– Dissertation –

Information Theoretical Analysis of Cooperative Orthogonal Relay Networks

Sebastian Vorköper

Rostock, 26. Dezember 2016

Information Theoretical Analysis of Cooperative Orthogonal Relay Networks

Dissertation
zur
Erlangung des akademischen Grades
Doktor-Ingenieur (Dr.-Ing.)
der Fakultät für Informatik und Elektrotechnik
der Universität Rostock

vorgelegt von
Sebastian Vorköper, geboren am 30. Januar 1980 in Stralsund

Rostock, den 26. Dezember 2016

urn:nbn:de:gbv:28-diss2017-0006-5

Gutachter:

Prof. Dr.-Ing. habil. Volker Kühn
Universität Rostock, Institut für Nachrichtentechnik

Prof. Dr.-Ing. habil. Tobias Weber
Universität Rostock, Institut für Nachrichtentechnik

Prof. Dr.-Ing. Gerhard Bauch
TU Hamburg-Harburg, Institut für Nachrichtentechnik

Tag der Einreichung: 18.12.2015
Tag der Verteidigung: 04.07.2016

Kurzfassung

Mit zunehmender Anzahl an mobilen Endgeräten, wie z.B. Smartphones oder Tablets, erhöht sich auch die zu übertragende und zu verarbeitende Datenrate in einem Mobilfunknetz. Weiterhin sollte ausreichend gewährleistet sein, dass die Benutzer immer und überall Zugang zum Netz des Mobilfunkanbieters erhalten. Datenrate, Durchsatz, Zugang und Quality of Service stellen demnach die zu erfüllenden Herausforderungen eines Mobilfunkanbieters dar.

Mehrantennensysteme können diese Aspekte mit zunehmender Miniaturiersierung der Endgeräte nur noch bedingt erfüllen. Zum einen bieten die Geräte nicht mehr genügend Platz um eine deutlich anwachsende Zahl von Antennen aufzunehmen. Zum anderen verringert sich die räumliche Distanz zwischen den Antennen, wodurch die Ausbreitungspfade stark miteinander korreliert sind und somit diverstitätssichernde Maßnahmen nicht voll ausgeschöpft werden können.

Um diese Probleme besser in den Griff zu bekommen, wird seit einiger Zeit an der Verwendung von sogenannten Relay-Knoten geforscht. Diese Relays können kooperativ, d.h. sich gegenseitig unterstützend, oder nicht-kooperativ agieren. Nicht-kooperative Relay-Knoten, sogenannte Repeater, erhöhen lediglich die Quality of Service bzw. den Abdeckungsbereich eines Mobilfunknetzes. Sie sind bereits in einigen Systemen verfügbar bzw. durchlaufen den Standardisierungsprozess (z.B. LTE-Advanced).

Um dennoch Diversitätsvorteile zu erzielen, bieten sich kooperativer Relay-Knoten an. Diese können sich zu einem virtuellen Antennenarray zusammenschließen und agieren nahezu identisch wie ein Mehrantennensystem.

In der vorliegenden Arbeit werden kooperative orthogonale Relay-Netze untersucht. Dazu werden die Eigenschaften bestimmter Relay-Protokolle wiedergegeben und miteinander verglichen. Weiterhin wird ein Framework auf Basis der erweiterten Informationsverkettung vorgestellt und herausgearbeitet, um praktische (ggf. fehlerbehaftete) orthogonale Relay-Netze semi-analytisch vorhersagen zu können.

Im Vergleich zu der sonst üblichen Monte-Carlo Simulation liefert dieses Framework in einem Bruchteil der Rechenzeit zuverlässige Ergebnisse bezüglich der Ende-zu-Ende Transinformation eines solchen Netzes.

Ein weiteres Augenmerk liegt auf der Leistungsfähigkeit von Relay-Netzen mit fehleranfälligen Relay-Knoten, d.h. mit Relay-Knoten die das empfangene Signal möglicherweise nicht fehlerfrei decodieren können. Hierbei wird besonders der Ansatz der Soft-ReCodierung verfolgt, um nach Möglichkeit die Information über mögliche Symbolfehler an die Destination weiterzureichen und dort mit der Information anderer Relay-Knoten zu einem fehlerfreien Ergebnis zu gelangen.

Für eine gute Ausnutzung der Soft-ReCodierung wird in dieser Arbeit eine verbesserte Berechnung der Kanal-Wahrscheinlichkeitsverhältnisse (channel likelihood-ratios) angegeben. Leider zeigt sich, dass der Soft-ReCodierungsansatz nicht in jedem Fall zu einer verbesserten Leistungsfähigkeit führt. Die Gründe hierfür werden ausführlich in dieser Arbeit diskutiert.

Abstract

With increasing number of mobile devices, such as smartphones or tablets, the amount of data which has to be transmitted and processed in a wireless network also increases. Furthermore, it should be sufficiently ensured that the user has always and everywhere access to the network of the mobile operator. Accordingly, data rate, throughput, access and Quality of Service represent the challenges which have to be conquered and achieved by a mobile service provider.

Multiple antenna systems are limited to meet these aspects with increasing miniaturization of the mobile devices. On the one hand, the devices do not offer enough space to accommodate more number of antennas. On the other hand the spatial distance between the antennas would have to be reduced with a growing number of antennas, such that the propagation paths are highly correlated and thus diversity achieving schemes can not be fully exploited.

To overcome these problems, the use of so-called relay nodes is in research focus for some time. These relays can operate cooperatively, i. e. are mutually supportive, or non-cooperatively. Non-cooperative relay nodes, so called repeaters, increase only the Quality of Service or the coverage area of a cellular network. They are already available in some systems, or are currently undergoing the standardization process (e.g. LTE-Advanced).

To achieve diversity advantages, cooperative relay nodes are an interesting option. They can join together in order to form a virtual antenna array and act almost the same as a multiple antenna system.

In the present work, cooperative orthogonal relay networks are investigated, including the properties and performance of certain relay protocols. Furthermore, a framework is presented and analyzed in depth, based on the extended information combining/concatenation technique, to semi-analytically predict practical (possibly non-error free decoding) orthogonal relay networks. Compared to the usual Monte

Carlo simulation, this framework provides reliably end-to-end mutual information of such a network in a fraction of computational time.

Another focus lies on the performance of relay networks with error-prone relay nodes, i.e. relay nodes which perhaps did not decode the received signal correctly. Here, in particular the approach of soft re-encoding is followed, in order to possibly forward the information of erroneous symbols towards the destination. At the destination, this information in addition to the information from other relay nodes maybe achieve error-free decoding.

For a better exploitation of the soft re-encoding, this dissertation contains an improved calculus of the channel log-likelihood ratios at the destination's decoder input. Unfortunately, the results show that contrary to popular belief soft re-encoding does not always lead to an improved performance. The reasons for this are discussed in detail in the presented work.

Acknowledgements

This dissertation represents the final statements of a long scientific journey, which I started as a Ph.D. student in June 2007 in the Institute of Communications Engineering at the University of Rostock. During the time I spent there, I have met so many interesting people and friends which I would like to express my special thanks to.

First I want to thank my principal supervisor Prof. Volker Kühn for many fruitful discussions and advises, not only in the scientific area, but also in normal life situations. Also, I want to thank him for proof-reading my doctoral dissertation and, thus, improving my rusty English. Furthermore, I thank Prof. Tobias Weber for being one of my co-referees. I always enjoyed discussions with him, especially in the field of software defined radios. My third referee Prof. Gerhard Bauch deserves many thanks as well, since it is not a matter of course to dedicate time reading a doctoral thesis.

Furthermore, I would like to thank my friends and colleagues for their (not always serious) discussions and support. Some of them (in no specific order) are Ralph Hänsel, Henryk Richter, Nico Palleit, Martin Garbe, Daniel Kern, Peter Bartmann, Stephan Lange, Hans-Dietrich Melzer, Aimal Khan, Frank Jeschke, Petra Westphal, Gundula König and Angelika Eggert.

Special thanks to my family, i.e. my wife, childrens, parents, brothers and sisters-in-law, which supported me throughout this long journey and which always strengthen me emotionally.

Rostock, December 2016

Contents

1 Introduction **1**

1.1 Motivation and historical notes 2
1.2 Problem statement 5
1.3 Structure of the dissertation 6

2 Information theory and error correction coding **9**

2.1 Information theory 10
2.1.1 Entropy 10
2.1.2 Mean mutual information 11
2.1.3 Channel capacity 12
2.2 Simplified communication system 16
2.3 Forward error correction coding 18
2.3.1 Convolutional codes 18
2.3.2 Code concatenation and iterative decoding schemes 24
2.3.3 Information processing characteristic 25

3 Relay networks and relaying protocols **33**

3.1 Orthogonal half-duplex relaying 34
3.2 System model 34
3.2.1 Max-flow min-cut theorem 35
3.3 Relaying protocols 38
3.3.1 Direct transmission 39
3.3.2 Non-cooperative multi-hop transmission 39
3.3.3 Amplify-and-forward 40
3.3.4 Decode-and-forward 42
3.3.5 Compress-and-forward 44
3.3.6 Some results on the different relaying protocols 45

4 Functions for information concatenation **55**

4.1 Introduction 56
4.2 The information combining technique 56
4.3 Serial information combining 57
4.3.1 Serial information combining for SMCs 59
4.3.2 Upper and lower bound for serial information combining . . 60

4.4 Parallel information combining 66
4.4.1 Parallel information combining for SMCs 66
4.4.2 Upper and lower bound for parallel information combining . 68

5 Information combining for relay networks utilizing ideal and good codes 71
5.1 Three-node orthogonal relay network 72
5.1.1 Simulation environment 72
5.1.2 Max-flow min-cut bounds 74
5.1.3 Upper and lower bound for amplify-and-forward 75
5.1.4 Upper and lower bound for decode-and-forward with repetition coding 78
5.1.5 Selection between AF and DF 80
5.1.6 Optimal combining at the destination with decoding errors at the relay 81

6 Information combining for relay networks utilizing practical codes 89
6.1 Three-node relay network 90
6.1.1 Source destination link and source relay link 93
6.1.2 Relay destination link 93
6.1.3 Performance of soft and hard re-encoding 103
6.1.4 IPC for soft and hard re-encoding 105
6.2 Decoding at the destination 108
6.2.1 Limiter function 108
6.2.2 Maximum ratio combining 112
6.2.3 Turbo decoding 116
6.2.4 Comparing the different decoding methods 116
6.3 Information combining for practical relay networks 120
6.3.1 Practical three-node relay network 120
6.3.2 Extended information processing characteristic 123
6.3.3 Practical four-node relay network 126

7 Conclusions and outlook 133
7.1 Conclusions 133
7.2 Outlook 134

Bibliography 137

Publications 147

A Appendix 149
A.1 Proof of the ideal coding scheme 149
A.2 Weak converse versus ideal coding scheme 151
A.3 Random variables, probabilities and probability functions 152
A.4 Some properties and distributions of random variables 153

A.4.1 Gaussian distribution, $\mathcal{N}(\mu_X, \sigma_X^2)$ . . . 153
A.4.2 Bernoulli distribution, $\mathcal{B}(\epsilon)$ . . . 154
A.4.3 Derived distributions, $Y = g(X)$ . . . 154
A.4.4 Sum of two random variables, $Z = X + Y$ . . . 155
A.4.5 Product of two random variables, $Z = X \cdot Y$ . . . 155
A.4.6 Linear function of a random variable, $Y = aX + b$ . . . 155
A.4.7 Derived distribution for $Y = \tanh(X/2)$ . . . 156
A.4.8 Derived distribution for $Y = |X|$ . . . 156
A.4.9 Modulo-2 sum density, $U_Z = U_X \oplus U_Y$ . . . 157
A.4.10 Boxplus and min-sum density . . . 157
A.5 Log-likelihood ratios and soft bits . . . 158

B Nomenclature **161**

C Notation **165**

D Theses **169**

1 Introduction

In this chapter the motivation for the presented work is given and in addition to the general historical development in the field of wireless communication, the reference is made to current problems and their possible solutions. Among these problems are the increasing demand for high data rates and permanent access. Both can be solved by introducing relays into the wireless network.

This will inevitably raise new questions like the performance increase or the prediction of the throughput by utilizing relays. Furthermore, questions like how to deal with channel impairments or which relaying strategy to use, must be solved.

1.1 Motivation and historical notes

At the end of the 20th century wireless telephones and wireless telephone systems, such as cellular CDMA or GSM cell sites emerged. During the same time the requirement to transmit data between the participants of such systems appeared and, thus, continuously increases the demand for higher and higher data rates. For example, according to the German *Federal Network Agency for Electricity, Gas, Telecommunications, Posts and Railway*, the demand for data rates increased from $\approx$ 100 million gigabyte in 2011 to $\approx$ 140 million gigabyte in 2012. The number of regular *universal mobile telecommunications system* (UMTS) and *long term evolution* (LTE) users/devices increased from $\approx$ 28.6 million in 2011 to $\approx$ 33.6 million in 2012.

Because of the increasing number of mobile devices, especially smartphones and tablet computers, the demand for better coverage, higher throughput and permanent access has to be met by the system providers. One solution to meet these requirements is to increase the transmit power. But, this will inevitably increase interference among mobile user equipments and decrease the battery life of these small devices.

Another possibility is to shorten the symbol duration to increase the data rate. However, this will increase the bandwidth. Since increasing the bandwidth demands for utilizing higher frequency, the pathloss will rapidly increase as well. Therefore, higher bandwidth would lead to higher throughput but less coverage and limited access. Furthermore, frequencies above $\approx$ 3 GHz are not suitable for mobile communication in scattered urban environment with the demand for a large area to cover. Though, increasing the transmit power would also increase coverage and/or results in a better signal-to-noise ratio, still the interference from neighboring mobile stations must be treated carefully [JAW07] or [ASW10]. Otherwise, the network remains interference-limited. All of the above mentioned problems make frequency bands below $\approx$ 3 GHz a limited and, therefore, expensive resource.

Another drawback of short symbol duration is its increased degree of intersymbol interference, coming from reflections, diffraction and scattering caused by multipath signal propagation. Utilizing orthogonal frequency division multiplexing [Wei09] resolves problems of intersymbol interference. Still the requirement for high bandwidth remains.

To overcome the limitations of the demand for larger bandwidth, multiple input multiple output schemes have been developed [FG98], [AT99] or [PF01] and are already standardized in IEEE802.11n, 3GPP LTE and *worldwide interoperability for*

microwave access (WiMAX). The idea is to utilize more than one antenna at the transmitter and receiver, such that the spatial dimension of the communication link is exploited. This allows the implementation of space-time codes [TSC98] or [GSS+03] and other diversity achieving schemes which will increase the reliability of the communication or improve the coverage. Among these schemes are antenna diversity [VA87],[Ala98] or [BHS+02] and beamforming [RFLT98]. Furthermore, utilizing multiple input multiple output makes it possible to multiplex more than one data stream onto the channel to achieve higher data rates. This is denoted as spatial multiplexing and is utilized for example in [Fos96] or [WFGV98].

Besides higher complexity as compared to single antenna systems, MIMO is not really suitable for small devices. Especially, the antenna placement is very difficult. Furthermore, due to the limited space between two or more antennas in small devices, spatial diversity maybe not achieved as the propagation paths are highly correlated. Utilizing higher frequencies will make this problem become less severe, but lead to the discussed problems from above.

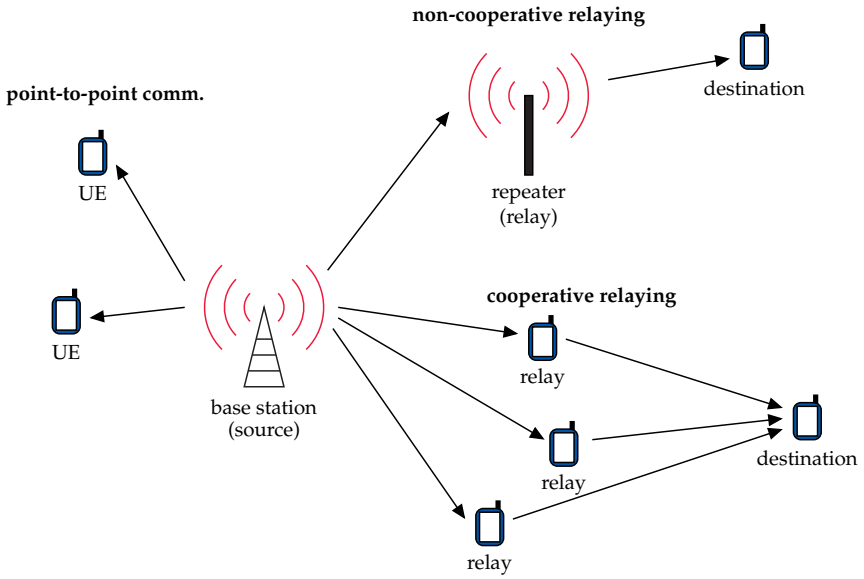


Fig. 1.1.: Typical base-station setup with upcoming cooperative relays in the downlink.

Utilizing intermediate nodes between the transmitter and the receiver can resolve most of the above problems. A so called relay network is depicted in **Fig. 1.1**[1]. Here, among the point-to-point system on the left side, two relay schemes (cooperative and non-cooperative) are depicted on the right side.

The point-to-point system is also called direct scheme. Here, the base station directly communicates with the *user equipment* (UE, mobile device). It is widely implemented in nowadays wireless communication systems for example UMTS,

[1] Although **Fig. 1.1** only shows the downlink, all given explanations are also valid for the uplink.

WiMAX, LTE and *wireless local area network* (WLAN). Some of these systems or their early implementations do not have standardized MIMO.

In the non-cooperative relay scheme, the base station (source) transmits its information to the repeater (relay). During a second time slot, the repeater will forward the amplified or processed information towards the UE (destination). The type of further processing depends on the implemented complexity of the repeater. If the repeater only amplifies the received signal prior to transmission, it works on *open systems interconnection* (OSI) layer 1 [ITN10] and is usually denoted as non-regenerative. The implementation and deployment for non-regenerative relays is simple and inexpensive, as there is only minimal impact on already standardized systems. The drawback relies in the amplification and transmission of the repeater's receiving noise towards the user equipment. To overcome this problem, layer 2 and layer 3 repeaters implement (de-) modulation and (de-) coding schemes, which eventually eliminates receiving noise. These repeaters are called regenerative relays, as they are regenerating the information prior re-transmission. Non-cooperative relay schemes extend coverage and, thus, provide better *Quality of Service* (QoS) and greater throughput especially for cell edge users. Mostly, repeaters are fixed stations, placed by the provider in appropriate locations, with no wired connection to the backhaul of the system provider. The drawback of these non-cooperative repeaters resides in no additional diversity since the degree of freedom is not changed. Furthermore, the transmission of information happens orthogonal in two time slots, which eventually decreases the throughput. Non-cooperative relay schemes are already standardized or are in the standardization process.

The next step in mobile communication is to utilize one or more repeating nodes as relays to establish a cooperative relay network. In such a network, the source (base station) transmits its information to one or more relay nodes[2]. Afterwards, during subsequent time slots, the relays cooperatively re-transmit the received and processed information towards the destination. Here, the cooperation itself depends on the repeater type, i.e. regenerative or non-regenerative.

The limitations with repeaters relies in a connection to a power supply, large antenna equipment and/or high maintenance requirements. Most notably is the permanent placement of a repeater at a fixed location, which can be cumbersome in changing network requirements. A solution to these problems is given by utilizing user equipments as relays, which is depicted in the lower right part of **Fig. 1.1**. Here, individual mobile phones can become relay nodes and will cooperatively forward information. Since users tend to move, a higher synchronization effort has to be taken into account and the routing/cooperation must be adapted permanently.

[2]If a direct connection is present, one relay is sufficient to form a cooperative relay network. Otherwise at least two relays are necessary.

Still, these drawbacks will be outreached by a better Quality of Service (reliability, throughput, coverage and so on) for individual users and less fixed costs for the service provider (energy, location, e.g.).

Since the cooperative relay network can form a virtual antenna array [DS03] or [LTW04], all the benefits from MIMO schemes are potentially present, besides the non-cooperative advantages. This includes beamforming gains, space-time coding diversity and higher data rates through multiplexing. These benefits of cooperative relay networks usually exceed the disadvantages, like implementation costs and coordination overhead.

Historically, Van der Meulen [Meu71] introduced cooperative relay networks by discovering upper and lower bounds on the capacity of the three-terminal relay channel. This channel consists of one source, one relay and one destination, with communication links connecting all three nodes. Although, the capacity of the general relay channel is still unknown, Cover and El Gamal [CG79] significantly improved capacity bounds from Van der Meulen. They also investigated different relaying strategies from the mutual information point of view, giving upper bounds on the capacity for *Decode-and-Forward* (DF) and *Compress-and-Forward* (CF). In [Lan02] Laneman introduced the *Amplify-and-Forward* (AF) relay protocol. Furthermore, together with Tse and Wornell, they studied cooperative diversity protocols and their outage behaviors in fading environments [LTW04]. Kramer, Gastpar and Gupta studied coding strategies that exploit node cooperation [KGG05] in case of DF and CF, delivering novel insights on information theoretic coding. Another issue regarding the three-node relay channel, was addressed by Høst Madsen and Zhang [HMZ05]. They studied upper and lower bounds on the outage and ergodic capacity in fading environment, taking into account practical constraints and optimizing power allocation. In addition to the above, there are many other scientists who work in the field of cooperative relay networks, making it impossible to mention all of them.

1.2 Problem statement

The aim of this dissertation is to compare orthogonal half-duplex relay networks with different relaying protocols. Here, special interest is given to practical relay systems with error-prone nodes, including different re-encoding strategies at these nodes, as well as different combining and decoding strategies at the destination. Among others, the major problem is the computational complexity and, therewith, the simulation time of such networks.

To tackle this problem, a new approach is introduced. Here, the relay network with its different practical codes and decoding and combining methods, will be completely described by the mutual information. Thus, the network itself is decomposed into its individual links, including encoder and decoder. These individual links can be characterized by their information processing characteristic, which simply represents the mutual information of the end-to-end channel, from encoder input to decoder output, with respect to the underlying communication channel. Afterwards, the information processing characteristics are concatenated with the help of the information combining technique in order to retain the mutual information of the complete relay network. This will eventually save simulation time, as the information processing characteristics can be evaluated and stored into tables beforehand. In order to cope with soft-input soft-output re-encoding at the relay nodes and turbo-decoding at the destination, the information processing characteristic is extended by an extra a-priori channel.

Regarding soft-input soft-output re-encoding at the relay, in this dissertation a solution is given to calculate the log-likelihood ratios at the destination's decoder input. This solution depends on the approximation of the so called boxplus-operation by the min-sum operation. Since the soft re-encoder can be regarded as a filter with boxplus-operations, the output distribution is approximated via the min-sum distribution. The latter one can be calculated semi-analytically instead of exhaustive Monte-Carlo simulations.

Furthermore, the information combining technique is also utilized in order to predict the achievable rate for a theoretical three-node relay system, with ideal codes. Here, the aim is to find different transmitting and combining methods in order to close the gap between the rates of known relaying protocols and the upper bound given by the max-flow min-cut theorem. To do so, the original information combining technique is extended by serial- and parallel concatenation/combining of mutual information in cases of symmetric memoryless channels.

1.3 Structure of the dissertation

Among this introduction together with the conclusion and the appendix, the dissertation contains 5 additional chapters with the following contents:

Chapter 2: Chapter 2, contains a small introduction to information theory, channel capacity, the simplified communication system and forward error correction coding. After introducing practical codes and its decoding methods, the

information processing characteristic for ideal, good and practical codes are recapitulated. All three are exhaustively utilized throughout this dissertation.

Chapter 3: Chapter 3 recapitulates the orthogonal half-duplex relay channel, including different relaying protocols like Amplify-and-Forward and Decode-and-Forward. In addition to the analysis of the individual relaying protocols and their achievable rates, this chapter provides the basis for the upcoming Chapters. Furthermore, this chapter also contains the max-flow min-cut upper bound for the achievable rate of the orthogonal half-duplex relay channel.

Chapter 4: Chapter 4 familiarizes with the information combining technique for serial and parallel concatenated channels. The information combining technique is extended to additionally work with symmetric memoryless channels. Furthermore, lower and upper bounds for the serial and parallel concatenation of mutual information in cases of binary symmetric channels and binary erasure channels are recapitulated or derived.

Chapter 5: Chapter 5 utilizes the information combining technique in order to predict the achievable rates for the different relaying protocols. Among the given examples, this chapter also tries to reach the max-flow min-cut bound by combining the information combining technique with the information processing characteristic of ideal codes in a semi-analytically approach.

Chapter 6: Chapter 6 is about relaying networks with practical codes, which maybe results in error-prone receiving nodes. Even with imperfect decoding results at the relays, a perfect decoding result could be achieved at the destination. Therefore, special attention is paid to the different coding and decoding strategies. This includes the hard-input hard-output and soft-input soft-output re-encoding at the relay nodes, as well as the different combining and decoding strategies at the destination.

2 Information theory and error correction coding

This chapter contains three major sections, introducing a few used notations and concepts in information-, coding- and communication theory.

The first section reviews parts of the information theory by Shannon, including the definition of entropy, mutual information and channel capacity for discrete memoryless sources and channels. Here, the focus relies on symmetric memoryless channels with binary or Gaussian distributed inputs, crossover and erasure probabilities or additive Gaussian noise.

Subsequently, the second section introduces the simplified communication model of a point-to-point system with binary input and superimposed Gaussian noise, i.e. the BIAWGN channel.

The last section introduces forward error correction channel coding, with special attention paid on convolutional codes and iterative decoding schemes utilizing soft-information. Here, the information processing characteristic which represents input/output behavior of ideal or practical codes, is introduced. It plays a major role throughout this dissertation, as it will be utilized in the upcoming chapters to predict the mutual information of practical and idealized relay networks.

2.1 Information theory

This section introduces entropy, mutual information and capacity for *symmetric memoryless channel*s (SMC), i.e. measures of either uncertainty, mutual dependency or rate of information. Especially the *binary input symmetric memoryless channel* (BISMC) and the *Gaussian input symmetric memoryless channel* (GISMC) are of great interest, as they are utilized exhaustively throughout this dissertation.

2.1.1 Entropy

The Shannon entropy of a discrete *random variable* (rv) X, with given *probability mass function* (pmf) $p_X(x) = \Pr(X = x)$, is defined by a non-negative function as [CT06]

$$H(X) = \sum_{x \in \mathbb{X}} p_X(x) \log_2 \frac{1}{p_X(x)} \text{ [bit]} \ , \tag{2.1}$$

providing the uncertainty in X, which takes on values $x \in \mathbb{X}$. It is measured in bits[1] and allows the quantification of the uncertainty of the outcome of an event.

The joint entropy $H(X,Y)$ of two discrete random variables X and Y with their joint pmf $p_{X,Y}(x,y) = \Pr(X = x, Y = y)$ is given by

$$H(X,Y) = \sum_{x \in \mathbb{X}} \sum_{y \in \mathbb{Y}} p_{X,Y}(x,y) \log_2 \frac{1}{p_{X,Y}(x,y)} \ , \tag{2.2}$$

whereas the conditional entropy (or equivocation) of these random variables is defined as

$$H(X|Y) = \sum_{x \in \mathbb{X}} \sum_{y \in \mathbb{Y}} p_{X,Y}(x,y) \log_2 \frac{1}{p_{X|Y}(x|y)} \ . \tag{2.3}$$

The joint entropy measures the uncertainty among a set of random variables, whereas the conditional entropy delivers the uncertainty of a random variable given another rv.

By inserting $p_{X,Y}(x,y) = p_X(x) p_{Y|X}(y|x)$ into (2.2), the joint entropy follows the chain rule of entropy [CT06, p. 17,22], i.e.

$$H(X,Y) = H(X) + H(Y|X) \ . \tag{2.4}$$

[1] The measurement unit can also be *nat* or *dit*, depending on the base of the logarithm.

If and only if (iff) the random variables X and Y are statistically independent, i.e. their joint pmf is given by $p_{X,Y}(x,y) = p_X(x)p_Y(y)$, then their individual entropies are additive and their joint entropy is calculated according to

$$H(X,Y) = H(X) + H(Y) \ . \tag{2.5}$$

2.1.2 Mean mutual information

The *mean mutual information* via the expectation over the joint probability mass function $p_{X,Y}(x,y)$ is defined [CT06, p. 20] by

$$\begin{aligned} I(X;Y) &= \sum_{x\in\mathbb{X}} \sum_{y\in\mathbb{Y}} p_{X,Y}(x,y) \log_2 \frac{p_{X,Y}(x,y)}{p_X(x)p_Y(y)} \\ &= H(X) - H(X|Y) \ . \end{aligned} \tag{2.6}$$

The mean mutual information in (2.6) of two random variables X and Y is always non-negative and only zero in case of independence of X and Y. Please note, that the *mean mutual information* is simply referred to as *mutual information* (MI) throughout this dissertation.

Other ways to calculate the MI are given by utilizing the marginal entropies and the conditional or joint entropies:

$$I(X;Y) = H(Y) - H(Y|X) \tag{2.7}$$

$$= H(X) + H(Y) - H(X,Y) \tag{2.8}$$

$$= H(X,Y) - H(X|Y) - H(Y|X) \ . \tag{2.9}$$

Please keep in mind, that this holds only for discrete random variables. For continuous random variables and corresponding differential entropies, the formulas almost look the same, but care has to be taken since some properties don't hold as in the discrete case. Still, the *differential mutual information* between two continuous random variables is defined equivalently to (2.6)

$$I(X;Y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} p_{X,Y}(x,y) \log_2 \frac{p_{X,Y}(x,y)}{p_X(x)p_Y(y)} \, \mathrm{d}x \, \mathrm{d}y \ , \tag{2.10}$$

replacing sum's by integral's. As can be seen from (2.10), continuous and discrete random variables and their corresponding probability density functions and probability mass functions will be denoted with the same variable names. Throughout

this dissertation, their distinction is either given by definition or obvious from the formulas in which they are utilized.

2.1.3 Channel capacity

Given the formula for the mutual information, one can calculate the channel capacity [CT06, p. 183]

$$C = \max_{p_X(x)} I(X;Y) \ , \tag{2.11}$$

by adopting the channel input and therewith the corresponding distribution $p_X(x)$, such that the MI is maximized. In case of discrete time channel in- and outputs, the channel capacity's measurement unit is *bits/symbol* or equivalently *bits/channel use* and it denotes the maximal amount of information which can be reliable transmitted over a given communication channel.

Channel capacities for different channel types

The following subsections contain different channel types, which are utilized throughout the dissertation. To be more precise, the focus relies on symmetric memoryless channels with discrete inputs and discrete or continuous outputs.

Binary erasure channel

The first channel model to start with is the *binary erasure channel* (BEC) as shown in **Fig. 2.1**. It has a binary input X, $x \in \{+1, -1\}$ with probabilities of occurrence

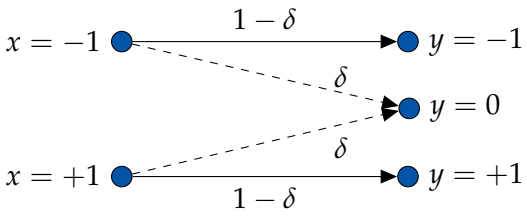


Fig. 2.1.: Binary erasure channel

$\Pr(X = +1) = \Pr(X = -1) = 1/2$ and a ternary output Y, $y \in \{+1, 0, -1\}$. Due to the uniform input distribution, maximum entropy of $H(X) = 1$ is achieved. For the output, the probability $\Pr(Y = 0)$ denotes the event of an erasure, i.e. the

probability of an information loss. The capacity for the binary erasure channel is calculated by [CT06, p. 188]

$$C_{\text{BEC}} = 1 - \delta \ , \tag{2.12}$$

with erasure probability

$$\delta = \Pr(Y = 0|X = +1) = \Pr(Y = 0|X = -1) \ . \tag{2.13}$$

Binary symmetric channel

In order to achieve maximum entropy of $H(X) = 1$, the *binary symmetric channel* (BSC), as depicted in **Fig. 2.2**, is described by the uniform distribution[2] of the channel input with $\Pr(X = +1) = \Pr(X = -1) = 1/2$ and the crossover (error) probability ϵ via

$$\begin{aligned} \epsilon &= \Pr(Y = -1|X = +1) = \Pr(Y = +1|X = -1) \\ 1 - \epsilon &= \Pr(Y = +1|X = +1) = \Pr(Y = -11|X = -1) \ , \end{aligned} \tag{2.14}$$

taken from the conditional probability. With (2.11), the channel capacity for the BSC

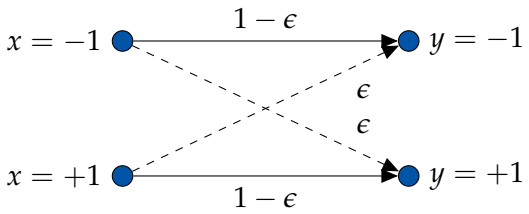


Fig. 2.2.: Binary symmetric channel

is given by

$$C_{\text{BSC}} = \max_{p_X(x)} \{H(Y) - H(Y|X)\} = 1 - H_b(\epsilon) \ , \tag{2.15}$$

where $H_b(\epsilon) = -\epsilon \log_2 \epsilon - (1 - \epsilon) \log_2(1 - \epsilon)$ denotes the binary entropy function.

[2] As for the binary erasure channel, the uniform input distribution is capacity achieving in case of the binary symmetric channel and other symmetric channels.

Binary input additive white Gaussian noise channel

The *binary input additive white Gaussian noise* (BIAWGN) channel, as depicted in **Fig. 2.3**, has a discrete input X, an additional noise term N and output Y. The

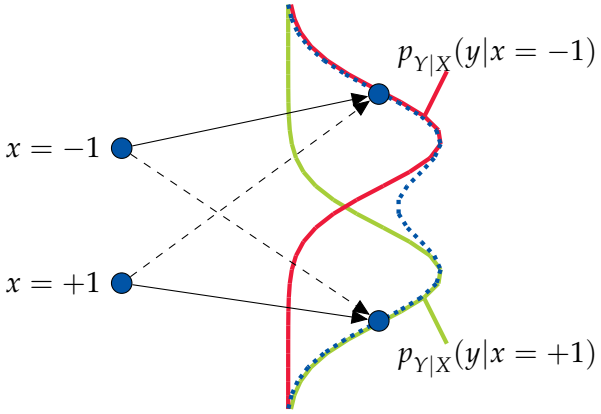


Fig. 2.3.: Binary (symmetric) input additive white Gaussian noise channel and the indication for the decomposed representation by BSCs

channel output is calculated via

$$Y = X + N \ , \tag{2.16}$$

where the input X, $x \in \{-1, +1\}$ is binary, the additive white Gaussian noise term $N \sim \mathcal{N}(0, \sigma_N^2)$ is Gaussian distributed, with zero-mean and variance σ_N^2 and the channel output Y, $y \in \mathbb{R}$ is real valued and bimodal Gaussian distributed.

Given the channel transition probabilities

$$p_{Y|X}(y|x=\pm 1) = \frac{1}{\sqrt{2\pi\sigma_N^2}} e^{-\frac{(y\mp 1)^2}{2\sigma_N^2}} \ , \tag{2.17}$$

the capacity of the BIAWGN channel is calculated according to (2.10) and (2.11)

$$C_{\text{BIAWGN}} = \max_{\Pr(X)} \int_{-\infty}^{+\infty} \sum_{i=1}^{2} \Pr(X = x_i) \cdot p_{Y|X}(y|x_i) \cdot \log_2 \frac{p_{Y|X}(y|x_i)}{p_Y(y)} \, \mathrm{d}y \ . \tag{2.18}$$

Gaussian input additive white Gaussian noise channel

A real valued channel is called as a *Gaussian input additive white Gaussian noise* (GIAWGN) channel, if the channel input $X \sim \mathcal{N}(0, \sigma_X^2)$ and the channel noise $N \sim \mathcal{N}(0, \sigma_N^2)$ are Gaussian distributed with zero-men and variance σ_X^2 and σ_N^2, respectively. Since the random variables X and N are independent, the channel output $Y = X + N$ is also Gaussian distributed, with $Y \sim \mathcal{N}(0, \sigma_X^2 + \sigma_N^2)$.

According to (2.10) and (2.11), the channel capacity for the GIAWGN channel can be calculated by

$$C_{\text{GIAWGN}} = \max_{p_X(x)} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} p_{Y|X}(y|x) \cdot \log_2 \frac{p_{Y|X}(y|x)}{p_X(x) \cdot p_Y(y)} \, \mathrm{d}x \, \mathrm{d}y$$
$$= \frac{1}{2} \log_2 (1 + \gamma) \quad , \tag{2.19}$$

maximizing the mutual information over the input distribution $p_X(x)$. In (2.19), γ denotes the *signal-to-noise ratio* (SNR)

$$\gamma = \frac{\sigma_X^2}{\sigma_N^2} \quad , \tag{2.20}$$

which is the ratio between the signal power and the noise power. As both are assumed to have zero-mean, it is the ratio between the variance of the channel input and the white Gaussian noise.

Some numerical results for the different channel models from the preceding subsections are depicted in **Fig. 2.4**. It can be seen, that in the lower SNR region, both, the

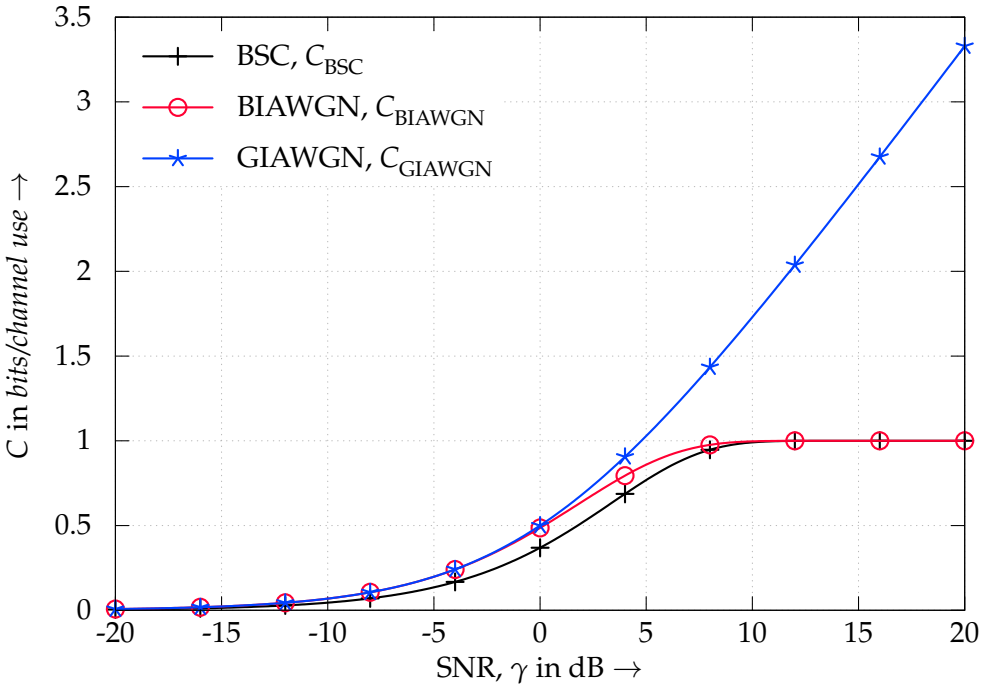


Fig. 2.4.: Channel capacities for time discrete channels with real valued channel in- and outputs

GIAWGN and the BIAWGN channel, almost perform the same. Therefore, in the region of a higher noise variance, a binary input seems to be sufficient to closely

reach the upper bound of the channel capacity, which is denoted by the GIAWGN channel. Above $\gamma = 0$dB, the curves diverge and the capacity for the BIAWGN channel saturates, due to the binary input alphabet.

The same saturation holds for the binary symmetric channel. Additionally, the capacity curve for the BSC proceeds below the BIAWGN curve. This follows from the hard decision, and, thereby, the loss of information over the channel output reliability.

2.2 Simplified communication system

The simplified wireless communication system[3], as depicted in **Fig. 2.5**, consists of a transmitter T, a receiver R and a channel in between. Please note, that in order

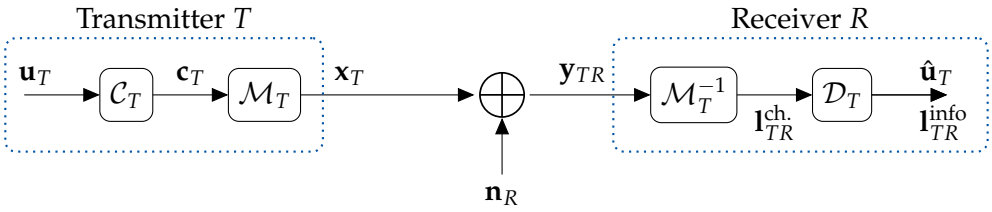


Fig. 2.5.: Communication system with additive white Gaussian noise

to denote the source $\mathcal{S}$, the relay $\mathcal{R}$ and the destination $\mathcal{D}$ the indices T, R will be substituted by $T \in \{\mathcal{S}, \mathcal{R}\}$ and $R \in \{\mathcal{R}, \mathcal{D}\}$ in the upcoming chapters. For now, the general description will utilize T and R.

The transmitter encodes the binary information sequence $\mathbf{u}_T = [u_{T,k}]_{K\times 1}$, $u_{T,k} \in \{0,1\}$ of length K and $k \in \{1,\ldots,K\}$ with the encoder $\mathcal{C}_T$. The encoder outputs the corresponding binary code sequence $\mathbf{c}_T = [c_{T,m}]_{M\times 1}$, $c_{T,m} \in \{0,1\}$ of length M. Subsequently, the encoded bits are modulated by $\mathcal{M}_T$ to generate the transmit signal $\mathbf{x}_T = [x_{T,m}]_{M\times 1}$, $x_{T,m} \in \{\mathbb{X}_T\}$. Throughout the whole dissertation, only real valued channel in- and outputs and only real valued channel impairments are of interest. Thus, $\mathbb{X}_T \in \mathbb{R}$ and the length of the transmit sequence is also M. Furthermore, except in some rare cases only *binary phase shift keying* (BPSK) is considered. The mapping for BPSK is performed via $\{0,1\} \rightarrow \{+1,-1\}$.

After the transmitter emits $\mathbf{x}_T$, the transmit signal is influenced by the noise $\mathbf{n}_R = [n_{R,m}]_{M\times 1}$, $n_{R,m} \in \{\mathbb{R}\}$. Thus, in case of an *additive white Gaussian noise* (AWGN)

[3]The simplified communication channel lacks some parts, like the source coder and others of a complete communication system.

channel between the transmitter T and the receiver R the channel output $\mathbf{y}_{TR}$ is given by

$$\mathbf{y}_{TR} = \mathbf{x}_T + \mathbf{n}_R \ , \tag{2.21}$$

where $N_R \sim \mathcal{N}(0, \sigma_N^2 / a_{TR}^2)$ denotes the real valued Gaussian distributed noise with zero-mean and variance σ_N^2 / a_{TR}^2. Here, instead of attenuating the channel input $\mathbf{x}_T$, the channel noise is amplified with the inverse of the path-loss attenuation factor $a_{TR} = \sqrt{d_{TR}^{-\alpha}}$, where d_{TR} denotes the distance between the two stations T, R and α is the path-loss exponent. The latter one ranges from 2, i.e. free space propagation, to 5, i.e. sub-urban propagation.

Hence, the receiver SNR is given by

$$\gamma_{TR} = \frac{a_{TR}^2 \cdot \sigma_{X_T}^2}{\sigma_N^2} \ , \tag{2.22}$$

with average transmit symbol power $\sigma_{X_T}^2$. Please note that throughout this dissertation, the average transmit symbol power will be normalized, i.e. $\sigma_{X_T}^2 = 1$ holds and, hence, all transmitting nodes will not exceed the given unit mean power constraint.

The receiver will demodulate the received signal $\mathbf{y}_{TR}$ with the symbol demodulator $\mathcal{M}_T^{-1}$ in order to generate *log-likelihood ratio*s (LLR) $\mathbf{l}_{TR}^{\text{ch.}} = \left[l_{TR,m}^{\text{ch.}} \right]_{M \times 1}$, $l_{TR,m}^{\text{ch.}} \in \{\mathbb{R}\}$ of length M.

The a-posteriori LLR value of the binary random variable C_T and its realization $c_{T,m} \in \{0, 1\}$, given the channel output sequence $\mathbf{y}_{TR}$, is defined by[4]

$$l_{TR,m}^{\text{ch.}} = \log \frac{p_{C_T|Y_{TR}}(c_{T,m} = 0 | y_{TR})}{p_{C_T|Y_{TR}}(c_{T,m} = 1 | y_{TR})} \ , \tag{2.23}$$

where the conditional pmf $p_{C_T|Y_{TR}}(c_{T,m}|y_{TR})$ denotes the probability that C_T takes on value $c_{T,m}$ given y_{TR}.

Afterwards, these LLR values are passed to the channel decoder $\mathcal{D}_T$ in order to retain a hard-decision $\hat{\mathbf{u}}_T = [\hat{u}_{T,k}]_{K \times 1}$, $\hat{u}_{T,k} \in \{0, 1\}$ or log-likelihood ratios $\mathbf{l}_{TR}^{\text{info}} = \left[l_{TR,k}^{\text{info}} \right]_{K \times 1}$, $l_{TR,k}^{\text{info}} \in \{\mathbb{R}\}$ of the estimated transmitter information.

Please note, that the description of (2.21) is a little unusual but simplifies the mathematical description substantially in the upcoming chapters. In most publications

[4] Detailed results for the LLR calculation in case of specific channel models is given in the upcoming chapters.

one would find $\mathbf{y}_{TR} = a_{TR} \cdot \mathbf{x}_T + \mathbf{n}_R$, with $N_R \sim \mathcal{N}(0, \sigma_N^2)$. In comparison to (2.21), both have the same SNR which is calculated for the AWGN channel by (2.22). Furthermore, σ_N^2 will be equal at all receiving nodes, and, therefore the SNR is dominated by the path-loss attenuation factor a_{TR} for each individual link.

Depending on the channel input, the channel capacity C_{TR} for (2.22) is calculated either via (2.11) with (2.6) or (2.10).

2.3 Forward error correction coding

The capacities, which were given in the previous Subsection 2.1.3 for the different channel models, define the upper bounds for codes with rates $R_c \leq C$. As stated in Shannon's famous channel coding theorem [Sha48], a reliable communication through the specific channel is possible at these rates with a vanishing small error probability.

Hence, in principle codes with code rates $R_c \leq C$ exists, which can communicate over a given channel and under the presence of noise with vanishing small error probability [Sha48]. These codes will be called *capacity achieving codes*. Among these codes are *ideal codes* and *good codes*.

The problem with capacity achieving codes relies in their requirement of infinite long code words, which cannot be encoded and decoded in a reasonable amount of time [Mac02, p. 183]. Thus, under the assumption of existing capacity achieving codes, they are unacceptable for practical considerations. Here, *practical codes*, i.e. codes which are not capacity achieving but decodeable in a time that scales as a function of the blocklength (preferably linearly) are in favor for wireless communication systems.

Among these codes are *convolutional codes* (CC) introduced by Peter Elias in 1955, Robert G. Gallager's *low-density parity-check codes* (LDPC) [Gal63] [MN97] and *turbo codes* (TC) [BGT93] by Berrou, Glavieux and Thitimajshima in 1993. In this dissertation, only CCs and their *parallel* concatenation, forming a parallel TC are considered.

2.3.1 Convolutional codes

Although Peter Elias introduced convolutional codes in 1955, they where not widely utilized until 1967, when Viterbi published an algorithm [Vit67] which was capable of efficiently decoding these types of codes. This algorithm became the well known

Viterbi algorithm. It is based on the maximum-likelihood sequence estimator (MLSE) principle, which looks for the most likely sequence out of all possible transmitted ones, from a noisy reception. Due to the structure of convolutional codes and their corresponding representation by a trellis, instead of comparing all possible sequences, the Viterbi-algorithm can remove less probable sequences per node at each time instance. Thus, it significantly reduces the computational complexity for binary sequences from 2^N to $N \cdot 2^{m+1}$, where N denotes the number of observed symbols and m the memory-length of the CC.

In contrast to the Viterbi-algorithm, which minimizes the sequence error rate, the BCJR algorithm minimizes the symbol error rate. The BCJR algorithm was developed and published in 1974 by Bahl, Coke, Jelinek and Raviv [BCJR74]. It also performs the decoding with the help of the trellis diagram, but the underlying principle is the *maximum-a-posteriori* (MAP) estimation. Here, the information sequence is chosen, which maximizes the a-posteriori probability conditioned on the received sequence. A more detailed description of the BCJR is given in Section 2.3.1.

Nowadays, convolutional codes are widely used in all areas of information transmission and storage. They are easy to implement, do not rely on constant block lengths and there exists different decoding methods, like the mentioned maximum-likelihood or the maximum-a-posteriori decoder.

Convolutional encoder

The general encoder without feedback for a convolutional code can be regarded as a time discrete *linear time invariant* system (LTI) with I input and J output sequences as depicted in **Fig. 2.6** and described in [Bos98, p. 228 ff.]. Here, the information sequence[5] $\mathbf{u} = \mathbf{u}_0, \mathbf{u}_1, \ldots = \left[\left[u_0^{(1)}, u_0^{(2)} \ldots u_0^{(I)}\right], \left[u_1^{(1)}, u_1^{(2)} \ldots u_1^{(I)}\right], \ldots\right]$ is partitioned into blocks $\mathbf{u}_k = \left[u_k^{(1)} \ldots u_k^{(i)} \ldots u_k^{(I)}\right]$, $u_k^{(i)} \in \{0,1\}$ with I bits each and afterwards encoded via the linear function $\mathbf{c}_k = f(\mathbf{u}_{k-m}, \mathbf{u}_{k-m+1}, \ldots, \mathbf{u}_k)$ into the code sequence $\mathbf{c} = \mathbf{c}_0, \mathbf{c}_1, \ldots = \left[\left[c_0^{(1)}, c_0^{(2)} \ldots c_0^{(J)}\right], \left[c_1^{(1)}, c_1^{(2)} \ldots c_1^{(J)}\right], \ldots\right]$ with code blocks $\mathbf{c}_k = \left[c_k^{(1)} \ldots c_k^{(j)} \ldots c_k^{(J)}\right]$, $c_k^{(j)} \in \{0,1\}$ and J bits each. Technically, the information blocks are shifted into registers of memory m and the encoding is per-

[5] To emphasize the in- and output of the encoder, a slightly change in the mathemetical notation of arrays is performed in this Subsection.

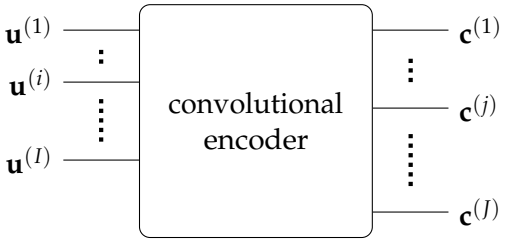


Fig. 2.6.: Convolutional encoder as a LTI system

formed by the convolution of the i-th information sequence with the corresponding generator sequence $\mathbf{g}_i^{(j)}$

$$\mathbf{c}^{(j)} = \sum_{i=1}^{I} \mathbf{u}^{(i)} \circledast \mathbf{g}_i^{(j)} \ , \tag{2.24}$$

in order to retain the j-th code sequence from a modulo-2 convolution, i.e. convolution in the Galois field $\mathbb{F}^2$. In (2.24), the generator sequence $\mathbf{g}_i^{(j)}$ for the j-th input and the i-th output is often given in the octal notation and all generator sequences are represented in a matrix notation, called the generator matrix. The number of how man times an input bit affects the encoder output during the encoding process is specified by the constraint length $L_c = m + 1$, whereas the code rate for the convolutional code is given by $R_c^{\mathrm{CC}} = I/J$.

For the upcoming Subsections and Chapters, the general description of convolutional codes is simplified to the specific case of $I = 1$ input blocks and code rate $R_c^{\mathrm{CC}} = 1/J$. Therefore, only one input sequence $\mathbf{u} = [u_k]_{K\times 1}$, $u_k \in \{0,1\}$ of length K and subindex k denoting the time index, is given.

