an energy allocation among the users. Energy allocation is performed in such a way that the destination has the same total energy after cooperation. Therefore, local CSI is necessary at the relays and full CSI of all links is required for the source. Those relay candidates that can guarantee a maximization of η are partitioned into a group that will later be used for retransmission.

In [104] a relay selection approach based on opportunistic feedback is investigated. After source transmission, the destination sends an `ACK` or a `NACK` to inform about successful or failed transmission. In case of a `NACK`, relay candidates send `Hello` messages to the source during a contention interval. Then, the source selects a relay that aids communication by sending parity information to the destination. If the destination is still not able to decode the source message, the source selects another relay that transmits in the next transmission phase.

Yang and Petropulu propose in [105] a relay selection approach that selects nodes which are, on the one hand, closer to the destination and which have, on the other hand, low power attenuation with respect to source nodes. Medium access is based on ALLIANCES[5]. This is a random access scheme that achieves high throughput by resolving collisions. However, information of the nodes' locations is required. Depending on the network structure and size, this drawback could lead to greater complexity. In [106] Petropulu and Lin consider a relay selection approach based on CSI. A network access point, for instance, could broadcast a set of relay candidates which is based on the received SNR. Additionally, relay selection could consider remaining battery power as in [96] and fairness issues, too.

Joint optimization for relay selection is proposed by Ng et al. in [107]. Convex optimization procedures are employed in order to optimize a utility function that depends on the user application. Under a power constraint and given channel rates, the relay that achieves the "best" value of the utility function is selected. Clearly, there is the major drawback of knowing the utility function.

In [108] relay candidates are divided into two groups – one set has a very good channel toward the destination but low interuser SNR, whereas the other set has a good interuser channel but low SNR toward the destination. The system uses coded cooperation. Cooperation regions are introduced as areas where the error probability can be reduced. Moreover, the authors talk of symmetric cooperation if a user has a good interuser channel and the channel quality from the user to the destination is comparable to the quality of the source-to-destination link. Asymmetric cooperation, on the contrary, appears when the quality of the user-to-destination channel is good.

3.2 Relay Area

The relay area comprises the region within the network where relay candidates, i.e., those nodes which have been able to receive the `RTS`/`CTS` messages, are located. This area is described by the intersection of the source's and the destination's transmission range and is illustrated in Fig. 3.1.

[5] ALLIANCES is the acronym for "allow improved access in the network via cooperation and energy savings."

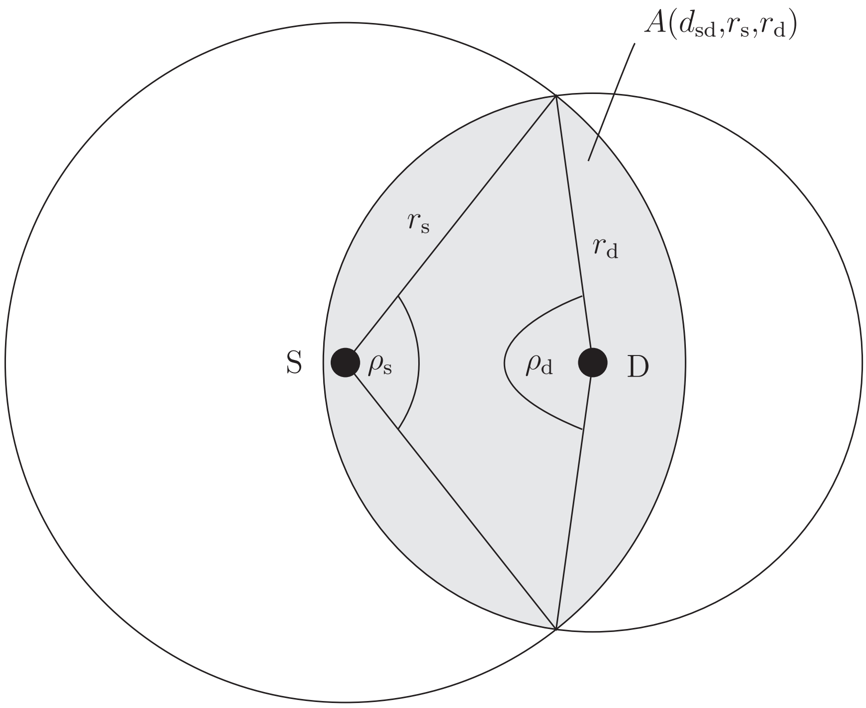

Figure 3.1: Relay area $A(d_{sd}, r_s, r_d)$.

It can be expressed as (cf. [109, p. 145])

$$A(d_{sd}, r_s, r_d) = \frac{1}{2} \left[r_s^2 \left(\rho_s - \sin(\rho_s) \right) + r_d^2 \left(\rho_d - \sin(\rho_d) \right) \right], \tag{3.14}$$

where d_{sd} is the distance between source and destination, r_s and r_d are the radii of the source's and the destination's transmission range, respectively, and ρ_s and ρ_d are given by

$$\rho_s = 2 \arccos \left(- \frac{r_d^2 - r_s^2 - d_{sd}^2}{2 d_{sd} r_s^2} \right)$$

and

$$\rho_d = 2 \arccos \left(- \frac{r_s^2 - r_d^2 - d_{sd}^2}{2 d_{sd} r_d^2} \right).$$

In practice, the network coverage, i.e., the source's and the destination's transmission ranges, depends on the transmit power and the sensitivity of each node. (3.14) contains the equation derived by Feeney et al. in [110]. By setting $r_s = r_d = 1$, we have after some algebraic manipulation

$$A(d_{sd}) = \pi - 2 \arcsin \left(\frac{d_{sd}}{2} \right) + d_{sd} \sqrt{1 - \left(\frac{d_{sd}}{2} \right)^2}. \tag{3.15}$$

Assuming that K nodes are uniformly distributed in the source's transmission range, the probability of finding at least one node in the relay area becomes

$$\Pr(d_{sd}) = 1 - \left(1 - \frac{A(d_{sd})}{\pi} \right)^K \tag{3.16}$$

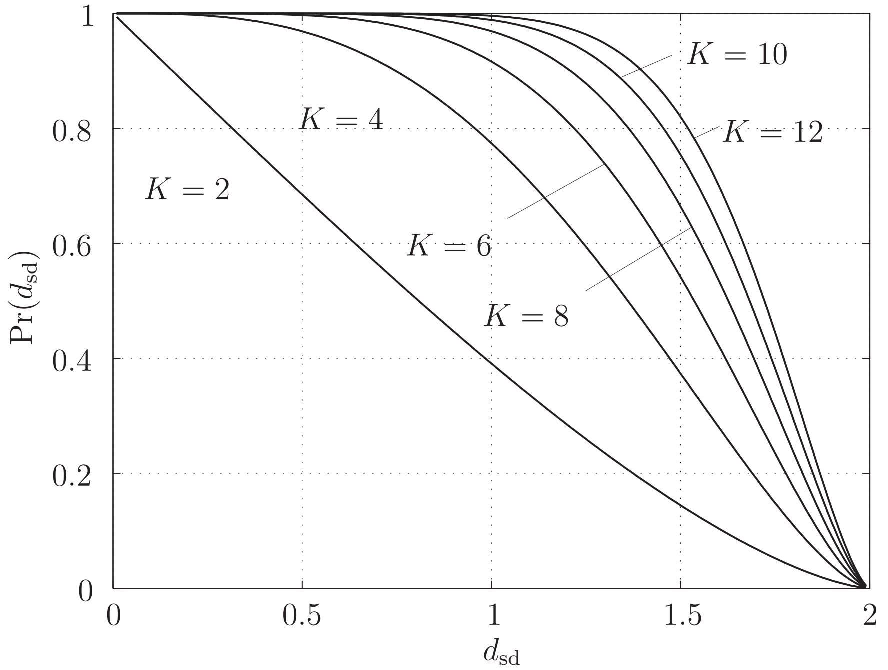

Figure 3.2: Probability of finding at least one relay candidate.

and is illustrated for $K = 2,4,6,8,10,12$ in Fig. 3.2. It is clear that a large number of relay candidates increases the probability of finding at least one relay candidate within the relay area. The relay area itself only depends on the transmission ranges and the distance between source and destination, not on the specific location of nodes (i.e., the relay area possesses the property of rotation invariance). This is due to the broadcast nature of the wireless channel.

3.3 Adaptive Relay Selection Protocol

3.3.1 General Description

The adaptive relay selection protocol (ARSP) is an adaptive and centralized protocol that selects from a group of relay candidates those who aid communication in order to guarantee a required bit error rate (BER) at the destination. The general functionality of the ARSP is illustrated in Fig. 3.3, where one relay has been selected, and briefly described in the following[6]:

1. The source sends an `RTS` frame in order to reserve the channel for an intended transmission and the destination answers with a `CTS` frame.

[6] This description assumes that there is a direct link between source and destination. However, this need not necessarily be the case. More information on that subject is given in Subsection 3.3.2.

2. Those relays that are able to receive the source's RTS frame and the destination's CTS frame ensure that a two-hop communication between source and destination is possible. Those relays lie in the relay area and serve as relay candidates.

3. The relays evaluate several conditions (i.e., outage probability from relay to destination, probability of decoding, remaining power, direction of movement, and willingness to cooperate). These conditions are combined to the intrinsic relay parameters (IRP) which are sent to the source (including a relay identifier).

4. The source ranks the relay candidates with respect to the so-called γ-coefficient. The γ-coefficient is a weighted sum of the IRP. After the ranking, the source evaluates the error probability and selects K relays in order to guarantee the required BER.

5. Hereafter, the source sends the relay table (RT), which contains – in an ordered manner – the selected relay identifiers.

6. The source transmits its message (data). The selected relays and the destination decode the source message.

7. The selected relays transmit subsequently during the next time slots. The highest ranked relay, i.e., the relay whose identifier is first in the RT starts followed by the second one and so on.

8. The destination combines the received signals by applying maximal ratio combining (MRC).

9. Finally, the destination sends an acknowledgment (ACK).

Since all nodes access the medium non-deterministically, medium access must be managed properly in order to avoid harmful interference. This is done by a modified version of the distributed coordination function (DCF) [111–113] and is discussed in detail in the next subsection.

3.3.2 System Model and Medium Access

Initially, the source sends an RTS message in order to reserve the channel and to inform the destination about an intended data transmission. If there is a direct path between source and destination, the destination replies with a CTS message. If there exists no direct path between source and destination, the source randomly selects one of its one-hop neighbors (relays) to forward its request. Clearly, this one-hop neighbor must have the destination as a one-hop neighbor as well. The knowledge of one-hop neighbors is obtained by periodical Hello messages.[7]

[7] More information on Hello messaging is given in Section 3.4, where we discuss several examples of the ARSP.

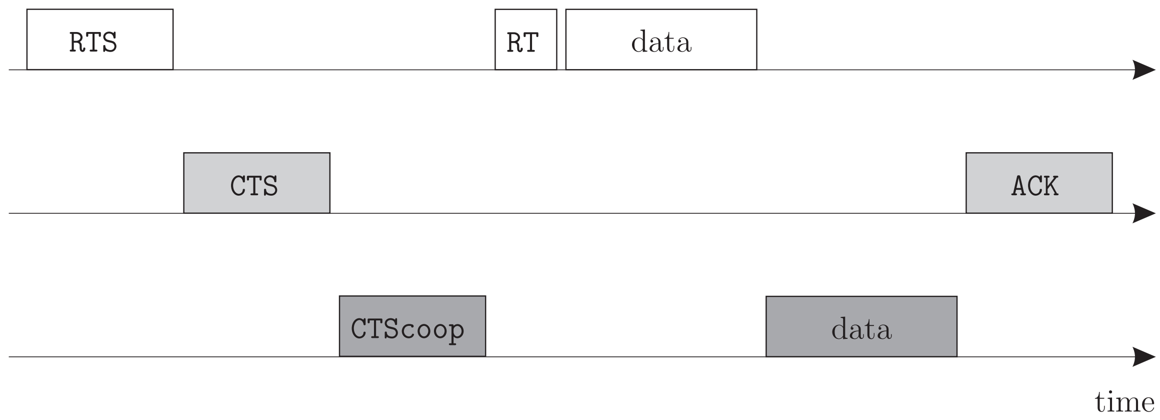

Figure 3.3: General functionality of the adaptive relay selection protocol (ARSP), where one relay has been selected.

Each node that is located in the relay area is able to receive the `RTS`/`CTS` messages. As soon as a node receives the source's `RTS` frame, it sets up its network allocation vector (NAV). If this node is also able to receive the destination's `CTS` frame, it is a relay candidate and starts to measure the IRP. Those nodes that have received either the `RTS` or the `CTS` frame – but not both – set up their NAV in order to not cause harmful interference. The IRP are discussed in detail in Subsection 3.3.3. Briefly, with knowledge of the IRP, the source is able to judge the suitability of a relay candidate for relaying. The IRP are sent from the relay candidates to the source by `CTScoop` messages that follow the destination's `CTS` frame. This is a major variation to the usual DCF with handshaking [112]. In order to clarify if cooperation is required, the destination sets up a flag within its `CTS` frame.

If no cooperation is required, all relay candidates set up their NAV and remain silent for the duration of source transmission which starts after a short interframe space (SIFS). If cooperation is required, the source waits a fixed time for possible relay responses (`CTScoop` frames). If no direct path between source and destination exists, but there is a one-hop neighbor to both nodes, the first `CTScoop` message informs the source that cooperation is required. Consequently, the source sets up its timer and the procedure continues as described before.

Data transmission is performed in $K+1$ phases, where K denotes the number of selected relays. In the first phase, the source broadcasts its message. Hence, the received signals after the first phase at the destination and the k-th relay candidate are[8]

$$y_{\mathrm{sd}} = h_{\mathrm{sd}} x_{\mathrm{s}} + n_{\mathrm{sd}} \tag{3.17}$$

$$y_{\mathrm{sr}_k} = h_{\mathrm{sr}_k} x_{\mathrm{s}} + n_{\mathrm{sr}_k}, \tag{3.18}$$

where x_{s} is the source message and n_{ij} is AWGN on the channel between node i and j. In the subsequent phases, the relays transmit one after another, where the sequence

[8] Again, we omit the time dependency since it becomes clear from the context.

of transmission is determined by the relay table RT. Therefore, after each phase, the signal at the destination is a degraded version of the signal sent by the selected relay. We have

$$y_{\mathrm{r}_k\mathrm{d}} = h_{\mathrm{r}_k\mathrm{d}} x_{\mathrm{r}_k} + n_{\mathrm{r}_k\mathrm{d}}, \tag{3.19}$$

where x_{r_k} is an estimated version of the source signal at the k-th relay and $h_{\mathrm{r}_k\mathrm{d}}$ is the channel gain between the k-th relay and the destination. The destination then performs MRC. The final output at the destination thus is

$$y_{\mathrm{d}} = \sum_{k=1}^{K+\kappa} \mathrm{g}_k y_{\mathrm{r}_k\mathrm{d}}, \tag{3.20}$$

where g_k are the weighting coefficients for the k-th signal and κ is a flag set to one if there is a direct link between source and destination and zero otherwise. In this case, MRC also takes into account the signal from the source with $y_{\mathrm{r}_{K+1}\mathrm{d}} = y_{\mathrm{sd}}$. In order to maximize the average receive SNR, the weighting coefficients are chosen to be [55, p. 329]

$$\mathrm{g}_k = \frac{y_{\mathrm{r}_k\mathrm{d}}}{\tilde{N}}, \tag{3.21}$$

where it is assumed that each channel has the same average noise power $\tilde{N}$. This leads to the summation of the individual receive SNR values. Finally, average power at each node is constrained by

$$\mathbb{E}(|x_k|^2) \leq P_k, \tag{3.22}$$

where $\mathbb{E}(\cdot)$ denotes expectation.

3.3.3 Intrinsic Relay Parameters

Selection of the most suitable relays is based on five intrinsic characteristics. These include the channel quality from the k-th relay candidate to the destination denoted by $p_{\mathrm{nout},k}$, the channel quality between the source and the k-th relay candidate expressed as $p_{\mathrm{dec},k}$, the remaining battery power $P_{\mathrm{rem},k}$, the direction of movement D_k, and the relay candidate's willingness to cooperate W_k.

The channel quality between the k-th relay candidate and the destination is measured by the probability that this channel is not in a deep fade, i.e., there is no outage event. That is, the instantaneous channel capacity is greater than or equal to the target rate R. We have

$$p_{\mathrm{nout},k} = \Pr\left(\log_2(1 + |h_{\mathrm{r}_k\mathrm{d}}|^2 \mathsf{SNR}_{\mathrm{r},k}) \geq R\right), \tag{3.23}$$

where $\mathsf{SNR}_{\mathrm{r},k}$ is the SNR at the destination after receiving a signal from the k-th relay. Recalling that $|h_{\mathrm{r}_k\mathrm{d}}|^2$ is exponentially distributed, the probability that no outage event occurs becomes

$$p_{\mathrm{nout},k} = \exp\left(-\frac{2^R - 1}{\sigma^2_{\mathrm{r}_k\mathrm{d}} \mathsf{SNR}_{\mathrm{r},k}}\right), \tag{3.24}$$

where $\sigma^2_{\mathrm{r}_k\mathrm{d}}$ is the mean value of $|h_{\mathrm{r}_k\mathrm{d}}|^2$.

So far, we have considered the channel quality between the relay candidates and the destination. Now, we deal with the channel quality between the source and the relay candidates. It is given by the ability of a relay candidate to decode the source signal. In accordance to [114], we define this ability as the probability that the instantaneous receive SNR at the k-th relay candidate SNR_k is above a certain threshold SNR $\mathsf{SNR}_{\mathrm{th}}$, which depends on the target rate R. Therefore, the relay can decode whenever

$$\mathsf{SNR}_k = |h_{\mathrm{sr}_k}|^2 \mathsf{SNR} > \mathsf{SNR}_{\mathrm{th}} \quad \Leftrightarrow \quad |h_{\mathrm{sr}_k}|^2 > \frac{\mathsf{SNR}_{\mathrm{th}}}{\mathsf{SNR}}, \tag{3.25}$$

Accordingly, the decoding probability of the k-th relay candidate becomes

$$p_{\mathrm{dec},k} = \exp\left(-\frac{\mathsf{SNR}_{\mathrm{th}}}{\sigma^2_{\mathrm{sr}_k}\mathsf{SNR}}\right), \tag{3.26}$$

where $\sigma^2_{\mathrm{sr}_k} = \mathbb{E}(|h_{\mathrm{sr}_k}|^2)$.

Power consumption is a crucial point in wireless networks and is one of the most important bottlenecks for system designers [48]. Obviously, high power consumption leads to an inefficient use of the mobile equipment. We model the remaining battery power by a random variable $P_{\mathrm{rem},k}$, where low values represent relay candidates that have already been active for a long time. An important aspect is that this parameter can be measured locally by the relay candidates. By modeling this parameter, a difficult scenario is also taken into account. Imagine there is only one relay candidate that can eventually serve as relay, but this mobile node has too low power so that it will itself become useless after cooperation. Should a node "sacrifice" itself for cooperation? The remaining battery power can be expressed as a ratio of the currently available power and the initial power of the relay candidate in order to have a relative consideration, which results in a parameter with values between 0 and 1. This approach is similar to [96].

Next, we consider the direction of movement of each relay candidate. Generally, the ARSP prioritizes relay candidates that move toward the destination. Hereby, it is assumed that a mobile node that is moving from the source to the destination is more useful than a relay that goes away from the destination. This is not in contrast to the well-known fact that DF performs better if the relay is located close to the source, because it is then often assumed that the relay node is placed on a straight line between source and destination. However, for the ARSP, we deal with a two-dimensional geometric model. The direction of movement is measured by a counter $\tilde{\vartheta}_k$. When a relay candidate receives the source's `RTS` frame, it starts an internal counter. This counter is stopped as soon as the relay candidate receives the destination's `CTS` frame. Relay candidates with large counter values model nodes that were moving away from the destination, whereas those nodes which possess lower counter values model nodes that were approaching to it. For a better understanding, consider the scenario illustrated in Fig. 3.4. Assume that source S and destination D do not move. The

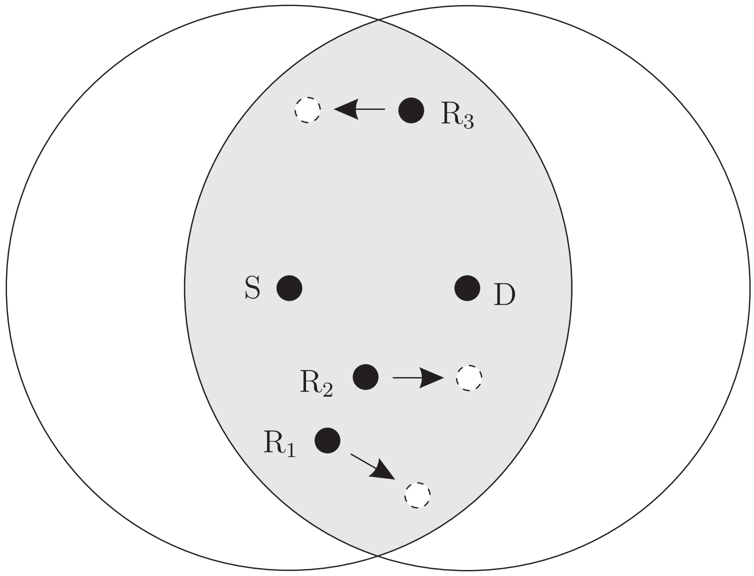

Figure 3.4: Direction of movement.

full circles represent the locations of the relay candidates R_1, R_2, and R_3 when they receive the RTS frame. The white circles with the dashed borderline show the location of the relay candidates after having received the CTS. For reasons of simplicity let us assume that all relay candidates move with the same velocity so that the geometrical distance between the relay candidates and the destination clearly indicates the time it takes for the relay candidates to be able to receive the CTS frame. We see that the counters satisfy $\tilde{\vartheta}_2 < \tilde{\vartheta}_1 < \tilde{\vartheta}_3$. The parameter that eventually describes the direction of movement is inversely proportional to the counter value, i.e., $D_k \propto 1/\tilde{\vartheta}_k$, and in the mentioned scenario the ARSP therefore prefers the relay candidate R_2 over R_1 and R_3. For the simulations, there are two possibilities how the counters can be implemented. The counters can be initiated by the `tic` function and terminated by the `toc` function in Matlab. However, we assume in Section 3.4 that all relay candidates receive the RTS frame at the same time and, hence, model D_k as a uniformly distributed random variable on the interval [0,1].

In order to have a more realistic scenario, it is advantageous to create a parameter that also includes characteristics like traffic load. This parameter represents the willingness W_k of the k-th relay candidate for cooperation and is modeled as a uniformly distributed random variable on [0,1] as well.

3.3.4 Relay Table

In Subsection 3.3.3, we described the intrinsic relay parameters that affect the selection process of our protocol. In the following paragraphs, we present how those parameters are processed in order to create a ranking of the relay candidates. This is basically done by a coefficient γ, which is a weighted sum of the intrinsic relay parameters. The source will finally use this coefficient in order to evaluate the most suitable relay candidates.

Formally, the γ-coefficient of the k-th relay candidate is denoted as

$$\gamma_k = x_1 p_{\text{nout},k} + x_2 p_{\text{dec},k} + x_3 P_{\text{rem},k} + x_4 D_k + x_5 W_k, \tag{3.27}$$

where $\mathbf{x} = [x_1 x_2 x_3 x_4 x_5]^T$ is a weighting vector which is equal for every relay candidate and depends on the traffic type of the network. It is obvious that the larger the γ-coefficient, the more suitable a relay is. The whole network that consists of K relay candidates is therefore described by

$$\begin{aligned}
\gamma_1 &= x_1 p_{\text{nout},1} + x_2 p_{\text{dec},1} + x_3 P_{\text{rem},1} + x_4 D_1 + x_5 W_1 \\
\gamma_2 &= x_1 p_{\text{nout},2} + x_2 p_{\text{dec},2} + x_3 P_{\text{rem},2} + x_4 D_2 + x_5 W_2 \\
&\vdots \\
\gamma_K &= x_1 p_{\text{nout},K} + x_2 p_{\text{dec},K} + x_3 P_{\text{rem},K} + x_4 D_K + x_5 W_K.
\end{aligned}$$

For a more compact description, we apply a matrix notation and get

$$\begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_K \end{bmatrix} = \boldsymbol{\pi} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_5 \end{bmatrix}, \tag{3.28}$$

where $\boldsymbol{\pi}$ is given by

$$\boldsymbol{\pi} = \begin{bmatrix} p_{\text{nout},1} & p_{\text{dec},1} & P_{\text{rem},1} & D_1 & W_1 \\ p_{\text{nout},2} & p_{\text{dec},2} & P_{\text{rem},2} & D_2 & W_2 \\ \vdots & & \ddots & & \vdots \\ p_{\text{nout},K} & p_{\text{dec},K} & P_{\text{rem},K} & D_K & W_K \end{bmatrix}. \tag{3.29}$$

Next, we have to consider how the weighting factors $\mathbf{x}$ are determined in order to take the network traffic load into account. This can be done by means of multiple criteria optimization [115]. However, for the ARSP we select the weighting vector intuitively. We deal with two types of traffic load and, therefore, distinguish between non-delay tolerant traffic and delay tolerant traffic. Noteworthy, that the coefficients x_4 and x_5 are not constrained by the type of data traffic.

Case 1 – non-delay tolerant traffic: Our main objective is to make sure that there is no outage at the destination. As a consequence, we prioritize the parameter that describes the relay-destination channel. Additionally, the remaining battery power at the relays is a critical component and we have the following constraints: $x_1 > x_2$ and $x_3 > x_2$. Hence, a possible weighting vector is $\mathbf{x} = [2\ 0.5\ 1\ 1\ 1]^T$. It was shown by simulations that the actual value of x_i, $i \in [1, \ldots, 5]$, does not have a significant influence on the decision making process as long as the above mentioned conditions are true. In this context, AF could be beneficial over DF in order to avoid time delays due to decoding and encoding.

Case 2 – delay tolerant traffic: If delays are tolerable, we might spend more time on decoding and encoding and we have the constraints $x_2 > x_1$ and $x_2 > x_3$. Accordingly, $\mathbf{x} = [1\ 2\ 1\ 1\ 1]^T$.

With information about the γ-coefficients, the source is now able to rank the relay candidates with respect to their suitability for cooperation. Now, the source starts an iterative process. First, it calculates the achieved BER without cooperation. If this BER is above the required BER, it evaluates if the requirement is met by additionally using the "best" relay candidate, i.e., the relay candidate with the largest γ-coefficient. If the BER is still above the requirement, the "second best" relay candidate is included in the calculations. The source continues until the required BER is obtained. The iteration process continues as long as the source has `CTScoop` frames to evaluate. This means that if the source has only received a single `CTScoop` frame, i.e., there is only one relay candidate, but the use of this relay candidate is not enough to satisfy the BER requirement, the protocol gives a warning about failing. Calculation of the BER is done by evaluating

$$\mathsf{BER} \leq \frac{a}{2} \sum_{k=1}^{K} \frac{\overline{\mathsf{SNR}}_k^{K-1}}{\prod\limits_{l=1, l\neq k}^{K} \left(\overline{\mathsf{SNR}}_k - \overline{\mathsf{SNR}}_l\right)} \left(1 - \sqrt{\frac{\overline{\mathsf{SNR}}_k}{2/b + \overline{\mathsf{SNR}}_k}}\right), \tag{3.30}$$

where a and b depend on the modulation scheme and the type of approximation used for the derivation of the BER. For the nearest neighbor approximation, a describes the number of nearest neighbors to a constellation point at the minimum distance divided by the number of bits that form a symbol ($\log_2(M)$). The parameter b is given by the multiplication of $\log_2(M)$ and a constant that is related to the ratio of minimum distance and average symbol energy [53, ch. 6.1.6]. In particular, for BPSK/QPSK we have $a = 1$ and $b = 2$. For more information, we refer the reader to Appendix B. K is the number of signals received at the destination (i.e., if there exists a direct path between source and destination, the number of selected relays will be $K - 1$), and $\overline{\mathsf{SNR}}_k$ is the average SNR of the k-th path. (3.30) is based on the fact that all paths are independent but not necessarily identically distributed. The proof is given in Appendix B. After the source has finished its calculations and has selected the relay candidates, it sends the relay table `RT`, which contains the relay identifiers and the order in which the selected relays have to transmit.

3.4 Examples

In this subsection, we demonstrate the functionality of the ARSP by simulations. For this purpose, we created a simulation environment based on the IEEE 802.11b standard (cf. [111]). The carrier frequency is 2.4 GHz. The transmission power is set to 10 mW and the pass loss exponent is $\alpha = 3$. All other parameters are taken in accordance to [111]. QPSK with Gray encoding is used at each transmitting node. All nodes are uniformly distributed over a square of size 100 m × 100 m. For the simulations, the locations of the nodes have been fixed in order to have a fair comparison between the different examples that we consider. The locations of the nodes are shown in Tab. 3.1. Node 2 has randomly been selected as source and node 6 is the destination. All other network nodes can act as potential relay candidates (cf. Fig. 3.5).

Table 3.1: Location of the nodes within the network (see Fig. 3.5).

node identifier (ID)	x [m]	y [m]
1	10.4378	50.6501
2 (source)	16.5728	4.1986
3	50.8175	89.8505
4	44.0026	84.5284
5	47.0419	60.6535
6 (destination)	85.0867	41.9677
7	54.3127	23.3511
8	94.5233	45.9014
9	12.7782	93.4829
10	30.6439	39.1191

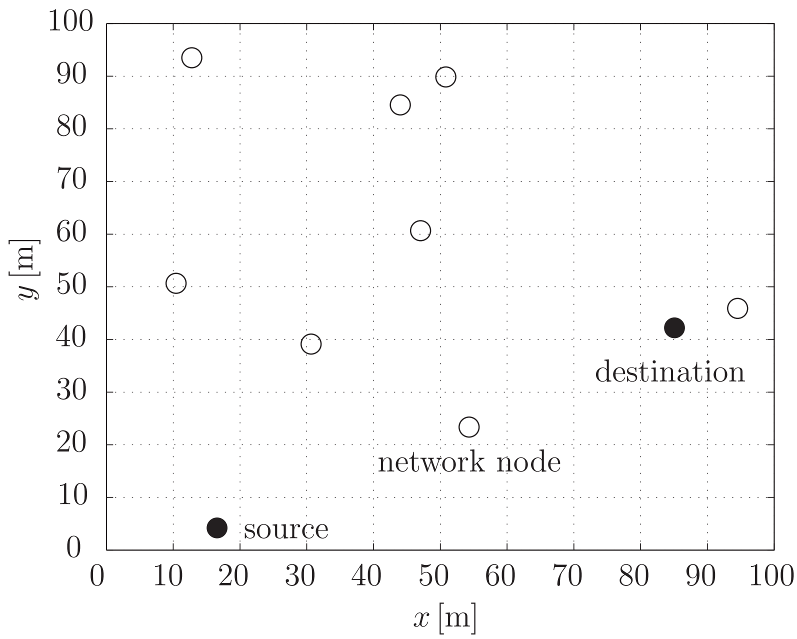

Figure 3.5: Network constellation used for the simulations.

Table 3.2: Sending of `Hello` messages.

time [s]	node ID	packet type
$1.00 \cdot 10^{-5}$	1	`Hello`
$3.73 \cdot 10^{-3}$	2	`Hello`
$7.45 \cdot 10^{-3}$	3	`Hello`
$1.12 \cdot 10^{-2}$	4	`Hello`
$1.48 \cdot 10^{-2}$	5	`Hello`
$1.86 \cdot 10^{-2}$	6	`Hello`
$2.23 \cdot 10^{-2}$	7	`Hello`
$2.60 \cdot 10^{-2}$	8	`Hello`
$2.97 \cdot 10^{-2}$	9	`Hello`
$3.35 \cdot 10^{-2}$	10	`Hello`
$1.50 \cdot 10^{-1}$	1	`Hello`
$1.53 \cdot 10^{-1}$	2	`Hello`
$1.57 \cdot 10^{-1}$	3	`Hello`
$1.61 \cdot 10^{-1}$	4	`Hello`
$1.64 \cdot 10^{-1}$	5	`Hello`
$1.68 \cdot 10^{-1}$	6	`Hello`
$1.72 \cdot 10^{-1}$	7	`Hello`
$1.76 \cdot 10^{-1}$	8	`Hello`
$1.79 \cdot 10^{-1}$	9	`Hello`
$1.83 \cdot 10^{-1}$	10	`Hello`

We consider four different examples. First, we start with the case of non-delay tolerant traffic (case 1 in Subsection 3.3.4), where a BER of 10^{-3} is required. Then, we investigate the delay tolerant case (case 2 in Subsection 3.3.4), where the required BER is also 10^{-3}. Third, we consider the delay tolerant case for a BER of 10^{-4} and, finally, for a BER of 10^{-2}. The first simulation example is explained in detail. The following examples are then discussed by a concluding statement.

Example 1: $\mathbf{x} = [2\,0.5\,1\,1\,1]^T$ and $\mathsf{BER} = 10^{-3}$

Simulations start with the nodes exchanging information about their one- and two-hop neighbors by sending `Hello` messages (Tab. 3.2). After that, the source (node 2) creates a message for the destination (node 6) at the time instant $t = 2$ s. The source sends an `RTS` and the destination sends a `CTS`. Those nodes that can act as relay candidates, i.e., that have been able to receive the `RTS` as well as the `CTS`, indicate this by sending a `CTScoop` packet. We see (cf. Tab. 3.3) that node 9 cannot act as relay candidate.

In a next step, the source calculates the achievable BER and checks if the requirement of $\mathsf{BER} = 10^{-3}$ can be met. For this purpose, it first evaluates the suitability of

Table 3.3: Sending of RTS, CTS, and CTScoop.

time [s]	node ID	packet type
2.000605	2	RTS
2.025967	6	CTS
2.026745	1	CTScoop
2.027693	3	CTScoop
2.028167	4	CTScoop
2.028641	5	CTScoop
2.029589	7	CTScoop
2.030063	8	CTScoop
2.031011	10	CTScoop

Table 3.4: Node identifiers (IDs) and γ-coefficients.

node ID	γ-coefficient
10	4.8869
3	4.1413
7	3.9749
8	3.8444
4	3.8219
1	3.6665
5	3.3935

the nodes to act as a relay. This is done by comparing the γ-coefficients of the relay candidates. Tab. 3.4 shows the node IDs and the corresponding γ-coefficients in an ordered manner. We see that node 10 is best suited for cooperation, followed by node 3. In our simulation the BER of direct transmission is only $\mathsf{BER} = 0.0597$, and the achievable BER if the "best" relay aids (node 10 in this case) becomes 0.0018. We see that in both cases the required BER cannot be achieved. With the help of two relays (node 10 and 3), a BER of $\mathsf{BER} = 3.73 \cdot 10^{-5}$ can be achieved, which meets the requirement. Hence, the source broadcasts the relay table to inform the nodes about the cooperation needs and data transmission continues. Finally, the destination acknowledges the reception of the transmitted data. This is shown in Tab. 3.5.

Fig. 3.6 illustrates the network constellation of example 1. We see a rather surprising result. For instance, the node located close to the destination has not been selected as relay, though it might be assumed that the path between this node and the destination is good. In contrast to that, node 3 (located at $(x,y) = (50.8175\,\mathrm{m}, 89.8505\,\mathrm{m})$) has been selected, which is pretty far away from both the source and the destination. Reason for this is that the ARSP is not based on a simple metric that only takes

Table 3.5: Sending of the relay table RT, data, and ACK.

time [s]	node ID	packet type
2.031809	2	RT
2.032171	2	data
2.072684	10	data
2.092940	3	data
2.113197	6	ack

the distance between nodes into account, but rather considers several aspects which are important for the overall network performance, for instance, remaining battery power.

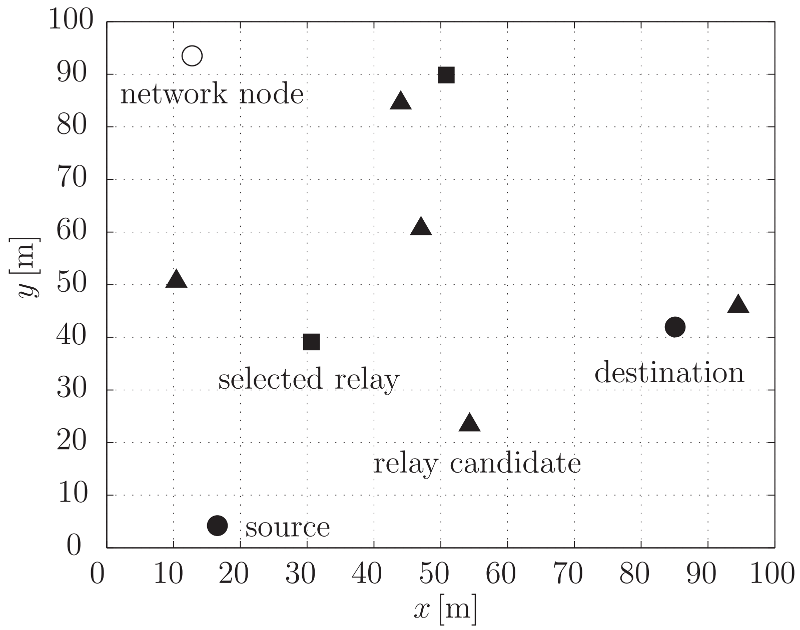

Figure 3.6: Network constellation of example 1 with $\mathbf{x} = [2\,0.5\,1\,1\,1]^T$ and $\mathsf{BER} = 10^{-3}$.

Example 2: $\mathbf{x} = [1\,2\,1\,1\,1]^T$ **and** $\mathsf{BER} = 10^{-3}$

This example demonstrates the performance of the ARSP for the delay tolerant case, i.e., $\mathbf{x} = [1\,2\,1\,1\,1]^T$, and a BER of 10^{-3}. The transmission procedure is the same as for example 1. The network constellation is shown in Fig. 3.7. Two relays are necessary to meet the required BER. The selected relays are node 7 and node 1 with the corresponding γ-coefficients 5.1611 and 4.8326, respectively. The achievable BER becomes $6.87 \cdot 10^{-4}$.

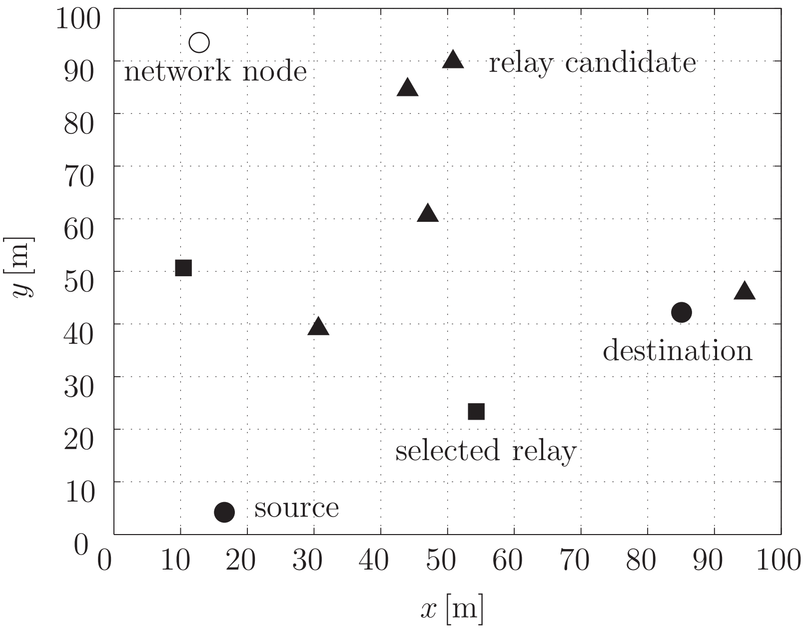

Figure 3.7: Network constellation of example 2 with $\mathbf{x} = [1\,2\,1\,1\,1]^T$ and $\mathsf{BER} = 10^{-3}$.

Example 3: $\mathbf{x} = [1\,2\,1\,1\,1]^T$ **and** $\mathsf{BER} = 10^{-4}$

In this example, we consider the delay tolerant case with a required BER of 10^{-4}. Though we consider different Rayleigh fading scenarios where the channel coefficients are random variables, we might expect that at least two relays are necessary to fulfill a BER of 10^{-4}. Indeed, if those two relays with the highest γ-coefficients are taken into account (node 10 with $\gamma = 4.8931$ and node 1 with $\gamma = 4.6519$, respectively), the achievable BER becomes $6.77 \cdot 10^{-4}$, which is not yet sufficient. If the next "best" relay with respect to its γ-coefficient is considered (node 7 with $\gamma = 4.6349$), the BER then is $2.41 \cdot 10^{-5}$. This value meets the requirement and the network is adapted accordingly. The network constellation for this example is illustrated in Fig. 3.8. We can easily see that now three relays have been selected. Interestingly, all selected relays are co-located.

Example 4: $\mathbf{x} = [1\,2\,1\,1\,1]^T$ **and** $\mathsf{BER} = 10^{-2}$

In a final example, we set the required BER to a value of 10^{-2}. Since this is not a difficult requirement, we may assume that the number of selected relays is rather low (if any relay is selected). The simulation results are depicted in Fig. 3.9. We immediately see that no relay node had to be selected. The achieved BER for direct transmission in this case is $4.39 \cdot 10^{-3}$, which obviously fulfills the requirement.

As already stated, one major advantage of the ARSP compared to other selection protocols is that the number of selected relays is not fixed. Therefore, the deployment of network resources is much more efficient. This can be particularly seen in example 4. The requirement of $\mathsf{BER} = 10^{-2}$ is met with direct transmission and all other

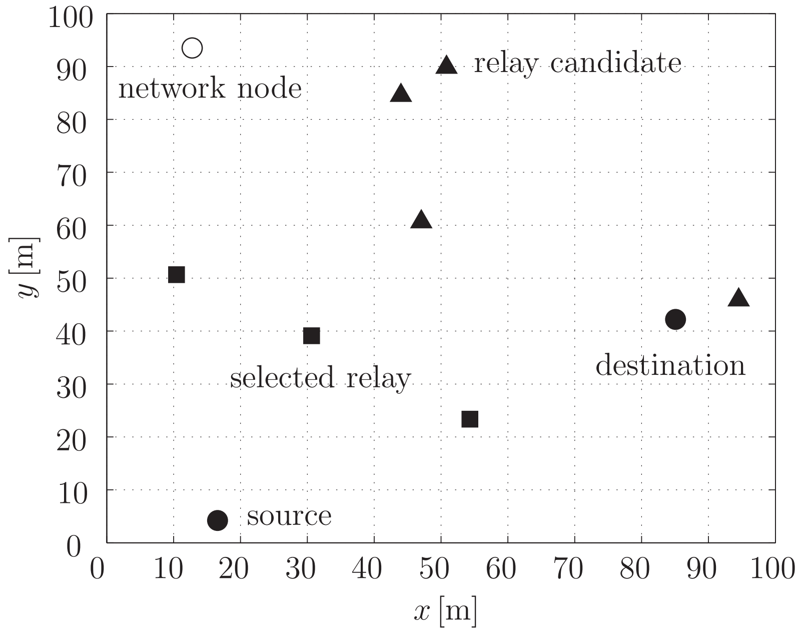

Figure 3.8: Network constellation of example 3 with $\mathbf{x} = [1\,2\,1\,1\,1]^T$ and $\mathsf{BER} = 10^{-4}$.

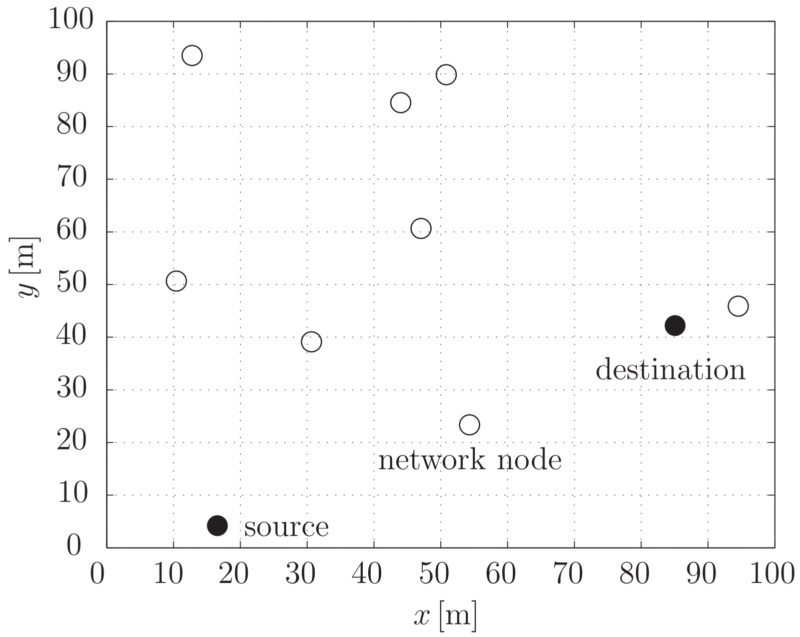

Figure 3.9: Network constellation of example 4 with $\mathbf{x} = [1\,2\,1\,1\,1]^T$ and $\mathsf{BER} = 10^{-2}$.

nodes can safe their energy. This prolongs the life-time of nodes in a network – an issue which is especially important in ad-hoc networks.

3.5 Shortcomings

In spite of having a lot of advantages, there are also some shortcomings of the ARSP. One drawback is an increase of the delay due to the additional packet type `CTScoop`. This packet type is necessary in order to control medium access and to give the source the possibility of creating a sorted list of suitable relay candidates. Data transmission is also performed in a time-division manner, which leads to additional delay times. This can be improved by employing STBC among the selected relay nodes. Information about the codes to use could be sent by the source when it broadcasts the relay table. The delay might be acceptable for rather small networks (only a handful of nodes), but becomes unbearable for large networks (several hundreds of nodes) and networks that cannot tolerate such delay times.

Another drawback is the use of pilot signals in order to measure the achievable BER at the destination. These pilot signals are contained in the `RTS` frame sent by the source. It is quite obvious that the amount of pilot signals increases with a decreasing required BER (this also leads to larger delay times). For practical implementations, this drawback can be weakened by the following action. The achievable BER at the destination is not calculated based on pilot signals sent by the source, but on the received SNR at the destination. The destination “simply” measures the SNR after having received the `RTS` frame and uses a look-up table with stored values in order to decide about the achievable BER.

Further improvements of the protocol performance are possible by applying rate adaptation or power allocation within the network. This can easily be done since the ARSP is a centralized protocol. Last but not least, it must be stated that the major aim of the ARSP is reliability which must be bought by increased delay times. So to speak, the question is: How much delay is acceptable for the sake of higher reliability?

4

Optimal Resource Allocation

> Willingness to compromise with others' ways of living and cooperation in common tasks, these make living happy and fruitful.
>
> *Sri Sathya Sai Baba*

Resource allocation is an important research area in wireless communications. It is evident that any allocation of resources such as frequency, time, or power should be done in a way that user requirements are met. Frequency is perhaps best suited to explain the idea and importance of optimal resource allocation. In the beginning of wireless communications, frequencies have been assigned to users on a permanent basis. However, two problems emerge with such an approach. First, if a user does not use its assigned spectrum, these frequencies cannot be allocated by any other user and, thus, are wasted. Second, performance of a user may be limited by the fact that not enough frequencies have been assigned to him (assuming that transmit power has been fixed). In either case, new emerging and bandwidth demanding multimedia applications lead to a bottleneck in network performance [116]. Obviously, a proper solution would be to assign spectrum to users dynamically. With respect to the aspects mentioned before this means that either temporarily unused spectrum can be allocated to other users or that a user's demand for more bandwidth can be served (see [117–119]). The issue of dynamic spectrum allocation led to new techniques like dynamic channel assignment, spectrum trading, and spectrum pooling to name a few [120–122].

In contrast to spectrum allocation, we focus on the optimization of transmit power and transmission time in a wireless network with Rayleigh fading that consists of one source, one relay, and one destination [123]. Optimization is based on Brent's method[1] (see Section 4.2). The applied optimization criteria are the instantaneous channel capacity and the delay-limited capacity, respectively. Source and relay are equipped with one antenna each. We note that an extension to MIMO terminals is straightforward and methods for this are well-known in literature. The usage of resource allocation implies longer battery lifetime and reduces the interference to other terminals in the network, which is especially important for ad-hoc and sensor networks.

In [124] the authors consider power allocation and use outage probability as optimization criterion. They show that an optimized allocation increases the system performance enormously. This is especially true for networks where the communication links are highly unbalanced with respect to the channel gains or where the number of hops is large. The authors derive another interesting outcome. They show that non-regenerative systems with power allocation achieve better results compared to regenerative systems where no optimized power allocation is applied. Optimization is performed by the use of the Lagrangian[2] multiplier method. This method needs the calculation of derivatives, which is not the case for Brent's method.

In [74] power and time are optimized. The authors consider the delay-limited capacity by assuming partial CSI at the transmitters and full CSI at the receivers. They demonstrate that in a relay network a nonzero delay-limited capacity is achievable in contrast to a network consisting of one source and one destination only. They further introduce an opportunistic transmission protocol where the relay is used depending on the channel gains. This protocol improves the delay-limited capacity enormously. The authors show that this protocol performs close to the cut-set bound.

Another publication that treats outage minimization with CSI at the transmitters is [73]. The authors basically investigate two approaches. In a first approach, the source and the relay have to transmit with constant powers. Both nodes can then ensure coherent summation of their signals at the destination by correcting their initial transmission phases. Furthermore, the correlation between the signal from the source and the signal from the relay can be adjusted in a way to further reduce the outage probability. Second, the source and the relay can adapt their corresponding power values from time slot to time slot. The authors derive a power control policy that shows significant gains over constant power transmission.

In [125] no CSI is assumed, but rather channel distribution information (CDI). It is shown that transmitter cooperation with decode-and-forward outperforms receiver

[1] Richard Brent, ∗ 1946. Australian mathematician and computer scientist. His root-finding algorithm known as Brent's method builds on earlier work by Theodorus Dekker, a Dutch mathematician born in 1927.

[2] Joseph-Louis Lagrange, originally Giuseppe Lodovico Lagrangia, ∗ January 25, 1736, † April 10, 1813. Italian-born mathematician and astronomer. Significant contributions to all fields of analysis, number theory, classical and celestial mechanics. He is one of the 72 honored French scientists whose names were put on plaques at the first stage of the Eiffel Tower. A lunar crater is also named after him.

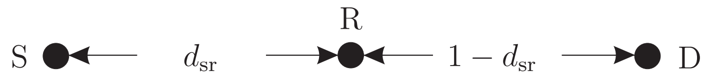

Figure 4.1: One-dimensional network geometry. The distance between source S and destination D is normalized to $d_{\text{sd}} = 1$. Furthermore, $d_{\text{sr}} = 1 - d_{\text{rd}}$.

cooperation and is capacity achieving under the constraint that the average transmit power is the same for all nodes. In contrast to that, if power is allocated optimally among the nodes, then receiver cooperation with compress-and-forward is beneficial to transmitter cooperation. The authors also examine the effects of large clusters on the performance, i.e., $K+1$ cooperating nodes either at the transmitter or at the receiver.[3] In particular, in a static channel, both cooperation schemes, i.e., transmitter cooperation without CSIT or receiver cooperation where power is allocated equally, provide no capacity gains. In a fading channel, however, a constant capacity gain can be achieved.

In [126] outage regions for energy-constrained multi-hop and adaptive multi-route networks with an arbitrary number of relay nodes are investigated. The authors derive optimal power allocation strategies in a sense that outage probability is minimized (depending on the distances between the nodes). Moreover, the metric of rate gain was introduced and it was shown that a combined strategy of direct transmission and adaptive multi-route outperforms multi-hop networks for all values of target rate R. It is stated that cooperation strategies are beneficial for low-rate systems where the main objective is a very low outage probability. The notion of outage region is also used in [127], where different network models either with repetition coding or parallel channel coding are examined.

More information about related work in the field of resource allocation in wireless relay networks can be found in [128–131] and the references therein.

In the following, we use a common path loss model, where the relation between the channel variances σ_i^2 and the distance d_i between two nodes is given by $\sigma_i^2 \propto d_i^{-\alpha}$, where α denotes the path loss exponent and $i \in \{\text{sd,sr,rd}\}$. We assume a one-dimensional network geometry, where the distance between source and destination is normalized to 1. The relay is placed on a straight line between source and destination. Accordingly, $d_{\text{sr}} = 1 - d_{\text{rd}}$ (see Fig. 4.1). We get $\sigma_{\text{sd}}^2 = 1, \sigma_{\text{sr}}^2 = d_{\text{sr}}^{-\alpha}$, $\sigma_{\text{rd}}^2 = (1 - d_{\text{sr}})^{-\alpha}$. On each channel, white Gaussian noise is added. Noise realizations are modeled as mutually independent, circularly-symmetric, complex Gaussian random variables with zero mean and variance 1. A network realization is described by the triple $\mathbf{h} = (|h_{\text{sd}}|^2, |h_{\text{sr}}|^2, |h_{\text{rd}}|^2)$. We assume full CSI at the receivers and partial CSI at the transmitters as it is the case in [74]. The relay operates in a half-duplex mode and uses decode-and-forward. For decode-and-forward the knowledge of CSI is of great importance in relay networks. If the source does not know the channel gain between

[3]The number of cooperating nodes was changed from M to $K+1$ to keep the thesis concise.

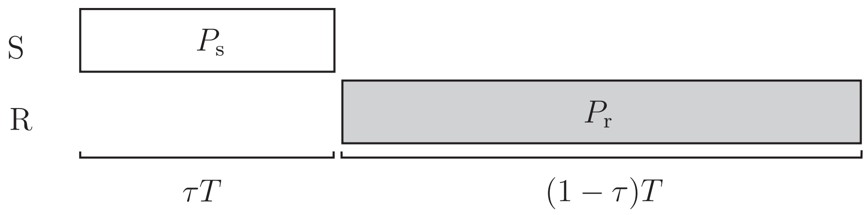

Figure 4.2: Basic transmission scheme. The first phase of duration τT is reserved for source transmission with power P_s and the second phase of duration $(1-\tau)T$ is reserved for relay transmission with power P_r.

itself and the relay, its only possibility is to transmit with a fixed transmit power (and, thus, with a fixed rate). However, if the instantaneous channel capacity falls below the source's transmission rate, reliable decoding at the relay cannot be ensured anymore. In this case, the relay decides not to cooperate in order to not waste any system resources. However, if CSI is available at the source, an adaptive allocation of power and time can, at least in accordance to a given power constraint, ensure reliable decoding at the relay. An allocation strategy $(\mathbf{P},\boldsymbol{\tau})$ is described by the power allocation vector $\mathbf{P} = (P_s,P_r)$ and the time allocation vector $\boldsymbol{\tau} = (\tau,1-\tau)$, where P_s denotes the source transmit power, P_r is the relay transmit power, and $\tau \in (0,1]$ denotes the time fraction used for source transmission, i.e., the time fraction $1-\tau$ is used for relay transmission. The length of one transmission block is T and, hence, the source transmits for the duration of τT and the relay transmits for the duration of $(1-\tau)T$. The basic transmission scheme is illustrated in Fig. 4.2.

4.1 Optimization Problem

Due to cooperation and taking CSI into account, it is possible to allocate network resources among source and relay in a way to optimize a certain criterion. In our case the design criteria are the instantaneous channel capacity and the delay-limited capacity. The position of the relay d_{sr} and the path loss exponent α are constant parameters of the optimization problem. The overall transmit power P_{tot} is given by

$$P_{tot} := \tau P_s + (1-\tau)P_r. \tag{4.1}$$

The overall aim is to allocate resources in a way that the capacity is maximized. Therefore,

$$\mathcal{C}^*(\mathbf{h}) = \max_{\tau} \max_{P_s} \{\mathcal{C}(P_{tot},\mathbf{h},\tau) : \tau P_s + (1-\tau)P_r = P_{tot}\} \tag{4.2}$$

$$\text{subject to} = \begin{cases} P_s \in [0,P_{tot}/\tau] \\ \tau \in (0,1] \end{cases}.$$

The optimization algorithm searches for the pair (P^*,τ^*) with respect to the source which maximizes capacity. Both values P^* as well as τ^* are fractions relative to the overall transmit power and the length of a transmission block, respectively. Summarized, the overall transmit power is kept constant and the optimal power allocation is found iteratively for different values of τ. Clearly, the step size of τ is critical for this step. If it is too small, the optimization takes too long. However, if it is too big, the result of the optimization may not well approximate the global optimum.

4.2 Optimization Algorithm

4.2.1 Description

For the optimization of power and time allocation, we use an algorithm based on Brent's method [132, 133]. Brent's method is a root-finding algorithm in numerical analysis which has the advantage that it does not require any derivatives of functions. The main idea behind Brent's method is to combine the secant method, the bisection method, and inverse quadratic interpolation. It is sometimes known as the van Wijngaarden[4]-Dekker-Brent method. If possible the secant method or inverse quadratic interpolation will be used, because these methods converge very fast. However, these methods are less reliable than the robust bisection method. As a consequence, whenever necessary, the algorithm selects the bisection method to increase reliability of convergence. Brent's method (or variations of it) is implemented in a lot of mathematical tool kits like Mathematica and Matlab.

We use a similar algorithm that combines golden section search (see Subsection 4.2.2) and parabolic interpolation (see Subsection 4.2.3). This leads to a robust optimization algorithm. Whenever possible we apply parabolic interpolation which converges faster than the golden section search. In cases where reliability is questionable, we aid stability by switching to the golden section search. For the usage of such an algorithm, several requirements have to be met [123]:

- The function that is optimized must be continuous with respect to the optimization variable.
- The function has to be unimodal in order to be able to find the extreme value. If there are more than one extreme values, then only one extreme value will be found. However, this need not to be the global optimum.
- Optimization can only be done with respect to one variable.

We will see later that these requirements are met when we maximize capacity and that the applied optimization algorithm produces reliable results. In the next subsections, we give a short survey on the principles of golden section search and parabolic interpolation. For more information, the reader is referred to the vast literature on

[4] Adriaan van Wijngaarden, ∗ November 2, 1916, † February 7, 1987. Dutch mathematician and computer scientist. He is regarded as the founding father of computer science in the Netherlands.

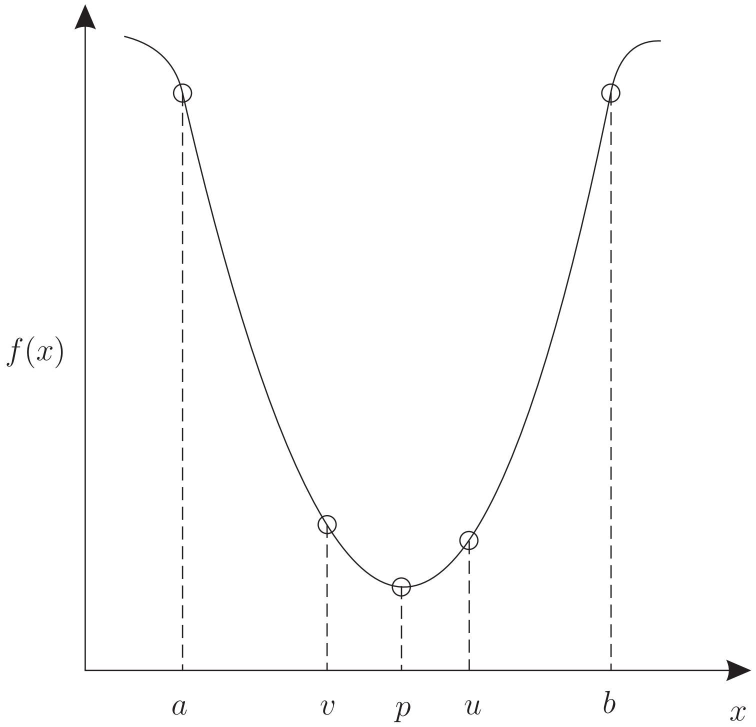

Figure 4.3: Golden section search.

that subjects in mathematics and computer science, especially [134] which gives a good explanation of the Brent's method.

4.2.2 Golden Section Search

Consider a unimodal function on the closed (bracketing) interval $[a,b]$. As a consequence, there exists exactly one $p \in [a,b]$ so that the function is decreasing in $[a,p)$ and increasing in $(p,b]$ (or vice versa). Let us assume in the following, that a function f has a minimum in the interval $[a,b]$ (see Fig. 4.3).

In order to find this minimum, we choose a point $v \in [a,b]$ and evaluate the function at this point. In the next step, the bracketing interval is $[v,b]$, and we divide this interval by choosing a new trial point u. Now, if $f(u) < f(v)$, the new bracketing interval will be $[u,b]$. Otherwise, if $f(u) > f(v)$, the new bracketing interval is $[a,u]$. The question is how trial points v and u are chosen. Let us assume that our first trial point v is determined in a way that the interval $[a,v]$ is a fraction c of the interval $[a,b]$,

$$c = \frac{v-a}{b-a} \quad \text{and} \quad 1-c = \frac{b-v}{b-a}, \tag{4.3}$$

and that the trial point u is determined in a way that the interval $[v,u]$ is an additional fraction d beyond v,

$$d = \frac{u-v}{b-a}. \tag{4.4}$$

The next bracketing interval is then either $c + d$ or $1 - c$ (in relation to the original one). If we equal both, so that we minimize the worst case possibility, we get

$$d = 1 - 2c. \tag{4.5}$$

In doing so, the point u becomes the point symmetric to v in the interval $[a,b]$, therefore, $|v - a| = |b - u|$. We see that this is only true, if u lies in the larger of the two intervals $[a,v]$ and $[v,b]$ (which means that $c < 0.5$). Since the algorithm of finding the minimum is iterative, the trial points v and u are found by applying the same strategy and, hence, $[v,u]$ should be the same fraction in $[v,b]$ as $[a,v]$ was in $[a,b]$. We get

$$\frac{d}{1-c} = c. \tag{4.6}$$

Combining (4.5) and (4.6) yields

$$c^2 - 3c + 1 = 0 \tag{4.7}$$

with the solution

$$c = \frac{3-\sqrt{5}}{2} \approx 0.38197. \tag{4.8}$$

This value is related to the golden ratio ϕ by

$$c = \frac{1}{\phi^2}, \tag{4.9}$$

with

$$\phi = \frac{1+\sqrt{5}}{2} \approx 1.61803. \tag{4.10}$$

Sometimes the golden ratio is considered to be $(\sqrt{5}-1)/2 \approx 0.61803$. Clearly, this is the multiplicative inverse of ϕ.

The great advantage of the golden section search is that it reliably finds the optimum of a unimodal function, even if this function behaves “uncooperatively” (which means that it progresses unsteadily).

4.2.3 Parabolic Interpolation

Parabolic interpolation converges much faster to an optimum than does golden section search. Whenever a function does not behave “uncooperatively,” which is the generic case for smooth functions, it is better to approximate the function by a parabola that brings us close to the optimum. Consider three points $(u,f(u))$, $(v,f(v))$, and $(w,f(w))$ on a function's f graph (see Fig. 4.4). These three points surely define a parabola (dashed line). The minimum x^* of that parabola can then be expressed as[5]

$$x^* = v - \frac{1}{2}\frac{(v-u)^2(f(v)-f(w)) - (v-w)^2(f(v)-f(u))}{(v-u)(f(v)-f(w)) - (v-w)(f(v)-f(u))}. \tag{4.11}$$

[5] Of course, this formula is only true if all three points are not collinear. If the points are collinear, the denominator will become zero.

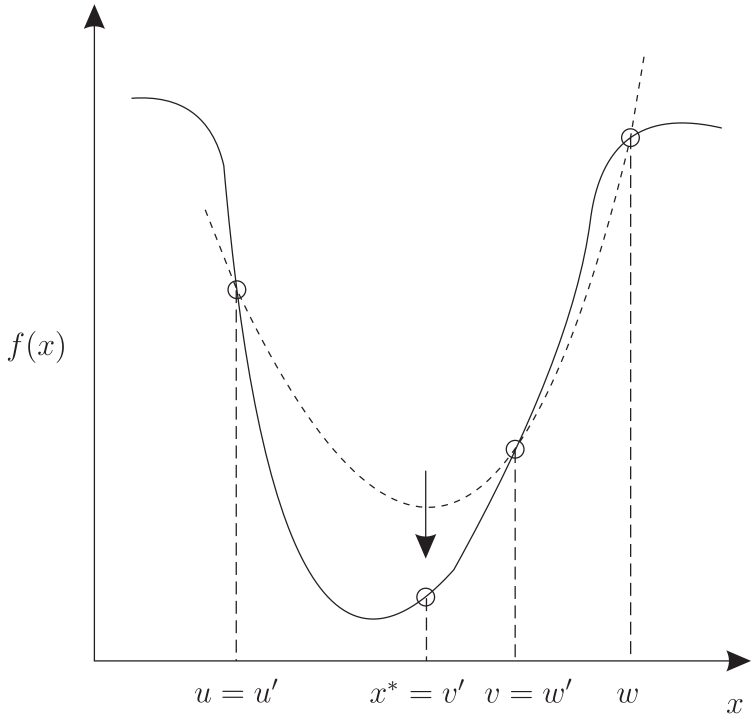

Figure 4.4: Parabolic interpolation.

The minimum of the parabola is determined and the minimum's abscissa is called x^*. In a next step the point u remains as u', the point v becomes the new w', and x^* becomes v'. A new parabola is drawn through the points $(u',f(u'))$, $(v',f(v'))$, and $(w',f(w'))$. Then, the minimum of the new parabola is evaluated and so on.

4.2.4 Example

We now give an example which demonstrates the operating mode of the optimization algorithm. The function

$$-f(x) = -\frac{1}{2}\min\{\log_2(1+rx), \log_2(1+qx) + \log_2(1+sz)\} \tag{4.12}$$

is minimized with respect to the variable x. As we will see later, $f(x)$ shows similarities to the instantaneous capacity of a multi-route relay network with decode-and-forward. We choose $q = 1$, $r = s = 8$, and $z = 2 - x$.[6] Fig. 4.5 illustrates the function and shows the first 5 iterations. It can easily be seen that $-f(x)$ is unimodal and that there exists exactly one minimum. The point 1 is the initial point that has been chosen with respect to the golden ratio. The points 2, 3 have also been found by applying golden section search. The points 4 and 5 have been chosen by the use of parabolic interpolation. The algorithm stops if the alteration of the functional value

[6] These values correspond to a scenario where the relay is placed half-way between source and destination, the path loss exponent is set to 3, and the time fraction equals 0.5.

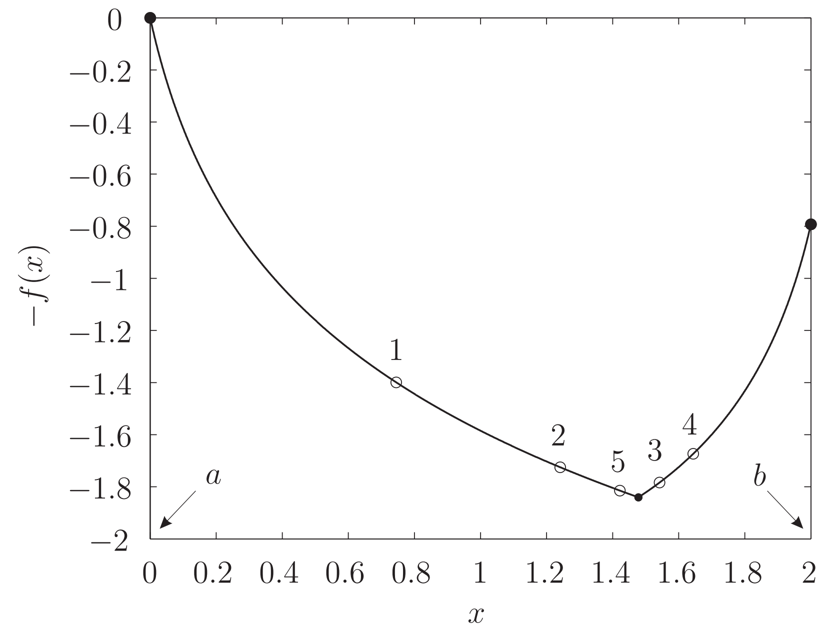

Figure 4.5: Operating mode and the first five iteration steps of the optimization algorithm on the interval $[a,b] = [0,2]$. Point 1 is the initial value. Points 2 and 3 have been found by golden section search, whereas points 4 and 5 have been found by parabolic interpolation (cf. Tab. 4.1).

of f between two subsequent iterations is lower than 10^{-4}. This takes 19 iterations for the considered example (see Tab. 4.1). Eventually, the minimum is evaluated as $-f(x) = -1.840419690427235$, which is achieved by $x = 1.478072174497307$.

4.3 Results

4.3.1 Instantaneous Channel Capacity

In this subsection, we examine the instantaneous channel capacity of multi-route and multi-hop networks when power and time are allocated optimally. We also introduce a figure of merit to compare the benefits of cooperation to direct transmission. We will see that cooperation does not always outperform direct transmission, even if resource allocation is applied. This leads to the definition of new "selective" protocols, where cooperation is only used if it performs better than direct transmission. This kind of protocols was also used in [74], where the authors call them "opportunistic."

The instantaneous channel capacity[7] of direct transmission clearly is[8]

$$\mathcal{C}_{\mathrm{DT}} = \log_2(1 + |h_{\mathrm{sd}}|^2 P_{\mathrm{s}}). \tag{4.13}$$

[7] Recall that we use the normalized channel capacity $\mathcal{C} = C/B$ (see 1.7).

[8] As stated in the beginning of Chapter 4, noise realizations have variance 1.

Table 4.1: Iterations of the optimization algorithm (cf. Fig. 4.5).

#	x	$-f(x)$	procedure
1	0.763932	−1.41507	`initial`
2	1.23607	−1.72237	`golden`
3	1.52786	−1.79703	`golden`
4	1.62936	−1.69104	`parabolic`
5	1.42066	−1.81411	`parabolic`
6	1.43883	−1.82254	`parabolic`
7	1.46288	−1.83355	`parabolic`
8	1.4877	−1.83243	`golden`
9	1.47309	−1.83817	`parabolic`
10	1.47463	−1.83887	`parabolic`
11	1.47962	−1.83917	`golden`
12	1.47763	−1.84022	`parabolic`
13	1.47728	−1.84006	`parabolic`
14	1.47799	−1.84038	`parabolic`
15	1.47861	−1.84	`golden`
16	1.47823	−1.84032	`golden`
17	1.47803	−1.8404	`parabolic`
18	1.47807	−1.84042	`parabolic`
19	1.47813	−1.8404	`golden`

This expression is of course maximized if we choose $P_{\mathrm{s}} = P_{\mathrm{tot}}$, which simply means that the source transmits with all the available power.

In order to derive the capacity expression for a multi-hop network, we have to apply the cut-set bound (max-flow min-cut theorem [52]). In a first block of duration τT the source sends its message to the relay. After decoding and reencoding, the relay sends its message to the destination in a second block of duration $(1-\tau)T$. The instantaneous channel capacity becomes

$$\mathcal{C}_{\mathrm{MH}} = \min\{\tau \log_2(1+|h_{\mathrm{sr}}|^2 P_{\mathrm{s}}),(1-\tau)\log_2(1+|h_{\mathrm{rd}}|^2 P_{\mathrm{r}})\}. \tag{4.14}$$

Its maximum for $\tau = 1/2$ is clearly limited by the weakest channel in the network, which is intuitively clear and shown mathematically in [75, ch. 2.3]. Since the destination only receives a message from the relay, it only has to know the codebook used by the relay. This weakens the requirements for a priori knowledge for the destination compared to multi-route networks (see below). As it is possible that direct transmission outperforms multi-hop networks, especially when one link in the multi-hop network is weak, it may be desirable to choose between both protocols. We call this protocol selective multi-hop. It has the great advantage that it does not always use cooperation, which saves the relay's resources. Capacity becomes

$$\mathcal{C}_{\text{sMH}} = \max\{\mathcal{C}_{\text{MH}}, \mathcal{C}_{\text{DT}}\}. \tag{4.15}$$

Next, we consider a multi-route network, where there is a direct link between source and destination. The source sends its message in the first block of duration τT to both the relay as well as the destination. During the second block of duration $(1-\tau)T$ the relay transmits to the destination. Source and relay apply parallel channel coding, which means that they use independently generated Gaussian codebooks. This leads to an accumulation of capacity. Another possibility is repetition coding, where both terminals use the same codebook. This has the advantage that the destination only has to know one codebook, however, repetition coding only leads to an accumulation of SNR.[9] The instantaneous channel capacity with parallel channel coding can be expressed as

$$\mathcal{C}_{\text{MR}} = \min\{\tau \log_2(1+|h_{\text{sr}}|^2 P_{\text{s}}), \tau \log_2(1+|h_{\text{sd}}|^2 P_{\text{s}}) + (1-\tau)\log_2(1+|h_{\text{rd}}|^2 P_{\text{r}})\}, \tag{4.16}$$

where the first expression in the min-function denotes the maximal rate at which the relay can decode the source signal and the second expression describes the maximal rate at which the destination can decode the source and the relay signal. Since power is divided between source and relay, the source cannot transmit with P_{tot}. However, due to the reduced transmit power of the source, deep fades on the source-relay link cannot be avoided generally. In this case, it might be beneficial to use direct transmission rather than cooperation. We call this protocol selective multi-route. Its instantaneous channel capacity is

$$\mathcal{C}_{\text{sMR}} = \max\{\mathcal{C}_{\text{MR}}, \mathcal{C}_{\text{DT}}\}. \tag{4.17}$$

Our overall aim is to maximize the instantaneous channel capacity of selective multi-hop and selective multi-route networks through power and time allocation. Accordingly,

$$\mathcal{C}^*_{\text{sMH}} = \max_{\tau} \max_{P_{\text{s}}} \mathcal{C}_{\text{sMH}} \quad \text{and} \quad \mathcal{C}^*_{\text{sMR}} = \max_{\tau} \max_{P_{\text{s}}} \mathcal{C}_{\text{sMR}}. \tag{4.18}$$

The instantaneous channel capacities with optimal resource allocation for multi-hop (MH) and multi-route (MR) networks in bit/s/Hz vs. the overall available transmit power P_{tot} in dB are illustrated in Fig. 4.6.[10] The distance between source and relay is $d_{\text{sr}} = 0.3$. Direct transmission (DT) is shown as reference case. It can be seen that for $P_{\text{tot}} = -10 \ldots 0$ dB, MR and MH have almost the same capacity. For values $P_{\text{tot}} > 0$ dB, the gap between MR and MH increases with increasing overall power. The reason for this is, that MR creates diversity at the destination in contrast to MH. At $P_{\text{tot}} \approx 11$ dB, DT and MH intersect. From that value on, DT outperforms MH, which highlights the great benefits of selective protocols.

[9] In general, parallel channel coding outperforms repetition coding. However, for low SNR values, it can be shown that parallel channel coding reduces to repetition coding. For more information see Chapter 6.

[10] Recall that the variance of the noise is equal to 1. Therefore, the SNR is simply given by P_{tot}. The value $P_{\text{tot}} = 0$ dB describes the fact that the overall transmit power equals the noise power.

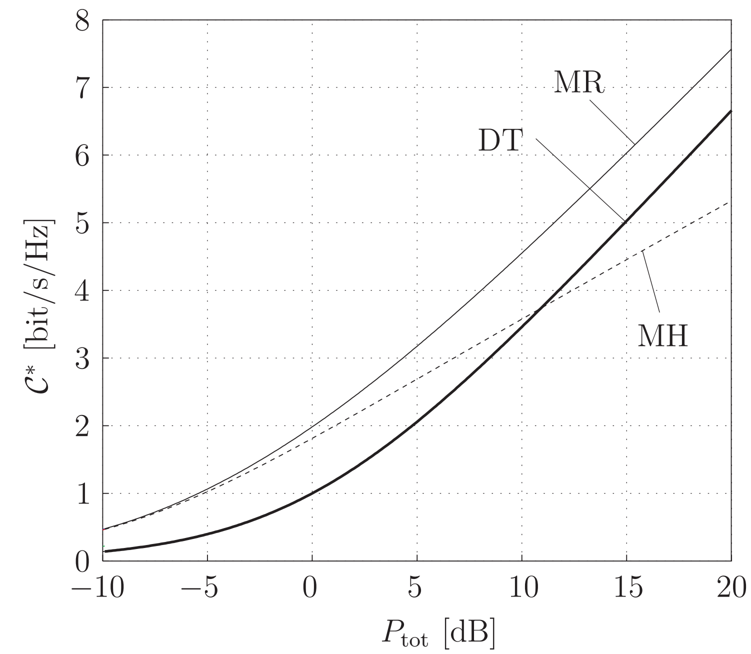

Figure 4.6: Instantaneous channel capacities with optimal resource allocation for multi-hop (MH) and multi-route (MR) in bit/s/Hz. The distance source-relay has been set to $d_{sr} = 0.3$. Capacity of direct transmission (DT) is illustrated as reference.

As mentioned before, Fig. 4.6 shows capacity results for optimal resource allocation. However, what were the optimal values for transmit power and transmit time in that case? The optimal power allocation for the source for $d_{sr} = 0.3$ is depicted in Fig. 4.7. Optimal power allocation P^* already includes optimal time allocation τ^* and is given by

$$P^* = \tau^* \frac{P_s}{P_{tot}}. \tag{4.19}$$

Power allocation is a monotonically increasing function in P_{tot} for $d_{sr} = 0.3$ (the unsteady course is due to simulations). Hence, the more power available, the more power is (relatively) allocated to the source. It is important to notice that this is not generally the case. For $d_{sr} = 0.8$, for instance, power allocation decreases with increasing P_{tot}. We see that there is always more power allocated to the source in the case of MR compared to MH. This becomes intuitively clear. Since there is no direct link between source and destination for the case of MH and the source only has to transmit its message to the relay (which is located *between* source and destination in our system model), less power can be allocated to the source. This is not true for MR. Here, diversity is created at the destination due to the fact that the destination receives the source message from the source directly and an estimated version of the source message from the relay. The larger distance between source and destination requires that more power is allocated to the source.

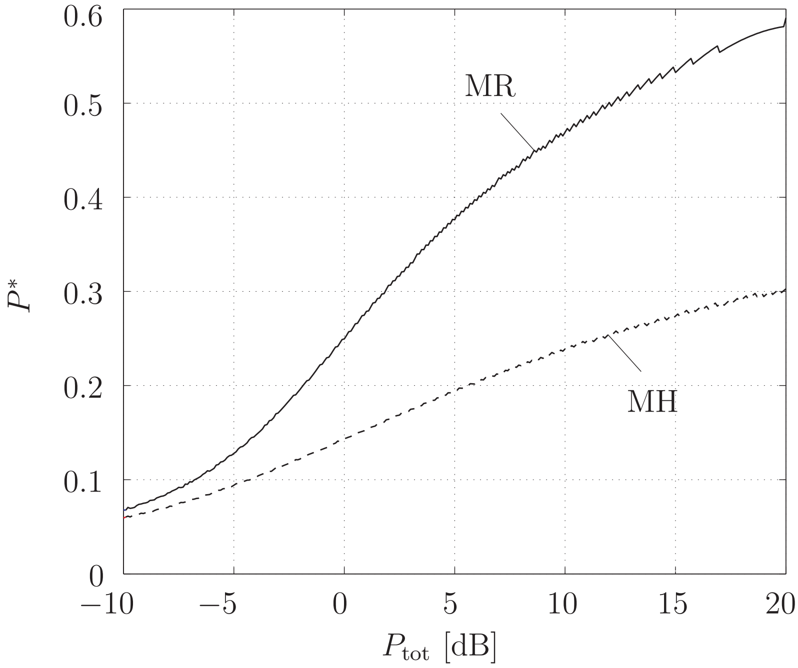

Figure 4.7: Optimal power allocation $P^* = \tau^* P_\mathrm{s}/P_\mathrm{tot}$ for the source. The distance source-relay has been set to $d_\mathrm{sr} = 0.3$.

Fig. 4.8 shows the optimal time allocation τ^* for the source. Up to a value of $P_\mathrm{tot} \approx 5$ dB, there is little difference between the time allocation for MR and MH. At a value of $P_\mathrm{tot} = 2.5$ dB, both curves intersect and time allocation for MR is larger than that for MH. Both curves increase with increasing P_tot when d_sr is chosen to be 0.3 (again, the unsteady course of the curve – especially for values of $P_\mathrm{tot} = -10 \ldots -5$ dB – is due to simulations). However, this is also not generally true. For $d_\mathrm{sr} = 0.8$, time allocation for MR shows a parabolic behavior in the range from -10 dB to 20 dB. For MH, time allocation is decreasing in this case.

Until now, we had a look at the instantaneous channel capacities of DT, MH and MR. In order to see the gains that cooperation achieves over DT in terms of capacity, we define the capacity gain as

$$G_{\mathcal{C},l}(P_\mathrm{tot}, d_\mathrm{sr}) := 10 \log_{10} \left(\frac{\mathcal{C}_l^*(P_\mathrm{tot}, d_\mathrm{sr})}{\mathcal{C}_\mathrm{DT}(P_\mathrm{tot})} \right) \quad [\mathrm{dB}], \tag{4.20}$$

where $l \in \{\mathrm{MR}, \mathrm{sMR}, \mathrm{MH}, \mathrm{sMH}\}$. It is obvious that for $l \in \{\mathrm{sMR}, \mathrm{sMH}\}$, the capacity gain becomes $G_{\mathcal{C},l}(P_\mathrm{tot}, d_\mathrm{sr}) \geq 0$ dB. The results are depicted in Fig. 4.9. Let us first have a look at the curves where $P_\mathrm{tot} = 1$ dB and $d_\mathrm{sr} = 0.3$, since we have considered power and time allocation for this source-relay distance previously. For MH a capacity gain of approximately 2.3 dB can be achieved and for MR we achieve a gain of 2.7 dB. For the case of MH, power allocation is approximately $P^* = 0.15$ and for MR we have $P^* = 0.28$ (see Fig. 4.7). In both cases, the optimal time allocation is $\tau^* \approx 0.31$

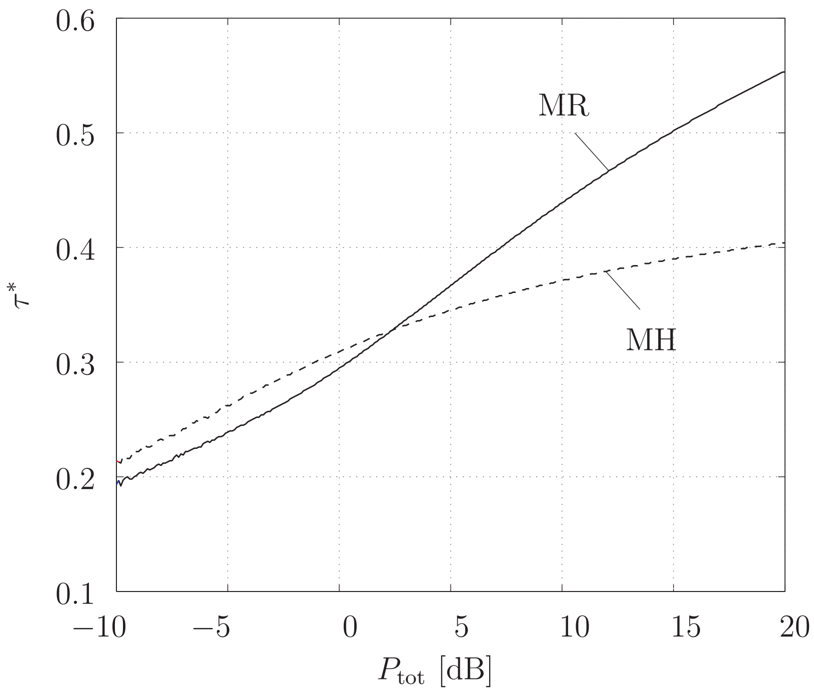

Figure 4.8: Optimal time allocation τ^* for the source. The distance source-relay has been set to $d_{\text{sr}} = 0.3$.