For example, **Fig. 2.7** shows a non-recursive non-systematic convolutional encoder with generator matrix $\mathbf{g} = [5;7]_8$. The encoder consists of two linear shift registers

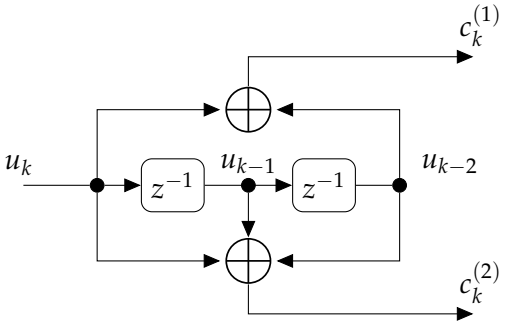


Fig. 2.7.: Non-recursive non-systematic convolutional encoder with $R_c^{\mathrm{CC}} = 1/2$

and two modulo-2 adders, performing the module-2 convolutions. Furthermore, the information bit u_k at time instance k is modulo-2 added (XOR) with the preceding

information bits u_{k-1} and u_{k-2}, in order to obtain the code bits $c_k^{(1)}$ and $c_k^{(2)}$. Thus, for this example, equation (2.24) is rewritten as

$$\mathbf{c}^{(1)} = \mathbf{u} \circledast \mathbf{g}^{(1)}$$
$$\mathbf{c}^{(2)} = \mathbf{u} \circledast \mathbf{g}^{(2)} \ ,$$

where $\mathbf{g}^{(1)} = 5_8 = [1\ 0\ 1]$ and $\mathbf{g}^{(2)} = 7_8 = [1\ 1\ 1]$. Due to the structure with two outputs, the code rate is given by $R_c^{\mathrm{CC}} = 1/2$.

Besides non-recursive convolutional encoders, without feedback in their encoding structures, there also exists recursive convolutional encoders. The main difference in between is, that the recursive convolutional encoder has a feedback loop around the shift registers. Furthermore, recursive and non-recursive convolutional encoders can be systematic. Here, one output bit (code bit) is set directly to one input bit (information bit), bypassing modulo-2 sums and shift registers.

Although *recursive systematic convolutional code* (RSC) codes seem to achieve better decoding performance especially in iterative decoding schemes [BG96], the focus in this dissertation is mostly on *non-systematic convolutional code* (NSC) codes. This is due to problems [WWKK08] with RSC codes in combination with error-prone relays and *soft-input soft-output* (SISO) re-encoding[6]. Here, soft re-encoding the estimated source's information sequence with a RSC code at the relay node will result in loosing almost the complete non-systematic information. In this case only systematic information is amplified and transmitted. This maybe result in a lower performance as compared to a relay network employing NSC codes.

Maximum-a-posteriori decoder

As already mentioned, the convolutional code can be decoded by the Viterbi algorithm via the *maximum-likelihood* criteria, or the BCJR algorithm via the *maximum-a-posteriori* principle. Both methods work on the trellis representation of the code. An example of a trellis for a convolutional code with rate $R_c^{\mathrm{CC}} = 1/2$, four states $S_k \in \{0, \ldots, 3\}$ at time instance k with two possible state transitions from state S_{k-1} to S_k each, is depicted in **Fig. 2.8**. Each of these transitions corresponds to one code word (code bit pair), which itself was generated by the current information bit and all bits in the register during the encoding process. Furthermore, the trellis shows the transitions at its end for a terminated code, given by additional tail bits which ensures a predefined state (in this example $S_6 = 0$) to which the encoder terminates. Utilizing this knowledge, the decoding result improves slightly.

[6] A detailed description of soft re-encoding is given later in Section 6.1.2 for the relay node

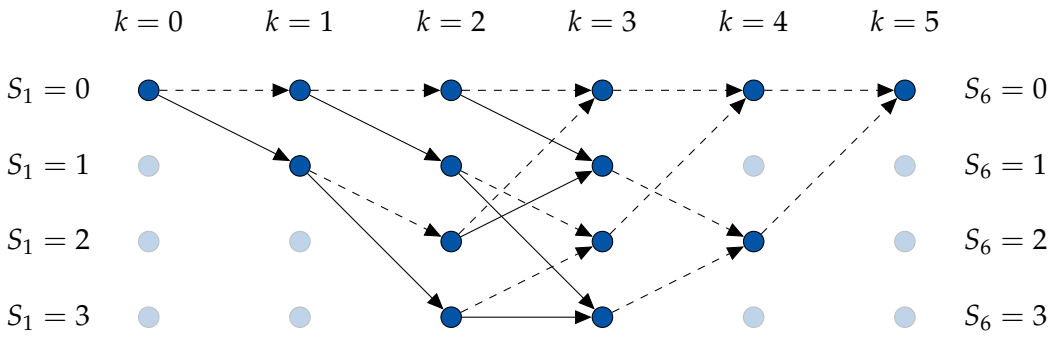


Fig. 2.8.: Trellis diagram for a terminated CC with $R_c = 1/2$

As stated earlier, the BCJR algorithm [BCJR74] is an effective implementation of the symbol-wise MAP principle. It is based on recursively computing forward and backward probabilities (α and β) and outputs soft-information $L(u_k|\mathbf{y})$ over the information bit u_k at time instance k via

$$L(u_k|\mathbf{y}) = \log \frac{\sum\limits_{(s',s),u_k=0} \alpha_{k-1}(s') \cdot \gamma_k(s',s) \cdot \beta_k(s)}{\sum\limits_{(s',s),u_k=1} \alpha_{k-1}(s') \cdot \gamma_k(s',s) \cdot \beta_k(s)} \quad , \tag{2.25}$$

based on the branch metric γ. In equation (2.25), s and s' denotes the corresponding state $s = S_k$ and previous state $s' = S_{k-1}$ at time k and $k-1$, respectively, whereas (s',s) defines the state transition from s' to s. Thus, the summation in the numerator is computed over all transitions (s',s) for which the message bit is $u_k = 0$. The same holds for the denominator with $u_k = 1$.

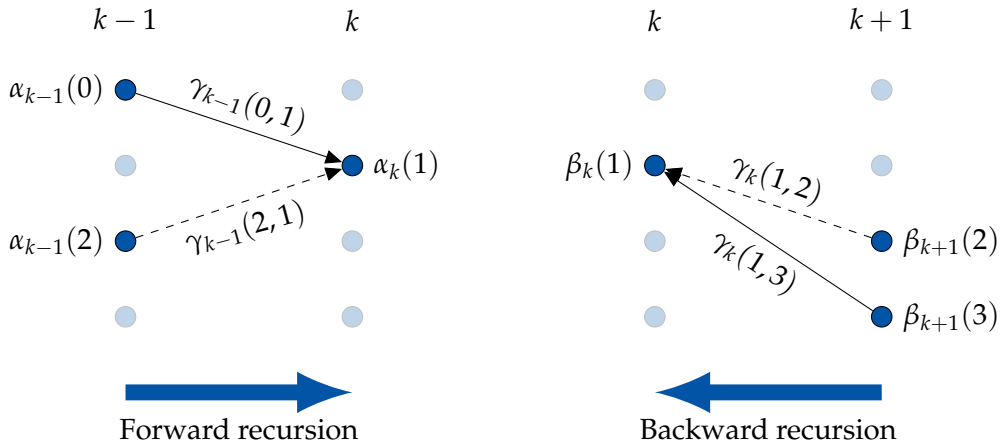


Fig. 2.9.: Forward and backward recursion example within the given trellis diagram from **Fig. 2.8**

To clarify the recursion for calculating α, β and γ, lets consider **Fig. 2.9**. It shows one example for possible transitions entering and leaving a single state inside a trellis. To be precisely, the figure depicts the states $S_{k-1} = 0$ and $S_{k-1} = 2$ at time instance

$k-1$ entering state $S_k = 1$ at time instance k in order to calculate $\alpha_k(S_k = 1)$ with the branch metrics $\gamma_{k-1}(S_{k-1} = 0, S_k = 2)$ and $\gamma_{k-1}(S_{k-1} = 2, S_k = 1)$ for the forward recursion. For the backward recursion, $\beta_k(S_k = 1)$ is given by $\beta_{k+1}(S_{k+1} = 2)$ with $\gamma_k(S_k = 1, S_{k+1} = 2)$ and $\beta_{k+1}(S_{k+1} = 3)$ with $\gamma_k(S_k = 1, S_{k+1} = 3)$.

Thus, the probability $\alpha_k(s)$ is calculated during the forward recursion via

$$\alpha_k(s) = \sum_{s'} \gamma_k(s', s) \cdot \alpha_{k-1}(s')$$

and $\beta_k(s)$ is computed during the backward recursion via

$$\beta_{k-1}(s') = \sum_{s} \gamma_k(s', s) \cdot \beta_k(s) \ ,$$

respectively. In both cases, the sum is over all possible encoder states. Forward and backward metrics are initialized at their respective beginning via

$$\alpha_0(s') = \begin{cases} 1, & s' = 0 \\ 0, & s' \neq 1 \end{cases}$$

and

$$\beta_K(s) = \begin{cases} 1, & s = 0 \\ 0, & s \neq 1 \end{cases} \ .$$

For $\beta_K(s)$, it is assumed that the code sequence was terminated and the required termination bits are present at the end of the information sequence. If the code is not terminated, then all states s at the end of the sequence can be regarded as equal probable and the backward metrics are initialized by $\beta_K(s) = 1/2^{L_c - 1}$. Another approach utilizes the knowledge of $\alpha_K(s)$ and assumes that the probabilities at the end from the forward recursion delivers the probabilities of the starting states for the backward recursion and, thus, $\beta_K(s) = \alpha_K(s)$ is set.

The branch transition probability is given by

$$\gamma_k(s', s) = \Pr(s|s') \cdot p_{Y|S',S}(\mathbf{y}_k|s', s) \tag{2.26}$$

whenever a transition (s', s) is possible. In cases where no transition inside the trellis between state s' and state s exists, $\Pr(s', s) = 0$ and, thus, $\gamma_k(s', s) = 0$ holds.

For a binary input additive white Gaussian noise channel with i.i.d distributed information bits, the a-priori probability is $\Pr(s|s') = 1/2$ and the conditional transition probability function is given by

$$p_{Y|S',S}(\mathbf{y}_k|s',s) = \prod_{n=1}^{N} \frac{1}{\sqrt{2\pi\sigma_N^2}} \cdot e^{-\frac{\left\|y_k^{(n)} - x_k^{(n)}\right\|^2}{2\sigma_N^2}}, \tag{2.27}$$

where $\mathbf{x}_k$ is the BPSK modulated version of the binary code bits $\mathbf{c}_k$ belonging to state transition (s', s) at time instance k.

2.3.2 Code concatenation and iterative decoding schemes

In [BGT93] Berrou, Glavieux and Thitimajshima proposed the combination of two FEC schemes encoding the same information sequence and decoding the outcome after transmission and reception via an iterative decoding scheme. This has the great advantage that the bit error rate performance is improved, although the two employed codes might only be decent for the underlying channel.

The utilized encoders can either be parallel or serial concatenated. At the receiver, the received sequence is iteratively decoded by a so called *turbo code* decoder, due to the similar structure of the decoder compared to a turbo engine. This is depicted in **Fig. 2.10**, for a given parallel concatenated code as shown in **Fig. 2.11**. The

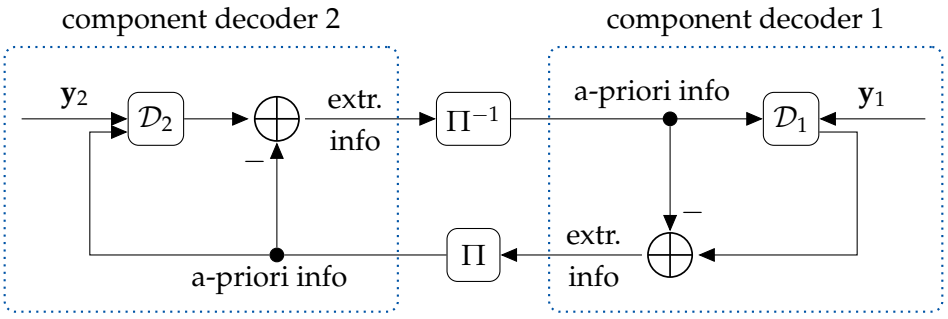


Fig. 2.10.: Iterative decoder for parallel concatenated codes

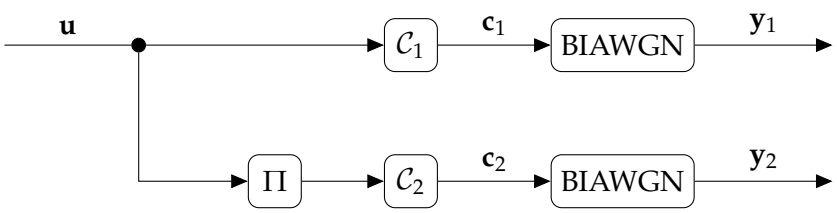


Fig. 2.11.: Parallel concatenated encoding scheme

turbo decoder receives two code words $\mathbf{y}_1$ and $\mathbf{y}_2$ over a single or two independent

channels. Since it is assumed, that there is no interference between these two code sequences, both can be decoded separately by their corresponding component decoder.

Thus, component decoder $\mathcal{D}_2$ tries to recover an interleaved version of the transmitted information sequence, which is also the decoding objective for component decoder $\mathcal{D}_1$ and vice versa. In order to improve the decoding result, both can share their extrinsic information in an iterative process. It is important to note, that the component decoders have to share log-likelihood ratio information and, therefore, must be based on soft-input soft-output decoders like the BCJR algorithm. The iterative exchange of reliability information continuous until a certain stopping criterion has been reached. This can be for example the maximal number of loops, which have to be performed. After the decoding process has finished, a hard-decision is made on either one of the component decoders' soft information sequence in order to retain the information bits. Usually, the LLR values from component decoder with the non-interleaved information sequence is selected.

2.3.3 Information processing characteristic

To measure the coding scheme performance of different encoder-decoder pairs with an underlying time invariant memoryless channel, Huettinger et al. proposed the *information processing characteristic* (IPC) in [HHJF01] and [HHFJ02]. Here, in order to gain the mutual information of practical codes for such a *superchannel*, exhaustive Monte-Carlo simulations are performed. A similar approach was taken by S. ten Brink developing the *extrinsic information transfer* (EXIT) chart analysis in [Bri99] and [Bri01]. The focus for the EXIT chart analysis relies in the development of a simple tool, predicting the behavior of soft-input soft-output turbo codes.

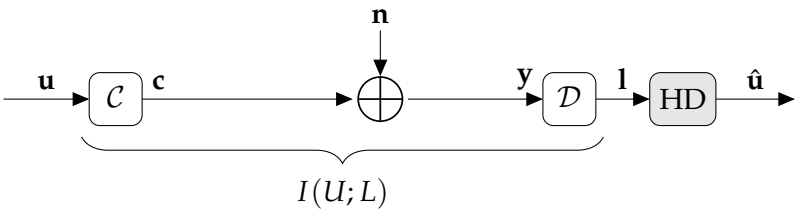


Fig. 2.12.: Simplified communication model for illustrating the generation of the information processing characteristic

Basically, the IPC measures the mean mutual information $I(U;L)$ or $I(U;\hat{U})$ between the input and output of a transmission system as depicted in **Fig. 2.12**. Hence, it measures the mutual information of the end-to-end channel and not the extrinsic

information of component codes, like it is the case for the EXIT chart analysis. The IPC for symbol-by-symbol soft-output decoding is given by [HL06]

$$\mathrm{IPC}_{\mathrm{SD}}^{\mathrm{info}}(C) = I(U;L) = \frac{1}{K}\sum_{k=1}^{K} I(u_k;l_k) \ , \tag{2.28}$$

where the i.i.d. information bits $\mathbf{u} = [u_k]_{K\times 1}$ are realizations of the binary random variable $U \in \{0,1\}$ and the LLR at the decoder output $\mathbf{l} = [l_k]_{K\times 1}$ are realizations of the random variable $L \in \{\mathbb{R}\}$. In (2.28), $I(u_k;l_k)$ denotes the mutual information per symbol between the encoder input $\mathbf{u}$ and decoder output $\mathbf{l}$. The mutual information depends on the encoder-decoder pair and the SNR respectively the capacity C of the underlying memoryless communication channel in between both. Equivalently to (2.28), the IPC after a hard-decision is calculated by $\mathrm{IPC}_{\mathrm{HD}}^{\mathrm{info}}(C) = I(U;\hat{U}) = \frac{1}{K}\sum_{k=1}^{K} I(u_k;\hat{u}_k)$.

In this dissertation, the IPC is utilized to predict the mutual information for the different links of a relay network. Owning this information, one can combine them in order to retain the overall MI performance of such a network. For this, there are basically three different classes of codes, including their corresponding IPCs, which have to be considered.

The first class consists of IPCs for *ideal coding schemes*, which have been defined by Huettinger, Huber, Johannesson and Fischer [HHJF01]. They represent the behavior of capacity achieving codes with MI equal to one (error-free decoding), but with a decoding error bounded away from zero (MI<1) once the code rate is above the channel capacity ($R_c > C$). This behavior can be compared with the result for the weak converse of the coding theorem [Fei58].

The second class of IPCs follows from the result for the strong converse of the coding theorem [Wol64] and will be donated as IPCs for *good coding schemes*. Similar to the ideal coding schemes, good coding schemes also utilizes capacity achieving codes. But, different to ideal coding schemes, good coding schemes cannot communicate at all once the rate is above the channel capacity. Here, the probability of decoding error is maximized (MI=0) and, thus, no information gain is achieved by decoding. Please note, that in this dissertation the term *good coding scheme* or, sometimes referred to as *good codes*, is different to the definition in [Mac00] or [PSS05]. In these publications, MacKay, Peleg and others utilizes the term of good codes indicating codes which are capacity achieving in case of $R_c < C$, but they do not provide clear information of how these codes behave in case of $R_c > C$.

Practical coding schemes compose the third class of information processing characteristics. These codes, sometimes in literature also referred to as bad codes [BSC10],

don't belong to neither class one or two, as they won't achieve capacity. Their encoding and decoding time is quasi-polynomial in the block length and their IPCs have to be retrieved by simulation as described above.

The information processing characteristics of class one and two are utilized to predict the mutual information of a three-node orthogonal relay network with capacity achieving codes. Here, the question which is tried to be answered is, if these codes can achieve the *max-flow min-cut* (MFMC) rate of this network. And, if not, how large is the gap between the upper bound given by MFMC and the actual achievable MI utilizing ideal and good codes.

Later in this dissertation, the IPCs of practical codes are employed to predict orthogonal relay networks with these types of codes.

Information processing characteristic for ideal codes

With the help of Fano's inequality [Fan61] and the rate-distortion theory [Sha59], the information processing characteristic of an ideal coding scheme can be specified as [HHJF01]

$$\mathrm{IPC}^{\mathrm{ideal}}(C) = I\left(U;\hat{U}\right) = \frac{1}{K} I\left(\mathbf{U};\hat{\mathbf{U}}\right) = \begin{cases} C/R_c & \text{for } C < R_c \\ 1 & \text{for } C \geq R_c \end{cases} . \tag{2.29}$$

Thus, for code rates below capacity $R_c \leq C$ an error free communication can be guaranteed, and $\mathrm{IPC}^{\mathrm{ideal}}(C) = 1$ holds. As soon as $R_c > C$, the information cannot be retrieved perfectly and the MI decreases with decreasing channel capacity C. Furthermore, in [HHFJ02] and [PSS05] it was proven that the IPC for an ideal coding scheme is identical for symbol-by-symbol and sequence-wise decoding, i.e. $I\left(U;\hat{U}\right) = \frac{1}{K} I\left(\mathbf{U};\hat{\mathbf{U}}\right)$ holds, with no benefit for soft outputs, since all symbols at the decoder output have the same reliability's. Thus, the channel between the encoder input and decoder output is a memoryless binary symmetric channel with no additional benefit for utilizing soft information as given by the LLR values [LH05]. Hence, the mutual information between $I\left(U;\hat{U}\right)$ and $I(U;L)$ is equal with equal reliable symbols $\hat{u}_k$.

A different prove for $I\left(U;\hat{U}\right) = \frac{1}{K} I\left(\mathbf{U};\hat{\mathbf{U}}\right)$ (at least in the case of $R_c \leq C$) was given by MacKay in [Mac00]. He showed that for optimally decoded capacity achieving codes, the thresholds at which the bit and word error probabilities become zero are equal.

Please note that the proof for the ideal coding scheme is shortly reviewed in the appendix A.1 on page 149. Furthermore, please note that the result $I(U;\hat{U}) = 1$ in case of $C = R_c$ in (2.29) holds due to the fact that a memoryless BSC is created [CT06, p. 208].

Furthermore, the ideal coding scheme is closely related to the weak converse of the coding theorem, which was proven by Feinstein [Fei58] for discrete symmetric memoryless channels by using Fano's inequality. The major difference between both is that the converse [CT06, p. 206] delivers a lower bound on the word error probability if the code word-length and, therewith, the alphabet size (possible number of code words) tends to infinity. In contrast, the ideal coding scheme was proven with the stronger form of Fano's inequality [CT06, p. 39] together with a binary alphabet. In fact, replacing the property of infinite alphabet size inside the prove of the weak converse with a binary alphabet, the decoding error probabilities for both (the ideal coding scheme and the weak converse to the coding theorem) are equal. This is shown in the Appendix A.2 on page 151.

Please note that, the IPC of an ideal coding scheme is the upper bound on mutual information for any other coding scheme utilizing any sort of decoding algorithm. Thus, the MI for sequence-wise or symbol-by-symbol decoding is less or equal to this bound [Lan05b]. Furthermore, in this dissertation and in the context of capacity achieving coding schemes, the IPCs are only regarded in the context of information symbols. For the information processing characteristics of code-symbol information, please refer to [LH05].

Information processing characteristic for good codes

In [Wol64], Wolfowitz established the strong converse of the coding theorem. It states that the probability of decoding error approaches one as the code-word length tends to infinity and the code rate is above capacity [Ari73][Rei66]. Therefore, the information processing characteristic under consideration of the strong converse is given by

$$\mathrm{IPC}^{\mathrm{good}}(C) = I(U;\hat{U}) = \frac{1}{K} I(\mathbf{U};\hat{\mathbf{U}}) = \begin{cases} 0 & \text{for } C < R_c \\ 1 & \text{for } C \geq R_c \end{cases}, \tag{2.30}$$

and will be denoted as the IPC for good codes. This is a step-function, where the mutual information is zero as long as $R_c > C$ and one for $R_c \leq C$ (with equality in case of BSCs [CT06, p. 208]). One can think of a perfect decoder, which only outputs perfectly decoded information sequences. Otherwise, in case of decoding errors,

the decoder is silent. Thus, the information processing characteristic for a good coding scheme is only a lower bound on the mutual information of any capacity achieving coding scheme.

A similar observation was made in [PSS05] and [PSS07], where the authors proved that the extrinsic information about the coded bits of any capacity achieving code is zero when the channel capacity is below the code rate. Thus, the receiver has only knowledge of the intrinsic information in front of the channel decoder and the channel decoder cannot provide any additional information. Unfortunately, the authors do not provide any insides on the error probability or mutual information of the information sequence or symbols after decoding.

Comparing ideal and good coding schemes, one can conclude that their information processing characteristics deliver upper and lower bounds on mutual information for the end-to-end channel in case of $R_c > C$. Once, $R_c \leq C$, reliable communication is possible and the upper and lower bounds are tight, i.e. $\mathrm{IPC}^{\mathrm{ideal}}(C) = \mathrm{IPC}^{\mathrm{good}}(C) = 1$. This is shown in **Fig. 2.13**, where the information

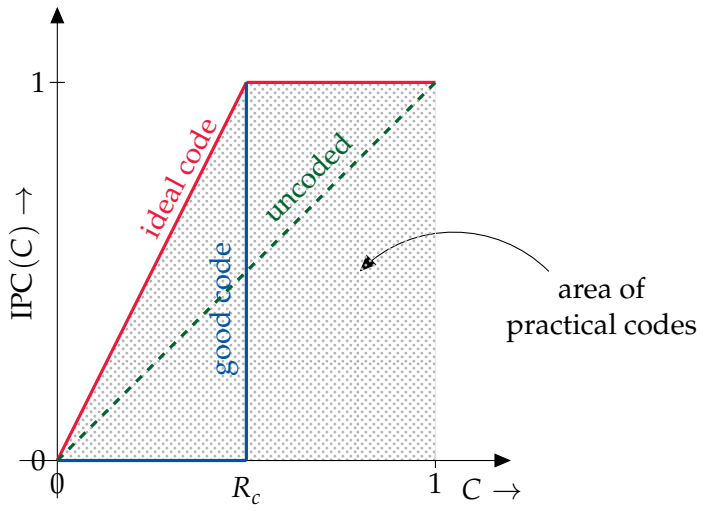


Fig. 2.13.: IPC for ideal and good codes

processing characteristics $\mathrm{IPC}^{\mathrm{ideal}}(C)$ and $\mathrm{IPC}^{\mathrm{good}}(C)$ are depicted together with the area in which the information processing characteristics $\mathrm{IPC}^{\mathrm{practical}}(C)$ of practical codes reside.

Information processing characteristic for practical coding schemes

Practical codes, as for example the non-recursive $R_c^{\mathrm{CC}} = 1/2$ and recursive $R_c^{\mathrm{CC}} = 1/4$ convolutional codes in **Fig. 2.14**, show totally different behavior compared to the capacity achieving ones. Their $\mathrm{IPC}^{\mathrm{info}}(C)$ is always below of $\mathrm{IPC}^{\mathrm{ideal}}(C)$, as long

as their code word length is reasonable large. Thus, all depicted codes reside inside the shaded area of **Fig. 2.13**, depending on their code rate R. As a contradiction, in some cases and only for codes with only a descent word length, the estimated mean IPC value can be above of the ideal coding schemes. This is due to the limitation in measuring the (exact) expectation in short sequences.

Besides the introduced soft-input soft-output decoders, where also exists so called *hard-input hard-output* (HIHO) decoders. These decoders are either a combination of SISO decoders with subsequent hard-decision, or decoders which only outputs binary values as for example the original Viterbi decoder [Vit67]. Taken the mutual information as a measurement for the decoding quality, it does matter if a SISO or HIHO decoder is applied. Here, the MI of a HIHO decoder is always smaller than for a SISO decoder, since the decision function destroys useful information about the decoding quality of the symbols. This relation in MI for SISO and HIHO decoders is shown in **Fig. 2.14** for example for a $R_c^{\mathrm{CC}} = 1/4$ recursive convolutional code with maximum-a-posteriori decoding and additional hard-decision in case of HD.

To better distinguish between the IPCs of HIHO and SISO decoding, the IPCs are indexed either with HD and SD, respectively. The *info* in $\mathrm{IPC}^{\mathrm{info}}(C)$ indicates that the IPC is measured over the information symbols from the encoder input compared to the decoder output. In cases where the IPC for the extrinsic or code symbols are measured, *extr* [7] or *code* would be attached. Thus, for example $\mathrm{IPC}_{\mathrm{SD}}^{\mathrm{info}}(C)$ is giving the measured mutual information of soft information symbols, whereas $\mathrm{IPC}_{\mathrm{HD}}^{\mathrm{code}}(C)$ denotes the MI of the hard decided code bits.

As an example for a near capacity achieving code, the IPC of the *repeat-accumulate code* (RAC) from [TB00], with underlying BIAWGN channel and capacity C_{BIAWGN}, is depicted in **Fig. 2.14**. It is a serially concatenated code, with an outer repetition code of rate $R_c^{\mathrm{RP}} = 1/2$ and an inner recursive non-systematic convolutional code code with rate $R_c^{\mathrm{CC}} = 1$. The decoder is an iterative decoder for serially concatenated codes. Although it approaches the Shannon limit by approximately 0.1 dB and is therefore close to the capacity bound, its IPC shows some different and interesting behavior compared to the strong converse IPC of good codes.

Since the decoder cannot decode the received sequence in case of $C_{\mathrm{BIAWGN}} \leq R_c$, the MI of the information symbols is almost zero. Only due to the few systematic information bits, which are inserted during the encoding process, the MI is slightly above zero. Once the channel capacity is above the code rate, the decoder is able to decode perfectly and the MI of the information symbols (bits) jumps to one.

[7] It will be obvious from the context, whether the extrinsic information is given for code or information symbols.

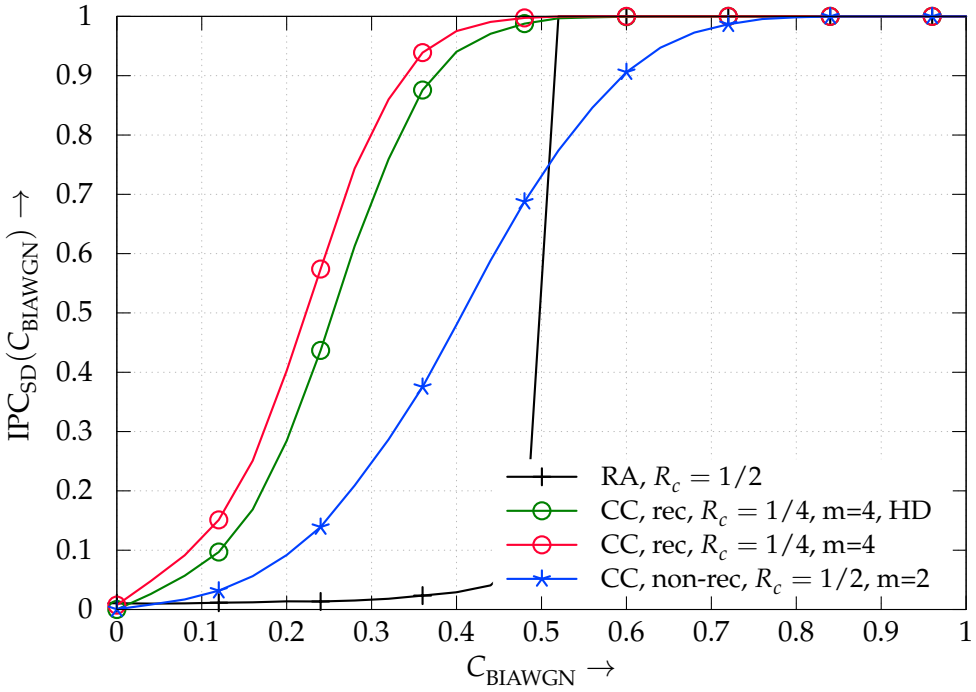


Fig. 2.14.: IPC over the information bits for two convolutional codes and one repeat-accumulate code

Summarizing, the repeat-accumulate code is a good example of a practical code with decoding performance close to the strong converse (good codes).

Comparing the repeat-accumulate code with the non-recursive $R_c^{\mathrm{CC}} = 1/2$ convolutional code, the latter one shows an interesting behavior. Although the CC code is by no means capacity achieving, the MI is above the RAC code in cases of $C_{\mathrm{BIAWGN}} < R_c$. This region is especially interesting for the development of concatenated codes, as the component decoders should operate above capacity and close to the ideal code characteristic, to obtain a coding scheme which operates near the channel capacity. A few more examples of codes which have a near ideal coding scheme characteristic are given in [HHJF01].

3 Relay networks and relaying protocols

This chapter deals with the orthogonal half-duplex three-node relay channel and its upper bound in channel capacity given by the *max-flow min-cut* (MFMC) theorem. Furthermore, different relaying strategies are introduced together with their achievable rates. Among these relaying protocols are *Amplify-and-Forward* (AF), *Decode-and-Forward* (DF) and *Compress-and-Forward* (CF). Afterwards, the rates for the different protocols are evaluated and compared against each other in different scenarios. Unfortunately, none of the given strategies reaches the upper bound for all simulation parameters (relay position etc.).

3.1 Orthogonal half-duplex relaying

Cover and El Gamal [CG79] investigated the non-orthogonal full-duplex three-node relay channel from the mutual information point of view. In theory, these non-orthogonal full-duplex relay channels show tremendous gains in capacity compared to point-to-point wireless communication systems.

In practice, such relay systems can hardly be implemented. Here, if the nodes will transmit and receive in the same frequency band and the same time slot, the transmit and receive signals will interfere with each other. Therefore, an interference canceler must be implemented. For the interference cancellation process, all characteristics needs to be known very precisely. Otherwise, the interference cancellation can be catastrophic. The problem is that the transmit signal is typically $\approx$ 100 to 150 dB stronger [LTW04] than the receive signal. This additionally demands expensive high-resolution fast sampling analog-to-digital converters and a sufficient electromagnetic isolation between the receiving and transmitting circuitry.

Another major problem in non-orthogonal full-duplex relay network is timing. All nodes must be precisely synchronized in time, in order to perform a non-orthogonal cooperative channel access. Either the nodes possess extremely accurate and, thus, expensive timers, or the network must be permanently synchronized.

So, in practice the benefits of non-orthogonal full-duplex relay networks disappear or can only be maintained if expensive equipment is utilized. Therefore, nowadays mostly orthogonal half-duplex relays are considered [MKP+09] in ongoing standardization processes. These relay nodes cannot transmit and receive at the same time or in the same frequency band, which will eventually degrade the throughput. Furthermore, in this dissertation it is assumed that source and relays transmit orthogonal during two separate time slots. the combination of both signals is performed at the destination, once the source's and the relay's signal has been received and stored.

3.2 System model

The system model of a three-node relay network is depicted in **Fig. 3.1**. It consists of a source $\mathcal{S}$, a destination $\mathcal{D}$ and a relay $\mathcal{R}$ in between. As the channel access is supposed to be orthogonal, the transmission of information is departed into two normalized time slots T_1 and T_2, with corresponding duration t_1, t_2 and their relationship $t_1 + t_2 = 1$.

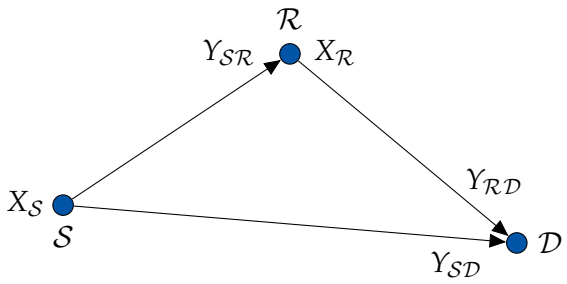


Fig. 3.1.: Structure of a three-node relay network

In the first time slot T_1, the source transmits its independent and identically distributed (i.i.d.) code symbols $X_{\mathcal{S}} \in \{\mathbb{R}\}$ with zero mean and variance $\sigma^2_{X_{\mathcal{S}}} = 1$ towards the relay and the destination.

If the channel condition between $\mathcal{S}$ and $\mathcal{D}$ is good, then the destination might be able to perfectly recover the source's information from the received signal $Y_{\mathcal{SD}} \in \{\mathbb{R}\}$. Otherwise, the relay can support the source by forwarding an altered version $X_{\mathcal{R}} \in \{\mathbb{R}\}$ of the received sequence $Y_{\mathcal{SR}} \in \{\mathbb{R}\}$ towards the destination during the second time slot T_2. Furthermore, utilizing relays does not only increase the Quality of Service, it can also increases the possible rate at which the source is communicating with the destination.

Depending on the chosen relay protocol, the size of the relay's transmit signal is not necessarily the same as of the source signal. Therefore, the time slot lengths maybe not equal.

Throughout this thesis, it is assumed that the source stays silent in T_2, in order to guarantee an orthogonal access during the second time slot. This assumption obviously degrades the throughput, but has some other advantages, like an easier radio transmitter design.

3.2.1 Max-flow min-cut theorem

One goal in this thesis is to semi-analytically calculate the mutual information of large practical relay networks and to evaluate how far this MI is from the capacity. The problem is that the capacity of the general relay network is still unknown. There is only the possibility to calculate an upper bound with the *max-flow min-cut* (MFMC) theorem from [FF56] [EFS56]. The theorem is a generalization of the Menger theorem [Men27] and loosely states that:

> *The maximum possible flow from the source to the sink is equal to the minimum capacity among all simple cut-sets.*

Please note, that the following description is given universally but will be restricted towards orthogonal half-duplex relay networks at the end of this subsection.

Depicted in **Fig. 3.2** is a simple directed graph $G = (\mathbb{V}, \mathbb{E})$, consisting of finite sets $\mathbb{V}$ and $\mathbb{E}$ including vertices and edges, respectively. The flow in the depicted network is from vertex T (source, transmitter) to vertex R (sink, receiver). Furthermore, each

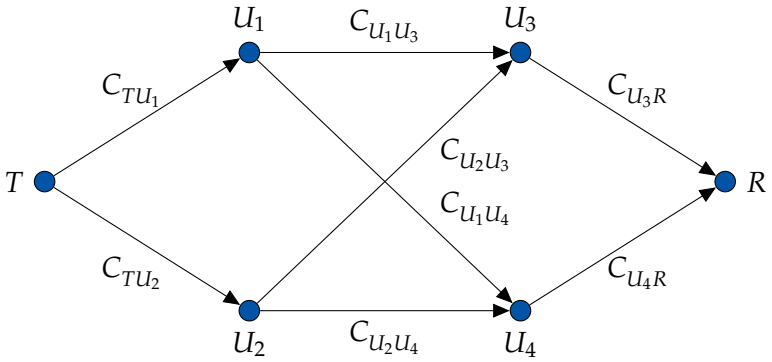


Fig. 3.2.: Directed graph with a single Source T and a single Destination R

edge has a capacity C_{uv}, which is the maximum flow from u to v with

$$\sum_{u:(u,v)\in\mathbb{E}} C_{uv} = \sum_{u:(v,u)\in\mathbb{E}} C_{vu} \ , \tag{3.1}$$

for each $v \in \mathbb{V} \setminus \{T, R\}$. Here, the assumption of maximum possible flows on all existing edges was made. Therefore (3.1) states, that the sum of capacities for all incoming edges must equal the sum of capacities for all leaving edges.

Now, in order to calculate the maximum capacity of the given network, one has to find the *bottleneck* in the graph, which is simply the *minimum cut set*, i.e. the cut with minimum weight or smallest capacity. For this, the graph is separated into two disjoint *cut-sets* $\mathbb{T} \subseteq \mathbb{V}$ and $\mathbb{R} \subseteq \mathbb{V}$ with $\mathbb{V} = \mathbb{T} \cup \mathbb{R}$, $\mathbb{T} = \mathbb{V} \setminus \mathbb{R}$ and $\mathbb{R} = \mathbb{V} \setminus \mathbb{T}$. The edges between the two disjoint sets, i.e. the edges through which the cut was performed, will be referred to as *cut edges*. Each cut has its own capacity, which is simply the sum of all cut edges

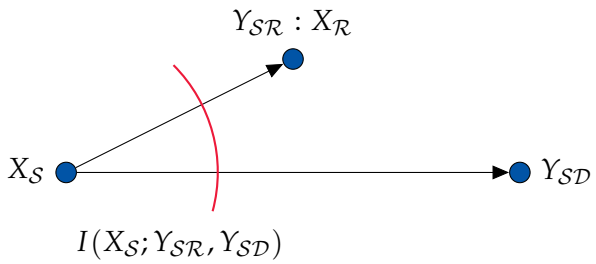
$$C(\mathbb{T}, \mathbb{R}) = \sum_{u\in\mathbb{T},\ v\in\mathbb{R}|(u,v)\in\mathbb{E}} C_{uv} \ . \tag{3.2}$$

After calculating the capacity for all *cut-sets*, one can determine the maximum capacity of the complete network with

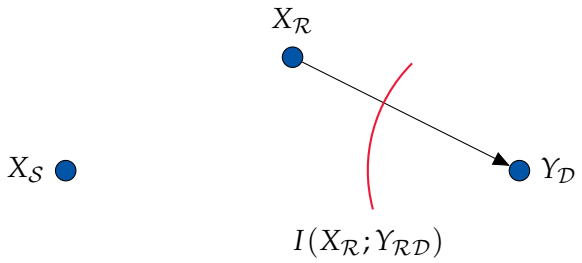
$$C = \min_{\forall(\mathbb{T},\mathbb{R})} C(\mathbb{T}, \mathbb{R}) \ , \tag{3.3}$$

by searching the minimum cut capacity.

Now, after the general description of the max-flow min-cut theorem was given, a relay network as depicted in **Fig. 3.3** with only three nodes, orthogonal access and half-duplex constrain is considered. This three-node relay network is the simplified version of the general three-node relay network, which is widely discussed in literature [CG79, KGG05]. The half-duplex constrain limits the relay not to receive



(a) Cut-set for the broadcast channel



(b) Cut-set for the multiple access channel

Fig. 3.3.: Max-flow min-cut theorem with both existing cut-sets for the three-node relay channel

and to transmit at the same time. So as discussed earlier, the transmission of information is departed into two time slots. Furthermore, the orthogonal access restricts the source to transmit only in the first time slot, whereas the relay transmits during the subsequent time slot.

The cut-sets for the half-duplex relay network with orthogonal access is depicted in **Fig. 3.3(a)** and **Fig. 3.3(b)**. Here, the first cut-set in **Fig. 3.3(a)** denotes the mutual information

$$I(X_{\mathcal{S}}; Y_{\mathcal{SR}}, Y_{\mathcal{SD}}) \tag{3.4}$$

between the source and the relay and also between the source and the destination. The second cut-set depicted in **Fig. 3.3(b)** delivers the mutual information

$$I(X_{\mathcal{R}}; Y_{\mathcal{RD}}) \tag{3.5}$$

of the $\mathcal{R} \rightarrow \mathcal{D}$ link.

Considering *orthogonal access* and the *half-duplex* constraint, the upper bound for the rate is given by

$$R_{\text{MFMC}} \leq \sup_{t_1,\, p_{X_\mathcal{S}, X_\mathcal{R}}(x_\mathcal{S}, x_\mathcal{R})} \min \left\{ \begin{array}{c} t_1 \cdot I(X_\mathcal{S}; Y_{\mathcal{SR}}, Y_{\mathcal{SD}}), \\ t_1 \cdot I(X_\mathcal{S}; Y_{\mathcal{SD}}) + (1 - t_1) \cdot I(X_\mathcal{R}; Y_{\mathcal{RD}}) \end{array} \right\} , \quad (3.6)$$

utilizing the MFMC theorem. Here, during T_1 the relay remains silent while the source transmits towards the relay and destination. In the second time slot T_2, the relay transmits towards the destination with a time duration of $t_2 = 1 - t_1$. Then, the destination can combine the sources information from time slot T_1 with the relays information from time slot T_2. The transmission rate R_{MFMC} in (3.6) serves as an upper bound for the three-node relay network with given constraints (orthogonal access and half-duplex) throughout this thesis.

3.3 Relaying protocols

In order to support the source transmitting information towards the destination and depending on the allowed complexity, the relay can choose from numerous types of relaying protocols. Such protocols have different advantages and disadvantages. For example, *Amplify-and-Forward* (AF) [LTW04] is beneficial if the relay was not able to decode the received sequence perfectly. Here, it forwards a scaled observation of the received signal towards the the destination. This maybe helps the destination to retain the original source's message if only some additional information is needed. This, amongst others, requires a second link with good channel conditions. If the source-relay link has good channel conditions and, therefore, the relay can perfectly decode the source's sequence, *Decode-and-Forward* (DF) [Lan02] should be chosen, as the relay possess the full source information and, thus, can perfectly support the source. Another relay protocol, where the relay forwards an observation of the received signal is *Compress-and-Forward* (CF) [CG79] and [LTW04]. Compared to AF, CF has a higher coding complexity as it will compress and quantize the observation before re-transmission.

All the above mentioned protocols will be discussed in the upcoming subsections. However, these subsections will lack a more in-depth description with all possible constraints. Only the mathematical foundations, which are needed for the upcoming chapters are introduced. Here, the major restriction is given by the half-duplex constraint and a quiet source during the second time slot.

3.3.1 Direct transmission

Although, not directly a relaying protocol, the *direct transmission* (DT) is given here for completeness and comparison reasons. As shown in Section 2.2 equation (2.21), the receive signal for the $\mathcal{S} \rightarrow \mathcal{D}$ AWGN link is characterized by

$$Y_{\mathcal{SD}} = X_{\mathcal{S}} + N_{\mathcal{D}} \ , \tag{3.7}$$

with signal-to-noise ratio from (2.22)

$$\gamma_{\mathcal{SD}} = \frac{a_{\mathcal{SD}}^2 \cdot \sigma_X^2}{\sigma_N^2} \tag{3.8}$$

and corresponding channel capacity $C_{\mathcal{SD}}$.

Since the relay is not present in this scenario, i.e. $t_2 = 0$, the source can always transmit and therefore, the capacity does not suffer from the loss due to the half-duplex constraint. As, in this thesis the channel access is assumed to be orthogonal, all other relaying protocols must meet the half-duplex constraint, which decreases the throughput. Thus, it is obvious that the information rate R of a relay network with at least one relay and orthogonal access first needs to overcome the capacity of the direct link, before it can be considered as a replacement for the direct transmission. Furthermore, time slot sizes and power constraints are additional requirements which have to be accounted for.

3.3.2 Non-cooperative multi-hop transmission

If there exists no direct link between source and destination, then the relay acts as a simple repeater. The advantage relies in shortened distances or in saving transmit power. Unfortunately, no diversity is achieved with this method, as compared to other relaying schemes.

Depending on the complexity of the relay node and under the restriction that the complete source information will be repeated, the relay can perform two major protocols. Either the relay just amplifies and forwards the received source signal, or, the relay decodes, re-encodes and forwards the source information. The first method is closely related to the *Amplify-and-Forward* (AF) relay protocol, whereas the second method is based on the *Decode-and-Forward* (DF) protocol. Both protocols will be discussed in the upcoming Subsections 3.3.3 and 3.3.4.

In case of non-cooperative *Multi-Hop Transmission* with AF (MHT-AF), the achievable rate is given by

$$R_{\mathrm{MHT-AF}} = \frac{1}{2} \cdot C\left(\gamma_{\mathcal{SRD}}^{\mathrm{MHT-AF}}\right) \tag{3.9}$$

with the corresponding SNR

$$\gamma_{\mathcal{SRD}}^{\mathrm{MHT-AF}} = \frac{\gamma_{\mathcal{SR}} \cdot \gamma_{\mathcal{RD}}}{1 + \gamma_{\mathcal{SR}} + \gamma_{\mathcal{RD}}} \ . \tag{3.10}$$

Comparing (3.10) with (3.16) from the upcoming Amplify-and-Forward relay protocol, equation (3.10) lacks the SNR of the direct link. Obviously, this is due to the absence of the direct link. Furthermore, the time slots T_1 and T_2 have equal time lengths.

In case of non-cooperative *Multi-Hop Transmission* with DF (MHT-DF), the rate is restricted to

$$R_{\mathrm{MHT-DF}} \leq \max_{p_{X_{\mathcal{S}},X_{\mathcal{R}}}(x_{\mathcal{S}},x_{\mathcal{R}})} \min\left\{t_1 \cdot C(\gamma_{\mathcal{SR}}), (1-t_1) \cdot C(\gamma_{\mathcal{RD}})\right\} \ , \tag{3.11}$$

where $t_1 = t_2$ is necessary in case of the upcoming Decode-and-Forward with repetition coding protocol (see Subsection 3.3.4). For Decode-and-Forward with incremental redundancy (see Subsection 3.3.4), the time slots must not be equal and, thus, t_1 can be optimized such that the mutual information is equal for both links.

3.3.3 Amplify-and-forward

Probably the most simplest, from the mathematical point of view, relaying protocol is *Amplify-and-Forward* (AF) [LTW04]. In AF, the relay receives the source's sequence and retransmits an amplified version towards the destination. For the three-node relay network in **Fig. 3.1**, the destination performs *maximum ratio combining* (MRC) to combine the channel outputs from time slot T_1 and T_2 in order to retain the complete received sequence

$$\begin{aligned} Y_{\mathcal{D}} =& \frac{a_{\mathcal{SD}}^2}{\sigma_N^2} Y_{\mathcal{SD}} + \frac{a_{\mathcal{SR}}^2 a_{\mathcal{RD}}^2}{\sigma_N^2 \left(a_{\mathcal{RD}}^2 \beta + a_{\mathcal{SR}}^2\right)} Y_{\mathcal{RD}} \\ =& \frac{a_{\mathcal{SD}}^2}{\sigma_N^2} \cdot \left(X_{\mathcal{S}} + N_{\mathcal{D}}^{T_1}\right) + \frac{a_{\mathcal{SR}}^2 a_{\mathcal{RD}}^2}{\sigma_N^2 \left(a_{\mathcal{RD}}^2 \beta + a_{\mathcal{SR}}^2\right)} \cdot \left(X_{\mathcal{R}} + N_{\mathcal{D}}^{T_2}\right) \ , \end{aligned} \tag{3.12}$$

with $N_{\mathcal{D}}^{T_1} \sim \mathcal{N}(0, \sigma_N^2/a_{\mathcal{SD}}^2)$, $N_{\mathcal{D}}^{T_2} \sim \mathcal{N}(0, \sigma_N^2/a_{\mathcal{RD}}^2)$ and $X_{\mathcal{R}}$ being the amplified version of the channel output $Y_{\mathcal{SR}}$, i.e.

$$\begin{aligned} X_{\mathcal{R}} &= \beta \cdot Y_{\mathcal{SR}} \\ &= \beta \cdot (X_{\mathcal{S}} + N_{\mathcal{R}}) \quad , \end{aligned} \tag{3.13}$$

with $N_{\mathcal{R}} \sim \mathcal{N}(0, \sigma_N^2/a_{\mathcal{SR}}^2)$ and the amplification factor [LTW04][Lan02, p. 92]

$$\beta = \sqrt{\frac{1}{1 + \sigma_N^2/a_{\mathcal{SR}}^2}} \quad . \tag{3.14}$$

Since the relay forwards a disturbed version of the source signal and, therefore, both time slots are equally long MRC can be applied at the destination. Unfortunately, equal time slots will degrade the rate R_{AF} for the Amplify-and-Forward relay network.

Combining (3.13) with equation (3.12) delivers

$$\begin{aligned} Y_{\mathcal{D}} = &\left(\frac{a_{\mathcal{SD}}^2}{\sigma_N^2} + \frac{\beta a_{\mathcal{SR}}^2 a_{\mathcal{RD}}^2}{\sigma_N^2 \left(a_{\mathcal{RD}}^2 \beta + a_{\mathcal{SR}}^2 \right)} \right) \cdot X_{\mathcal{S}} \\ &+ \frac{a_{\mathcal{SD}}^2}{\sigma_N^2} N_{\mathcal{D}}^{T_1} + \frac{a_{\mathcal{SR}}^2 a_{\mathcal{RD}}^2}{\sigma_N^2 \left(a_{\mathcal{RD}}^2 \beta + a_{\mathcal{SR}}^2 \right)} \left(N_{\mathcal{D}}^{T_2} + \beta \cdot N_{\mathcal{R}} \right) \quad . \end{aligned} \tag{3.15}$$

By transposing and neglecting constant factors, it is easy to show that the corresponding SNR is calculated through

$$\gamma_{\mathcal{D}}^{\mathrm{AF}} = \gamma_{\mathcal{SD}} + \frac{\gamma_{\mathcal{SR}} \cdot \gamma_{\mathcal{RD}}}{1 + \gamma_{\mathcal{SR}} + \gamma_{\mathcal{RD}}} \quad , \tag{3.16}$$

which is a well-known result in literature, e.g. [LW00].

In case of Amplify-and-Forward, the relay is a linear processor with the information rate [LTW04]

$$R_{\mathrm{AF}} = \frac{1}{2} \cdot C\left(\gamma_{\mathcal{D}}^{\mathrm{AF}} \right) \quad , \tag{3.17}$$

where C is any capacity function from (2.18) or (2.19), depending on the source's channel input, i.e. binary or Gaussian distributed. Furthermore, the factor $1/2$ in front of the right side in equation (3.17) is the loss in throughput from the half-duplex constraint, due to equal time slots lengths.

Amplify-and-Forward has some major disadvantages in comparison with the other relay protocols. For example, in AF the relay amplifies and forwards not only the desired signal, but also the received noise. Furthermore, the continuous signals from T_1 and T_2, which were received at the relay and destination, have to be converted from analog to digital for further processing. Thus, due to limited amount of memory, the benefits of AF vanishes as information gets lost during the sampling and quantization process.

3.3.4 Decode-and-forward

To overcome the shortcomings of AF, *Decode-and-Forward* (DF) [CG79] and [Lan02] decodes and re-encodes the received sequence from the source. This, on the one hand, saves memory since only the information sequence needs to be stored. On the other hand, the receiver noise is eliminated after successful decoding.

Once, the pure information sequence is recovered, the relay can re-encode this sequence for the second hop. This can be done in two ways:

1. *Decode-and-Forward with repetition coding* (DF-RC): The relay utilizes the same code as the source and, thus, the destination can *maximum ratio combine* (MRC) the relay's signal with the source's signal before decoding. This implies, that both time slots T_1 and T_2 have equal lengths.