(see Fig. 4.7). Next, consider the curves where $P_{\text{tot}} = 10$ dB. Generally, we can state that the achieved gains are less compared to those where $P_{\text{tot}} = 1$ dB. It can even be seen that for $d_{\text{sr}} \in (0{,}0.2)$ and $d_{\text{sr}} \in (0.8{,}1)$, respectively, DT outperforms MH. Generally spoken, MR always performs better than MH. This is due to the fact that MR – as explained before – creates diversity at the destination which is not the case for MH. Additionally, all curves are symmetric to $d_{\text{sr}} = 0.5$ and we can conclude that cooperation with resource allocation is especially preferable if the distances between source and relay as well as relay and destination are almost the same. All in all, we demonstrated that dependent on the relay location and the overall transmit power P_{tot}, MR and MH achieve remarkable gains in comparison to DT. However, for high values of P_{tot}, there are relay locations where DT outperforms MH. Generally, we can state that capacity gains increase with decreasing overall system power. This clearly shows that relaying is beneficial for low overall transmission powers [123].

4.3.2 Delay-limited Capacity

We have examined the instantaneous channel capacity in the previous subsection. However, as the name already states, this capacity expression is not suitable if we want to make statements about the performance of transmission over a channel in average. If we further want to deal with delay-constrained applications, such as voice or video, where delays are tolerable only to a certain degree, the ergodic capacity is not suitable

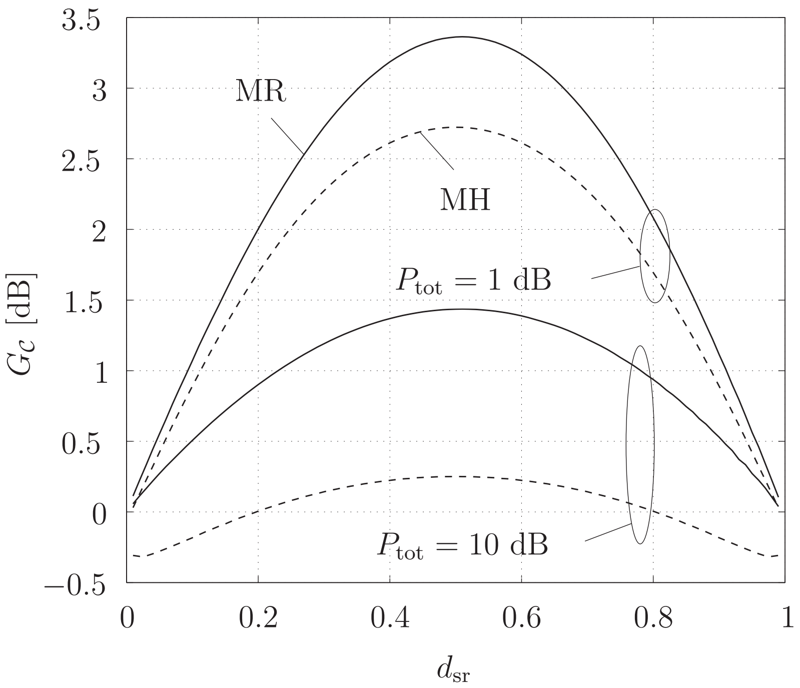

Figure 4.9: Capacity gain $G_{\mathcal{C}}$ of multi-route and multi-hop over direct transmission when optimized power allocation P^* and optimized time allocation τ^* are used.

anymore. For that purpose, the delay-limited capacity can be used. The delay-limited capacity is the channel capacity of a quasi-static fading channel when an outage probability of $p_{\text{out}} = \epsilon = 0$ is required.[11] Hence, it describes the maximal transmission rate that can be achieved for each network realization $\mathbf{h} = (|h_{\text{sd}}|^2,|h_{\text{sr}}|^2,|h_{\text{rd}}|^2)$ [58].

We are especially interested in the delay-limited capacity that can be achieved with a given average transmit power constraint. If we want to make sure that a capacity $\mathcal{C}_0$ is achievable for every channel realization, a certain minimal transmit power $P_{\text{tot}}(\mathcal{C}_0,\mathbf{h})$ is necessary. This is illustrated for a delay-limited capacity of $\mathcal{C}_0 = 3$ bit/s/Hz in Fig. 4.10. A selective multi-route cooperation scheme has been used for the simulation and the parameters were set to $d_{\text{sr}} = 0.2$ and $\alpha = 4$. We see 20 different realizations and the corresponding values of P_{tot} in dB that are required in order to achieve the given $\mathcal{C}_0$.

In a next step, we average the values for the required transmit powers over all channel realizations and get

$$\overline{P}_{\text{tot}}(\mathcal{C}_0) = \mathbb{E}_{\mathbf{h}}(P_{\text{tot}}(\mathcal{C}_0,\mathbf{h})). \tag{4.21}$$

By doing so, we are able to derive the delay-limited capacities for different parameter settings. In particular, Fig. 4.11 depicts the delay-limited capacity over the averaged

[11]Recall that the delay-limited capacity is a special case of the ϵ-outage capacity.

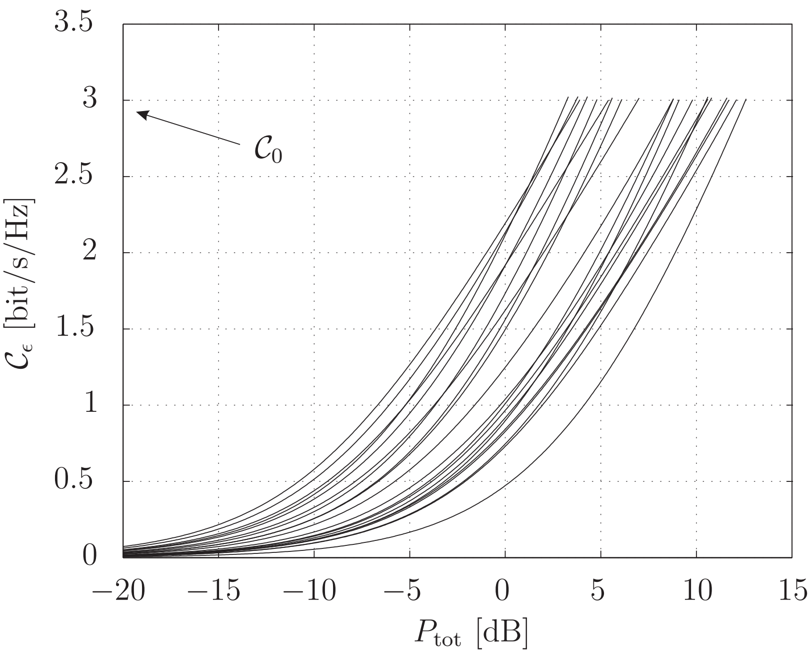

Figure 4.10: Delay-limited capacity vs. total transmit power for selective multi-route and the parameters $d_{\text{sr}} = 0.2$ and $\alpha = 4$.

total transmit power for different relay locations. For the simulations, we considered 10000 channel realizations. It can be seen that the capacity depends on the location of the relay. This leads to two different interpretations of the figure. Either one fixes the averaged total transmit power or the delay-limited capacity. The first viewpoint leads to a capacity gain if the relay is moved from the source towards the middle of the source-destination distance. If the relay is placed half-way between source and destination, i.e., $d_{\text{sr}} = 0.5$, the capacity gain becomes maximal. If the source-relay distance is further increased, the capacity gain decreases. For instance, consider an average total transmit power of $\overline{P}_{\text{tot}} = 5$ dB. Then the maximal delay-limited capacity is $\mathcal{C}_\epsilon \approx 2.6$ bit/s/Hz. Whereas, if we place the relay at $d_{\text{sr}} = 0.1$, the delay-limited capacity is $\mathcal{C}_\epsilon \approx 1.9$ bit/s/Hz. The ladder viewpoint leads to power savings. Take a value of $\mathcal{C}_\epsilon = 3$ bit/s/Hz. If the relay is placed at a distance of $d_{\text{sr}} = 0.1$, then an average total transmit power of $\overline{P}_{\text{tot}} \approx 9.5$ dB is required. However, if the relay is located at $d_{\text{sr}} = 0.5$, approximately 2.5 dB can be saved and an average total transmit power of $\overline{P}_{\text{tot}} \approx 7$ dB is necessary.

Since we are still interested in an optimal resource allocation, the total transmit power and the total transmission time in order to guarantee a certain delay-limited capacity must be distributed over the source and the relay in an appropriate fashion. Therefore, the allocation strategy is also averaged over all channel realizations. In case of power allocation, this can be expressed mathematically as

$$\mathbb{E}_{\mathbf{h}}(P^*) = \mathbb{E}_{\mathbf{h}}\left(\frac{\tau^* P_{\text{s}}}{P_{\text{tot}}}\right) \tag{4.22}$$

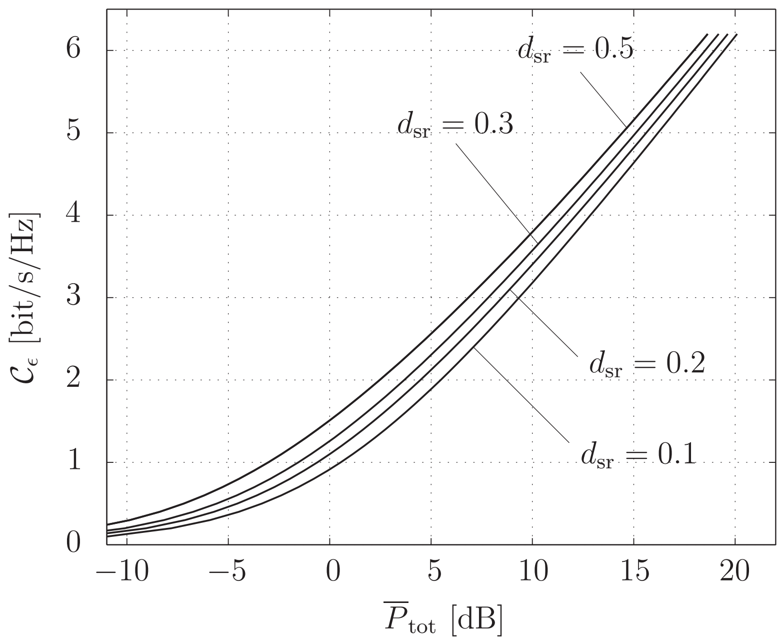

Figure 4.11: Delay-limited capacity vs. averaged total transmit power for different relay locations.

for the source and as

$$\mathbb{E}_{\mathbf{h}}(1 - P^*) = \mathbb{E}_{\mathbf{h}}\left(\frac{(1-\tau^*)P_r}{P_{tot}}\right) \tag{4.23}$$

for the relay. Fig. 4.12 shows the averaged optimal power allocation for the relay vs. the averaged total transmit power for two different relay locations ($d_{sr} = 0.2$ and $d_{sr} = 0.5$). For the simulations we again used 10000 channel realizations. Noteworthy, that even if only 1000 channel realizations are used, the results match pretty nicely. It is intuitively clear that the curve for $d_{sr} = 0.5$ is below the curve for $d_{sr} = 0.2$, since then the relay itself requires less power in order to transmit to the destination.

In some applications, however, the available transmit power may not be enough to support the target delay-limited capacity (and, thus, to provide an outage probability of $p_{out} = 0$). Depending on the application, we may allow some amount of outage events. It is then obviously preferable to not transmit at all in such cases. This leads to power savings and the averaged total transmit power is minimized at the expense of an increased outage probability. For instance, assume that a value of $\mathcal{C}_0 = 1$ bit/s/Hz has to be provided. It is clear that in this case the average transmit power is decreased once we increase the outage probability. This is shown for the selective multi-route protocol in Fig. 4.13 for two different relay locations ($d_{sr} = 0.2$ and $d_{sr} = 0.6$). We averaged over 1000 channel realizations for the simulations and set the path loss exponent to $\alpha = 4$. It can be seen that for a relay location of $d_{sr} = 0.6$, a much lower outage probability can be achieved compared to the relay location of $d_{sr} = 0.2$. This is in line with the conclusions we could draw from Fig. 4.11, where we examined

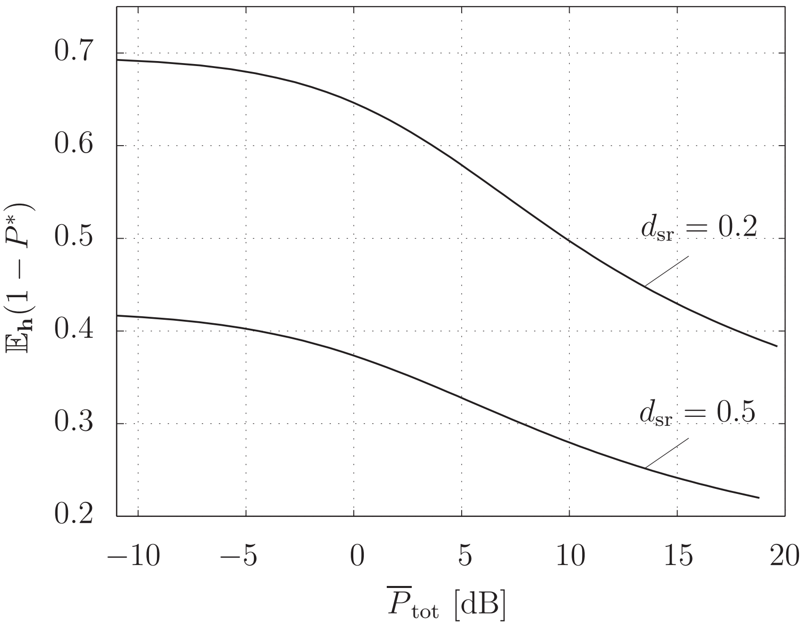

Figure 4.12: Averaged optimal power allocation for the relay vs. averaged total transmit power for $d_{sr} = 0.2$ and $d_{sr} = 0.5$.

the issue of delay-limited capacity gain dependent on different relay locations. For an outage probability of $p_{out} = 10^{-2}$, for instance, approximately 1.7 dB of total averaged transmit power can be saved.

Lastly, the great potential of optimal resource allocation (power and time) can be seen in Fig. 4.14, where we depicted the outage probabilities of selective multi-route (sMR) and direct transmission (DT) vs. the averaged total transmit power for $\mathcal{C}_0 = 1$ bit/s/Hz, $d_{sr} = 0.2$, and $\alpha = 4$. Noteworthy, the curve for direct transmission is not simulated, but directly drawn from analysis. With (4.13) and Rayleigh fading, we have

$$\begin{aligned} p_{out} &= \Pr(\mathcal{C}_{DT} < \mathcal{C}_0) \\ &= \Pr\left(|h_{sd}|^2 < \frac{2^{\mathcal{C}_0}-1}{P_s}\right) \\ &= 1-\exp\left(-\frac{2^{\mathcal{C}_0}-1}{\sigma_{sd}^2 P_s}\right). \end{aligned} \tag{4.24}$$

If we again consider an outage probability of $p_{out} = 10^{-2}$, we see that enormous power savings can be achieved of up to 22 dB. These savings increase even more, when the target outage probability is further decreased. These simulation results match perfectly with those presented in [74].

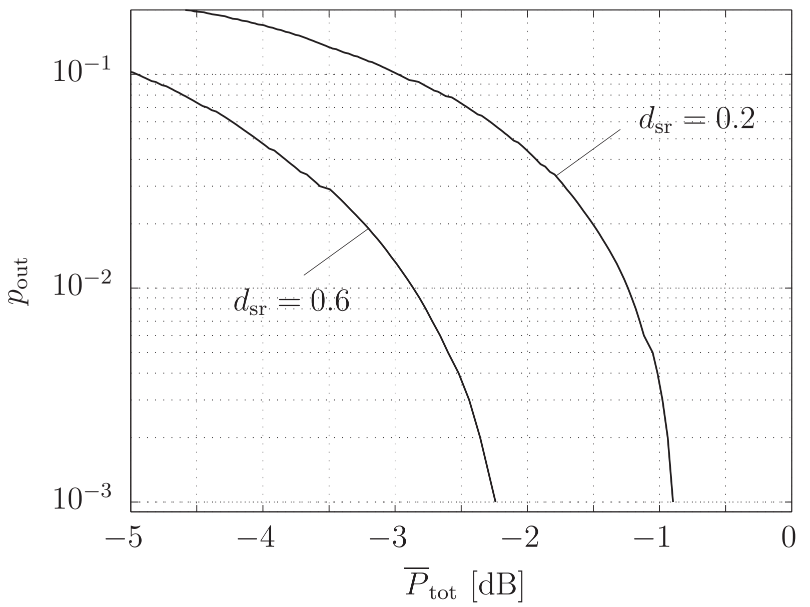

Figure 4.13: Outage probability vs. averaged total transmit power for $d_{\text{sr}} = 0.2$ and $d_{\text{sr}} = 0.6$.

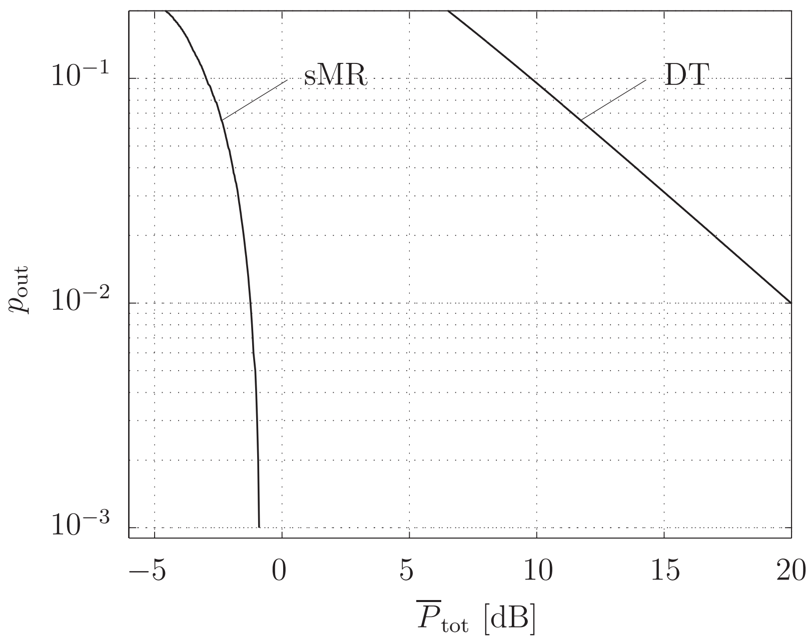

Figure 4.14: Outage probability of selective multi-route and direct transmission vs. averaged total transmit power for $\mathcal{C}_0 = 1$ bit/s/Hz, $d_{\text{sr}} = 0.2$, and $\alpha = 4$.

5

Combining Receiver

> Life is like a sewer. What you get out of it depends on what you put into it.
>
> *Tom Lehrer*

5.1 Introduction

In the previous chapters we have discussed the basic cooperative networks and cooperation strategies. Furthermore, we dealt with relay selection approaches and optimal resource allocation. Hence, an obvious extension is to consider what happens exactly at the destination in a relay network. As mentioned before, the wireless communications channel is characterized by many scattered rays arriving at the destination. Destructive and/or constructive superposition of these rays leads to multipath fading which causes great signal fluctuations of the received signal strength. These effects can be mitigated by the use diversity combining at the destination [53].

Basically, there are three main diversity combining techniques. These are selection combining (SC), equal gain combining (EGC), and maximal ratio combining (MRC). The most valuable paper on combining techniques is "Linear Diversity Combining Techniques" by D. G. Brennan [135]. It gives a comprehensive and well-structured overview as well as a quantitative performance analysis of each technique. By employing SC, only the strongest branch is selected for further processing. For EGC, all branches are co-phased, equally weighted, and then summed up. In contrast to that, the weighting for MRC is performed with respect to the individual channel gains. This means that branches with higher signal strength possess a larger weight. The

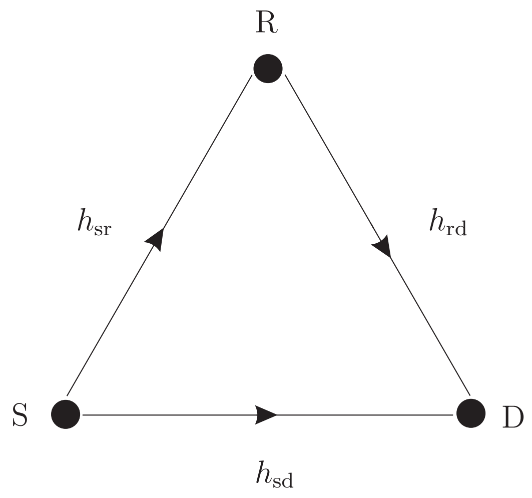

Figure 5.1: Statistically symmetric relay network with source S, relay R, and destination D. Statistically symmetric means that in average all channel coefficients are equal.

principles of combining are widely discussed in literature [53–55]. Nonetheless, the topic of diversity combining receivers is still of great practical relevance. Therefore, we propose a hybrid combining technique that switches between SC and EGC based on an SNR threshold β. This approach leads to a better error performance and has less complexity compared to a receiver structure based on MRC.

In the following sections, we compare those three combining techniques and design a receiver structure for a dual-diversity wireless relay network [136, 137]. For the sake of analysis, we assume a statistically symmetric relay network consisting of one source S, one relay R, and one destination D. Statistically symmetric in that case refers to the fact that in average all channel coefficients h_i, $i \in \{\text{sd,sr,rd}\}$, are equal. With respect to the path loss model presented in Subsection 2.1.1, where $\sigma_i^2 \propto d_i^{-\alpha}$, this means that the distances between all nodes are equal (equilateral triangle, see Fig. 5.1). All distances d_i are normalized to 1, so that no further path loss considerations are necessary. In addition, each branch is represented by a slowly varying flat Rayleigh fading channel. Moreover, each branch is perturbed by AWGN with average power $\tilde{N}$. Let source and relay transmit with equal power P. Then the average SNR of a single branch is given by $\overline{\mathsf{SNR}}_i = \sigma_i^2 \cdot P/\tilde{N}$. Each terminal is equipped with a single antenna and cannot receive and transmit simultaneously. To cope this restriction, a transmission block is divided into two sub-blocks of equal length. Like before, the first sub-block is reserved for the source transmission and the second sub-block is reserved for the relay transmission. As we are interested in proper combining at the destination, we assume that the relay is able to perform some kind of error detection and correction so that an error-free refreshed signal is transmitted from the relay to the destination.

5.2 Relative Comparison

We compare the commonly used combining strategies SC, EGC, and MRC with respect to SNR gains for different degrees of branch unbalance in a dual-diversity communications system. In a first step, we give expressions on the cumulative distribution functions and probability density functions of SC and EGC. As we will show later, the asymptotic gain of MRC and SC, i.e., the SNR gain for a high degree of branch unbalance, are the same. This is one reason why we omit the cumulative distribution function and the probability density function of MRC here. Another reason is the fact that we concentrate on a hybrid approach of SC and EGC due to complexity issues.

5.2.1 Combiner Output Signal-to-Noise Ratio

Let us consider EGC first. As mentioned before, the incoming signals are co-phased, equally weighted, and then summed up. Accordingly, for a dual-diversity relay network the output SNR of EGC is given by [53, 138]

$$\mathsf{SNR}_{\mathrm{egc}} = \frac{(|h_{\mathrm{sd}}| + |h_{\mathrm{rd}}|)^2}{2} \frac{P}{\tilde{N}}. \tag{5.1}$$

In order to give an expression of the cumulative distribution function for $\mathsf{SNR}_{\mathrm{egc}}$, we first have to calculate the cumulative distribution function of the two Rayleigh distributed random variables $|h_{\mathrm{sd}}|+|h_{\mathrm{rd}}|$. This was done, e.g., in [139]. After a proper transformation of random variables as it was done in [138], the cumulative distribution function of $\mathsf{SNR}_{\mathrm{egc}}$ becomes

$$\begin{aligned} F_{\mathsf{SNR}_{\mathrm{egc}}}(\mathsf{SNR}) \;=\; & 1 - \frac{\overline{\mathsf{SNR}}_{\mathrm{sd}} e^{-(2\mathsf{SNR}/\overline{\mathsf{SNR}}_{\mathrm{sd}})} + \overline{\mathsf{SNR}}_{\mathrm{rd}} e^{-(2\mathsf{SNR}/\overline{\mathsf{SNR}}_{\mathrm{rd}})}}{\overline{\mathsf{SNR}}_{\mathrm{sd}} + \overline{\mathsf{SNR}}_{\mathrm{rd}}} \\ & - \frac{2\sqrt{2\overline{\mathsf{SNR}}_{\mathrm{sd}}\overline{\mathsf{SNR}}_{\mathrm{rd}}\pi\mathsf{SNR}}}{(\overline{\mathsf{SNR}}_{\mathrm{sd}} + \overline{\mathsf{SNR}}_{\mathrm{rd}})^{3/2}} e^{-2\mathsf{SNR}/(\overline{\mathsf{SNR}}_{\mathrm{sd}}+\overline{\mathsf{SNR}}_{\mathrm{rd}})} \\ & \cdot \left[1 - Q\left(2\sqrt{\frac{\overline{\mathsf{SNR}}_{\mathrm{sd}}\mathsf{SNR}}{\overline{\mathsf{SNR}}_{\mathrm{rd}}(\overline{\mathsf{SNR}}_{\mathrm{sd}} + \overline{\mathsf{SNR}}_{\mathrm{rd}})}}\right)\right. \\ & \left. - Q\left(2\sqrt{\frac{\overline{\mathsf{SNR}}_{\mathrm{rd}}\mathsf{SNR}}{\overline{\mathsf{SNR}}_{\mathrm{sd}}(\overline{\mathsf{SNR}}_{\mathrm{sd}} + \overline{\mathsf{SNR}}_{\mathrm{rd}})}}\right)\right], \end{aligned} \tag{5.2}$$

where $\overline{\mathsf{SNR}}_i$, $i \in \{\mathrm{sd},\mathrm{rd}\}$, denotes the average SNR per symbol on branch i and $Q(\cdot)$ denotes the Gaussian Q-function defined as [55]

$$Q(z) = \frac{1}{\sqrt{2\pi}} \int_z^{\infty} e^{-x^2/2} \mathrm{d}x. \tag{5.3}$$

The probability density function can easily be derived by differentiating (5.2) with respect to SNR, which leads to the following expression [138]:

$$\begin{aligned} f_{\mathsf{SNR}_{\mathrm{egc}}}(\mathsf{SNR}) &= \frac{2\left(\overline{\mathsf{SNR}}_{\mathrm{sd}} e^{-(2\mathsf{SNR}/\overline{\mathsf{SNR}}_{\mathrm{sd}})} + \overline{\mathsf{SNR}}_{\mathrm{rd}} e^{-(2\mathsf{SNR}/\overline{\mathsf{SNR}}_{\mathrm{rd}})}\right)}{(\overline{\mathsf{SNR}}_{\mathrm{sd}} + \overline{\mathsf{SNR}}_{\mathrm{rd}})^2} \\ &\quad + \sqrt{\frac{2\pi \overline{\mathsf{SNR}}_{\mathrm{sd}} \overline{\mathsf{SNR}}_{\mathrm{rd}}}{\mathsf{SNR}}} \frac{e^{-2\mathsf{SNR}/(\overline{\mathsf{SNR}}_{\mathrm{sd}} + \overline{\mathsf{SNR}}_{\mathrm{rd}})}}{(\overline{\mathsf{SNR}}_{\mathrm{sd}} + \overline{\mathsf{SNR}}_{\mathrm{rd}})^{3/2}} \left(\frac{4\mathsf{SNR}}{\overline{\mathsf{SNR}}_{\mathrm{sd}} + \overline{\mathsf{SNR}}_{\mathrm{rd}}} - 1\right) \\ &\quad \cdot \left[1 - Q\left(2\sqrt{\frac{\overline{\mathsf{SNR}}_{\mathrm{sd}} \mathsf{SNR}}{\overline{\mathsf{SNR}}_{\mathrm{rd}}} (\overline{\mathsf{SNR}}_{\mathrm{sd}} + \overline{\mathsf{SNR}}_{\mathrm{rd}})}\right)\right. \\ &\quad \left. - Q\left(2\sqrt{\frac{\overline{\mathsf{SNR}}_{\mathrm{rd}} \mathsf{SNR}}{\overline{\mathsf{SNR}}_{\mathrm{sd}}} (\overline{\mathsf{SNR}}_{\mathrm{sd}} + \overline{\mathsf{SNR}}_{\mathrm{rd}})}\right)\right] \end{aligned} \tag{5.4}$$

Next, we consider SC. The difference between EGC and SC is, that SC does not combine the two incoming branches, but rather selects the branch with higher SNR for further signal processing. This has the advantage that no co-phasing of the two signals is required. However, SNR monitoring is indispensable for SC in order to have a suitable selection criterion. Consequently, the output SNR of SC is given by

$$\mathsf{SNR}_{\mathrm{sc}} = \max\left\{|h_{\mathrm{sd}}|^2, |h_{\mathrm{rd}}|^2\right\} \frac{P}{\tilde{N}}. \tag{5.5}$$

The cumulative distribution function of $\mathsf{SNR}_{\mathrm{sc}}$ for independent but not necessarily identically distributed branches is well-known and can be found in, e.g., [53, 55]. We get

$$F_{\mathsf{SNR}_{\mathrm{sc}}}(\mathsf{SNR}) = \left(1 - e^{-\mathsf{SNR}/\overline{\mathsf{SNR}}_{\mathrm{sd}}}\right)\left(1 - e^{-\mathsf{SNR}/\overline{\mathsf{SNR}}_{\mathrm{rd}}}\right). \tag{5.6}$$

Again, differentiating (5.6) relative to SNR finally yields the probability density function of $\mathsf{SNR}_{\mathrm{sc}}$:

$$\begin{aligned} f_{\mathsf{SNR}_{\mathrm{sc}}}(\mathsf{SNR}) &= \frac{1}{\overline{\mathsf{SNR}}_{\mathrm{sd}}} e^{-\mathsf{SNR}/\overline{\mathsf{SNR}}_{\mathrm{sd}}} + \frac{1}{\overline{\mathsf{SNR}}_{\mathrm{rd}}} e^{-\mathsf{SNR}/\overline{\mathsf{SNR}}_{\mathrm{rd}}} \\ &\quad - \left(\frac{1}{\overline{\mathsf{SNR}}_{\mathrm{sd}}} + \frac{1}{\overline{\mathsf{SNR}}_{\mathrm{rd}}}\right) e^{-\mathsf{SNR}\left(\frac{1}{\overline{\mathsf{SNR}}_{\mathrm{sd}}} + \frac{1}{\overline{\mathsf{SNR}}_{\mathrm{rd}}}\right)} \end{aligned} \tag{5.7}$$

5.2.2 Signal-to-Noise Ratio Gain

There are two types of performance gains in diversity systems, namely diversity gain and SNR gain[1]. Diversity gain for relay networks was intensively investigated. For instance, consider [37], where the authors define the so-called diversity order as

$$d(R) := -\lim_{\mathsf{SNR}\to\infty} \frac{\log p_{\mathrm{out}}(R, \mathsf{SNR})}{\log(\mathsf{SNR})}. \tag{5.8}$$

[1] In [53, p. 192] SNR gain is referred to as array gain.

It describes the slope of the outage probability curve over SNR for large values of SNR. Diversity order of 3, for example, means that increasing the SNR by 10 dB reduces the outage probability by a factor of 10^3. Generally, it can be stated that the higher the diversity order, the higher the robustness of a communications system to fading. The dependence of the diversity order on the transmission rate R leads to the definition of the multiplexing gain r [140] as

$$r := \lim_{\mathsf{SNR}\to\infty} \frac{R}{\log(\mathsf{SNR})}. \tag{5.9}$$

This is the asymptotic slope of the rate curve over SNR in bits/s/Hz per 3 dB [141]. A similar metric for wireless relay networks was proposed in [142], where the authors considered SNR-vs.-R curves.

In contrast to this, we concentrate on the SNR gain in this section, which is – for a dual-diversity relay system – defined as

$$\Delta_{\mathsf{SNR}} := \frac{\overline{\mathsf{SNR}}_l}{\max\{\overline{\mathsf{SNR}}_{\mathrm{sd}},\overline{\mathsf{SNR}}_{\mathrm{rd}}\}}, \tag{5.10}$$

where $\overline{\mathsf{SNR}}_l$ represents the average SNR of the combining schemes and $l \in \{\mathrm{mrc,egc,sc}\}$. There exist other definitions of SNR gain in literature, e.g., in [143], where the authors refer to SNR gain as the ratio of the SNR of direct transmission and the SNR of various cooperative protocols for the same outage probability. This work was extended in [144].

It is well-known that the average SNR of MRC is the sum of the individual average SNR values [53–55]. Therefore, we have

$$\overline{\mathsf{SNR}}_{\mathrm{mrc}} = \overline{\mathsf{SNR}}_{\mathrm{sd}} + \overline{\mathsf{SNR}}_{\mathrm{rd}}. \tag{5.11}$$

The average SNR of EGC can be calculated by averaging SNR over the probability density function $f_{\mathsf{SNR}_{\mathrm{egc}}}(\mathsf{SNR})$. Accordingly,

$$\overline{\mathsf{SNR}}_{\mathrm{egc}} = \int_0^\infty \mathsf{SNR}\, f_{\mathsf{SNR}_{\mathrm{egc}}}(\mathsf{SNR})\,\mathrm{d}\mathsf{SNR} \tag{5.12}$$

and after proper algebraic manipulation we get

$$\overline{\mathsf{SNR}}_{\mathrm{egc}} = \frac{1}{2}\overline{\mathsf{SNR}}_{\mathrm{sd}} + \frac{1}{2}\overline{\mathsf{SNR}}_{\mathrm{rd}} + \frac{\pi}{4}\sqrt{\overline{\mathsf{SNR}}_{\mathrm{sd}}\overline{\mathsf{SNR}}_{\mathrm{rd}}}. \tag{5.13}$$

The factor $\pi/4$ in (5.13) is typical for Rayleigh fading, where $(\mathbb{E}(|h_i|))^2 = \pi/4 \cdot \mathbb{E}(|h_i|^2) = \pi/4 \cdot \sigma_i^2$.

In order to calculate the average SNR of SC, we have to apply the same techniques as for the average SNR of EGC, i.e., averaging SNR over the probability density function $f_{\mathsf{SNR}_{\mathrm{sc}}}(\mathsf{SNR})$. This finally yields

$$\overline{\mathsf{SNR}}_{\mathrm{sc}} = \overline{\mathsf{SNR}}_{\mathrm{sd}} + \overline{\mathsf{SNR}}_{\mathrm{rd}} - \frac{\overline{\mathsf{SNR}}_{\mathrm{sd}}\overline{\mathsf{SNR}}_{\mathrm{rd}}}{\overline{\mathsf{SNR}}_{\mathrm{sd}} + \overline{\mathsf{SNR}}_{\mathrm{rd}}}. \tag{5.14}$$

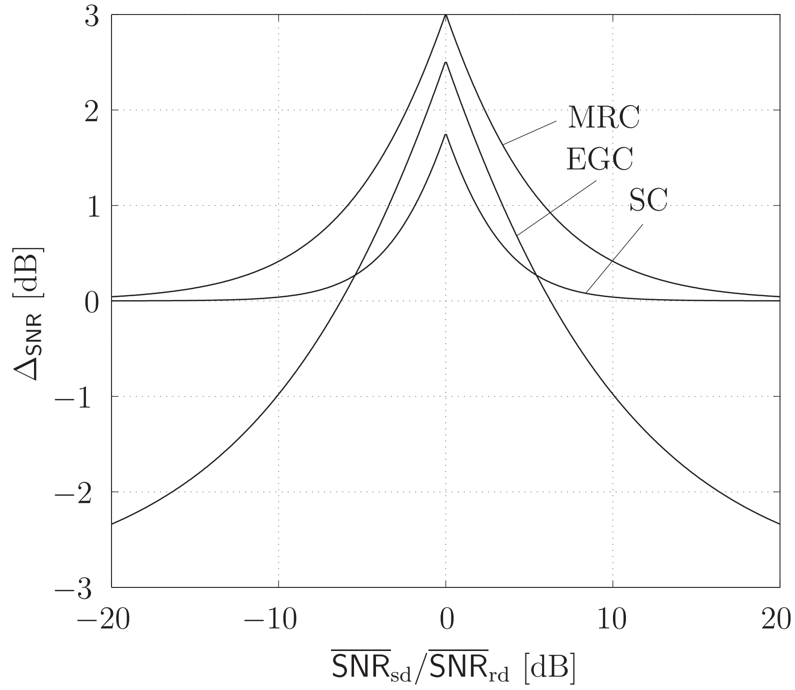

Figure 5.2: Comparison of the SNR gain Δ_{SNR} of selection combining (SC), equal gain combining (EGC), and maximal ratio combining (MRC) with respect to branch unbalance $\overline{\mathsf{SNR}}_{\mathrm{sd}}/\overline{\mathsf{SNR}}_{\mathrm{rd}}$ [dB].

Fig. 5.2 illustrates the SNR gain Δ_{SNR} [dB] for SC, EGC, and MRC over the branch unbalance $\overline{\mathsf{SNR}}_{\mathrm{sd}}/\overline{\mathsf{SNR}}_{\mathrm{rd}}$ [dB]. For *no* branch unbalance, i.e., $\overline{\mathsf{SNR}}_{\mathrm{sd}}/\overline{\mathsf{SNR}}_{\mathrm{rd}} = 0$ dB, we obtain the well-known results for dual-diversity and independent and identically distributed (i.i.d.) branches. Then, SNR gain for SC becomes 1.8 dB, for EGC we get 2.5 dB, and for MRC the SNR gain is 3 dB. Since the maximal SNR gain of a dual-diversity communications system without fading is 3 dB, it can easily be seen that MRC performs optimal if the system does not suffer any branch unbalance. However, MRC requires the knowledge of channel state information which is a challenging task, especially for time-variant channels. Furthermore, MRC outperforms all other combining strategies independent of the branch unbalance. This is due to the weighting factor which is proportional to the channel quality. EGC outperforms SC only for low values of branch unbalance. This aspect is further discussed in Subsection 5.2.3 which deals with the asymptotic behavior of the SNR gain for the different combining techniques.

The interception point between the SNR gain of SC and EGC can be calculated by simply equating (5.13) and (5.14). This leads to a fourth-order equation with respect to branch unbalance $\overline{\mathsf{SNR}}_{\mathrm{sd}}/\overline{\mathsf{SNR}}_{\mathrm{rd}}$ and can be solved by applying Ferrari's method (see Appendix C, where the principle of the method is described in detail) which is implemented in most mathematical tools. In our case, we have

$$\begin{aligned} \overline{\mathsf{SNR}}_{\mathrm{sd}}/\overline{\mathsf{SNR}}_{\mathrm{rd}} &= 3.488 \quad \rightarrow \quad 5.42\,\mathrm{dB} \\ \overline{\mathsf{SNR}}_{\mathrm{sd}}/\overline{\mathsf{SNR}}_{\mathrm{rd}} &= 0.287 \quad \rightarrow -5.42\,\mathrm{dB}, \end{aligned}$$

which corresponds to the results illustrated in Fig. 5.2.

5.2.3 Asymptotic Behavior

A short glance at Fig. 5.2 reveals that the SNR gain possesses an asymptotic behavior for high branch unbalances. The asymptotic values can easily be derived by letting $\overline{\mathsf{SNR}}_{\mathrm{sd}}/\overline{\mathsf{SNR}}_{\mathrm{rd}} \rightarrow \infty$. Since all transmitted signals are power constrained, this can be done by letting $\overline{\mathsf{SNR}}_{\mathrm{rd}} \rightarrow 0$ and keeping $\overline{\mathsf{SNR}}_{\mathrm{sd}}$ fixed. Afterwards, the division by $\max\{\overline{\mathsf{SNR}}_{\mathrm{sd}},\overline{\mathsf{SNR}}_{\mathrm{rd}}\}$ leads to the effect that the asymptotes are independent of single average SNR values.

SC and MRC both tend to 0 dB as the branch unbalance increases. This means that the SNR gain of these combining strategies is always higher than that of a single branch transmission system. For the case of SC, this is due to the fact that only the branch with higher SNR is selected for further signal processing and, therefore, the worst we can do is at least as good as a single branch transmission system. For the case of MRC, the reason is the way the weighting factor is determined. Each branch is weighted with its individual channel gain, i.e., strong channels have a larger weighting factor than weak channels. As a consequence, weak channels are "filtered out" for high degrees of branch unbalance and, again, we perform at least as good as a single branch transmission system. This is not true for EGC anymore. EGC shows a different behavior, since both branches are weighted equally. This means that if one of the two branches is very strong and the other one is very weak, the latter becomes more or less noise and only increases the noise level with respect to the strong branch. Worst case is doubling the noise power, which eventually leads to an asymptotic value of -3 dB. Then, a dual-diversity communications system that employs EGC is degraded to a single branch transmission system with half the SNR. We will come back to this aspect in Subsection 5.3.2, where we investigate the error performance of our hybrid combining receiver.