2. *Decode-and-Forward with incremental redundancy* (DF-IR): The relay utilizes a different code for the forward error correction, which generates additional parity bits. Afterwards, both code sequences are concatenated at the destination, forming a larger code with lower rate and better decoding capabilities.

In both cases, error-free decoding of the source's information is required at the relay to avoid error propagation. This requirement should be kept in mind in the following subsections.

Decode-and-forward with repetition coding

If the relay applies the same code as the source then the destination can maximum ratio combine both signals prior to decoding. This leads to a modified version of equation (3.12) with

$$Y_{\mathcal{D}} = \left(\frac{a_{\mathcal{SD}}^2}{\sigma_N^2} + \frac{a_{\mathcal{RD}}^2}{\sigma_N^2} \right) \cdot X_{\mathcal{S}} + \frac{a_{\mathcal{SD}}^2}{\sigma_N^2} \cdot N_{\mathcal{D}}^{T_1} + \frac{a_{\mathcal{RD}}^2}{\sigma_N^2} \cdot N_{\mathcal{D}}^{T_2} \tag{3.18}$$

since $X_{\mathcal{R}} = X_{\mathcal{S}}$ holds. The resulting signal-to-noise ratio follows from MRC and is given by

$$\gamma_{\mathcal{D}}^{\text{DF-RC}} = \gamma_{\mathcal{SD}} + \gamma_{\mathcal{RD}} \ . \tag{3.19}$$

As successful decoding at the relay is required (see definition) the information rate is upper bounded by [Lan02, p. 93]

$$R_{\text{DF-RC}} \leq \frac{1}{2} \cdot \max_{p_{X_{\mathcal{S}},X_{\mathcal{R}}}(x_{\mathcal{S}},x_{\mathcal{R}})} \min\left\{I(X_{\mathcal{S}}, X_{\mathcal{R}}; Y_{\mathcal{D}}), I(X_{\mathcal{S}}; Y_{\mathcal{SR}})\right\} \ . \tag{3.20}$$

Here, the first part in (3.20) denotes the mutual information

$$I(X_{\mathcal{S}}, X_{\mathcal{R}}; Y_{\mathcal{D}}) = C\left(\gamma_{\mathcal{D}}^{\text{DF-RC}}\right) \ , \tag{3.21}$$

for the maximum ratio combined links $\mathcal{S} \to \mathcal{D}$ and $\mathcal{R} \to \mathcal{D}$, which was stated in (3.18) with corresponding SNR from (3.19). The second part in (3.20) delivers the mutual information of the $\mathcal{S} \to \mathcal{R}$ link, i.e.

$$I(X_{\mathcal{S}}; Y_{\mathcal{SR}}) = C(\gamma_{\mathcal{SR}}) \tag{3.22}$$

the maximum rate at which the relay can decode without producing errors. In equation (3.22), $\gamma_{\mathcal{SR}}$ denotes the corresponding signal-to-noise ratio for that link.

Decode-and-forward with incremental redundancy

In contrast to DF-RC, in *Decode-and-Forward with incremental redundancy* (DF-IR) the relay re-encodes the information with a code, different from the source's code. Thus, additional redundancy is generated, which, with the source's parity sequence, creates a longer code with better decoding capabilities. This is exploited by the destination during the decoding procedure.

Since, with *incremental redundancy*, both code words are transmitted over two independent channels, the mutual information of the two links can be accumulated [Gal68, p. 149, p. 343], if the relay has perfectly retained the source's information. Therefore, the rate (3.6) simplifies to

$$\begin{aligned} R_{\text{DF-IR}} \leq \max_{p_{X_{\mathcal{S}},X_{\mathcal{R}}}(x_{\mathcal{S}},x_{\mathcal{R}})} \min\{&t_1 \cdot I(X_{\mathcal{S}}; Y_{\mathcal{SR}}), \\ &t_1 \cdot I(X_{\mathcal{S}}; Y_{\mathcal{SD}}) + (1-t_1) \cdot I(X_{\mathcal{R}}; Y_{\mathcal{RD}})\} \end{aligned} \tag{3.23}$$

with

$$I(X_{\mathcal{S}};Y_{\mathcal{SR}}) = C(\gamma_{\mathcal{SR}}) \tag{3.24}$$

$$I(X_{\mathcal{S}};Y_{\mathcal{SD}}) = C(\gamma_{\mathcal{SD}})$$

$$I(X_{\mathcal{R}};Y_{\mathcal{RD}}) = C(\gamma_{\mathcal{RD}}) \ . \tag{3.25}$$

The mutual information statement on the left in equation (3.23), i.e. the capacity for the $\mathcal{S} \rightarrow \mathcal{R}$ link (3.24), results from the error-free decoding constraint at the relay, which has to be fulfilled. If the relay can perfectly decode the source's information it supports the source and thus, the MI statement on the right in (3.23) holds.

If both terms in 3.23 are identical, i.e. the resulting mutual informations on both sides are equal, than the optimum with respect to t_1 is reached.

3.3.5 Compress-and-forward

In DF, the $\mathcal{S} \rightarrow \mathcal{R}$ link limits the mutual information of the entire relay network, if the channel conditions of this link is worse as compared to the direct transmission. The *Compress-and-Forward* (CF) protocol [CG79] and [LTW04] tries to overcome limitations of the AF and DF protocols. If the relay cannot decode the source's information correctly, then the relay transmits a compressed and quantized version of its observation over $Y_{\mathcal{SR}}$ towards $\mathcal{D}$. Among the methods, which the relay can employ are Wyner-Ziv [WZ76], rate-distortion coding [CT91] or standard source coding.

According to [HMZ05], the achievable rate for the half-duplex Compress-and-Forward is

$$R_{\mathrm{CF}} \leq t_1 \cdot C\left(\gamma_{\mathcal{D}}^{\mathrm{CF}}\right) \ , \tag{3.26}$$

with the signal-to-noise ratio calculated by

$$\gamma_{\mathcal{D}}^{\mathrm{CF}} = \gamma_{\mathcal{SD}} + \frac{\gamma_{\mathcal{SR}}}{1+\sigma_w^2} \tag{3.27}$$

and the *compression noise* power

$$\sigma_w^2 = \frac{\gamma_{\mathcal{SR}} + \gamma_{\mathcal{SD}} + 1}{\left[(1+\gamma_{\mathcal{RD}})^{(1-t_1)/t_1} - 1\right] \cdot (1+\gamma_{\mathcal{SD}})} \ . \tag{3.28}$$

3.3.6 Some results on the different relaying protocols

All introduced relaying protocols have different advantages and disadvantages. Among them are their computational complexity or their error rate performance. In this subsection, the relay protocols are compared via their information rate R for the three-node relay network.

For the simulation, the relay is placed at different positions on a normalized 2D-grid, i.e. from (-1,-1) up to (1,1). Furthermore, the source, at position (0,0), and the destination, at position (1,0), are fixed and the underlying Gaussian input additive white Gaussian noise channels have a signal-to-noise ratio of

$$10\log_{10}(\gamma_{\mathcal{SD}}) = -10\,\text{dB}\ . \tag{3.29}$$

Since the distance for the direct link is given by $d_{\mathcal{SD}} = 1$ and, therewith, the path-loss attenuation factor is $a^2_{\mathcal{SD}} = \sqrt{d^{-\alpha}_{\mathcal{SD}}} = 1$, the noise variance $\sigma^2_N = 1/\gamma_{\mathcal{SD}}$ holds as a reference for all other SNRs with $\gamma_{\mathcal{SR}} = a^2_{\mathcal{SR}}/\sigma^2_N$ and $\gamma_{\mathcal{RD}} = a^2_{\mathcal{RD}}/\sigma^2_N$. Furthermore, a path-loss exponent of $\alpha = 4$ was assumed for all links during the simulation.

As a reference, the source has the possibility to transmit during both time slots if the relay is not present. Hence, the information rate is equivalent to the channel capacity of the direct link (2.19) by

$$C_{\text{GIAWGN}}(\gamma_{\mathcal{SD}}) = \frac{1}{2}\log_2(1+\gamma_{\mathcal{SD}}) \approx 0.07\,\text{bits/symbol} \tag{3.30}$$

with the given simulation parameters. The result is depicted as a red plane in **Fig. 3.4**. This plane also holds as the lower bound on the information rate. Therefore, only relaying protocols and relay positions, which deliver a higher information rate than the capacity of the direct link, should be considered for deployment. Please note, that the information rate will not alter at the different relay positions, since the relay does not participate at the transmission during the before mentioned case.

Also depicted in **Fig. 3.4** is the information rate

$$R_{\text{MFMC}} \leq \frac{1}{2}\cdot\min\left\{\begin{array}{c}\frac{1}{2}\log_2(1+\gamma_{\mathcal{SD}}+\gamma_{\mathcal{SR}})\,,\\ \frac{1}{2}\log_2(1+\gamma_{\mathcal{SD}})+\frac{1}{2}\log_2(1+\gamma_{\mathcal{RD}})\end{array}\right\} \tag{3.31}$$

for the MFMC upper bound from (3.6), under the constraint of equal time slot lengths $t_1 = t_2 = 1/2$. This constraint is, for example, necessary for the Amplify-and-Forward protocol and the Decode-and-Forward with repetition coding protocol.

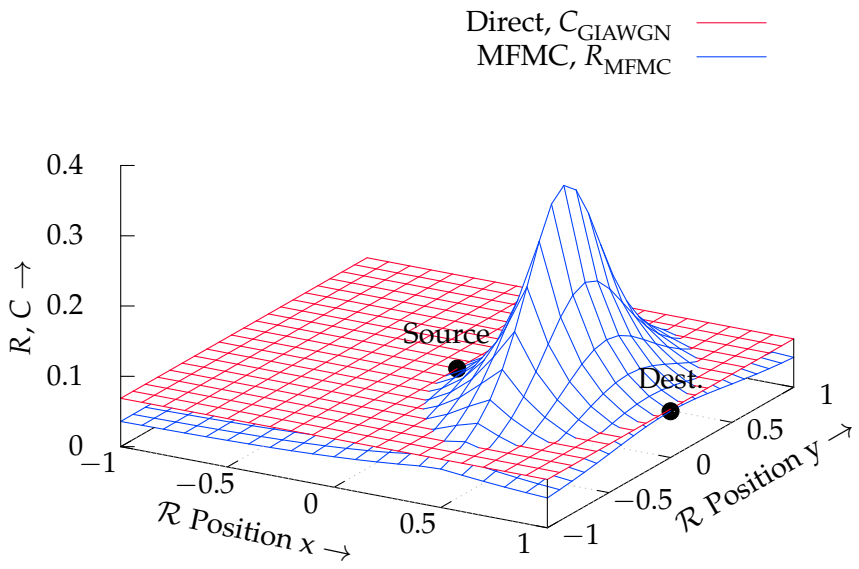


Fig. 3.4.: Rate and capacity for a three-node relay network with $\alpha = 4$, $\gamma_{\mathcal{SD}} = -10$dB. Position of the source and the destination nodes are fixed, while the relay's position was altered.

Furthermore, GIAWGN channels with channel capacities from (2.19) have been utilized in (3.31).

It can bee seen, that if the relay is in a certain region, the rate R_{MFMC} is less than the capacity $C_{\text{GIAWGN}}(\gamma_{\mathcal{SD}})$ of the direct link. In the worst case, for example at relay position (-1,-1), not only the $\mathcal{R} \rightarrow \mathcal{D}$ link is poor, but also the $\mathcal{S} \rightarrow \mathcal{R}$ link. Thus, by forcing the utilization of the relay and equal time slot lengths, the information rate is almost half of the direct link capacity and the relay does not considerably contribute additional source information to the destination. Therefore, it can be seen, that the relay is only useful[1] if it is positioned somewhere between the source and the destination, which will eventually shorten the distance between the source and the destination. With these shortened distances and the corresponding higher signal-to-noise ratios, the relay can increase[2] the possible information rate. Choosing a good position, for example at position (0.5,0), the rate can be five times higher than for the direct link.

Loosing the constraint of equal time slot length, one can optimize the information rate by altering $t_1 = 1 - t_2$. Therefore, in areas where the relay does not contribute to the information rate, only the source should be selected for transmitting the

[1] Please note, that this statement will not hold for a relay network with an additional *automatic repeat request* (ARQ) process. Here, the relay can be beneficial at any position, depending on the ARQ mechanism.

[2] The improvement of the information rate depends on the reference SNR $\gamma_{\mathcal{SD}}$.

information. In this situation, one can set $t_1 = 1$ and $t_2 = 0$, which is basically turning the relay off and spending the complete time for the source. In other areas, it can be beneficial spending more time for the relay than for the source.

The challenge is to optimize the transmission time slots T_1 and T_2, which is a non-trivial task [HMZ05]. In fact, even for the simple three-node relay network no closed form solution exists for all relay protocols. Therefore, the optimization was done via the bisection or binary search method. The result is depicted in **Fig. 3.5**. It can be seen, that a variable time slot allocation can significantly improve

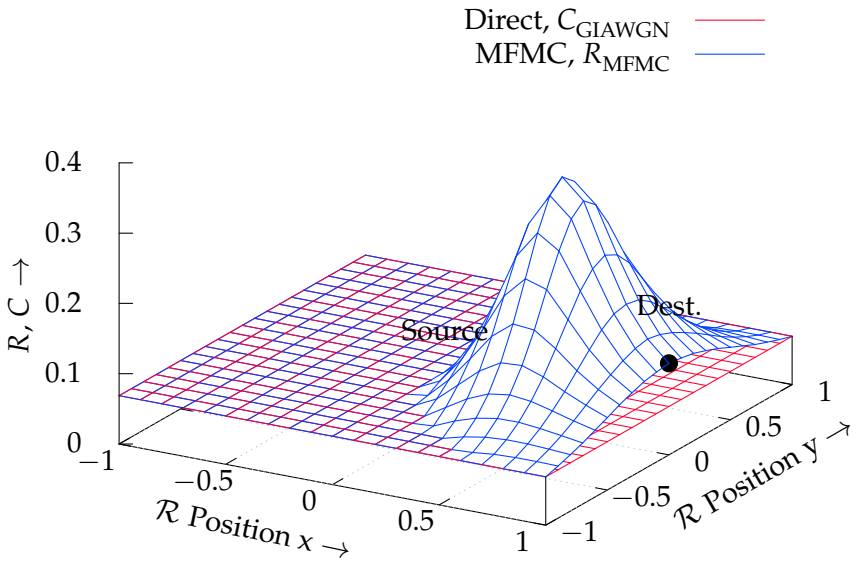


Fig. 3.5.: Rate of a three-node relay network with optimized time slot lengths and $\alpha = 4$, $\gamma_{\mathcal{SD}} = -10$dB. Position of the source and the destination nodes are fixed, while the relay's position was altered.

the information rate of the relay network. Here, the red plane for the capacity of the direct link is always equal or below the blue surface of the information rate. Comparing the information rates in **Fig. 3.4** with **Fig. 3.5**, the elevation in the latter one possesses a larger volume and, therefore, maintains an equal or higher rate at any relay position.

The result for the corresponding time slot lengths $t_1 = 1 - t_2$ is depicted in **Fig. 3.6**. So, depending on the relay's position, the source's transmit time slot T_1 will either be larger or smaller than the relay's transmit time slot T_2. Generally, for the given setup and simulation parameters, the relay is especially useful, if it is placed between the source and the destination or shortly behind the destination.

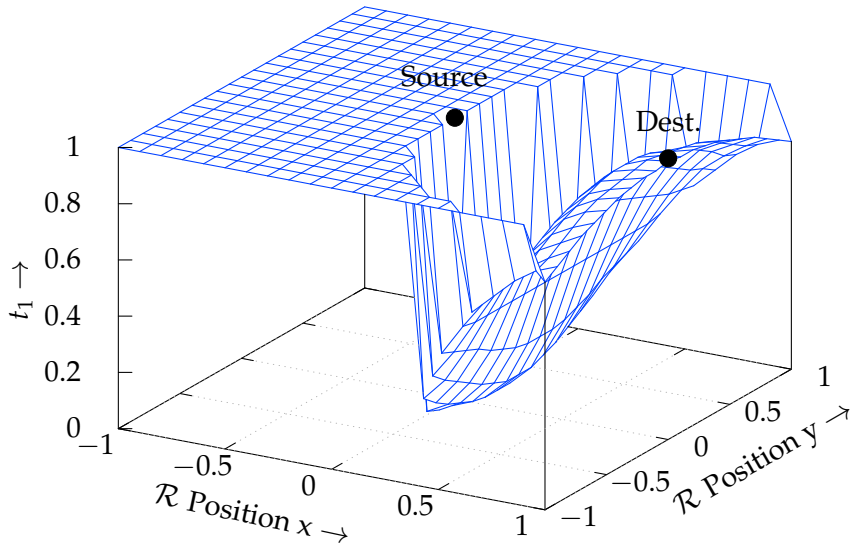


Fig. 3.6.: Time t_1 spend for the transmission. Only for a relay position between $\mathcal{S}$ and $\mathcal{D}$, the time $t_1 < 1$ which means that the relay is beneficial and is allowed to transmit.

The performance comparison of the different relaying protocols are depicted in **Figures 3.7**, **3.8** and **3.9**. Here, the relay's position was altered, moving the relay on a straight line from the source towards the destination. The parameters are equivalent as before, except that the path-loss exponent α and the SNR has been changed to discuss different aspects of the relaying protocols. Furthermore, the time slots lengths have either been optimized (opt. t_1) or were fixed ($t_1 = 1/2$). For better illustration, the curves for the resulting two sets of rates are separated into two graphs. Here, the top figure always shows the simulated rates for the optimized time slot lengths, whereas the bottom figure depicts the results for equal time slot lengths. Additional to these curves are results for relaying protocols which depend on equal time slot lengths and, thus, cannot be optimized. Among these protocols are Amplify-and-Forward, Compress-and-Forward and Decode-and-Forward with repetition coding.

In **Fig. 3.7**, the path-loss exponent and the SNR are given by $\alpha = 4$ and $10 \cdot \log_{10}(\gamma_{\mathcal{SD}}) = -10\text{dB}$. With this set of parameters, it can be seen that most relaying protocols deliver a rate above the capacity of the direct link from relay position $0 \leq d_{\mathcal{SR}} \leq 1$. The only exceptions are given by the Multi-Hop Transmission protocols and the Decode-and-Forward protocol shown in the bottom part of **Fig. 3.7**.

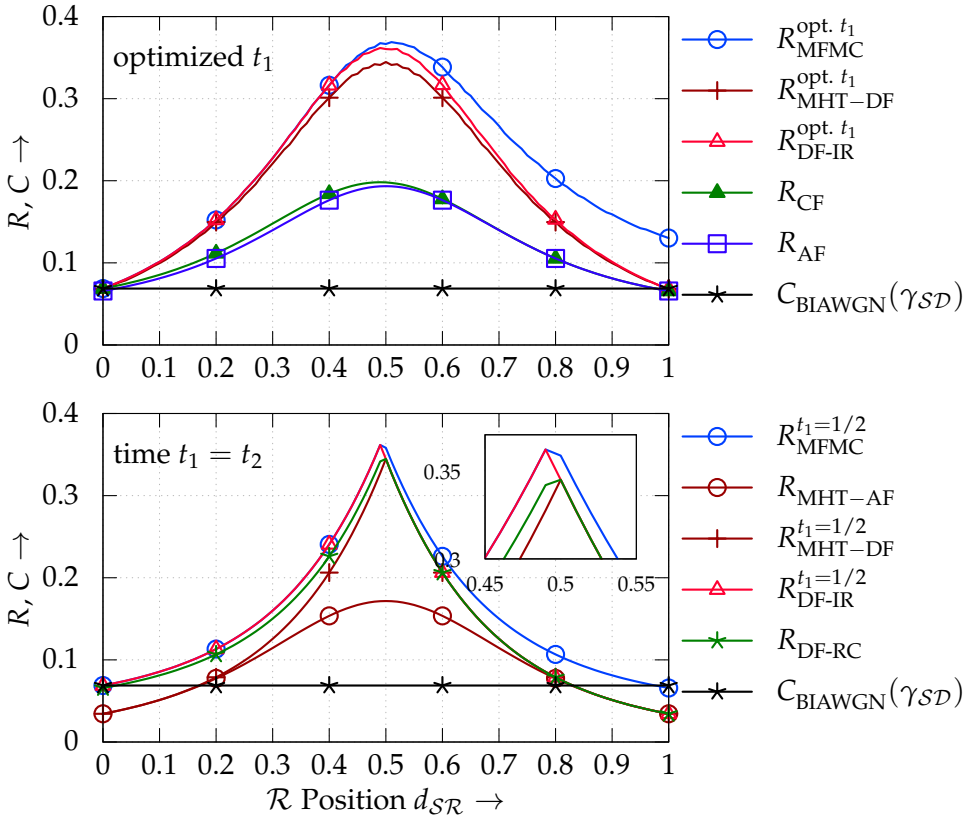


Fig. 3.7.: Different rates for the three-node relay network with $\alpha = 4$, $10 \cdot \log_{10}(\gamma_{\mathcal{SD}}) = -10$dB. The position of the relay was altered on a straight line between source and destination. Top: Results for optimized time slot lengths. Bottom: Results for equal time slot lengths.

Furthermore, please keep in mind that the optimized rate $R^{\text{opt. } t_1}_{\text{MFMC}}$, determined by the max-flow min-cut theorem, in the upper part of **Fig. 3.7**. always holds as the upper bound for all other optimized relay protocols. Equivalently, the rate $R^{t_1=1/2}_{\text{MFMC}}$ determines the upper bound for all non-optimized protocols, like CF, AF and all protocols in the lower part of **Fig. 3.7**. Obviously, it should be noted that the rates for optimized time slot lengths are always superior compared to rates of non-optimized time slot lengths and equal relay protocols.

Recapitulating equation (3.6) and (3.23), it can be seen that both include the term $t_1 \cdot I(X_{\mathcal{S}}; Y_{\mathcal{SD}}) + (1 - t_1) \cdot I(X_{\mathcal{R}}; Y_{\mathcal{RD}})$ in their rate calculation. Thus, the rates for MFMC and DF-IR (optimized and non-optimized) are equal as long as this part dominates the minimum function. This is the case for a good $\mathcal{S} \rightarrow \mathcal{R}$ link, i.e. the relay is close to the source. Once the relay moves towards the destination, the rate for DF-IR is limited by the capacity of the $\mathcal{S} \rightarrow \mathcal{R}$ link and $R_{\text{DF-IR}}$ diverges from the upper bound R_{MFMC}. This is due to the constraint that the relay has to perfectly decode the source's signal.

Since the same constraint is made for the DF-RC protocol, the rates $R_{\text{DF-RC}}$ and $R_{\text{DF-IR}}^{t_1=1/2}$ for DF-RC and the non-optimized DF-IR are, in this situation, equal (see lower part in **Fig. 3.7**). On the other side, in case of a good first hop (relay is close to the source) $R_{\text{DF-RC}} < R_{\text{DF-IR}}^{t_1=1/2}$ holds. Here, the available degree of freedom is not exploited by the DF-RC protocol and, therefore, the rate $R_{\text{DF-IR}}^{t_1=1/2}$ for the non-optimized Decode-and-Forward with incremental redundancy protocol is higher than the rate $R_{\text{DF-RC}}$ for the Decode-and-Forward with repetition coding protocol. Comparing equation (3.21) with (3.25) delivers the inequality

$$\frac{1}{2} \cdot C(\gamma_{\mathcal{SD}} + \gamma_{\mathcal{RD}}) \leq \frac{1}{2} \cdot C(\gamma_{\mathcal{SD}}) + \frac{1}{2} \cdot C(\gamma_{\mathcal{RD}}) \ , \tag{3.32}$$

given equal time slot lengths for both DF protocols.

Furthermore, the rate $R_{\text{DF-IR}}^{\text{opt. } t_1}$ is always equal and, in most relay positions $0 \leq d_{\mathcal{SR}} \leq 1$, above the capacity of the direct link. This is ensured by optimizing t_1 and, therewith, t_2. For example, if $d_{\mathcal{SR}} = 1$ then only the source will transmit, whereas the relay is quiet ($t_1 = 1$ and $t_2 = 0$) and, thus, Decode-and-Forward with incremental redundancy achieves the same rate as the capacity of the direct link. In case of non-optimized time slot lengths ($t_1 = t_2$), the rates $R_{\text{DF-IR}}^{t_1=1/2}$ and $R_{\text{DF-RC}}$ fall below the capacity of the direct link once the relay is close to the destination. In the special case of $d_{\mathcal{SR}} = 1$, the source will transmit half the time, whereas the relay is quiet during the second time slot. Thus, in this situation, the achievable rates from the DF-IR and DF-RC protocols are only half of the capacity of the direct link.

Comparing the rate $R_{\text{MHT-DF}}^{\text{opt. } t_1}$ for the Multi-Hop Transmission protocol with Decode-and-Forward and optimized t_1 with the rate $R_{\text{DF-IR}}^{\text{opt. } t_1}$, $R_{\text{MHT-DF}}^{\text{opt. } t_1}$ is always below $R_{\text{DF-IR}}^{\text{opt. } t_1}$. This is due to the missing direct link. But in case of equal time slot lengths, all rates $R_{\text{MHT-DF}}^{t_1=1/2}$, $R_{\text{DF-IR}}^{t_1=1/2}$ and $R_{\text{DF-RC}}^{t_1=1/2}$ are equal after the capacity of the $\mathcal{S} \to \mathcal{R}$ link is dominant.

Furthermore, the rate curves for Amplify-and-Forward (upper graph) and MHT-AF (lower graph) posses similar progress. Again, since the direct link is missing for the MHT protocol, the rate $R_{\text{MHT-AF}}$ is below R_{AF}. Also, please note that the time slot lengths cannot be optimized in case of MHT-AF. Interestingly, the rates for DF-IR and DF-RC are only slightly higher than for the MHT-DF protocol. Here, only in case of a good $\mathcal{S} \to \mathcal{R}$ link visible gains can be observed.

Unfortunately, for the complete range of relay positions, none of the introduced relaying protocols can reach the upper bounds $R_{\text{MFMC}}^{\text{opt. } t_1}$ (optimized t_1) and $R_{\text{MFMC}}^{t_1=1/2}$ (non-optimized t_1), given by the max-flow min-cut theorem. Only for certain relay positions, some of the relaying protocols are equal to the MFMC bound. Even worse, their exists a gap between all given rates and the upper bound in areas where

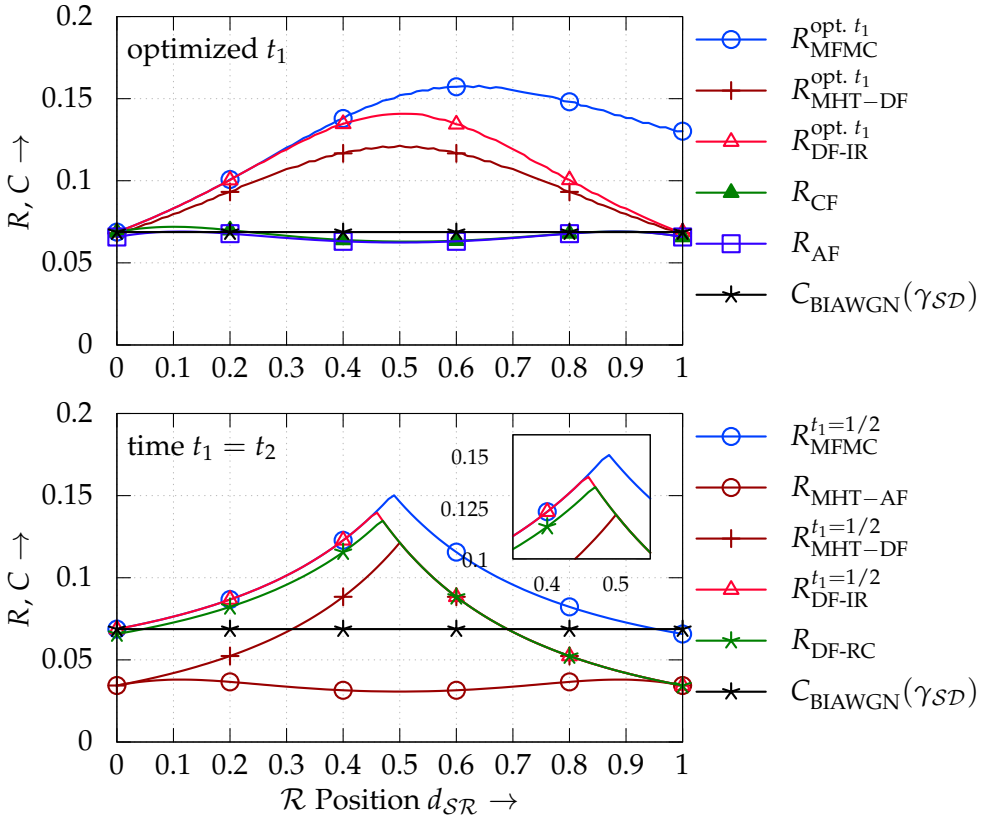


Fig. 3.8.: Different rates for the three-node relay network with $\alpha = 2$, $10 \cdot \log_{10}(\gamma_{\mathcal{SD}}) = -10$dB. The position of the relay was altered on a straight line between source and destination. Top: Results for optimized time slot lengths. Bottom: Results for equal time slot lengths.

the relay is moving away from the source and before it reaches the destination. Until now, there does not exist a relaying protocol which is known to close this gap. Depending on the underlying channel parameters this gap can be smaller or larger.

Fortunately in some areas and under special conditions, the upper bound can be reached by some known relaying protocols. For example, as depicted in **Fig. 3.7** Decode-and-Forward with incremental redundancy performs equally well until the relay reaches half the distance between the source and the destination. Then, no other relaying protocol gets close to $R^{t_1=1/2}_{\text{MFMC}}$ until the relay is near the destination. Here, CF and AF from the upper part of **Fig. 3.7** are reaching the upper bound for the non-optimized MFMC theorem from the lower part of **Fig. 3.7**.

In Chapter 5 a method is introduced to predict the achievable rate of the a three-node relay network with the help of the information combining technique and ideal codes. Although, a rate is determined which closes the gap towards the upper bound in certain areas, still the upper bound cannot be reached in all cases.

Changing the path-loss exponent from $\alpha = 4$ to $\alpha = 2$, one will obtain the results as depicted in **Fig. 3.8**. Comparing the rates from **Fig. 3.7** with **Fig. 3.8**, one will notice an decrease in throughput. This is remarkable since a change from $\alpha = 4$ to $\alpha = 2$, i.e. from a relative urban environment with lots of scattering and shadowing effects, to a free space propagation should retain a better rate performance. But, because of the normalized distance between source and destination ($d_{\mathcal{SD}} = 1$), a higher path-loss exponent results in a larger path-loss attenuation factor. Therefore, the simulated rates in case of $\alpha = 2$ are less than for $\alpha = 4$. In a simulation environment, which represents a *real world* scenario, the propositions for both scenarios will switch, but the general assumptions and results for the comparison still holds.

Looking at **Fig. 3.8**, the gap between the rates of the different relay protocols and the max-flow min-cut bound is increased, as compared to the results in **Fig. 3.7**. But, similar to **Fig. 3.7**, the rates of all optimized protocols are above the capacity of the direct link (for all relay positions). Furthermore, if the relay is close to the source, then the optimized and non-optimized DF-IR have the same rates as their counterparts for the MFMC upper bound, until the DF-IR rate performance is limited by the capacity of the first hop. Shortly after, the non-optimized DF-IR, DF-RC and MHT-DF will perform equally well. In the optimized case and due to the missing direct link, the rate for MHT-DF is still below the DF-IR rate over the complete range of $0 \leq d_{\mathcal{SR}} \leq 1$.

Comparing the rate R_{AF} for the Amplify-and-Forward protocol with the rate R_{CF} for the Compress-and-Forward protocol reveals that both perform nearly the same as long as the signal-to-noise ratio is low (see **Fig. 3.7** and **Fig. 3.8**). Once the SNR increases, both curves diverges and Compress-and-Forward is superior to Amplify-and-Forward (see **Fig. 3.9**). Interestingly, in **Fig. 3.8** the rates for AF, CF and MHT-DF will decrease with the relay moving towards the destination. Reaching half the distance the rate slightly increases before decreasing again.

For a given scenario with $\alpha = 2$ and a SNR of $10 \cdot \log_{10}(\gamma_{\mathcal{SD}}) = 10\text{dB}$, it can be seen in **Fig. 3.9** that the rates of most relaying protocols are below the capacity of the direct link. Only the optimized rates from the max-flow min-cut theorem and the Decode-and-Forward with incremental redundancy protocol can achieve a better performance throughout all relay positions.

In case of $t_1 = 1/2$ and a relay position close to the source, the rates from $R_{\text{MFMC}}^{t_1=1/2}$, $R_{\text{DF-IR}}^{t_1=1/2}$ and R_{CF} are slightly above the capacity of the direct link. Since in this situation the term $1/2 \cdot I(X_{\mathcal{S}}; Y_{\mathcal{SD}}) + 1/2 \cdot I(X_{\mathcal{R}}; Y_{\mathcal{RD}})$ from the MFMC theorem

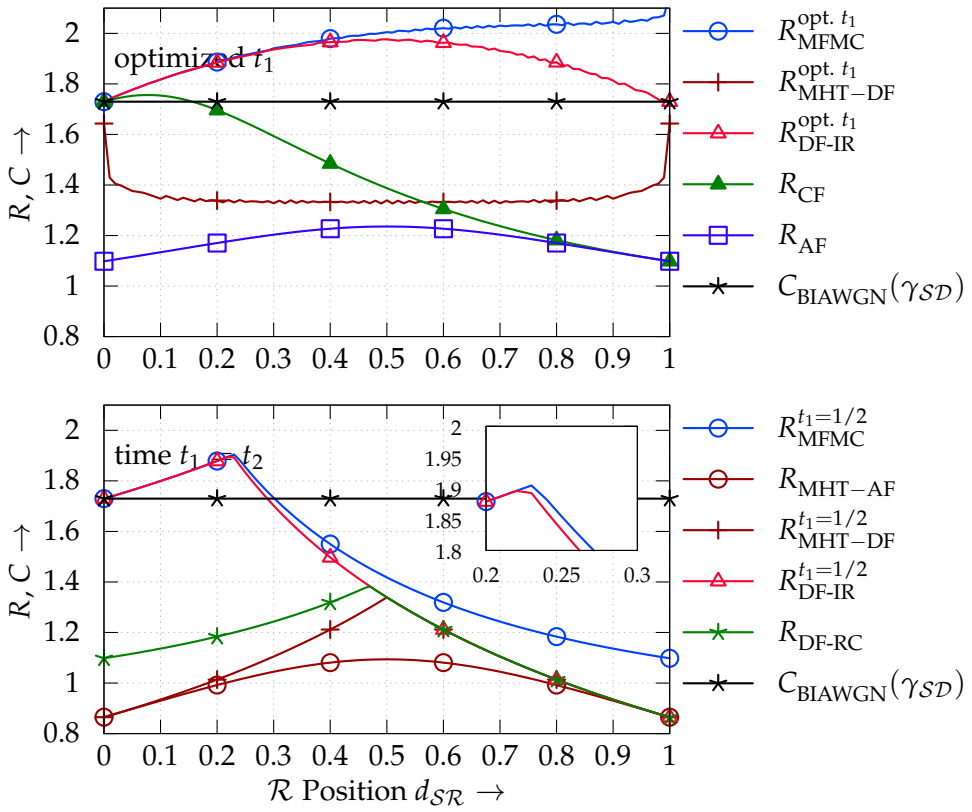


Fig. 3.9.: Different rates for the three-node relay network with $\alpha = 2$, $10 \cdot \log_{10}(\gamma_{\mathcal{SD}}) = 10\text{dB}$. The position of the relay was altered on a straight line between source and destination. Top: Results for optimized time slot lengths. Bottom: Results for equal time slot lengths.

and the DF-IR protocol is dominating the *min*-function and the capacity $C(\gamma_{\mathcal{RD}})$ of the $\mathcal{R} \rightarrow \mathcal{D}$ link is greater (or equal) than for the direct link, the inequality

$$1/2 \cdot C(\gamma_{\mathcal{SD}}) + 1/2 \cdot C(\gamma_{\mathcal{RD}}) \geq C(\gamma_{\mathcal{SD}}) \tag{3.33}$$

is applicable. This is also valid for **Fig. 3.7** and **Fig. 3.8**.

Once the relay increases its distance from the source, all non-optimized rates are below $C(\gamma_{\mathcal{SD}})$. In this situation the minimum function from equation (3.6), (3.20) and (3.23) is dominated by the capacity of the $\mathcal{S} \rightarrow \mathcal{R}$ link, or, in case of MFMC by the MI $I(X_{\mathcal{S}}; Y_{\mathcal{SR}}, Y_{\mathcal{SD}})$.

4 Functions for information concatenation

This chapter introduces functions for calculating the mutual information of serial and parallel concatenated channels, with the help of the *information combining technique* (ICT). These functions are essential for the upcoming Chapters 5 and 6 to determine the MI of orthogonal relay channels with ideal and practical coding schemes.

To do so, the *serial* and *parallel* information combining technique for symmetric memoryless channels is introduced and upper and lower bounds are presented. These bounds will partially coincidence to the upper and lower bounds, which have been already introduced by I. Land and J. Huber in [Lan05a] and [HL06] and were utilized, for example, in [LHHH05a, CMH10] to evaluate the performance of serial or parallel concatenated coding schemes.

4.1 Introduction

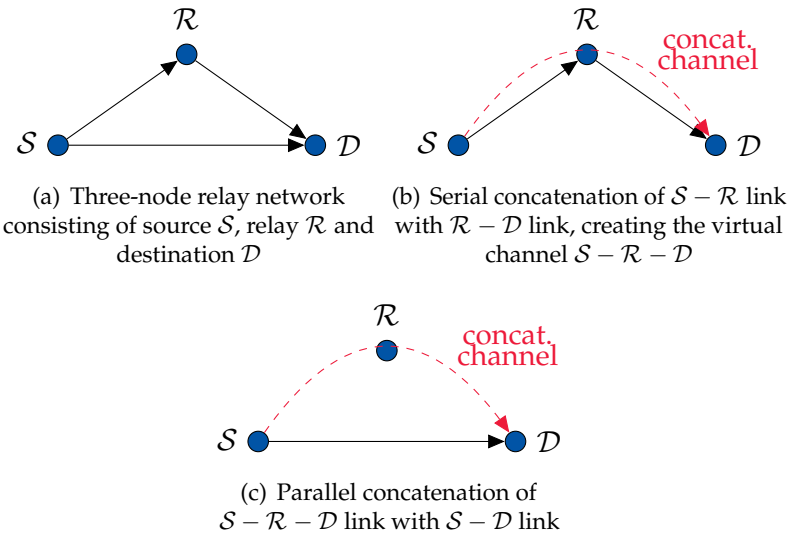
Calculating the true capacity of the general relay network is still an unsolved problem. However, upper bounds exists which where determined utilizing the max-flow min-cut theory as shown earlier in Section 3.2.1. The question which arises is *how to approach these bounds, especially with capacity achieving codes* and probably more importantly *how large is the gap utilizing non-ideal coding schemes*?

To answer these two questions, first, the information combining technique originally introduced by Ingmar Land et al. in [LHHH03, LHHH05a, Lan05b, HL06] and others will be utilized and extended in order to predict the mutual information of a relay network. The prediction is done for capacity achieving codes in the next chapter and for practical codes in Chapter 6.

This chapter is organized in three major parts. First, Section 4.2 gives an overview of the information combining technique, whereas the subsequent Sections 4.3 and 4.4 are delivering the analytical formulation for serial and parallel combining and concatenation of mutual information, including their lower and upper bounds.

4.2 The information combining technique

In order to estimate the mutual information of the complete relay network, i.e. from the input at the source towards the output of the destination, the network has to be decomposed into its different parts. This is depicted in **Fig. 4.1**. First, the mean



(a) Three-node relay network consisting of source $\mathcal{S}$, relay $\mathcal{R}$ and destination $\mathcal{D}$

(b) Serial concatenation of $\mathcal{S}-\mathcal{R}$ link with $\mathcal{R}-\mathcal{D}$ link, creating the virtual channel $\mathcal{S}-\mathcal{R}-\mathcal{D}$

(c) Parallel concatenation of $\mathcal{S}-\mathcal{R}-\mathcal{D}$ link with $\mathcal{S}-\mathcal{D}$ link

Fig. 4.1.: Simple three-node relay network with the two different types of concatenating mutual information

mutual information of the link $\mathcal{S} \rightarrow \mathcal{R} \rightarrow \mathcal{D}$ is calculated via the *serial information combining function* (SICF). Then, the outcome is parallel combined with the mutual information of the direct link $\mathcal{S} \rightarrow \mathcal{D}$ via the *parallel information combining function* (PICF), in order to get the complete mean MI for the network. Both methods, i.e. serial and parallel combining of mutual information for binary input symmetric memoryless channels [LHHH03, LHHH05a], are reviewed in the following two sections. Furthermore, the information combining technique will be extended for the general case of symmetric memoryless channels.

There are three major benefits of utilizing information combining for relay networks:

1. The mutual information of almost any orthogonal relay network, including the different relaying protocols, can be predicted. Sure, for simple networks and protocols like Decode-and-Forward or Amplify-and-Forward the rates can directly be calculated and the information combining technique will only increase the computational complexity. But for more advanced protocols and, more importantly, larger networks with many relaying nodes, the information combining technique can be utilized in a semi-analytical manner to precisely predict the maximum rate.

2. The information combining technique can help to predict the maximum rate in cases of erroneous relays. Thus it would be possible to find new relaying protocols, which maybe perform superior as the known protocols.

3. Especially, for relay networks with practical codes, the information combining technique in combination with the information processing characteristic can help to predict the behavior of such networks. Here, the prediction is less computational complex than with Monte Carlo simulations.

4.3 Serial information combining

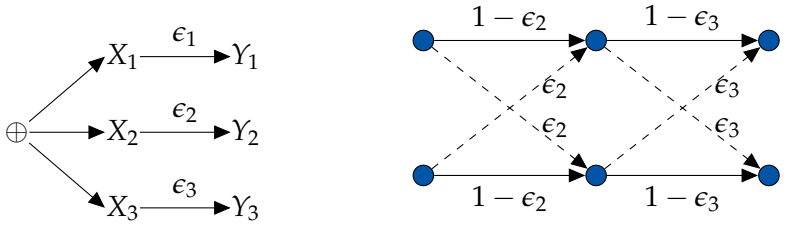
The serial information combining technique is motivated by the information combining for *single parity-check* (SPC) codes, described in [Lan05a] and [HL06]. Here, for example a SPC code of length 3 is defined by the check constraint between the code bits $X_k \in \{0, 1\}$ via

$$0 = X_1 \oplus X_2 \oplus X_3 \ .$$

Furthermore, the code bits will be transmitted over three independent binary symmetric channels[1] with corresponding error probability ϵ_i and channel outputs $Y_k \in \{0,1\}$. This is shown in **Fig. 4.2(a)**. Now, in order to calculate the extrinsic probability of $X_1 = X_2 \oplus X_3 = 0$, both symbols X_2 and X_3 must be either 0 or 1

$$\begin{aligned} \Pr_{\text{ext}} &= \Pr(X_1 = 0 | Y_2, Y_3) \\ &= \Pr(X_2 \oplus X_3 = 0 | Y_2, Y_3) \\ &= \epsilon_2 \cdot \epsilon_3 + (1-\epsilon_2) \cdot (1-\epsilon_3) \ . \end{aligned} \tag{4.1}$$

Looking at the last line in equation (4.1) reveals a serial concatenation of two BSCs, which is depicted in **Fig. 4.2(b)**. One can conclude that the parity-check constraint



(a) Single parity check code of length $N = 3$ [LHHH05b]

(b) Serial concatenation of two binary symmetric channels

Fig. 4.2.: Parity-check constraint versus serial concatenation

for a SPC code of any length forms a Markov chain which can be regarded as a serial concatenation of channels [HL06, p. 231]. The combining of mutual information for channels which fulfill a parity-check constraint is denoted as serial information combining.

Though, the multi-hop transmission in a relay network is a serial concatenation of channels that will not fulfill a check constraint. But, they also form a Markov chain for which the end-to-end MI can be calculated and, hence, the serial information combining was the initial point in utilizing the information combining technique for relay channels. Although, the area of application is different in both cases (relay networks vs. turbo codes), the mathematical description almost stays the same.

The only difference appears in the fact that the second channel input for two serial concatenated channels is not necessarily binary distributed. This permits the use of the original upper and lower bounds of the *serial information combining function* for discrete symmetric memoryless channel.

[1] For this example and without loss of generality, BPSK modulation and demodulation is neglected due to better understanding.

4.3.1 Serial information combining for SMCs

To motivate serial information combining of discrete symmetric memoryless channels, lets consider the 2-hop communication from $\mathcal{S}$ over $\mathcal{R}$ towards $\mathcal{D}$ which is shown in **Fig. 4.1(b)**. Every link distorts the source message and the question, which needs to be answered is, how much information from the source is received at the destination.

The problem is equivalent to the one that was given by the non-cooperative multi-hop transmission in Subsection 3.3.2. Here, a more general description, independent of specific channel realizations, is introduced. In the following subsections, this general description is then limited to cases with symmetric channels like BSCs and BECs, which in turn will deliver upper and lower bounds on the MI, respectively.

The serial concatenation of the two channels $X_\mathcal{S} \to Y_{\mathcal{SR}}$ with $X_\mathcal{R} \to Y_{\mathcal{RD}}$ (conditioned[2] on $X_\mathcal{R} = Y_{\mathcal{SR}}$) and the calculation of mutual information for this new *virtual* channel $X_\mathcal{S} \to X_\mathcal{R} \to Y_{\mathcal{RD}}$ is limited by the *data processing inequality* [CT06] given by $X_\mathcal{S} \perp Y_{\mathcal{RD}} | X_\mathcal{R}$, i.e. the conditional independence between $X_\mathcal{S}$ and $Y_{\mathcal{RD}}$ given $X_\mathcal{R}$. It loosely states, that no post-processing of the channel output $Y_{\mathcal{RD}}$ can increase the mutual information $I(X_\mathcal{S}; Y_{\mathcal{RD}})$ above the value $I(X_\mathcal{S}; X_\mathcal{R})$. Therefore $I(X_\mathcal{S}; Y_{\mathcal{RD}}) \leq I(X_\mathcal{S}; X_\mathcal{R})$ and $I(X_\mathcal{S}; Y_{\mathcal{RD}}) \leq I(X_\mathcal{R}; Y_{\mathcal{RD}})$ holds. Thus, the maximum mutual information between the channel input $X_\mathcal{S}$ and the channel output $Y_{\mathcal{RD}}$ is upper bounded by

$$I(X_\mathcal{S}; Y_{\mathcal{RD}}) = \min \{ I(X_\mathcal{S}; X_\mathcal{R}), I(X_\mathcal{R}; Y_{\mathcal{RD}}) \} \quad . \tag{4.2}$$

Equation (4.2) requires the relay and the destination to perfectly decode the received sequence. This implies that the maximum rate, with which the source transmits is at most equal or below $I(X_\mathcal{S}; Y_{\mathcal{RD}})$. In the relay case, it might be of interest to transmit with rates above $I(X_\mathcal{S}; Y_{\mathcal{RD}})$ since the mutual information of the direct link can be exploited as well. Although the direct and relay link can be erroneous, combining them maybe lead to an error-free communication. Thus, the rate at which the source transmits can be greater than the individual channel capacities. This was already stated by the max-flow min-cut theorem.

Now, the question which needs to be answered is, what is the maximum amount of mutual information $I(X_\mathcal{S}; Y_{\mathcal{RD}})$ for the serial concatenated $\mathcal{S} \to \mathcal{R} \to \mathcal{D}$ link, in case of an erroneous relay node and due upper and lower bounds on that MI exist.

[2] In this case, the relay only forwards the received signal without any further processing and, thus, no additional calculation has to be considered for the serial concatenation. Of course, in the upcoming chapters the relay's signal processing will be included.

To obtain $I(X_\mathcal{S}; Y_{\mathcal{RD}})$, the joint probability mass function $p_{X_\mathcal{S}, Y_{\mathcal{RD}}}(x_\mathcal{S}, y_{\mathcal{RD}})$ has to be evaluated. This is given by

$$\begin{aligned} p_{X_\mathcal{S}, Y_{\mathcal{RD}}}(x_\mathcal{S}, y_{\mathcal{RD}}) &= \sum_{x_\mathcal{R} \in \mathbb{X}_\mathcal{R}} p_{X_\mathcal{S}, X_\mathcal{R}, Y_{\mathcal{RD}}}(x_\mathcal{S}, x_\mathcal{R}, y_{\mathcal{RD}}) \\ &= \sum_{x_\mathcal{R} \in \mathbb{X}_\mathcal{R}} p_{X_\mathcal{S}}(x_\mathcal{S}) \cdot p_{X_\mathcal{R}|X_\mathcal{S}}(x_\mathcal{R}|x_\mathcal{S}) \cdot p_{Y_{\mathcal{RD}}|X_\mathcal{R}, X_\mathcal{S}}(y_{\mathcal{RD}}|x_\mathcal{R}, x_\mathcal{S}) \\ &= \sum_{x_\mathcal{R} \in \mathbb{X}_\mathcal{R}} p_{X_\mathcal{S}}(x_\mathcal{S}) \cdot p_{X_\mathcal{R}|X_\mathcal{S}}(x_\mathcal{R}|x_\mathcal{S}) \cdot p_{Y_{\mathcal{RD}}|X_\mathcal{R}}(y_{\mathcal{RD}}|x_\mathcal{R}) \\ &= p_{X_\mathcal{S}}(x_\mathcal{S}) \cdot p_{Y_{\mathcal{RD}}|X_\mathcal{S}}(y_{\mathcal{RD}}|x_\mathcal{S}) \ , \end{aligned} \tag{4.3}$$

i.e. summing the joint pmf $p_{X_\mathcal{S}, X_\mathcal{R}, Y_{\mathcal{RD}}}(x_\mathcal{S}, x_\mathcal{R}, y_{\mathcal{RD}})$ over $X_\mathcal{R}$, applying Bayes rule and the conditional independence $X_\mathcal{S} \perp Y_{\mathcal{RD}}|X_\mathcal{R}$ of the Markov chain, where $p_{Y_{\mathcal{RD}}|X_\mathcal{R}}(y_{\mathcal{RD}}|x_\mathcal{R}) = p_{Y_{\mathcal{RD}}|X_\mathcal{R}, X_\mathcal{S}}(y_{\mathcal{RD}}|x_\mathcal{R}, x_\mathcal{S})$ holds.

Next, (4.3) can be inserted in (2.6) or (2.10) to retain $I(X_\mathcal{S}; Y_{\mathcal{RD}})$ for example via

$$\begin{aligned} I(X_\mathcal{S}; Y_{\mathcal{RD}}) &= \mathcal{I}^{\text{ser}}\left(p_{X_\mathcal{S}}(x_\mathcal{S}), p_{X_\mathcal{R}|X_\mathcal{S}}(x_\mathcal{R}|x_\mathcal{S}), p_{Y_{\mathcal{RD}}|X_\mathcal{R}}(y_{\mathcal{RD}}|x_\mathcal{R})\right) \\ &= \sum_{y_{\mathcal{RD}} \in \mathbb{Y}_{\mathcal{RD}}} \sum_{x_\mathcal{S} \in \mathbb{X}_\mathcal{S}} p_{X_\mathcal{S}, Y_{\mathcal{RD}}}(x_\mathcal{S}, y_{\mathcal{RD}}) \log_2 \frac{p_{X_\mathcal{S}, Y_{\mathcal{RD}}}(x_\mathcal{S}, y_{\mathcal{RD}})}{p_{X_\mathcal{S}}(x_\mathcal{S}) p_{Y_{\mathcal{RD}}}(y_{\mathcal{RD}})} \ . \end{aligned} \tag{4.4}$$

In equation (4.4), the function $\mathcal{I}^{\text{ser}}(\ldots)$ is called *serial information combining function* (SICF) and determines the overall mutual information for the serial concatenation of two channels. Please note that the function input parameters can change depending on the type of channels which are connected and if the upper or lower bound is evaluated. This will be made clear in the upcoming subsections.

4.3.2 Upper and lower bound for serial information combining

In Section 4.3.1, the general description of calculating the mutual information for serial concatenated channels was given. This description did not depend on any specific input or output distribution and, thus, holds for any memoryless channel with discrete arbitrary distributions.