5.3 Receiver Structure

In this section, we describe a new hybrid combining receiver that selects dynamically between SC and EGC on the basis of an SNR threshold criterion. As we have seen before, EGC outperforms SC for a low branch unbalance. As the branch unbalance increases, the performance of EGC compared to SC gets worse. With respect to SNR gain, the interception point where SC outperforms EGC is given by $|\overline{\mathsf{SNR}}_{\mathrm{sd}}/\overline{\mathsf{SNR}}_{\mathrm{rd}}| =$ 5.42 dB (see Subsection 5.2.2). With respect to error performance, the issue gets more complicated, since there exists no linear mapping of SNR gains to the error probability. Nevertheless, we can still state that there is an interception point where SC achieves a lower error probability compared to EGC depending on the branch unbalance. However, determination of this interception point gets more involved. In Subsection 5.3.1 we describe the new hybrid receiver structure, whereas the issue of error probability is discussed in detail in Section 5.3.2.

5.3.1 Description

For practical implementation issues it is sometimes preferable to use EGC and/or SC instead of MRC, even if MRC achieves a better performance. The reason for this is that EGC and SC do not require an estimation of the channel state information as it is the case for MRC, where each branch is weighted by a factor that is proportional to its channel gain. That is why we focus on a hybrid combining receiver that alters between the usage of EGC and SC depending on an SNR threshold criterion. However, EGC suffers a great SNR gain degradation for a high degree of branch unbalance as can be seen in Fig. 5.2. It is, hence, obvious that we also take SC into account for the design of our receiver.

The principle of the receiver structure can be explained as follows. For a low branch unbalance, i.e., if the branches from the source and the relay have approximately the same quality, EGC is the preferred combining strategy, whereas for a high branch unbalance, i.e., one of the branches either from the source or from the relay suffers great signal fluctuations due to fading, it is beneficial to select SC and exclude the worse branch from further signal processing. The great advantage of that receiver structure is that we will always achieve a better performance compared to a transmission system where the receiver only exploits one branch. Recall that this is only the case when we select dynamically between EGC and SC. If we only rely on EGC, we will perform worse compared to a single branch transmission system for a high branch unbalance. This will be discussed in more detail in Subsection 5.3.2.

Fig. 5.3 illustrates the structure of the hybrid combining receiver. The destination receives two signals in orthogonal time slots. The first signal it receives is from the source, the second one comes from the relay and is a refreshed version of the original source signal. The receiver monitors the SNR value on each branch.[2] In a next step, the receiver calculates the ratio of the two SNR values, i.e.,

$$\theta = \frac{\mathsf{SNR}_{\mathrm{sd}}}{\mathsf{SNR}_{\mathrm{rd}}}. \tag{5.15}$$

Thereafter, the absolute value of this ratio expressed in dB is determined as

$$\Theta = |10 \log_{10}(\theta)|. \tag{5.16}$$

[2]Note that we use the instantaneous SNR values for the practical implementation rather than average values.

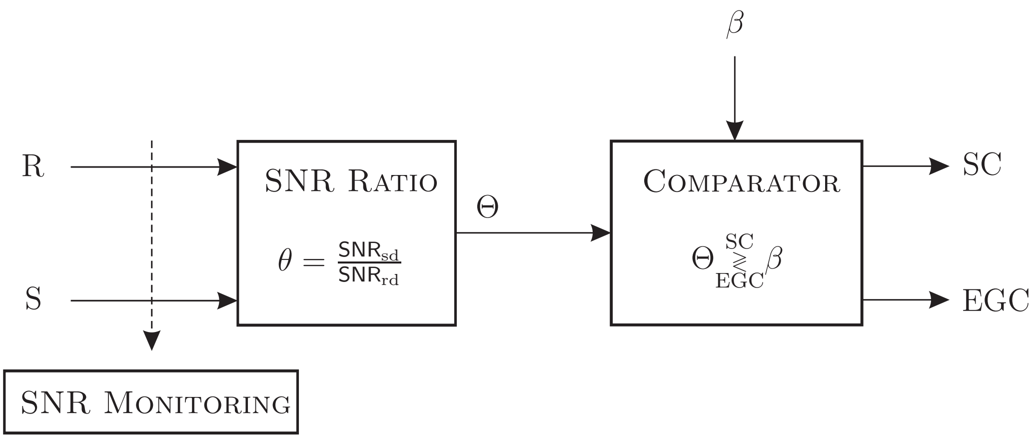

Figure 5.3: Structure of a dual-diversity combining receiver that selects between selection combining (SC) and equal gain combining (EGC) on the basis of an SNR criterion.

The usage of the log-function and the determination of the absolute value are necessary in order to cope with the problem of branch unbalance. It is not important if the branch from the source or the branch from the relay is the strong one. The only aspect is indeed the branch unbalance. The value Θ represents the input to a comparator, where Θ is compared to a threshold value β. If Θ is greater than the threshold value β, SC will be beneficial and will be selected as combining strategy. If, however, Θ is lower than the threshold value β, EGC will be selected. As already mentioned in the introduction of Section 5.3, the crucial point of this receiver structure is – apart from the practical challenges in measuring the true SNR of a branch[3] – the determination of the threshold value β. We could, indeed, determine a threshold based on the SNR gain, i.e., $|\overline{\mathsf{SNR}}_{\mathrm{sd}}/\overline{\mathsf{SNR}}_{\mathrm{rd}}| = 5.42$ dB. However, this will not guarantee that we always achieve the best error performance due to the non-linear mapping of the SNR gain to the error probability. This issue is further discussed in the following subsection.

5.3.2 Error Performance

The error performance of combining strategies for several fading characteristics was widely investigated in literature. We state results for SC and EGC that are of special interest for further analysis of the hybrid receiver structure and concentrate especially on the bit error probability of binary phase shift keying (BPSK) in the following. The interested reader is referred to the publications [138, 145–148] and the references therein.

[3] See Section 5.4 for more information on that topic.

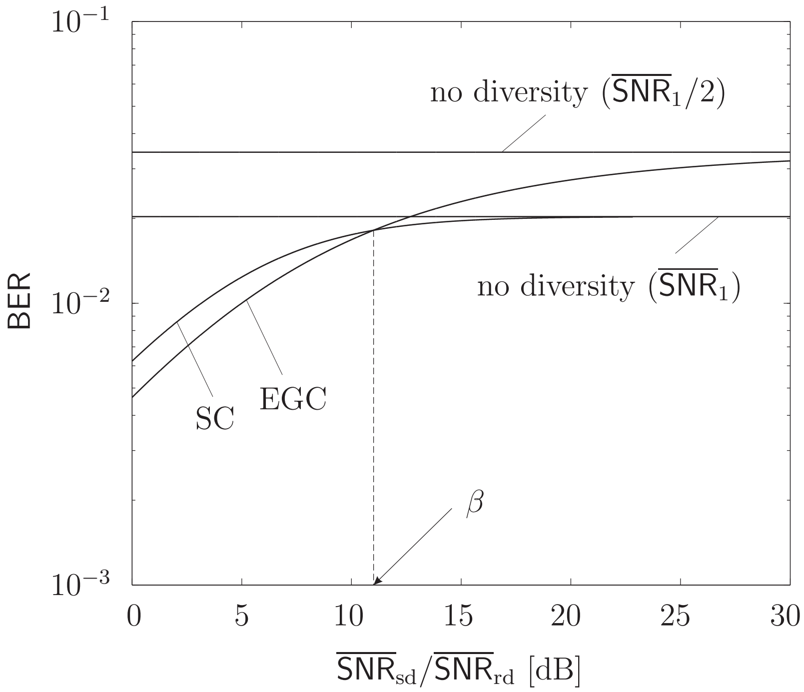

Figure 5.4: Bit error rate (BER) of selection combining (SC) and equal gain combining (EGC) for BPSK with respect to branch unbalance $\overline{\mathsf{SNR}}_{\mathrm{sd}}/\overline{\mathsf{SNR}}_{\mathrm{rd}}$ [dB]. Parameter $\overline{\mathsf{SNR}}_{\mathrm{sd}}$ was set to 5 dB.

The BER of EGC for BPSK for two independent but not identically distributed branches is given by [146, 148]

$$\mathsf{BER}_{\mathrm{egc}} = \frac{1}{2}\left(1 - \frac{\sqrt{\overline{\mathsf{SNR}}_{\mathrm{sd}}(\overline{\mathsf{SNR}}_{\mathrm{sd}}+2)} + \sqrt{\overline{\mathsf{SNR}}_{\mathrm{rd}}(\overline{\mathsf{SNR}}_{\mathrm{rd}}+2)}}{\overline{\mathsf{SNR}}_{\mathrm{sd}} + \overline{\mathsf{SNR}}_{\mathrm{rd}} + 2}\right). \tag{5.17}$$

In the case of SC, the error probability for BPSK for two independent branches can be expressed as [145]

$$\begin{aligned}\mathsf{BER}_{\mathrm{sc}} &= \frac{1}{2}\left(1 - \sqrt{\frac{\overline{\mathsf{SNR}}_{\mathrm{sd}}}{\overline{\mathsf{SNR}}_{\mathrm{sd}}+1}} - \sqrt{\frac{\overline{\mathsf{SNR}}_{\mathrm{rd}}}{\overline{\mathsf{SNR}}_{\mathrm{rd}}+1}}\right.\\ &\qquad \left. + \sqrt{\frac{\overline{\mathsf{SNR}}_{\mathrm{sd}}\overline{\mathsf{SNR}}_{\mathrm{rd}}}{\overline{\mathsf{SNR}}_{\mathrm{sd}}\overline{\mathsf{SNR}}_{\mathrm{rd}} + \overline{\mathsf{SNR}}_{\mathrm{sd}} + \overline{\mathsf{SNR}}_{\mathrm{rd}}}}\right).\end{aligned} \tag{5.18}$$

The error probabilities of EGC and SC vs. branch unbalance are depicted in Fig. 5.4 for $\overline{\mathsf{SNR}}_{\mathrm{sd}} = 5$ dB and Fig. 5.5 for $\overline{\mathsf{SNR}}_{\mathrm{sd}} = 10$ dB. It can easily be seen that the BER – as expected – increases with increasing branch unbalance. This behavior corresponds to the fact that the SNR gain of both combining strategies decreases with increasing branch unbalance (cf. Fig. 5.2). The intersection point between EGC and SC denotes

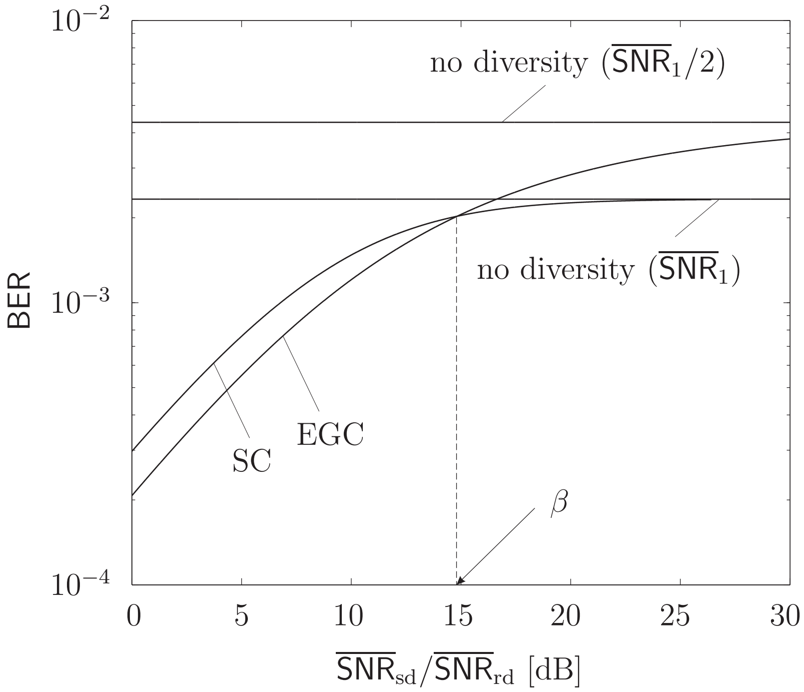

Figure 5.5: Bit error rate (BER) of selection combining (SC) and equal gain combining (EGC) for BPSK with respect to branch unbalance $\overline{\mathsf{SNR}}_{\mathrm{sd}}/\overline{\mathsf{SNR}}_{\mathrm{rd}}$ [dB]. Parameter $\overline{\mathsf{SNR}}_{\mathrm{sd}}$ was set to 10 dB.

the threshold value β. For $\overline{\mathsf{SNR}}_{\mathrm{sd}} = 5$ dB, we get $\beta = 11.01$ dB, and for $\overline{\mathsf{SNR}}_{\mathrm{sd}} = 10$ dB, we get $\beta = 14.83$ dB. Moreover, we see that both BER curves tend to an asymptotic value for a high degree of branch unbalance. For SC this asymptote corresponds to a single branch transmission system with an average SNR of $\overline{\mathsf{SNR}} = \overline{\mathsf{SNR}}_{\mathrm{sd}}$. This can be made intuitively clear by the fact that the SNR gain of SC tends to zero for a high degree of branch unbalance. The asymptote for EGC is determined by a single branch transmission system with an average SNR of $\overline{\mathsf{SNR}} = \overline{\mathsf{SNR}}_{\mathrm{sd}}/2$. This becomes obvious if we take a closer look to the SNR gain again. For EGC, the SNR gain tends to -3 dB for a high degree of branch unbalance. This factor of 1/2 contributes to the average SNR expression. BER then equals that of a non-diversity receiver, which is

$$\mathsf{BER}_{\mathrm{no-div}} = \frac{1}{2}\left(1 - \sqrt{\frac{\overline{\mathsf{SNR}}}{\overline{\mathsf{SNR}} + 1}}\right) \tag{5.19}$$

with the values for $\overline{\mathsf{SNR}}$ given above. The values for the asymptotic error probabilities can also be derived from (5.17) and (5.18) by letting $\overline{\mathsf{SNR}}_{\mathrm{rd}} \to 0$.

5.4 Practical Implementation Issues

Due to complexity issues, we skipped MRC in our hybrid receiver structure. The advantage of EGC and SC over MRC is that no estimation of channel state information is required. In contrast to EGC, where SNR monitoring is not necessary, it is indispensable for SC to be able to select the strongest branch. Measuring true SNR of a branch, i.e.,

$$|h_i|^2 \cdot P/\tilde{N},$$

is a complex and practically challenging task. A beneficial approach is to measure the total power of the received signal, i.e.,

$$|h_i|^2 \cdot P + \tilde{N},$$

which is equivalent if the noise power on each branch is considered to be equal [53, 55].

Another issue is the derivation of the threshold β as a function of $\overline{\mathsf{SNR}}_{\mathrm{sd}}$ and $\overline{\mathsf{SNR}}_{\mathrm{rd}}$. We can see in Fig. 5.4 and Fig. 5.5 that the threshold β is strongly varying depending on the parameter $\overline{\mathsf{SNR}}_{\mathrm{sd}}$ and does not only dependent on the branch unbalance Θ.[4] Unfortunately, the calculation of the interception point between the error probabilities of EGC and SC gets involved and there exists no closed-form solution to this problem. A practical approach is to find the threshold β depending on the branch unbalance Θ and the parameter $\overline{\mathsf{SNR}}_{\mathrm{sd}}$ by simulations and store the different results in a look-up table. All the receiver now has to do is the following. It measures the SNR of the source branch and stores the value. Next, it measures the SNR of the relay branch and calculates the ratio Θ of both values. With knowledge of $\overline{\mathsf{SNR}}_{\mathrm{sd}}$, the receiver can look up the threshold β. If β is lower than Θ, it will be beneficial to use SC. On the contrary, if β is greater than Θ, the usage of EGC will be beneficial. By applying this simple algorithm, we take the variability of β with respect to the parameter $\overline{\mathsf{SNR}}_{\mathrm{sd}}$ into account. The algorithm is depicted in Fig. 5.6.

An exemplary look-up table for $\overline{\mathsf{SNR}}_{\mathrm{sd}} = -2, \ldots, 30$ dB is shown in Tab. 5.1. The corresponding curve of β vs. $\overline{\mathsf{SNR}}_{\mathrm{sd}}$ is illustrated in Fig. 5.7. It can be seen that β increases with increasing $\overline{\mathsf{SNR}}_{\mathrm{sd}}$. Furthermore, for low values of $\overline{\mathsf{SNR}}_{\mathrm{sd}}$, the threshold β possesses a strongly non-linear behavior. For large values of $\overline{\mathsf{SNR}}_{\mathrm{sd}}$, however, β can be approximated by a linear curve given by

$$\beta \approx \overline{\mathsf{SNR}}_{\mathrm{sd}} + 4\,\mathrm{dB}. \tag{5.20}$$

This approximation is shown as dashed curve in Fig. 5.7.

[4] This is in contrast to the SNR gain discussed in Subsection 5.2.2.

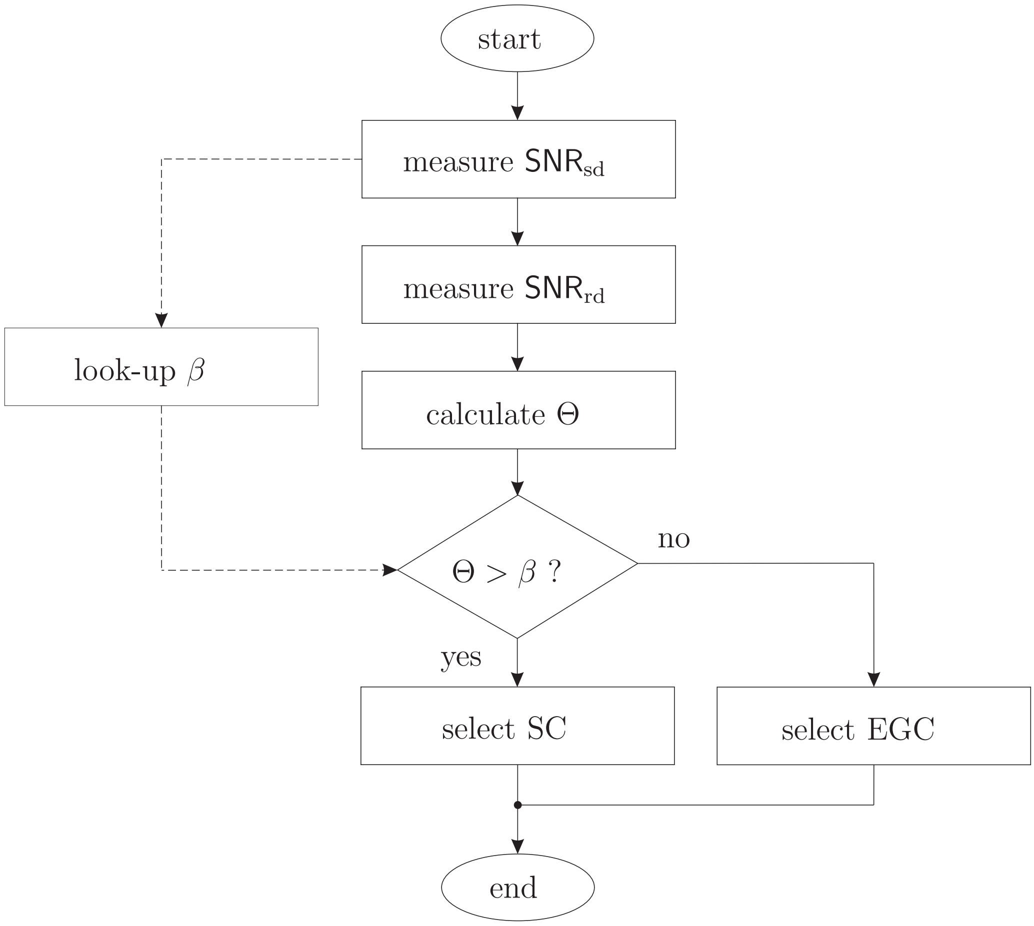

Figure 5.6: Algorithm for the hybrid receiver structure.

Table 5.1: Look-up table for threshold β [dB] and $\overline{\mathsf{SNR}}_{sd}$ [dB] (cf. Fig. 5.7).

$\overline{\mathsf{SNR}}_{sd}$ [dB]	β [dB]	$\overline{\mathsf{SNR}}_{sd}$ [dB]	β [dB]	$\overline{\mathsf{SNR}}_{sd}$ [dB]	β [dB]
−2	7.77	9	14.01	20	24.22
−1	8.09	10	14.83	21	25.20
0	8.43	11	15.70	22	26.19
1	8.84	12	16.58	23	27.18
2	9.29	13	17.51	24	28.17
3	9.82	14	18.43	25	29.17
4	10.37	15	19.37	26	30.16
5	11.01	16	20.35	27	31.16
6	11.70	17	21.29	28	32.15
7	12.41	18	22.26	29	33.15
8	13.18	19	23.23	30	34.15

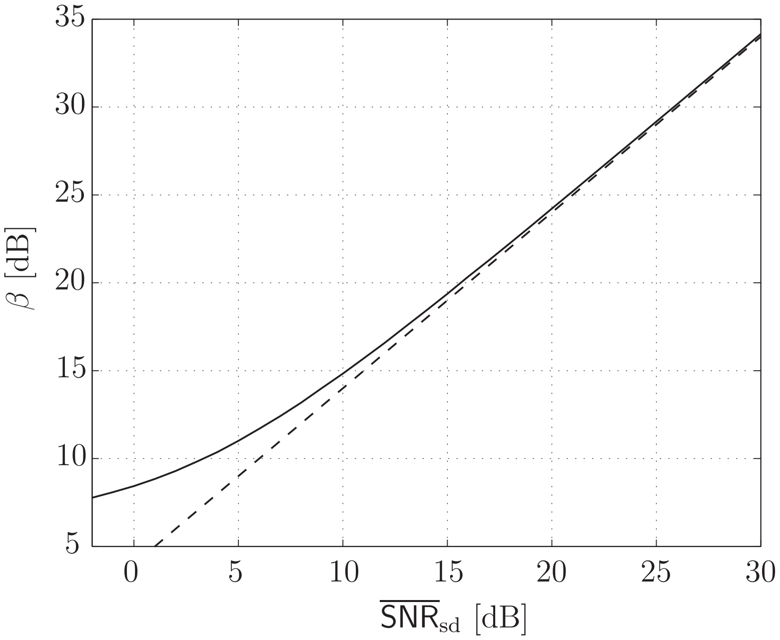

Figure 5.7: SNR threshold β [dB] vs. $\overline{\mathsf{SNR}}_{\text{sd}}$ [dB].

6

Incremental Relaying

Two roads diverged in a wood, and
I – I took the one less traveled by,
And that has made all the
difference.

Robert Frost

6.1 Introduction and System Model

Diversity as defined in (5.8) leads to an increased exponential decay rate in the error probability with increasing SNR and therefore becomes more evident in the high SNR regime. For example, a diversity order of 2 describes a decrease of the outage probability proportional to 10^{-2} when SNR of the system is increased by 10 dB [37]. Drawback of this performance metric is, however, that SNR cannot be increased arbitrarily. This is especially the case for applications such as ad-hoc and sensor networks, where the limited resource power (or energy) plays an important role in the network design [149]. Consequently, from a practical point of view, the low SNR regime is of much more interest.

Shannon capacity which describes the maximal transmission rate for an arbitrarily small probability of error (under an average power constraint) is not a useful metric anymore, since we consider Rayleigh block fading where errors are inevitable at any nonzero transmission rate. In a strict sense, Shannon capacity of these channels equals 0 and a more suitable metric has to be found. That is why ϵ-outage capacity was defined as the maximal transmission rate for which the outage probability is not larger than a given target error rate ϵ [60, 150]. Outage probability is considered here, as it gives a good approximation of the error probability in coded systems with sufficiently long block size [151]. The ϵ-outage capacity for a frequency division cooperative system for low SNR values was investigated in [152]. There, it was shown that a bursty version of the amplify-and-forward (BAF) protocol (see Subsection 6.2.3) achieves the optimal performance[1] and that the ϵ-outage capacities for the non-coherent and the coherent scenario are the same.

In this chapter, we investigate an incremental relaying protocol (IR). It was first described by Laneman et al. in [37]. In IR networks, a source first transmits its information to the destination. Due to the broadcast nature of the wireless channel, the relay is able to receive the source signal as well. Now, the destination sends a one-bit acknowledgment (`ACK`) to the relay and the source if it is able to decode the source signal reliably. If this is not the case, then the destination sends a negative acknowledgment (`NACK`) to indicate failure of transmission. When the relay receives the `NACK`, it forwards an alternate version of the source information to the destination. The version that the relay transmits to the destination depends on the relaying strategy, i.e., DF or AF, and on the coding strategy, i.e., repetition coding or parallel channel coding[2]. The destination then combines both signals by using maximal ratio combining. Note that we are interested in the optimal performance of the protocol and do not care about complexity issues at the moment (cf. Chapter 5 for more information). We stress that by using jointly designed but independent codebooks (i.e., parallel channel coding), it is generally possible to achieve better results. However, in the low SNR regime, parallel channel coding can be deduced to repetition coding. For that reason, repetition coding is optimal for low SNR values [152]. The major problem in analyzing an incremental relaying protocol is the fact that the overall transmission rate of the system is a random variable which depends on the channel conditions between the network nodes. This problem was solved by defining a long-term average rate $\bar{R}$ in [37]. However, for a given SNR, there are several values of target rate R that lead to the same $\bar{R}$ (see Fig. 6.1). The authors solved this problem by selecting the smallest rate R, i.e., the rate that leads to the highest degree of reliability. Another scheme that was proposed in literature and that can also have a variable transmission rate that depends on the channel conditions was investigated in [153], where the authors dealt with hybrid automatic repeat request (ARQ) with a constant outage probability. This is achieved by dynamically adapting the transmission rate.

We first assume a perfect feedback channel and a three-node network which consists of one source, one relay, and one destination, and investigate the ϵ-outage capacity for different cooperation strategies, i.e., DF and BAF. We then compare the results to the ϵ-outage capacity of the cut-set bound. Especially, the results for DF (and some

[1] Optimal in this context means that the ϵ-outage capacity of the described bursty amplify-and-forward protocol equals the ϵ-outage capacity of the cut-set bound.

[2] The difference between repetition coding and parallel channel coding is the following. For repetition coding, the source and the relay employ the same codebook. For parallel channel coding, the source and the relay employ independent codebooks.

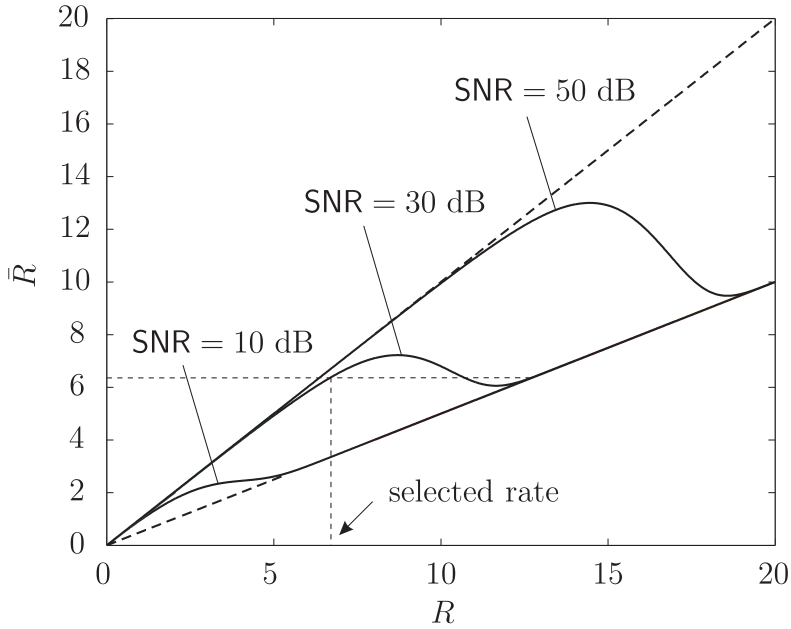

Figure 6.1: Mapping of R to $\bar{R}$. Dashed lines illustrate upper and lower bounds. The upper bound is given by $\bar{R} = R$ and the lower bound is given by $\bar{R} = R/2$. In [37] the smallest R out of those is selected that determine the same $\bar{R}$.

variations of it) are compared to those of transmit diversity in Subsection 6.2.2. In Section 6.3 our results are extended to networks with an arbitrary number of relays. The strong assumption of perfect feedback is weakened in Section 6.4, where we model imperfect feedback links as binary symmetric channels. After analyzing the one-relay and the two-relay case, our findings are extended again to networks with an arbitrary number of relays.

6.2 Perfect Feedback Channel

6.2.1 Decode-and-Forward

As already mentioned, we consider incremental relaying (IR) as a cooperation protocol that exploits the availability of a one-bit feedback from the destination in form of an `ACK`/`NACK` signal [79]. For the following description, confer to Fig. 6.2. One transmission block is divided into two sub-blocks of equal length. Note that the initial transmission rate is R. Hence, the transmission rate within each sub-block is set to $2R$ in order to have the same amount of information (number of bits) transmitted compared to the case where a source transmits over the whole block with rate R (direct transmission). Now, if the source transmission was successful at the end of the first sub-block, information has been transmitted over half a block and we get a total rate of $2R$. Since we are concerned with block fading, this automatically means that

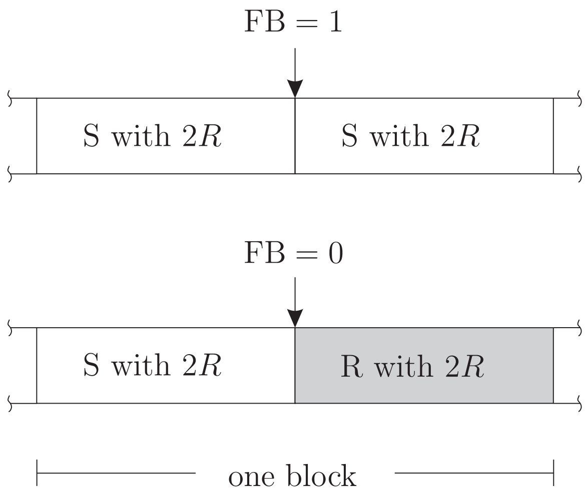

Figure 6.2: Transmission model for incremental relaying. If the source-destination link is not in outage (feedback FB $= 1$), the source transmits during the second sub-block, too. If the source-destination link is in outage (feedback FB $= 0$), the relay aids communication during the second sub-block.

during the second sub-block the source can transmit its next message which will then be sent successfully to the destination. As a consequence, there is no need for the relay to transmit during the second sub-block. However, if the source transmission failed during the first sub-block, the relay transmits over the second sub-block and the overall rate becomes R. For the analysis, we define $\mathcal{X}$ as the event that the source transmission failed, i.e.,

$$\mathcal{X} := \{h_{\text{sd}} : |h_{\text{sd}}|^2 < g\},$$

where we used the definition

$$g(R,\mathsf{SNR}) := \frac{2^{2R} - 1}{\mathsf{SNR}} \tag{6.1}$$

and dropped the dependence on R and SNR for the sake of brevity. Similarly, we define

$$\begin{aligned} \mathcal{Y} &:= \left\{h_{\text{sr}} : |h_{\text{sr}}|^2 < g\right\} \\ \mathcal{Z} &:= \left\{(h_{\text{sd}},h_{\text{rd}}) : |h_{\text{sd}}|^2 + |h_{\text{rd}}|^2 < g\right\}. \end{aligned}$$

There are two cases in which an outage of the system is declared. First, both the source transmission to the destination as well as the source transmission to the relay failed. Second, the relay was able to decode the source signal, but the accumulation of SNR from the source transmission and the relay transmission at the destination is not large enough to exceed a required minimum threshold for decoding. Dropping the dependence on R and SNR for simplicity, the outage probability is given by

$$\begin{aligned} p_{\text{out}}^{(\text{DF})} &= \Pr(\mathcal{X})\Pr(\mathcal{Y})\Pr(\mathcal{Z}|\mathcal{X}\mathcal{Y}) + \Pr(\mathcal{X})\Pr(\mathcal{Y}^c)\Pr(\mathcal{Z}|\mathcal{X}\mathcal{Y}^c) \\ &= \Pr(\mathcal{X})\Pr(\mathcal{Y}) + \Pr(\mathcal{Y}^c)\Pr(\mathcal{Z}), \end{aligned}$$

where $\mathcal{Y}^c$ describes the complement of $\mathcal{Y}$, $\Pr(\mathcal{Z}|\mathcal{X}\mathcal{Y}) = 1$, and $\Pr(\mathcal{X})\Pr(\mathcal{Z}|\mathcal{X}\mathcal{Y}^c) = \Pr(\mathcal{Z})$ due to $\mathcal{Z} \subseteq \mathcal{X}$. Adopting our system model yields

$$\begin{aligned} p_{\text{out}}^{(\text{DF})} &= \Pr\left(|h_{\text{sd}}|^2 < g\right)\Pr\left(|h_{\text{sr}}|^2 < g\right) \\ &\quad + \Pr\left(|h_{\text{sr}}|^2 \geq g\right)\Pr\left(|h_{\text{sd}}|^2 + |h_{\text{rd}}|^2 < g\right). \end{aligned} \tag{6.2}$$

This expression on outage probability can be calculated by applying Lemma 1 (see Appendix A) which deals with the sum of independent exponentially distributed random variables. Since we are interested in the low SNR regime, we have to ensure that the rate is adapted according to the SNR, in order to be able to apply the above mentioned lemma. If the condition $g \to 0$ for $\mathsf{SNR} \to 0$ is met, outage probability for small values of SNR can then be expressed as

$$\begin{aligned} \lim_{\substack{\epsilon\to 0 \\ \mathsf{SNR}\to 0}} \frac{p_{\text{out}}^{(\text{DF})}}{g^2} &= \lim_{\substack{\epsilon\to 0 \\ \mathsf{SNR}\to 0}} \left\{ \frac{\Pr(|h_{\text{sd}}|^2 < g)}{g}\frac{\Pr(|h_{\text{sr}}|^2 < g)}{g} \right. \\ &\quad \left. + \frac{\Pr(|h_{\text{sd}}|^2 + |h_{\text{rd}}|^2 < g)}{g^2}\frac{\Pr(|h_{\text{sr}}|^2 \geq g)}{1} \right\} \\ &= \frac{1}{\sigma_{\text{sd}}^2}\frac{1}{\sigma_{\text{sr}}^2} + \frac{1}{2\sigma_{\text{sd}}^2\sigma_{\text{rd}}^2}\cdot 1 \\ &= \frac{2\sigma_{\text{rd}}^2 + \sigma_{\text{sr}}^2}{2\sigma_{\text{sd}}^2\sigma_{\text{sr}}^2\sigma_{\text{rd}}^2}. \end{aligned} \tag{6.3}$$

Here, $\epsilon \to 0$ implies $g \to 0$, which again means that the rate is adapted in accordance to the SNR. With (1.12), the ϵ-outage capacity eventually becomes

$$\mathcal{C}_{\epsilon}^{(\text{DF})} = \frac{1}{2}\log_2\left(1 + \mathsf{SNR}\sqrt{\frac{2\sigma_{\text{sd}}^2\sigma_{\text{sr}}^2\sigma_{\text{rd}}^2\epsilon}{2\sigma_{\text{rd}}^2 + \sigma_{\text{sr}}^2}}\right). \tag{6.4}$$

Note, that this expression is not the ϵ-outage capacity for incremental relaying, as it does not include the variable transmission rate that occurs for incremental relaying on a long-term perspective. To account for that, the average amount of sub-blocks required for transmission must be taken into account. If the source transmission was successful, we need only one sub-block, no matter whether the relay is able to decode the source signal or not. However, if the source transmission failed, we then must transmit over two sub-blocks. Again, the number of sub-blocks required for transmission does not depend on the ability of the relay to decode the source signal. Let us define a random variable N denoting the number of transmission phases. The average of N becomes $\mathbb{E}(N) = 1 + \Pr(\mathcal{X})$. The ϵ-outage capacity of incremental relaying for decode-and-forward, denoted by the superscript (iDF), can now be written as

$$\mathcal{C}_{\epsilon}^{(\text{iDF})} = \frac{2}{\mathbb{E}(N)}\mathcal{C}_{\epsilon}^{(\text{DF})} = \frac{1}{\mathbb{E}(N)}\log_2\left(1 + \mathsf{SNR}\sqrt{\frac{2\sigma_{\text{sd}}^2\sigma_{\text{sr}}^2\sigma_{\text{rd}}^2\epsilon}{2\sigma_{\text{rd}}^2 + \sigma_{\text{sr}}^2}}\right). \tag{6.5}$$

The factor $2/\mathbb{E}(N)$ accounts for the possible reduction of required transmission phases. If we only need one transmission phase, i.e., half a block (see Fig. 6.2), we obtain a gain of 2. If we need two phases, i.e., the whole block, we are at least as good as a relay network without feedback where the relay always transmits if it has been able to decode the source message. Therefore, we get the bounds

$$1 \leq \frac{\mathcal{C}_\epsilon^{(\mathrm{iDF})}}{C_\epsilon^{(\mathrm{DF})}} \leq 2. \tag{6.6}$$

Assume a one-dimensional geometry, where the relay is located on a straight line between the source and the destination. Accordingly, $d_{\mathrm{rd}} = 1 - d_{\mathrm{sr}}$. Moreover, let all distances be normalized to d_{sd} so that $\sigma_{\mathrm{sd}}^2 = 1$. We then get

$$\mathcal{C}_\epsilon^{(\mathrm{iDF})} = \frac{1}{\mathbb{E}(N)} \log_2 \left(1 + \mathsf{SNR} \sqrt{\frac{2\epsilon}{2d_{\mathrm{sr}}^\alpha + (1-d_{\mathrm{sr}})^\alpha}} \right). \tag{6.7}$$

We are now interested in the optimal relay location, i.e., the relay location that maximizes ϵ-outage capacity. We get

$$d_{\mathrm{sr}}^* = \arg\max_{d_{\mathrm{sr}}} \mathcal{C}_\epsilon^{(\mathrm{iDF})} = \arg\min_{d_{\mathrm{sr}}} \Psi(d_{\mathrm{sr}}), \tag{6.8}$$

where $\Psi(d_{\mathrm{sr}}) = 2d_{\mathrm{sr}}^\alpha + (1-d_{\mathrm{sr}})^\alpha$. It can easily be seen that the optimal relay location is independent of SNR and the target outage probability ϵ. Differentiating $\Psi(d_{\mathrm{sr}})$ with respect to d_{sr} and setting the result equal to 0,

$$\frac{\partial \Psi(d_{\mathrm{sr}})}{\partial d_{\mathrm{sr}}} = 0, \tag{6.9}$$

we get

$$d_{\mathrm{sr}}^* = \frac{1}{1 + \sqrt[\alpha-1]{2}} < 0.5. \tag{6.10}$$

The fact that d_{sr}^* is bounded by 0.5 corresponds to results for decode-and-forward presented in [39]. The authors demonstrate that decode-and-forward performs better if the relay is located closer to the source than to the destination. For free space propagation, e.g., we have $d_{\mathrm{sr}}^*(\alpha = 2) = 1/3$ and for $\alpha = 3$, $d_{\mathrm{sr}}^*(\alpha = 3) = \sqrt{2} - 1 \approx 0.4142$. The optimal relay location d_{sr}^* vs. path loss factor α is depicted in Fig. 6.3. We see that d_{sr}^* is a monotonically increasing function in α. For the worst channel condition, i.e., $\alpha \to \infty$, the relay should be located half-way between source and destination, which is also clear from an intuitive point of view.

Note, that (6.5) is only valid if the condition

$$\Pr\left(\text{“source transmission fails”}\right) \approx \frac{2^{2R} - 1}{\sigma_{\mathrm{sd}}^2 \mathsf{SNR}} \geq \epsilon \tag{6.11}$$

is true.[3] Since we want to achieve a target outage probability of ϵ, it is immediately evident that the outage probability of source transmission must be higher than ϵ. Consequently, this condition inherits a proper design criterion: If ϵ is given and the SNR and the channel state are known, rate can be adapted accordingly.

[3] We applied a Taylor series approximation here.

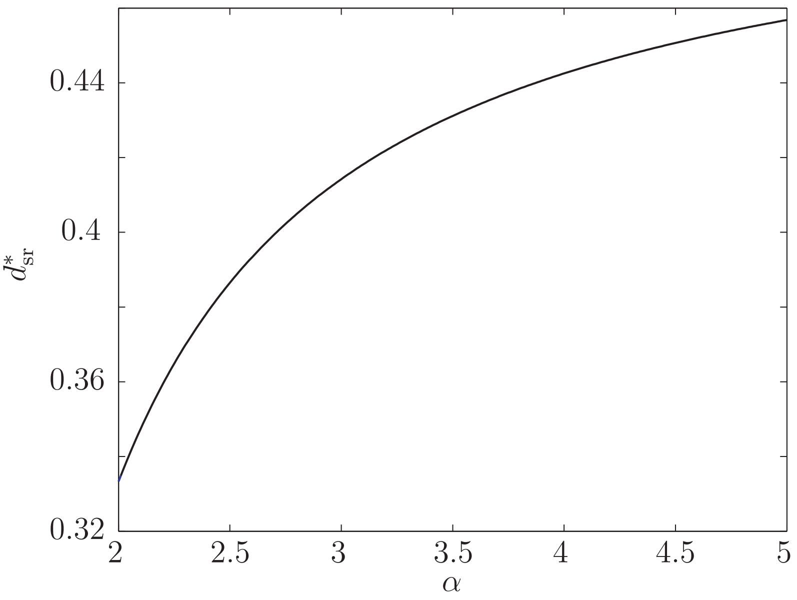

Figure 6.3: Optimal source-relay distance d^*_{sr} vs. path-loss factor α. For $\alpha \to \infty$ d^*_{sr} tends asymptotically to $d^*_{\mathrm{sr}} = 0.5$. $d^*_{\mathrm{sr}}(\alpha = 2) = 1/3$.

6.2.2 Decode-and-Forward vs. Transmit Diversity

One method that creates diversity is multiple-input multiple-output (MIMO) [2–4], where multiple transmit and receive antennas are placed in a way that each antenna faces a different (and most likely independent) channel. However, due to size, cost, and hardware constraints, the maximal number of antennas is limited. If suitable signal processing algorithms are applied for the signal transmission, e.g., Alamouti coding for two transmit antennas [154], enormous performance gains are possible. In this subsection, we compare transmit diversity with two transmit and one receive antenna to cooperative protocols in a network with one source, one relay, and one destination. Especially, we will consider simple decode-and-forward, adaptive decode-and-forward, and incremental relaying with decode-and-forward. For the simple decode-and-forward protocol, the relay always transmits during the second sub-block independent of success or failure of prior source transmission. For the sake of analysis, the relay is supposed to fully decode the source signal. For adaptive decode-and-forward, we allow the relay to be somewhat more intelligent. If it is able to decode the source signal, it retransmits during the second sub-block. If it is not able to decode the source signal, it remains silent. As performance metric we use the ϵ-outage capacity. The following paragraphs are mainly due to [155, 156].

Let us consider a two-antenna transmit diversity system where channel state information is not available at the transmitter. The Alamouti coding scheme which achieves a diversity order of 2 in that case (full order) and has the optimal outage performance for i.i.d. Rayleigh fading channels is applied. Both transmit antennas transmit over the whole transmission block of duration T and with power $P/2$ each. The instantaneous channel capacity for Gaussian inputs then is [37]

$$\mathcal{C}_{\mathrm{TD}} = \log_2\left(1 + ||\mathbf{h}||_2^2 \frac{\mathsf{SNR}}{2}\right), \tag{6.12}$$

where $||\cdot||_2$ is the Euclidean norm,

$$||\mathbf{x}||_2^2 = \sum_i |x_i|^2,$$

and $\mathbf{h} := (h_{\mathrm{sd}}^{(1)}, h_{\mathrm{sd}}^{(2)})^T$ denotes the channel coefficients from antenna 1 and antenna 2, respectively. Channel coefficients are considered to be i.i.d.

If the instantaneous channel capacity cannot serve a required target rate R, an outage event is declared. Therefore,

$$\begin{aligned} p_{\mathrm{out}}^{(\mathrm{TD})}(R) &= \Pr\left(\mathcal{C}_{\mathrm{TD}}(\mathsf{SNR}) < R\right) && (6.13)\\ &= \Pr\left(||\mathbf{h}||_2^2 < \frac{2^R - 1}{\mathsf{SNR}/2}\right) && (6.14)\\ &= F\left(\frac{2^R - 1}{\mathsf{SNR}/2}\right), && (6.15) \end{aligned}$$

where $F_W(w) = \Pr(W < w)$ denotes the cumulative distribution function of the random variable W.[4] Here, W is the sum of two exponentially distributed variables U_0 and U_1 (cf. Lemma 1 in Appendix A). By rearranging the expression on outage probability, the ϵ-outage capacity can be expressed as

$$\mathcal{C}_{\epsilon}^{(\mathrm{TD})} = \log_2\left(1 + \frac{\mathsf{SNR}}{2} F^{-1}(\epsilon)\right). \tag{6.16}$$

In a last step, $F^{-1}(\epsilon)$ has to be determined. This is also done by applying Lemma 1. As we assume a path loss model, it becomes evident that the square magnitudes of both channel gains have the same average value, i.e.,

$$\mathbb{E}(|h_{\mathrm{sd}}^{(1)}|^2) = \mathbb{E}(|h_{\mathrm{sd}}^{(2)}|^2) = \sigma_{\mathrm{sd}}^2.$$

Then, $F^{-1}(\epsilon) = \sqrt{2\epsilon}\sigma_{\mathrm{sd}}^2$ and the ϵ-outage capacity eventually becomes

$$\mathcal{C}_{\epsilon}^{(\mathrm{TD})} = \log_2\left(1 + \mathsf{SNR}\sqrt{\frac{\epsilon}{2}}\sigma_{\mathrm{sd}}^2\right). \tag{6.17}$$

In the following, we investigate the simple decode-and-forward (sDF) protocol in detail. In order to account for the half-duplex constraint, the overall transmission block is divided into two sub-blocks of equal length $T/2$. During the first sub-block

[4] We sometimes skip the subscript W in this subsection for reasons of readability.

the source broadcasts its information to the relay and the destination with power P. In the subsequent sub-block, the relay transmits to the destination with power P, too. The destination then performs maximal ratio combining. If the relay is supposed to fully decode the source signal, the instantaneous channel capacity for Gaussian inputs equals

$$\mathcal{C}_{\mathrm{sDF}} = \frac{1}{2}\min\{\log_2(1+|h_{\mathrm{sr}}|^2\mathsf{SNR}), \log_2(1+||\tilde{\mathbf{h}}||_2^2\mathsf{SNR})\}, \tag{6.18}$$

where $\tilde{\mathbf{h}} := (h_{\mathrm{sd}}, h_{\mathrm{rd}})^T$ describes the channel coefficients between source and destination and relay and destination, respectively. The expression $\log_2(1+|h_{\mathrm{sr}}|^2\mathsf{SNR})$ in the min-function is the maximal rate at which the relay can decode, whereas the expression $\log_2(1+||\tilde{\mathbf{h}}||_2^2\mathsf{SNR})$ describes the maximal rate at which the destination can decode the combination of the source and the relay transmission. The factor 1/2 in front of the min-function takes the half-duplex constraint into account. This equation is well-known in literature and can be found, e.g., in [37].

Accordingly, the outage probability is given by the event that either the relay was not able to decode the source signal reliably or – if the relay could decode – the destination cannot decode the combination of both signals. Hence, it is straightforward to show that for the low SNR regime[5]

$$p_{\mathrm{out}}^{(\mathrm{sDF})}(R) = F\left(\frac{2^{2R}-1}{\mathsf{SNR}}\right), \tag{6.19}$$

where $F_W(w) = \Pr(W < w)$ is the cumulative distribution function of an exponentially distributed random variable W. To sum up, for the simple decode-and-forward protocol, the outage behavior is strongly determined by the quality of the source-to-relay channel and the ability of the relay to (fully) decode the source signal. We can obtain the ϵ-outage capacity by rearranging the equation on outage probability and have

$$\mathcal{C}_{\epsilon}^{(\mathrm{sDF})} = \frac{1}{2}\log_2\left(1+\mathsf{SNR}\,F^{-1}(\epsilon)\right). \tag{6.20}$$

By exploiting the fact that the cumulative distribution function of an exponentially distributed random variable W with mean σ_w^2 satisfies

$$\lim_{\xi\to 0}\frac{1}{g(\xi)}F_W(g(\xi)) = \frac{1}{\sigma_w^2}, \quad g(\xi)\to 0 \text{ as } \xi\to 0, \tag{6.21}$$

$F^{-1}(\epsilon) = \epsilon\sigma_{\mathrm{sr}}^2$. Inserting in (6.20) finally yields

$$\mathcal{C}_{\epsilon}^{(\mathrm{sDF})} = \frac{1}{2}\log_2\left(1+\mathsf{SNR}\epsilon\sigma_{\mathrm{sr}}^2\right). \tag{6.22}$$

[5]The result was presented in [37] for large values of SNR. The Taylor approximation made there can be adopted to the low SNR regime if we consider an additional condition on the rate R. This is similar to the calculation procedure described in Subsection 6.2.1.

Simple decode-and-forward has the following disadvantages. First, the relay has to fully decode the source signal and, second, the relay always has to transmit. In order to overcome these drawbacks one could allow the relay to only partially decode the source signal. However, analysis of this protocol becomes involved. Another solution is that we allow the relay to decide itself upon retransmission of the source signal. On the one hand, if the relay was not able to decode reliably, it remains silent during the second sub-block. On the other hand, if the relay was able to decode, it transmits during the second sub-block. The instantaneous channel capacity of this adaptive decode-and-forward (aDF) protocol becomes

$$\mathcal{C}_{\mathrm{aDF}} = \begin{cases} \frac{1}{2}\log_2(1+|h_{\mathrm{sd}}|^2\mathsf{SNR}) & \text{if relay cannot decode} \\ \frac{1}{2}\log_2(1+||\tilde{\mathbf{h}}||_2^2\mathsf{SNR}) & \text{if relay can decode} \end{cases}, \tag{6.23}$$

where the event "relay cannot decode" is given by

$$|h_{\mathrm{sr}}|^2 < \frac{2^{2R}-1}{\mathsf{SNR}},$$

and the event "relay can decode" is given by

$$|h_{\mathrm{sr}}|^2 \geq \frac{2^{2R}-1}{\mathsf{SNR}}.$$

These events have a simple communication theoretic interpretation. Once the received SNR is above a certain minimum required threshold value, the relay can decode. Of course, the threshold value is determined by the target rate R.

Derivation of the outage probability yields (cf. [37])

$$p_{\mathrm{out}}^{(\mathrm{aDF})}(R) \approx \frac{2\sigma_{\mathrm{rd}}^2+\sigma_{\mathrm{sr}}^2}{2\sigma_{\mathrm{sd}}^2\sigma_{\mathrm{sr}}^2\sigma_{\mathrm{rd}}^2}\left(\frac{2^{2R}-1}{\mathsf{SNR}}\right)^2, \tag{6.24}$$

where the same arguments as for (6.19) hold. Rearranging leads to an ϵ-outage capacity of

$$\mathcal{C}_{\epsilon}^{(\mathrm{aDF})} \approx \frac{1}{2}\log_2\left(1+\mathsf{SNR}\sqrt{\frac{2\sigma_{\mathrm{sd}}^2\sigma_{\mathrm{sr}}^2\sigma_{\mathrm{rd}}^2\epsilon}{2\sigma_{\mathrm{rd}}^2+\sigma_{\mathrm{sr}}^2}}\right). \tag{6.25}$$

In general, the adaptive decode-and-forward protocol has an advantage in diversity order compared to the simple decode-and-forward protocol. This can be seen by the $1/\mathsf{SNR}^2$ dependence of the outage probability, which results in a diversity order of 2. The simple decode-and-forward protocol, in contrast to that, achieves a diversity order of 1 [37].

We now illustrate our results and compare the investigated cooperative relaying protocols to transmit diversity. All distances have been normalized to the source-to-destination distance, i.e., $d_{\mathrm{sd}} = 1$. First, let us consider Fig. 6.4. Here, the ϵ-outage capacities in bit/s/Hz vs. SNR in dB are illustrated. An outage probability of $\epsilon = 10^{-4}$

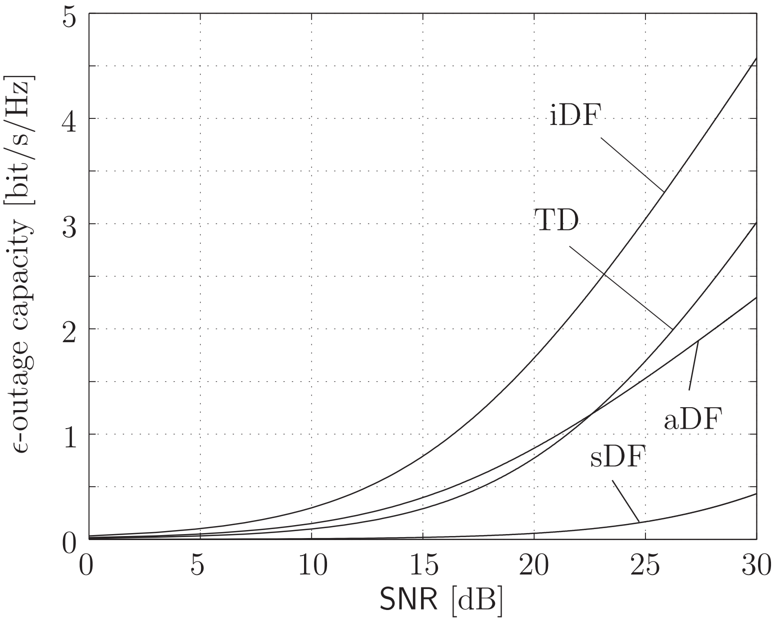

Figure 6.4: ϵ-outage capacity in bit/s/Hz vs. SNR in dB for an outage probability of $\epsilon = 10^{-4}$ and $\alpha = 3$. The relay has been placed at $d_{\mathrm{sr}} = 0.5$.

was selected and the path loss factor was set to $\alpha = 3$. In addition to that, the relay was placed in the middle of a straight line between source and destination, i.e., $d_{\mathrm{sr}} = 0.5$. It can be seen that for high values of SNR, simple decode-and-forward (sDF) shows the weakest performance. Adaptive decode-and-forward (aDF) and transmit diversity (TD) achieve approximately the same values of ϵ-outage capacity for SNR values of up to 23 dB. From then on transmit diversity outperforms adaptive decode-and-forward with increasing SNR. The best performance by far shows incremental relaying with decode-and-forward (iDF). This protocol outperforms the other protocols for the whole considered SNR range.

In order to compare the ϵ-outage capacities for different relay locations, we use the following definition.

Definition 1 *The ratio between the ϵ-outage capacities of cooperative relaying protocols and the ϵ-outage capacity of transmit diversity for the same value of ϵ is defined as*

$$\Delta(\epsilon) := \frac{\mathcal{C}_\epsilon^{(i)}}{\mathcal{C}_\epsilon^{(\mathrm{TD})}}, \tag{6.26}$$

where $i \in \{\mathrm{sDF,aDF,iDF}\}$.

It can easily be seen that a cooperative relaying protocol outperforms transmit diversity whenever the ratio is greater than 1.

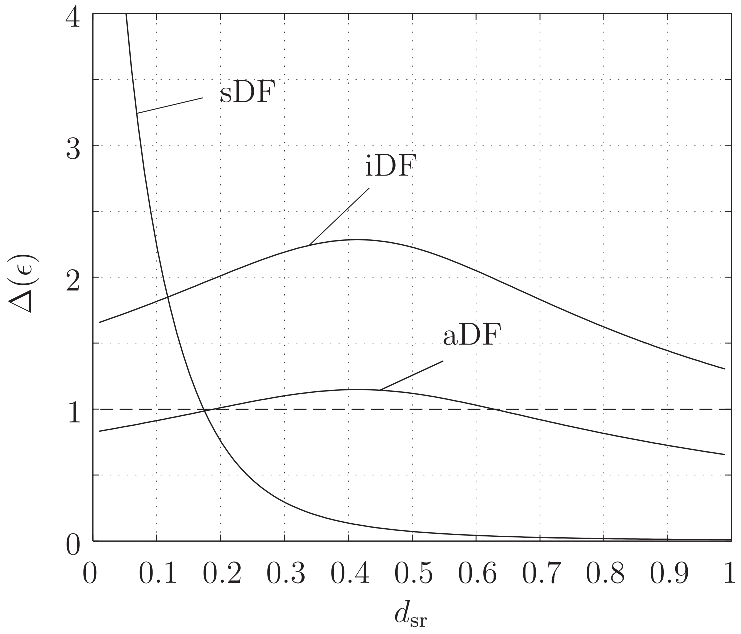

Figure 6.5: Ratio between ϵ-outage capacity of relaying protocols and transmit diversity vs. source-relay distance d_{sr} for $\epsilon = 10^{-4}$, $\mathsf{SNR} = 20$ dB, and $\alpha = 3$. Curve sections that are above the dashed line indicate regions in which the cooperative protocols outperform transmit diversity.

Fig. 6.5 shows the results for $\epsilon = 10^{-4}$, $\mathsf{SNR} = 20$ dB, and $\alpha = 3$. Simple decode-and-forward outperforms transmit diversity only when the relay is located close to the source and up to a relay location of $d_{\text{sr}} \approx 0.18$. An interesting aspect is the fact that $\Delta(\epsilon)$ tends to infinity for $d_{\text{sr}} \to 0$. Though this seems a bit strange from a first point of view, there is a reasonable explanation. In the considered system model, the variances of the channel gains are proportional to $d^{-\alpha}$. As we have seen in (6.22), the ϵ-outage capacity of simple decode-and-forward only depends on σ_{sr}^2 and, thus, the illustrated behavior makes sense. Adaptive decode-and-forward slightly outperforms transmit diversity for the given parameter setting and $d_{\text{sr}} \in [0.18; 0.61]$. The best performance is achieved by incremental relaying with decode-and-forward, which is beneficial to transmit diversity for all relay locations between the source and the destination.

Until now, we have dealt with a one-dimensional geometry. Though this gives good hints for the understanding of the performance of different protocols, it does not represent a practical mobile communications system, where terminals most likely change location all the time. Therefore, a two-dimensional geometry can give much more insight into the performance of the investigated protocols. In Fig. 6.6 regions are shown in which the ϵ-outage capacities of the cooperative relaying protocols are larger than the ϵ-outage capacity of transmit diversity. The following parameter setting was used: $\epsilon = 10^{-4}$, $\mathsf{SNR} = 20$ dB, and $\alpha = 3$. As expected from the previous results,

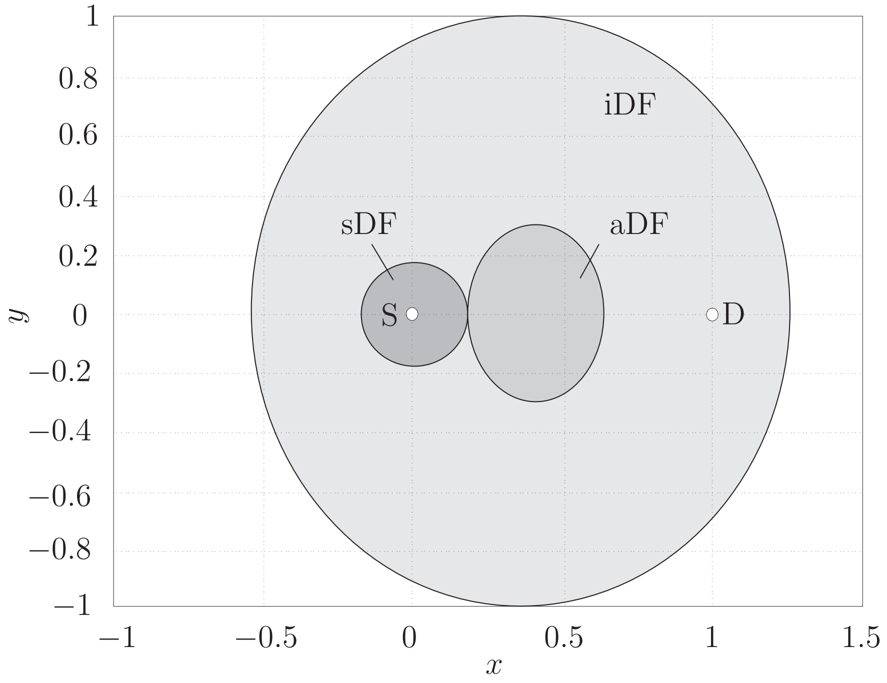

Figure 6.6: Regions in which ϵ-outage capacity of the cooperative relaying protocols outperforms ϵ-outage capacity of transmit diversity for $\epsilon = 10^{-4}$, $\mathsf{SNR} =$ 20 dB, and $\alpha = 3$.

simple decode-and-forward is beneficial if the relay is located close to the source. This is, however, not the case for adaptive decode-and-forward. We see that the region where adaptive decode-and-forward is beneficial to transmit diversity is larger. Especially, the region is located between the source and the destination with being slightly closer to the source.[6] The ϵ-outage capacity region for incremental relaying with decode-and-forward is pretty large. As can be seen, even if the relay is located further away from the source than the destination (even behind the destination), transmit diversity can be outperformed. For this simulation, we selected a rather small outage probability for the source-to-destination transmission (only 2ϵ). It is obvious that as the outage probability is increased, this region will decrease in size. This effect is illustrated in Fig. 6.7. It depicts the region where the ϵ-outage capacity of incremental relaying with decode-and-forward outperforms the ϵ-outage capacity of transmit diversity for different values of Pr(“source transmission fails”) as a function of ϵ (precisely we have 100ϵ, 1000ϵ, and 10000ϵ). It can be seen that the region gets smaller with increasing outage probability of the source-to-destination channel. Since ϵ has been set to 10^{-4}, the highest value we are allowed to choose is 10000ϵ (cf. (6.11)). For this case, incremental relaying with decode-and-forward turns into adaptive decode-and-forward. Hence, it achieves the same performance (which can

[6] This is due to the fact that decode-and-forward in general works better if the relay is closer to the source.

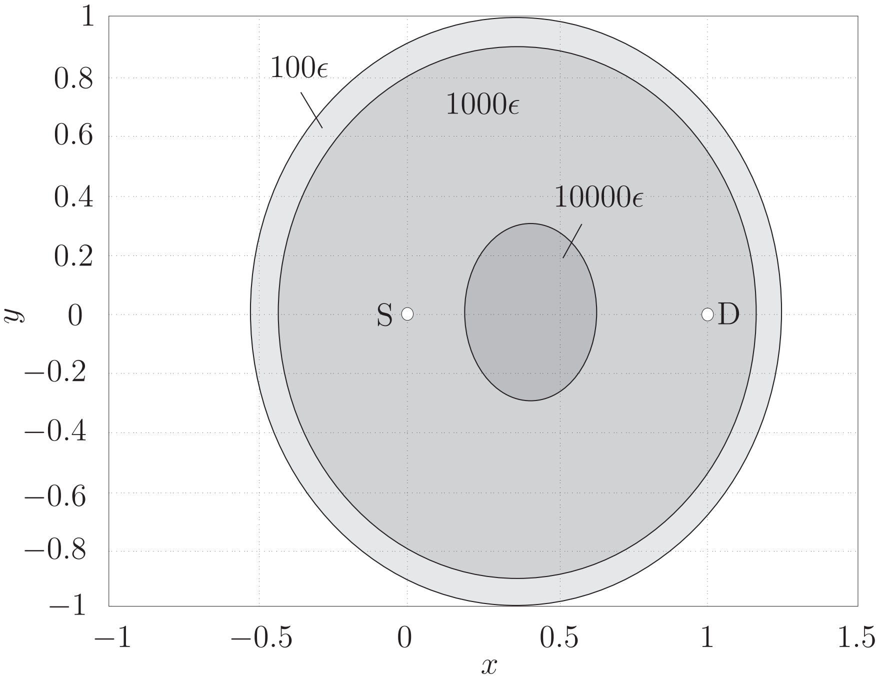

Figure 6.7: Regions where ϵ-outage capacity of incremental relaying with decode-and-forward outperforms ϵ-outage capacity of transmit diversity for different values of Pr(“source transmission fails”) as a function of ϵ. Parameters are $\epsilon = 10^{-4}$, $\mathsf{SNR} = 20$ dB, and $\alpha = 3$.

be seen by comparing the region for adaptive decode-and-forward in Fig. 6.6 to the region for incremental relaying with decode-and-forward and 10000ϵ in Fig. 6.7).

6.2.3 Bursty Amplify-and-Forward

Avestimehr and Tse derived the ϵ-outage capacity of the fading relay channel without feedback in [152]. They showed that the “original” version of amplify-and-forward is not applicable for low values of SNR, since then the relay actually amplifies noise, which complicates decoding at the destination. In order to overcome this drawback, they proposed a bursty version of amplify-and-forward (BAF) and showed that this protocol is outage optimal for the frequency division duplex channel without feedback. This confirms results presented by Verdú in [157]. He revises the fact that the capacity of an ideal bandlimited additive white Gaussian noise channel can be approached by pulse position modulation with a very low duty cycle in the low power regime. This fact dates back to a publication by Golay in 1949 [158]. Additionally, it was demonstrated in [159] that BAF is also outage optimal for a wide class of independent channels. It is pointed out, that this is true if the distribution functions of the channels are smooth.[7]

[7] For more information on the smoothness of the distribution functions the interested reader is referred to [159].

With respect to the related work mentioned in the previous paragraph, we derive the ϵ-outage capacity of an incremental relaying protocol with BAF [80]. We first investigate a three-node network consisting of one source, one relay, and one destination, where again the source and the relay transmit in orthogonal time slots. The main ideas of AF and BAF with incremental relaying are illustrated in Fig. 6.8, respectively. The overall transmission block is divided into two sub-blocks of equal length. During the first sub-block, the source broadcasts its signal with power P to the destination and the relay (Fig. 6.8a). After that, the destination sends a one-bit feedback (FB) indicating success or failure of source transmission. Depending on the feedback either the source transmits its next message or the relay retransmits an amplified version of its own receive signal, i.e., of the source's first message corrupted with noise. As stated before, AF possesses poor performance in the low SNR regime. Performance can be improved enormously if the source and the relay transmit bursts during their corresponding sub-blocks, i.e., both transmit only for a fraction of $(\tau T)/2$ and with power P/τ ($\tau \to 0$) in order to meet the average power constraint (Fig. 6.8b).[8] This is then comparable to pulse position modulation with a very low duty cycle (see [157]).

We now derive the ϵ-outage capacity of BAF with incremental relaying. The way is similar to the one presented in Subsection 6.2.1 (see also [79]). First, we derive an expression for the ϵ-outage capacity without feedback and then introduce a pre-log factor that takes feedback into account. The instantaneous channel capacity for a half-duplex relay channel with BAF can be written as

$$\mathcal{C}_{\mathrm{BAF}}(\mathsf{SNR},\tau) = \frac{\tau}{2}\log_2\left(1+\frac{\mathsf{SNR}}{\tau}\left(|h_{\mathrm{sd}}|^2+\frac{|h_{\mathrm{rd}}|^2|h_{\mathrm{sr}}|^2}{|h_{\mathrm{rd}}|^2+|h_{\mathrm{sr}}|^2+\tau/\mathsf{SNR}}\right)\right). \qquad (6.27)$$

In contrast to the expression given in [152, 159], we consider an additional pre-log factor of 1/2 due to the half-duplex constraint and use the logarithm to the base 2 in order to express capacity in bit/s/Hz. We define

$$\mathrm{A}(\mathbf{h},\tau) := |h_{\mathrm{sd}}|^2+\frac{|h_{\mathrm{rd}}|^2|h_{\mathrm{sr}}|^2}{|h_{\mathrm{rd}}|^2+|h_{\mathrm{sr}}|^2+\tau/\mathsf{SNR}} \qquad (6.28)$$

and drop the dependence on $\mathbf{h} = (|h_{\mathrm{sd}}|^2,|h_{\mathrm{sr}}|^2,|h_{\mathrm{rd}}|^2)$ and τ in the following for the sake of description. An outage is declared whenever $\mathcal{C}_{\mathrm{BAF}}(\mathsf{SNR},\tau)$ is smaller than the target rate R. Accordingly,

$$p_{\mathrm{out}}^{(\mathrm{BAF})} = \Pr\left(\mathrm{A} < \frac{2^{2R/\tau}-1}{\mathsf{SNR}/\tau}\right).$$

Since we are interested in a target error rate that approaches zero in the low SNR regime, i.e., $\epsilon \to 0$ for $\mathsf{SNR} \to 0$, we have to choose τ in a suitable fashion, so that the right hand side within the $\Pr(\cdot)$ expression goes to zero. A proper choice of τ

[8] Note that here the time fraction τ tends to 0 in contrast to Chapter 4, was it was optimized in order to maximize capacity.

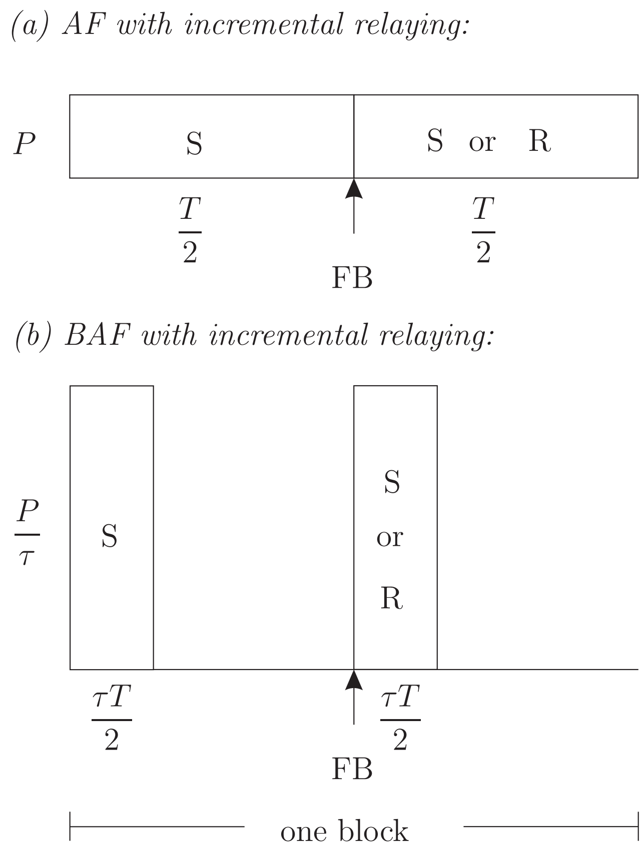

Figure 6.8: Transmission model for incremental relaying with AF and BAF. If the source transmission succeeded (feedback FB = 1), the source transmits during the second sub-block, too. If the source transmission failed (feedback FB = 0), the relay aids communication during the second sub-block.

was given in [152] to be $\tau = \sqrt{R\,\mathsf{SNR}}$. Inserting this into the above equation yields $\sqrt{R/\mathsf{SNR}} \to 0$. Hence, outage probability becomes

$$p_{\text{out}}^{(\text{BAF})} = \Pr\left(\text{A} < \tilde{g}\right)$$

where $\tilde{g}$ is given by

$$\tilde{g} = \sqrt{\frac{R}{\mathsf{SNR}}}\left(2^{2\sqrt{R/\mathsf{SNR}}} - 1\right).$$

In order to derive the ϵ-outage capacity, we apply Lemma 2 (see Appendix A). We can write

$$\lim_{\substack{\epsilon\to 0\\ \mathsf{SNR}\to 0\\ \tilde{g}\to 0}} \frac{p_{\text{out}}^{(\text{BAF})}}{\tilde{g}^2} = \frac{\sigma_{\text{rd}}^2 + \sigma_{\text{sr}}^2}{2\sigma_{\text{sd}}^2\sigma_{\text{rd}}^2\sigma_{\text{sr}}^2}. \tag{6.29}$$

The ϵ-outage capacity in bit/s/Hz of BAF *without* incremental relaying becomes

$$\mathcal{C}_{\epsilon}^{(\text{BAF})} \approx \frac{1}{2}\log_2\left(1 + \mathsf{SNR}\sqrt{\frac{2\sigma_{\text{sd}}^2\sigma_{\text{rd}}^2\sigma_{\text{sr}}^2\epsilon}{\sigma_{\text{rd}}^2 + \sigma_{\text{sr}}^2}}\right), \tag{6.30}$$

where we used the approximation[9]

$$\frac{x}{\log_2(e)} \approx \frac{1}{2}\log_2(1+x)\,. \tag{6.31}$$

As mentioned before, the variability of the transmission rate is not considered in (6.30). This variability is due to the feedback from the destination to the source and the relay. To account for that, the average amount of transmitted sub-blocks required for one source message must be considered. If the source transmission was successful during the first sub-block, only one sub-block is required (independent of the relay). However, if the source transmission failed, the relay transmits during the second sub-block. If the destination is still not able to decode after the second sub-block, an outage event will be declared. As in Subsection 6.2.1, $\mathbb{E}(N)$ describes the average amount of transmission phases required for one specific message. We can now express the ϵ-outage capacity of BAF *with* incremental relaying denoted by the superscript (iBAF) as

$$\mathcal{C}_\epsilon^{(\text{iBAF})} = \frac{2}{\mathbb{E}(N)}\mathcal{C}_\epsilon^{(\text{BAF})} \tag{6.32}$$

$$\approx \frac{1}{\mathbb{E}(N)}\log_2\left(1+\mathsf{SNR}\sqrt{\frac{2\sigma_{\text{sd}}^2\sigma_{\text{rd}}^2\sigma_{\text{sr}}^2\epsilon}{\sigma_{\text{rd}}^2+\sigma_{\text{sr}}^2}}\right). \tag{6.33}$$

The factor $2/\mathbb{E}(N)$ in (6.32) describes possible savings in the required amount of sub-blocks for transmitting a specific source message. If only one sub-block is required, i.e., source transmission was successful, a gain of 2 can be achieved ($\mathbb{E}(N) = 1$), since then the source can transmit its next message after reception of the positive feedback from the destination ($\text{FB} = 1$)[10]. If both sub-blocks are required for transmitting one and the same message, i.e., source transmission failed and the relay aids communication ($\text{FB} = 0$), we perform at least as good as the BAF relaying protocol without feedback ($\mathbb{E}(N) = 2$).

Furthermore, if we consider a one-dimensional geometry, where the relay is placed on a straight line between the source and the destination, and the path loss model presented in Subsection 6.2.1, it can easily be verified that the optimal relay location that maximizes the ϵ-outage capacity is $d_{\text{sr}}^* = 0.5$ independent of the path loss factor α [80].

6.2.4 Comparison to Cut-Set Bound

In this subsection we compare the performance of incremental relaying employing either DF or BAF to the cut-set bound[11] (CSB) [52, Theorem 14.10.1, p. 445]. For that purpose, we first derive the ϵ-outage capacity of the CSB and then define the

[9] This approximation is related to the approximation $\ln(1+x) \approx x$ for small values of x.
[10] Recall that we assume block fading.
[11] The cut-set bound is sometimes also referred to as the max-flow min-cut theorem.

comparison ratio $\Delta(\epsilon)$ comparable to (6.26). Since the CSB is an upper bound on the flow of information in any network that consists of multiple terminals, it clearly is an upper bound to incremental relaying. Hence, the best we could do is to achieve the cut-set bound.

The cut-set bound of the relay channel with Gaussian codebooks is

$$\mathcal{C}_{\text{CSB}} = \min\{\log_2(1 + (|h_{\text{sd}}|^2 + |h_{\text{sr}}|^2)\mathsf{SNR}), \log_2(1 + (|h_{\text{sd}}|^2 + |h_{\text{rd}}|^2)\mathsf{SNR})\}. \quad (6.34)$$

We now follow exactly the same steps that we used for incremental relaying in order to get an expression of the ϵ-outage capacity. The outage probability in the low SNR regime, where again the condition $g \to 0$ for $\mathsf{SNR} \to 0$ must be met, becomes

$$\begin{aligned}\lim_{\substack{\epsilon\to 0\\ \mathsf{SNR}\to 0}} \frac{p_{\text{out}}^{(\text{CSB})}}{g^2} &= \lim_{\substack{\epsilon\to 0\\ \mathsf{SNR}\to 0}} \left\{ \frac{\Pr(|h_{\text{sd}}|^2 + |h_{\text{sr}}|^2 < g)}{g^2} \right. \\ &\quad \left. + \frac{\Pr(|h_{\text{sd}}|^2 + |h_{\text{sr}}|^2 \geq g)}{1} \frac{\Pr(|h_{\text{sd}}|^2 + |h_{\text{rd}}|^2 < g)}{g^2} \right\} \\ &= \frac{1}{2\sigma_{\text{sd}}^2\sigma_{\text{sr}}^2} + 1 \cdot \frac{1}{2\sigma_{\text{sd}}^2\sigma_{\text{rd}}^2} \\ &= \frac{\sigma_{\text{rd}}^2 + \sigma_{\text{sr}}^2}{2\sigma_{\text{sd}}^2\sigma_{\text{sr}}^2\sigma_{\text{rd}}^2}. \end{aligned} \quad (6.35)$$

For the ϵ-outage capacity we then have

$$\mathcal{C}_{\epsilon}^{(\text{CSB})} \geq \frac{1}{1+\epsilon} \log_2\left(1 + \mathsf{SNR}\sqrt{\frac{2\sigma_{\text{sd}}^2\sigma_{\text{sr}}^2\sigma_{\text{rd}}^2\epsilon}{\sigma_{\text{rd}}^2 + \sigma_{\text{sr}}^2}}\right), \quad (6.36)$$

where we have applied $\mathbb{E}(N) \leq 1+\epsilon$. This can be explained as follows. Since our aim is to have an overall outage probability lower then or equal to ϵ, the outage probability for the first sub-block clearly is higher than ϵ. Hence, $\Pr(\mathcal{X}) \geq \epsilon$, and we get a tighter upper bound on the ϵ-outage capacity by setting $\mathbb{E}(N) \leq 1 + \epsilon$.

In order to compare the ϵ-outage capacities of incremental relaying protocols to the ϵ-outage capacity of the CSB, we use the following performance criterion.

Definition 2 *The ratio between the ϵ-outage capacities of incremental relaying protocols and the ϵ-outage capacity of the cut-set bound for the same value of ϵ is defined as*

$$\Delta(\epsilon) := \frac{\mathcal{C}_{\epsilon}^{(i)}}{\mathcal{C}_{\epsilon}^{(\text{CSB})}} \leq 1, \quad (6.37)$$

where $i \in \{\text{iDF,iBAF}\}$.

Let us first consider incremental relaying with decode-and-forward. Applying (6.5) and (6.36), we get

$$\Delta(\epsilon) \approx \sqrt{\frac{\sigma_{\text{rd}}^2 + \sigma_{\text{sr}}^2}{2\sigma_{\text{rd}}^2 + \sigma_{\text{sr}}^2}} = \sqrt{\frac{\left(\frac{d_{\text{sr}}}{d_{\text{rd}}}\right)^{\alpha} + 1}{2\left(\frac{d_{\text{sr}}}{d_{\text{rd}}}\right)^{\alpha} + 1}}, \quad (6.38)$$

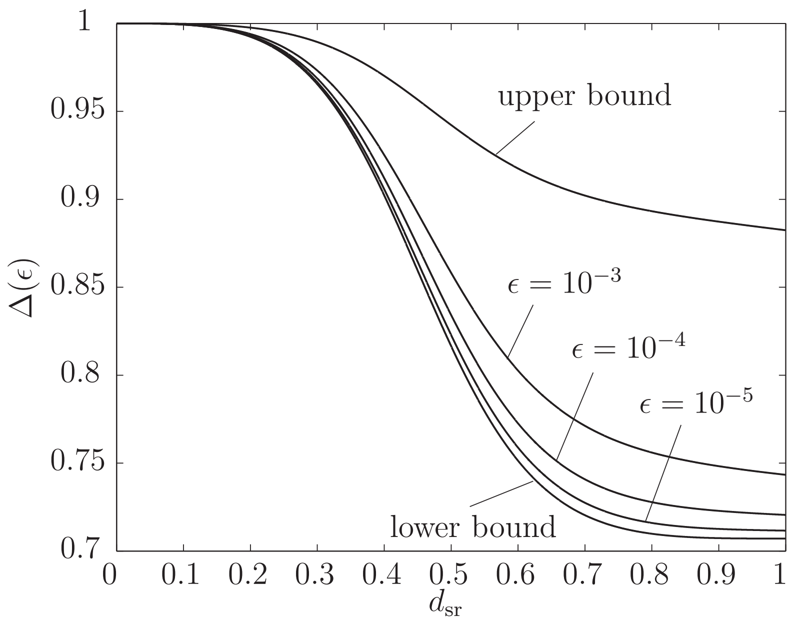

Figure 6.9: Ratio $\Delta(\epsilon)$ of the ϵ-outage capacity of incremental relaying with decode-and-forward to the cut-set bound for $g \to 0$ and $\alpha = 3$.

where we used the approximation $\ln(1+x) \approx x$ for small values of x and $\Pr(\mathcal{X}) \approx \epsilon$. It can readily be seen that $\Delta(\epsilon) \in [1/\sqrt{2}, 1]$ for $g \to 0$. The value $1/\sqrt{2}$ describes the case when the relay is placed close to the destination, whereas the value 1 represents the case when the relay is located close to the source.

Fig. 6.9 illustrates the ratio $\Delta(\epsilon)$ of the ϵ-outage capacity of incremental relaying for decode-and-forward to the ϵ-outage capacity of the CSB for $g \to 0$. The protocol shows its weakest performance when the relay is placed close to the destination. However, when it is located close to the source, incremental relaying with decode-and-forward is optimal in a sense that its ϵ-outage capacity achieves that of the CSB.