In cases, where the channels have binary inputs, one can easily derive upper and lower bounds for the end-to-end MI of serial concatenated channels. This might be of interest in cases where channels with ideal or good codes are concatenated, since these codes always deliver a binary decision. Also, in cases where intermediate nodes perform a hard decision on the information sequence before processing and forwarding them towards other stations, binary channel inputs are created. In

all cases, the underlying memoryless channel can be arbitrary but the end-to-end channel (including encoder and decoder) must have binary inputs.

Upper bound for serial information combining

The proof[3] for the upper bound in mutual information of serial concatenated symmetric channels is inspired by the generalization [CS89][RU08, p. 242] of Mrs. Gerber's lemma [WZ73][Wyn73]. A different approach was taken in [Lan05b]. Both, [RU08] and [Lan05b] proofed the existence of an upper bound for two independent BISMCs, which are connected via a (single parity) check-equation. Since, this check-equation can be regarded as a serial connection (see Section 4.3) the proof will also be valid in cases of serially connected BISMCs. As the generalized Mrs. Gerber's

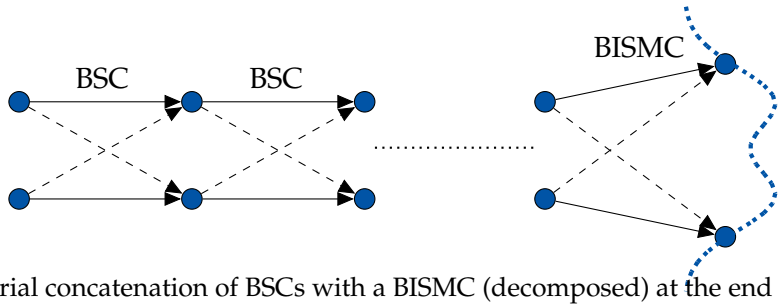


Fig. 4.3.: Serial concatenation of BSCs with a BISMC (decomposed) at the end

lemma relies on channels with binary inputs, the first and all subsequent serial connected channels must be binary symmetric channels with the exception of a BISMC for the last hop as shown in **Fig. 4.3**. Thus, the proof via Mrs. Gerber's lemma in the context of serial concatenation of channels is limited to this special case only.

However, in the case of two and more serially concatenated channels, the output of the preceding hop determines the input of the subsequent one. Thus, the subsequent hop is dependent upon the preceding one and the subsequent channel can be a symmetric memoryless channel with an arbitrary input. The general formulation for the true mutual information for the case of two channels was already stated in Subsection 4.3.1. For the upper bound in MI of a BISMC serially connected with a SMC no such proof exists and, here, only a conjecture can be stated. Please note, that the following description is given for the case of two serial connected channels. The generalization towards more than two channels can be performed by recursion and will be neglected here.

[3] The proof only holds for the serial connection of symmetric binary-input binary-output channels followed by a possible binary input symmetric memoryless channel at the end. For other types of symmetric channels, only a conjecture can be stated.

The conjecture is based on different assumptions and facts. For example, it is proven that every binary input symmetric memoryless channel can be decomposed into multiple BSC sub-channels [HL06], each with a different crossover-probability $\epsilon_y = \Pr(Y = +|y||X = -1) = \Pr(Y = -|y||X = +1)$, $y \in \mathbb{Y}$ and corresponding channel capacity (see **Fig. 2.3** on page 14)

$$C_{\mathrm{BSC},y} = 1 + \epsilon_y \log_2 \epsilon_y + (1 - \epsilon_y) \log_2 (1 - \epsilon_y) \ . \tag{4.5}$$

The expectation of all BSC-capacities in (4.5) will eventually result in the capacity of the original BISMC

$$C_{\mathrm{BISMC}} = \sum_{y \in \mathbb{Y}} p_Y(y) C_{\mathrm{BSC},y} \ . \tag{4.6}$$

So, in the context of connecting two symmetric channels with a binary input at the source, the first hop and the overall channel can be decomposed into multiple BSCs. The problem resides in the possibility of decomposing the second and all subsequent hops, which are eventually SMCs. Obviously, decomposing a SMC into multiple BSC sub-channels (see **Fig. 4.4**) and taking the expectation over the capacity of these sub-channels will most likely not deliver the same capacity as the original channel. For example, the expectation of the BSC sub-channels capacities can at maximum only be one, whereas the channel capacity of the original SMC can be greater ($C_{\mathrm{SMC}} > 1$).

To justify the following conjecture, one can deduce that the mutual information of the overall end-to-end channel is restricted by the binary input at the source. Thus, the MI of the end-to-end channel cannot be greater than one.

If the SMC with arbitrary input could be decomposed into multiple BSCs, then Mrs. Gerber's lemma is also applicable on the BSC sub-channel bases as shown in **Fig. 4.4**.

[4] In general, symmetric memoryless channels cannot be decomposed into binary symmetric channels with an equivalent expected channel capacity. Please, refer to the text for an justification in the specific case of a BIAWGN channel at the beginning.

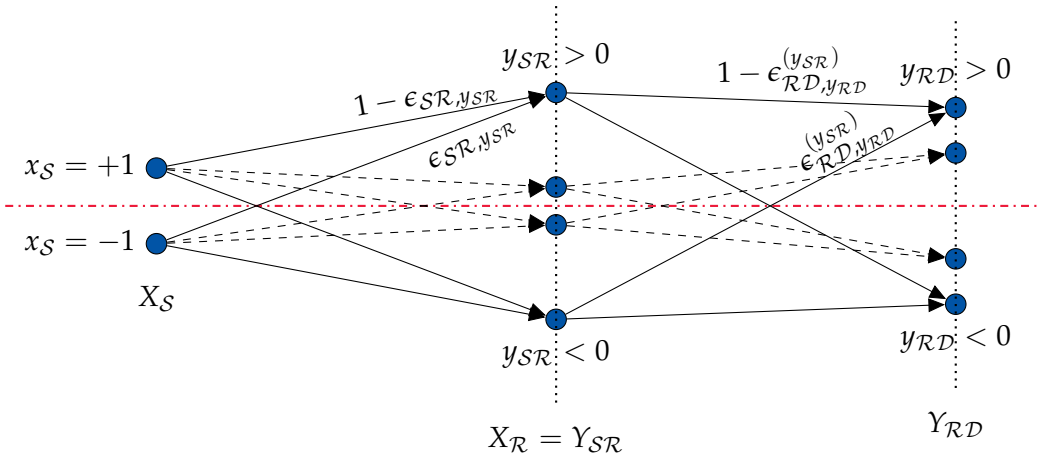


Fig. 4.4.: Decomposition of, for example, binary input additive white Gaussian noise and symmetric memoryless channel [4]into binary symmetric channels and corresponding crossover probabilities. After decomposition, the generalized Mrs. Gerber's lemma is potentially applicable.

Here, the first hop (BISMC) is decomposed [KV09b] into a set of BSCs with crossover-probabilities $\epsilon_{\mathcal{SR},y_{\mathcal{SR}}}$, $y_{\mathcal{SR}} \in \mathbb{Y}_{\mathcal{SR}}$ according to

$$\epsilon_{\mathcal{SR},y_{\mathcal{SR}}} = \begin{cases} \frac{1}{1+\frac{p_{Y_{\mathcal{SR}}|X_{\mathcal{S}}}(y_{\mathcal{SR}}|x_{\mathcal{S}}=+1)}{p_{Y_{\mathcal{SR}}|X_{\mathcal{S}}}(y_{\mathcal{SR}}|x_{\mathcal{S}}=-1)}} & \text{if } y_{\mathcal{SR}} > 0 \\ \frac{1}{1+\frac{p_{Y_{\mathcal{SR}}|X_{\mathcal{S}}}(y_{\mathcal{SR}}|x_{\mathcal{S}}=-1)}{p_{Y_{\mathcal{SR}}|X_{\mathcal{S}}}(y_{\mathcal{SR}}|x_{\mathcal{S}}=+1)}} & \text{if } y_{\mathcal{SR}} < 0 \\ 0.5 & \text{if } y_{\mathcal{SR}} = 0 \end{cases} . \tag{4.7}$$

Now, each symmetric channel output pair of the first hop determines a sub-binary input symmetric memoryless channel at the second hop. Therefore, given this binary input, the second hop can also be decomposed into multiple BSCs with crossover-probabilities $\epsilon^{(y_{\mathcal{SR}})}_{\mathcal{RD},y_{\mathcal{RD}}}$, $y_{\mathcal{RD}} \in \mathbb{Y}_{\mathcal{RD}}$ according to

$$\epsilon^{(y_{\mathcal{SR}})}_{\mathcal{RD},y_{\mathcal{RD}}} = \begin{cases} \frac{1}{1+\frac{p_{Y_{\mathcal{RD}}|Y_{\mathcal{SR}}}(y_{\mathcal{RD}}|y_{\mathcal{SR}}>0)}{p_{Y_{\mathcal{RD}}|Y_{\mathcal{SR}}}(y_{\mathcal{RD}}|y_{\mathcal{SR}}<0)}} & \text{if } y_{\mathcal{RD}} > 0 \\ \frac{1}{1+\frac{p_{Y_{\mathcal{RD}}|Y_{\mathcal{SR}}}(y_{\mathcal{RD}}|y_{\mathcal{SR}}<0)}{p_{Y_{\mathcal{RD}}|Y_{\mathcal{SR}}}(y_{\mathcal{RD}}|y_{\mathcal{SR}}>0)}} & \text{if } y_{\mathcal{RD}} < 0 \\ 0.5 & \text{if } y_{\mathcal{SR}} = 0 \text{ or } y_{\mathcal{RD}} = 0 \end{cases} , \tag{4.8}$$

depending on the first channel output $y_{\mathcal{SR}} = x_{\mathcal{R}}$.

Once, both channels have been decomposed into sub-channels via (4.7) and (4.8), Mrs. Gerber's lemma can be applied through

$$\begin{aligned}\epsilon_{\mathcal{SRD},y_{\mathcal{RD}}} \geq \sum_{y_{\mathcal{SR}}\in\mathbb{Y}_{\mathcal{SR}}} & \frac{p_{Y_{\mathcal{SR}}}(y_{\mathcal{SR}})}{p_{Y_{\mathcal{RD}}}(y_{\mathcal{RD}})} \\ & \cdot \frac{1}{2}\cdot\left(p_{Y_{\mathcal{RD}}|Y_{\mathcal{SR}}}(y_{\mathcal{RD}}|y_{\mathcal{SR}}) + p_{Y_{\mathcal{RD}}|Y_{\mathcal{SR}}}(y_{\mathcal{RD}}|-y_{\mathcal{SR}})\right) \\ & \cdot \left[(1-\epsilon_{\mathcal{SR},y_{\mathcal{SR}}})\cdot\epsilon^{(y_{\mathcal{SR}})}_{\mathcal{RD},y_{\mathcal{RD}}} + \epsilon_{\mathcal{SR},y_{\mathcal{SR}}}\cdot(1-\epsilon^{(y_{\mathcal{SR}})}_{\mathcal{RD},y_{\mathcal{RD}}})\right] \ . \end{aligned} \tag{4.9}$$

Here, the individual crossover-probabilities from the sub-channels of the first hop are serially concatenated with the crossover-probabilities for every sub-channel from the second hop. According to Mrs. Gerber's lemma, the entropy at the input of a binary channel lower bounds the entropy at the output. In the case of serial channel concatenation, the crossover-probability $\epsilon_{\mathcal{SRD},y_{\mathcal{RD}}}$ is lower bounded as shown in (4.9).

Therefore, after calculating the mutual information per concatenated sub-channel with

$$\begin{aligned} I^{\text{BSC}}_{y_{\mathcal{RD}}}(X_{\mathcal{S}};Y_{\mathcal{RD}}) = & \\ 1+\epsilon_{\mathcal{SRD},y_{\mathcal{RD}}}\log_2\epsilon_{\mathcal{SRD},y_{\mathcal{RD}}} & + (1-\epsilon_{\mathcal{SRD},y_{\mathcal{RD}}})\log_2(1-\epsilon_{\mathcal{SRD},y_{\mathcal{RD}}}) \ , \end{aligned} \tag{4.10}$$

the MI of the serial concatenation is upper bounded by the expectation of (4.10). Finally, the expectation over $p_{Y_{\mathcal{RD}}}(y_{\mathcal{RD}})$ of the concatenated sub-channels is given by

$$I^{\text{SMC}}(X_{\mathcal{S}};Y_{\mathcal{RD}}) = \sum_{y_{\mathcal{RD}}\in\mathbb{Y}_{\mathcal{RD}}} p_{Y_{\mathcal{RD}}}(y_{\mathcal{RD}})\, I^{\text{BSC}}_{y_{\mathcal{RD}}}(X_{\mathcal{S}};Y_{\mathcal{RD}}) \tag{4.11}$$

and the upper bound is donated by

$$\begin{aligned} I(X_{\mathcal{S}};Y_{\mathcal{RD}}) &\leq \mathcal{I}^{\text{ser,up}}_{\text{SMC}}(I(X_{\mathcal{S}};X_{\mathcal{R}}), I(X_{\mathcal{R}};Y_{\mathcal{RD}})) \\ &= I^{\text{SMC}}(X_{\mathcal{S}};Y_{\mathcal{RD}}) \ . \end{aligned} \tag{4.12}$$

Please note, although MIs are function parameters of $\mathcal{I}^{\text{ser,up}}_{\text{SMC}}(I(X_{\mathcal{S}};X_{\mathcal{R}}), I(X_{\mathcal{R}};Y_{\mathcal{RD}}))$, the calculation is performed via the probability distributions of the two corresponding channels. Furthermore, as stated earlier, decomposing the first hop and taking the expectation as shown in (4.5) and (4.6) will lead to $I(X_{\mathcal{S}};X_{\mathcal{R}})$.

Considering, that channel input and output between source and relay, as well as relay and destination are binary with alphabet $\{-1,\ +1\}$, then the concatenation of both channels is depicted in **Fig. 4.2(b)**.

Here, two BSCs are serially concatenated, where each BSC has uniformly and independent distributed binary inputs and binary outputs with their corresponding random variables $X_{\mathcal{S}}$, $X_{\mathcal{R}}$ and $Y_{\mathcal{RD}}$. It is still assumed, that the relay's output $X_{\mathcal{R}}$ equals the relay's input $Y_{\mathcal{SR}}$, i.e. $X_{\mathcal{R}} = Y_{\mathcal{SR}}$. Furthermore, the crossover (error) probabilities of both BSCs are denoted by $\epsilon_{\mathcal{SR}}$ and $\epsilon_{\mathcal{RD}}$, respectively.

It is obvious that the concatenation of both BSCs results in an overall BSC with input $X_{\mathcal{S}}$ and output $Y_{\mathcal{RD}}$ and crossover probability $\epsilon_{\mathcal{SRD}}$

$$\epsilon_{\mathcal{SRD}} = (1 - \epsilon_{\mathcal{SR}}) \cdot \epsilon_{\mathcal{RD}} + \epsilon_{\mathcal{SR}} \cdot (1 - \epsilon_{\mathcal{RD}}) \ . \tag{4.13}$$

The mutual information for this virtual channel is calculated with the help of the binary entropy function $H_b(\epsilon_{\mathcal{SRD}})$ from Section 2.1.3 via

$$\begin{aligned} I(X_{\mathcal{S}}; Y_{\mathcal{RD}}) &= \mathcal{I}_{\text{BSC}}^{\text{ser,up}}(\epsilon_{\mathcal{SR}}, \epsilon_{\mathcal{RD}}) \\ &= 1 - [-\epsilon_{\mathcal{SRD}} \cdot \log_2(\epsilon_{\mathcal{SRD}}) - (1 - \epsilon_{\mathcal{SRD}}) \cdot \log_2(1 - \epsilon_{\mathcal{SRD}})] \ . \end{aligned} \tag{4.14}$$

This result corresponds to the case of information combining of two BSCs for a *single parity-check* (SPC) code as derived in [HL06]. Furthermore, equation (4.14) holds as the upper bound on mutual information for serial concatenated binary symmetric channels [LHHH05a, Theorem 3].

If only the mutual informations of the separate BSC links are known, the relationship $I = 1 - H_b(\epsilon)$ and its inverse $\epsilon = H_b^{-1}(1 - I)$ between the mutual information I and crossover probability ϵ, can be utilized.

Lower bound for serial information combining

According to the data processing theorem, the end-to-end mutual information $I(X_{\mathcal{S}}; Y_{\mathcal{RD}})$ is always smaller or equal to the MI of the individual channels $I(X_{\mathcal{S}}; X_{\mathcal{R}})$ and $I(X_{\mathcal{R}}; Y_{\mathcal{RD}})$. Thus, $I(X_{\mathcal{S}}; Y_{\mathcal{RD}}) \leq I(X_{\mathcal{S}}; X_{\mathcal{R}})$ and $I(X_{\mathcal{S}}; Y_{\mathcal{RD}}) \leq I(X_{\mathcal{R}}; Y_{\mathcal{RD}})$ holds. In case of symmetric channels with binary input, the MI is greater or equal than zero and smaller or equal to one, i.e. $0 \leq I(X_{\mathcal{S}}; X_{\mathcal{R}}) \leq 1$ and $0 \leq I(X_{\mathcal{R}}; Y_{\mathcal{RD}}) \leq 1$ is given. It is now shown that

$$\begin{aligned} I(X_{\mathcal{S}}; Y_{\mathcal{RD}}) &\geq \mathcal{I}^{\text{ser,low}}(I(X_{\mathcal{S}}; X_{\mathcal{R}}), I(X_{\mathcal{R}}; Y_{\mathcal{RD}})) \\ &= I(X_{\mathcal{S}}; X_{\mathcal{R}}) \cdot I(X_{\mathcal{R}}; Y_{\mathcal{RD}}) \ , \end{aligned} \tag{4.15}$$

i.e. the MI for the overall channel is always greater or equal to the product of the MI of the individual channels. From (4.15) follows that $I(X_{\mathcal{S}}; X_{\mathcal{R}}) \leq \frac{I(X_{\mathcal{S}}; Y_{\mathcal{RD}})}{I(X_{\mathcal{R}}; Y_{\mathcal{RD}})} \leq 1$. The

converse $I(X_\mathcal{S};X_\mathcal{R}) > \frac{I(X_\mathcal{S};Y_{\mathcal{RD}})}{I(X_\mathcal{R};Y_{\mathcal{RD}})} > 1$ is a violation of the constraints. Equivalently to (4.15), $I(X_\mathcal{R};Y_{\mathcal{RD}}) \leq \frac{I(X_\mathcal{S};Y_{\mathcal{RD}})}{I(X_\mathcal{S};X_\mathcal{R})} \leq 1$ also holds.

Hence, the mutual information of two serial concatenated binary symmetric channels is lower bounded by the product of the MI from the individual channels. Furthermore, the lower bound of (4.15) is equivalent to the lower bound of mutual information for serial concatenated binary input symmetric memoryless channels as stated in [LHHH05a, Theorem 3].

4.4 Parallel information combining

Equivalently to the serial information combining technique, which was motivated by the information combining for single parity-check codes, the parallel information combining technique is motivated by the information combining for *repetition* codes [Lan05a]. One can imagine that the relay in **Fig. 4.1(a)** has perfectly recovered the source's information before retransmission and, thus, the destination receives two estimates of the same message over two different channels. This is depicted in **Fig. 4.1(c)**.

4.4.1 Parallel information combining for SMCs

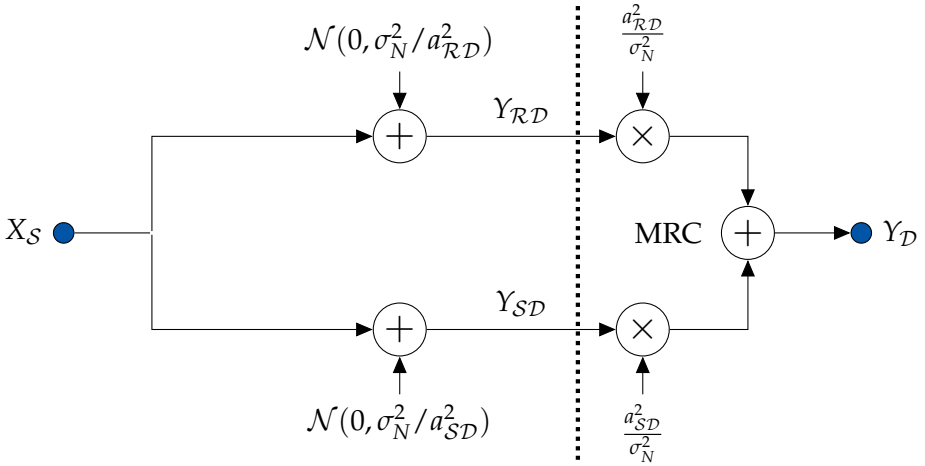


Fig. 4.5.: Optimal parallel concatenation of two SMCs, performing maximum ratio combining

In **Fig. 4.5**, the parallel concatenation of two channels is shown. This can be regarded as Decode-and-Forward with repetition coding, since the input $X_\mathcal{S}$ is transmitted over two independent links with independent noise realizations. The optimum combining for the considered case, given discrete symmetric memoryless channels with additive white Gaussian noise, is performed by *maximum ratio combining*

[Bre03] the channel outputs $Y_{\mathcal{SD}}$ and $Y_{\mathcal{RD}}$, for the $\mathcal{S} \to \mathcal{D}$ and the $\mathcal{S} \to \mathcal{R} \to \mathcal{D}$ links, via

$$Y_{\mathcal{D}} = \frac{a_{\mathcal{SD}}^2}{\sigma_N^2} \cdot Y_{\mathcal{SD}} + \frac{a_{\mathcal{RD}}^2}{\sigma_N^2} \cdot Y_{\mathcal{RD}} \ . \tag{4.16}$$

In (4.16), it was assumed, that the relay has perfect knowledge over $X_{\mathcal{S}}$ and, thus, transmits the same information as the source. Therefore, only the noise at the destination, i.e. the noise on the second hop for the $\mathcal{S} \to \mathcal{R} \to \mathcal{D}$ link, needs to be considered during the MRC. This assumption was taken in order to avoid multiple descriptions of the optimum combining, which clearly depends on the processing inside the relay node. For example, if the relay performs DF-RC and, additionally, allowing the relay to transmit even in the case of decoding errors, then the optimal combination also depends on these errors. This problem will be addressed in the upcoming Chapter 6.

The mutual information $I(X_{\mathcal{S}}; Y_{\mathcal{D}})$ for (4.16), between the channel input $X_{\mathcal{S}}$ and the combined channel output $Y_{\mathcal{D}}$, is given by

$$I(X_{\mathcal{S}}; Y_{\mathcal{D}}) = \sum_{y_{\mathcal{D}} \in \mathbb{Y}_{\mathcal{D}}} \sum_{x_{\mathcal{S}} \in \mathbb{X}_{\mathcal{S}}} p_{X_{\mathcal{S}}, Y_{\mathcal{D}}}(x_{\mathcal{S}}, y_{\mathcal{D}}) \log_2 \frac{p_{X_{\mathcal{S}}, Y_{\mathcal{D}}}(x_{\mathcal{S}}, y_{\mathcal{D}})}{p_{X_{\mathcal{S}}}(x_{\mathcal{S}}) p_{Y_{\mathcal{D}}}(y_{\mathcal{D}})} \ , \tag{4.17}$$

depending on the joint and marginal probability mass functions $p_{X_{\mathcal{S}}, Y_{\mathcal{D}}}(x_{\mathcal{S}}, y_{\mathcal{D}})$, with

$$p_{X_{\mathcal{S}}, Y_{\mathcal{D}}}(x_{\mathcal{S}}, y_{\mathcal{D}}) = p_{X_{\mathcal{S}}}(x_{\mathcal{S}}) \cdot p_{Y_{\mathcal{D}}|X_{\mathcal{S}}}(y_{\mathcal{D}}|x_{\mathcal{S}}) \ , \tag{4.18}$$

and $p_{X_{\mathcal{S}}}(x_{\mathcal{S}})$ for the channel input and $p_{Y_{\mathcal{D}}}(y_{\mathcal{D}})$ for the channel output, respectively. Utilizing the shorthand notation (A.24) and (A.22) for the scaling and the sum of random variables from appendix A.4.6 and A.4.4, the conditional probability mass function $p_{Y_{\mathcal{D}}|X_{\mathcal{S}}}(y_{\mathcal{D}}|x_{\mathcal{S}})$ is calculated via

$$\begin{aligned} p_{Y_{\mathcal{D}}|X_{\mathcal{S}}}(y_{\mathcal{D}}|x_{\mathcal{S}}) = \mathcal{F}^{+} \Bigg(& \mathcal{F}^{\triangleright} \left(p_{Y_{\mathcal{SD}}|X_{\mathcal{S}}}(y_{\mathcal{SD}}|x_{\mathcal{S}}), \frac{a_{\mathcal{SD}}^2}{\sigma_N^2} \right), \\ & \mathcal{F}^{\triangleright} \left(p_{Y_{\mathcal{RD}}|X_{\mathcal{S}}}(y_{\mathcal{RD}}|x_{\mathcal{S}}), \frac{a_{\mathcal{RD}}^2}{\sigma_N^2} \right) \Bigg) \ . \end{aligned} \tag{4.19}$$

Utilizing the total probability theorem, the marginal pmf $p_{Y_{\mathcal{D}}}(y_{\mathcal{D}})$, for the combined channel output, is given by

$$p_{Y_{\mathcal{D}}}(y_{\mathcal{D}}) = \sum_{x_{\mathcal{S}} \in \mathbb{X}_{\mathcal{S}}} p_{X_{\mathcal{S}}}(x_{\mathcal{S}}) p_{Y_{\mathcal{D}}|X_{\mathcal{S}}}(y_{\mathcal{D}}|x_{\mathcal{S}}) \ , \tag{4.20}$$

summing up over all possible channel inputs $X_{\mathcal{S}}$. Please note that, due to the orthogonal access with equal time slot lengths, the rate for maximum ratio combining is restricted to half of the channel capacity,

$$R_{\text{MRC}} \leq \frac{1}{2} \max_{p_{X_{\mathcal{S}}}(x_{\mathcal{S}})} \{I(X_{\mathcal{S}}; Y_{\mathcal{D}})\} \ , \tag{4.21}$$

which conforms with Decode-and-Forward with repetition coding (3.20) and, hence, under the assumption of an error-free relay, i.e. $X_{\mathcal{R}} = X_{\mathcal{S}}$.

The *parallel information combining function* (PICF) $\mathcal{I}^{\text{par}}$ for calculating the mutual information for two parallel concatenated symmetric memoryless channels, can now be described by (4.17)

$$I(X_{\mathcal{S}}; Y_{\mathcal{D}}) = \mathcal{I}^{\text{par}}\left(p_{X_{\mathcal{S}}}(x_{\mathcal{S}}), p_{Y_{\mathcal{SD}}|X_{\mathcal{S}}}(y_{\mathcal{SD}}|x_{\mathcal{S}}), p_{Y_{\mathcal{RD}}|X_{\mathcal{S}}}(y_{\mathcal{RD}}|x_{\mathcal{S}})\right) \ , \tag{4.22}$$

where the function arguments are denoted by the conditional pmfs of the scaled[5] random variables. Again, the input parameters can change depending on the channel types and if upper or lower bounds are considered.

4.4.2 Upper and lower bound for parallel information combining

As equivalent to the serial concatenation of binary input symmetric memoryless channels there exists upper and lower bounds for parallel concatenated BISMCs. These bounds have been derived and proven in [LHHH05a]. However, a short sketch of the result is recapitulated here.

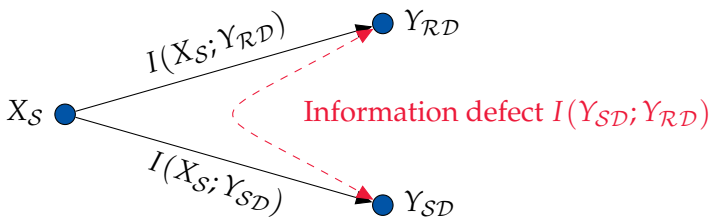


Fig. 4.6.: Representation of the information defect $I(Y_{\mathcal{SD}}; Y_{\mathcal{RD}})$

The MI for the parallel concatenation of two channels is given by

$$\begin{aligned} I(X_{\mathcal{S}}; Y_{\mathcal{D}}) &= I(X_{\mathcal{S}}; Y_{\mathcal{SD}}, Y_{\mathcal{RD}}) \\ &= I(X_{\mathcal{S}}; Y_{\mathcal{SD}}) + I(X_{\mathcal{S}}; Y_{\mathcal{RD}}) - I(Y_{\mathcal{SD}}; Y_{\mathcal{RD}}) \ , \end{aligned} \tag{4.23}$$

[5] The optimal scaling is performed with the scaling factors for maximum ratio combining.

the sum of the MI for the individual channels, i.e. the direct and the relay link, subtracting the MI of the *information defect* $I(Y_{\mathcal{SD}};Y_{\mathcal{RD}})$ [LHHH05a]. The information defect as shown in **Fig. 4.6** can be interpreted as follows: Since the direct and the relay link are independent, they form a memoryless broadcast channel, and, therefore their MI will sum up. But, the destination receives both messages and, thus, the sum in mutual information is reduced by the information defect. Here, if $Y_{\mathcal{RD}}$ is already known, than receiving $Y_{\mathcal{SD}}$ will only deliver little additional information about $X_{\mathcal{S}}$. The same holds for the opposite case, where $Y_{\mathcal{SD}}$ is already known and $Y_{\mathcal{RD}}$ is received. This concept is known as the *co-information* [Bel03] or *multi-information/interaction-information* [McG54] concept. Thus, as shown in (4.23), the sum of the MI of the individual channels has to be reduced by $I(Y_{\mathcal{SD}};Y_{\mathcal{RD}})$ in order to calculate the MI of parallel concatenated channels.

Since $Y_{\mathcal{SD}}$ and $Y_{\mathcal{RD}}$ form a Markov chain ($Y_{\mathcal{SD}} \rightarrow X_{\mathcal{S}} \rightarrow Y_{\mathcal{RD}}$, or equivalently $Y_{\mathcal{RD}} \rightarrow X_{\mathcal{S}} \rightarrow Y_{\mathcal{SD}}$, see **Fig. 4.6**) with $Y_{\mathcal{SD}} \perp Y_{\mathcal{RD}}|X_{\mathcal{S}}$, the corresponding MI is zero if conditioned on $X_{\mathcal{S}}$, i.e. $I(Y_{\mathcal{SD}};Y_{\mathcal{RD}}|X_{\mathcal{S}}) = 0$ holds. Thus, in order to calculate $I(Y_{\mathcal{SD}};Y_{\mathcal{RD}})$, serial information combining is applied and equation (4.23) can be rewritten as

$$\begin{aligned} I(X_{\mathcal{S}};Y_{\mathcal{D}}) &= I(X_{\mathcal{S}};Y_{\mathcal{SD}},Y_{\mathcal{RD}}) \\ &= I(X_{\mathcal{S}};Y_{\mathcal{SD}}) + I(X_{\mathcal{S}};Y_{\mathcal{RD}}) - \mathcal{I}^{\text{ser}}(I(X_{\mathcal{S}};Y_{\mathcal{SD}}),I(X_{\mathcal{S}};Y_{\mathcal{RD}}))\ , \end{aligned} \tag{4.24}$$

with equality [LHHH05a] for the MI of the information defect and the MI of the serial concatenation.

Given the lower (4.15) and upper (4.12) bounds for the serial information combining technique, the information defect is bounded by

$$\begin{gathered} \mathcal{I}^{\text{ser,up}}(I(X_{\mathcal{S}};Y_{\mathcal{SD}}),I(X_{\mathcal{S}};Y_{\mathcal{RD}})) \\ \geq I(Y_{\mathcal{SD}};Y_{\mathcal{RD}}) \geq \\ \mathcal{I}^{\text{ser,low}}(I(X_{\mathcal{S}};Y_{\mathcal{SD}}),I(X_{\mathcal{S}};Y_{\mathcal{RD}}))\ . \end{gathered}$$

Thus, the mutual information for the parallel concatenation of two channels must be inside

$$\begin{gathered} I(X_{\mathcal{S}};Y_{\mathcal{SD}}) + I(X_{\mathcal{S}};Y_{\mathcal{RD}}) - \mathcal{I}^{\text{ser,up}}(I(X_{\mathcal{S}};Y_{\mathcal{SD}}),I(X_{\mathcal{S}};Y_{\mathcal{RD}})) \\ \leq I(X_{\mathcal{S}};Y_{\mathcal{D}}) \leq \\ I(X_{\mathcal{S}};Y_{\mathcal{SD}}) + I(X_{\mathcal{S}};Y_{\mathcal{RD}}) - \mathcal{I}^{\text{ser,low}}(I(X_{\mathcal{S}};Y_{\mathcal{SD}}),I(X_{\mathcal{S}};Y_{\mathcal{RD}}))\ . \end{gathered} \tag{4.25}$$

With this result, upper and lower bounds for the mutual information of parallel information combining, similar as for the serial information combining, can be defined.

Upper bound for parallel information combining

As was shown in (4.25), the upper bound is given by

$$\begin{aligned} I(X_{\mathcal{S}};Y_{\mathcal{D}}) &= \mathcal{I}^{\text{par,up}}(I(X_{\mathcal{S}};Y_{\mathcal{SD}}), I(X_{\mathcal{S}};Y_{\mathcal{RD}})) \\ &\leq I(X_{\mathcal{S}};Y_{\mathcal{SD}}) + I(X_{\mathcal{S}};Y_{\mathcal{RD}}) - \mathcal{I}^{\text{ser,low}}(I(X_{\mathcal{S}};Y_{\mathcal{SD}}), I(X_{\mathcal{S}};Y_{\mathcal{RD}})) \ , \end{aligned} \tag{4.26}$$

where the information defect function is given by

$$\mathcal{I}^{\text{ser,low}}(I(X_{\mathcal{S}};Y_{\mathcal{SD}}), I(X_{\mathcal{S}};Y_{\mathcal{RD}})) = I(X_{\mathcal{S}};Y_{\mathcal{SD}}) \cdot I(X_{\mathcal{S}};Y_{\mathcal{RD}}) \ ,$$

i.e. the MI for the serial concatenation of two binary erasure channels.

Lower bound for parallel information combining

The lower bound for the mutual information of parallel concatenated channels is given by

$$\begin{aligned} I(X_{\mathcal{S}};Y_{\mathcal{D}}) &= \mathcal{I}^{\text{par,low}}(I(X_{\mathcal{S}};Y_{\mathcal{SD}}), I(X_{\mathcal{S}};Y_{\mathcal{RD}})) \\ &\geq I(X_{\mathcal{S}};Y_{\mathcal{SD}}) + I(X_{\mathcal{S}};Y_{\mathcal{RD}}) - \mathcal{I}^{\text{ser,up}}(I(X_{\mathcal{S}};Y_{\mathcal{SD}}), I(X_{\mathcal{S}};Y_{\mathcal{RD}})) \ , \end{aligned} \tag{4.27}$$

where the information defect $\mathcal{I}^{\text{ser,up}}(I(X_{\mathcal{S}};Y_{\mathcal{SD}}), I(X_{\mathcal{S}};Y_{\mathcal{RD}}))$ is depending on the different channel types. For example, in case of two binary symmetric channels

$$\mathcal{I}^{\text{ser,up}}_{\text{BSC}}(I(X_{\mathcal{S}};Y_{\mathcal{SD}}), I(X_{\mathcal{S}};Y_{\mathcal{RD}})) = 1 - H_b((1-\epsilon_{\mathcal{SD}}) \cdot \epsilon_{\mathcal{SR}} + \epsilon_{\mathcal{SD}} \cdot (1-\epsilon_{\mathcal{SR}}))$$

from equation (4.14) with crossover probabilities

$$\begin{aligned} \epsilon_{\mathcal{SD}} &= H_b^{-1}(1 - I(X_{\mathcal{S}};Y_{\mathcal{SD}})) \\ \epsilon_{\mathcal{SR}} &= H_b^{-1}(1 - I(X_{\mathcal{S}};Y_{\mathcal{SR}})) \ , \end{aligned}$$

holds.

5 Information combining for relay networks utilizing ideal and good codes

In this chapter, the mutual information of orthogonal relay channels is evaluated with the help of the *information combining technique* (ICT). This is done for ideal, i.e. capacity achieving codes, applied to some relaying protocols known from the last Chapter 4.

Here, in this thesis, the focus relies on the evaluation of relay networks, as done in [KV09b] and [KV09a]. The results of this chapter will coincide with the results from Section 3.3.6 since the ICT is only a different tool to predict the mutual information performance of relay channels. In the next Chapter 6, the models are extended towards practical relay networks, with practical codes and their corresponding *information transfer characteristic* (ITC).

5.1 Three-node orthogonal relay network

After the information combining technique has been introduced, one can apply this technique onto the relay system. In order to predict the mutual information of such a network, the different relaying protocols have to be considered during the calculation. This means that each relay protocol needs to be handled separately, i.e. utilizing serial information combining functions (SICF) and parallel information combining functions (PICF) individually.

5.1.1 Simulation environment

Unlike in the preceding and following chapters, the simulation environment and all its parameters is presented at this early stage of this chapter. This will eventually improve the readability and simplify the process of describing and comparing the simulation results. Please note that, by changing the parameters like the underlying channel model or the path-loss attenuation factor, only the shape or characteristic but not the relation between the rates (curves) is influenced. Furthermore, the descriptions of the different protocols will be general and only the simulation results depend on the specific parameters.

So, in order to show the effectiveness of the information combining technique for relay networks the simulation setup in **Fig. 5.1** is considered. The distance between

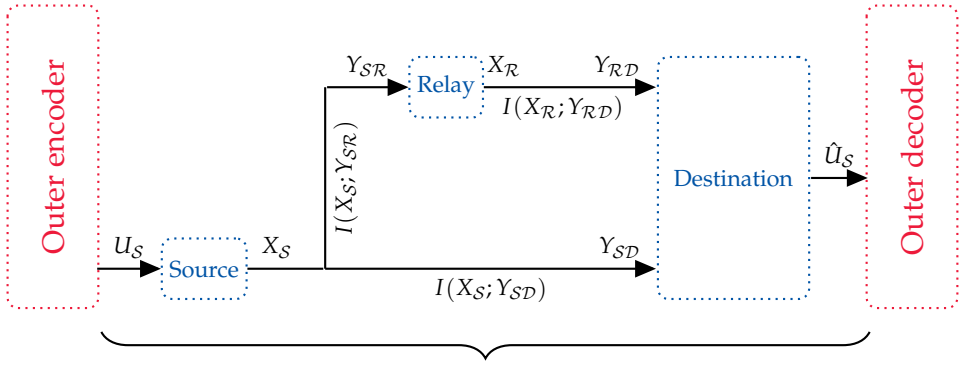


Fig. 5.1.: Simulation setup for the three-node relay network with BSC channels

source and destination is normalized, such that $d_{SD} = 1$ holds. Furthermore, the relay's position was altered on a straight line from $d_{SR} = -0.2$ to $d_{SR} = 1.2$, i.e. moving the relay from behind the source towards the destination and behind. The path-loss exponent was chosen to be $\alpha = 2$ and the receive SNR $\gamma_{SD} = -5$dB of

the direct link serves as a reference (due to the normalization of the distance and $\sigma_N^2 = 1/\gamma_{\mathcal{SD}}$).

For the other channels, the following equations hold

$$\begin{aligned} Y_{\mathcal{SR}} &= X_{\mathcal{S}} + N_{\mathcal{R}} \text{ , with } N_{\mathcal{D}} \sim \mathcal{N}\left(0, \sigma_N^2 / a_{\mathcal{SR}}^2\right) \\ Y_{\mathcal{SD}} &= X_{\mathcal{S}} + N_{\mathcal{D}}^{T_1} \text{ , with } N_{\mathcal{D}}^{T_1} \sim \mathcal{N}\left(0, \sigma_N^2 / a_{\mathcal{SD}}^2\right) \\ Y_{\mathcal{RD}} &= X_{\mathcal{R}} + N_{\mathcal{D}}^{T_2} \text{ , with } N_{\mathcal{D}}^{T_2} \sim \mathcal{N}\left(0, \sigma_N^2 / a_{\mathcal{RD}}^2\right) \text{ ,} \end{aligned}$$

with corresponding signal-to-noise ratios given by $\gamma_{\mathcal{SD}} = a_{\mathcal{SD}}^2/\sigma_N^2 = 1/\sigma_N^2$, $\gamma_{\mathcal{SR}} = a_{\mathcal{SR}}^2/\sigma_N^2$ and $\gamma_{\mathcal{RD}} = a_{\mathcal{RD}}^2/\sigma_N^2$.

The information is transmitted orthogonal during two equal time slots T_1 and T_2. In the first time slot, the source broadcasts its information towards the silent relay and the destination. Here, the source's transmit signal consists of i.i.d. BPSK modulated binary random symbols $X_{\mathcal{S}} \in \{-1, +1\}$. During the second time slot, the relay forwards the received and processed information to the destination, whereas the source keeps silent. The transmit signal of the relay node can either be $X_{\mathcal{R}} \in \{-1, +1\}$ or $X_{\mathcal{R}} \in \{\mathbb{R}\}$. This depends on the chosen relaying protocol. Finally, the destination receives two disturbed copies of the same information. Since each station transmits with unit mean power, $\mathrm{E}\{X_{\mathcal{S}}\} = \mathrm{E}\{X_{\mathcal{R}}\} = 1$ always holds.

With the source's BPSK modulated transmit signal, the channels between source and destination and source and relay are binary input additive white Gaussian noise channels. Therefore, their mutual information $I(X_{\mathcal{S}}; Y_{\mathcal{SD}})$ and $I(X_{\mathcal{S}}; Y_{\mathcal{SR}})$ can be calculated with (2.18). The same applies for the $\mathcal{R} \to \mathcal{D}$ link, as long as the channel input is BPSK modulated.

The corresponding mutual informations of the different links have the form

$$I(X_{\mathcal{S}}; Y_{\mathcal{SR}}, Y_{\mathcal{SD}}) = C_{\text{BIAWGN}}(\gamma_{\mathcal{SR}} + \gamma_{\mathcal{SD}}) \tag{5.1}$$

$$I(X_{\mathcal{S}}; Y_{\mathcal{SR}}) = C_{\text{BIAWGN}}(\gamma_{\mathcal{SR}}) \tag{5.2}$$

$$I(X_{\mathcal{S}}; Y_{\mathcal{SD}}) = C_{\text{BIAWGN}}(\gamma_{\mathcal{SD}}) \tag{5.3}$$

$$I(X_{\mathcal{R}}; Y_{\mathcal{RD}}) = C_{\text{BIAWGN}}(\gamma_{\mathcal{RD}}) \text{ ,} \tag{5.4}$$

where $C_{\text{BIAWGN}}(\ldots)$ denotes the channel capacity of a BIAWGN channel (2.18).

Furthermore, **Fig. 5.1** includes the rate $R_{\text{Protocol}}^{\text{pred,up/low}}$ (which will achieve error-free communication) and, therewith, the capacity of the complete relay channel under the constraint of a specific relaying *Protocol*. Although, the rates for the main

known protocols like Amplify-and-Forward or Decode-and-Forward can be given directly, the utilization of $R_{\text{Protocol}}^{\text{pred,up/low}}$ will become clear in the upcoming sections. In these sections, the known rates will be calculated via the information combining technique for relay networks and, in this context, $R_{\text{Protocol}}^{\text{pred,up}}$ denotes the upper bound of the rate, whereas $R_{\text{Protocol}}^{\text{pred,low}}$ holds for the lower one. To find $R_{\text{Protocol}}^{\text{pred,up}}$ and $R_{\text{Protocol}}^{\text{pred,low}}$, the code rate R_c, which for now will be the same for the source's and relay's encoder, is altered from zero to one ($0 \leq R_c \leq 1$) and the mutual information $I\left(U_{\mathcal{S}}; \hat{U}_{\mathcal{S}}\right)$ is determined for each code rate via the ICT. Afterwards, the rate $R_{\text{Protocol}}^{\text{pred,up/low}}$ will be given by

$$R_{\text{Protocol}}^{\text{pred,up/low}} = \frac{1}{2} \cdot \max_{0 \leq R_c \leq 1} \left\{ R_c \cdot I\left(U_{\mathcal{S}}; \hat{U}_{\mathcal{S}}\right) \right\} \quad . \tag{5.5}$$

Please note that, the mutual information $I\left(U_{\mathcal{S}}; \hat{U}_{\mathcal{S}}\right)$ in equation (5.5) is determined via the information combining technique and one or more information processing characteristic functions. The ICT may be delivering upper and lower bounds on the MI, which then denotes the upper and lower bound on the rate, whereas, the IPCs are functions of the code rate R_c. The factor $1/2$ in (5.5) is due to the half-duplex and orthogonal access constraint.

5.1.2 Max-flow min-cut bounds

The upper bound of achievable rates for an orthogonal three-node relay network with equal time slot lengths is calculated by the max-flow min-cut theorem (3.6) via

$$R_{\text{MFMC}}^{\text{IR}} = \frac{1}{2} \min \left\{ \begin{array}{c} I(X_{\mathcal{S}}; Y_{\mathcal{SR}}, Y_{\mathcal{SD}}), \\ I(X_{\mathcal{S}}; Y_{\mathcal{SD}}) + I(X_{\mathcal{R}}; Y_{\mathcal{RD}}) \end{array} \right\} \quad , \tag{5.6}$$

and $I(X_{\mathcal{S}}; Y_{\mathcal{SR}}, Y_{\mathcal{SD}}) = C_{\text{BIAWGN}}(\gamma_{\mathcal{SR}} + \gamma_{\mathcal{SD}})$. Hence, the maximum code rate, which includes the inner and outer coding scheme (see **Fig. 5.1**) and which achieves error-free transmission in a relay system, is determined by $R_{\text{MFMC}}^{\text{IR}}$.

Removing the ability for the source and the relay to transmit different code words, i.e. the relay utilizes the same code as the source (repetition coding), formula (5.6) further degrades to

$$R_{\text{MFMC}}^{\text{RC}} = \frac{1}{2} \min \left\{ \begin{array}{c} I(X_{\mathcal{S}}; Y_{\mathcal{SR}}, Y_{\mathcal{SD}}), \\ I\big(X_{\mathcal{S}}, X_{\mathcal{R}}; \underbrace{Y_{\mathcal{SD}}, Y_{\mathcal{RD}}}_{Y_{\mathcal{D}}}\big) \end{array} \right\} \quad , \tag{5.7}$$

with channel capacity $I(X_{\mathcal{S}}, X_{\mathcal{R}}; Y_{\mathcal{D}}) = C_{\text{BIAWGN}}(\gamma_{\mathcal{SD}} + \gamma_{\mathcal{RD}})$.

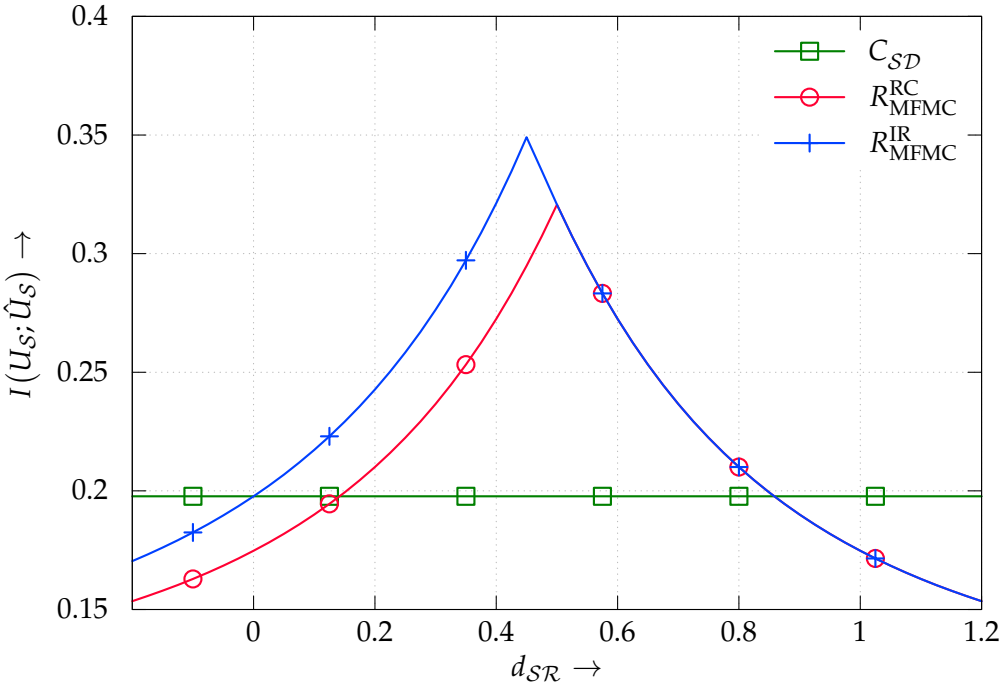


Fig. 5.2.: Calculated upper bounds for the achievable rates via the max-flow min-cut theorem

The plots for both rates, i.e. MFMC with incremental redundancy and repetition coding, are depicted in **Fig. 5.2**. As it can be seen, the limitation of repetition coding can reduce the maximum possible rate if the relay is at a position where the rate is determined by the second part of (5.6) and (5.7). Here, the rate is dominated by the $\mathcal{S} \rightarrow \mathcal{D}$ and $\mathcal{R} \rightarrow \mathcal{D}$ channels and $I(X_\mathcal{S}; Y_{\mathcal{SD}}) + I(X_\mathcal{R}; Y_{\mathcal{RD}}) \geq I(X_\mathcal{S}, X_\mathcal{R}; Y_\mathcal{D})$ holds. Furthermore, utilizing a relay over a wide range of relay positions can achieve a higher rate than for the direct transmission. This was already discussed in Chapter 3.

The depicted rates in **Fig. 5.2** will be permanently shown in the next figures to hold as a reference and for comparing the performance of the different relay protocols. In order to improve the readability of the following pictures, these curves will be slightly gray and plotted without any further description or labeling.

5.1.3 Upper and lower bound for amplify-and-forward

In case of AF the relay amplifies and forwards the received source's sequence towards the destination. Thus, the first hop is a BIAWGN channel, whereas the second hop is given by a channel with a bimodal Gaussian distributed input and additive Gaussian distributed noise.

Since both channels, i.e. direct and indirect link have binary inputs, the prediction of the rate R_{AF} for the AF protocol can be established by utilizing the parallel information combining function. To reduce the computational complexity, upper and lower bounds for the PICF, tightly surrounding the true mutual information, can be utilized. Thus, in order to calculate the upper bound, the following optimization problem is considered

$$R_{\mathrm{AF}}^{\mathrm{pred,up}} = \frac{1}{2} \cdot \max_{0 \leq R_c \leq 1} \left\{ R_c \cdot I\left(U_{\mathcal{S}}; \hat{U}_{\mathcal{S}}\right) \right\} \tag{5.8}$$

solve

$$I(X_{\mathcal{S}}; Y_{\mathcal{RD}}) = \mathcal{I}^{\mathrm{ser,up}}\left(p_{X_{\mathcal{S}}}(x_{\mathcal{S}}), p_{X_{\mathcal{R}}|X_{\mathcal{S}}}(x_{\mathcal{R}}|x_{\mathcal{S}}), p_{Y_{\mathcal{RD}}|X_{\mathcal{R}}}(y_{\mathcal{RD}}|x_{\mathcal{R}})\right) \tag{5.9}$$

$$I(X_{\mathcal{S}}; Y_{\mathcal{D}}) = \mathcal{I}^{\mathrm{par,up}}(I(X_{\mathcal{S}}; Y_{\mathcal{SD}}),\ I(X_{\mathcal{S}}; Y_{\mathcal{RD}})) \tag{5.10}$$

$$I\left(U_{\mathcal{S}}; \hat{U}_{\mathcal{S}}\right) = \mathrm{IPC}^{\mathrm{ideal}}(I(X_{\mathcal{S}}; Y_{\mathcal{D}})) \tag{5.11}$$

subject to

$$R_c \in \mathbb{R} \text{ and } 0 \leq R_c \leq 1 \ .$$

Here, in (5.9) the mutual information of the complete $\mathcal{S} \rightarrow \mathcal{R} \rightarrow \mathcal{D}$ link is calculated via the serial information combining function. Afterwards, the outcome of the SICF is parallel concatenated with the MI of the direct link via the PICF in (5.10). The resulting MI $I(X_{\mathcal{S}}; Y_{\mathcal{D}})$ is passed into the information processing characteristic of an ideal code to determine the mutual information of the information symbols after decoding. Please note that, the IPC itself is a function of the code rate R_c (see Subsection 2.3.3). Multiplying the outcome with the code rate, optimizing for the maximum and multiplying with the factor $1/2$ in (5.8), is giving the maximum rate at which an error-free communication is established. This rate includes R_c of the inner encoder/decoder pair and the rate at which the outer encoder/decoder pair must communicate. For the lower bound $R_{\mathrm{AF}}^{\mathrm{pred,low}}$, equation (5.9) and (5.10) inside the optimization problem must be replaced by $I(X_{\mathcal{S}}; Y_{\mathcal{RD}}) = \mathcal{I}^{\mathrm{ser,low}}\left(p_{X_{\mathcal{S}}}(x_{\mathcal{S}}), p_{X_{\mathcal{R}}|X_{\mathcal{S}}}(x_{\mathcal{R}}|x_{\mathcal{S}}), p_{Y_{\mathcal{RD}}|X_{\mathcal{R}}}(y_{\mathcal{RD}}|x_{\mathcal{R}})\right)$ and $\mathcal{I}^{\mathrm{par,low}}(I(X_{\mathcal{S}}; Y_{\mathcal{SD}}),\ I(X_{\mathcal{R}}; Y_{\mathcal{RD}}))$ from equation (4.15) and (4.27).