Next consider incremental relaying with BAF. The ratio $\Delta(\epsilon)$ becomes

$$\Delta(\epsilon) \leq \frac{1+\epsilon}{\mathbb{E}(N)} = \frac{1+\epsilon}{1+\Pr(\text{“source transmission fails”})}. \tag{6.39}$$

The outage probability of source transmission in the low SNR regime can easily be derived. We get

$$\begin{aligned}
\Pr(\text{“source transmission fails”}) &= \Pr\left(\frac{\tau}{2}\log_2\left(1+|h_{\text{sd}}|^2\frac{\mathsf{SNR}}{\tau}\right) < R\right) \\
&\approx \Pr\left(|h_{\text{sd}}|^2 < \frac{\log_2(e)\,R}{\mathsf{SNR}}\right) \\
&\approx \frac{\log_2(e)\,R}{\sigma_{\text{sd}}^2\mathsf{SNR}},
\end{aligned}$$

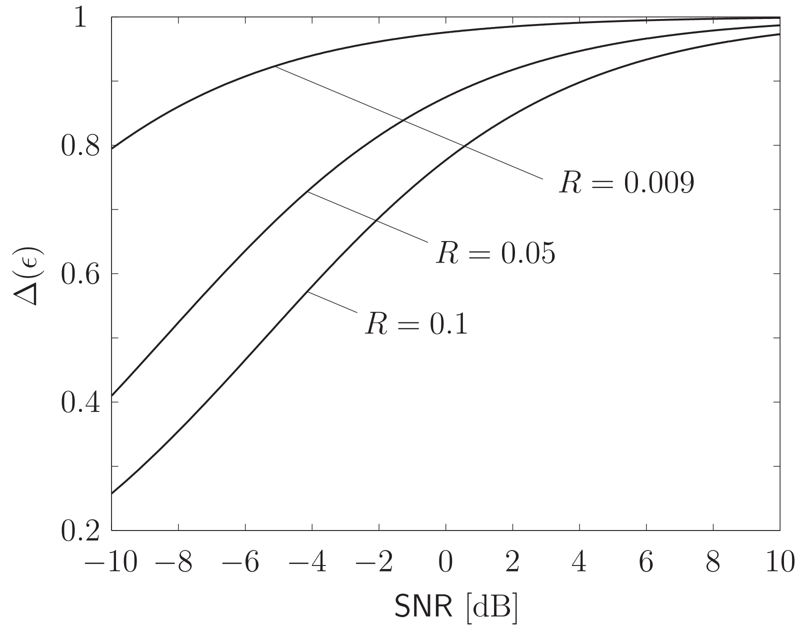

Figure 6.10: Ratio $\Delta(\epsilon)$ vs. SNR for $\epsilon = 0.001$ and different values of rate R in bit/s/Hz. The distance source-destination was normalized to 1.

where we again set $\tau = \sqrt{R\,\mathsf{SNR}}$ (cf. Subsection 6.2.3), let $\sqrt{R/\mathsf{SNR}} \to 0$, and used the approximation given in (6.31). Since Pr("source transmission fails") must be higher than ϵ, we get an upper bound on the target error rate of

$$\epsilon \leq \frac{\log_2(e)\,R}{\sigma_{\mathrm{sd}}^2 \mathsf{SNR}}. \tag{6.40}$$

Fig. 6.10 illustrates the ratio $\Delta(\epsilon)$ vs. SNR in dB for $\epsilon = 0.001$. The distance source-destination was normalized to 1, i.e., $\sigma_{\mathrm{sd}}^2 = 1$. Obviously, $\Delta(\epsilon)$ is a monotonically increasing function in SNR. We see that the values of $\Delta(\epsilon)$ for a given SNR will decrease, if the rate R is increased.

6.3 Extension to K Relays

In this section, we consider parallel relay networks that consist of one source, one destination, and an arbitrary number of K relays (see Fig. 2.6). We divide one transmission block into $K+1$ sub-blocks of equal length, i.e., the duration of each sub-block equals $T/(K+1)$. In order to have the same amount of information transmitted compared to direct transmission, where one source transmits with rate R over the whole block length, the initial transmission rate now is $(K+1)R$. After the source transmission, the destination informs all transmitting nodes about success or failure of transmission by sending a one-bit feedback. If the source transmission was successful, the next sub-block is occupied by the source that starts transmitting its next message. If the source transmission failed, the first relay will transmit during the second sub-block.[12] Then, the destination accumulates both SNR values and tries to decode. Again, the destination indicates whether the combined transmission was successful or not by a one-bit feedback. If transmission was successful, the source starts transmitting. If it was not, then the second relay transmits. This procedure continues until the K-th relay has transmitted. If the destination is still not able to decode, an outage is declared. We can immediately conclude that such a procedure will lead to a maximal gain of $K + 1$ (compared to a cooperative protocol without incremental relaying), if source transmission in the first sub-block is successful. If all relays have to transmit, then the gain reduces to 1.

We stress that there are protocols in literature where several sources transmit through several half-duplex relays to several destinations. However, the main focus of our work is on interference-free transmission and a proper analysis of feedback. The interested reader is referred to [160] and the references therein. First, incremental relaying with decode-and-forward is investigated. After that, we deal with incremental relaying with bursty amplify-and-forward before considering the cut-set bound.

The calculation of the outage probability of incremental relaying for an arbitrary number of relays gets involved. Normally, one would have to investigate all possibilities of how information can be sent from the source over the relays to the destination. In a network with K relays, this leads to 2^K different cuts. A general expression on outage probability for parallel relay networks with selection combining at the destination was derived in [161] and is given in Appendix D. In this case, an outage is declared when all connections from the source to the relay via all possible relays fail. For decode-and-forward, this means that either the source transmission to the destination and to the relays fails or, if a relay was able to decode the source signal, the transmission from the relay to the destination fails. Clearly, this scheme performs worse than a scheme that employs maximal ratio combining at the destination, since only the SNR of the strongest branch is considered for decoding rather than the accumulation of all incoming branches.

In contrast to that, we apply MRC and simplify the calculation for incremental relaying with decode-and-forward by making the assumption that either all K relays can decode the source message or none can decode. A lower bound on the outage probability is then given by

$$p_{\text{out}}^{(\text{DF})} \geq \Pr(|h_{\text{sd}}|^2 < g_K) \prod_{k=1}^{K} \Pr(|h_{\text{sr}_k}|^2 < g_K) + \prod_{k=1}^{K} \Pr(|h_{\text{sr}_k}|^2 \geq g_K) \Pr(|h_{\text{sd}}|^2 + \sum_{k=1}^{K} |h_{\text{r}_k\text{d}}|^2 < g_K),$$

[12] The ordering of the relay nodes can be done with respect to several performance characteristics. Confer to Chapter 3 for more information on that topic.

where we used

$$g_K(R,\mathsf{SNR}) = \frac{2^{(K+1)R} - 1}{\mathsf{SNR}} \tag{6.41}$$

and again dropped the dependence on R and SNR for the sake of description. Applying Lemma 1 yields

$$\lim_{\substack{\epsilon\to 0\\ \mathsf{SNR}\to 0\\ g_K\to 0}} \frac{p_{\text{out}}^{(\text{DF})}}{g_K^{K+1}} \geq \frac{(K+1)!\prod_{k=1}^{K}\sigma_{\text{r}_k\text{d}}^2 + \prod_{k=1}^{K}\sigma_{\text{sr}_k}^2}{(K+1)!\sigma_{\text{sd}}^2\prod_{k=1}^{K}\sigma_{\text{r}_k\text{d}}^2\sigma_{\text{sr}_k}^2}. \tag{6.42}$$

Manipulation eventually leads to an upper bound on the ϵ-outage capacity of incremental relaying with decode-and-forward of

$$\mathcal{C}_\epsilon^{(\text{iDF})} \leq \frac{1}{\mathbb{E}_K(N)}\log_2\left(1 + \mathsf{SNR}\sqrt[K+1]{\frac{(K+1)!\sigma_{\text{sd}}^2\prod_{k=1}^{K}\sigma_{\text{r}_k\text{d}}^2\sigma_{\text{sr}_k}^2\epsilon}{(K+1)!\prod_{k=1}^{K}\sigma_{\text{r}_k\text{d}}^2 + \prod_{k=1}^{K}\sigma_{\text{sr}_k}^2}}\right), \tag{6.43}$$

where

$$\mathbb{E}_K(N) = 1 + \sum_{k=1}^{K}\Pr(\mathcal{Z}_k) \quad \text{and} \quad \mathcal{Z}_k = \{|h_{\text{sd}}|^2 + \sum_{l=1}^{k-1}|h_{\text{r}_l\text{d}}|^2 < g_K\}. \tag{6.44}$$

The event $\mathcal{Z}_k$ describes the accumulation of SNR at the destination and takes into account that the relays transmit in a successive manner [79].

The basic transmission model for incremental relaying with bursty amplify-and-forward is shown in Fig. 6.11. Again, the main idea is that the destination transmits negative feedbacks (FB = 0) until it has accumulated sufficient SNR to decode. Therefore, either the source S or the first relay R_1 transmits in the second sub-block depending on the success or failure of the source transmission during the first sub-block. In the third sub-block either the source S or the first relay R_1 or the second relay R_2 transmits depending on whether the previously accumulated SNR was sufficiently high so that the destination could decode and so on. Once the destination has accumulated enough SNR to decode, it indicates that no more relay transmissions are required by transmitting a positive feedback (FB = 1). When this happens, the source occupies the next sub-block and starts transmitting its new message. An outage is declared, when the SNR at the destination is still not sufficient to decode after the K-th relay has transmitted.

The instantaneous channel capacity of bursty amplify-and-forward with K relays can be expressed as

$$\mathcal{C}_{\text{BAF}}(\mathsf{SNR},\tau) = \frac{\tau}{K+1}\log_2\left(1 + \frac{\mathsf{SNR}}{\tau}\text{A}_K\right), \tag{6.45}$$

where we used the substitution

$$\text{A}_K := |h_{\text{sd}}|^2 + \sum_{k=1}^{K}\frac{|h_{\text{r}_k\text{d}}|^2|h_{\text{sr}_k}|^2}{|h_{\text{r}_k\text{d}}|^2 + |h_{\text{sr}_k}|^2 + \tau/\mathsf{SNR}}$$

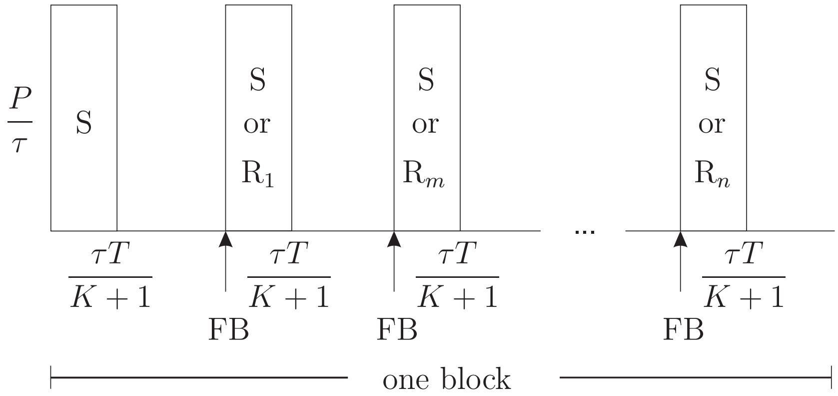

Figure 6.11: Transmission model for BAF with incremental relaying and an arbitrary number of relay nodes. R_m, $m \in \{1,2\}$, and R_n, $n \in \{1,2,\ldots,K\}$, describe the transmitting relay depending on prior transmissions.

and again dropped the dependence on the $(K+1)$-tuple $\mathbf{h}_K = (|h_{\mathrm{sd}}|^2,|h_{\mathrm{sr}_k}|^2,|h_{\mathrm{r}_k\mathrm{d}}|^2)$, $k = 1,\ldots,K$, and τ for the sake of description. For small values of SNR, a proper approximation is given by

$$\mathcal{C}_{\mathrm{BAF}}(\mathsf{SNR},\tau) \approx \frac{\mathsf{SNR}}{K+1} \log_2(e)\, \mathrm{A}_K. \tag{6.46}$$

Eventually, the outage probability becomes

$$p_{\mathrm{out}}^{(\mathrm{BAF})} \approx \Pr\left(\mathrm{A}_K < \tilde{g}_K\right), \tag{6.47}$$

where we used $\tilde{g}_K = (K+1)R/(\log_2(e)\,\mathsf{SNR})$. Due to the structure of A_K, the solution gets involved. However, there exists an accurate approximation. We apply the inequality

$$\min\{x,y\} \geq \frac{xy}{x+y+\delta}, \quad x,y \in \mathbb{R}^+,$$

where δ is an arbitrarily small and positive number, to upper bound A_K. By defining

$$\tilde{\mathrm{A}}_K := |h_{\mathrm{sd}}|^2 + \sum_{k=1}^{K} \min\{|h_{\mathrm{r}_k\mathrm{d}}|^2,|h_{\mathrm{sr}_k}|^2\},$$

we get

$$p_{\mathrm{out}}^{(\mathrm{BAF})} \geq \Pr\left(\tilde{\mathrm{A}}_K < \tilde{g}_K\right). \tag{6.48}$$

Using results given in [152] and applying (6.31) finally yields

$$\lim_{\substack{\epsilon\to 0\\ \mathsf{SNR}\to 0\\ \tilde{g}_K\to 0}} \frac{p_{\mathrm{out}}^{(\mathrm{BAF})}}{\tilde{g}_K^{K+1}} \geq \frac{\prod_{k=1}^{K}(\sigma^2_{\mathrm{r}_k\mathrm{d}} + \sigma^2_{\mathrm{sr}_k})}{(K+1)!\sigma^2_{\mathrm{sd}} \prod_{k=1}^{K} \sigma^2_{\mathrm{r}_k\mathrm{d}}\sigma^2_{\mathrm{sr}_k}}$$

and, therefore, the ϵ-outage capacity of bursty amplify-and-forward *without* incremental relaying is upper bounded by

$$\mathcal{C}_{\epsilon}^{(\mathrm{BAF})} \leq \frac{1}{K+1} \log_2 \left(1 + \mathsf{SNR} \sqrt[K+1]{\frac{(K+1)! \sigma_{\mathrm{sd}}^2 \prod_{k=1}^{K} \sigma_{\mathrm{r}_k\mathrm{d}}^2 \sigma_{\mathrm{sr}_k}^2 \epsilon}{\prod_{k=1}^{K} (\sigma_{\mathrm{r}_k\mathrm{d}}^2 + \sigma_{\mathrm{sr}_k}^2)}} \right). \tag{6.49}$$

The ϵ-outage capacity *with* incremental relaying then is derived by introducing a factor $(K+1)/\mathbb{E}_K(N)$ [80]. Accordingly,

$$\mathcal{C}_{\epsilon}^{(\mathrm{iBAF})} \leq \frac{1}{\mathbb{E}_K(N)} \log_2 \left(1 + \mathsf{SNR} \sqrt[K+1]{\frac{(K+1)! \sigma_{\mathrm{sd}}^2 \prod_{k=1}^{K} \sigma_{\mathrm{r}_k\mathrm{d}}^2 \sigma_{\mathrm{sr}_k}^2 \epsilon}{\prod_{k=1}^{K} (\sigma_{\mathrm{r}_k\mathrm{d}}^2 + \sigma_{\mathrm{sr}_k}^2)}} \right). \tag{6.50}$$

In the above equation, $\mathbb{E}_K(N)$ denotes the average amount of required sub-blocks in order to send a specific source message to the destination, which is given by

$$\mathbb{E}_K(N) = 1 + \sum_{k=1}^{K} \Pr \left(|h_{\mathrm{sd}}|^2 + \sum_{l=1}^{k-1} \frac{|h_{\mathrm{r}_l\mathrm{d}}|^2 |h_{\mathrm{sr}_l}|^2}{|h_{\mathrm{r}_l\mathrm{d}}|^2 + |h_{\mathrm{sr}_l}|^2 + \tau/\mathsf{SNR}} < \tilde{g}_K \right). \tag{6.51}$$

This practically means that the destination accumulates SNR until it is able to decode the source signal. If the destination is not able to decode after the K-th relay has transmitted, an outage event is declared.

In [159] a bursty amplify-and-forward protocol is described where only the "best" relay is selected for transmission rather than all relays. Noteworthy, that this scheme leads to additional overhead due to relay selection and possesses a slightly worse performance with respect to outage probability.

For the cut-set bound, the instantaneous channel capacity for Gaussian inputs is upper bounded by

$$\mathcal{C}_{\mathrm{CSB}} \leq \min\{\log_2(1 + (|h_{\mathrm{sd}}|^2 + \sum_{k=1}^{K} |h_{\mathrm{sr}_k}|^2)\mathsf{SNR}), \log_2(1 + (|h_{\mathrm{sd}}|^2 + \sum_{k=1}^{K} |h_{\mathrm{r}_k\mathrm{d}}|^2)\mathsf{SNR})\}, \tag{6.52}$$

where we only considered the cut for the broadcast channel and the cut for the multiple access channel and neglected any mix-terms. With

$$\lim_{\substack{\epsilon \to 0 \\ \mathsf{SNR} \to 0 \\ g_K \to 0}} \frac{p_{\mathrm{out}}^{\mathrm{CSB}}}{g_K^{K+1}} \leq \frac{\prod_{k=1}^{K} \sigma_{\mathrm{r}_k\mathrm{d}}^2 + \prod_{k=1}^{K} \sigma_{\mathrm{sr}_k}^2}{(K+1)! \sigma_{\mathrm{sd}}^2 \prod_{k=1}^{K} \sigma_{\mathrm{r}_k\mathrm{d}}^2 \sigma_{\mathrm{sr}_k}^2} \tag{6.53}$$

the ϵ-outage capacity can be written as

$$\mathcal{C}_{\epsilon}^{(\mathrm{CSB})} \geq \frac{1}{1+K\epsilon} \log_2 \left(1 + \mathsf{SNR} \sqrt[K+1]{\frac{(K+1)! \sigma_{\mathrm{sd}}^2 \prod_{k=1}^{K} \sigma_{\mathrm{r}_k\mathrm{d}}^2 \sigma_{\mathrm{sr}_k}^2 \epsilon}{\prod_{k=1}^{K} \sigma_{\mathrm{r}_k\mathrm{d}}^2 + \prod_{k=1}^{K} \sigma_{\mathrm{sr}_k}^2}} \right). \tag{6.54}$$

The pre-log factor $1/(1+K\epsilon)$ is a straightforward extension to (6.36).

We now compare the performance of incremental relaying protocols to that of the cut-set bound. First, we consider incremental relaying with decode-and-forward. Inserting (6.43) and (6.54) into Definition 2 leads to

$$\Delta(\epsilon) \approx \frac{1+K\epsilon}{\mathbb{E}_K(N)} \cdot \frac{\prod_{k=1}^{K} \sigma_{\mathrm{r}_k\mathrm{d}}^2 + \prod_{k=1}^{K} \sigma_{\mathrm{sr}_k}^2}{(K+1)! \prod_{k=1}^{K} \sigma_{\mathrm{r}_k\mathrm{d}}^2 + \prod_{k=1}^{K} \sigma_{\mathrm{sr}_k}^2}. \tag{6.55}$$

For incremental relaying with bursty amplify-and-forward, we get

$$\Delta(\epsilon) \leq \frac{1+K\epsilon}{\mathbb{E}_K(N)}.$$

In contrast to the one-relay case, it can be seen that the ratio between the ϵ-outage capacities depends on the relay locations, which determine the average amount of required sub-blocks, i.e., $\mathbb{E}_K(N)$. Note that the expressions for $\mathbb{E}_K(N)$ for decode-and-forward and bursty amplify-and-forward differ due to the different ways of SNR accumulation at the destination. For decode-and-forward, $\mathbb{E}_K(N)$ is given in (6.44), and for bursty amplify-and-forward, $\mathbb{E}_K(N)$ is given in (6.51).

If we compare the result of decode-and-forward to that of bursty amplify-and-forward, we see that the limiting factor for decode-and-forward is the term $(K+1)!$ in the denominator of (6.43). Therefore, if the number of relays is increased, the relative loss compared to bursty amplify-and-forward will increase.

6.4 Imperfect Feedback Channel

In the previous section, the one-bit feedback from the destination is perfectly received at the relay and the source. For that reason, each node knows exactly what to do after receiving the feedback. Even more important, each node always does the right thing, which simplifies analysis enormously. This means that there will never be any kind of collision due to simultaneous channel access by the source and the relay. However, this is not true anymore, if the feedback is considered to be imperfect [162].

With the introduction of imperfect feedback, numerous transmission scenarios are thinkable that lead to a reduced performance, e.g.:

- Assume that the source transmission failed. Though the relay should transmit during the second sub-block, it remains silent. Finally, this leads to an outage event.
- After failed source transmission, the source retransmits its message and the relay does not, even if the relay-destination channel is of better quality. This results in a lower decoding probability at the destination and, thus, reduces performance.
- Generally, the feedback channel from the destination to the relay differs from the one to the source. Therefore, it is possible that both nodes receive different

information about success or failure of the source transmission in the first sub-block. Hence, collisions can occur when both terminals access the channel in the second sub-block.[13]

The questions addressed in this section are the following: What happens if the one-bit feedback from the destination is imperfectly received at the source and the relay? Especially, is the ϵ-outage capacity influenced by imperfect feedback? We briefly summarize our findings in some words. The quality of the feedback link has a strong influence on the average amount of required sub-blocks which determine the pre-log factor, i.e., the scaling factor in front of the log-function of capacity expressions [163]. It is reasonable to model the one-bit feedback channel as a binary symmetric channel. With this setting, we are able to quantify the pre-log factor and, consequently, the ϵ-outage capacity of various cooperative protocols with incremental relaying.

It is evident that imperfect feedback only influences the average amount of required sub-blocks $\mathbb{E}(N)$ and not the log-expression. This is due to the fact that the ϵ-outage capacity of incremental relaying is derived by using a "baseline model", i.e., a similar network that employs the same cooperative strategy but no feedback. Feedback then is introduced by a scaling factor which depends on the success or failure of the source transmission (for the one-relay case). For the case of multiple relays, treatment of feedback gets more involved (see Subsection 6.4.2). Hence, the ϵ-outage capacity will be reduced, if the average amount of required sub-blocks increases. To sum up, in order to investigate the influence of imperfect feedback on the ϵ-outage capacity of an incremental relaying protocol, it is sufficient to analyze the average amount of required sub-blocks $\mathbb{E}(N)$.

In the following, we make the useful assumption that the destination knows if it has been able to decode reliably. This means that there is no such thing as "destination transmits a positive acknowledgment, though it has not been able to decode." Moreover, we use the following notation. The probability that the source-destination transmission was successful is denoted by P_{SD}. Accordingly, the probability that the source-destination transmission was *not* successful is described by $\bar{P}_{\mathrm{SD}} = 1 - P_{\mathrm{SD}}$. $P_{\mathrm{R}_k\mathrm{D}}$ is the probability that the destination can decode after the transmission of the k-th relay ($k = 1, \ldots, K$). This also includes prior transmissions, e.g., consider $P_{\mathrm{R}_1\mathrm{D}}$. This is the probability that the destination can decode after combining the source and the first relay transmissions. Combining here refers again to maximal ratio combining, i.e., an accumulation of SNR. For decode-and-forward, the destination can decode whenever

$$|h_{\mathrm{sd}}|^2 + |h_{\mathrm{rd}}|^2 \geq \frac{2^{2R} - 1}{\mathsf{SNR}}, \tag{6.56}$$

[13]Recall that we consider a TDMA-like transmission scheme, where the source and the relay transmit in orthogonal sub-blocks.

where it is assumed that the relay received the source transmission reliably. For amplify-and-forward, the destination can decode if

$$|h_{\mathrm{sd}}|^2 + \frac{|h_{\mathrm{sr}}|^2 |h_{\mathrm{rd}}|^2}{|h_{\mathrm{sr}}|^2 + |h_{\mathrm{rd}}|^2 + 1/\mathsf{SNR}} \geq \frac{2^{2R} - 1}{\mathsf{SNR}}. \tag{6.57}$$

Accordingly, $\bar{P}_{\mathrm{R}_k\mathrm{D}}$ describes the probability that the destination *cannot* decode after the transmission of the k-th relay. With $(\mathbf{XY})_{l,m}$ we denote the element of the l-th row and the m-th column of the matrix product $\mathbf{XY}$. Thus, for

$$\mathbf{X} = \begin{bmatrix} x_0 & x_1 \\ x_2 & x_3 \end{bmatrix}, \quad \mathbf{Y} = \begin{bmatrix} y_0 & y_1 \\ y_2 & y_3 \end{bmatrix},$$

we have

$$(\mathbf{XY})_{2,1} = x_2 y_0 + x_3 y_2.$$

The Hadamard[14] product of two matrices $\mathbf{X}$ and $\mathbf{Y}$, i.e., the entry-wise product, is given by

$$\mathbf{X} \circ \mathbf{Y} = \begin{bmatrix} x_0 y_0 & x_1 y_1 \\ x_2 y_2 & x_3 y_3 \end{bmatrix}.$$

Since we consider a one-bit feedback, it makes sense to model the feedback channel as binary symmetric channel. We define

$$\begin{aligned} p &:= \Pr(\mathtt{ACK}|\mathtt{ACK}) = \Pr(\mathtt{NACK}|\mathtt{NACK}) \\ 1 - p &:= \Pr(\mathtt{NACK}|\mathtt{ACK}) = \Pr(\mathtt{ACK}|\mathtt{NACK}). \end{aligned}$$

Note that we are mostly interested in the reliability of feedback rather than its failure. Therefore, we use p as the probability of correct transmission. The network model for the one-relay case is shown in Fig. 6.12. For the sake of analysis, we assume that the source and the relay face the same feedback channel from the destination.

6.4.1 One- and Two-Relay Case

We first consider the one-relay case. There are two constellations when only one sub-block is required for successful transmission. Either the source transmission was successful and `ACK` was received correctly or the source transmission was successful and `NACK` was received incorrectly. In addition to that, we have two required sub-blocks for the following cases. The source transmission was successful and `ACK` was received incorrectly or the source transmission was not successful and `NACK` was received correctly. Summarizing, this can be written as

$$\begin{aligned} \mathbb{E}(N) &= P_{\mathrm{SD}} p + \bar{P}_{\mathrm{SD}}(1-p) + 2P_{\mathrm{SD}}(1-p) + 2\bar{P}_{\mathrm{SD}} p \\ &= (2\bar{P}_{\mathrm{SD}} - 1)p + 2 - \bar{P}_{\mathrm{SD}}. \end{aligned}$$

[14] Jacques Salomon Hadamard, ∗ December 8, 1865, † October 17, 1963. French mathematician. Important contributions to the fields of number theory, complex function theory, and partial differential equations.

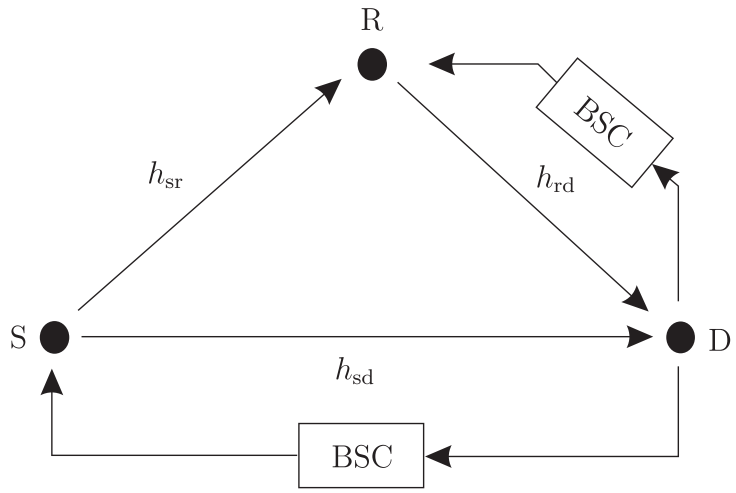

Figure 6.12: Network for incremental relaying with imperfect feedback modeled as binary symmetric channel (BSC).

This represents a linear equation of $\mathbb{E}(N)$ in p depending on the parameter $\bar{P}_{\mathrm{SD}}$. It can easily be concluded that all curves are located in a rectangular box bounded by $\mathbb{E}(N) = 1$, $\mathbb{E}(N) = 2$, $p = 0$, and $p = 1$. Fig. 6.13 shows the array of curves for different values of $\bar{P}_{\mathrm{SD}}$. The extreme values of $\bar{P}_{\mathrm{SD}}$ are $\bar{P}_{\mathrm{SD}} = 0$ and $\bar{P}_{\mathrm{SD}} = 1$, which bound the regions in which we cannot find any curves (gray area). It can be seen that $\mathbb{E}(N)$ increases or decreases in p depending on the parameter $\bar{P}_{\mathrm{SD}}$. In order to investigate this behavior, we differentiate $\mathbb{E}(N)$ with respect to p and get

$$\frac{\mathrm{d}\mathbb{E}(N)}{\mathrm{d}p} = 2\bar{P}_{\mathrm{SD}} - 1 \begin{cases} < 0 \text{ for } \bar{P}_{\mathrm{SD}} \in [0; 0.5) \\ > 0 \text{ for } \bar{P}_{\mathrm{SD}} \in (0.5; 1] \end{cases} .$$

For $p = 1$, we have "perfect" feedback and the average amount of required sub-blocks becomes $\mathbb{E}(N) = 1 + \bar{P}_{\mathrm{SD}}$ (see Section 6.2). For $p = 0$, which simply means that each observation of the feedback is wrong with probability 1, we get $\mathbb{E}(N) = 2 - \bar{P}_{\mathrm{SD}}$. Another interesting fact is that all curves intersect at $(p = 0.5; \mathbb{E}(N) = 1.5)$. Why is this the case? A probability of $p = 0.5$ means that observation of feedback is worthless. As a consequence, from a long-term perspective, the best thing for the relay to do is to transmit in block i, to remain silent in block $i + 1$, to transmit in block $i + 2$, and so on. This strategy finally leads to an average amount of required sub-blocks of $\mathbb{E}(N) = 1.5$ independent of $\bar{P}_{\mathrm{SD}}$. Expressed in other words, the relay scrambles in each block if it should transmit or not.

We summarize our findings in a few words.

- If $\bar{P}_{\mathrm{SD}} < 0.5$, then $\mathbb{E}(N)$ will decrease with increasing p. Hence, if the source-destination channel is reliable (i.e., $\bar{P}_{\mathrm{SD}} \to 0$), the average amount of required sub-blocks $\mathbb{E}(N)$ will decrease, when the feedback channel gets more and more reliable (i.e., $p \to 1$).

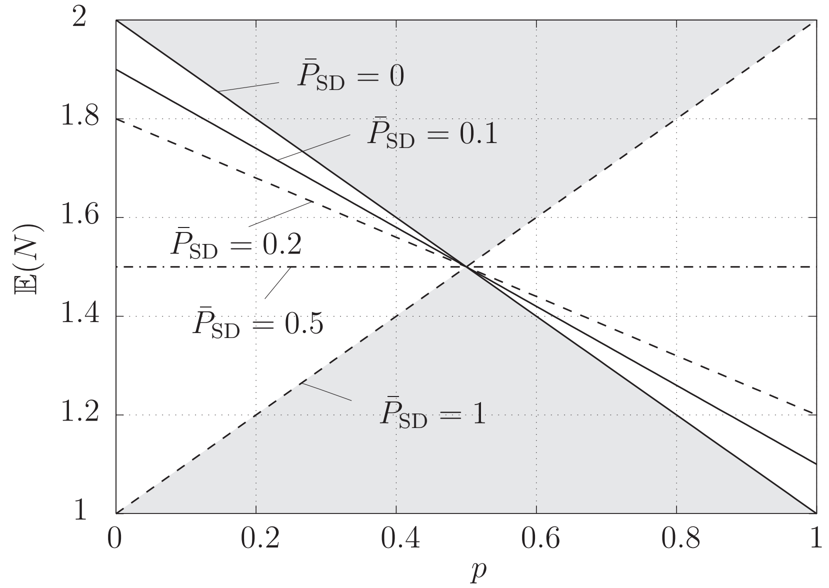

Figure 6.13: Average amount of required sub-blocks $\mathbb{E}(N)$ vs. reliability p of the feedback channel for the one-relay case.

- If the source-destination channel is not reliable (i.e., $\bar{P}_{SD} \to 1$), the average amount of required sub-blocks $\mathbb{E}(N)$ will increase, when the feedback channel gets more and more reliable (i.e., $p \to 1$). This is also clear from an intuitive viewpoint. If source-destination transmission fails pretty often and the relay receives information from the destination about success or failure of this transmission correctly, the relay has to aid communications more often, hence, $\mathbb{E}(N)$ will increase with increasing p.

- Consider the case $\bar{P}_{SD} = 0.5$ (illustrated as dash-dotted line in Fig. 6.13). The relay should transmit in every second block (from a long-term perspective, since every second source transmission fails). Therefore, $\mathbb{E}(N)$ becomes 1.5.

The average amount of required sub-blocks can also be written in matrix notation as

$$\mathbb{E}(N) = \underbrace{[1, 2]}_{\mathbf{K}_2} \underbrace{\begin{bmatrix} p & 1-p \\ 1-p & p \end{bmatrix}}_{\mathbf{P}} \underbrace{\begin{bmatrix} P_{SD} \\ \bar{P}_{SD} \end{bmatrix}}_{\mathbf{S}} = \mathbf{K}_2\mathbf{P}\mathbf{S}. \tag{6.58}$$

The vector $\mathbf{K}_2$ denotes the possible amount of required sub-blocks. The feedback link is represented by the matrix $\mathbf{P}$ and $\mathbf{S} = [P_{SD}\ \bar{P}_{SD}]^T$ is a vector describing the success or failure of the source transmission.

Next, we consider the two-relay case. Obviously, the maximum amount of required sub-blocks is 3. Equally to the case of one relay, one sub-block is required if the source transmission was successful and `ACK` was received correctly or if the source transmission was not successful and `NACK` was received incorrectly. If source transmission was successful and `ACK` was received incorrectly, we need two sub-blocks if the combination of the source and the first relay transmissions was successful and the `ACK` was received correctly or if the combined transmissions were not successful and the `NACK` was received incorrectly. Additionally, if the source transmission was not successful and `NACK` was received correctly, two sub-blocks are required if the combination of the source and the first relay transmissions was successful and the `ACK` was received correctly or if the combined transmissions were not successful and the `NACK` was received incorrectly. For the following four constellations, three sub-blocks are required:

- The source-destination transmission succeeded, `ACK` was received incorrectly, combined transmissions of the source and the first relay succeeded, and `ACK` again was received incorrectly.
- The source-destination transmission succeeded, `ACK` was received incorrectly, combined transmissions of the source and the first relay failed, and `NACK` was received correctly.
- The source-destination transmission failed, `NACK` was received correctly, combined transmissions of the source and the first relay succeeded, and `ACK` was received incorrectly.
- The source-destination transmission failed, `NACK` was received correctly, combined transmissions of the source and the first relay failed, and `NACK` again was received correctly.

All these considerations can be summarized in the equation

$$\begin{aligned}
\mathbb{E}_2(N) \;=\; & P_{\mathrm{SD}}p + \bar{P}_{\mathrm{SD}}(1-p) \\
& +2P_{\mathrm{SD}}(1-p)P_{\mathrm{R_1D}}p \\
& +2P_{\mathrm{SD}}(1-p)\bar{P}_{\mathrm{R_1D}}(1-p) \\
& +2\bar{P}_{\mathrm{SD}}pP_{\mathrm{R_1D}}p \\
& +2\bar{P}_{\mathrm{SD}}p\bar{P}_{\mathrm{R_1D}}(1-p) \\
& +3P_{\mathrm{SD}}(1-p)P_{\mathrm{R_1D}}(1-p) \\
& +3P_{\mathrm{SD}}(1-p)\bar{P}_{\mathrm{R_1D}}p \\
& +3\bar{P}_{\mathrm{SD}}pP_{\mathrm{R_1D}}(1-p) \\
& +3\bar{P}_{\mathrm{SD}}p\bar{P}_{\mathrm{R_1D}}p.
\end{aligned}$$

Similar to the one-relay case, there exists a compact matrix notation of $\mathbb{E}_2(N)$. It is shown in Fig. 6.14, where $\mathbf{K}_3 = [1\,2\,3]$ clearly is of dimension (1×3). The matrix $\mathbf{P}$ again denotes the feedback channel and is given in (6.58). The matrices $\mathbf{S}$ and $\mathbf{R}_1$

$$\mathbb{E}_2(N) = \mathbf{K}_3 \begin{bmatrix} \mathbf{PS} \\ (\mathbf{PS})_{2,1} \end{bmatrix} \circ \begin{bmatrix} 1 \\ \mathbf{PR}_1 \end{bmatrix}$$

Figure 6.14: Matrix notation of $\mathbb{E}_2(N)$ for the case of two relays.

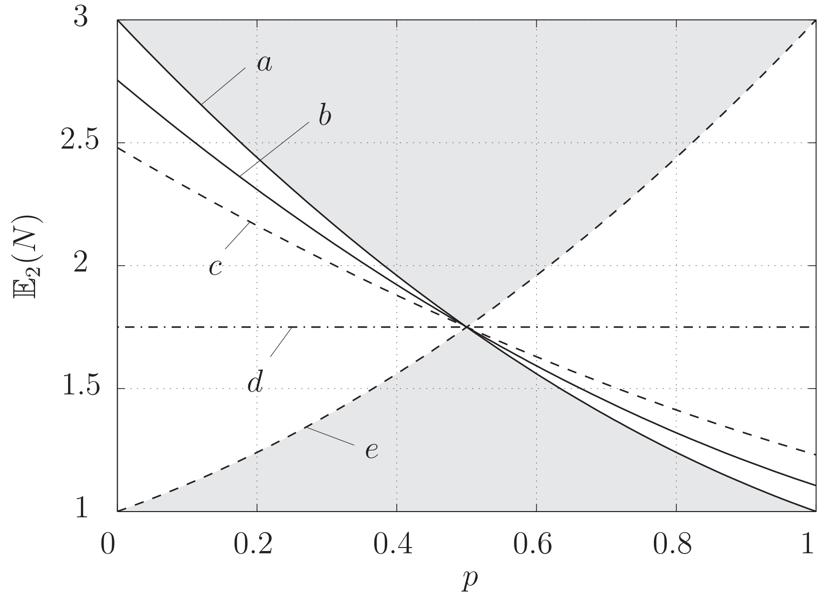

Figure 6.15: Average amount of required sub-blocks $\mathbb{E}_2(N)$ vs. reliability p of the feedback channel for the two-relay case.

describe the decoding probability at the destination after the source transmission and after combination of the source and the relay transmission, respectively. Therefore,

$$\mathbf{S} = \begin{bmatrix} P_{\mathrm{SD}} \\ \bar{P}_{\mathrm{SD}} \end{bmatrix}, \quad \mathbf{R}_1 = \begin{bmatrix} P_{\mathrm{R_1D}} \\ \bar{P}_{\mathrm{R_1D}} \end{bmatrix}.$$

The matrix products $\mathbf{PS}$ and $\mathbf{PR}_1$ have dimension (2×1) each. Finally, $(\mathbf{PS})_{2,1}$ is the element of the second row and the first column of the matrix product $\mathbf{PS}$, i.e., $(\mathbf{PS})_{2,1} = P_{\mathrm{SD}}(1-p) + \bar{P}_{\mathrm{SD}}p$. The dimension of the Hadamard product of the two matrices then is (3×1).

The average amount of required sub-blocks $\mathbb{E}_2(N)$ vs. the reliability p of the feedback channel is illustrated in Fig. 6.15. As for the one-relay case, we are able to bound the area in which all curves are located. The area is represented by a rectangular box bounded by $\mathbb{E}_2(N) = 1$, $\mathbb{E}_2(N) = 3$, $p = 0$, and $p = 1$. Of course, we now have a different parameter set consisting of $\bar{P}_{\mathrm{SD}}$ and $\bar{P}_{\mathrm{R_1D}}$. The values used in Fig. 6.15 are listed in Tab. 6.1. Similar to Fig. 6.13, the gray area illustrates the region where no curves can be found. It can easily be seen that we do not have straight lines anymore. Interestingly, all curves intercept for $p = 0.5$ (as for the one-relay case). The average amount of required sub-blocks for this value becomes $\mathbb{E}_2(N) = 1.75$.

Table 6.1: Values for $\bar{P}_{\mathrm{SD}}$ and $\bar{P}_{\mathrm{R_1D}}$ in Fig. 6.15.

tag	$\bar{P}_{\mathrm{SD}}$	$\bar{P}_{\mathrm{R_1D}}$
a	0	0
b	0.1	0.05
c	0.2	0.15
d	0.5	0.5
e	1	1

6.4.2 Generalization

It is obvious that as the number of relay nodes is increased, the amount of possible combinations for $\mathbb{E}_K(N)$ is increased as well.[15] We have already seen that even for the case of two relays, a compact matrix notation cannot be derived at a first short glance. For that reason, we have to find a construction rule that allows us to derive an easy mathematical description of large networks. In the following, we extend the results of the previous subsection to cooperative networks with an arbitrary number of relay nodes.

The calculation of $\mathbb{E}_K(N)$ can be described by a binary tree. Therefore, we are able to derive a simple construction rule for generalized networks with an arbitrary number of relay nodes. To show that, consider Fig. 6.16. Subfigure a) illustrates the one-relay case and subfigure b) shows the binary tree for the case of two relays. First, we introduce the notion of "levels." We see that for the one-relay case, there exists only one level, and for the two-relay case, the number of levels equals 2. Hence, the number of levels corresponds directly to the number of relays. Moreover, we introduce two kinds of blocks. A "positive block" that treats successful transmission (e.g., P_{SD} and $P_{\mathrm{R_1D}}$) and the possibility of successful or failed positive acknowledgment (`ACK`). And a "negative block" that deals with failed transmission (e.g., $\bar{P}_{\mathrm{SD}}$ and $\bar{P}_{\mathrm{R_1D}}$) and the possibility of successful or failed negative acknowledgment (`NACK`). We talk of a "path," when we consider the multiplication of a decoding probability with the corresponding `ACK` or `NACK`. Both kinds of blocks appear in each level. Now, the average amount of required sub-blocks $\mathbb{E}_K(N)$ can be derived by applying the following construction rule which gives the different summands of $\mathbb{E}_K(N)$.