Certainly, the rate $R_{\mathrm{AF}}^{\mathrm{pred,up}} = \frac{1}{2} \cdot I(X_{\mathcal{S}}; Y_{\mathcal{D}})$ can be also predicted directly by solving only (5.9) and (5.10), and without performing the optimization problem. This is due to the ideal coding scheme and its corresponding IPC. Here, solving $R_c \cdot I\left(U_{\mathcal{S}}; \hat{U}_{\mathcal{S}}\right)$ in (5.11) with (2.29) will always yield $I(X_{\mathcal{S}}; Y_{\mathcal{D}})$, which is independent of the code rate R_c. But, to achieve a consistent description throughout this chapter, the optimization problem was included. Although, optimization is not necessary in some relaying protocol cases, like AF and DF with ideal coding schemes, it is important especially in the case for predicting relay networks with erroneous relay nodes.

In equation (5.10), the MI $I(X_{\mathcal{S}};Y_{\mathcal{SD}})$ of the direct link is given by the respective channel capacity $C_{\text{BIAWGN}}(\gamma_{\mathcal{SD}})$. Contrarily, the MI $I(X_{\mathcal{S}};Y_{\mathcal{RD}})$ in (5.9) of the $\mathcal{S}\to\mathcal{R}\to\mathcal{D}$ link is more complex to evaluate. Here, the relay scales the received signal with $\beta=\frac{1}{\sqrt{1+\sigma_N^2/a_{\mathcal{SR}}^2}}$ in order to obtain unit mean transmit power. Hence, the relay's transmit signal $X_{\mathcal{R}}$ is bimodal Gaussian distributed

$$p_{X_{\mathcal{R}}}(x_{\mathcal{R}})=\frac{1}{2}\left(\mathcal{N}\left(+\beta,\frac{\beta^2\cdot\sigma_N^2}{a_{\mathcal{SR}}^2}\right)+\mathcal{N}\left(-\beta,\frac{\beta^2\cdot\sigma_N^2}{a_{\mathcal{SR}}^2}\right)\right)\ , \tag{5.12}$$

with individual means $\pm\beta$ and common variance $\frac{\beta^2\cdot\sigma_N^2}{a_{\mathcal{SR}}^2}$. The distribution function of the channel output at the destination is given by the convolution of $p_{X_{\mathcal{R}}}(x_{\mathcal{R}})$ with $p_{N_{\mathcal{D}}^{T_2}}\left(n_{\mathcal{D}}^{T_2}\right)$, i.e. the distribution for the noise $N_{\mathcal{D}}^{T_2}$. Thus, the conditional probability density functions for the $\mathcal{S}\to\mathcal{R}\to\mathcal{D}$ link are given via

$$\begin{aligned}p_{X_{\mathcal{R}}|X_{\mathcal{S}}}(x_{\mathcal{R}}|x_{\mathcal{S}})&=\mathcal{F}^{\triangleright}\left(\mathcal{F}^{+}\left(p_{X_{\mathcal{S}}}(x_{\mathcal{S}}),p_{N_{\mathcal{R}}}(n_{\mathcal{R}})\right),\beta\right)\\ p_{Y_{\mathcal{RD}}|X_{\mathcal{R}}}(y_{\mathcal{RD}}|x_{\mathcal{R}})&=\mathcal{F}^{+}\left(p_{X_{\mathcal{R}}}(x_{\mathcal{R}}),p_{N_{\mathcal{D}}^{T_2}}\left(n_{\mathcal{D}}^{T_2}\right)\right)\ ,\end{aligned} \tag{5.13}$$

i.e. the pdf for the scaled (A.24) sum and, in the second line, the sum of two random variables (A.22). The latter one denotes the conditional pdf between the destination's input conditioned on the relay's output. Utilizing the SICF with (4.3) will deliver the MI $I(X_{\mathcal{S}};Y_{\mathcal{RD}})$ for the $\mathcal{S}\to\mathcal{R}\to\mathcal{D}$ link via

$$I(X_{\mathcal{S}};Y_{\mathcal{RD}})=\mathcal{I}^{\text{ser}}\left(p_{X_{\mathcal{S}}}(x_{\mathcal{S}}),p_{X_{\mathcal{R}}|X_{\mathcal{S}}}(x_{\mathcal{R}}|x_{\mathcal{S}}),p_{Y_{\mathcal{RD}}|X_{\mathcal{R}}}(y_{\mathcal{RD}}|x_{\mathcal{R}})\right)\ ,$$

where $X_{\mathcal{S}}\in\{+1,-1\}$ is BPSK modulated.

As depicted in **Fig. 5.3**, the true rate given in (3.17) and (3.16)

$$R_{\text{AF}}=\frac{1}{2}\cdot C\left(\frac{\gamma_{\mathcal{SR}}\cdot\gamma_{\mathcal{RD}}}{\gamma_{\mathcal{SR}}+\gamma_{\mathcal{RD}}+1}+\gamma_{\mathcal{SD}}\right) \tag{5.14}$$

is tightly predicted by

$$R_{\text{AF}}^{\text{pred,low}}\le R_{\text{AF}}\le R_{\text{AF}}^{\text{pred,up}}\ , \tag{5.15}$$

i.e. the upper and lower bound from (5.8), including the factor $1/2$ for the orthogonal access with equal time slot lengths.

Please note, that the rate R_{AF} for the AF relay protocol could also be calculated by utilizing the parallel information combining function for symmetric memoryless channels. This would result in an exact, within the range of very small numerical errors, prediction of R_{AF} given by (5.14). Of course the PICF for SMCs would involve a more complex numerical computation and, here, the focus relies in

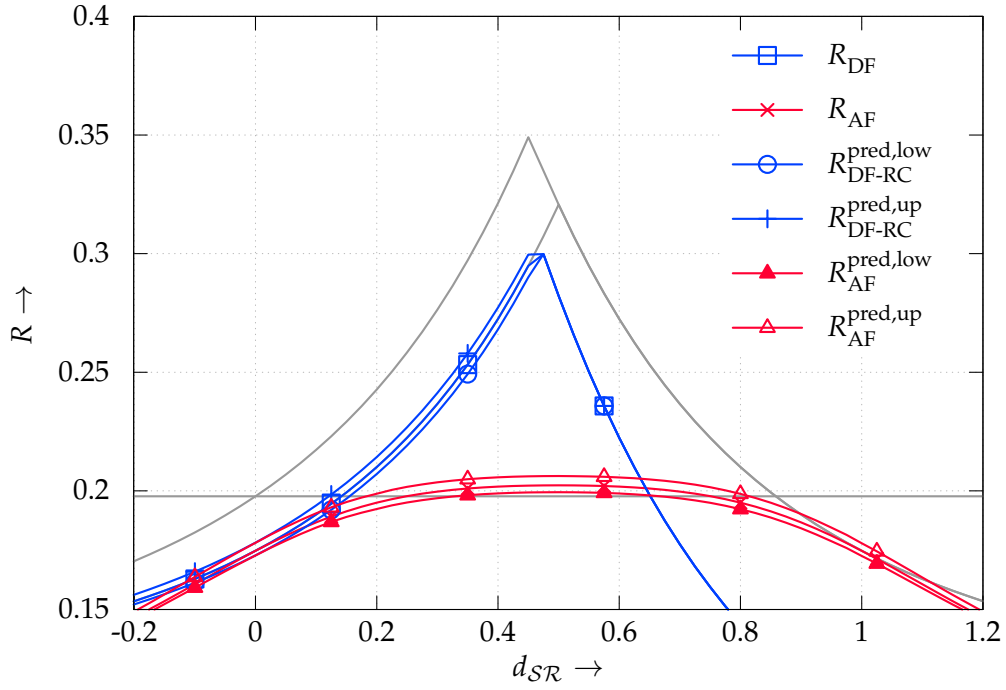


Fig. 5.3.: Predicted mutual information and calculated rates for a orthogonal three-node relay network with AF and DF

showing how tightly the upper and lower bound enfold the true rate. Thus, the outcome is not shown in the figure.

5.1.4 Upper and lower bound for decode-and-forward with repetition coding

For the Decode-and-Forward with repetition coding protocol, the relay has to perfectly decode the received sequence before forwarding a re-encoded version (same code-word as the source) towards the destination. This limits the rate of the DF-RC protocol to the capacity of the $\mathcal{S} \to \mathcal{R}$ link, once this capacity dominates the

minimum function in (3.20). Thus, to predict the upper bound on the rate for the DF-RC protocol the following optimization problem can be resolved

$$R_{\text{DF-RC}}^{\text{pred,up}} = \frac{1}{2} \cdot \max_{0 \le R_c \le 1} \left\{ R_c \cdot I\left(U_{\mathcal{S}}; \hat{U}_{\mathcal{S}}\right) \right\} \tag{5.16}$$

solve

$$I(X_{\mathcal{S}}, X_{\mathcal{R}}; Y_{\mathcal{D}}) = \mathcal{I}^{\text{par,up}}(I(X_{\mathcal{S}}; Y_{\mathcal{SD}}),\ I(X_{\mathcal{R}}; Y_{\mathcal{RD}})) \tag{5.17}$$

$$I(X_{\mathcal{S}}; Y_{\mathcal{D}}) = \min\left\{I(X_{\mathcal{S}}, X_{\mathcal{R}}; Y_{\mathcal{D}}),\ I(X_{\mathcal{S}}; Y_{\mathcal{SR}})\right\} \tag{5.18}$$

$$I\left(U_{\mathcal{S}}; \hat{U}_{\mathcal{S}}\right) = \text{IPC}^{\text{ideal}}(I(X_{\mathcal{S}}; Y_{\mathcal{D}})) \tag{5.19}$$

subject to

$$R_c \in \mathbb{R} \text{ and } 0 \le R_c \le 1$$

Furthermore, please consider **Fig. 5.4** in helping for the description.

For the optimization problem, the relay is supposed to perfectly decode the received sequence. Therefore, the channel capacities (5.3) and (5.4) of the direct and the relay link can be parallel concatenated. Afterwards, the complete information of the relay channel is calculated via the minimum of the outcome of the PICF and the capacity of the $\mathcal{S} \to \mathcal{R}$ link. At the end, the mutual information of the information symbols is given by the information processing characteristic of an ideal coding scheme. Again, the information processing characteristic function in (5.19) is a function of the code rate R_c.

The predicted upper rate for the Decode-and-Forward with repetition coding protocol, at which the combination of inner and outer capacity achieving coding schemes can perfectly communicate, is then given by (5.16). Equivalently, the lower rate is calculated by changing (5.17) to $\mathcal{I}^{\text{par,low}}(\ldots)$.

The predicted upper and lower rates

$$R_{\text{DF-RC}}^{\text{pred,low}} \le R_{\text{DF-RC}} \le R_{\text{DF-RC}}^{\text{pred,up}} \tag{5.20}$$

tightly enfold the rate

$$R_{\text{DF-RC}} = \frac{1}{2} \cdot \min\left\{ I(X_{\mathcal{S}}, X_{\mathcal{R}}; Y_{\mathcal{D}}), I(X_{\mathcal{S}}, X_{\mathcal{R}}) \right\} \tag{5.21}$$

for the DF-RC given by equation (3.20). This is shown in **Fig. 5.3**.

Comparing the Amplify-and-Forward and Decode-and-Forward with repetition coding protocols, it can be seen that the latter one achieves a higher rate as long as the relay is close to the source. In this region, and due to the restriction of perfect decoding of the relay the minimum in (5.21) is dictated by $I(X_{\mathcal{S}}, X_{\mathcal{R}}; Y_{\mathcal{D}})$. Further-

more, the rate $R_{\text{DF-RC}}$ corresponds to the rate of $R_{\text{MFMC}}^{\text{RC}}$. As soon as $I(X_{\mathcal{S}}, X_{\mathcal{R}}; Y_{\mathcal{D}}) \geq I(X_{\mathcal{S}}, X_{\mathcal{R}})$, the rate drops to the MI of the $\mathcal{S} \rightarrow \mathcal{R}$ link.

For the given simulation setup and at a distance of approximately $d_{\mathcal{SR}} \approx 0.65$, the AF protocol outperforms the DF-RC protocol. This proposition is also valid comparing with the capacity of the direct link for $d_{\mathcal{SR}} \geq 0.75$. Hence, a selection between the Amplify-and-Forward and Decode-and-Forward with repetition coding protocol seems desirable and feasible.

5.1.5 Selection between AF and DF

For the selection between AF and DF-RC, the relay either perfectly decodes the source sequence (as for DF) and then transmits a re-encoded version (with the same source code) towards the destination. Or, in case of erroneous decoding, it forwards (3.13) its normalized received symbols $X_{\mathcal{R}} = \beta \cdot Y_{\mathcal{SR}}$, with $\beta = 1/\sqrt{1 + (1/\gamma_{\mathcal{SR}})}$. This mechanism is depicted in **Fig. 5.4**, where the relay switches depending on the decoding result.

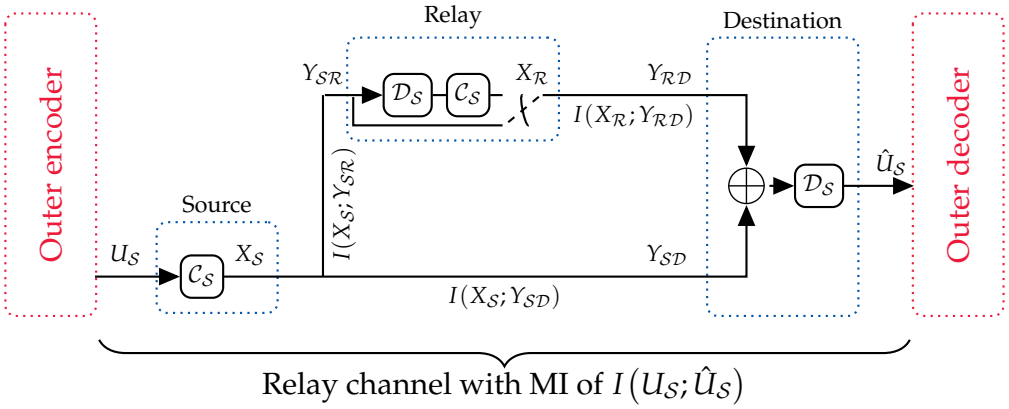


Fig. 5.4.: Two hop, orthogonal relaying scheme, with maximum ratio combining of code symbols before the destinations decoder input.

Therefore, the maximum rate is given by

$$R_{\text{AF\&DF-RC}} = \max\left\{R_{\text{AF}},\ R_{\text{DF-RC}}\right\}\ , \tag{5.22}$$

i.e. the maximum of the AF and DF-RC rate. The same holds for the upper and lower bounds on the predicted rates given by (5.8) and (5.16).

The result is depicted in **Fig. 5.6**. Here, the selection between AF and DF-RC always achieves the best performance throughout all relay positions as compared to their individual protocols.

5.1.6 Optimal combining at the destination with decoding errors at the relay

Until now, it was silently assumed that the source's and the relay's channel output symbols $Y_{\mathcal{RD}}$ and $Y_{\mathcal{SD}}$ are maximum ratio combined before the destination's decoder input (see **Fig. 5.4**). This is possible, since the transmission is orthogonal and, hence, the destination receives two separate sequences of the same information. But the combination can also be performed after the destination's decoder outputs $\hat{U}_{\mathcal{SRD}}$ and $\hat{U}_{\mathcal{SD}}$ for the source and relay link. This is depicted in **Fig. 5.5**. Here, the

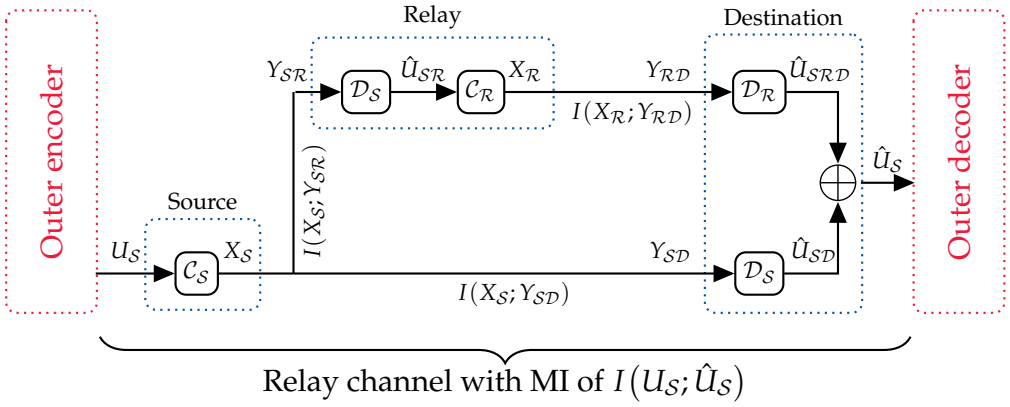


Fig. 5.5.: Two hop, orthogonal relaying scheme, with maximum ratio combining of information symbols after the destinations decoders.

combination of information after decoding can have some advantages as compared to the combination of channel outputs. For example, it is now advantageous to re-encode and re-transmit an erroneous information sequence[1] from the relay node. As this sequence is not maximum ratio combined at the destinations decoder input, it might not further distort the source's transmit sequence. This is an important issue since one error in the information sequence at the relay node can lead to a complete different code word $X_{\mathcal{R}}$, e.g. recursive convolutional codes in practical coding schemes. Hence, the combination with $X_{\mathcal{S}}$, respectively the maximum ratio combining of $Y_{\mathcal{RD}}$ with $Y_{\mathcal{SD}}$, results in descent performance. To overcome this issue, the relay should apply a forward error correction for the second hop and the destination maybe capable to perfectly recover the relay's information sequence. This sequence might have the same errors, which where present at the relay node. But combining this sequence with the decoded source's sequence can improve the overall performance and lead to an increased rate of the complete relay channel. On a side note, passing $\hat{U}_{\mathcal{SRD}}$ as a-priori information for $\mathcal{D}_{\mathcal{S}}$ or vice versa ($\hat{U}_{\mathcal{SD}}$

[1] By erroneous information sequence, it is meant that the recovered information at the relay node after $\mathcal{D}_{\mathcal{S}}$ might contain errors. Still, parts of $\hat{U}_{\mathcal{SR}}$ are useful for further processing at the destination in combination with the information from the direct link. Thus, the idea is to re-encode and re-transmit the relay's information.

into $\mathcal{D}_{\mathcal{R}}$), as for example in a turbo decoder, does not improve the decoding result since ideal coding schemes are utilized in this chapter.

Now, in contrast to DF, the source is allowed to apply a code with a code rate larger than $I(X_{\mathcal{S}};Y_{\mathcal{SR}})$ and, inherently, larger than $I(X_{\mathcal{S}};Y_{\mathcal{SD}})$. Hence, error-free decoding is neither achieved at the relay, nor at the destination node for the direct link. But, since the chosen code rate maybe smaller than $I(X_{\mathcal{R}};Y_{\mathcal{RD}})$, the relay's information $\hat{U}_{\mathcal{SR}}$ could be perfectly recovered ($\hat{U}_{\mathcal{SRD}} = \hat{U}_{\mathcal{SR}}$) at the destination via $\mathcal{D}_{\mathcal{R}}$. Even if errors occurred at the relay node, the combination of $\hat{U}_{\mathcal{SRD}}$ with $\hat{U}_{\mathcal{SD}}$ can lead to a larger mutual information of $I(U_{\mathcal{S}};\hat{U}_{\mathcal{S}})$ as compared to the DF case. The MI $I(U_{\mathcal{S}};\hat{U}_{\mathcal{S}})$ determines the code rate of the outer coding scheme and, therewith, of the relay channel by multiplying with R_c of the inner coding scheme.

The following optimization problem

$$R_{\text{after}}^{\text{pred,up}} = \frac{1}{2} \cdot \max_{0 \le R_c \le 1} \left\{ R_c \cdot I(U_{\mathcal{S}};\hat{U}_{\mathcal{S}}) \right\} \tag{5.23}$$

solve

$$I(U_{\mathcal{S}};\hat{U}_{\mathcal{SR}}) = \text{IPC}^{\text{ideal}}(I(X_{\mathcal{S}};Y_{\mathcal{SR}})) \tag{5.24}$$

$$I(U_{\mathcal{S}};\hat{U}_{\mathcal{SD}}) = \text{IPC}^{\text{ideal}}(I(X_{\mathcal{S}};Y_{\mathcal{SD}})) \tag{5.25}$$

$$I(\hat{U}_{\mathcal{SR}};\hat{U}_{\mathcal{SRD}}) = \text{IPC}^{\text{ideal}}(I(X_{\mathcal{R}};Y_{\mathcal{RD}})) \tag{5.26}$$

$$I(U_{\mathcal{S}};\hat{U}_{\mathcal{SRD}}) = \mathcal{I}^{\text{ser,up}}\left(I(U_{\mathcal{S}};\hat{U}_{\mathcal{SR}}),\ I(\hat{U}_{\mathcal{SR}};\hat{U}_{\mathcal{SRD}})\right) \tag{5.27}$$

$$I(U_{\mathcal{S}};\hat{U}_{\mathcal{S}}) = \mathcal{I}^{\text{par,up}}\left(I(U_{\mathcal{S}};\hat{U}_{\mathcal{SD}}),\ I(U_{\mathcal{S}};\hat{U}_{\mathcal{SRD}})\right) \tag{5.28}$$

subject to

$$R_c \in \mathbb{R} \text{ and } 0 \le R_c \le 1$$

summarizes the mentioned steps of this subsection.

To determine the rate $R_{\text{after}}^{\text{pred,up}}$ at which the relay system can communicate without error, the mutual information after decoding at the different nodes must be evaluated. This is done in (5.24), (5.25) and (5.26) for the $\mathcal{S} \to \mathcal{R}$, $\mathcal{S} \to \mathcal{D}$ and $\mathcal{R} \to \mathcal{D}$ links with the help of information processing characteristics for ideal coding schemes. Please keep in mind that the IPCs are functions of the channel capacities from (5.2), (5.3) and (5.4), and the code rate R_c.

The concatenation of $I(U_{\mathcal{S}};\hat{U}_{\mathcal{SR}})$ and $I(\hat{U}_{\mathcal{SR}};\hat{U}_{\mathcal{SRD}})$ via the serial information combining function in (5.27) and the combination of the outcome with the direct link in (5.28) via the parallel information combining function leads to the upper bound in mutual information $I(U_{\mathcal{S}};\hat{U}_{\mathcal{S}})$ of the relay network in **Fig. 5.5**. Multiplying $I(U_{\mathcal{S}};\hat{U}_{\mathcal{S}})$ with R_c and solving the optimization problem by finding the maximum

in (5.23) will eventually deliver the upper and lower bound on the maximum rate R_{after} denoted by

$$R_{\text{after}}^{\text{pred,low}} \leq R_{\text{after}} \leq R_{\text{after}}^{\text{pred,up}} \quad . \tag{5.29}$$

Again, the lower bound is given by replacing the SICF and PICF in (5.27) and (5.28) with its lower counterpart.

The result is shown in **Fig. 5.6**. As can be seen for cases where the relay is close to the

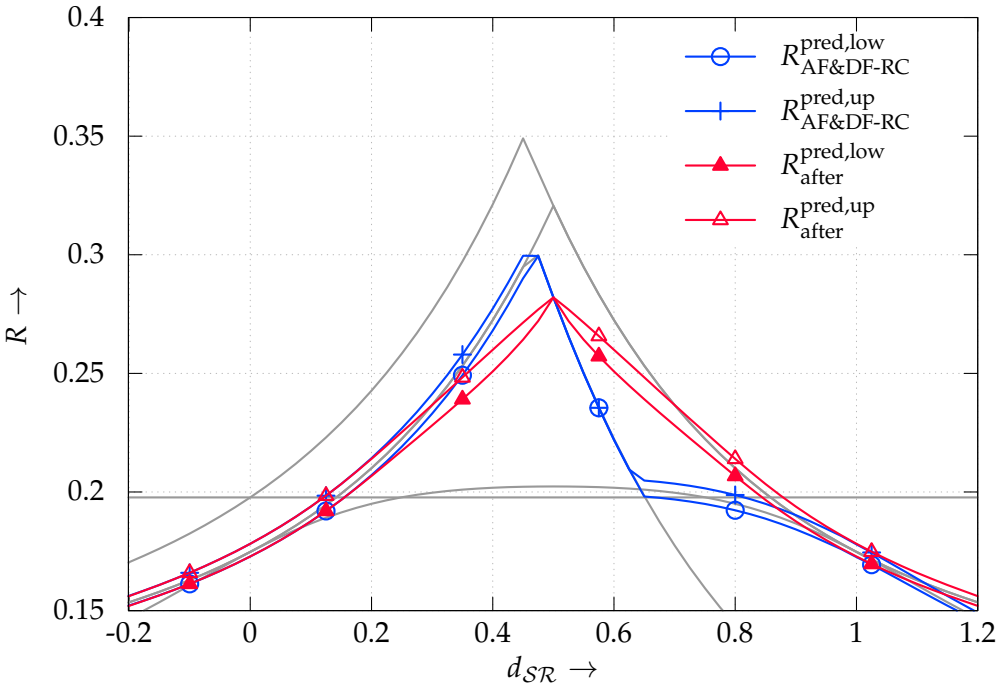


Fig. 5.6.: Predicted mutual information for a three-node relay network with the selection between the AF and DF-RC protocol before and after the destination's decoder(s)

source ($d_{\mathcal{SR}} \leq 0.5$), maximum ratio combining the decoder outputs (information symbols) at the destination node leads to a smaller rate as compared to MRC the channel outputs (code symbols), i.e. $R_{\text{after}}^{\text{pred}} < R_{\text{AF\&DF-RC}}$ holds. Please recapitulate **Fig. 5.5** and **Fig. 5.4** for the difference of the two methods.

Although, in this situation the relay can perfectly decode in both schemes, the mutual information for the indirect link of the scheme from **Fig. 5.5** is limited to $\min\left\{I\left(U_{\mathcal{S}};\hat{U}_{\mathcal{SR}}\right),\ I\left(\hat{U}_{\mathcal{SR}};\hat{U}_{\mathcal{SRD}}\right)\right\}$. Parallel combining this result with the MI $I\left(U_{\mathcal{S}};\hat{U}_{\mathcal{SD}}\right)$ of the direct link will only slightly improve the performance.

This is different for the scheme shown in **Fig. 5.4**. There, combining the channel outputs leads to a small increase of the mutual information of code symbols. But

since the slope of the ideal code is above that of the direct transmission, a small increase of MI at the decoder input produces a larger increase of MI at the decoder output. Thus, the mutual information performance is equivalent to the Decode-and-Forward with repetition coding scheme and above that of combining the information at the destination.

At the position of $d_{\mathcal{SR}} \geq 0.5$, the relay's information sequence can be perfectly recovered at the destination and, therefore, the first hop is the limiting factor in $\min\left\{I\left(U_{\mathcal{S}};\hat{U}_{\mathcal{SR}}\right),\ I\left(\hat{U}_{\mathcal{SR}};\hat{U}_{\mathcal{SRD}}\right)\right\}$. But, since one can re-encode the relay's information sequence, the recovered but erroneous (compared to the original source's information) sequence can improve the overall performance. This is not possible for maximum ratio combining the channel output and, hence, the performance drops to the equivalent of the Amplify-and-Forward relaying scheme.

Again, please note that it is essential to combine the decoder and not the channel outputs, if the relay is allowed to re-encode erroneous information. In the case of errors, the resulting code word and, hence, transmit sequence $X_{\mathcal{R}}$ can be completely different from $X_{\mathcal{S}}$. Thus, maximum ratio combining of $X_{\mathcal{R}}$ and $X_{\mathcal{S}}$ can lead to a severe performance degradation. For practical codes, a method is introduced in Subsection 6.2 which utilizes the decoding performance of the relay node in order to lower the performance degradation. This is possible, since the constraint lengths of many practical codes are limited and, thus, errors in the code sequence, caused by errors in the information sequence, are limited in their spreading.

Comparison of MRC before and after the destination's decoder and ideal coding schemes

In this section, the benefits and drawbacks of maximum ratio combining of the channel outputs versus combining the destination's decoder outputs will be discussed. For this, let's recapitulate **Fig. 5.4** and **Fig. 5.5**.

The first figure shows an orthogonal two-hop three-node relay network with MRC of channel (code) symbols ($Y_{\mathcal{SD}}$ with $Y_{\mathcal{RD}}$) before decoding at the destination. This is possible since, in the case of error-free decoding, the relay node applies the same encoder as the source in order to (re)produce the same code word. Or, in case of errors, the relay forwards the received source message like in Amplify-and-Forward.

For the second figure, the setup is similar but with maximum ratio combining of information symbols at the destination's decoder outputs. Surely, this setup has an

increased complexity due to the additional decoder at the destination. But, the relay can now apply a different encoder as compared to the source and, furthermore, the relay is not restricted by achieving error-free decoding. Thus, the relay is allowed to re-encode any received sequence, regardless of decoding errors.

Since maximum ratio combining of channel symbols requires the employment of two equal encoding schemes at the source and the relay node and to perform a fair comparison, the relay's encoder $\mathcal{C}_{\mathcal{R}}$ will be equivalent to the source's encoder $\mathcal{C}_{\mathcal{S}}$ in both setups (**Fig. 5.4** and **Fig. 5.5**). Even though, applying a different code with different code rate for the $\mathcal{R} \rightarrow \mathcal{D}$ link is the outstanding benefit of the second setup.

To perform the comparison of both schemes, the mutual informations $I(X_{\mathcal{S}}; Y_{\mathcal{SD}})$, $I(X_{\mathcal{S}}; Y_{\mathcal{SR}})$ and $I(X_{\mathcal{R}}; Y_{\mathcal{RD}})$ of the three different communication channels have been altered via altering their corresponding channel SNRs or, to be precisely, via their corresponding crossover probabilities. This also implies a change of the MI $I(U_{\mathcal{S}}; \hat{U}_{\mathcal{SD}})$, $I(U_{\mathcal{S}}; \hat{U}_{\mathcal{SR}})$ and $I(\hat{U}_{\mathcal{SR}}; \hat{U}_{\mathcal{SRD}})$ for the different end-to-end channels as stated in (5.24), (5.25) and (5.26) for the second setup. Furthermore, for simplicity, the communication links are binary symmetric channels. This has the benefit that in order to calculate the MI of the $\mathcal{S} \rightarrow \mathcal{R} \rightarrow \mathcal{D}$ link, simple information combining of BSCs (4.14) on page 65 can be applied without changing the quantitative conclusions.

In the case of serial concatenation of a BISMC with a subsequent SMC, the calculations will become cumbersome as shown in Section 4.3.1 on page 59. The problem resides in the fact, that the capacity of the symmetric memoryless channel is not only depending on the additive noise, but also on the distribution of the channel input. Thus, with changing capacity of the first hop, a change in the mutual information of the second hop is inevitably. Furthermore, the required information processing characteristics of ideal or good codes are based on channels with uniform binary input distributions and, until now, no theory exists for channels with other types of inputs (non-binary or non-uniform input distributions). Thus, the analysis is based on binary symmetric channels. On a side note, please note that in Subsection 6.1.4 an extended IPC for practical soft-input soft-output codes is introduced, which inherently deals with SMCs.

The goal, as already stated, is to predict the maximum achievable rate of the network such that the communication between source and destination will become error-free.

Before to start, please note that in a scenario where the first hop of the indirect link is perfect ($I(X_{\mathcal{S}}; Y_{\mathcal{SR}}) = 1$), i.e. in a scenario where the source communicates with

the destination over two independent (parallel) channels without any further signal processing on the channels, combining the channel (MRC before) or the decoder (MRC after) outputs at the destination will not make a difference regarding the mutual information $I(U_S; \hat{U}_S)$. Here, both setups will deliver the same result.

In the case of a first hop with $I(X_S; Y_{SR}) < 1$ and in case of decoding errors either on the first[2] or second hop of the $S \to R \to D$ link, the combination will cause different results on the MI for both maximum ratio combining methods. This is depicted in **Fig. 5.7**. Here, two different scenarios have been simulated.

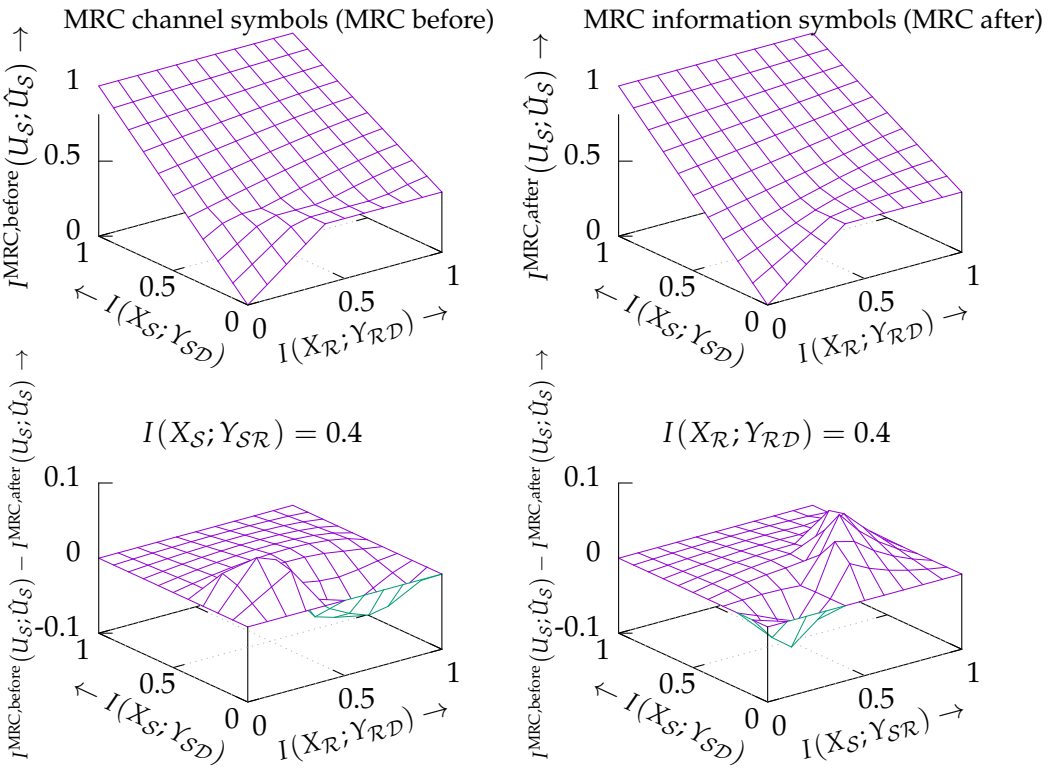


Fig. 5.7.: Mutual informations and their differences in case of maximum ratio combining of channel or information symbols at the destination. Either the MI of the first or second hop is limited ($I(X_S; Y_{SR}) = 0.4$ or $I(X_R; Y_{RD}) = 0.4$).

In the first scenario, the mutual information $I(X_S; Y_{SR}) = 0.4$ of the communication channel has been fixed. Thus, the relay node is not capable to perfectly decode the received sequence until the code rate corresponds to that MI and, hence, link errors occur on the first hop at the relay node. This is depicted in the two upper plots in **Fig. 5.7**. Here, the mutual informations $I(X_S; Y_{SD})$ and $I(X_R; Y_{RD})$ of the two remaining channels have been altered. Furthermore, in case of $I^{\text{MRC,before}}(U_S; \hat{U}_S)$

[2]Please note, that in the case of Decode-and-Forward with repetition coding, the relay is restricted by the perfect decoding constraint. Thus, the relay is not allowed to make errors, which automatically influences the maximum code rate the source is allowed to apply.

(MRC before) the optimization problems (5.8) and (5.16) for calculating (5.22), i.e. switching between AF and DF-RC have been performed. For $I^{\text{MRC,after}}(U_{\mathcal{S}};\hat{U}_{\mathcal{S}})$ (MRC after), the optimization problem (5.23) was utilized. Thus, the mutual information $I(X_{\mathcal{S}};Y_{\mathcal{SR}}) = 0.4$ cannot be regarded as the maximum possible code rate for these two setups.

In the simulation results, it can be seen that there exists a slight difference between the corresponding mutual informations $I^{\text{MRC,before}}(U_{\mathcal{S}};\hat{U}_{\mathcal{S}})$ and $I^{\text{MRC,after}}(U_{\mathcal{S}};\hat{U}_{\mathcal{S}})$ for the end-to-end channels. This difference is depicted in the lower left plot. Here, if the direct link and the $\mathcal{R} \rightarrow \mathcal{D}$ link have poor link qualities ($I(X_{\mathcal{S}};Y_{\mathcal{SD}}) < 0.4$ and $I(X_{\mathcal{R}};Y_{\mathcal{RD}}) < 0.4$), then MRC before the destination's decoder is superior. Although, combining the channel outputs at the destination only leads to a small increase of the mutual information for the channel symbols, the MI after the decoder is increased due to the slope inside the information processing characteristic of the ideal code. In comparison, combining the decoder outputs, each individual decoder already recovered as much as possible information which is less or equal to what can be achieved by combining the channel outputs.

As soon as the $\mathcal{R} \rightarrow \mathcal{D}$ link quality increases ($I(X_{\mathcal{R}};Y_{\mathcal{RD}}) > 0.4$), MRC of information symbols after decoding at the destination will become superior. In this case, the relay node is allowed to re-encode even under the case of decoding errors, since the restriction to equal transmit code words is abrogated. Then, the coding scheme for the second hop maybe capable to recover the full information of the relay node (including the errors). This is definitely better than forwarding the relay's sequence as in AF and, therefore, dealing with additional distortion introduced by the second hop. Hence, in this region, maximum ratio combining of the information symbols is the better option.

Please note that in case of a non-existing direct link, i.e. $I(X_{\mathcal{S}};Y_{\mathcal{SD}}) = 0$, the overall mutual information is limited to the minimum of $I(X_{\mathcal{R}};Y_{\mathcal{RD}})$ and $I(X_{\mathcal{S}};Y_{\mathcal{SR}}) = 0.4$. This is also the sharp decision line for deciding which combining scheme to use, as long as the direct link has no further influence in the overall mutual information.

For the second scenario with $I(X_{\mathcal{R}};Y_{\mathcal{RD}}) = 0.4$ of the second hop, and the simulation result depicted in the lower right part of **Fig. 5.7**, the situation from above changes. Here, as long as $I(X_{\mathcal{S}};Y_{\mathcal{SR}}) \geq 0.4$ and $I(X_{\mathcal{S}};Y_{\mathcal{SD}}) \geq 0.4$, the maximum possible code rate will be below $I(X_{\mathcal{R}};Y_{\mathcal{RD}}) \leq 0.4$ and, thus, the relay has perfect knowledge of the source's information sequence and DF-RC will be applied. Furthermore, in this situation, the best option for the relaying scheme is to combine the channel symbols.

Once, the relay node cannot perfectly recover the source's information, one should perform maximum ratio combining of the decoder output symbols. The corresponding result for the difference in mutual information $I^{\text{MRC,before}}\left(U_{\mathcal{S}};\hat{U}_{\mathcal{S}}\right) - I^{\text{MRC,after}}\left(U_{\mathcal{S}};\hat{U}_{\mathcal{S}}\right)$ is plotted in the lower right part of **Fig. 5.7**.

Closely looking at the simulation result, one could conclude that a sharp decision line for choosing MRC before or after decoding exists. This maybe true as long as only the combination of DF-RC and AF with MRC of channel symbols is compared with MRC of information symbols. The problem resides in finding of this decision line, which represents the intersection of MI for both setups. For example, looking at **Fig. 5.6**, there is a small area (relay position at $0.45 < d_{\mathcal{SR}} < 0.5$) in which DF-RC (constraint on $R_c = I(X_{\mathcal{S}};Y_{\mathcal{SR}})$) is still better than MRC after.

6 Information combining for relay networks utilizing practical codes

Until now, only relay networks with capacity achieving codes have been considered. Their performance was evaluated for the different relaying protocols with the max-flow min-cut theorem and the information combining technique.

In this chapter, the focus relies on practical relay networks, i.e. relay networks with practical codes, and orthogonal access. Their performance cannot be evaluated by the MFMC theorem since it only delivers an upper bound on the information rate. Thus, the beforehand introduced information combining technique for relay networks is now utilized in order to semi-analytically predict the mutual information of the practical relay network.

After showing the effectiveness of the information combining technique for small relay schemes with only one relay between source and destination, the system model will be extended towards a network with two relays. Here, the semi-analytically calculated mutual information is always compared with the mutual information obtained from the Monte Carlo simulation. The extension towards more relays and hops is straight forward.

Additionally, in this chapter a method is introduced which calculates the channel LLRs at the destination's input under the assumption of a soft-input soft-output re-encoder at the relay node. Then the performance of hard-input hard-output versus soft-input soft-output re-encoding of the complete $\mathcal{S} \rightarrow \mathcal{R} \rightarrow \mathcal{D}$ link is investigated. This will eventually give some more insights over the performance of the complete relay channel.

6.1 Three-node relay network

The relay protocols in Section 3.3 have been reviewed under the constraint of capacity achieving codes. Thus, the results from the preceding chapters only deliver upper bounds on the achievable rates. Furthermore, these bounds where tightly predicted by the information combining technique for relay networks. The problem is, that the results do not give much insights on how to cope with practical codes for the communication. Unfortunately, this problem cannot be solved via the max-flow min-cut theorem.

To evaluate a communication system, either in terms of bit error rate (BER) or mutual information (MI) performance, exhaustive Monte Carlo simulations are usually performed. This method of repeating simulations in order to obtain a numerical result is often very time intensive and computationally complex. Therefore, to overcome this problem and to predict the mutual information of a relay network directly, the information transfer characteristic (see Section 2.3.3) is included in the information combining technique. Thus, for example, one can semi-analytically calculate the MI of each link and afterwards combine the outcome to predict the MI of a complete relay network.

In this section, the introduced relaying protocols are now combined with practical codes, which will be referred to as *practical relaying*. Among others, these codes can be convolutional codes, with either maximum ratio combining or iterative turbo decoding at the destination. Also, unlike relaying protocols with capacity achieving codes, the upcoming schemes cannot guarantee error-free decoding neither at the relay nor at the destination side.

The system model for an orthogonal practical two-hop relaying network with a single relay is depicted in **Fig. 6.1**. Due to the orthogonality, the communication between source and destination is departed into two not necessarily equal time slots T_1 and T_2. During the first time slot T_1, the source encodes the binary information sequence $\mathbf{u}_\mathcal{S} = \left[u_{\mathcal{S},j}\right]_{J\times 1}$, $u_{\mathcal{S},j} \in \{0,1\}$ with the code[1] $\mathcal{C}_\mathcal{S}$, generating the encoded and BPSK modulated (BPSK, *binary phase shift keying*) code sequence $\mathbf{x}_\mathcal{S} = [x_{\mathcal{S},k}]_{K\times 1}$, $x_{\mathcal{S},k} \in \{+1,-1\}$, before transmitting $\mathbf{x}_\mathcal{S}$ towards the relay and the destination.

The relay receives $\mathbf{y}_{\mathcal{SR}} = [y_{\mathcal{SR},k}]_{K\times 1}$, $y_{\mathcal{SR},k} \in \{\mathbb{R}\}$, i.e.

$$\mathbf{y}_{\mathcal{SR}} = \mathbf{x}_\mathcal{S} + \mathbf{n}_\mathcal{R} \tag{6.1}$$

[1] In this thesis, exclusively convolutional codes will be regarded.

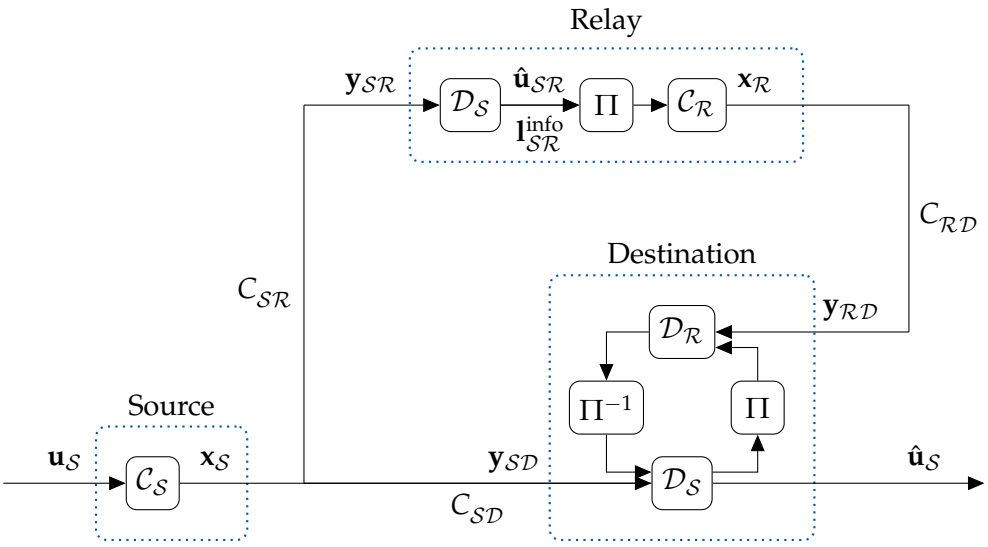


Fig. 6.1.: Orthogonal relaying scheme, with two hops, a single relay and a turbo decoder at the destination.

which forms a BIAWGN channel with a channel capacity of $C_{\text{BIAWGN}}(\gamma_{\mathcal{SR}})$ with $N_{\mathcal{R}} \sim \mathcal{N}(0, \sigma_N^2/a_{\mathcal{SR}}^2)$. After decoding, the relay either obtains log-likelihood ratios $\mathbf{l}^{\text{info}}_{\mathcal{SR}} = \left[l^{\text{info}}_{\mathcal{SR},j}\right]_{J\times 1}$, $l^{\text{info}}_{\mathcal{SR},j} \in \{\mathbb{R}\}$ or hard decisions $\hat{\mathbf{u}}_{\mathcal{SR}} = \left[\hat{u}_{\mathcal{SR},j}\right]_{J\times 1}$, $\hat{u}_{\mathcal{SR},j} \in \{0,1\}$ of the source's information bits. These estimates are interleaved and re-encoded with the code $\mathcal{C}_{\mathcal{R}}$, before forwarding the resulting code sequence $\mathbf{x}_{\mathcal{R}} = [x_{\mathcal{R},m}]_{M\times 1}$, $x_{\mathcal{R},m} \in \{\mathbb{R}\}$ towards the destination in the second time slot T_2.

The relay's encoder $\mathcal{C}_{\mathcal{R}}$ is either a *soft-input soft-output* (SISO) or *hard-input hard-output* (HIHO) encoder, depending on the type of its input sequence. This, in case of non-binary inputs, requires some further processing (scaling) after the relay's encoder output, since the output $\mathbf{x}_{\mathcal{R}}$ is supposed to have unit variance (unit mean power).

Furthermore, the relay is allowed to apply a different code to the estimated source sequence. Thus, in case of a different code rate but unchanged modulation, the code word lengths $K \neq M$ are not equal and, therefore, implies different time slot lengths, i.e. $t_1 \neq t_2$. Here, the source's code rate $R^{\text{CC}}_{\mathcal{C}_{\mathcal{S}}}$ and the relay's code rate $R^{\text{CC}}_{\mathcal{C}_{\mathcal{R}}}$ are given by

$$R^{\text{CC}}_{\mathcal{C}_{\mathcal{S}}} = J/K$$

and

$$R^{\text{CC}}_{\mathcal{C}_{\mathcal{R}}} = J/M$$

respectively.

The destination receives the source's message during the first time slot T_1 via the direct link and during the second time slot T_2 via the relay link, i.e.

$$\mathbf{y}_{\mathcal{SD}} = \mathbf{x}_{\mathcal{S}} + \mathbf{n}_{\mathcal{D}}^{T_1} \ , \tag{6.2}$$

with $\mathbf{y}_{\mathcal{SD}} = [y_{\mathcal{SD},k}]_{K\times 1}$, $y_{\mathcal{SD},k} \in \{\mathbb{R}\}$ and

$$\mathbf{y}_{\mathcal{RD}} = \mathbf{x}_{\mathcal{R}} + \mathbf{n}_{\mathcal{D}}^{T_2} \ , \tag{6.3}$$

with $\mathbf{y}_{\mathcal{RD}} = [y_{\mathcal{RD},m}]_{M\times 1}$, $y_{\mathcal{RD},m} \in \{\mathbb{R}\}$. The noise sequences are Gaussian distributed with $N_{\mathcal{D}}^{T_1} \sim \mathcal{N}\left(0, \sigma_N^2/a_{\mathcal{SD}}^2\right)$ and $N_{\mathcal{D}}^{T_2} \sim \mathcal{N}\left(0, \sigma_N^2/a_{\mathcal{RD}}^2\right)$.

It is obvious that the source's message via the relay link has been additionally processed by the relay and, thus, extra treatment of this received sequence has to be taken before the destination can perform any kind of combining or processing. These steps will be precisely discussed in the following subsections.

After the reception, the destination is capable of performing maximum ratio combining (MRC) before the decoder input or after the decoders outputs. The decision, where the combination takes place is depending on the way, the relay processes the received source's information. Here, MRC before the decoder input can only be performed, iff. the relay did not interleave the decoded source's sequence and iff. the utilized code does not differ from the source's code. If one of these conditions is not fulfilled, then the decoder outputs at the destination are maximum ratio combined. Please keep in mind that in the second case two decoders, i.e. one for the direct link and one for the relay link, have to be implemented. This is not necessary in the first case, where only one decoder is utilized (see **Fig. 5.4** and **Fig. 5.5**).

A third option to retrieve the source's message at the destination node is to perform iterative turbo decoding (see **Fig. 6.1**). Here, since source and relay links are independently carrying the same information a parallel turbo code is established and, thus, an iterative decoding process at the destination is possible (see for example [Tho08] and [SLV08]).

All three combining and decoding options will be discussed in the following subsections. Furthermore, a comparison of these techniques is given, regarding the mutual information and bit error rate performance. This is needed, as all three techniques perform differently, depending on the link qualities of the channels. Given some information about the channel qualities, like the signal-to-noise ratios, an adaptive selection of the best combining and decoding technique can be made beforehand. To reduce the computational complexity and, thus, to perform this selection "on time", the information combining technique is utilized. This allows the selection of

the different combining and coding schemes in advance and differently for each transmitted frame.

6.1.1 Source destination link and source relay link

As already mentioned, the destination receives the source's message during T_1 through the direct link. Since all decoders in the relay network are *soft-input soft-output* decoders, they require *log-likelihood ratio*s (LLRs) for an optimal decoding performance (see Section 2.3). Assuming a binary input additive white Gaussian noise channel (BIAWGN), the symbolwise log-likelihood ratios (A.38) for (6.2) are calculated via

$$\begin{aligned} l^{\text{ch.}}_{\mathcal{SD},k} &= \log\left(\frac{p_{Y_{\mathcal{SD}}|X_{\mathcal{S}}}(y_{\mathcal{SD},k}|x_{\mathcal{S},k}=+1)}{p_{Y_{\mathcal{SD}}|X_{\mathcal{S}}}(y_{\mathcal{SD},k}|x_{\mathcal{S},k}=-1)}\right) \\ &= \frac{2a^2_{\mathcal{SD}}}{\sigma^2_N}\cdot y_{\mathcal{SD},k}\ , \end{aligned} \tag{6.4}$$

given the noise variance $\sigma^2_N/a^2_{\mathcal{SD}}$ at the destination, which includes the path-loss attenuation factor $a_{\mathcal{SD}}$. In (6.4), index k denotes the time instance k and superscript ch. points out that these LLR values belong to the channel output.