1. *Positive block:*

- If a block ends with $p = \Pr(\mathtt{ACK}|\mathtt{ACK})$, then the corresponding path is terminated and multiplied by the level number. For the one-relay case, this is the path $P_{\mathrm{SD}}p$.

[15] We denote the average amount of required sub-blocks as $\mathbb{E}_K(N)$ for the general case.

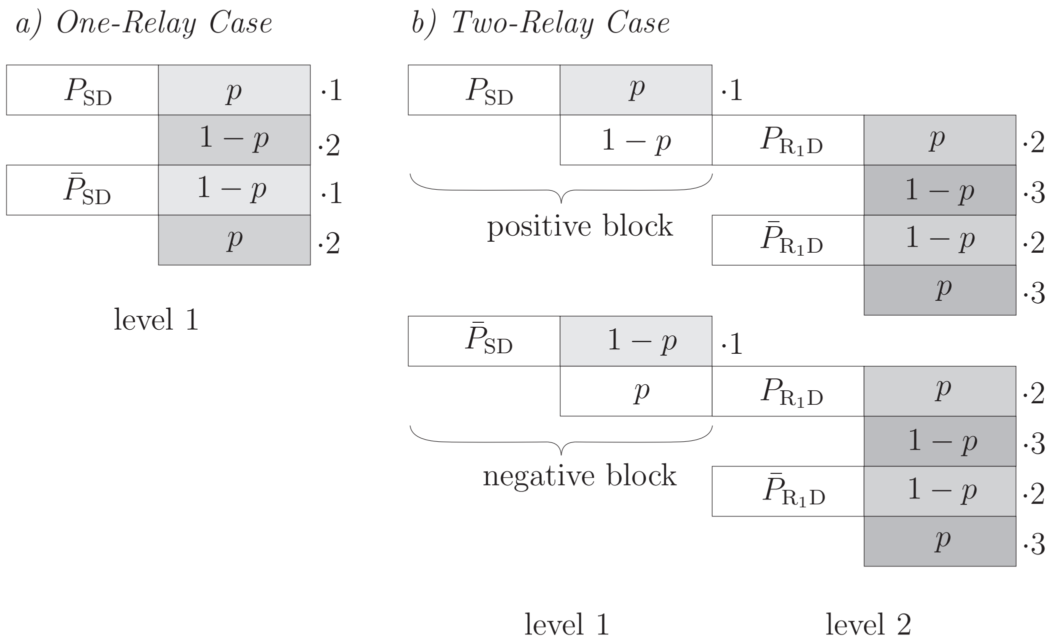

Figure 6.16: Binary tree-based construction in order to calculate the average amount of required sub-blocks $\mathbb{E}_K(N)$. a) One-relay case and b) two-relay case. Number of levels corresponds to the number of relays in the network.

For the two-relay case, these are the paths $P_{\mathrm{SD}}p$ (which is multiplied by 1) as well as $P_{\mathrm{SD}}(1-p)P_{\mathrm{R_1D}}p$ and $\bar{P}_{\mathrm{SD}}pP_{\mathrm{R_1D}}p$ (which are multiplied by 2).

- If a block ends with $1-p = \Pr(\texttt{NACK}|\texttt{ACK})$, a new level is added, i.e., a new positive and a new negative block are added. The construction continues until the highest level is reached. (The highest level corresponds to the number of relays in the network.) Then, the last path is multiplied by a factor that is equal to the highest level number plus 1. For the one-relay case, this is the path $P_{\mathrm{SD}}(1-p)$ which is multiplied by 2. For the two-relay case, these are the paths $P_{\mathrm{SD}}(1-p)P_{\mathrm{R_1D}}(1-p)$ and $\bar{P}_{\mathrm{SD}}pP_{\mathrm{R_1D}}(1-p)$ which are multiplied by 3.

2. Negative block:

- If a block ends with $1-p = \Pr(\texttt{ACK}|\texttt{NACK})$, then the corresponding path is terminated and multiplied by the level number. For the one-relay case, this is the path $\bar{P}_{\mathrm{SD}}(1-p)$ (which is multiplied by 1). For the two-relay case, these are the paths $\bar{P}_{\mathrm{SD}}(1-p)$ (which is multiplied by 1) as well as $P_{\mathrm{SD}}(1-p)\bar{P}_{\mathrm{R_1D}}(1-p)$ and $\bar{P}_{\mathrm{SD}}p\bar{P}_{\mathrm{R_1D}}(1-p)$ (which are multiplied by 2).

- If a block ends with $p = \Pr(\texttt{NACK}|\texttt{NACK})$, a new level is added, i.e., a new positive and a new negative block are added. The construction continues until the highest level is reached. Then, the last path is multiplied by a factor that is equal to the highest level number plus 1. For the one-relay case, this is the path $\bar{P}_{\mathrm{SD}}p$

$$\mathbb{E}_K(N) = \mathbf{K}_{K+1} \begin{bmatrix} \mathbf{PS} \\ (\mathbf{PS})_{2,1} \\ \vdots \\ (\mathbf{PS})_{2,1} \end{bmatrix} \circ \begin{bmatrix} 1 \\ \mathbf{PR}_1 \\ (\mathbf{PR}_1)_{2,1} \\ \vdots \\ (\mathbf{PR}_1)_{2,1} \end{bmatrix} \circ \begin{bmatrix} 1 \\ 1 \\ \mathbf{PR}_2 \\ (\mathbf{PR}_2)_{2,1} \\ \vdots \\ (\mathbf{PR}_2)_{2,1} \end{bmatrix} \circ \cdots \circ \begin{bmatrix} 1 \\ \vdots \\ \mathbf{PR}_{K-1} \end{bmatrix}$$

Figure 6.17: Matrix notation of $\mathbb{E}_K(N)$ for the case of K relays.

which is multiplied by a factor 2. For the two-relay case, these are the paths $P_{\mathrm{SD}}(1-p)\bar{P}_{\mathrm{R_1D}}p$ and $\bar{P}_{\mathrm{SD}}p\bar{P}_{\mathrm{R_1D}}p$ which are multiplied by 3.

With this rather simple construction rule, we can describe the average amount of required sub-blocks for networks with an arbitrary number of relay nodes. An interesting fact occurs for $p = 0.5$. For the one-relay case, we have $\mathbb{E}(N) = 1.5$. For two relays, the average amount of required sub-blocks becomes $\mathbb{E}_2(N) = 1.75$. It can easily be verified that due to the binary tree-based construction rule explained before, the limit for $K \to \infty$ tends to $\mathbb{E}_\infty(N) = 2$. This can directly be seen if we consider the geometric series $\sum_{k=0}^{\infty} \frac{1}{2^k} = 2$. To conclude, if the feedback link is unreliable, i.e., $p = 0.5$, and the network is large, the transmission strategy of the relays should be as follows. Each source message is retransmitted by one and only one relay, which clearly leads to $\mathbb{E}_\infty(N) = 2$. This is in line with results presented in [91, 159], where an opportunistic relay protocol is proposed that selects only one relay for cooperation.

We are now able to derive a compact matrix notation for $\mathbb{E}_K(N)$ by extending the results illustrated in Fig. 6.16. The key is to exploit the binary tree-based construction rule and to keep in mind that – apart from the final level – there are 2^k paths per level that are terminated and multiplied by the level number, which means that they are not considered for further calculations. Note that k here is the number of relays. The final level consists of 2^{k+1} paths which are all terminated. The result is shown in Fig. 6.17.

The Hadamard product of K vectors must be calculated. Note that we do not have to consider $K+1$ vectors, in spite of having $K+1$ transmitting nodes. The reason is that after the last relay has transmitted, an outage is declared if the accumulated SNR is still not high enough and there will not be any further transmissions for this specific source message, i.e., success or failure of the last relay's transmission does not influence the average amount of required sub-blocks anymore. The transition matrices, i.e., $\mathbf{PS}$ and $\mathbf{PR}_k$, $k = 1, \ldots, K$, respectively, are shifted downwards from left to right. Hence, we have a kind of a lower triangular structure, which is due to the fact that per level some paths are terminated. The resulting Hadamard product is multiplied by a $1 \times (K+1)$ vector $\mathbf{K}_{K+1}$ that accounts for the different factors each terminated level is multiplied with.

The amount of summands z of $\mathbb{E}_K(N)$ gives insight into the computational effort of the construction rule. It can easily be verified that for one relay we have $z = 4$, for two relays we have $z = 10$, for three relays we have $z = 22$ and so on. As a consequence, if we add one relay to the network, we have to add 1 to the preceding number of summands and multiply the result by 2. This is due to the fact that by adding a new relay the number of levels is increased by one and, additionally, the number of blocks per level is doubled. Hence,

$$z_{k+1} = (z_k + 1) \cdot 2. \tag{6.59}$$

The drawback of this notation is that it does not allow us to calculate the number of summands for a network with k relays without knowing the result for the network with $k-1$ relays. By induction it can readily be shown that

$$z_k = 2^{k+1} + 2^k - 2 = 3 \cdot 2^k - 2. \tag{6.60}$$

In this subsection we discussed the idea of imperfect feedback for incremental relaying networks. We could extend the results for the one- and the two-relay case to networks with an arbitrary number of relays. A compact matrix notation for the average amount of required sub-blocks in order to transmit a specific source message has been derived and the computational effort has been treated by giving an expression on the amount of summands. It is evident that the larger the network becomes, the amount of multiplications per summand increases.

Combining the results of this subsection with those in Section 6.3, we are able to express the ϵ-outage capacities of incremental relaying networks that perform either decode-and-forward or bursty amplify-and-forward with imperfect feedback from the destination.

7

Conclusions

> We may have all come on different ships, but we're in the same boat now.
>
> *Martin Luther King Jr.*

The field of cooperative communications in wireless networks is a vivid research area. We gave a detailed overview over the impairments of wireless channels and discussed the issue of half-duplex vs. full-duplex nodes. We also reviewed the basic network models for cooperative communications including multi-hop, variations of multi-route, and incremental relaying networks. A thorough analysis with respect to cooperation strategies was performed as well. We stress that the main focus of this dissertation was the investigation of incremental relaying networks. In particular, we derived the ϵ-outage capacities of incremental relaying networks with decode-and-forward and bursty amplify-and-forward.

In this final chapter, we summarize our contributions and highlight some fields for further research.

7.1 Contributions

Chapter 3 developed an adaptive relay selection protocol, where the network performance is improved by intelligently selecting a set of relays. In contrast to most known selection protocols, the amount of selected relays is not fixed, but rather adapted to the performance requirements. In the case of the adaptive relay selection protocol, the performance metric is the bit error rate at the destination. The selection process is centralized and performed by the source based on five intrinsic parameters. These are the channel quality from the source to a relay and from a relay to the destination, the remaining battery power, the direction of movement of the relay, and, finally, the relay's willingness to participate to the cooperative process. Each relay transmits its parameters to the source. For that purpose, we proposed a new medium access control scheme by introducing an additional frame. The source then evaluates these parameters and creates a list from the most suitable to the least suitable relay. Applying this list, the source is able to determine the amount of required relays and transmits a so-called relay table containing the ranked relay identifiers. With this relay table, each relay knows if it should aid communications and when it is allowed to access the channel.

In Chapter 4 we dealt with the issue of optimal time and power allocation. We used the instantaneous channel capacity and the delay-limited capacity as optimization criteria and solved the optimization problem by applying an algorithm based on Brent's method. The basic idea was to balance the aspects of speed of convergence and reliability. Whenever possible, parabolic interpolation is used, since it converges very fast. If we are not sure about the reliability of the result, we switch to the robust method of the golden section search. We demonstrated that there are cases in which cooperation with optimal resource allocation does not outperform direct transmission. For that purpose, selective protocols where considered, where cooperation is applied once it outperforms direct transmission. Generally, we can state that the capacity gains achieved by cooperation over direct transmission increase with decreasing overall system power.

A new hybrid combining dual-diversity receiver was presented in Chapter 5. The receiver selects dynamically between the combining strategies SC and EGC on the basis of an SNR threshold criterion. For reasons of complexity, MRC was not considered for the receiver, since it requires estimation of channel gains. We showed that it is reasonable to switch between SC and EGC for different degrees of branch imbalance. For a low branch imbalance, EGC is beneficial to SC. However, as the degree of branch imbalance increases, SC becomes more preferable. The crucial point is to find the exact value of the threshold. Derivation gets involved, since the threshold value does not only depend on the degree of branch imbalance, but also on the true SNR values of each branch. There exists no closed-form solution to this problem. For that circumstance, we created a look-up table by simulations, that takes the aforementioned issues into account.

Chapter 6 contains our main theoretical results. We looked at the metric of ϵ-outage capacity for incremental relaying networks where either decode-and-forward or bursty amplify-and-forward is employed. We first focused on perfect feedback from the destination and derived expressions for the one-relay case. The results where then compared to the cut-set bound. We found that for decode-and-forward, incremental relaying shows the weakest performance if the relay is located close to the destination. If the relay is located close to the source, however, incremental relaying becomes optimal in a sense that it achieves the cut-set bound. In contrast to that, the ratio of the ϵ-outage capacities of bursty amplify-and-forward and the cut-set bound is independent of the relay location and only depends on the target error rate and the quality of the source-to-destination transmission. The results are extended to larger networks with an arbitrary number of relays. We demonstrated that the performance of decode-and-forward is – in contrast to bursty amplify-and-forward – limited by a factor that depends on the amount of relays. We conclude Chapter 6 by considering imperfect feedback channels. For that purpose, the feedback channel is modeled as binary symmetric channel. We showed that imperfect feedback influences the pre-log factor of capacity expressions. After considering the one- and the two-relay case, we generalized our findings to networks with an arbitrary number of relays.

7.2 Further Research

There are numerous fields of further research that may extend the work of this dissertation. We focus on four major aspects in the following.

We have briefly given a review of synchronization approaches in this thesis. In addition, we have assumed a reasonable amount of synchronization among network nodes for our analysis. A next step would be a test-bed implementation, where such an assumption does not hold anymore. Obviously, synchronization does not come at no cost and will definitely decrease performance. Synchronization methods are especially important for distributed cooperative networks.

The wireless vision of ubiquitous access and seamless connectivity leads directly to problems related to human health. There are studies available that state that the influence of electromagnetic emissions due to the usage of mobile phones might lead to an increased risk of brain tumors on a long-term perspective [164]. One major step to solve this problem is the decrease of emitted powers of mobile phones. This is the basic idea behind green radio [165]. One possibility to achieve the reduction of the transmission powers of mobile stations is the installation of multi-hop cellular networks as described in [166]. The advantages of such network structures are – apart from the already mentioned health issues – lower interference to other users, an increased uplink capacity, an intelligent interference management via wired backbone, the optimization of the traffic load sharing, and an extension of battery lifetime.

We have mainly dealt with three-node networks throughout the dissertation and extended some results to large networks with an arbitrary number of relays. A final step is the consideration of what happens if the network size goes to infinity, i.e., in our setting of parallel networks, if the number of relays tends to infinity. Due to our interference-free TDMA-based transmission model, the time occupied by a relay to transmit information tends to zero as the number of nodes increases. The question to address is in what way the employed relay strategy influences the performance of the network. Though analysis gets involved, we suspect that such an investigation would give deep insight into the efficiency of large networks and determine – at least to a certain extent – a suitable number of relays for practical implementations.

So far, a lot of research has been carried out in the area of one-relay networking or parallel relay networks. Though there has been some research on multi-stage relay networks where, e.g., several relay networks are cascaded, there remain still a lot of open problems. Is it beneficial to group some relays into small clusters? The cluster head could then increase its decoding probability by exploiting data from the other relays in the cluster. Eventually, only the cluster head could be used for further transmission. Clearly, such an approach saves network resources. However, what about the cluster size? Who is destined to be the cluster head? How many clusters are useful in a specific setting, and so forth. This research area is especially important in a cellular environment where relays have been set up to aid communications.

A

Probability Preliminaries

In this appendix, we state two important lemmas that are useful for the derivation of the ϵ-outage capacities of relaying networks. Both lemmas approximate the cumulative distribution function of different combinations of exponentially distributed random variables. Given a random variable U, its cumulative distribution function is denoted by $F_U(\cdot)$ and its probability density function is denoted by $f_U(\cdot)$, where $F_U(u) = \Pr(U \leq u) = \int_{-\infty}^{u} f_U(x)\mathrm{d}x$. We now briefly recall the probability density function of an exponential random variable U [167, 168]. It is

$$f_U(u) = \begin{cases} 0 & : \quad u < 0 \\ \lambda_U e^{-\lambda_U u} & : \quad u \geq 0 \end{cases}, \tag{A.1}$$

where $\lambda_U > 0$ and $\mathbb{E}(U) = 1/\lambda_U := \sigma_u^2$. The definition $\mathbb{E}(U) := \sigma_u^2$ is due to the fact that we consider Rayleigh fading in this thesis. As a matter of fact, channel gain h is modeled as independent, zero-mean, circularly-symmetric complex random variable with variance σ_u^2. Consequently, the square magnitude $|h|^2 := U$ is exponentially distributed with mean σ_u^2. This may not be mistaken with the variance of an exponentially distributed variable which is given by $1/\lambda_U^2$.

Lemma 1 *Let $W = \sum_{k=0}^{K} U_k$, where U_k are independent exponentially distributed random variables with mean σ_k^2. If $g(\xi)$ is a continuous function at $\xi = 0$ and $g(\xi) \to 0$ as $\xi \to 0$, then the cumulative distribution function $F_W(\cdot)$ of W satisfies*

$$\lim_{\xi \to 0} \frac{1}{g(\xi)^{K+1}} F_W(g(\xi)) = \frac{1}{(K+1)! \prod_{k=0}^{K} \sigma_k^2}. \tag{A.2}$$

Proof 1 *We provide a sketch of the proof, which can be found in [57, app. B.2]. The idea is to upper bound the* lim sup *(limit superior) and to lower bound the* lim inf *(limit inferior) without assuming that the limit exists. If the bounds are equal, it can be concluded that the limit exists and its value is given by the corresponding bounds.*

Lemma 2 *Let U, V, and W be independent exponentially distributed random variables with mean σ_u^2, σ_v^2, and σ_w^2. If $g(\xi)$ is a continuous function at $\xi = 0$ and $g(\xi) \to 0$ as $\xi \to 0$, then*

$$\lim_{\xi \to 0} \frac{1}{g(\xi)^2} \Pr\left(U + \frac{VW}{V + W + \xi} < g(\xi)\right) = \frac{\sigma_v^2 + \sigma_w^2}{2\sigma_u^2 \sigma_v^2 \sigma_w^2}. \tag{A.3}$$

Proof 2 *We again provide a sketch of the proof. The complete proof is given in [152]. There, the authors first show that*

$$\lim_{\xi \to 0} \frac{1}{g(\xi)} \Pr\left(\frac{VW}{V + W + \xi} < g(\xi)\right) = \frac{\sigma_v^2 + \sigma_w^2}{\sigma_v^2 \sigma_w^2}. \tag{A.4}$$

With this result,

$$\begin{aligned} \Pr\left(U + \frac{VW}{V + W + \xi} < g(\xi)\right) &= \Pr(U + r_\xi < g(\xi)) \\ &= \int_0^{g(\xi)} \Pr(r_\xi < g(\xi) - U) f_U(u)\,\mathrm{d}u \end{aligned}$$

can be solved, where we used the substitution

$$r_\xi := \frac{VW}{V + W + \xi}, \tag{A.5}$$

and Lemma 2 is proved.

B

Multipath Error Probability

In this appendix, we prove two results which were used in Chapter 3. We first give an expression on the probability density function of K uncorrelated exponentially distributed random variables and then derive the error probability in a multipath Rayleigh fading environment. In Chapter 3, the random variables are denoted as SNR_k. For the sake of description, we use $U_k := \mathsf{SNR}_k$ and $\mathbb{E}(U_k) = \overline{u}_k$ in the following.

Lemma 3 *Let U_k be an exponentially distributed random variable. The probability density function of the sum of K independent random variables, $U = \sum_{k=1}^{K} U_k$, is given by*

$$f_U^{(K)}(u) = \sum_{k=1}^{K} \frac{\overline{u}_k^{K-2}}{\prod\limits_{l=1,l\neq k}^{K} (\overline{u}_k - \overline{u}_l)} \exp\left(-\frac{u}{\overline{u}_k}\right), \quad u \geq 0,$$

where $\overline{u}_k$ denotes the mean value of the random variable U_k.

Proof 3 *The probability density function of the sum of K independent random variables is given by the convolution of the corresponding probability density functions. Accordingly,*

$$f_U^{(K)}(u) = f_1(u_1) * f_2(u_2) * \cdots * f_K(u_K), \tag{B.1}$$

where the probability density function for each u_k is given by

$$f_k(u_k) = \frac{1}{\overline{u}_k} \exp\left(-\frac{u_k}{\overline{u}_k}\right), \quad u_k \geq 0, \quad k = 1, \cdots, K. \tag{B.2}$$

The following derivation is based on mathematical induction. Lemma 3 is certainly true for $K = 1$. The induction hypothesis for $K = n$ is

$$f_U^{(n)}(u) = \sum_{k=1}^{n} \frac{\overline{u}_k^{n-2}}{\prod\limits_{l=1,l\neq k}^{n} (\overline{u}_k - \overline{u}_l)} \exp\left(-\frac{u}{\overline{u}_k}\right). \tag{B.3}$$

Now, we have to show that this is also true for $K = n+1$. We start by writing the probability density function of $n+1$ random variables as

$$\begin{aligned} f_U^{(n+1)}(u) &= f_1(u_1) * f_2(u_2) * \cdots * f_{n+1}(u) \\ &= f_U^{(n)}(u) * f_{n+1}(u), \end{aligned}$$

where we set $u_{n+1} = u$. Using the induction hypothesis, we get:

$$\begin{aligned} f_U^{(n+1)}(u) &= \sum_{k=1}^{n} \frac{\overline{u}_k^{n-2}}{\prod\limits_{l=1,l\neq k}^{n} (\overline{u}_k - \overline{u}_l)} \exp\left(-\frac{u}{\overline{u}_k}\right) * \frac{1}{\overline{u}_{n+1}} \exp\left(-\frac{u}{\overline{x}_{n+1}}\right) \\ &= \int_0^u \sum_{k=1}^{n} \frac{\overline{u}_k^{n-2}}{\prod\limits_{l=1,l\neq k}^{n} (\overline{u}_k - \overline{u}_l)} \exp\left(-\frac{\tau}{\overline{u}_k}\right) \frac{1}{\overline{u}_{n+1}} \exp\left(-\frac{u-\tau}{\overline{u}_{n+1}}\right) \mathrm{d}\tau \\ &= \sum_{k=1}^{n} \frac{\overline{u}_k^{n-2}}{\prod\limits_{l=1,l\neq k}^{n} (\overline{u}_k - \overline{u}_l)} \frac{1}{\overline{u}_{n+1}} \exp\left(-\frac{u}{\overline{u}_{n+1}}\right) \int_0^u \exp\left(-\tau\left(\frac{1}{\overline{u}_k} - \frac{1}{\overline{u}_{n+1}}\right)\right) \mathrm{d}\tau \\ &= \sum_{k=1}^{n} \frac{\overline{u}_k^{n-2}}{\prod\limits_{l=1,l\neq k}^{n} (\overline{u}_k - \overline{u}_l)} \exp\left(-\frac{u}{\overline{u}_{n+1}}\right) \frac{-\overline{u}_k}{\overline{u}_{n+1} - \overline{u}_k} \left[\exp\left(-u\left(\frac{1}{\overline{u}_k} - \frac{1}{\overline{u}_{n+1}}\right)\right) - 1\right] \\ &= \sum_{k=1}^{n} \frac{\overline{u}_k^{n-1}}{\prod\limits_{l=1,l\neq k}^{n+1} (\overline{u}_k - \overline{u}_l)} \left(\exp\left(-\frac{u}{\overline{u}_k}\right) - \exp\left(-\frac{u}{\overline{u}_{n+1}}\right)\right) \\ &= \sum_{k=1}^{n} \frac{\overline{u}_k^{n-1}}{\prod\limits_{l=1,l\neq k}^{n+1} (\overline{u}_k - \overline{u}_l)} \exp\left(-\frac{u}{\overline{u}_k}\right) - \sum_{k=1}^{n} \frac{\overline{u}_k^{n-1}}{\prod\limits_{l=1,l\neq k}^{n+1} (\overline{u}_k - \overline{u}_l)} \exp\left(-\frac{u}{\overline{u}_{n+1}}\right) \\ &= \sum_{k=1}^{n} \frac{\overline{u}_k^{n-1}}{\prod\limits_{l=1,l\neq k}^{n+1} (\overline{u}_k - \overline{u}_l)} \exp\left(-\frac{u}{\overline{u}_k}\right) + \exp\left(\frac{-u}{\overline{u}_{n+1}}\right) \frac{\overline{u}_{n+1}^{n-1}}{\prod\limits_{l=1,l\neq n+1}^{n+1} (\overline{u}_{n+1} - \overline{u}_l)} \end{aligned}$$

And finally,

$$f_U^{(n+1)}(u) = \sum_{k=1}^{n+1} \frac{\overline{u}_k^{n-1}}{\prod\limits_{l=1,l\neq k}^{n+1} (\overline{u}_k - \overline{u}_l)} \exp\left(-\frac{u}{\overline{u}_k}\right).$$

During the last steps, we used the identity

$$-\sum_{k=1}^{n} \frac{\overline{u}_k^{n-1}}{\prod\limits_{l=1,l\neq k}^{n+1} (\overline{u}_k - \overline{u}_l)} = \frac{\overline{u}_{n+1}^{n-1}}{\prod\limits_{l=1,l\neq n+1}^{n+1} (\overline{u}_{n+1} - \overline{u}_l)}. \tag{B.4}$$

In order to show this, we make the change of variables $x := \overline{u}_{n+1}$. *Hence, we get*

$$\frac{\overline{u}_{n+1}^{n-1}}{\prod_{l=1,l\neq n+1}^{n+1}(\overline{u}_{n+1}-\overline{x}_l)} = \frac{x^{n-1}}{\prod_{l=1,l\neq n+1}^{n+1}(x-\overline{u}_l)} = \frac{A(x)}{B(x)}. \tag{B.5}$$

We now apply the concept of partial fractions and get

$$\frac{A(x)}{B(x)} = \frac{x^{n-1}}{(x-\overline{u}_1)(x-\overline{u}_2)\cdots(x-\overline{u}_n)} = \frac{A_1}{(x-\overline{u}_1)} + \frac{A_2}{(x-\overline{u}_2)} + \cdots + \frac{A_n}{(x-\overline{u}_n)}. \tag{B.6}$$

The coefficient A_1 *is obtained by multiplying* $A(x)/B(x)$ *with* $(x-\overline{u}_1)$ *and setting* $x = \overline{u}_1$. *Thus,*

$$A_1 = \left[\frac{A(x)}{B(x)}(x-\overline{u}_1)\right]_{x=\overline{u}_1} = \frac{\overline{u}_1^{n-1}}{(\overline{u}_1-\overline{u}_2)\cdots(\overline{u}_1-\overline{u}_n)}. \tag{B.7}$$

The same procedure is repeated for the coefficients $A_2, \ldots, A_n$. *We finally have*

$$\frac{A(x)}{B(x)} = -\sum_{k=1}^{n} \frac{\overline{u}_k^{n-1}}{\prod_{l=1,l\neq k}^{n+1}(\overline{u}_k-\overline{u}_l)} \tag{B.8}$$

and the proof is completed.

Next, we will use a result given in [114] to derive the error probability in a multipath Rayleigh fading environment where the channel gains are independent but not identically distributed.

Lemma 4 *The error probability in a multipath Rayleigh fading environment with independent but not identically distributed channel gains is given by*

$$\mathsf{BER} \leq \frac{a}{2}\sum_{k=1}^{K} \frac{\overline{u}_k^{K-1}}{\prod_{l=1,l\neq k}^{K}(\overline{u}_k-\overline{u}_l)} \left(1-\sqrt{\frac{\overline{u}_k}{2/b+\overline{u}_k}}\right), \tag{B.9}$$

where a *and* b *depend on the modulation scheme.*

Proof 4 *According to [114], the error probability can be expressed as*

$$\mathsf{BER} \leq \underbrace{\left[h(u)aQ(\sqrt{bu})\right]_{u=u_a}^{u_b}}_{\mathcal{I}_1} + \frac{a\sqrt{b}}{2\sqrt{2\pi}} \underbrace{\int_{u_a}^{u_b} \frac{1}{\sqrt{u}} h(u) \exp\left(-\frac{b}{2}u\right) \mathrm{d}u}_{\mathcal{I}_2}, \tag{B.10}$$

where

$$h(u) = \int f_U^{(K)}(u)\,\mathrm{d}u = -\sum_{k=1}^{K} \frac{\overline{u}_k^{K-1}}{\prod_{l=1,l\neq k}^{K}(\overline{u}_k-\overline{u}_l)} \exp\left(-\frac{u}{\overline{u}_k}\right). \tag{B.11}$$

For the sake of readability, we introduce the abbreviation

$$\Psi(K,\overline{u}_k,\overline{u}_l) = \sum_{k=1}^{K} \frac{\overline{u}_k^{K-1}}{\prod\limits_{l=1,l\neq k}^{K} (\overline{u}_k - \overline{u}_l)}. \tag{B.12}$$

Now, consider the first summand $\mathcal{I}_1$ *in* (B.10)*. By setting* $u_a = 0$ *and* $u_b = \infty$*, which are reasonable values for communication systems, and by using the facts that* $Q(\infty) = 0$ *and* $Q(0) = 1/2$*, we get*

$$\mathcal{I}_1 = \frac{a}{2}\Psi(K,\overline{u}_k,\overline{u}_l). \tag{B.13}$$

The integral $\mathcal{I}_2$ *can be solved by using a suitable substitution. We first write it down applying* (B.11) *and* (B.12)*. Accordingly,*

$$\mathcal{I}_2 = -\Psi(K,\overline{u}_k,\overline{u}_l) \int\limits_{u_a}^{u_b} \frac{1}{\sqrt{u}} \exp\left(-u\left(\frac{1}{\overline{u}_k} + \frac{b}{2}\right)\right) \,\mathrm{d}u. \tag{B.14}$$

In order to solve this integral, we first define

$$\frac{\psi^2}{2} := u\left(\frac{1}{\overline{u}_k} + \frac{b}{2}\right), \tag{B.15}$$

and get after some algebraic manipulation

$$\mathcal{I}_2 = -\Psi(K,\overline{u}_k,\overline{u}_l)\sqrt{2\pi}\sqrt{\frac{\overline{u}_k}{2+\overline{u}_k b}}, \tag{B.16}$$

where again we set $u_a = 0$ *and* $u_b = \infty$*. Combining* (B.10), (B.13), *and* (B.16) *finally yields*

$$\mathsf{BER} \leq \frac{a}{2}\Psi(K,\overline{u}_k,\overline{u}_l)\left(1 - \sqrt{\frac{\overline{u}_k}{2/b+\overline{u}_k}}\right) \tag{B.17}$$

and the proof is completed.

C

Ferrari's Method

The first one to present a closed-form solution to a biquadratic (also quartic[1]) equation was the Italian mathematician Lodovico Ferrari[2]. His solution was published by his teacher Gerolamo Cardano[3] in 1545 in his work *Ars magna de Regulis Algebraicis*. Another solution was proposed by Leonhard Euler[4] in 1738 with the intention to find a general formula to solve equations of higher even degrees. Niels Henrik Abel[5] proved in 1824, that this is impossible.

The following derivation is mostly due to [169, ch. 3.8.3] and [170]. The equation

$$z^4 + a_3 z^3 + a_2 z^2 + a_1 z + a_0 = 0 \tag{C.1}$$

can be solved by first eliminating the cubic term. This can be done by the substitution

$$z := x - \frac{a_3}{4}. \tag{C.2}$$

This yields

$$x^4 + px^2 + qx + r = 0, \tag{C.3}$$

where

$$p = a_2 - \frac{3}{8}a_3^2 \tag{C.4}$$

$$q = a_1 - \frac{1}{2}a_2 a_3 + \frac{1}{8}a_3^3 \tag{C.5}$$

$$r = a_0 - \frac{1}{4}a_1 a_3 + \frac{1}{16}a_2 a_3^2 - \frac{3}{256}a_3^4. \tag{C.6}$$

[1]Sometimes the term 'biquadratic' is only used for quartic equations that have no odd powers, i.e., $z^4 + a_2 z^2 + a_0 = 0$.

[2]Lodovico Ferrari, ∗ February 2, 1522, † October 5, 1565. Italian mathematician. Was a servant of Gerolamo Cardano.

[3]Gerolamo Cardano, ∗ September 24, 1501, † September 21, 1576. Italian mathematician, physician, and astrologer.

[4]Leonhard Euler, ∗ April 15, 1707, † September 18, 1783. Swiss mathematician and physicist. Contributed to fields of infinitesimal calculus and graph theory.

[5]Niels Henrik Abel, ∗ August 5, 1802, † April 6, 1829. Norwegian mathematician.

Now, this quartic equation can be solved if we are able to put it in a general form that allows us to factorize it. This means that the quartic can be written as the difference of two squared terms, i.e., $P^2 - Q^2 = (P+Q)(P-Q)$. Therefore, as a first step, we add and subtract $x^2u + u^2/4$ and get

$$(x^4 + x^2u + \frac{1}{4}u^2) - x^2u - \frac{1}{4}u^2 + px^2 + qx + r = 0, \tag{C.7}$$

which can be rewritten as

$$(x^2 + \frac{1}{2}u)^2 - [(u-p)x^2 - qx + (\frac{1}{4}u^2 - r)] = 0. \tag{C.8}$$

We see that the first term is already a perfect square P^2 with

$$P = x^2 + \frac{1}{2}u. \tag{C.9}$$

The second term will become a perfect square Q^2 if u is chosen in an appropriate way, i.e.,

$$Q^2 = (u-p)\left(x^2 - \frac{q}{u-p}x + \frac{\frac{1}{4}u^2 - r}{u-p}\right) = (u-p)\left(x - \sqrt{\frac{\frac{1}{4}u^2 - r}{u-p}}\right)^2. \tag{C.10}$$

Hence, it is required that

$$2\sqrt{\frac{\frac{1}{4}u^2 - r}{u-p}} = \frac{q}{u-p}, \tag{C.11}$$

which gives after some manipulation

$$q^2 = 4(u-p)(\frac{1}{4}u^2 - r). \tag{C.12}$$

This eventually leads to the expression

$$Q = \sqrt{u_1 - p}x - \frac{q}{2\sqrt{u_1 - p}}, \tag{C.13}$$

where u_1 denotes one of the three solutions to (C.10). We can conclude that Q is linear in x, whereas P is quadratic in x. This means that $P+Q$ as well as $P-Q$ is quadratic in x and can easily be solved using the quadratic formula and we finally get all four solutions to the original biquadratic equation.

Now, plugging the expressions for p, q, and r (see (C.4), (C.5), and (C.6)) into (C.12) gives the resolvent cubic equation

$$y^3 - q_2y^2 + (a_1a_3 - 4a_0)y + (4a_2a_0 - a_1^2 - a_3^2a_0) = 0, \tag{C.14}$$

where we applied the substitution $u := y - a_3^2/8$. The four solutions z_1, z_2, z_3, and z_4 of the original quartic equation are then given by the roots of the quadratic equation

$$x^2 + \frac{1}{2}(a_3 \pm \sqrt{a_3^2 - 4a_2 + 4y_1})x + \frac{1}{2}(y_1 \pm \sqrt{y_1^2 - 4a_0}) = 0 \tag{C.15}$$

with y_1 being a real root of (C.14). Finally,

$$z_1 = -\frac{1}{4}a_3 + \frac{1}{2}R + \frac{1}{2}D \tag{C.16}$$

$$z_2 = -\frac{1}{4}a_3 + \frac{1}{2}R - \frac{1}{2}D \tag{C.17}$$

$$z_3 = -\frac{1}{4}a_3 - \frac{1}{2}R + \frac{1}{2}E \tag{C.18}$$

$$z_4 = -\frac{1}{4}a_3 - \frac{1}{2}R - \frac{1}{2}E, \tag{C.19}$$

where

$$R = \sqrt{\frac{1}{4}a_3^2 - a_2 + y_1} \tag{C.20}$$

$$D = \begin{cases} \sqrt{\frac{3}{4}a_3^2 - R^2 - 2a_2 + \frac{1}{4}(4a_3a_2 - 8a_1 - a_3^3)R^{-1}} & \text{for} \quad R \neq 0 \\ \sqrt{\frac{3}{4}a_3^2 - 2a_2 + 2\sqrt{y_1^2 - 4a_0}} & \text{for} \quad R = 0 \end{cases} \tag{C.21}$$

$$E = \begin{cases} \sqrt{\frac{3}{4}a_3^2 - R^2 - 2a_2 - \frac{1}{4}(4a_3a_2 - 8a_1 - a_3^3)R^{-1}} & \text{for} \quad R \neq 0 \\ \sqrt{\frac{3}{4}a_3^2 - 2a_2 - 2\sqrt{y_1^2 - 4a_0}} & \text{for} \quad R = 0 \end{cases}. \tag{C.22}$$

D

Outage Probability for SC

In this appendix, a general expression on outage probability for a parallel relay network with an arbitrary number of K relays is derived, where the destination performs selection combining (SC). The results are then used to calculate the outage probability for independent channels in a Rayleigh fading environment. It is mainly due to [161].

Define the outage event $\mathcal{A}_{ij} = \{$"transmission from node i to node j fails"$\}$. We use $\mathcal{A}_{ij}^c$ to denote the complement of $\mathcal{A}_{ij}$. The joint probability that k, $0 \le k \le K$, connections from the source to K relays fail can be written as

$$\sum_{1 \le i_1 < \ldots < i_k \le K} \Pr(\mathcal{A}_{\mathrm{s}i_1}, \ldots, \mathcal{A}_{\mathrm{s}i_k}) = \sum_{1 \le i_1 < \ldots < i_k \le K} \Pr\left(\bigcap_{l=1}^{k} \mathcal{A}_{\mathrm{s}i_l}\right), \tag{D.1}$$

where $1 \le i_1 < \ldots < i_k \le K$ describes all $\binom{K}{k}$ combinations. In a next step, we have to treat the fact that all $K - k$ relays, that could decode the source signal, cannot transmit reliably to the destination. Hence, when k channels to the relays fail, an outage occurs with probability

$$\sum_{1 \le i_1 < \ldots < i_k \le K} \Pr\left(\bigcap_{l=1}^{k} \mathcal{A}_{\mathrm{s}i_l} \bigcap_{\substack{m=1 \\ m \notin \{i_1, \ldots, i_k\}}}^{K} (\mathcal{A}_{\mathrm{s}m}^c\, \mathcal{A}_{m\mathrm{d}})\right). \tag{D.2}$$

Lastly, direct transmission from the source to the destination must fail, too. We denote this outage event as $\mathcal{A}_{\mathrm{sd}}$. The general expression of outage probability p_{out} then finally becomes

$$p_{\mathrm{out}} = \sum_{k=0}^{K} \sum_{1 \le i_1 < \ldots < i_k \le K} \Pr\left(\mathcal{A}_{\mathrm{sd}} \bigcap_{l=1}^{k} \mathcal{A}_{\mathrm{s}i_l} \bigcap_{\substack{m=1 \\ m \notin \{i_1, \ldots, i_k\}}}^{K} (\mathcal{A}_{\mathrm{s}m}^c\, \mathcal{A}_{m\mathrm{d}})\right). \tag{D.3}$$

As mentioned in Section 6.3, we examined selection combining here in contrast to [171, 172], where maximal ratio combining was investigated. This means that we experience no accumulation of SNR at the destination and transmission fails if the individual SNR of each incoming branch is below a required threshold SNR $\mathsf{SNR}_{\mathrm{th}}$.

It is quite realistic that different branches from the source to the relays and from the relays to the destination are statistically independent. This assumption leads to an outage probability expression of

$$p_{\mathrm{out}} = \Pr\left(\mathcal{A}_{\mathrm{sd}}\right) \sum_{k=0}^{K} \sum_{1 \leq i_1 < \ldots < i_k \leq K} \prod_{l=1}^{k} \Pr\left(\mathcal{A}_{\mathrm{s}i_l}\right) \prod_{\substack{m=1 \\ m \notin \{i_1,\ldots,i_k\}}}^{K} \Pr\left(\mathcal{A}_{\mathrm{s}m}^{c}\right) \Pr\left(\mathcal{A}_{m\mathrm{d}}\right). \qquad \text{(D.4)}$$

The outage probability from node i to node j in a Rayleigh fading environment is given in, e.g., [55]. We have

$$\Pr(\mathsf{SNR}_{ij} \leq \mathsf{SNR}_{\mathrm{th}}) = \Pr(\mathcal{A}_{ij}) = 1 - \exp\left(-\frac{\mathsf{SNR}_{\mathrm{th}}}{\overline{\mathsf{SNR}}_{ij}}\right), \qquad \text{(D.5)}$$

where $\mathsf{SNR}_{\mathrm{th}}$ describes the required SNR for reliable communications and $\overline{\mathsf{SNR}}_{ij}$ is the average SNR at node j. Inserting (D.5) into (D.4) yields

$$\begin{aligned} p_{\mathrm{out}} \;=\; & \left(1 - \exp\left(-\frac{\mathsf{SNR}_{\mathrm{th}}}{\overline{\mathsf{SNR}}_{\mathrm{sd}}}\right)\right) \sum_{k=0}^{K} \sum_{1 \leq i_1 < \ldots < i_k \leq K} \prod_{l=1}^{k} \left(1 - \exp\left(-\frac{\mathsf{SNR}_{\mathrm{th}}}{\overline{\mathsf{SNR}}_{\mathrm{s}i_l}}\right)\right) \\ & \prod_{\substack{m=1 \\ m \notin \{i_1,\ldots,i_k\}}}^{K} \exp\left(-\frac{\mathsf{SNR}_{\mathrm{th}}}{\overline{\mathsf{SNR}}_{\mathrm{s}m}}\right) \left(1 - \exp\left(-\frac{\mathsf{SNR}_{\mathrm{th}}}{\overline{\mathsf{SNR}}_{m\mathrm{d}}}\right)\right). \end{aligned} \qquad \text{(D.6)}$$

Notations & Symbols

Notations

$_x$	index
x	variable
$x!$	factorial of x
x^*	optimization of x
$\hat{x}$	estimation of x
$\mathcal{X}$	set
$\mathcal{X}^c$	complement of $\mathcal{X}$
$\mathbf{x}$	vector
$\mathbf{x}^T$	transposed vector
$\mathbf{X}$	matrix
$(\mathbf{XY})_{l,m}$	element of the l-th row and the m-th column of the matrix product $\mathbf{XY}$
$\mathbf{X} \circ \mathbf{Y}$	Hadamard product of $\mathbf{X}$ and $\mathbf{Y}$
X	random variable
$f_X(\cdot)$	probability density function of random variable X
$F_X(\cdot)$	cumulative distribution function of random variable X
$\overline{x}$	mean value
σ	standard deviation
σ^2	variance
$\mathcal{N}(\mu,\sigma^2)$	normal distribution with mean μ and variance σ^2
$\arg\max$	argument of the maximum
$\arg\min$	argument of the minimum

$\mathbb{E}(\cdot)$	expectation operator
$\in$	element of
j	imaginary unit with the property $j^2 = -1$
$\liminf$	limit inferior
$\limsup$	limit superior
$\log_2$	logarithm function to the base 2
max	maximum function
min	minimum function
$O(\cdot)$	Landau symbol
$\Pr(\cdot)$	probability
$Q(\cdot)$	Gaussian Q-function
$\mathbb{R}^+$	set of all positive real numbers
$x \propto y$	x proportional to y
$\mathcal{X} \subseteq \mathcal{Y}$	$\mathcal{X}$ is subset of or included in $\mathcal{Y}$
$\bigcap_{l=1}^{L} \mathcal{X}_l$	intersection of $\mathcal{X}_l$ ($\mathcal{X}_1 \cap \mathcal{X}_2 \cap \ldots \cap \mathcal{X}_L$)
$\binom{n}{k}$	binomial coefficient ("n choose k")
$\lvert \cdot \rvert$	magnitude
$\lvert\lvert \cdot \rvert\rvert_2$	Euclidean norm
$*$	convolution operator
$:=$	definition

Symbols

a	amplification factor
$A(\cdot, \cdot, \cdot)$	intersection area of two circles
B	bandwidth
B_c	coherence bandwidth
B_D	Doppler spread
BER	bit error rate
c	speed of light in free space
C	channel capacity
$\mathcal{C}$	normalized capacity C/B
$\mathcal{C}_\epsilon$	ϵ-outage capacity
CIR	carrier-to-interference ratio
d	diversity order
d_{ij}	distance between node i and node j
D	direction of movement
D	destination

E	total energy
E_r	remaining energy
$\mathbf{E}$	energy allocation
f	frequency
f_0	carrier frequency
f_D	Doppler frequency shift
$f_{D,\max}$	maximum Doppler frequency shift
g_k	weighting coefficient for the k-th signal
$G_\mathcal{C}$	capacity gain
G_d	antenna gain at the destination
G_s	antenna gain at the source
h_{ij}	channel gain between node i and j
$h(X)$	differential entropy of X
$h(X\|Y)$	conditional differential entropy of X given Y
$I(X;Y)$	mutual information of X and Y
K	number of relays
L	number of (received) paths
M	number of antennas at the source and/or the destination
$n(t)$	additive white Gaussian noise
N	number of transmission phases (i.e., sub-blocks)
$\tilde{N}$	average noise power
N_0	one-sided noise power spectral density
$\mathcal{N}$	set of neighboring nodes
p_dec	decoding probability
p_out	outage probability
p_nout	probability for no outage $(1-p_\mathrm{out})$
P	average signal power
$P(\cdot)$	power allocation function
P_d	receive power at the destination
P_r	relay transmit power
P_rem	remaining battery power

$P_{\mathrm{R}_k\mathrm{D}}$	probability destination can decode after transmission of the k-th relay
P_{s}	source transmit power
P_{SD}	probability source-destination transmission succeeded
P_{tot}	total transmit power
$\bar{P}_{\mathrm{R}_k\mathrm{D}}$	probability destination cannot decode after transmission of the k-th relay
$\bar{P}_{\mathrm{SD}}$	probability source-destination transmission failed
$\mathtt{PL}$	path loss
$\mathbf{P}$	power allocation vector
q	channel quality
r	multiplexing gain
r_k	radius of the k-th node's transmission range
R	target transmission rate
R_k	k-th relay
$\bar{R}$	average (long-term) transmission rate
S	source
SNR	signal-to-noise ratio
t	(discrete) time
T	block length
T_c	coherence time
T_D	delay spread
v	speed of the moving receiver
W	willingness to cooperate
$x(t)$	discrete-time transmit signal
$\mathcal{X}$	support set of random variable X
$y(t)$	discrete-time receive signal
α	path loss exponent
β	decision threshold value
γ	weighted sum of the intrinsic relay parameters
$\Delta(\epsilon)$	ratio of ϵ-outage capacities

Δ_{SNR}	SNR gain
ϵ	target error probability
η	optimization criterion
θ	angle between direction of electro-magnetic wave and motion
ϑ	counter value
Θ	branch unbalance
κ	flag
λ	wavelength
λ_X	parameter of exponential random variable X
$\boldsymbol{\pi}$	matrix of intrinsic relay parameters
ρ	angle
τ	time fraction
$\tau_i(t)$	time delay of the i-th path
$\boldsymbol{\tau}$	time allocation vector
ϕ	golden ratio ($\approx 1.61803398\ldots$)
φ	phase of channel gain
Ω	weighting coefficient

Abbreviations

3G	third-generation
3GPP	3rd generation partnership project

A

ACK	acknowledgment
AF	amplify-and-forward
AMR	adaptive multi-route
ARQ	automatic repeat request
ARSP	adaptive relay selection protocol
AWGN	additive white Gaussian noise

B

B3G	beyond-third-generation
BAF	bursty amplify-and-forward
BC	broadcast channel
BER	bit error rate
BPSK	binary phase shift keying
BSC	binary symmetric channel

C

CDI	channel distribution information
CDMA	code division multiple access
CF	compress-and-forward
CIR	carrier-to-interference ratio
CoNET	cooperative network working group
CSI	channel state information
CSIR	channel state information at the receiver

CSIT	channel state information at the transmitter
CTS	clear-to-send

D

DCF	distributed coordination function
DF	decode-and-forward
DSDV	destination sequenced distance vector
DT	direct transmission

E

EGC	equal gain combining

F

FDMA	frequency division multiple access

I

ID	identifier
i.i.d.	independent and identically distributed
IP	internet protocol
IR	incremental relaying
IRP	intrinsic relay parameter

L

LOS	line of sight
LTE	long term evolution

M

MAC	multiple access channel
MH	multi-hop
MIMO	multiple-input multiple-output
MR	multi-route
MRC	maximal ratio combining

N

NACK	negative acknowledgment
NAV	network allocation vector
NLOS	non-line of sight

O

OFDM	orthogonal frequency division multiplex
OFDMA	orthogonal frequency division multiple access

Q

QoS	quality of service
QPSK	quadrature phase shift keying

R

RT	relay table
RTS	request-to-send

S

SC	selection combining
SIFS	short interframe space
SNR	signal-to-noise ratio
STBC	space-time block coding
STC	space-time code

T

TDMA	time division multiple access

W

WiMAX	worldwide inter-operability for microwave access
WWRF	wireless world research forum

Bibliography

[1] IEEE Communications Society, *A Brief History of Communications.* IEEE, Piscataway, NJ, 2002.

[2] G. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications, Kluwer Academic Publishers*, vol. 6, pp. 311–335, 1998.

[3] İ. E. Telatar, "Capacity of multi-antenna gaussian channels," *Bell Labs Journal*, vol. 10, no. 6, November 1999.

[4] ——, "Capacity of multi-antenna gaussian channels," *European Transactions on Telecommunications*, vol. 10, pp. 585–595, November/December 1999.

[5] R. Kahn, "The organization of computer resources into a packet radio network," *IEEE Transactions on Communications*, vol. 25, no. 1, pp. 169–178, January 1977.

[6] R. Kahn, S. Gronemeyer, J. Burchfiel, and R. Kunzelman, "Advances in packet radio technology," *Proceedings of the IEEE*, vol. 66, no. 11, pp. 1468–1496, November 1978.

[7] R. Nabar, H. Bölcskei, and F. Kneubühler, "Fading relay channels: Performance limits and spaceŰtime signal design," *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 6, pp. 1099–1109, August 2004.

[8] H. Bölcskei, R. U. Nabar, O. Oyman, and A. J. Paulraj, "Capacity scaling laws in MIMO relay networks," *IEEE Trans. Wireless Communications*, vol. 5, no. 6, pp. 1433–1444, June 2006.

[9] A. Ganesan and A. Sayeed, "A virtual input-output framework for transceiver analysis and design for multipath fading channels," *IEEE Transactions on Communications*, vol. 51, no. 7, pp. 1149–1160, July 2003.

[10] K. Liu, V. Raghavan, and A. Sayeed, "Capacity scaling and spectral efficiency in wide-band correlated mimo channels," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2504–2626, October 2003.

[11] T. Renk, H. Jaekel, and F. Jondral, "Relay-Netzwerke – Kooperative Kommunikation unter mobilen Teilnehmern," *Telekommunikation Aktuell*, vol. 61, no. 08-09, pp. 1–19, August/September 2007.

[12] T. Renk, C. Kloeck, D. Burgkhardt, and F. K. Jondral, "Cooperative communications in wireless networks – a requested relaying protocol," *16th IST Mobile & Wireless Communications Summit*, Budapest, Hungary, 2007.

[13] TerraNet AB, "http://www.terranet.se," February 2010.

[14] F. Fitzek and M. Katz, *Cooperation in Wireless Networks: Principles and Applications*. Springer, 2006.

[15] C. Politis *et al.*, "Cooperative networks for the future wireless world," *IEEE Communications Magazine*, vol. 42, no. 9, pp. 70–79, September 2004.

[16] C. Eklund *et al.*, "IEEE standard 802.16: A technical overview of the WirelessMAN air interface for broadband wireless access," *IEEE Communications Magazine*, vol. 40, no. 6, pp. 98–107, June 2002.

[17] A. Ghosh *et al.*, "Broadband wireless access with WiMAX/802.16: Current performance benchmarks and future potential," *IEEE Communications Magazine*, vol. 43, no. 2, pp. 129–136, February 2005.

[18] J. Andrews, A. Ghosh, and R. Muhamed, *Fundamentals of WiMAX*. Prentice-Hall, 2007.

[19] S. Peters and R. Heath Jr., "The future of WiMAX: Multihop relaying with IEEE 802.16j," *IEEE Communications Magazine*, vol. 47, no. 1, pp. 104–111, January 2009.

[20] O. Oyman, J. Laneman, and S. Sandhu, "Multihop relaying for broadband wireless mesh networks: From theory to practice," *IEEE Communications Magazine*, vol. 45, no. 11, pp. 116–122, November 2007.

[21] V. Venkatkumar, T. Wirth, T. Haustein, and E. Schulz, "Relaying in Long Term Evolution: Indoor full frequency reuse," *European Wireless 2009*, Aarlborg, Denmark, May 2009.

[22] T. Wirth, V. Venkatkumar, T. Haustein, E. Schulz, and R. Halfmann, "LTE-Advanced relaying for outdoor range extension," *IEEE 70th Vehicular Technology Conference (VTC-Fall)*, Anchorage, Alaska, September 2009.

[23] E. van der Meulen, "Transmission of information in a t-terminal discrete memoryless channel," *Department of Statistics, University of California, Berkeley, CA, Technical Report*, 1968.

[24] ——, "Three-terminal communication channels," *Advances in Applied Probability*, vol. 3, no. 1, pp. 120–154, 1971.

[25] T. Cover and A. El Gamal, "Capacity theorems for the relay channel," *IEEE Transactions on Information Theory*, vol. 25, no. 5, pp. 572–584, September 1979.

[26] M. Aref, "Information flow in relay networks," Ph.D. dissertation, Stanford University, 1981.

[27] E. van der Meulen, "A survey of multi-way channels in information theory: 1961-1976," *ieeeit*, vol. 23, pp. 1–37, 1977.

[28] D. Slepian and J. Wolf, "Noiseless coding of correlated information sources," *IEEE Transactions on Information Theory*, vol. 19, no. 4, pp. 471–480, July 1973.

[29] N. Gaarder and J. Wolf, "The capacity region of a multiple-access discrete memoryless channel can increase with feedback," *IEEE Transactions on Information Theory*, vol. 21, no. 1, pp. 100–102, January 1975.

[30] A. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the receiver," *IEEE Transactions on Information Theory*, vol. 22, no. 1, pp. 1–11, January 1976.

[31] F. Willems, "Informationtheoretical results for the discrete memoryless multiple access channel," Ph.D. dissertation, Katholieke Universiteit Leuven, Belgium, 1982.

[32] V. Tarokh, N. Seshadri, and A. Calderbank, "Space-time codes for high data rate wireless communication: performance criterion and code construction," *IEEE Transactions on Information Theory*, vol. 44, no. 2, pp. 744–765, March 1998.

[33] V. Tarokh, H. Jafarkhani, and A. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Transactions on Information Theory*, vol. 45, no. 5, pp. 1456–1467, July 1999.

[34] B. Schein, "Distributed coordination in network information theory," Ph.D. dissertation, Massachusetts Institute of Technology, October 2001.

[35] A. Sendonaris, E. Erkip, and B. Aazhang, “User cooperation diversity – part I: System description,” *IEEE Transactions on Communications*, vol. 51, pp. 1927–1938, 2003.

[36] ——, “User cooperation diversity – part II: Implementation aspects and performance analysis,” *IEEE Transactions on Communications*, vol. 51, pp. 1939–1948, 2003.

[37] J. Laneman, D. Tse, and G. Wornell, “Cooperative diversity in wireless networks: Efficient protocols and outage behavior,” *IEEE Transactions on Information Theory*, vol. 50, no. 12, pp. 3062–3080, December 2004.

[38] J. 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, 2003.

[39] G. Kramer, M. Gastpar, and P. Gupta, “Cooperative strategies and capacity theorems for relay networks,” *IEEE Transactions on Information Theory*, vol. 51, no. 9, pp. 3037–3063, September 2005.

[40] A. Høst-Madsen and J. Zhang, “Capacity bounds and power allocation for wireless relay channels,” *IEEE Transactions on Information Theory*, vol. 51, no. 6, pp. 2020–2040, June 2005.

[41] J. Boyer, D. Falconer, and H. Yanikomeroglu, “A theoretical characterization of the multihop wireless communications channel without diversity,” *International Symposium on Personal, Indoor and Mobile Radio communications*, San Diego, California, USA, September/October 2004.

[42] ——, “Multihop diversity in wireless relaying channels,” *IEEE Transactions on Communications*, vol. 52, no. 10, pp. 1820–1830, 2004.

[43] A. Bletsas, “Intelligent antenna sharing in cooperative diversity wireless networks,” Ph.D. dissertation, Massachusetts Institute of Technology, September 2005.

[44] A. Bletsas and A. Lippman, “Implementing cooperative diversity antenna arrays with commodity hardware,” *IEEE Communications Magazine*, vol. 44, no. 12, pp. 33–40, December 2006.

[45] A. Bletsas, A. Khisti, and M. Z. Win, “Low complexity virtual antenna arrays using cooperative relay selection,” *International Wireless Communications and Mobile Computing Conference*, pp. 461–466, Vancouver, British Columbia, Canada, July 2006.

[46] L. Xie and P. Kumar, "A network information theory for wireless communication: Scaling laws and optimal operation," *IEEE Transactions on Information Theory*, vol. 50, no. 5, pp. 748–767, May 2004.

[47] ——, "An achievable rate for the multiple-level relay channel," *IEEE Transactions on Information Theory*, vol. 51, no. 4, pp. 1348–1358, April 2005.

[48] A. Goldsmith and S. Wicker, "Design challenges for energy-constrained ad hoc wireless networks," *IEEE Wireless Communications Magazine*, vol. 9, no. 4, pp. 8–27, August 2002.

[49] S. Cui and A. Goldsmith, "Cross-layer design in energy-constrained networks using cooperative mimo techniques," *EURASIP Journal on Applied Signal Processing, Special Issue on Advances in Signal Processing-based Cross-layer Designs*, vol. 86, no. 8, pp. 1804–1814, August 2002.

[50] O. K. Tonguz and G. Ferrari, *Ad Hoc Wireless Networks; A Communication-Theoretic Perspective.* John Wiley & Sons, Ltd, 2006.

[51] J. Chang and L. Tassiulas, "Maximum lifetime routing in wireless sensor networks," *IEEE/ACM Transactions on Networking*, vol. 12, no. 4, pp. 609–619, August 2004.

[52] T. Cover and J. Thomas, *Elements of Information Theory.* John Wiley & Sons, Inc., Wiley Series in Telecommunications, New York Chichester Brisbane Toronto Signapore, 1991.

[53] A. Goldsmith, *Wireless Communications.* Cambridge University Press, 2005.

[54] J. Proakis, *Digital Communications*, 4th ed. New York: McGraw–Hill, 2001.

[55] T. Rappaport, *Wireless Communications: Principles & Practice.* New Jersey: Prentice Hall, 1996.

[56] E. Biglieri, J. Proakis, and S. Shamai (Shitz), "Fading channels: Information-theoretic and communications aspects," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2619–2692, October 1998.

[57] J. Laneman, "Cooperative diversity in wireless networks: Algorithms and architectures," Ph.D. dissertation, Massachusetts Institute of Technology, September 2002.

[58] S. Hanly and D. Tse, "Multi-access fading channels–Part II: Delay-limited capacities," *IEEE Transactions on Information Theory*, vol. 44, no. 7, pp. 2816–2831, November 1998.

[59] A. Goldsmith and P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Transactions on Information Theory*, vol. 43, no. 6, pp. 1986–1992, November 1995.

[60] L. H. Ozarow, S. Shamai (Shitz), and A. D. Wyner, "Information theoretic considerations for cellular mobile radio," *IEEE Transactions on Vehicular Technology*, vol. 43, no. 2, pp. 359–377, May 1994.

[61] S. Verdú and T. Han, "A general formula for channel capacity," *IEEE Transactions on Information Theory*, vol. 40, no. 4, pp. 1147–1157, July 1994.

[62] R. Berry and R. Gallager, "Communication over fading channels with delay constraints," *IEEE Transactions on Information Theory*, vol. 48, no. 5, pp. 1135–1149, May 2002.

[63] A. Sadek, "Cross-layer design for cooperative communications and networking," Ph.D. dissertation, University of Maryland, 2007.

[64] W. Jakes, Jr., *Microwave Mobile Communications.* Wiley, New York, 1974.

[65] M. Pätzold, *Mobile Fading channels.* Wiley, New York, 2002.

[66] D. Parsons, *The Mobile Radio Propagation Channel.* Wiley, New York, 1994.

[67] B. Sklar, "Rayleigh fading channels in mobile digital communication systems Part I: Characterization," *IEEE Communications Magazine*, vol. 35, no. 7, pp. 90–100, July 1997.

[68] H. Friis, "A note on a simple transmission formula," *Proceedings of the IRE*, vol. 34, no. 5, pp. 254–256, May 1946.

[69] E. Biglieri, G. Caire, and G. Taricco, "Coding for the fading channel: a survey," *Signal Processing for Multimedia (J.S. Byrnes)*, vol. 80, no. 7, pp. 1135–1148, 2000.

[70] G. Kramer, I. Marić, and R. Yates, "Cooperative communications," *Foundations and Trends in Networking*, vol. 1, no. 3-4, pp. 271–425, 2006.

[71] E. Biglieri, R. Calderbank, A. Constantinides, A. Goldsmith, A. Paulraj, and V. Poor, *MIMO Wireless Communications.* Cambridge University Press, 2007.

[72] M. Kiessling, J. Speidel, and M. Reinhardt, "Ergodic capacity of MIMO channels with statistical channel state information at the transmitter," *ITG Workshop on Smart Antennas*, pp. 79–86, Munich, Germany, March 2004.

[73] N. Ahmed, M. Khojastepour, and B. Aazhang, "Outage minimization and optimal power control for the fading relay channel," *IEEE Information Theory Workshop (ITW)*, pp. 458–462, San Antonio, Texas, USA, October 2004.

[74] D. Gündüz and E. Erkip, "Opportunistic cooperation by dynamic resource allocation," *IEEE Transactions on Wireless Communications*, vol. 6, no. 4, pp. 1446–1454, April 2004.

[75] R. Gallager, *Information Theory and Reliable Communication.* John Wiley & Sons, Inc., New York London Sydney Toronto, 1968.

[76] T. Renk, C. Kloeck, H. Jaekel, and F. Jondral, "Error performance of multihopping half-duplex relay networks," *IEEE International Symposium on Wireless Pervasive Computing (ISWPC)*, Santorini, Greece, May 2008.

[77] D. Brennan, "Linear diversity combining techniques," *Proceedings of the IRE*, vol. 47, no. 1, pp. 1075–1102, 1959.

[78] P. Herhold, E. Zimmermann, and G. Fettweis, "Cooperative multi-hop transmission in wireless networks," *Journal on Computer Networks*, vol. 49, no. 3, pp. 299–324, 2005.

[79] T. Renk, H. Jaekel, F. Jondral, D. Gündüz, and A. Goldsmith, "Outage capacity of incremental relaying at low signal-to-noise ratios," *IEEE 70th Vehicular Technology Conference (VTC-Fall)*, Anchorage, Alaska, September 2009.

[80] ——, "Outage capacity of bursty amplify-and-forward with incremental relaying," *submitted to International Symposium on Information Theory*, Austin, Texas, USA, June 2010.

[81] T. Hunter and A. Nosratinia, "Cooperative diversity through coding," *International Symposium on Information Theory*, Lausanne, Switzerland, July 2002.

[82] ——, "Diversity through coded cooperation," *IEEE Transactions on Wireless Communications*, vol. 5, no. 2, pp. 283–289, February 2006.

[83] M. Khojastepour, A. Sabharwal, and B. Aazhang, "Lower bounds on the capacity of gaussian relay channel," *38th Annual Conference on Information Sciences and Systems (CISS)*, Princeton, New Jersey, USA, March 2004.

[84] R. Hu and J. Li, "Practical compress-forward in user cooperation: Wyner-Ziv cooperation," *International Symposium on Information Theory*, Seattle, Washington, USA, July 2006.

[85] M. Dohler and H. Aghvami, "Wireless cooperative communication networks," *IEEE International Conference on Communications (ICC)*, Tutorial, Glasgow, Scotland, June 2007.

[86] F. Ng, J. Hwu, C. Mo, and L. Xiaohua, "Asynchronous space-time cooperative communications in sensor and robotic networks," *International Conference on Mechatronics & Automation*, Niagara Falls, Canada, July 2005.

[87] S. Wei, "Diversity-multiplexing tradeoff of asynchronous space-time-coded cooperative diversity," *42nd Annual Allerton Conference on Communication, Control, and Computing*, Monticello, IL, September/October 2004.

[88] M. Montenegro Toro, "Relay selection in cooperative networks," Master's thesis, Universität Karlsruhe (TH) and Politecnico di Torino, August 2008.

[89] A. Bletsas and A. Khisti, "The diversity-multiplexing tradeoff in cooperative protocols with relay selection," *MSRI Workshop on the Mathematics of Relaying*, Berkeley, California, USA, April 2006.

[90] A. Bletsas, A. Lippman, and D. Reed, "A simple distributed method for relay selection in cooperative diversity wireless networks, based on reciprocity and channel measurements," *IEEE 61st Vehicular Technology Conference (VTC-Spring 2006)*, Stockholm, Sweden, May/June 2005.

[91] A. Bletsas, A. Khisti, D. Reed, and A. Lippman, "A simple cooperative diversity method based on network path selection," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 3, pp. 659–672, March 2006.

[92] V. Streng, H. Yanikomeroglu, and D. Falconer, "Coverage enhancement through two-hop relaying in cellular radio system," *IEEE Wireless Communications and Networking Conference (WCNC)*, Orlando, Florida, USA, March 2002.

[93] J. Chu, R. Adve, and A. Eckford, "Relay selection for low-complexity coded cooperation," *IEEE Global Telecommunications Conference (GLOBECOM 03)*, Washington, D.C., USA, November 2007.

[94] A. Nosratinia and T. Hunter, "Grouping and partner selection in cooperative wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 2, pp. 369–378, February 2007.

[95] K. Hwang and Y. Ko, "An efficient relay selection algorithm for cooperative networks," *IEEE 66th Vehicular Technology Conference (VTC-Fall 2007)*, Baltimore, Maryland, USA, September 2007.

[96] C. Chen, B. Zheng, X. Zhao, and Z. Yan, "A novel weighted cooperative routing algorithm based on distributed relay selection," *2nd International Symposium on Wireless Pervasive Computing (ISWPC)*, San Juan, Puerto Rico, February 2007.

[97] Y. Li, B. Vucetic, Z. Chen, and J. Yuan, "An improved relay selection scheme with hybrid relaying protocols," *2nd International Symposium on Wireless Pervasive Computing (ISWPC)*, San Juan, Puerto Rico, February 2007.

[98] A. del Coso and C. Ibars, "Distributed antenna channels with regenerative relaying: relay selection and asymptotic capacity," *EURASIP Journal on Wireless Communications and Networking*, vol. 3, no. 9, July 2007.

[99] A. Nosratinia, T. Hunter, and A. Hedayat, "Cooperative communication in wireless networks," *IEEE Communications Magazine*, vol. 42, no. 10, pp. 74–80, 2004.

[100] T. Oechtering and H. Boche, "Relay selection in bidirectional relay communication," *IEEE 8th Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Helsinki, Finland, June 2007.

[101] V. Mahinthan, L. Cai, J. Mark, and X. Shen, "Partner selection based on optimal power allocation in cooperative diversity systems," *IEEE Transactions on Vehicular Technology*, vol. 57, no. 1, pp. 511–520, January 2008.

[102] M. Souryal and N. Moayeri, "Channel-adaptive relaying in mobile ad hoc networks with fading," *IEEE Second Annual IEEE Communications Society Conference on Sensor and Ad Hoc Communications and Networks (SECON)*, Santa Clara, California, USA, September 2005.

[103] J. Gómez-Vilardebó and A. Pérez-Neira, "Energy allocation, relay selection and ordering in orthogonal relay networks," *Fortieth Asilomar Conference on Signals, Systems and Computers*, pp. 196–200, Pacific Grove, California, USA, October/November 2006.

[104] K. Caleb, R. Heath, and S. Vishwanath, "Hybrid-ARQ in multihop networks with opportunistic relay selection," 2007. [Online]. Available: http://www.citebase.org/abstract?id=oai:arXiv.org:0704.0802

[105] H. Yang and A. Petropulu, "Alliances with optimal relay selection," *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1265–1268, Honolulu, Hawaii, USA, April 2007.

[106] 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, December 2005.

[107] T. Ng, W. Yu, J. Zhang, and A. Reid, "Joint optimization of relay strategies and resource allocations in cooperative cellular networks," *40th Annual Conference on Information Sciences and Systems (CISS)*, Princeton, New Jersey, USA, March 2006.

[108] Z. Lin, E. Erkip, and A. Stefanov, "Cooperative regions and partner choice in coded cooperative systems," *IEEE Transactions on Communications*, vol. 54, no. 7, pp. 1323–1334, July 2006.

[109] I. Bronstein, K. Semendjajew, G. Musiol, and H. Mühlig, *Taschenbuch der Mathematik*, 5th ed. Verlag Harri Deutsch, Thun, Frankfurt am Main, 2001.

[110] L. M. Feeney, D. Hollos, M. Kubisch, S. Mengesha, and H. Karl, "A geometric derivation of the probability of finding a relay in multi-rate networks," *3rd IFIP Conference on Networking*, 2004.

[111] *Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications, ISO/IEC 8802-11:2005.*

[112] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 3, pp. 535–547, March 2000.

[113] R. Litjens *et al.*, "Performance analysis of wireless LANs: an integrated packet/flow level approach," *18th International Teletraffic Congress (ITC18)*, Berlin, Germany, August/September 2003.

[114] P. Herhold, E. Zimmermann, and G. Fettweis, "A simple cooperative extension to wireless relaying," *Int. Zurich Seminar on Communications (IZS)*, February 2004.

[115] R. Steuer, *Multiple Criteria Optimization: Theory, Computation, and Application.* Krieger Publishing Company, 1989.

[116] T. Renk *et al.*, "Spectrum measurements supporting reconfiguration in heterogeneous networks," *16th IST Mobile & Wireless Communications Summit*, Budapest, Hungary, 2007.

[117] ——, "Bio-inspired algorithms for dynamic resource allocation in cognitive wireless networks," *Second International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CrownCom)*, Orlando, FL, July/August 2007.

[118] T. Renk, V. Blaschke, and F. Jondral, "Time-dependent statistical analysis of measurements for the evaluation of vacant spectrum bands," *XXIX. U.R.S.I. General Assembly*, Chicago, IL, August 2008.

[119] T. Renk *et al.*, "Bio-inspired algorithms for dynamic resource allocation in cognitive wireless networks," *Springer – Mobile Networks and Applications*, vol. 13, pp. 431–441, 2008.

[120] J. Zander, "Radio resource management in future wireless networks: requirements and limitations," *IEEE Communications Magazine*, vol. 35, no. 8, pp. 30–36, august 1997.

[121] T. Weiss, "OFDM-basiertes Spectrum Pooling," *Forschungsberichte aus dem Institut für Nachrichtentechnik der Universität Karlsruhe (TH)*, December 2004.

[122] C. Kloeck, "Auction-based medium access control," *Forschungsberichte aus dem Institut für Nachrichtentechnik der Universität Karlsruhe (TH)*, October 2007.

[123] T. Renk, D. Iankov, and F. Jondral, "Adaptive resource allocation in wireless relay networks," *IEEE 69th Vehicular Technology Conference (VTC-Spring)*, Barcelona, Spain, April 2009.

[124] M. O. Hasna and M.-S. Alouini, "Optimal power allocation for relayed transmissions over rayleigh-fading channels," *IEEE Transactions on Wireless Communications*, vol. 3, no. 6, pp. 1999–2004, November 2004.

[125] C. Ng and A. Goldsmith, "Capacity and power allocation for transmitter and receiver cooperation in fading channels," *IEEE International Conference on Communications (ICC)*, pp. 3741–3746, Istanbul, Turkey, June 2006.

[126] T. Renk, H. Jaekel, and F. Jondral, "Outage regions and optimal power allocation for wireless relay networks," *IEEE Information Theory Workshop (ITW)*, Taormina, Italy, October 2009.

[127] ——, "Outage regions of cooperative networks for different coding schemes," *4th International Symposium on Communications, Control and Signal Processing (ISCCSP)*, Limassol, Cyprus, March 2010.

[128] Y. Liang and V. Veeravalli, "Resource allocation for wireless relay channel," *38th Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, November 2004.

[129] S. Serbetli and A. Yener, "Power allocation strategies for relay assisted communications," *WICAT Workshop on Cooperative Communications*, Brooklyn, New York, USA, October 2005.

[130] K. Bakanoglu, D. Gündüz, and E. Erkip, "Dynamic resource allocation for the broadband relay channel," *Forty-First Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, California, USA, November 2007.

[131] T. Renk and F. Jondral, "Location-dependent performance of cooperative networks with power allocation," *IEEE Symposium on Computers and Communications (ISCC)*, Sousse, Tunisia, July 2009.

[132] R. P. Brent, *Algorithms for Minimization Without Derivatives.* Englewood Cliffs, New Jersey: Prentice-Hall, 1973.

[133] D. Anderson, "Algorithms for minimization without derivatives," *IEEE Transactions on Automatic Control*, vol. 19, pp. 632–633, October 1974.

[134] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed. Cambridge University Press, 1992.

[135] D. G. Brennan, "Linear diversity combining techniques," *Proceedings of the IRE*, vol. 47, no. 1, pp. 1075–1102, June 1959.

[136] P. Chakravarty, T. Renk, and F. Jondral, "Combining receiver for wireless relay networks based on signal-to-noise ratio criterion," *5th Karlsruhe Workshop on Software Radios (WSR)*, Karlsruhe, Germany, March 2008.

[137] T. Renk, H. Jaekel, and F. Jondral, "A novel combining receiver for a dual-diversity wireless relay network," *The 12th International Symposium on Wireless Personal Multimedia Communications (WPMC)*, Sendai, Japan, September 2009.

[138] X. Qi, M.-S. Alouini, and Y.-C. Ko, "Closed-form analysis of dual-diversity equal-gain combining over rayleigh fading channels," *IEEE Transactions on Wireless Communications*, vol. 2, no. 6, pp. 1120–1125, November 2003.

[139] S. W. Halpern, "The effect of having unequal branch gains in practical predetection diversity systems for mobile radio," *IEEE Transactions on Vehicular Technologies*, vol. 26, pp. 94–105, February 1977.

[140] L. Zheng and D. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple antenna channels," *IEEE Transactions on Information Theory*, vol. 49, pp. 1073–1096, May 2003.

[141] A. Lozano and N. Jindal, "Transmit diversity v. spatial multiplexing in modern MIMO systems," 2009. [Online]. Available: http://arXiv.org/abs/0811.3887

[142] T. Renk, H. Jaekel, C. Kloeck, and F. Jondral, "Signal-to-noise ratio requirements for wireless relay networks," *3rd International Symposium on Communications, Control and Signal Processing (ISCCSP)*, St. Julians, Malta, March 2008.

[143] T. Renk, H. Jaekel, and F. Jondral, "Signal-to-noise ratio gains for cooperative relaying strategies," *IEEE International Symposium on Wireless Communication Systems (ISWCS)*, Reykjavik, Iceland, October 2008.

[144] T. Renk and F. Jondral, "Upper and lower bound on signal-to-noise ratio gains for cooperative relay networks," *43rd Annual Conference on Information Sciences and Systems (CISS)*, Baltimore, MD, March 2009.

[145] R. K. Mallik, M. Z. Win, and J. H. Winters, "Performance of predetection dual diversity in correlated rayleighfading: EGC and SD," *Global Telecommunications Conference*, San Francisco, CA, November/December 2000.

[146] ——, "Performance of dual-diversity predetection EGC in correlated rayleigh fading with unequal branch SNRs," *IEEE Transactions on Communications*, vol. 50, no. 7, pp. 1041–1044, July 2002.

[147] S. Ikki and M. H. Ahmed, "Performance analysis of cooperative diversity using equal gain combining (EGC) technique over rayleigh fading channels," *IEEE International Conference on Communications (ICC)*, Glasgow, Scotland, June 2007.

[148] Q. T. Zhang, "Probability of error for equal-gain combiners over rayleigh channels: Some closed-form solutions," *IEEE Transactions on Communications*, vol. 45, no. 3, pp. 270–273, March 1997.

[149] A. Goldsmith and S. Wicker, "Design challenges for energy-constrained ad-hoc wireless networks," *IEEE Wireless Communications*, vol. 9, no. 4, pp. 8–27, 2002.

[150] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication.* Cambridge University Press, New York, 2005.

[151] G. Caire, G. Taricco, and E. Biglieri, "Optimum power control over fading channels," *IEEE Transactions on Information Theory*, vol. 45, no. 5, pp. 1468–1489, July 1999.

[152] A. Avestimehr and D. Tse, "Outage capacity of the fading relay channel in the low-snr regime," *IEEE Transactions on Information Theory*, vol. 53, no. 4, pp. 1401–1415, April 2007.

[153] P. Wu and N. Jindal, "Performance of hybrid-arq in block-fading channels: A fixed outage probability analysis," available at: http://arxiv.org/pdf/0811.4191, November 2008.

[154] S. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1451–1458, October 1998.

[155] T. Renk, H. Jaekel, and F. Jondral, "Comparing the outage capacity of transmit diversity and incremental relaying," *IEEE 70th Vehicular Technology Conference (VTC-Fall)*, Anchorage, Alaska, September 2009.

[156] T. Renk, H. Jaekel, F. Jondral, and A. Goldsmith, "Do decode-and-forward relaying protocols beat transmit diversity," *European Wireless*, Lucca, Italy, April 2010.

[157] S. Verdú, "A general formula for channel capacity," *IEEE Transactions on Information Theory*, vol. 48, no. 6, pp. 1319–1343, June 2002.

[158] M. Golay, "Note on the theoretical efficiency of information reception with ppm," *Proceedings of the IRE*, vol. 37, no. 9, p. 1031, September 1949.

[159] G. Atia, M. Sharif, and V. Saligrama, "On optimal outage in relay channels with general fading distributions," *IEEE Transactions on Information Theory*, vol. 53, no. 10, pp. 3786–3797, October 2007.

[160] S. Cui, A. Haimovich, O. Somekh, and H. Poor, "Opportunistic relaying in wireless networks," *IEEE Transactions on Information Theory*, vol. 55, no. 11, pp. 5121–5137, November 2009.

[161] T. Renk, H. Jaekel, C. Kloeck, and F. Jondral, "General expression of outage probability in cooperative networks," *European Transactions on Telecommunications*, vol. 20, pp. 537–548, 2008.

[162] T. Renk *et al.*, "Outage capacity of incremental relaying with imperfect feedback," *submitted to International Symposium on Information Theory*, Austin, Texas, USA, June 2010.

[163] B. Rankov and A. Wittneben, "Spectral efficient signaling for half-duplex relay channels," *Conference Record of the Thirty-Ninth Asilomar Conference on Signals, Systems and Computers*, pp. 1066–1071, 2005.

[164] C. Richardson *et al.*, "The interphone study: Design, epidemiological methods, and description of the study population," *European Journal of Epidemiology*, vol. 22, pp. 647–664, July 2007.

[165] D. Ezri and S. Shimi, "Green cellular – optimizing the cellular network for minimal emission from mobile stations," 2009. [Online]. Available: http://arxiv.org/abs/0906.5289

[166] S. Cho, E. Jang, and J. Cioffi, "Handover in multihop cellular networks," *IEEE Communications Magazine*, vol. 47, no. 7, pp. 64–73, July 2009.

[167] F. Jondral and A. Wiesler, *Wahrscheinlichkeitsrechnung und stochastische Prozesse*, 2nd ed. B. G. Teubner: Stuttgart, Leipzig, Wiesbaden, 2002.

[168] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3rd ed. New York: McGraw-Hill, 1991.

[169] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 9th ed. New York: Dover, 1972.

[170] W. Mathworld, "Quartic equation," January, 19, 2010. [Online]. Available: http://mathworld.wolfram.com/QuarticEquation.html

[171] Y. Zhao, R. Adve, and T. Lim, “Outage probability at arbitrary SNR with cooperative diversity,” *IEEE Communications Letters*, vol. 9, no. 8, pp. 700–702, 2005.

[172] N. Beaulieu and J. Hu, “A closed-form expression for the outage probability of decode-and-forward relaying in dissimilar rayleigh fading channels,” *IEEE Communications Letters*, vol. 10, no. 12, pp. 813–815, 2006.

Supervised Theses

Aymen Benzid	*Untersuchung von Nachbarkanalinterferenzen in Abhängigkeit verschiedener Parameter* (Studienarbeit)
Prasanjit Chakravarty	*Empfängerstrukturen für Relay-Netzwerke* (Master's Thesis)
Samuel Hatorangan	*Analyse des Durchsatzes und der Verzögerung in Relay-Netzwerken* (Studienarbeit)
Samuel Hatorangan	*Resource Optimization in Cognitive Radio Coexistence* (Master's Thesis)
Dimitar Iankov	*Adaptive Ressourcenbelegung in Relay-basierten Netzwerken* (Studienarbeit)
Mauricio Montenegro	*Relay Selection in Cooperative Networks* (Master's Thesis)

Index

S

T

U

V

W

Sponsorship

The work of this dissertation was supported in part by the EU-funded project *End-to-End Reconfigurability*.

The research visit at the Wireless Systems Laboratory at Stanford University was supported in part by the Karlsruhe House of Young Scientists (KHYS) of the Karlsruhe Institute of Technology (KIT).

Curriculum Vitae

Personal Details

Date of birth:	February 11, 1981
Place of birth:	Kronach, Germany
Citizenship:	German

Education

11/2005–11/2010:	Scientific Assistant, Karlsruhe Institute of Technology
09/2008–12/2008:	Visiting Researcher, Stanford University
10/2004–10/2005:	Ph.D. Qualification, Universität Karlsruhe (TH)
12/2004:	B.Sc. of Honours, Open University of London
10/2001–10/2004:	Dipl.-Ing. (BA), Berufsakademie Ravensburg
09/1991–09/2000:	Frankenwald-Gymnasium Kronach

Awards

Best Student Paper Award finalist, VTC Fall 2009
KHYS scholarship for a research visit at Stanford University 2008
URSI Young Scientist Award 2008
ICST Student Registration Grant, CrownCom 2007