The same procedure is taken for the LLR calculation for the $\mathcal{S}\to\mathcal{R}$ link. Here, the channel is again a BIAWGN channel and, therefore, the LLRs for the received signal (6.1) are calculated as well with

$$\begin{aligned} l^{\text{ch.}}_{\mathcal{SR},k} &= \log\left(\frac{p_{Y_{\mathcal{SR}}|X_{\mathcal{S}}}(y_{\mathcal{SR},k}|x_{\mathcal{S},k}=+1)}{p_{Y_{\mathcal{SR}}|X_{\mathcal{S}}}(y_{\mathcal{SR},k}|x_{\mathcal{S},k}=-1)}\right) \\ &= \frac{2a^2_{\mathcal{SR}}}{\sigma^2_N}\cdot y_{\mathcal{SR},k}\ . \end{aligned} \tag{6.5}$$

In (6.5), $\sigma^2_N/a^2_{\mathcal{SR}}$ denotes the noise variance at the relay node including the path-loss attenuation factor $a_{\mathcal{SR}}$ for the source relay link.

6.1.2 Relay destination link

Until now, the determination of the LLRs for the $\mathcal{S}\to\mathcal{R}$ link and the $\mathcal{S}\to\mathcal{D}$ link was simple, as the channel input distribution is given by only two symbols $x_{\mathcal{S},k}\in\{+1,-1\}$ which are equal probable. To calculate the log-likelihood ratios at the destination for the relay link, a more sophisticated approach needs to be taken. Here it is essential to consider the output type and, therewith, the output distribution of the relay's encoder into the calculation [Vor13].

In case of an *hard-input hard-output* (HIHO) re-encoder (see Subsection 2.3.1), the encoder output is binary and the transmit signal $\mathbf{x}_{\mathcal{R}}$, $x_{\mathcal{R},m} \in \{+1, -1\}$ is BPSK modulated. Thus, the log-likelihood ratios are given by

$$\begin{aligned} l^{\text{ch.}}_{\mathcal{RD},m} &= \log\left(\frac{p_{Y_{\mathcal{RD}}|X_{\mathcal{R}}}(y_{\mathcal{RD},m}|x_{\mathcal{R},m}=+1)}{p_{Y_{\mathcal{RD}}|X_{\mathcal{R}}}(y_{\mathcal{RD},m}|x_{\mathcal{R},m}=-1)}\right) \\ &= \frac{2a^2_{\mathcal{RD}}}{\sigma^2_N}\cdot y_{\mathcal{RD},m}\ , \end{aligned} \tag{6.6}$$

with the noise variance $\sigma^2_N / a^2_{\mathcal{RD}}$ at the destination.

If the relay applies a *soft-input soft-output* (SISO) re-encoder, then the encoder $\mathcal{C}_{\mathcal{R}}$ delivers a non-binary but real valued sequence of log-likelihood ratios $\mathbf{l}^{\text{code}}_{\mathcal{R}} \in \{\mathbb{R}\}$, which are directly calculated (via the encoder) from the interleaved decoder output LLRs $\mathbf{l}^{\text{info}}_{\mathcal{R}} = \Pi\left(\mathbf{l}^{\text{info}}_{\mathcal{SR}}\right)$ given by $\mathcal{D}_{\mathcal{S}}$. This is depicted in **Fig. 6.2**. Here, the

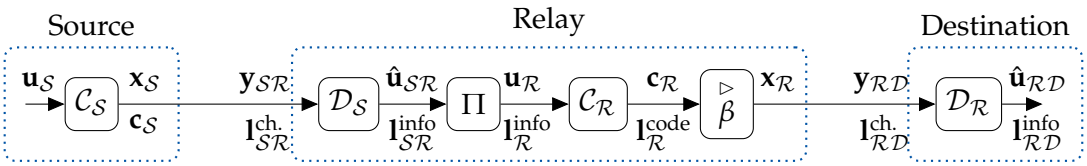


Fig. 6.2.: Schematic overview of the $\mathcal{S} \to \mathcal{R} \to \mathcal{D}$ link

interleaving of the LLRs ($\mathbf{l}^{\text{info}}_{\mathcal{SR}}$) is essential to break the correlation among preceding and succeeding soft symbols. This is beneficial in order to cope with error-bursts on the consecutive channel and it ensures the independence of cascaded channels. Surely, the correlation is re-established after de-interleaving at the destination. But for the $\mathcal{R} \to \mathcal{D}$ link, neighboring LLRs are uncorrelated as long as the transmit sequence is large enough and the interleaver is perfectly random.

Furthermore, it is noteworthy for the upcoming paragraphs that interleaving does not change the approximately bimodal Gaussian distribution. Thus, the pdf of the output sequence $\mathbf{l}^{\text{info}}_{\mathcal{R}}$ is identical to the pdf of the input sequence $\mathbf{l}^{\text{info}}_{\mathcal{SR}}$. Assuming no specific order in sign or magnitude of the LLRs, i.e. the LLRs are perfectly random distributed, the distribution of $\mathbf{l}^{\text{code}}_{\mathcal{R}}$ after the encoder output $\mathcal{D}_{\mathcal{S}}$ can be determined. This will be shown later (see for example equation (6.13) and (6.15)). For now, the general description of the relay node from **Fig. 6.2** will be continued.

After re-encoding, the relay has different options to alternate the LLR signal $\mathbf{l}^{\text{code}}_{\mathcal{R}}$ before normalizing the transmit power. It can either leave the LLR values untouched or apply the tanh-function, generating soft-bits.

In the first case, the encoder output sequence $\mathbf{l}_{\mathcal{R}}^{\text{code}} \in \{\mathbb{R}\}$ needs to be scaled with $\beta = 1/\,\mathrm{E}\left\{\mathbf{l}_{\mathcal{R}}^{\text{code}}\right\}$

$$\mathbf{x}_{\mathcal{R}} = \beta \cdot \mathbf{l}_{\mathcal{R}}^{\text{code}} \ , \tag{6.7}$$

in order to retain unit variance $\sigma_{X_{\mathcal{R}}}^2 = 1$ of $\mathbf{x}_{\mathcal{R}} \in \{\mathbb{R}\}$. Thus, the relay's transmit sequence $\mathbf{x}_{\mathcal{R}}$ is a scaled version of the encoder output LLRs. In the second case, i.e. transmitting soft-bits (see Appendix A.5), the output sequence is given by

$$\mathbf{x}_{\mathcal{R}} = \beta \cdot \tanh\left(\frac{\mathbf{l}_{\mathcal{R}}^{\text{code}}}{2}\right) \ , \tag{6.8}$$

with the scaling factor $\beta = 1/\,\mathrm{E}\left\{\tanh\left(\frac{\mathbf{l}_{\mathcal{R}}^{\text{code}}}{2}\right)\right\}$.

Again, the problem of the SISO encoder and the subsequent forwarding of LLRs or soft-bits is that the relay's transmit signal is, in both cases, non-binary anymore. Therefore, the definition of LLRs in (6.6) has to be extended. A solution to this problem is given by reinterpreting the non-binary random values $x_{\mathcal{R},m} \in \{\mathbb{R}\}$ as a binary signal with a time specific attenuation $\tilde{a}_{\mathcal{R},m}$ factor[2], i.e.

$$x_{\mathcal{R},m} = \tilde{a}_{\mathcal{R},m} \cdot \tilde{x}_{\mathcal{R},m} \ , \tag{6.9}$$

where $\tilde{a}_{\mathcal{R},m} = |x_{\mathcal{R},m}|$, $\tilde{a}_{\mathcal{R},m} \in \{\mathbb{R}^+\}$ is non-negative real valued and $\tilde{x}_{\mathcal{R},m} = \mathrm{sign}\,(x_{\mathcal{R},m})$, $\tilde{x}_{\mathcal{R},m} \in \{+1, -1\}$ is BPSK modulated.

The $\mathcal{R} \to \mathcal{D}$ receive signal (6.3) can then be rewritten as

$$\begin{aligned} y_{\mathcal{R}\mathcal{D},m} &= x_{\mathcal{R},m} + n_{\mathcal{D},m}^{T_2} \\ &= \tilde{a}_{\mathcal{R},m} \cdot \tilde{x}_{\mathcal{R},m} + n_{\mathcal{D},m}^{T_2} \ . \end{aligned} \tag{6.10}$$

Since the coefficient $\tilde{a}_{\mathcal{R},m}$ is unknown at the destination, the LLRs for (6.10) maybe evaluated by applying the total probability theorem and averaging w.r.t. to the unknown magnitude via

$$l_{\mathcal{R}\mathcal{D},m} = \log\left(\frac{\sum\limits_{\tilde{a}_{\mathcal{R},m} \in \mathbb{R}^+} p_{Y_{\mathcal{R}\mathcal{D}}|\tilde{X}_{\mathcal{R}},\tilde{A}_{\mathcal{R}}}\left(y_{\mathcal{R}\mathcal{D},m}|\tilde{x}_{\mathcal{R},m} = +1, \tilde{a}_{\mathcal{R},m}\right) \cdot p_{\tilde{A}_{\mathcal{R}}}\left(\tilde{a}_{\mathcal{R}}\right)}{\sum\limits_{\tilde{a}_{\mathcal{R},m} \in \mathbb{R}^+} p_{Y_{\mathcal{R}\mathcal{D}}|\tilde{X}_{\mathcal{R}},\tilde{A}_{\mathcal{R}}}\left(y_{\mathcal{R}\mathcal{D},m}|\tilde{x}_{\mathcal{R},m} = -1, \tilde{a}_{\mathcal{R},m}\right) \cdot p_{\tilde{A}_{\mathcal{R}}}\left(\tilde{a}_{\mathcal{R}}\right)}\right) \ , \tag{6.11}$$

i.e. calculating the expected value for the conditional probability density by integrating over $\tilde{a}_{\mathcal{R}}$.

[2] Please do not confuse the path-loss attenuation factor $a_{\mathcal{R}\mathcal{D}}$ with the coefficient $\tilde{a}_{\mathcal{R}}$.

Still the distribution $F_{\tilde{A}_\mathcal{R}}(\tilde{a}_\mathcal{R})$ for the discrete random variable $\tilde{A}_\mathcal{R} \in \{\mathbb{R}^+\}$ is unknown and will be derived in the following paragraphs. To do so, an assumption over the distribution $F_{L^{\text{info}}_{\mathcal{SR}}}\left(l^{\text{info}}_{\mathcal{SR}}\right)$ of the LLR sequence from the output of the relay's decoder $\mathcal{D}_\mathcal{S}$ has to be made. Here, it is assumed that the maximum-a-posteriori decoder delivers a bimodal Gaussian distributed sequence with probability density function

$$p_{L^{\text{info}}_{\mathcal{SR}}}\left(l^{\text{info}}_{\mathcal{SR}}\right) = \frac{1}{2}\left(\mathcal{N}\left(\mu_{L^{\text{info}}_{\mathcal{SR}}}, \sigma^2_{L^{\text{info}}_{\mathcal{SR}}}\right) + \mathcal{N}\left(-\mu_{L^{\text{info}}_{\mathcal{SR}}}, \sigma^2_{L^{\text{info}}_{\mathcal{SR}}}\right)\right) \quad . \tag{6.12}$$

with mean $\mu_{L^{\text{info}}_{\mathcal{SR}}}$ and variance $\sigma^2_{L^{\text{info}}_{\mathcal{SR}}}$.

Please note, although the maximum-a-posteriori decoder (see Section 2.3.1) emits a sequence of values, the distribution of these values does not follow the typical assumption of log-likelihood ratios. Here, especially the assumption of $\mu_{L^{\text{info}}_{\mathcal{SR}}} = \sigma^2_{L^{\text{info}}_{\mathcal{SR}}}/2$ is not fulfilled. The drift between the measured mean and the assumption of $\mu_{L^{\text{info}}_{\mathcal{SR}}} = \sigma^2_{L^{\text{info}}_{\mathcal{SR}}}/2$ increases especially in cases, where the decoding result improves. This is depicted at the bottom of **Fig. 6.3** for a non-recursive non-systematic convolutional encoder with generator polynomial $\mathbf{g} = [25;\ 33;\ 37]_8$ and a MAP decoder. The picture also provides the conditional pdfs for two cases with different channel noise variances $\sigma^2_N = 1.0$ and $\sigma^2_N = 0.1$. In the upper left figure, all conditional pdfs are nearly the same. This is different for the upper right figure. Here, with decreasing channel noise variance and, therefore, increasing $\sigma^2_{L^{\text{info}}_{\mathcal{SR}}}$ the conditional pdf given by the LLR assumption and the measured pdf diverges. Only the conditional pdf with measured mean and variance tightly predicts the measured conditional pdf from the decoder output. Still, the assumption of a bimodal Gaussian distribution (given by (6.12)) at the decoder output closely holds in case of a MAP decoder and decent channel noise.

A closer view on the topic (prediction of pdfs at the decoder output) was made in [ASR04], which is based on [RT02]. But the approaches taken in both publications are very complex and their application to the given problem will not greatly improve the outcome compared to the Gaussian approximation. So, in this thesis the decoder output is assumed to be bimodal Gaussian distributed.

With $p_{L^{\text{info}}_{\mathcal{SR}}}\left(l^{\text{info}}_{\mathcal{SR}}\right)$ one can derive the output density of the soft-input soft-output encoder. Since the interleaver does not change the distribution, the distribution of $L^{\text{info}}_{\mathcal{R}}$ is the same as for $L^{\text{info}}_{\mathcal{SR}}$, i.e. $p_{L^{\text{info}}_{\mathcal{R}}}\left(l^{\text{info}}_{\mathcal{R}}\right) = p_{L^{\text{info}}_{\mathcal{SR}}}\left(l^{\text{info}}_{\mathcal{SR}}\right)$ from (6.12) holds.

The next step is to calculate the densities at the relay's encoder output. For this, a simple SISO convolutional encoder is depicted in **Fig. 6.4**. Here, one can see that more than one density, i.e. one for every code sequence, must be evaluated.

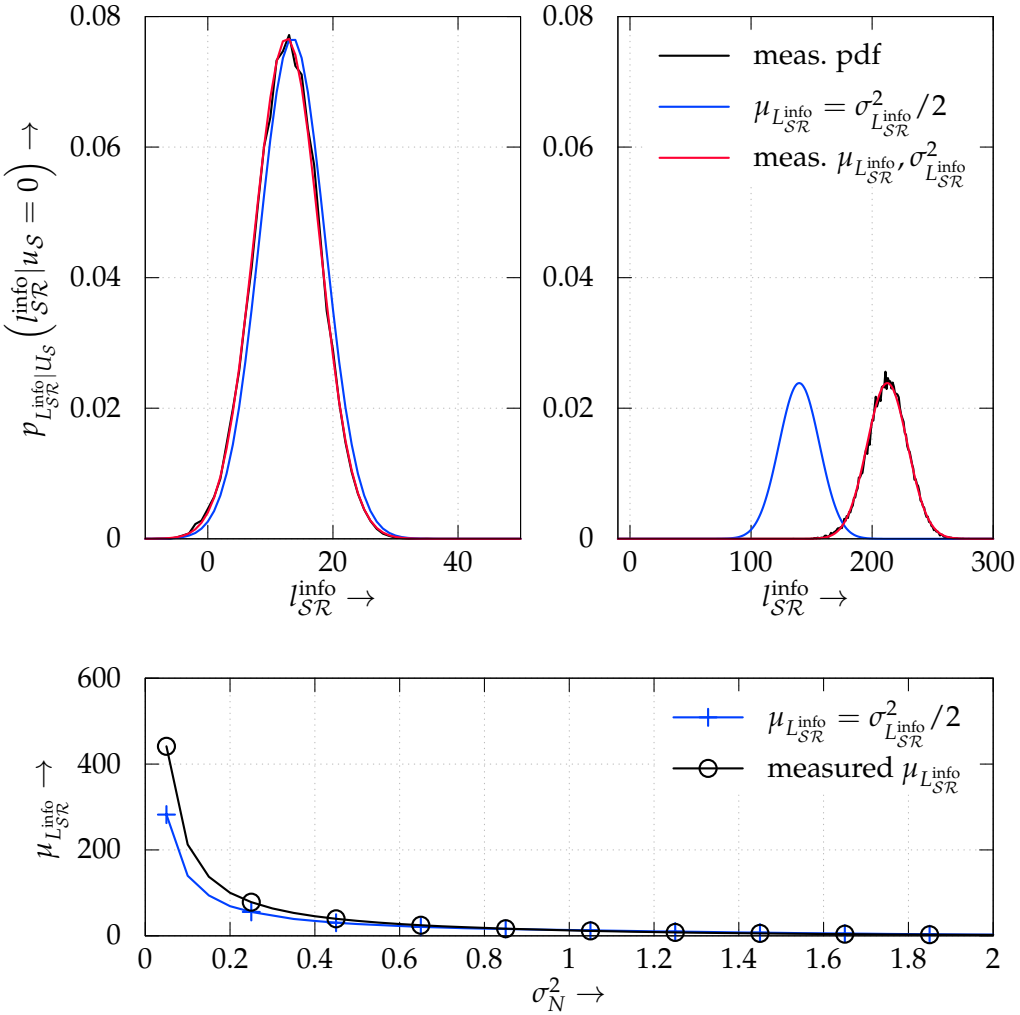


Fig. 6.3.: Conditional probability density functions versus the LLR values at the output of the MAP decoder. Left figure with $\sigma_N^2 = 1.0$, $\sigma^2_{L_{\mathcal{SR}}^{\text{info}}} \approx 27.54$ and $\mu_{L_{\mathcal{SR}}^{\text{info}}} \approx 12.57$. Right figure with $\sigma_N^2 = 0.1$, $\sigma^2_{L_{\mathcal{SR}}^{\text{info}}} \approx 282.99$ and $\mu_{L_{\mathcal{SR}}^{\text{info}}} \approx 212.42$. The figure below depicts the drift between the measured mean and the assumption of $\mu_{L_{\mathcal{SR}}^{\text{info}}} = \sigma^2_{L_{\mathcal{SR}}^{\text{info}}}/2$.

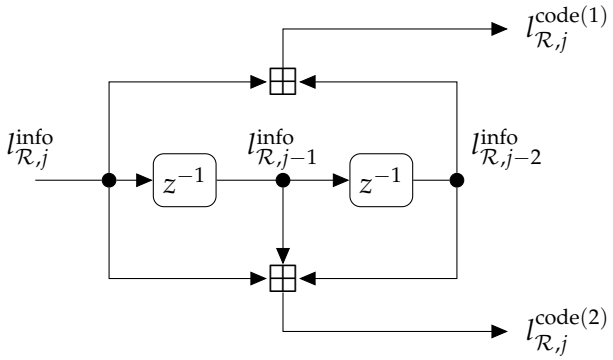


Fig. 6.4.: Soft-input soft-output encoder based on boxplus operations

The number of densities is given by the number of decoder outputs. Thus, for the example encoder from **Fig. 6.4**, there exists two outputs and, thus, two densities. Furthermore, each code LLR sequence $\mathbf{l}_{\mathcal{R}}^{\text{code}(1,2)}$ is given by one or more *boxplus* operations. For example, $l_{\mathcal{R},j}^{\text{code}(1)} = l_{\mathcal{R},j}^{\text{info}} \boxplus l_{\mathcal{R},j-2}^{\text{info}}$ holds. Therefore, the distribution for the random variable $L_{\mathcal{R}}^{\text{code}(1)}$ is based on the boxplus density (see Appendix A.4.10)

$$p_{L_{\mathcal{R},j}^{\text{code}(1)}}\left(l_{\mathcal{R},j}^{\text{code}(1)}\right) = \mathcal{F}^{\boxplus}\left(p_{L_{\mathcal{R},j}^{\text{info}}}\left(l_{\mathcal{R},j}^{\text{info}}\right), p_{L_{\mathcal{R},j-2}^{\text{info}}}\left(l_{\mathcal{R},j-2}^{\text{info}}\right)\right) \ , \tag{6.13}$$

given the density for $L_{\mathcal{R}}^{\text{info}}$. In (6.13), $\mathcal{F}^{\boxplus}(\ldots)$ denotes the function for calculating the output pdf resulting after the boxplus-operation. Since the random variables $L_{\mathcal{R},j}^{\text{info}}$, $L_{\mathcal{R},j-1}^{\text{info}}$ and $L_{\mathcal{R},j-2}^{\text{info}}$ are equally distributed, equation (6.13) can be rewritten as

$$p_{L_{\mathcal{R}}^{\text{code}(1)}}\left(l_{\mathcal{R}}^{\text{code}(1)}\right) = \mathcal{F}^{\boxplus}\left(p_{L_{\mathcal{R}}^{\text{info}}}\left(l_{\mathcal{R}}^{\text{info}}\right), p_{L_{\mathcal{R}}^{\text{info}}}\left(l_{\mathcal{R}}^{\text{info}}\right)\right) \ . \tag{6.14}$$

The same approach is taken for $p_{L_{\mathcal{R}}^{\text{code}(2)}}\left(l_{\mathcal{R}}^{\text{code}(2)}\right)$ with

$$p_{L_{\mathcal{R}}^{\text{code}(2)}}\left(l_{\mathcal{R}}^{\text{code}(2)}\right) = \mathcal{F}^{\boxplus}\left(\mathcal{F}^{\boxplus}\left(p_{L_{\mathcal{R}}^{\text{info}}}\left(l_{\mathcal{R}}^{\text{info}}\right), p_{L_{\mathcal{R}}^{\text{info}}}\left(l_{\mathcal{R}}^{\text{info}}\right)\right), p_{L_{\mathcal{R}}^{\text{info}}}\left(l_{\mathcal{R}}^{\text{info}}\right)\right) \ , \tag{6.15}$$

where $\mathcal{F}^{\boxplus}(\ldots)$ is called multiple times in order to get the pdf for three inputs. Thus, the output density for more than one boxplus operation is calculated recursively with no specific order of the input. Since the mean and variance of the output sequence decreases after each boxplus operation, the output density for a recursive code is given by a Dirac delta distribution [WWKK08]. The same holds for non-recursive codes with an increasing number of boxplus operations.

This is depicted in **Fig. 6.5**. Here, the initial mean and variance of the random variable L where $\mu_L = 20$ and $1/\sigma_L = 40$. One can see that with increasing boxplus-operations, the mean and variance decreases. After almost 15 operations in the given scenario, the mean and variance are close to zero and the Dirac impulse is created.

An approximation [WW96] of the boxplus operation and its corresponding density is given by the min-sum operation and density in Appendix A.4.10. This approximation can be evaluated semi-analytically and is therefore utilized throughout this dissertation for performing simulations. Equivalently to $\mathcal{F}^{\boxplus}(\ldots)$, the short hand notation $\mathcal{F}^{\tilde{\boxplus}}(\ldots)$ is utilized for calculating the output density in case of the min-sum approximation.

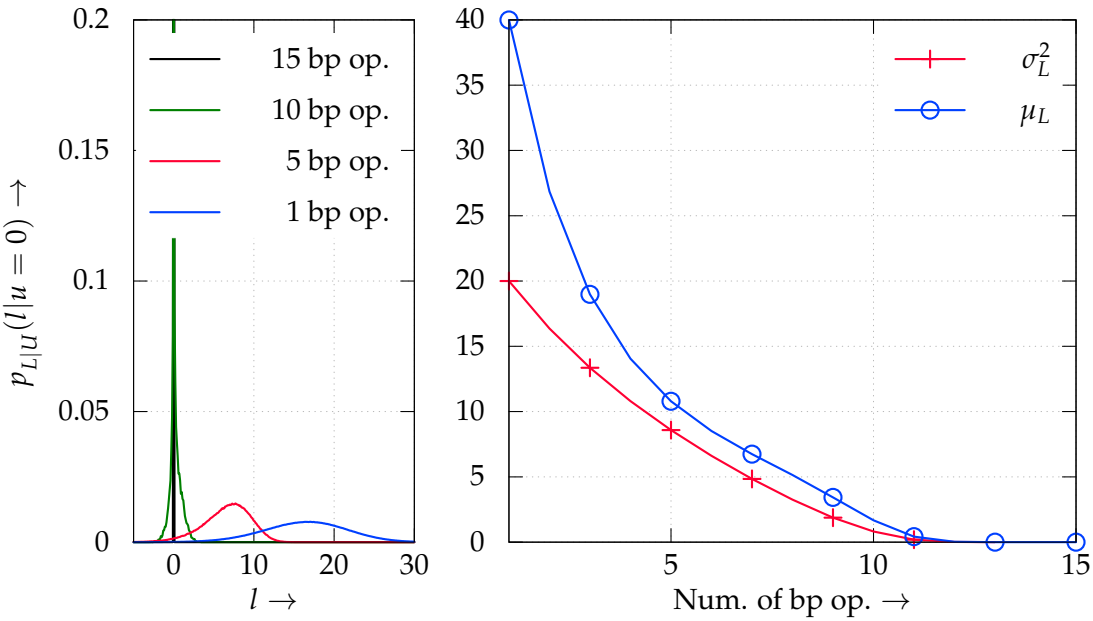


Fig. 6.5.: Conditional pdfs, mean and variance for the LLR values after a certain number of boxplus operations (bp op.)

In **Fig. 6.6**, the conditional probability density functions of a non-recursive non-systematic CC soft-encoder with $\mathbf{g} = [5;7]_8$ are depicted. The condition is taken on the true code sequence $\mathbf{c}_S$, which was generated from a non-disturbed binary sequence $\mathbf{u}_S$ and with the help of a hard-input hard-output encoder with the same parameters as above. It can be observed that the outcome for the approximation of the boxplus density via the min-sum density increases with increasing $\sigma^2_{L_\mathcal{R}^{\text{info}}}$. Furthermore, the figure justifies the separate treatment of the distributions for $L_\mathcal{R}^{\text{code}(1)}$ and $L_\mathcal{R}^{\text{code}(2)}$. Both are different due to the different numbers of boxplus-operations.

After calculating the densities at the output of the soft-encoder, they have to be linearly transformed to include the scaling factor β as used in (6.7) or (6.8). This is done by the scaling function $\mathcal{F}^{\triangleright}(\ldots)$ (see Appendix A.4.6) via

$$p_{X_\mathcal{R}^{\text{code}(1)}}\left(x_\mathcal{R}^{\text{code}(1)}\right) = \mathcal{F}^{\triangleright}\left(f_{L_\mathcal{R}^{\text{code}(1)}}\left(l_\mathcal{R}^{\text{code}(1)}\right), \beta\right) \tag{6.16}$$

for the first decoder output sequence and with

$$p_{X_\mathcal{R}^{\text{code}(2)}}\left(x_\mathcal{R}^{\text{code}(2)}\right) = \mathcal{F}^{\triangleright}\left(f_{L_\mathcal{R}^{\text{code}(2)}}\left(l_\mathcal{R}^{\text{code}(2)}\right), \beta\right) \tag{6.17}$$

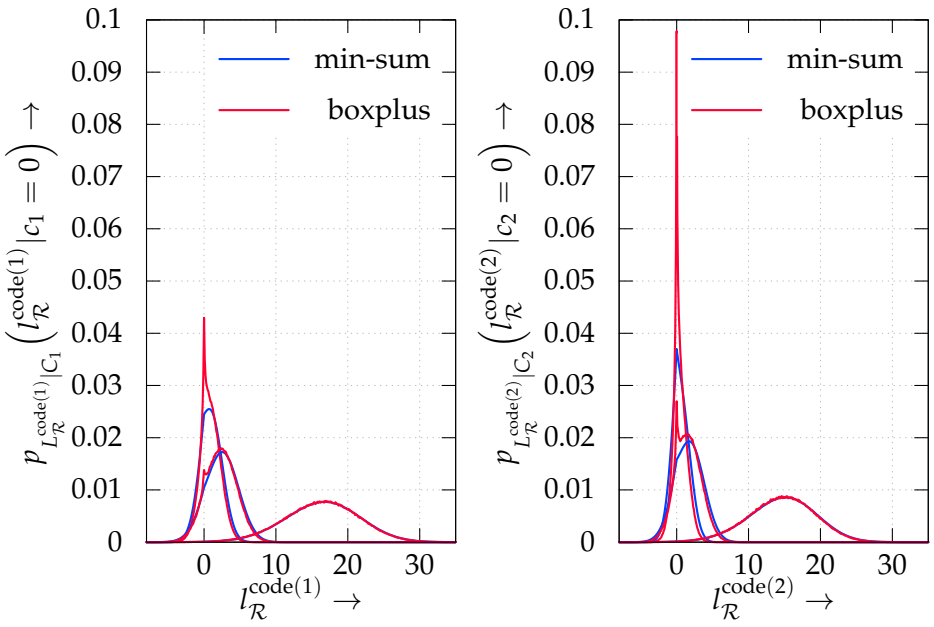


Fig. 6.6.: Measured and predicted (via min-sum) conditional pdfs for different variances at the soft-encoder output. Dotted line: $\sigma^2_{L_{\mathcal{R}}^{\text{info}}} = 4$. Dashed line: $\sigma^2_{L_{\mathcal{R}}^{\text{info}}} = 8$. Solid line: $\sigma^2_{L_{\mathcal{R}}^{\text{info}}} = 40$. Left figure for c_1. Right figure for c_2.

for the second output, respectively. It is obvious that the transformation has to be performed for each code sequence separately, but with the same scaling factor β which was already used in (6.7) and (6.8).

The relationship between relay's transmit sequence $\mathbf{x}_{\mathcal{R}}$ and the encoder output sequences $\mathbf{x}_{\mathcal{R}}^{\text{code}(1)}$ and $\mathbf{x}_{\mathcal{R}}^{\text{code}(2)}$ is (for example) given by

$$\mathbf{x}_{\mathcal{R}} = \left[\ldots, x_{\mathcal{R},j-1}^{\text{code}(2)}, x_{\mathcal{R},j}^{\text{code}(1)}, x_{\mathcal{R},j}^{\text{code}(2)}, x_{\mathcal{R},j+1}^{\text{code}(1)}, \ldots\right] \quad . \tag{6.18}$$

This strictly depends on the order in which the relay will transmit the symbols.

If the relay is supposed to transmit soft bits, then the boxplus density has to be additionally transformed with $\mathcal{F}^{\curlywedge}(\ldots)$, representing the derived distribution for the tanh $\left(L_{\mathcal{R}}^{\text{code}(1)}/2\right)$ function, before performing the scaling via

$$p_{X_{\mathcal{R}}^{\text{code}(1)}}\left(x_{\mathcal{R}}^{\text{code}(1)}\right) = \mathcal{F}^{\triangleright}\left(\mathcal{F}^{\curlywedge}\left(p_{L_{\mathcal{R}}^{\text{code}(1)}}\left(l_{\mathcal{R}}^{\text{code}(1)}\right)\right), \beta\right) \quad . \tag{6.19}$$

The function $\mathcal{F}^{\lambda}(\ldots)$ is given in the Appendix A.4.7. The same approach is taken for $p_{L_{\mathcal{R}}^{\text{code}(2)}}\left(l_{\mathcal{R}}^{\text{code}(2)}\right)$ in order to get $p_{X_{\mathcal{R}}^{\text{code}(2)}}\left(x_{\mathcal{R}}^{\text{code}(2)}\right)$.

Once the densities for $X_{\mathcal{R}}^{\text{code}(1)}$ and $X_{\mathcal{R}}^{\text{code}(2)}$ are calculated, one can easily derive the corresponding folded densities (see Appendix A.4.8) $p_{\tilde{A}_{\mathcal{R}}^{\text{code}(1)}}\left(\tilde{a}_{\mathcal{R}}^{\text{code}(1)}\right)$ and $p_{\tilde{A}_{\mathcal{R}}^{\text{code}(2)}}\left(\tilde{a}_{\mathcal{R}}^{\text{code}(2)}\right)$ for the attenuation factors $\tilde{A}_{\mathcal{R}}^{\text{code}(1)}$ and $\tilde{A}_{\mathcal{R}}^{\text{code}(2)}$. They represent the distribution of the relay's transmit sequence (per the encoder's code output) and their knowledge is mandatory to calculate the true channel LLR values at the destination's decoder input according to (6.10) and (6.11).

The densities for the attenuation factors are calculated via

$$p_{\tilde{A}_{\mathcal{R}}^{\text{code}(1)}}\left(\tilde{a}_{\mathcal{R}}^{\text{code}(1)}\right) = \mathcal{F}^{|X_{\mathcal{R}}^{\text{code}(1)}|}\left(p_{X_{\mathcal{R}}^{\text{code}(1)}}\left(x_{\mathcal{R}}^{\text{code}(1)}\right)\right) \quad , \tag{6.20}$$

and

$$p_{\tilde{A}_{\mathcal{R}}^{\text{code}(2)}}\left(\tilde{a}_{\mathcal{R}}^{\text{code}(2)}\right) = \mathcal{F}^{|X_{\mathcal{R}}^{\text{code}(2)}|}\left(p_{X_{\mathcal{R}}^{\text{code}(2)}}\left(x_{\mathcal{R}}^{\text{code}(2)}\right)\right) \quad , \tag{6.21}$$

where $\mathcal{F}^{|X_{\mathcal{R}}^{\text{code}(1)}|}(\ldots)$ denotes the density function for the derived distribution of an absolute value. Thus, applying (6.20) and (6.21) will deliver the density for $\tilde{A}_{\mathcal{R}} = |X_{\mathcal{R}}|$ as needed according to equation (6.10).

Under the assumption of $\sigma^2_{L_{\mathcal{R}}^{\text{info}}} = 8$, the conditional densities of the attenuation $\tilde{A}_{\mathcal{R}}^{\text{code}(1)}$ and $\tilde{A}_{\mathcal{R}}^{\text{code}(2)}$ are depicted in **Fig. 6.7**. With a channel noise variance $\sigma_N^2 = 0.1$ on the second hop, the conditional densities $p_{Y_{\mathcal{R}\mathcal{D}}|C_{\mathcal{R}}}(y_{\mathcal{R}\mathcal{D}}|c_{\mathcal{R}} = 0)$ (for code(1) and code(2)) at the destination's decoder input of a non-recursive non-systematic convolutional code with $\mathbf{g} = [5;7]_8$, are depicted on the right side of **Fig. 6.7**. These are the conditional pdfs, from which the LLR values will be calculated. It can be observed that the pdfs will slope towards the zero line. Thus, the LLR values at the channel output are not bimodal Gaussian distributed as for a simple BIAWGN channel. Here, due to the soft-encoding process, small LLRs are more likely than larger ones.

Once the distributions of the attenuation $\tilde{A}_{\mathcal{R}}^{\text{code}}$ is calculated, formulas (6.20) and (6.21) are inserted into (6.11) in order to individually (per re-encoder output) calcu-

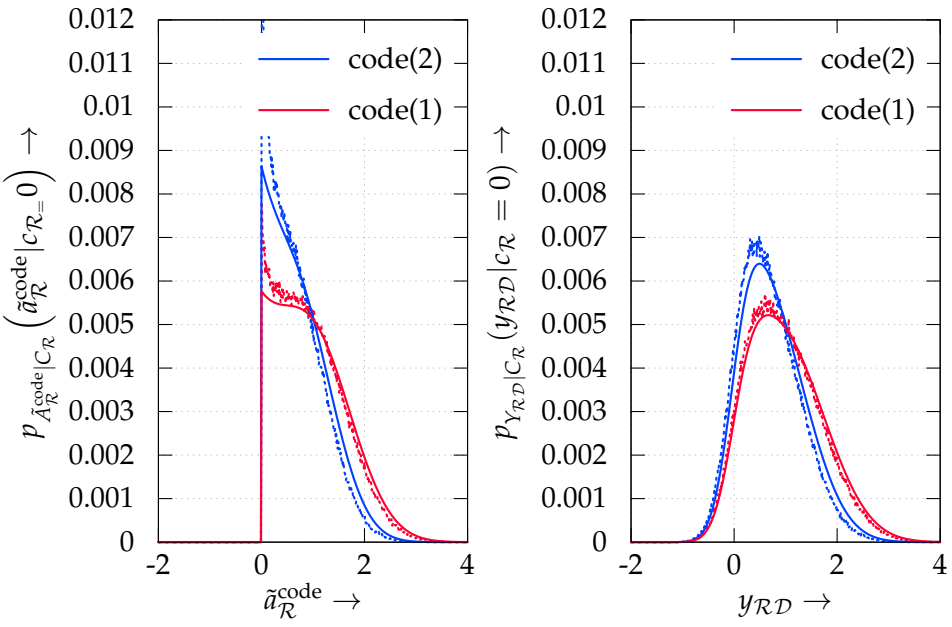


Fig. 6.7.: Measured (dashed line) and predicted (solid line) conditional pdfs of the attenuation factor (at the relay) (left figure) and the received channel output at the destination (right figure).

late the log-likelihood ratios at the destinations decoder input. Thus, depending on the code sequence (code(1) or code(2)), the LLRs are given by

$$L_{\mathcal{RD},m} = \log\left(\frac{\sum_{\tilde{a}_{\mathcal{R},m}\in\mathbb{R}^+} p_{Y_{\mathcal{RD}}|\tilde{X}_{\mathcal{R}},\tilde{A}_{\mathcal{R}}}\left(y_{\mathcal{RD},m}|\tilde{x}_{\mathcal{R},m}=+1,\tilde{a}_{\mathcal{R},m}\right)\cdot p_{\tilde{A}_{\mathcal{R}}}\left(\tilde{a}_{\mathcal{R}}\right)}{\sum_{\tilde{a}_{\mathcal{R},m}\in\mathbb{R}^+} p_{Y_{\mathcal{RD}}|\tilde{X}_{\mathcal{R}},\tilde{A}_{\mathcal{R}}}\left(y_{\mathcal{RD},m}|\tilde{x}_{\mathcal{R},m}=-1,\tilde{a}_{\mathcal{R},m}\right)\cdot p_{\tilde{A}_{\mathcal{R}}}\left(\tilde{a}_{\mathcal{R}}\right)}\right) \quad , \quad (6.22)$$

To conclude, in order to evaluate (6.22) the destination only needs to receive knowledge over the mean $\mu_{L_{\mathcal{SR}}^{\text{info}}}$ or variance $\sigma^2_{L_{\mathcal{SR}}^{\text{info}}}$ of the relay's decoded log-likelihood ratio sequence. If both are received, a better estimation of the distribution for the relay's decoder output sequence can be made. Eventually, this will strongly improve the prediction of the LLR distribution at the destination's input. Furthermore, the described method is valid for any soft-input soft-output re-encoder for which the output distribution can be evaluated from its input distribution. This holds, for example, for any recursive or non-recursive convolutional soft re-encoder, since these encoders can be modeled by a series of boxplus operations.

6.1.3 Performance of soft and hard re-encoding

In this subsection, the performance of the soft-input soft-output encoder is evaluated and compared with the performance of a hard-input hard-output system. For this, the system model from **Fig. 6.2** is utilized. Please note, that the decoding results and the given assumptions in this subsection can be different from the assumptions made for the three-node relay network in Section 6.2.4. This is due to the missing direct link throughout this subsection. But the results can deliver insights, on where soft re-encoding at the relay might be beneficial compared to a hard re-encoding given an overall relay network.

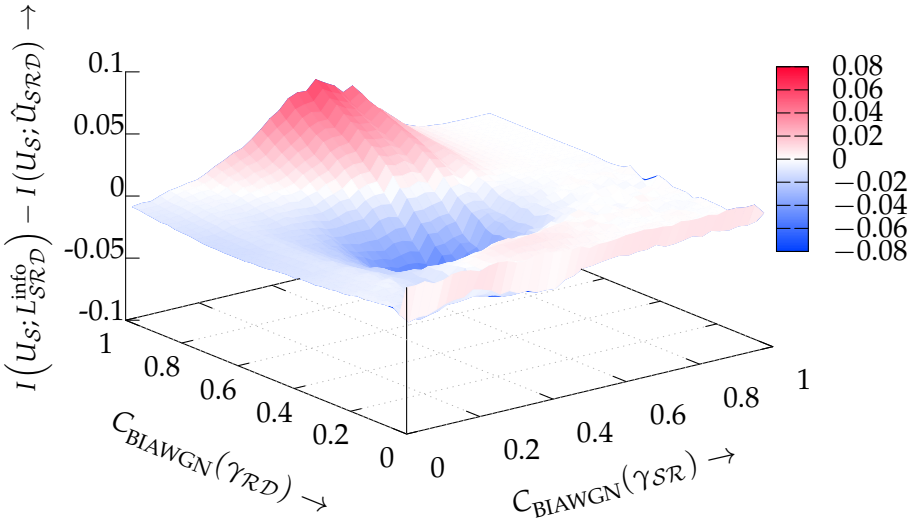


Fig. 6.8.: Difference in MI between SISO and HIHO at the relay.

The difference in mutual information between $I\left(U_\mathcal{S}; L^{\text{info}}_{\mathcal{SRD}}\right)$ and $I\left(U_\mathcal{S}; \hat{U}_{\mathcal{SRD}}\right)$ with $L^{\text{info}}_{\mathcal{SRD}} = \Pi^{-1}\left(L^{\text{info}}_{\mathcal{RD}}\right)$ and $\hat{U}_{\mathcal{SRD}} = \Pi^{-1}\left(\hat{U}_{\mathcal{RD}}\right)$, for SISO and HIHO re-encoding and transmitting soft-bits (6.8) at the relay node, is depicted in **Fig. 6.8**. Here, $\mathcal{C}_\mathcal{S}$ is a non-recursive non-systematic CC with $R_c^{\text{CC}} = 1/2$ and generator polynomials $\mathbf{g} = [5;7]_8$. For the relays encoder also a non-recursive non-systematic CC but with $R_c^{\text{CC}} = 1/3$ and polynomials $\mathbf{g} = [5;7;7]_8$ has been utilized. The difference in MI maybe not an adequate measurement, as it is not relative to the optimum achievable MI, but it is sufficient for the purpose of finding areas in which soft or hard re-encoding can be superior.

If the channel capacity of the $\mathcal{S} \rightarrow \mathcal{R}$ link is weak ($C_{\text{BIAWGN}}(\gamma_{\mathcal{SR}}) < 0.2$), then the relay cannot decode the source's information, resulting in a small $I\left(U_\mathcal{S}; \hat{U}_{\mathcal{SR}}\right) \approx 0$ or high BER and, thus, small LLR values at the relay's decoder output. Therefore,

the distribution for $X_{\mathcal{R}}$ after SISO re-encoding and forming soft-bits is concentrated around zero. This is depicted on the left side in **Fig. 6.9**.

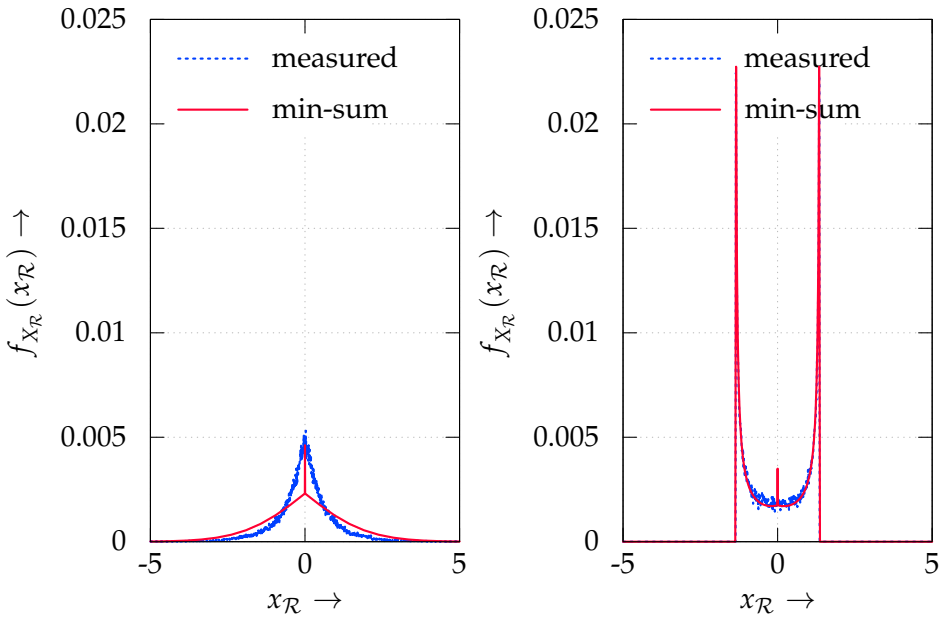


Fig. 6.9.: Measured and predicted (via min-sum) pmfs of the relay's transmit signal $\mathbf{x}_{\mathcal{R}}$. Left side: $C_{\text{BIAWGN}}(\gamma_{\mathcal{SR}}) = 0.025$. Right side: $C_{\text{BIAWGN}}(\gamma_{\mathcal{SR}}) = 0.5$.

In this situation, the relay cannot further improve the communication between the source and destination. Therefore, the difference $I\left(U_{\mathcal{S}}; L^{\text{info}}_{\mathcal{SRD}}\right) - I\left(U_{\mathcal{S}}; \hat{U}_{\mathcal{SRD}}\right)$ is concentrated around zero. Thus, neither HIHO nor SISO encoding does additionally increase the mutual information $I\left(U_{\mathcal{S}}; \hat{U}_{\mathcal{SD}}\right)$ of the source's information at the destination from the direct link.

On a side note, it is worth to mention, that a HIHO encoder with its resulting BPSK distribution at the relay's output, can achieve a higher MI $I\left(U_{\mathcal{R}}; \hat{U}_{\mathcal{RD}}\right)$ of the relay's information sequence at the destination. Still, this depends on the channel quality of the second hop and will not increase the knowledge of the source's information.

In case of a decent first hop ($C_{\text{BIAWGN}}(\gamma_{\mathcal{SR}}) = 0.3 \ldots 0.6$), the relay partly decodes the source's information. This means, that the MAP decoder outputs LLR values with a high variance in their amplitudes. Re-encoding these LLRs and forming soft-bits results in a distribution of the relay's transmit signal as depicted on the right side of **Fig. 6.9**. Contrarily, in case of HIHO re-encoding at the relay, all symbols will get the same power independently of their reliability.

If $C_{\text{BIAWGN}}(\gamma_{\mathcal{RD}})$ is low, then both re-encoding methods will have a weak performance and a resulting BER around 0.5. Once the second hop gets decent, HIHO re-encoding can achieve a better performance than SISO re-encoding as depicted in **Fig. 6.8**. Here, depending on the utilized practical encoding/decoding scheme, the loss in case of a hard-decision can be compensated by a better decoding performance of the second hop. In case of SISO re-encoding and a strong second hop ($C_{\text{BIAWGN}}(\gamma_{\mathcal{RD}}) > 0.7$), the reliability of individual symbols of the relay's decoder will eventually be preserved after decoding at the destination. This will lead to a higher MI as compared to the loss by the hard-decision during a HIHO re-encoding. This loss of information cannot be recovered at the destination.

If the relay's decoder perfectly decodes the source's information, then both re-encoding methods achieve identical performance as the relay's transmit signal is equivalent in both cases.

Fig. 6.8 will also give a hint, at which channel conditions and code combinations to use soft- or hard re-encoding methods in a relay network. For sure, this means that any code combination $\mathcal{C}_{\mathcal{S}}$ and $\mathcal{C}_{\mathcal{R}}$ has to be simulated.

6.1.4 IPC for soft and hard re-encoding

To reduce this simulation effort in predicting the mutual information performance of soft and hard decision re-encoding, a solution was given in [VK11c]. This solution is based on the evaluation of the information processing characteristics. Utilizing the IPC in combination with the information combining technique, lower and upper bounds for any possible channel conditions and code combinations can be evaluated (or predicted) with almost no computational effort.

The problem with the IPC is its dependency towards a specific channel input distribution. In case of a HIHO re-encoder at the relay, the channel input is binary distributed and the standard measurement tool as depicted in **Fig. 2.12** can be used. For SISO re-encoding, the channel input distribution depends on the distribution of the information symbols, the re-encoding process and the subsequent scaling/altering of the code symbols. Thus, all these components have to be included during the measurement of the IPC for a specific processing characteristic.

Therefore, the simulation setup as depicted in **Fig. 6.10** is now utilized. Here, the first hop was completely characterized by its IPC $\text{IPC}^{\text{info}}(C_{\text{BIAWGN}}(\gamma_{\mathcal{SR}}))$, which had been evaluated and stored beforehand for different codes. Knowing the capacity $C_{\text{BIAWGN}}(\gamma_{\mathcal{SR}})$ of the first hop, the information LLRs $\mathbf{l}_{\mathcal{SR}}^{\text{info}}$ at the relay's decoder output can be modeled as bimodal Gaussian distributed ran-

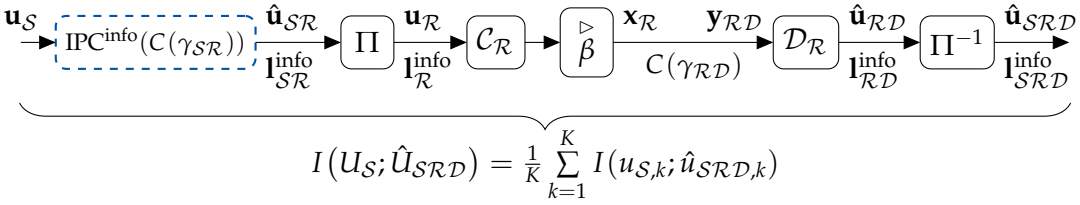


Fig. 6.10.: Schematic overview of the $\mathcal{S} \to \mathcal{R} \to \mathcal{D}$ link

dom variables with mean and variance depending on the mutual information of $I\left(U_{\mathcal{S}}; L^{\mathrm{info}}_{\mathcal{SR}}\right) = \mathrm{IPC}^{\mathrm{info}}_{\mathrm{SD}}(C_{\mathrm{BIAWGN}}(\gamma_{\mathcal{SR}}))$ [3]. Generating $\mathbf{l}^{\mathrm{info}}_{\mathcal{SR}}$ according to the IPC on the first hop, the analysis is performed by only implementing the second hop.

This saves simulation time as the second hop only needs to be simulated once per $\mathcal{C}_{\mathcal{R}}$ and not for any possible code combination between $\mathcal{C}_{\mathcal{S}}$ and $\mathcal{C}_{\mathcal{R}}$. Afterwards, the outcome mutual information can be "concatenated" with the IPC from the first hop.

One major problem is that the outcome

$$I\left(L^{\mathrm{info}}_{\mathcal{R}}; L^{\mathrm{info}}_{\mathcal{RD}}\right) = \mathrm{IPC}^{\mathrm{info}}_{\mathrm{SD}}\left(C(\gamma_{\mathcal{RD}}), I\left(U_{\mathcal{S}}; L^{\mathrm{info}}_{\mathcal{SR}}\right)\right)$$

of the second hop is three-dimensional, as the distribution of the random variable $L^{\mathrm{info}}_{\mathcal{R}}$ depends on the preceding hop. Therefore, the complete IPC, i.e. the MI $I\left(U_{\mathcal{S}}; L^{\mathrm{info}}_{\mathcal{SRD}}\right)$ of the $\mathcal{S} \to \mathcal{R} \to \mathcal{D}$ link, depends on two independent variables and is also three-dimensional.

Under the assumption that $L^{\mathrm{info}}_{\mathcal{R}}$ is bimodal Gaussian distributed with variance $\sigma^2_{L^{\mathrm{info}}_{\mathcal{R}}} = C^{-1}_{\mathrm{BIAWGN}}\left(I\left(U_{\mathcal{S}}; L^{\mathrm{info}}_{\mathcal{SR}}\right)\right)$ and mean[4] $\mu_{L^{\mathrm{info}}_{\mathcal{R}}} = \sigma^2_{L^{\mathrm{info}}_{\mathcal{R}}}/2$, a complete IPC is depicted in **Fig. 6.11**. Here, the relay node utilized a non-recursive non-systematic convolutional code with $\mathbf{g} = [25; 33; 37]_8$ and $R^{\mathrm{CC}}_c = 1/3$, before transmitting soft-bits towards the destination. Again, the MI between the relays encoder input and destinations decoder output is only depending of $\sigma^2_{L^{\mathrm{info}}_{\mathcal{R}}}$ and $\mu_{L^{\mathrm{info}}_{\mathcal{R}}}$ and, therefore, independent of the source code.

Now, in order to predict the MI $I\left(U_{\mathcal{S}}; L^{\mathrm{info}}_{\mathcal{SRD}}\right)$ of the complete channel, the outcome $I\left(U_{\mathcal{S}}; L^{\mathrm{info}}_{\mathcal{SR}}\right) = \mathrm{IPC}^{\mathrm{info}}_{\mathrm{SD}}(C_{\mathrm{BIAWGN}}(\gamma_{\mathcal{SR}}))$ of the first hop and the capacity

[3] Remember, that SD in $\mathrm{IPC}^{\mathrm{info}}_{\mathrm{SD}}(C_{\mathrm{BIAWGN}}(\gamma_{\mathcal{SR}}))$ stands for soft-decision. See Section 2.3.3

[4] Although, it was shown in subsection 6.1.2 that the assumption of $\mu_{L^{\mathrm{info}}_{\mathcal{R}}} = \sigma^2_{L^{\mathrm{info}}_{\mathcal{R}}}/2$ is not tight for the random variable $L^{\mathrm{info}}_{\mathcal{R}}$, still it is a common approximation in literature.

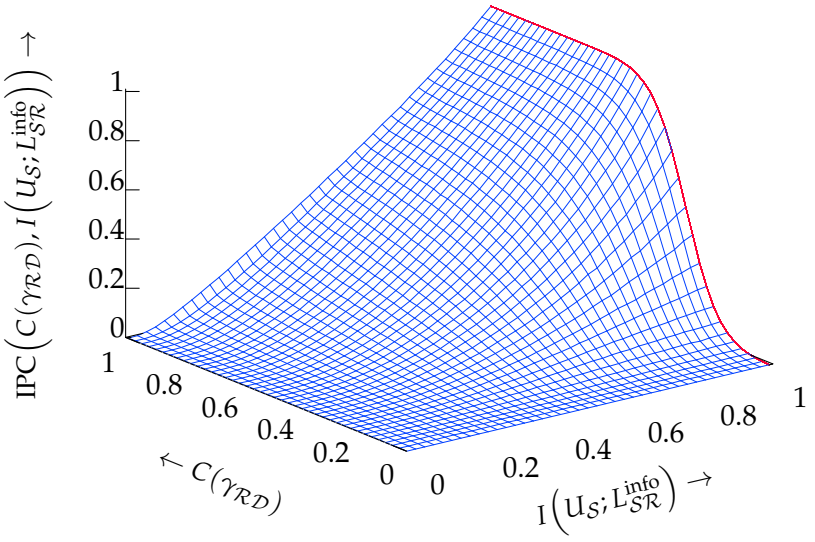


Fig. 6.11.: IPC for SISO re-encoding and transmitting soft-bits (red curve also holds for HIHO re-encoding)

$C_{\text{BIAWGN}}(\gamma_{\mathcal{RD}})$ of the second hop deliver the function elements for the three-dimensional IPC, calculating

$$I\left(U_{\mathcal{S}}; L^{\text{info}}_{\mathcal{SRD}}\right) = \text{IPC}^{\text{info}}_{\text{SD}}\left(C_{\text{BIAWGN}}(\gamma_{\mathcal{RD}}), I\left(U_{\mathcal{S}}; L^{\text{info}}_{\mathcal{SR}}\right)\right) \ . \tag{6.23}$$

If more than two hops are present, the mutual information of the complete channel is calculated recursively. Thus, no more than three dimensions are necessary.

From **Fig. 6.11**, one can conclude that in case of a perfect second hop ($C(\gamma_{\mathcal{RD}}) = 1$), the MI only depends on the first hop, which is linearly increasing as $I\left(U_{\mathcal{S}}; L^{\text{info}}_{\mathcal{SRD}}\right) = I\left(U_{\mathcal{S}}; L^{\text{info}}_{\mathcal{SR}}\right)$ holds. Furthermore, in case of a perfect first hop ($I\left(U_{\mathcal{S}}; L^{\text{info}}_{\mathcal{SR}}\right) = 1$), the mutual information of HIHO and SISO re-encoding will be equivalent and $I\left(U_{\mathcal{R}}; L^{\text{info}}_{\mathcal{RD}}\right) = I\left(L^{\text{info}}_{\mathcal{R}}; L^{\text{info}}_{\mathcal{RD}}\right) = \text{IPC}^{\text{info}}_{\text{SD}}(C(\gamma_{\mathcal{RD}}, 1))$ holds (see red curve in **Fig. 6.11**).

In case of HIHO re-encoding at the relay, the prediction of $I\left(U_{\mathcal{S}}; L^{\text{info}}_{\mathcal{SRD}}\right)$ can be easily performed via the serial information combining technique

$$I\left(U_{\mathcal{S}}; L^{\text{info}}_{\mathcal{SRD}}\right) = \mathcal{I}^{\text{ser,up}}\left(\text{IPC}^{\text{info}}_{\text{HD}}(C_{\text{BIAWGN}}(\gamma_{\mathcal{SR}})), \text{IPC}^{\text{info}}_{\text{SD}}(C_{\text{BIAWGN}}(\gamma_{\mathcal{RD}}))\right) \ , \tag{6.24}$$

given the IPCs for the first and second hop. So, no three-dimensional IPC for the complete channel is necessary.

6.2 Decoding at the destination

Since the destination receives the same information through different channels, different decoding methods can be applied. Among these methods are maximum ratio combining and turbo decoding of the received sequences. In both cases, the methods must be altered or extended in order to cope with the propagated errors from the relay node.

6.2.1 Limiter function

One major problem within the relay network is the transfer of false information to the successive nodes. This results from erroneous decoding at the preceding nodes, which cannot be recovered at the following nodes.

To cope with these errors, Sneessens et al. [SLV08] and, independently, Thobaben [Tho08] introduced proper decoding algorithms. In [SLV08], the authors modified the maximum-a-posteriori algorithm, whereas Thobaben introduced the so called *limiter function* (LF). This function is equivalent to the before mentioned, but has the advantage not to modify existing BCJR decoders. Thus, the adjustment of the channel or extrinsic LLRs is performed externally before or after the decoder, as for example depicted in **Fig. 6.14**, **Fig. 6.15** and **Fig. 6.17**.

The input-output relation of the log-likelihood ratios is given by the limiter function according to

$$
\begin{aligned}
L_O &= \mathcal{L}(L_I, \epsilon) \\
&= \log\left(\frac{\Pr(L_I|\hat{U}_{\mathcal{SR}}=0)\Pr(\hat{U}_{\mathcal{SR}}=0|U_{\mathcal{S}}=0)+\Pr(L_I|\hat{U}_{\mathcal{SR}}=1)\Pr(\hat{U}_{\mathcal{SR}}=1|U_{\mathcal{S}}=0)}{\Pr(L_I|\hat{U}_{\mathcal{SR}}=1)\Pr(\hat{U}_{\mathcal{SR}}=1|U_{\mathcal{S}}=1)+\Pr(L_I|\hat{U}_{\mathcal{SR}}=0)\Pr(\hat{U}_{\mathcal{SR}}=0|U_{\mathcal{S}}=1)}\right) \\
&= \log\left(\frac{(1-\epsilon)\cdot e^{+L_I/2}+\epsilon\cdot e^{-L_I/2}}{(1-\epsilon)\cdot e^{-L_I/2}+\epsilon\cdot e^{+L_I/2}}\right) \\
&= \operatorname{sign}(L_I)\cdot\min\left(|L_I|,|L_\epsilon|\right)+\log\left(1+e^{-|L_I+L_\epsilon|}\right)-\log\left(1+e^{-|L_I-L_\epsilon|}\right) \\
&= \operatorname{sign}(L_I)\cdot\min\left(|L_I|,|L_\epsilon|\right)+\log\left(\frac{1+e^{-|L_I+L_\epsilon|}}{1+e^{-|L_I-L_\epsilon|}}\right) \quad , \qquad (6.25)
\end{aligned}
$$

with ϵ being the mean error-probability of the relay's soft-output decoded sequence and with the derived LLR

$$L_\epsilon = \log\left(\frac{1-\epsilon}{\epsilon}\right) \quad . \tag{6.26}$$

Equation (6.25) is equivalent to the min-sum operation (Appendix A.4.10) with an additional correction term [CV05] in order to tightly approximate the boxplus operation, i.e.

$$L_O = L_I \boxplus l_\epsilon$$

holds. The main difference of equation (6.25) to the min-sum operation in (A.33) and besides the correction term, is the omitted second sign-term $\text{sign}(L_\epsilon)$. Here, it is unnecessary to include, since ϵ is always positive and $0 \leq \epsilon \leq 0.5$ holds. Thus, L_ϵ in (6.26) is also always positive, $0 \leq L_\epsilon \leq +\infty$ holds and $\text{sign}(L_\epsilon) = +1$. Therefore, the limiter function limits the absolute value of the output L_O to L_ϵ (in case of $L_\epsilon < L_I$), which is the LLR given by the error probability[5] $\epsilon = \Pr\left(\hat{U}_{\mathcal{SR}} \neq U_{\mathcal{S}}\right)$, i.e. the error probability at the relay's decoder output. This ensures that the mutual information at the destination, after decoding the relay's sequence, is limited by the MI regarding the source's information bits being only partly recovered at the relay node. Thus, $I\left(U_{\mathcal{S}}; \hat{U}_{\mathcal{SD}}\right) = \min\left\{I\left(U_{\mathcal{S}}; \hat{U}_{\mathcal{SR}}\right), I\left(U_{\mathcal{R}}; \hat{U}_{\mathcal{RD}}\right)\right\}$ with $I\left(U_{\mathcal{S}}; \hat{U}_{\mathcal{SR}}\right) = 1 - H_b(\epsilon)$ holds.

Recapitulating the major difference between a hard-input hard-output and a soft-input soft-output encoder, a SISO encoder will consider the soft-information over the information bits at its input during the encoding process, whereas the HIHO encoder destroys this reliability information. Thus, applying a HIHO encoder, the coded bits are all equal likely and the limiter function must be applied at the destination. This should not be the case for a three-node relay system utilizing a SISO encoder at the relay node. However, since the relay node normalizes the transmit power, the amplitude of the encoder output LLR values are increased or decreased. This will eventually alter their reliability and the destination should apply the limiter function in case of SISO encoding as well.

Utilizing a HIHO encoder at the relay node, the limiter function is not optimal but sufficiently good (as shown by the simulation results later). The optimal solution is given by applying an individual error-probability ϵ_k with each LLR value $l_{I,k}$ in (6.25) after the destination's decoder. Here, ϵ_k must be retrieved at the relay node for every information bit and, afterwards, forwarded towards the destination. Truly,

[5] Please note, that ϵ is the mean error probability over the decoded sequence. Furthermore, consideration of correlation among neighboring information bits would involve a more sophisticated approach which will not necessarily improve the outcome and, thus, will be neglected here.

to forward the precise information of ϵ_k, a perfect link is required, but not feasible. Thus, only the mean error-probability ϵ is utilized at the destination.

Furthermore, in the case of SISO re-encoding, the procedure in (6.25) has one major disadvantage. The limiter function at the destination can destroy soft information of the information bits, which has been retrieved at the relay node and forwarded to the destination. This is especially the case for certain information bits with an error probability lower than the average error probability ϵ, i.e. the corresponding absolute LLR values represent an error-probability lower than the mean one. Here, the limiter function would truncate the absolute values of the LLRs in correspondence to the mean error probability of the relay node. This maybe decrease the decoding performance at the destination, since the LF destroys parts of the soft information. The distribution of the (truncated) limiter functions LLRs L_O compared with the non-modified LLRs L_I is depicted in **Fig. 6.12**.

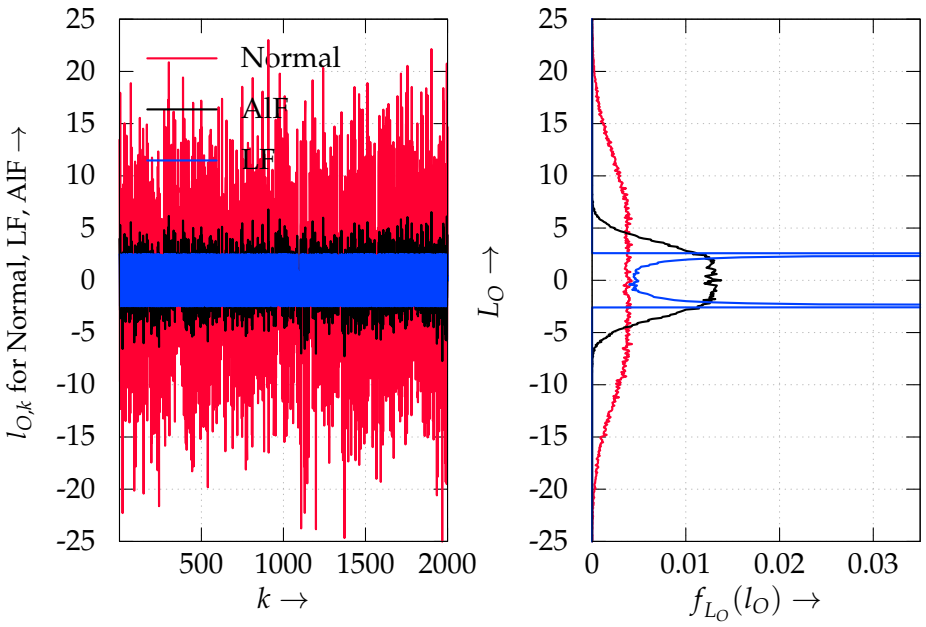


Fig. 6.12.: Values and distribution of the input LLR rv L_I and the output LLRs of the limiter function and the alpha function. The setup is equivalent as described in Subsection 6.1.3 with the relay node transmitting soft-values. The first and second hop capacities are given by ($C_{\text{BIAWGN}}(\gamma_{\mathcal{SR}}) = 0.5$ and $C_{\text{BIAWGN}}(\gamma_{\mathcal{RD}}) = 0.5$).

To overcome this problem, the author in [Huy11] proposed a so called *heuristically modified iterative decoder*. The aim is to replace the limiter function and to linearly modify the LLR values from the source's component decoder $\mathcal{D}_{\mathcal{R}}$ via a coefficient

α in the feedback path of the turbo-decoder. Since α is unknown, it is chosen heuristically, which is a major drawback of the method.

To solve this drawback, the idea of linear scaling the LLR values is extended in this thesis. To find a method, which delivers a decent and plausible coefficient α, one can utilize the fact that LLRs are sufficient statistics.

Thus, the idea is to find α such that the linear scaled LLRs $\mathbf{l}$ have the same mean error probability (calculated via $\mathrm{E}\left\{\frac{1}{1+e^{|\mathbf{l}|}}\right\}$) as given after applying the limiter function, i.e.

$$\mathrm{E}\left\{\frac{1}{1+e^{|\mathcal{L}(\mathbf{l}_I,\epsilon)|}}\right\} = \mathrm{E}\left\{\frac{1}{1+e^{|\mathbf{l}_O|}}\right\} \tag{6.27}$$

should hold, with the newly defined *alpha function* (AlF) by

$$\begin{aligned} l_O &= \mathcal{A}(l_I,\epsilon) \\ &= \alpha \cdot l_I \ . \end{aligned} \tag{6.28}$$

To achieve the equality in (6.27), the scaling factor α must be given by

$$\alpha_k = \frac{|\mathcal{L}(l_{I,k},\epsilon)|}{|l_{I,k}|} \tag{6.29}$$

per individual soft-bit at time index k and $l_{O,k} = \alpha_k \cdot l_{I,k}$ holds. Surly, this would lead to the same LLR values $l_{O,k}$ of (6.25) compared to (6.28) and no further improvement is achieved.

If the equality constraint in (6.27) is weakened by an approximation $\mathrm{E}\left\{\frac{1}{1+e^{|\mathcal{L}(\mathbf{l}_I,\epsilon)|}}\right\} \approx \mathrm{E}\left\{\frac{1}{1+e^{|\mathcal{A}(\mathbf{l}_I,\epsilon)|}}\right\}$ one maybe calculate α via

$$\alpha = \frac{\mathrm{E}\left\{|\mathcal{L}(\mathbf{l}_I,\epsilon)|\right\}}{\mathrm{E}\left\{|\mathbf{l}_I|\right\}} \ . \tag{6.30}$$

The difference between $\mathrm{E}\left\{\frac{1}{1+e^{|\mathcal{L}(\mathbf{l}_I,\epsilon)|}}\right\}$ and $\mathrm{E}\left\{\frac{1}{1+e^{|\mathcal{A}(\mathbf{l}_I,\epsilon)|}}\right\}$ is depicted in **Fig. 6.13**. It can be seen that the difference of the error, which is made by (6.30), is small as compared to the total possible error probability of 0.5. This, at least for the simulated setups, is justifying the approach taken in (6.30). Furthermore as depicted in **Fig. 6.12** and in the case of a descent first and second hop ($C_{\mathrm{BIAWGN}}(\gamma_{\mathcal{SR}}) = 0.5$ and $C_{\mathrm{BIAWGN}}(\gamma_{\mathcal{RD}}) = 0.5$), the distribution by the alpha function is now a scaled version of the distribution by the input LLR rv L_I, whereas the distribution of the limiter function is concentrated at some edges since the soft-bits are truncated. Applying the AlF can, in certain situations, improve the decoding result at the

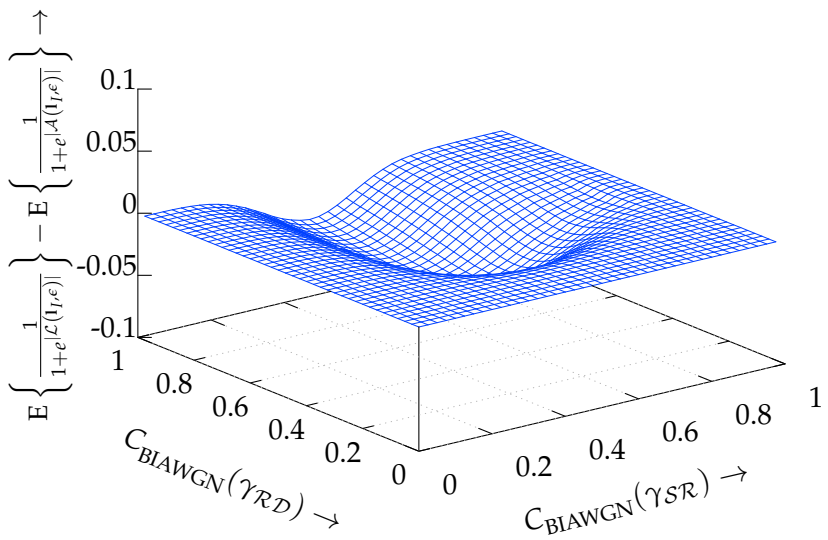


Fig. 6.13.: Difference between the mean error probability (calculated via $\mathrm{E}\left\{\frac{1}{1+e^{|\mathrm{I_O}|}}\right\}$) of the limiter function and the alpha function. The setup is equivalent as described in Subsection 6.1.3 with the relay node transmitting soft-values.

destination of a relay network. This is shown in the upcoming Subsection 6.2.4 on page 116.

In both cases, i.e. the limiter function and the alpha function, the destination must have knowledge of the decoding error probability ϵ from the relay node. This knowledge must be obtained at the relay node, either from the information processing characteristics (see Section 2.3.3) via

$$\epsilon = H_b^{-1}\left(1 - \mathrm{IPC}_{\mathrm{HD}}^{\mathrm{info}}(C_{\mathrm{BIAWGN}}(\gamma_{\mathcal{SR}}))\right) \ , \tag{6.31}$$

which where evaluated beforehand by Monte Carlo simulation, or from the LLR values through estimation [Lan05b] via

$$\epsilon = \mathrm{E}\left\{\frac{1}{1+e^{|\mathrm{I}_{\mathcal{SR}}^{\mathrm{info}}|}}\right\} \ . \tag{6.32}$$

Afterwards ϵ is forwarded to the destination with almost no overhead.

6.2.2 Maximum ratio combining

Since the source's information is transmitted through at least two independent links, the destination is capable of utilizing different decoding strategies. Due to

the orthogonal channel access, one strategy is to maximum ratio combine either the channel outputs or the decoder outputs at the destination. Both methods are

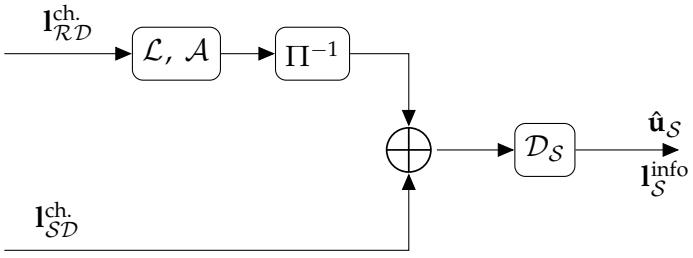


Fig. 6.14.: MRC before decoding at the destination

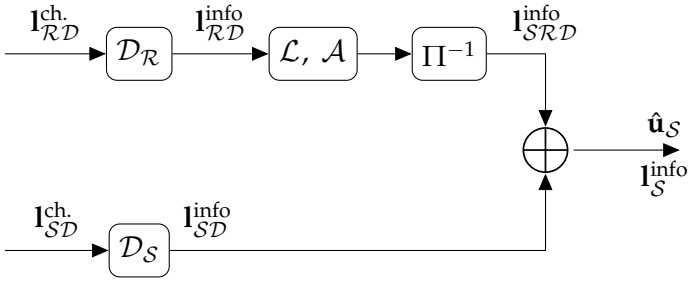


Fig. 6.15.: MRC after decoding at the destination

depicted in **Fig. 6.14** and **Fig. 6.15**.

The advantage of combining the channel outputs relies in the utilization of only one decoder and, therewith, saving resources. But, since the channel LLRs are combined, the transmitter has to utilize the same encoding strategy. This can lead to performance degradations which will be described in Subsection 6.2.4.

In case of combining the decoder outputs, different encoding schemes and interleavers can be used, which can lead to better BER performance. Here, the drawback is given by using more than one decoder and, thus, an increased resource requirement. Still, the computational complexity is less than for the turbo decoder in Subsection 6.2.3.

In **Fig. 6.14**, the received LLR values $\mathbf{l}^{\text{ch.}}_{\mathcal{SD}}$ from the direct link (from time slot T_1) are superimposed with the adjusted and de-interleaved LLR values $\mathbf{l}^{\text{ch.}}_{\mathcal{RD}}$ from the relay link (from time slot T_2). Afterwards, the resulting LLR sequence is decoded by the destination's decoder.

As mentioned before, the LLRs from the indirect link (relay path) have to be adjusted according to the relay's decoding error. In case of combining channel LLRs, the destination only possess knowledge about the error probability ϵ of the information sequence $\hat{\mathbf{u}}_{\mathcal{SR}}$, but the error probability ϵ^{code} of the re-encoded code sequence $\mathbf{x}_{\mathcal{R}}$ is

needed. Here, the relay's code sequence error probability ϵ^{code} is larger than the relay's information sequence error probability ϵ. One might imagine, that one error in the information sequence causes multiple errors in the code sequence, depending on the code rate $R_{\mathcal{C}_\mathcal{R}}^{\text{CC}}$ and the encoder structure.

For example, lets assume that the relay utilizes a $R_c^{\text{CC}} = 1/2$ HIHO non-recursive non-systematic convolutional encoder with $\mathbf{g} = [8;5]_8$. The code bits $c_j^{\text{code}(1)}$ and $c_j^{\text{code}(2)}$ are obtained via modulo-2 sum according to

$$\begin{aligned} c_j^{\text{code}(1)} &= \hat{u}_{\mathcal{SR},j} \oplus \hat{u}_{\mathcal{SR},j-2} \\ &= \left(u_{\mathcal{S},j} \oplus e_{\mathcal{S},j}\right) \oplus \left(u_{\mathcal{S},j-2} \oplus e_{\mathcal{S},j-2}\right) \\ &= \left(u_{\mathcal{S},j} \oplus u_{\mathcal{S},j-2}\right) \oplus \underbrace{\left(e_{\mathcal{S},j} \oplus e_{\mathcal{S},j-2}\right)}_{\mathbf{e}_\mathcal{S}^{\text{code}(1)}} \end{aligned} \tag{6.33}$$

with $\hat{u}_{\mathcal{SR},j} = u_{\mathcal{S},j} \oplus e_{\mathcal{S},j}$ and

$$c_j^{\text{code}(2)} = \left(u_{\mathcal{S},j} \oplus u_{\mathcal{S},j-2} \oplus u_{\mathcal{S},j-2}\right) \oplus \underbrace{\left(e_{\mathcal{S},j} \oplus e_{\mathcal{S},j-2} \oplus e_{\mathcal{S},j-2}\right)}_{\mathbf{e}_\mathcal{S}^{\text{code}(2)}} , \tag{6.34}$$

respectively. In (6.33) and (6.34), $e_{\mathcal{S},j} \in \{0,1\}$ denotes the bit error, which was made with a probability of ϵ. To evaluate the mean error probabilities $\epsilon^{\text{code}(1)}$ and $\epsilon^{\text{code}(2)}$ of the code bit sequences $\mathbf{c}^{\text{code}(1)}$ and $\mathbf{c}^{\text{code}(2)}$, one can perform some combinatorics with the help of a table:

$e_{\mathcal{S},j}$	$e_{\mathcal{S},j-1}$	$e_{\mathcal{S},j-2}$	$\mathbf{e}_\mathcal{S}^{\text{code}(1)}$	$\epsilon^{\text{code}(1)}$	$\mathbf{e}_\mathcal{S}^{\text{code}(2)}$	$\epsilon^{\text{code}(2)}$
0	0	0	0	$(1-\epsilon)^2$	0	$(1-\epsilon)^3$
0	0	1	1	$(1-\epsilon)\epsilon$	1	$(1-\epsilon)^2\epsilon$
0	1	0	-	-	1	$(1-\epsilon)^2\epsilon$
0	1	1	-	-	0	$(1-\epsilon)\epsilon^2$
1	0	0	1	$(1-\epsilon)\epsilon$	1	$(1-\epsilon)^2\epsilon$
1	0	1	0	ϵ^2	0	$(1-\epsilon)\epsilon^2$
1	1	0	-	-	0	$(1-\epsilon)\epsilon^2$
1	1	1	-	-	1	ϵ^3

According to the table, the error probabilities for $\mathbf{c}^{\text{code}(1)}$ and $\mathbf{c}^{\text{code}(2)}$ are given by

$$\begin{aligned} \epsilon^{\text{code}(1)} &= 2 \cdot (1-\epsilon) \cdot \epsilon \\ \epsilon^{\text{code}(2)} &= \epsilon^3 + 3 \cdot (1-\epsilon)^2 \cdot \epsilon \ . \end{aligned} \tag{6.35}$$

Theses error probabilities are utilized in the limiter function and the alpha function via $\mathcal{L}\left(\mathbf{I}_{\mathcal{RD}}^{\text{ch.}}, \epsilon^{\text{code}}\right)$ and $\mathcal{A}\left(\mathbf{I}_{\mathcal{RD}}^{\text{ch.}}, \epsilon^{\text{code}}\right)$, as shown in **Fig. 6.14**. It is also noteworthy

that the appropriate error probabilities must be applied. Here, for example $\epsilon^{\text{code}(1)}$ is used for all channel LLRs connected to the first output of the relay's convolutional encoder, whereas $\epsilon^{\text{code}(2)}$ must be applied to the other LLRs from the second output, respectively.

Depicted in **Fig. 6.16** are the simulated and predicted error probabilities $\epsilon^{\text{code}(1)}$ and $\epsilon^{\text{code}(2)}$, which results from an increasing information error probability ϵ. From

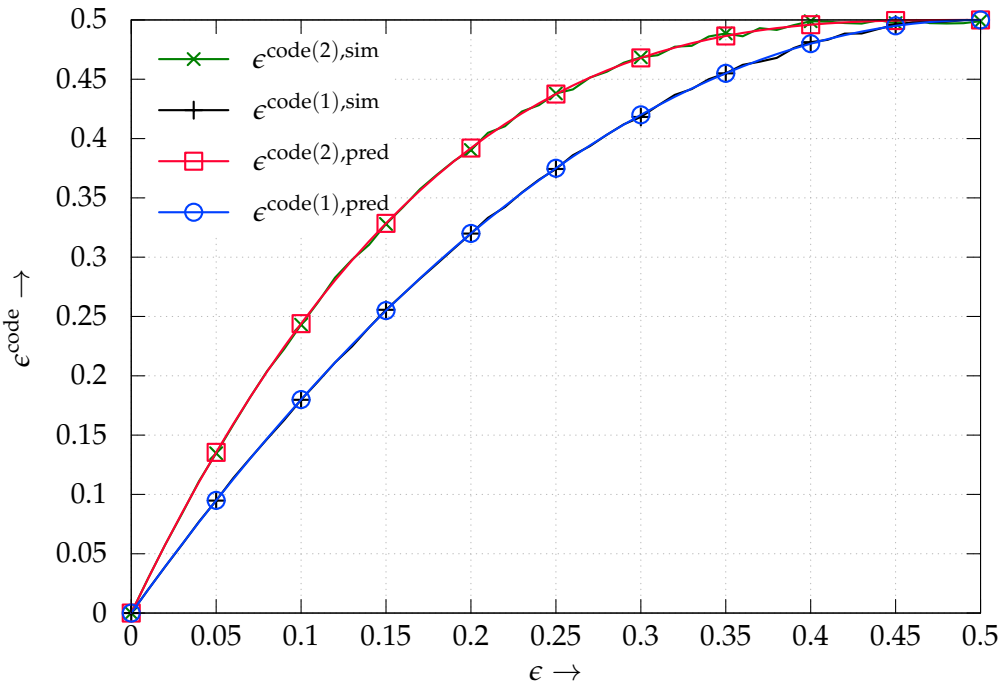


Fig. 6.16.: Measured and predicted code error probabilities $\epsilon^{\text{code}(1)}$ and $\epsilon^{\text{code}(2)}$

the simulation results and from equation (6.35) it can be seen that the code error probability depends on the number of modulo-2 sums during the encoding process. Thus, the error propagates for a recursive convolutional code and $\epsilon^{\text{code}} = 0.5$ holds for the non-systematic part (if $\epsilon > 0$). Furthermore, it is obvious that the error probability cannot decrease during the encoding process and, therefore, $\epsilon^{\text{code}} \geq \epsilon$ is always true.

The second maximum ratio combining method which is depicted in **Fig. 6.15** involves two component decoders. One for the direct link and one for the relay link. The output LLR values of $\mathcal{D}_{\mathcal{R}}$ are adjusted by either the limiter function $\mathcal{L}\left(\mathrm{l}^{\text{ch.}}_{\mathcal{RD}}, \epsilon\right)$ or the alpha function $\mathcal{A}\left(\mathrm{l}^{\text{ch.}}_{\mathcal{RD}}, \epsilon\right)$ and afterwards combined with the output LLRs of $\mathcal{D}_{\mathcal{S}}$. Both functions depend on the error probability ϵ of the source's information sequence at the relay's decoder output.

6.2.3 Turbo decoding

Under the assumption, that the relay can decode the source's message perfectly and, thus, forwards an interleaved and re-encoded version towards the destination, a distributed parallel concatenated coding scheme is obtained. Here, the destination receives two versions of the same information through independent channels. Fortunately, this can be decoded by utilizing a turbo decoder at the destination.

In case of decoding errors at the relay, the decoding performance dramatically decreases as the component decoders exchange extrinsic information about different information bits. The performance decrease can be limited utilizing the LF or the AlF, which have been introduced before.

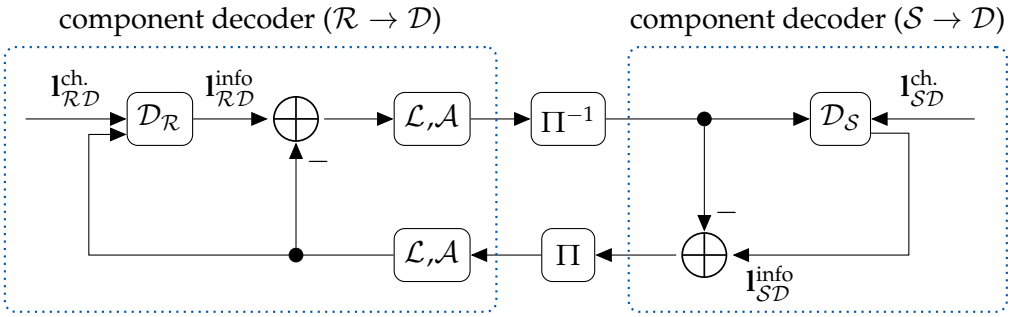


Fig. 6.17.: The turbo-decoder at the destination, inclusive the limiter function or alpha function at the relay's decoder component

This is depicted in **Fig. 6.17**. Here, the LLRs at the decoder output from the relay link will be adapted via $\mathcal{L}$ or $\mathcal{A}$ according to the error probability of the first hop. This is essential to regulate the iterative decoding process such that the decoding result will be closer to the source's information as to the relay's information. One can think of the extreme situation, where the $\mathcal{S} \rightarrow \mathcal{D}$ link is decent, the $\mathcal{S} \rightarrow \mathcal{R}$ link is extremely poor and the $\mathcal{R} \rightarrow \mathcal{D}$ link is perfect. Thus, the relay node forwards a re-encoded information sequence with an error probability of nearly $\epsilon \simeq 0.5$. As the relay link is perfect, the corresponding component decoder without the limiter or alpha function at the destination will output (infinite) large LLR values. These values will dominate the iterative decoding process and the destination will decide wrongly on the relay information. Utilizing the LF or AlF, will decrease the absolute (infinite) large LLRs such that they meet the error probability. Therefore, the turbo-decoder at the destination will now trust more on the direct link component decoder.

6.2.4 Comparing the different decoding methods

In this subsection, the bit error rate performance for the three decoding methods at the relay, including the limiter function and the alpha function, is evaluated through

Monte-Carlo simulation. To do so, the relay-node is moved on a straight line from the source towards the destination, i.e. from $d_{\mathcal{SR}} = 0.1$ towards $d_{\mathcal{SR}} = 0.9$. This is identical to the simulation setup in Section 5.1, except that now practical codes and maximum-a-posteriori decoders are utilized.

Equivalently to Subsection 6.1.3, the encoder $\mathcal{C}_{\mathcal{S}}$ at the source node employs a non-recursive non-systematic convolutional encoder with $\mathbf{g} = [5;7]_8$. At the relay node, the encoder $\mathcal{C}_{\mathcal{R}}$ is also a non-recursive non-systematic convolutional encoder, but with generator polynomial $\mathbf{g} = [5;7;7]_8$.

Furthermore, the SNR of the direct link is set to $\gamma_{\mathcal{SD}} = 1/\sigma_N^2 = 2\text{dB}$ ($d_{\mathcal{SD}} = 1$). Therefore, the receiving noise at the relay and destination node follows a distribution of $\mathcal{N}(0, \sigma_N^2/a_{\mathcal{SR}}^2)$ and $\mathcal{N}(0, \sigma_N^2/a_{\mathcal{RD}}^2)$, respectively. Please keep in mind that the path-loss attenuation is depending on the distance (see Section 2.2).

Depicted in **Fig. 6.18** are the simulation results for the different decoding and combining methods, i.e. turbo decoding at the destination, MRC after the destination's decoder outputs and MRC of the channel LLR values before the destination's decoder input.

Please note that the following results and their interpretations are only valid for this specific setup. A general statement involves the simulation of all possible relay positions, including different path-loss exponents and all possible code combinations. Of course, this would lead to a large amount of data which needs to be evaluated, but is out of scope in this dissertation. Still, some interesting performance behavior can be interpreted with the given setup.

Results for turbo decoding

The result for turbo decoding is depicted in the upper left figure in **Fig. 6.18**.

If the relay can perfectly decode the source's information sequence ($d_{\mathcal{SR}} < 0.35$), there is almost no difference between HIHO and SISO re-encoding with resulting soft-bits. Please note, although the BER is zero in this case, the magnitudes of the relay's decoder output LLRs are not identical. But, since their absolute values are large and in case of codes with small memories, the corresponding soft-bits will be closely binary distributed. In case of transmitting LLR values, the output distribution is bimodal Gaussian distributed and, thus, lower values will be more affected by the destinations noise. This will eventually degrade the BER performance.

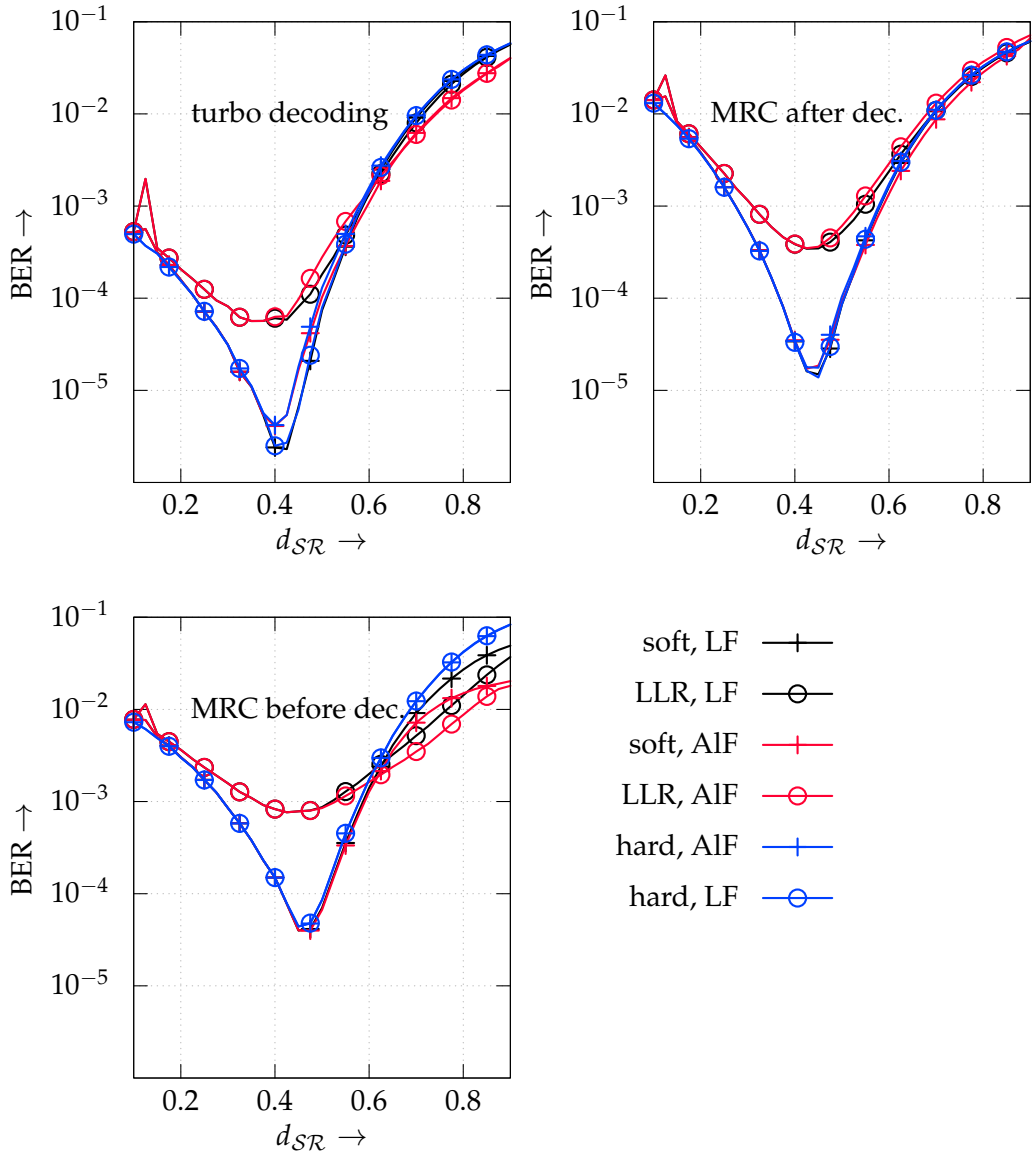


Fig. 6.18.: BER for the practical three-node relay network with SISO or HIHO re-encoding and different handling of propagated errors from the relay node. Upper left Figure: turbo code. Upper right Figure: MRC. Lower left Figure: MRC of the channel LLRs at the decoder input.

Once the relay starts making errors ($0.35 < d_{\mathcal{SR}} < 0.45$), there are differences in the BER between the hard, soft and LLR transmit signals together with the limiter function and alpha function. Here, transmitting soft-bits and utilizing the LF at the destination slightly outperforms the other methods. A possible explanation is that the relays transmit sequence contains enough amplified sign symbols (due to the tanh-function) and only a few attenuated symbols with less reliability (for example, see right figure in **Fig. 6.9**). The turbo decoder in combination with the limiter function seems to better cope with these weak symbols as compared with

the alpha function. Thus, scaling the LLRs inside the turbo decoder via the AlF is less effective as with the LF.

Soon, as the relay moves towards the destination ($d_{\mathcal{SR}} > 0.55$) and, thus, the relay's decoding result decreases, the AlF with soft-bits and LLRs performance slightly increases compared with the others. One explanation to this is that the alpha function does not limit the LLRs to a certain error probability as compared to the limiter function. Here, rare but strong relay transmit values, i.e. values of high certainty, will be retained during the iterative decoding process and possibly increase the overall performance.

Results for MRC after decoding

Utilizing the same codes as for the turbo scheme, depicted in the upper right graph of **Fig. 6.18** is the simulation result for maximum ratio combining the destination's decoder output LLRs (as shown in **Fig. 6.15**).

Here, except in case of transmitting LLR values, all other combinations perform equally well until the relay's decoding errors increases ($d_{\mathcal{SR}} > 0.45$). Once the relay's decoder rarely outputs large LLR values (values with high certainty), the alpha function at the destination slightly improves the overall decoding result.

Again, transmitting LLR values is least efficient until the relay is close to the destination. But, unlike as for the turbo scheme, transmitting LLRs does not improve the BER in the MRC scheme.

Comparing the MRC scheme with the turbo decoding scheme reveals the following assumptions. As long as the relay node can nearly perfectly retain the source's information sequence, turbo decoding at the destination has the lowest BER. Once the relay's decoding errors increases, both schemes nearly achieve the same BER.

Results for MRC prior decoding

In case of maximum ratio combining the channel LLRs prior decoding (as shown in **Fig. 6.14**), the relay has to utilize the same code $\mathcal{C}_{\mathcal{R}} = \mathcal{C}_{\mathcal{S}}$ and also the same interleaver. Thus, in this scenario, the relay's encoder $\mathcal{C}_{\mathcal{R}}$ is equivalent to $\mathcal{C}_{\mathcal{S}}$.

The problem with this method is that the code symbols are strongly correlated after re-encoding at the relay and one would expect a low efficiency [Wei13]. This is true

until the relay's position is close to the destination ($d_{\mathcal{SR}} > 0.6$). Here, this scheme outperforms the turbo and MRC scheme, although it utilizes a weaker second code $\mathcal{C}_{\mathcal{R}}$. This is depicted in the bottom left graph in **Fig. 6.18**. Additionally, this scheme also achieves a higher throughput since the code rate for $\mathcal{C}_{\mathcal{R}}$ is increased and, therefore, the second time slot T_2 is decreased. Thus, the same information can be transmitted in less time.

Using LLRs together with the alpha function seems to improve the decoding result, since more transmit power is spent towards symbols with high certainty. After reception, these strong symbols are then superimposed with the source's signal and afterwards processed by the destination's decoder, where they enhance the BER.

6.3 Information combining for practical relay networks

Until now, the mutual information of the $\mathcal{S} \rightarrow \mathcal{R} \rightarrow \mathcal{D}$ link with soft- and hard-decision re-encoding at the relay node, as well as the destinations behavior in case of maximum ratio combining and turbo-decoding has been studied separately. So all different parts of the practical relay network has been introduced and their mutual information was evaluated.

Now, in order to predict the MI behavior of a complete relay network, these parts are investigated together. This is done with the help of the information combining technique and the information processing characteristic for practical coding schemes, which, in case of convolutional codes, are described in Subsection 2.3.3 and 6.1.4.

For this, the network itself was decomposed into its different parts. Namely, virtual channels have been created, which include the encoder, the corresponding channel and the decoder. Each of these virtual channels are simple transmission systems for which the MI performance was evaluated. The input-output behavior of such a system is a function of the encoder-decoder pair, the underlying channel and the available a-priori information.

6.3.1 Practical three-node relay network

With the help of the IPC, one can easily predict three-node relay networks with maximum ratio combining after the destination's decoder (see Subsection 6.2.2).

This is done separately in case of either hard-input hard-output or soft-input soft-output re-encoding.

Relay network with hard-input hard-output re-encoder

If the relay utilizes a HIHO re-encoder, then the prediction is given by

$$\begin{aligned} I\left(U_{\mathcal{S}};\hat{U}_{\mathcal{S}}\right) = \mathcal{I}^{\mathrm{par}}\,(& \\ & \min\left\{\mathcal{I}^{\mathrm{ser}}\left(\mathrm{IPC}_{\mathrm{HD}}^{\mathrm{info}}(C(\gamma_{\mathcal{SR}})),\ \mathrm{IPC}_{\mathrm{SD}}^{\mathrm{info}}(C(\gamma_{\mathcal{RD}}))\right),\ I\left(U_{\mathcal{S}};\hat{U}_{\mathcal{SR}}\right)\right\}, \\ & \mathrm{IPC}_{\mathrm{SD}}^{\mathrm{info}}(C(\gamma_{\mathcal{SD}}))\ )\ , \end{aligned} \tag{6.36}$$

i.e. the serial concatenation of mutual information from the $\mathcal{S} \to \mathcal{R}$ link with the $\mathcal{R} \to \mathcal{D}$ link and the MI degradation (min-function) due to the limiter function ($I\left(U_{\mathcal{S}};\hat{U}_{\mathcal{SR}}\right) = \mathrm{IPC}_{\mathrm{HD}}^{\mathrm{info}}(C(\gamma_{\mathcal{SR}})) = 1 - H_b(\epsilon)$) and therewith the data processing theorem. Here, because of the limiter function, the MI of the $\mathcal{S} \to \mathcal{R} \to \mathcal{D}$ link is limited by the MI of the $\mathcal{S} \to \mathcal{R}$ link. Afterwards, the outcome is parallel combined with the MI of the direct link.

Since the upper bound for $\mathcal{I}^{\mathrm{ser}}$ is given by $\mathcal{I}^{\mathrm{ser,up}}$, the degradation of mutual information by the limiter function can be neglected. The proof is as follows:

$$\underbrace{I\left(U_{\mathcal{S}};\hat{U}_{\mathcal{SR}}\right)}_{1-H_b(\epsilon_{\mathcal{SR}})} \geq \underbrace{\mathcal{I}^{\mathrm{ser}}\left(\mathrm{IPC}_{\mathrm{HD}}^{\mathrm{info}}(C(\gamma_{\mathcal{SR}})),\ \mathrm{IPC}_{\mathrm{SD}}^{\mathrm{info}}(C(\gamma_{\mathcal{RD}}))\right)}_{1-H_b(\epsilon_{\mathcal{SRD}})}$$

$$\begin{aligned} \epsilon_{\mathcal{SR}} &\leq (1-\epsilon_{\mathcal{SR}})\epsilon_{\mathcal{RD}} + \epsilon_{\mathcal{SR}}(1-\epsilon_{\mathcal{RD}}) \\ \epsilon_{\mathcal{SR}} &\leq \epsilon_{\mathcal{SR}} + \epsilon_{\mathcal{RD}} - 2\epsilon_{\mathcal{SR}}\epsilon_{\mathcal{RD}}\ . \end{aligned}$$

Since $\epsilon_{\mathcal{SR}} \leq 0.5$ and $\epsilon_{\mathcal{RD}} \leq 0.5$ hold, the crossover probability for the $\mathcal{S} \to \mathcal{R}$ link cannot be greater than the crossover probability of the $\mathcal{S} \to \mathcal{R} \to \mathcal{D}$ link.

Therefore, equation (6.36) is rewritten as

$$\begin{aligned} I\left(U_{\mathcal{S}};\hat{U}_{\mathcal{S}}\right) = \mathcal{I}^{\mathrm{par}}\,(& \\ & \mathcal{I}^{\mathrm{ser}}\left(\mathrm{IPC}_{\mathrm{HD}}^{\mathrm{info}}(C(\gamma_{\mathcal{SR}})),\ \mathrm{IPC}_{\mathrm{SD}}^{\mathrm{info}}(C(\gamma_{\mathcal{RD}}))\right), \\ & \mathrm{IPC}_{\mathrm{SD}}^{\mathrm{info}}(C(\gamma_{\mathcal{SD}}))\ )\ . \end{aligned} \tag{6.37}$$

Relay network with soft-input soft-output re-encoder

In case of soft-input soft-output re-encoding, the prediction is slightly different. Here, as stated earlier in Subsection 6.1.4, the mutual information for the concatenation of the first with the second hop cannot be processed by the serial information combining function and two separate information processing characteristics. Instead a three-dimensional IPC must be generated (see for example **Fig. 6.11**), in which the mutual information for the serial concatenation can be taken from (see equation (6.23)).

Now, the MI of the $\mathcal{S} \rightarrow \mathcal{R} \rightarrow \mathcal{D}$ link is given by

$$I\left(U_{\mathcal{S}}; L^{\text{info}}_{\mathcal{SRD}}\right) = \min\left\{\text{IPC}^{\text{info}}_{\text{SD}}\left(C(\gamma_{\mathcal{RD}}), I\left(U_{\mathcal{S}}; L^{\text{info}}_{\mathcal{SR}}\right)\right), I\left(U_{\mathcal{S}}; \hat{U}_{\mathcal{SR}}\right)\right\} \quad . \tag{6.38}$$

Replacing the serial concatenation by inserting (6.38) into (6.36) leads to

$$I\left(U_{\mathcal{S}}; \hat{U}_{\mathcal{S}}\right) = \mathcal{I}^{\text{par}}\left(I\left(U_{\mathcal{S}}; L^{\text{info}}_{\mathcal{SRD}}\right),\ \text{IPC}^{\text{info}}_{\text{SD}}(C(\gamma_{\mathcal{SD}}))\right) \quad , \tag{6.39}$$

the parallel concatenation of the MI for the direct link with the MI taken from the three dimensional IPC (see for example **Fig. 6.11**) from the relay link.

Since the relay is capable of transmitting soft-bits, a higher MI on the second hop may be achieved. The problem, as stated earlier in Subsection 6.2.1, is that the best the destination can do to cope with propagated errors is to apply the limiter or alpha function. Thus, the destination limits the maximum possible mutual information of the indirect link in equation (6.38) towards the MI $I\left(U_{\mathcal{S}}; \hat{U}_{\mathcal{SR}}\right)$ of the first hop. This is also the answer to the results for MRC after decoding in **Fig. 6.18** (upper right). Here, HIHO and SISO re-encoding will perform equally well, as soon as the relay gets close to the destination and, thus, the $\mathcal{S} \rightarrow \mathcal{R}$ link limits the BER performance.

As an example, the simulation setup from Subsection 6.2.4 is utilized with $\gamma_{\mathcal{SD}} = -4\text{dB}$. The simulation result and the prediction of the MI via the information combining technique is shown in **Fig. 6.19**. Here, upper and lower bounds on the MI have been predicted in case of SISO re-encoding and soft-bits at the relay node. These bounds, i.e. $I^{\text{pred,up}}\left(U_{\mathcal{S}}; L^{\text{info}}_{\mathcal{S}}\right)$ and $I^{\text{pred,low}}\left(U_{\mathcal{S}}; L^{\text{info}}_{\mathcal{S}}\right)$, tightly enclose the measured MI $I^{\text{sim}}\left(U_{\mathcal{S}}; L^{\text{info}}_{\mathcal{S}}\right)$ from the Monte-Carlo simulation.

The advantage of predicting the MI of the relay network via the information processing characteristic and information combining technique is due to its low computational complexity. Once the information processing characteristics are given for all utilized codes, the prediction only depends on the channel capacity which is

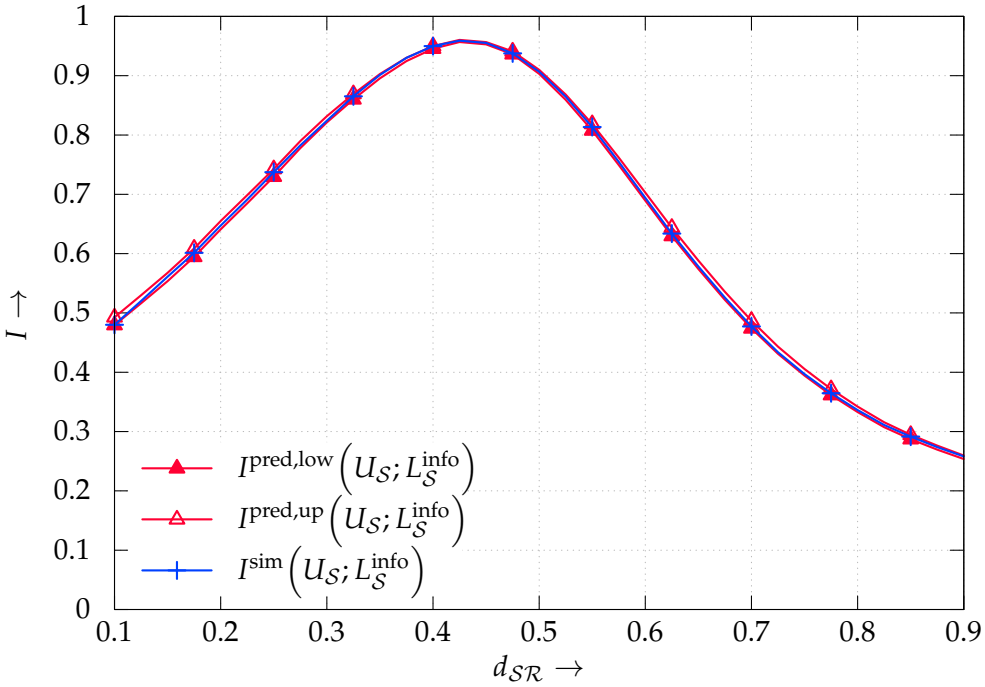


Fig. 6.19.: Predicted and simulated MI for a three-node relay network with SISO re-encoding.

given by the receive SNR. So, instead of doing exhaustive Monte-Carlo simulation for all possible code and SNR combinations, only simple combining/concatenation tasks have to be performed.

In order to predict the mutual information of relay networks with turbo decoding at the destination, the information processing characteristic has to be extended, including the a-priori knowledge of the second component decoder. This is done in the following subsection.

6.3.2 Extended information processing characteristic

In this section, the idea of the IPC from section 2.3.3 is extended such that it includes the a-priori information [VK11a]. This is needed in order to determine the mutual information of an iterative decoding process inside the destination, for example utilizing the turbo-decoder. Basically, the a-priori information represents the possible message from the other component decoder. Since this decoder may belong to a second relay, the exchanged a-priori LLR values $\mathbf{l}^{\text{a-pri.}}$ are adjusted according to the limiter function or the alpha function. The adjustment is performed according to the error probability ϵ which is present at the other relay node, i.e. the error probability of the first hop of this relay node.

Therefore, the extended information processing characteristic is a function of the channel capacity C, the mutual information $I^{\text{a-pri.}}$ of the a-priori channel (with noise variance $\sigma^2_{N^{\text{a-pri.}}}$) and its limitation according to the error probability ϵ

$$\begin{aligned} \text{IPC}^{\text{info}}_{\text{ext,SD}}\left(C, I^{\text{a-pri.}}, \epsilon\right) &= I\left(U; L | I^{\text{a-pri.}}, \epsilon\right) \\ &= \frac{1}{K} \sum_{k=1}^{K} I\left(u_k; l_k | I^{\text{a-pri.}}, \epsilon\right) \quad . \end{aligned} \tag{6.40}$$

Obviously, $\text{IPC}^{\text{info}}_{\text{SD}}(C) = \text{IPC}^{\text{info}}_{\text{ext,SD}}(C, 0, 0)$ holds. Furthermore, for the simulation it is noteworthy that the a-priori channel delivers channel output values with mean MI $I^{\text{a-pri.}}$ which are transformed into log-likelihood ratios. Afterwards, these LLRs are scaled according to the limiter or alpha function, retaining the decoder a-priori LLRs $\mathbf{l}^{\text{a-pri.}}$. One might think of the scaled LLRs of a second component decoder.

The simulation setup to determine the information processing characteristic of the different encoder-decoder pairs is depicted in **Fig. 6.20**. Additional to the

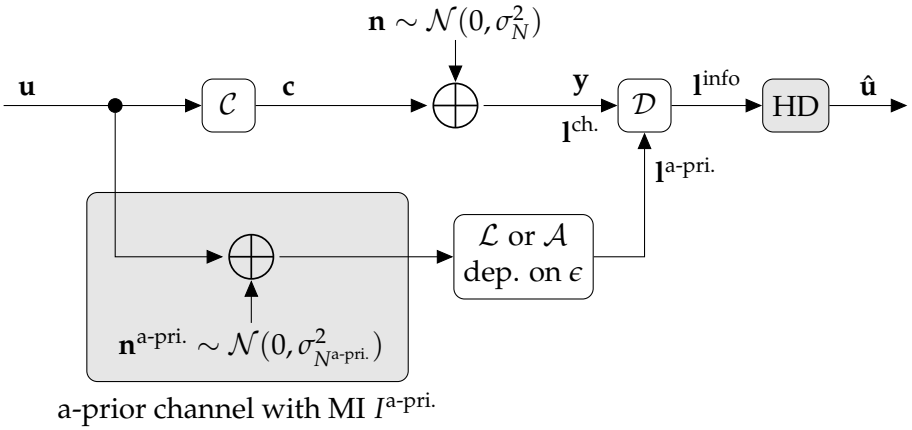


Fig. 6.20.: Simulation setup for the extended information processing characteristic

simulation setup from Section 2.3.3 is the existence of an a-priori channel. Here, the i.i.d. information bits $\mathbf{u} = [u_k]_{K\times 1}$ are transmitted over a BIAWGN channel with a noise variance $\sigma^2_{N^{\text{a-pri.}}}$ matching the a-priori mutual information $I^{\text{a-pri.}}$. The generated LLR values at the a-priori channel output are passed through the limiter function or the alpha function in order to be adjusted according to ϵ. Afterwards, the decoder $\mathcal{D}$ utilizes the LLR values from the direct channel and the a-priori channel to retain an estimate of $\mathbf{u}$.

As an example, the mutual information between $\mathbf{u}$ and its (soft) estimate $\mathbf{l}$ at the decoder output, i.e. the information processing characteristic, is depicted in **Fig. 6.21** for a simple convolutional code with rate $R_c^{\text{CC}} = 1/2$ and memory $m = 2$. In case of $\epsilon = 0$ (upper left figure), it can be observed that the extended IPC is dominated by the channel capacity C and the a-priori MI $I^{\text{a-pri.}}$. If $C = 0$, then the IPC is linear

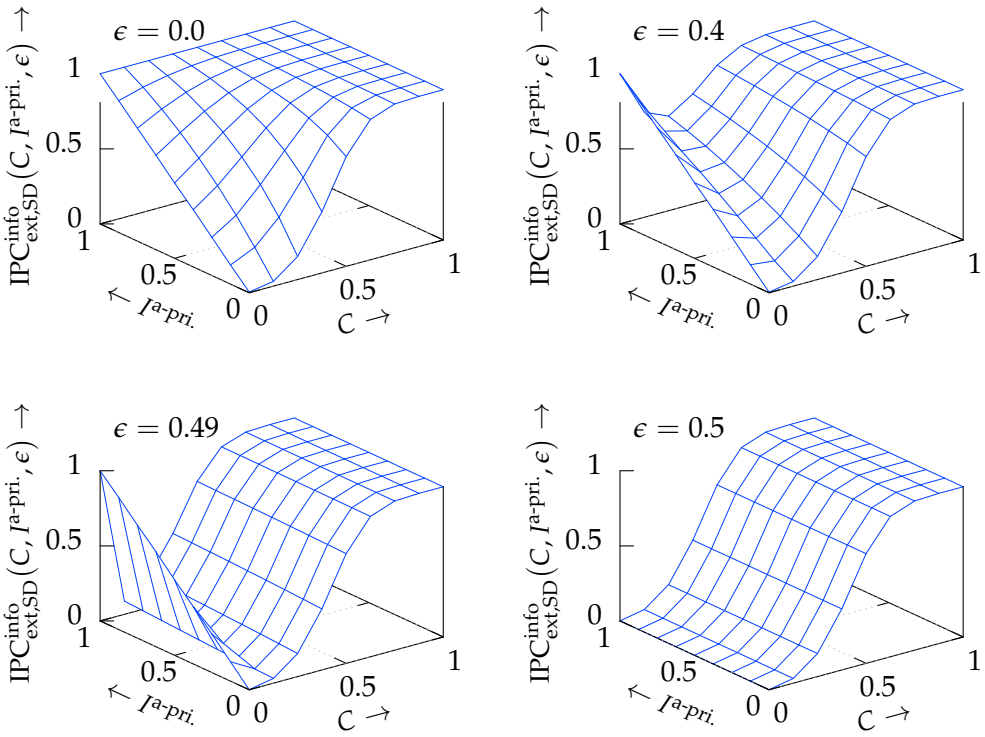


Fig. 6.21.: Extended IPC for a non-recursive convolutional code with rate $R_c = 1/2$ memory $m = 2$ and different a-priori channels with different error probabilities ϵ

increasing with $I^{\text{a-pri.}}$, since the only source of information is the a-priori channel. On the other extreme, if $I^{\text{a-pri.}} = 0$, then the IPC is determined by the channel capacity and the channel code without a-priori knowledge. This case corresponds to the traditional information processing characteristic and results in a highly non linear increasing function.

For $\epsilon = 0.5$ (lower right figure), no a-priori information is given at the decoder input and, therefore, the extended IPC is solely determined by the channel capacity C and the channel code.

More interestingly are the cases for $0.0 < \epsilon < 0.5$. Here, one can see that the IPC is increasing at lower channel capacity C in combination with increasing a-priori information $I^{\text{a-pri.}}$. This is due to the observation that the mean absolute values of the LLR values $\mathbf{l}^{\text{ch.}}$ for $\mathbf{y}$ at the channel output is less than the mean absolute values of $\mathbf{l}^{\text{a-pri.}}$ for the a-priori channel. In other words: Although the error probability for the a-priori channel is large (for example $\epsilon = 0.4$), the error probability of the direct channel is even larger (C is low). Therefore, the a-priori channel dominates the mutual information of the IPC in this region.

6.3.3 Practical four-node relay network

Until now, only the three-node relay network has been considered. But the prediction of mutual information via the information combining technique is not limited to any possible number of relays [VK11a]. Sure, more than one relay node will involve additional considerations, for example the order of the component decoders during the turbo decoding process at the destination, or the selection of the best combining method. Practically, an increase in relay nodes and the restriction towards orthogonal access also requires more physical resources and stronger consideration of resource mappings.

Another issue predicting the MI is the computational complexity of large relay networks. Any additional node increases the necessary computing time. But, in comparison with the ICT via Monte-Carlo simulation, the cost to predict the MI is considerably low.

In this subsection, the network is extended by an additional relay, establishing a four node-relay network. In order to ensure an orthogonal access, the transmission of information is now departed into three (not necessarily equal long) time slots T_1, T_2 and T_3. Here, the source encodes the i.i.d. binary information sequence $\mathbf{u}$ of

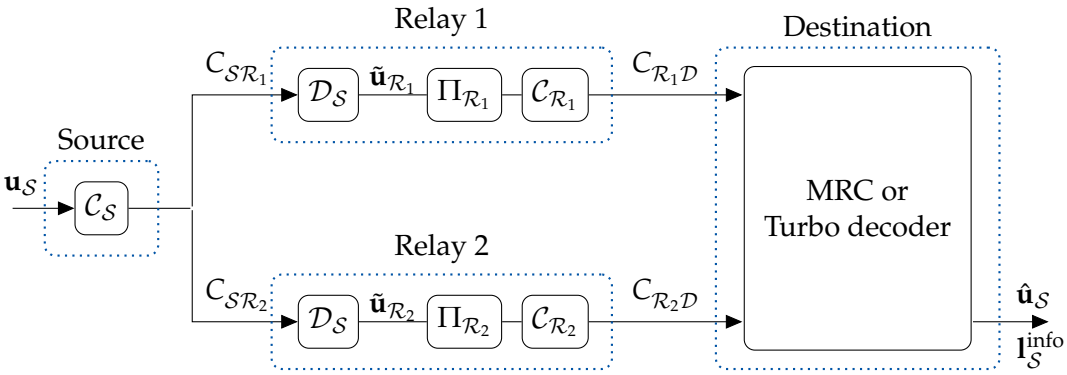


Fig. 6.22.: Practical four-node relay network

length K using the code $\mathcal{C}_\mathcal{S}$. Afterwards, the coded and BPSK modulated transmit sequence is broadcasted during the first time slot T_1 towards relay $\mathcal{R}_1$ and relay $\mathcal{R}_2$. At booth relays, the received message is decoded by $\mathcal{D}_\mathcal{S}$ and a hard decision is taken at the decoder output. This implies, that subsequently the relay nodes perform a HIHO re-encoding. Thus, without loss of generality, the upcoming description will be simplified.

The decided binary sequence $\tilde{\mathbf{u}}_{\mathcal{R}_m}$ is then interleaved with a relay specific interleaver $\Pi_{\mathcal{R}_m}$. Afterwards, the interleaved sequence is re-encoded with a relay specific

encoder $\mathcal{C}_{\mathcal{R}_m}$, which must not be equivalent to the encoder of the other relay node. Both relays forward their decided, interleaved and re-encoded sequence in time slot T_2 (for $\mathcal{R}_1$) and T_3 (for $\mathcal{R}_2$) towards the destination.

At the destination, either maximum ratio combining (see **Fig. 6.15**) or turbo decoding (see **Fig. 6.17**) is performed. In both figures, i.e. **Fig. 6.15** and **Fig. 6.17**, only the component decoder associated to the relay link utilizes the limiter function. For the scenario with two and also more relays, each component decoder has to utilize the limiter function. This can be best described in case of maximum ratio combining. Here, before the destination's decoder outputs can be combined, both log-likelihood ratio sequences need to be scaled according to the error probability of the specific relay [VK11a]. Otherwise, the destination will most likely "trust" the relay with the better channel capacity in between. This relay maybe not the relay with the lowest error probability over the source's information sequence.

MI prediction for turbo decoding

In order to determine the mutual information of the error-prone four-node relay network, the extended information processing characteristic from Subsection 6.3.2 is utilized to predict the behavior of the component decoders. The prediction is performed iteratively by taking the following steps:

1. Calculate error probability $\epsilon_{\mathcal{R}_m}$ of $\mathcal{R}_m$ for the first hops given the $\mathrm{IPC}_{\mathrm{HD}}^{\mathrm{info}}(C_{\mathrm{BIAWGN}}(\gamma_{\mathcal{SR}_m}))$ and initialize the extrinsic information $I_{\mathcal{D}_{\mathcal{R}_m}}^{\mathrm{extr}} = 0$

2. Iteratively calculate

   - $I_{\mathcal{D}_{\mathcal{R}_1}} = \mathcal{I}^{\mathrm{ser}}\left(\epsilon_{\mathcal{R}_1},\ \mathrm{IPC}_{\mathrm{ext,SD}}^{\mathrm{info}}\left(C_{\mathrm{BIAWGN}}(\gamma_{\mathcal{R}_1\mathcal{D}}), I_{\mathcal{D}_{\mathcal{R}_2}}^{\mathrm{extr}}, \epsilon_{\mathcal{R}_1}\right)\right)$
   - $I_{\mathcal{D}_{\mathcal{R}_1}}^{\mathrm{extr}} = \mathcal{I}^{\mathrm{ser}}\left(\epsilon_{\mathcal{R}_1},\ \mathrm{IPC}_{\mathrm{ext,SD}}^{\mathrm{extr}}\left(C_{\mathrm{BIAWGN}}(\gamma_{\mathcal{R}_1\mathcal{D}}), I_{\mathcal{D}_{\mathcal{R}_2}}^{\mathrm{extr}}, \epsilon_{\mathcal{R}_1}\right)\right)$
   - $I_{\mathcal{D}_{\mathcal{R}_2}} = \mathcal{I}^{\mathrm{ser}}\left(\epsilon_{\mathcal{R}_2},\ \mathrm{IPC}_{\mathrm{ext,SD}}^{\mathrm{info}}\left(C_{\mathrm{BIAWGN}}(\gamma_{\mathcal{R}_2\mathcal{D}}), I_{\mathcal{D}_{\mathcal{R}_1}}^{\mathrm{extr}}, \epsilon_{\mathcal{R}_2}\right)\right)$
   - $I_{\mathcal{D}_{\mathcal{R}_2}}^{\mathrm{extr}} = \mathcal{I}^{\mathrm{ser}}\left(\epsilon_{\mathcal{R}_2},\ \mathrm{IPC}_{\mathrm{ext,SD}}^{\mathrm{extr}}\left(C_{\mathrm{BIAWGN}}(\gamma_{\mathcal{R}_2\mathcal{D}}), I_{\mathcal{D}_{\mathcal{R}_1}}^{\mathrm{extr}}, \epsilon_{\mathcal{R}_2}\right)\right)$

3. Choose the decoder with the highest mutual information and determine $I^{\mathrm{pred}}\left(U_{\mathcal{S}}; L_{\mathcal{S}}^{\mathrm{info}}\right) = \max\left\{I_{\mathcal{D}_{\mathcal{R}_1}}, I_{\mathcal{D}_{\mathcal{R}_2}}\right\}$

As for the component codes during the turbo decoding process, the extrinsic information $I_{\mathcal{D}_{\mathcal{R}_1}}^{\text{extr}}$ from $\mathcal{D}_{\mathcal{R}_1}$ becomes the a-priori information for $\mathcal{D}_{\mathcal{R}_2}$ and vice versa. The extended IPC $\text{IPC}_{\text{ext,SD}}^{\text{extr}}$ denotes the IPC of the extrinsic information, i.e. $\mathbf{l}^{\text{extr}} = \mathbf{l} - \mathbf{l}^{\text{a-pri.}}$, the decoder output LLR values $\mathbf{l}$ removed by the a-priori LLR values $\mathbf{l}^{\text{a-pri.}}$.

Since decoding errors can occur at both relay nodes, the destination has to select the component decoder with the least unreliable information sequence. This can be done by looking at the average MI of both decoder outputs and choosing the component decoder with the highest mutual information. Without the introduced method, the destination can only randomly choose the decoder.

MI prediction for MRC

Predicting the mutual information of the error-prone network with maximum ratio combining at the destination is slightly easier than for turbo decoding. The steps which are needed to take are given by

1. Calculate error probability $\epsilon_{\mathcal{R}_m}$ of $\mathcal{R}_m$ for the first hops given the $\text{IPC}_{\text{HD}}^{\text{info}}(C_{\text{BIAWGN}}(\gamma_{\mathcal{S}\mathcal{R}_m}))$

2. Calculate the MI of the $\mathcal{S} \to \mathcal{R}_1 \to \mathcal{D}$ and $\mathcal{S} \to \mathcal{R}_2 \to \mathcal{D}$ link

   - $I_{\mathcal{D}_{\mathcal{R}_1}} = \mathcal{I}^{\text{ser}}\left(\epsilon_{\mathcal{R}_1}, \text{IPC}_{\text{SD}}^{\text{info}}(C_{\text{BIAWGN}}(\gamma_{\mathcal{R}_1\mathcal{D}}))\right)$

   - $I_{\mathcal{D}_{\mathcal{R}_2}} = \mathcal{I}^{\text{ser}}\left(\epsilon_{\mathcal{R}_2}, \text{IPC}_{\text{SD}}^{\text{info}}(C_{\text{BIAWGN}}(\gamma_{\mathcal{R}_2\mathcal{D}}))\right)$

3. Calculate upper and lower bound on mutual information $I^{\text{pred,up}}\left(U_{\mathcal{S}}; L_{\mathcal{S}}^{\text{info}}\right) = \mathcal{I}^{\text{par,up}}\left(I_{\mathcal{D}_{\mathcal{R}_1}}, I_{\mathcal{D}_{\mathcal{R}_2}}\right)$ and $I^{\text{pred,low}}\left(U_{\mathcal{S}}; L_{\mathcal{S}}^{\text{info}}\right) = \mathcal{I}^{\text{par,low}}\left(I_{\mathcal{D}_{\mathcal{R}_1}}, I_{\mathcal{D}_{\mathcal{R}_2}}\right)$ utilizing parallel information combining

The idea is to calculate the MI from the sources encoder input to the destinations component decoder output via the serial information combining function. Afterwards, one can determine the MI at the destinations output after MRC with the help of the parallel information combining function.

Please note, that in step 3 upper and lower bounds are calculated. A direct prediction of $I^{\text{pred,up}}\left(U_{\mathcal{S}}; L_{\mathcal{S}}^{\text{info}}\right)$ via $\mathcal{I}_{\text{SMC}}^{\text{par}}\left(I_{\mathcal{D}_{\mathcal{R}_1}}, I_{\mathcal{D}_{\mathcal{R}_2}}\right)$ is also possible, but involves a higher computational complexity as densities have to be considered.

Simulation result

In order to compare the semi-analytical analysis via the information combining technique of the practical error-prone four-node relay network with results from Monte-Carlo simulations, source and relay nodes employ different encoders. Here, for example, the source $\mathcal{S}$ utilizes a non-recursive convolutional code with code-rate $R_{c,\mathcal{S}}^{\mathrm{CC}} = 1/3$, memory $m = 4$ and generator polynomial $\mathbf{g} = [25;33;37]_8$, whereas relay $\mathcal{R}_1$ utilizes a non-recursive convolutional code with code-rate $R_{c,\mathcal{R}_1}^{\mathrm{CC}} = 1/3$, memory $m = 2$ and $\mathbf{g} = [5;7;7]_8$ and relay $\mathcal{R}_2$ a non-recursive convolutional code with code-rate $R_{c,\mathcal{R}_2}^{\mathrm{CC}} = 1/2$, memory $m = 2$ and $\mathbf{g} = [7;5]_8$. All decoders at the receiving nodes are maximum-a-posteriori decoders as described in Section 2.3.1. Furthermore, the relay nodes take a hard-decision at the decoder output before interleaving.

Due to the different codes, which are utilized at the nodes, maximum ratio combining is not possible prior decoding at the destination. Thus, only the decoder outputs can be combined in case of MRC (see **Fig. 6.15**).

Depicted in **Fig. 6.23**, **Fig. 6.24** and **Fig. 6.25** are the simulation results of the practical error-prone relay network. Here, the SNR $\gamma_{\mathcal{SR}_1}$ for the link $\mathcal{S} \rightarrow \mathcal{R}_1$ was altered such that the channel capacity increases from 0 to 1 ($0 \leq C_{\mathrm{BIAWGN}}(\gamma_{\mathcal{SR}_1}) \leq 1$). The capacities for the remaining links have been fixed to $C_{\mathrm{BIAWGN}}(\gamma_{\mathcal{SR}_2}) = 0.3$, $C_{\mathrm{BIAWGN}}(\gamma_{\mathcal{R}_1\mathcal{D}}) = 0.4$ and $C_{\mathrm{BIAWGN}}(\gamma_{\mathcal{R}_2\mathcal{D}}) = 0.6$.

Shown in **Fig. 6.23** is the result for the network with turbo decoding at the destination (see **Fig. 6.17**, but with LFs for both component decoders). Here, the number of iterations at the turbo decoder was fixed to 10.

The simulated mutual informations $I_{\mathcal{D}_{\mathcal{R}_1}}^{\mathrm{sim}}\left(U_{\mathcal{S}}; L_{\mathcal{S}}^{\mathrm{info}}\right)$ and $I_{\mathcal{D}_{\mathcal{R}_2}}^{\mathrm{sim}}\left(U_{\mathcal{S}}; L_{\mathcal{S}}^{\mathrm{info}}\right)$ for the two component decoders intersect at $C_{\mathcal{SR}_1} \approx 0.3$. As the capacities for the $\mathcal{S} \rightarrow \mathcal{R}_2 \rightarrow \mathcal{D}$ channels are fixed and the a-priori information from $\mathcal{D}_1$ is low (due to the LF), the MI $I_{\mathcal{D}_{\mathcal{R}_2}}^{\mathrm{sim}}\left(U_{\mathcal{S}}; L_{\mathcal{S}}^{\mathrm{info}}\right)$ for $\mathcal{D}_2$ does only increase slightly. In contrast, the mutual information $I_{\mathcal{D}_{\mathcal{R}_1}}^{\mathrm{sim}}\left(U_{\mathcal{S}}; L_{\mathcal{S}}^{\mathrm{info}}\right)$ of $\mathcal{D}_1$ increases non-linearly with increasing capacity $C_{\mathcal{SR}_1}$, until it reaches its possible maximum. This maximum is determined by the error-probabilities of both relay nodes and the second hop channel capacities.

In **Fig. 6.23**, one can observe that the information combining technique predicts the maximum mutual information $I^{\mathrm{pred}}\left(U_{\mathcal{S}}; L_{\mathcal{S}}^{\mathrm{info}}\right)$, and, therefore helps the destination to choose the decoder output with the highest MI. In the given example, decoder $\mathcal{D}_{\mathcal{R}_2}$ is the best choice if $C_{\mathcal{SR}_1} < 0.3$. Otherwise, for $C_{\mathcal{SR}_1} > 0.3$ decoder $\mathcal{D}_{\mathcal{R}_1}$ should be selected. Without the introduced method for predicting the MI of a relay network

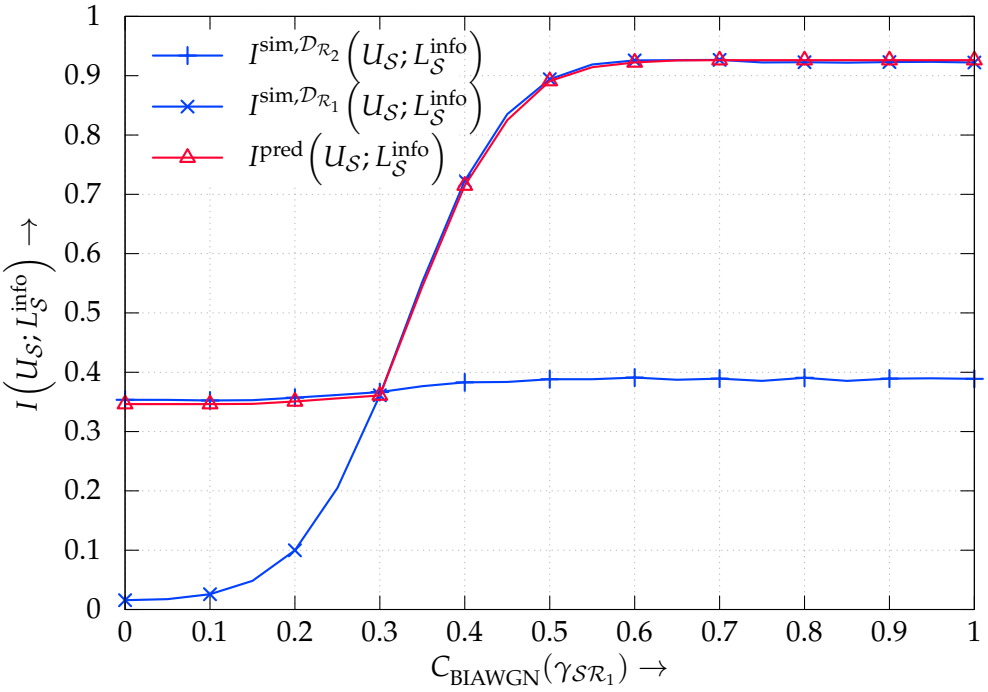


Fig. 6.23.: Measured and predicted MI for the practical four-node relay network with turbo-decoding

via the ICT, the selection of the correct component decoder would be randomly as the destination maybe not posses the needed information about the decoding quality.

Given the mutual information $I^{\text{pred}}\left(U_{\mathcal{S}};L_{\mathcal{S}}^{\text{info}}\right)$ one can estimate the possible bit error rate which could be achieved for the given relay setup via

$$\text{BER}^{\text{pred}} = \frac{1}{2}\cdot\text{erfc}\left(\sqrt{\frac{1}{2\cdot\sigma_N^2}}\right) \quad , \tag{6.41}$$

where σ_N^2 is the noise variance of a virtual BIAWGN channel with a channel capacity corresponding to $I^{\text{pred}}\left(U_{\mathcal{S}};L_{\mathcal{S}}^{\text{info}}\right)$. The result is depicted in **Fig. 6.24**. Although, the model of a BIAWGN channel for the virtual channel does not conform with the true end-to-end channel between source input and destination output, it can be observed that the predicted bit error rate BER^{pred} is quite close to the simulated bit error rate $\text{BER}_{\mathcal{D}_{\mathcal{R}_1}}^{\text{sim}}$ and $\text{BER}_{\mathcal{D}_{\mathcal{R}_2}}^{\text{sim}}$.

The simulation results in case of maximum ratio combining at the destination is depicted in **Fig. 6.25**. Here, the upper and lower bounds from the parallel information combining, i.e. $I^{\text{pred,up}}\left(U_{\mathcal{S}};L_{\mathcal{S}}^{\text{info}}\right)$ and $I^{\text{pred,low}}\left(U_{\mathcal{S}};L_{\mathcal{S}}^{\text{info}}\right)$, tightly embrace the

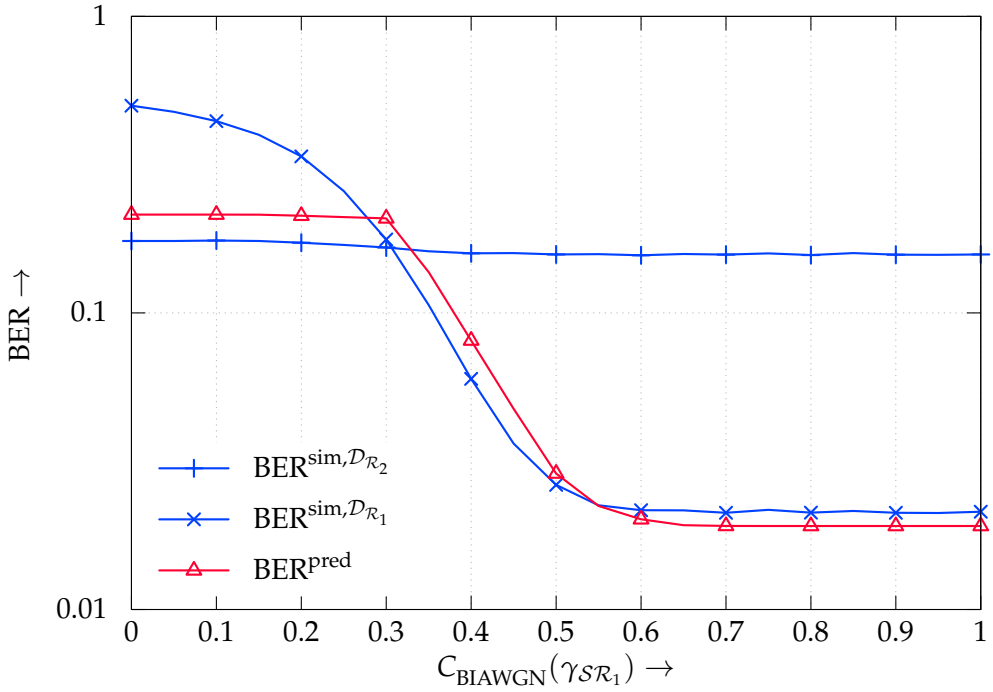


Fig. 6.24.: Measured and predicted BER for the practical four-node relay network with turbo-decoding

simulated MI $I^{\text{sim}}\left(U_{\mathcal{S}}; L_{\mathcal{S}}^{\text{info}}\right)$. The simulated mutual information $I^{\text{no adj}}\left(U_{\mathcal{S}}; L_{\mathcal{S}}^{\text{info}}\right)$ for the relay network without the LLR adjustment, i.e. in case of highly mismatched LLR combining, is clearly below the possible MI of the relay network. This demonstrates that considering the performance and, therewith, the error rate of the first hop is inevitable.

Comparing the mutual information resulting from the turbo-decoding in **Fig. 6.23** with the MI from maximum ratio combining in **Fig. 6.25**, shows that turbo-decoding will lead to a slightly higher MI at approximately $C_{\mathcal{SR}_1} > 0.4$. Before, i.e. $C_{\mathcal{SR}_1} < 0.4$, the MI in case of MRC is larger.

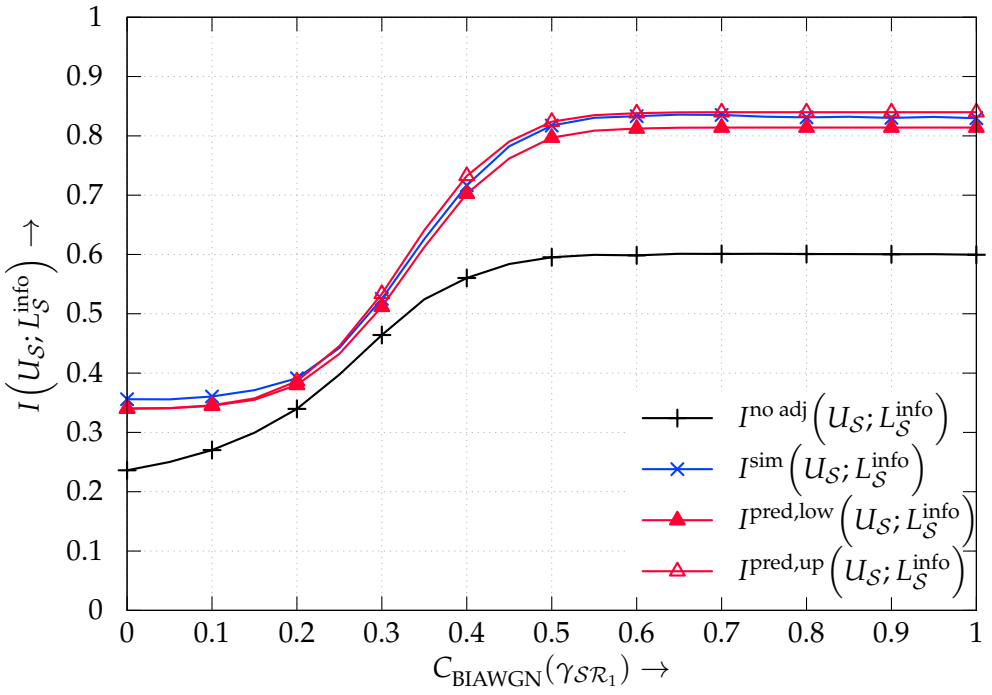


Fig. 6.25.: Measured and predicted MI for the practical four-node relay network with MRC

7 Conclusions and outlook

7.1 Conclusions

The present dissertation is to be understood as a binding link between the information theoretic consideration of relay networks and their practical implementation and evaluation.

Different information processing characteristic [LHHH05b] have been revised to describe the coding scheme behavior of ideal, good and practical codes. Furthermore, the information combining technique from the same authors have been utilized and extended, to evaluate the mutual information of relay networks. For this, the general description for serial and parallel concatenation of discrete symmetric memoryless channel was given and afterwards lower and upper bounds for binary input symmetric memoryless channels have been recapitulated or newly introduced.

New upper bounds have been found for the orthogonal three-node relay network with error-prone relay nodes. Here, the source operates at rates above the direct and indirect link. But it can be shown with the help of the information combining technique and the information processing characteristic for ideal codes, that the combination of the direct and indirect link after the destination's decoders output will deliver enough mutual information for error-free decoding with an outer coding scheme.

Furthermore, different aspects of relay networks with practical coding schemes have been investigated. Among them, especially the performance gain or loss of soft-output re-encoding at the relay nodes were of interest. Here, in the case of a relay scenario, soft output decoding is not always better than hard output decoding, regarding the mutual information performance. The condition of the soft-output decoded received sequence and the condition of the subsequent channels are limiting the overall decoding result at the destination. Here, as long as the second hop is of good quality, i.e. high signal-to-noise ratio, soft re-encoding will be beneficial. If the first hop is good, the soft-values from the decoder output are closely distributed compared to their hard-decided counterparts. Thus, in that case there will be no performance difference between hard-input hard-output and soft-input soft-output re-encoding.

Still an open question which was solved was the calculation of the probability mass function at the destinations decoder input in case of soft-input soft-output re-encoding at the relay node. Here, it was shown that the destination can precisely predict the decoder input distribution's and therewith the log-likelihood ratios with the structure of the convolutional encoder and only two parameters from the relay node. These two parameters are the mean and variance for the LLRs at the relay's decoder output. It was also shown that the sequences of the relay's convolutional encoder outputs have to be considered separately at the destination.

Furthermore, since an analytically expression for calculating the distribution of the boxplus operation is unknown, the min-sum density was utilized to retain an approximation. This approximation was shown to be precise, especially with increasing variance of the LLR values.

To predict the mutual information performance of a relay network, the information combining technique and the information processing characteristics for practical codes have been used. The described method is far less computational complex as Monte Carlo simulation of such networks. Thus, the introduced method can be employed for example in automatic repeat request schemes. In order to achieve a tight prediction also in the case of soft-input soft-output re-encoding, the information processing characteristic was extended such that the performance of the first hop is included. Unfortunately, this leads to a three-dimensional set of MI data points and special care has to be taken using them in conjunction with the ICT.

7.2 Outlook

The delivered framework for predicting the mutual information of practical orthogonal relay networks allows for further research in the field of automatic repeat request. Here, the service or backbone provider or even the nodes itself could easily and quickly evaluate possible code combinations and node cooperations in order to reliable transmit the information from the source towards the destination, under the constraint of least amount of time, energy or other efforts. Still, such an ARQ solution maybe not optimal, since a closed form solution does not exist and performing simulations over all possible scenarios cannot be done in a reasonable amount of time for large networks. But, even a non-optimal solution can reduce the load inside the wireless network dramatically.

Furthermore, the prediction should be extended in order to include non-orthogonal, non-synchronized (e.g. in time and frequency) relay networks as well. This introduces new problems. Here, for example the question of parallel information

combining for non-synchronized and overlapping sequences needs to be answered. Additionally, information processing characteristics for codes with unsynchronized input signals or signals with strong correlations have to be evaluated.

Bibliography

[Ala98] ALAMOUTI, S.: A simple transmit diversity technique for wireless communications. In: *Selected Areas in Communications, IEEE Journal on* 16 (1998), Oct, Nr. 8, S. 1451–1458. http://dx.doi.org/10.1109/49.730453. – DOI 10.1109/49.730453. – ISSN 0733–8716

[Ana01] ANASTASOPOULOS, A.: A comparison between the sum-product and the min-sum iterative detection algorithms based on density evolution. In: *Global Telecommunications Conference, 2001. GLOBECOM '01. IEEE* Bd. 2, 2001, S. 1021 –1025 vol.2

[Ari73] ARIMOTO, S.: On the converse to the coding theorem for discrete memoryless channels (Corresp.). In: *Information Theory, IEEE Transactions on* 19 (1973), May, Nr. 3, S. 357–359. http://dx.doi.org/10.1109/TIT.1973.1055007. – DOI 10.1109/TIT.1973.1055007. – ISSN 0018–9448

[ASR04] AVUDAINAYAGAM, Arun ; SHEA, John M. ; ROONGTA, Abhinav: On approximating the density function of reliabilities of the max-log-map decoder. In: *Fourth IASTED International Multi-Conference on Wireless and Optical Communications (CSAi04)*, 2004, 358–363

[ASW10] AL-SHATRI, H. ; WEBER, T.: Optimizing power allocation in interference channels using D.C. programming. In: *Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt), 2010 Proceedings of the 8th International Symposium on*, 2010, S. 360–366

[AT99] AR, Emre T. ; TELATAR, I. E.: Capacity of Multi-antenna Gaussian Channels. In: *European Transactions on Telecommunications* 10 (1999), S. 585–595

[BCJR74] BAHL, L. ; COCKE, J. ; JELINEK, F. ; RAVIV, J.: Optimal decoding of linear codes for minimizing symbol error rate (Corresp.). In: *Information*

Theory, IEEE Transactions on 20 (1974), Nr. 2, S. 284–287. http://dx.doi.org/10.1109/TIT.1974.1055186. – DOI 10.1109/TIT.1974.1055186. – ISSN 0018–9448

[Bel03] BELL, A. J.: The Co-Information Lattice. Nara, Japan, April 2003, S. 921–926

[BG96] BERROU, C. ; GLAVIEUX, A.: Near optimum error correcting coding and decoding: turbo-codes. In: *Communications, IEEE Transactions on* 44 (1996), Oct, Nr. 10, S. 1261–1271. http://dx.doi.org/10.1109/26.539767. – DOI 10.1109/26.539767. – ISSN 0090–6778

[BGT93] BERROU, C. ; GLAVIEUX, A. ; THITIMAJSHIMA, P.: Near Shannon limit error-correcting coding and decoding: Turbo-codes. 1. In: *Communications, 1993. ICC '93 Geneva. Technical Program, Conference Record, IEEE International Conference on* Bd. 2, 1993, S. 1064–1070 vol.2

[BHS+02] BOSSERT, M. ; HUEBNER, A. ; SCHUEHLEIN, F. ; HAAS, H. ; COSTA, E.: On cyclic delay diversity in OFDM based transmission schemes. In: *journal?* (2002)

[Bos98] BOSSERT, M.: *Kanalcodierung*. Teubner, 1998 http://books.google.de/books?id=BmxGCF9wsGcC. – ISBN 9783519161431

[Bre03] BRENNAN, D. G.: Linear diversity combining techniques. In: *Proceedings of the IEEE* 91 (2003), Nr. 2, S. 331–356. http://dx.doi.org/10.1109/JPROC.2002.808163. – DOI 10.1109/JPROC.2002.808163. – ISSN 0018–9219

[Bri99] BRINK, S. ten: Convergence of iterative decoding. In: *Electronics Letters* 35 (1999), may, Nr. 10, S. 806 –808. http://dx.doi.org/10.1049/el:19990555. – DOI 10.1049/el:19990555. – ISSN 0013–5194

[Bri01] BRINK, S. ten: Convergence behavior of iteratively decoded parallel concatenated codes. In: *Communications, IEEE Transactions on* 49 (2001), oct, Nr. 10, S. 1727 –1737. http://dx.doi.org/10.1109/26.957394. – DOI 10.1109/26.957394. – ISSN 0090–6778

[BSC10] BENNATAN, Amir ; SHAMAI, Shlomo ; CALDERBANK, A. R.: In Praise of Bad Codes for Multi-Terminal Communications. In: *CoRR* abs/1008.1766 (2010). http://arxiv.org/abs/1008.1766

[BT08] BERTSEKAS, D.P. ; TSITSIKLIS, J.N.: *Introduction To Probability*. 2. Athena Scientific, 2008 (Optimization and computation series). http://books.google.com.tr/books?id=yAy-PQAACAAJ,/bib/bertsekas/bertsekas2008introduction/15723_188652923XProbability2nd.pdf. – ISBN 9781886529236

[CG79] COVER, T. ; GAMAL, A.E.: Capacity theorems for the relay channel. In: *Information Theory, IEEE Transactions on* 25 (1979), September, Nr. 5, S. 572–584. – ISSN 0018–9448

[CMH10] CHEN, H. ; MAUNDER, R.G. ; HANZO, L.: An Exit-Chart Aided Design Procedure for Near-Capacity N-Component Parallel Concatenated Codes. In: *Global Telecommunications Conference (GLOBECOM 2010), 2010 IEEE*, 2010. – ISSN 1930–529X, S. 1–5

[CS89] CHAYAT, N. ; SHAMAI, S.: Extension of an entropy property for binary input memoryless symmetric channels. In: *Information Theory, IEEE Transactions on* 35 (1989), Sep, Nr. 5, S. 1077–1079. http://dx.doi.org/10.1109/18.42224. – DOI 10.1109/18.42224. – ISSN 0018–9448

[CT91] COVER, Thomas M. ; THOMAS, Joy A.: *Elements of information theory*. New York, NY, USA : Wiley-Interscience, 1991. – ISBN 0–471–06259–6

[CT06] COVER, Thomas M. ; THOMAS, Joy A.: *Elements of Information Theory 2nd Edition (Wiley Series in Telecommunications and Signal Processing)*. Wiley-Interscience, 2006. – ISBN 0471241954

[CV05] CLEVORN, T. ; VARY, P.: The box-minus operator and its application to low-complexity belief propagation decoding. In: *Vehicular Technology Conference, 2005. VTC 2005-Spring. 2005 IEEE 61st* Bd. 1, 2005. – ISSN 1550–2252, S. 687–691 Vol. 1

[DS03] DOHLER, Mischa ; SHAW, George B.: Virtual Antenna Arrays. 2003. – Forschungsbericht

[EFS56] ELIAS, P. ; FEINSTEIN, A. ; SHANNON, C.: A note on the maximum flow through a network. In: *Information Theory, IEEE Transactions on* 2 (Dec 1956), Nr. 4, S. 117–119. – ISSN 0018–9448

[Fan61] FANO, R.: *Transmission of Information: A Statistical Theory of Communications*. Cambridge, MA : The MIT Press, 1961

[Fei58] FEINSTEIN, Amiel: *Foundations of information theory*. http://opac.inria.fr/record=b1082815. Version: 1958 (McGraw-Hill electrical and electronic engineering series)

[FF56] FORD, L. R. ; FULKERSON, D. R.: Maximal Flow through a Network. In: *Canadian Journal of Mathematics* 8 (1956), 399–404. http://www.rand.org/pubs/papers/P605/

[FG98] FOSCHINI, G. J. ; GANS, M. J.: On limits of wireless communications in a fading environment when using multiple antennas. In: *Wireless Personal Communications* 6 (1998), S. 311–335

[Fos96] FOSCHINI, Gerard J.: Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas. In: *Bell Labs Technical Journal* 1 (1996), Nr. 2, 41–59. http://dx.doi.org/10.1002/bltj.2015. – DOI 10.1002/bltj.2015. – ISSN 1538–7305

[Gal63] GALLAGER, Robert G.: *Low-Density Parity-Check Codes*. 1963

[Gal68] GALLAGER, Robert G.: *Information Theory and Reliable Communication*. New York, NY, USA : John Wiley & Sons, Inc., 1968. – ISBN 0471290483

[GSS+03] GESBERT, D. ; SHAFI, M. ; SHIU, Da shan ; SMITH, P.J. ; NAGUIB, A.: From theory to practice: an overview of MIMO space-time coded wireless systems. In: *Selected Areas in Communications, IEEE Journal on* 21 (2003), Apr, Nr. 3, S. 281–302. http://dx.doi.org/10.1109/JSAC.2003.809458. – DOI 10.1109/JSAC.2003.809458. – ISSN 0733–8716

[Hag94] HAGENAUER, Joachim: SOFT IS BETTER THAN HARD. In: *Communicationses, Coding and Cryptology* (1994), May. http://www.lnt.e-technik.tu-muenchen.de/veroeffentlichungen/1994/ccc94h.pdf. – Kluwer Publication

[HH05] HUETTINGER, Simon ; HUBER, Johannes B.: Information Processing Characteristics and BER-Bounds for Concatenated Coding Schemes. In: *European Transactions on Telecommunications (ETT)* 16 (2005), March, Nr. 2, 137-149. http://www.lit.lnt.de/papers/ett_05_ipc_paper.pdf

[HHFJ02] HUETTINGER, Simon ; HUBER, Johannes ; FISCHER, Robert ; JOHANNESSON, Rolf: Soft-Output-Decoding: Some Aspects from Information

Theory. In: *in Proc. Int. ITG Conf. on Source and Channel Coding*, 2002, S. 81–90

[HHJF01] HUETTINGER, Simon ; HUBER, Johannes ; JOHANNESSON, Rolf ; FISCHER, Robert: Information processing in soft-output decoding. In: *in Proc. Allerton Conf. on Communications, Control, and Computing*, 2001

[HL06] HUBER, J. ; LAND, I.: Information Combining. In: *Foundation and Trends in Communications and and Information Theory* 3 (2006), November, Nr. 3, S. 226–330

[HMZ05] HØST-MADSEN, A. ; ZHANG, J.: Capacity bounds and power allocation for wireless relay channels. 51 (2005), Nr. 6, S. 2020–2040

[Höh13] HÖHER, P.A.: *Grundlagen der digitalen Informationsübertragung: Von der Theorie zu Mobilfunkanwendungen*. Springer Fachmedien Wiesbaden, 2013 `https://books.google.de/books?id=rzggBAAAQBAJ`. – ISBN 9783834822147

[Huy11] HUYNH, K.Q.: *Analysis of Relay Channels and Improved Iterative Decoders for Turbo-coded Decode-and-forward Relay Channels*. Chalmers University of Technology, 2011 (Technical report: L). `http://books.google.de/books?id=0wlFygAACAAJ`

[ITN10] IWAMURA, Mikio ; TAKAHASHI, Hideaki ; NAGATA, Satoshi: Relay Technology in LTE-Advanced. In: *NTT DOCOMO Technical Journal* 12 (2010), sep, Nr. 2, S. 29–36

[JAW07] JINDAL, N. ; ANDREWS, J.G. ; WEBER, S.: Energy-Limited vs. Interference-Limited Ad Hoc Network Capacity. In: *Signals, Systems and Computers, 2007. ACSSC 2007. Conference Record of the Forty-First Asilomar Conference on*, 2007. – ISSN 1058–6393, S. 148–152

[KGG05] KRAMER, G. ; GASTPAR, M. ; GUPTA, P.: Cooperative Strategies and Capacity Theorems for Relay Networks. 51 (2005), Nr. 9, S. 3037–3063

[Lan02] LANEMAN, J. N.: Cooperative Diversity in Wireless Networks: Algorithms and Architectures / Massachusetts Institute of Technology, Cambridge, MA. 2002. – Forschungsbericht

[Lan05a] LAND, I.: *Reliability Information in Channel Decoding: Practical Aspects and Information Theoretical Bounds*. Shaker, 2005 (Digital communications). – ISBN 9783832241551

[Lan05b] LAND, Ingmar ; HÖHER, Prof. Dr.-Ing. Peter A. (Hrsg.): *Reliability Information in Channel Decoding: Practical Aspects and Information Theoretical Bounds*. Shaker Verlag GmbH, Germany, 2005

[LH05] LAND, I. ; HUBER, J.: Information processing in ideal coding schemes with code-symbol decoding. In: *Information Theory, 2005. ISIT 2005. Proceedings. International Symposium on*, 2005, S. 1646–1650

[LHHH03] LAND, I. ; HÜTTINGER, S. ; HÖHER, P. ; HUBER, J.: Bounds on Information Combining. In: *International Symposium on Turbo Codes & Related Topics*. Brest, France, September 2003, S. 39–42

[LHHH05a] LAND, I. ; HUETTINGER, S. ; HOEHER, P.A. ; HUBER, J.B.: Bounds on information combining. In: *Information Theory, IEEE Transactions on* 51 (2005), Nr. 2, S. 612–619. `http://dx.doi.org/10.1109/TIT.2004.840883`. – DOI 10.1109/TIT.2004.840883. – ISSN 0018–9448

[LHHH05b] LAND, Ingmar ; HUETTINGER, S. ; HOEHER, P.A. ; HUBER, J.: Bounds on mutual information for simple codes using information combining. In: *Annales des Telecommunications* 60 (2005), Nr. 1/2, S. 184–214. – ISSN 0003–4347

[LTW04] LANEMAN, J.N. ; TSE, D.N.C. ; WORNELL, Gregory W.: *Cooperative diversity in wireless networks: Efficient protocols and outage behavior*

[LW00] LANEMAN, J. N. ; WORNELL, G. W.: Energy-efficient antenna sharing and relaying for wireless networks. In: *Proc. WCNC Wireless Communications and Networking Confernce 2000 IEEE* Bd. 1, 2000, S. 7–12

[Mac00] MACKAY, David J. C.: *On Threshold of Codes*. `www.cs.toronto.edu/~mackay/theorems.ps.gz`. Version: September 2000

[Mac02] MACKAY, David J. C.: *Information Theory, Inference & Learning Algorithms*. New York, NY, USA : Cambridge University Press, 2002. – ISBN 0521642981

[McG54] MCGILL, W.: Multivariate information transmission. In: *Information Theory, Transactions of the IRE Professional Group on* 4 (1954), Nr. 4,

S. 93–111. http://dx.doi.org/10.1109/TIT.1954.1057469. – DOI 10.1109/TIT.1954.1057469. – ISSN 2168–2690

[Men27] MENGER, Karl: Zur allgemeinen Kurventheorie. In: *Fundamenta Mathematicae* 10 (1927), S. 96–115

[Meu71] MEULEN, E. C. d.: Three-terminal communication channels. In: *Adv. Appl. Probab.* 3 (1971), 120–154. http://ieeexplore.ieee.org/xpls/askIeee.jsp?Doctitle=Three-terminalcommunicationchannels&Volume=3&issue=&page=120-154&authors=E.C.vanderMeulen,&Year=1971&JournalBook=Adv.Appl.Probab.&displayID=15601

[MKP+09] MOGENSEN, P.E. ; KOIVISTO, T. ; PEDERSEN, K.I. ; KOVACS, I.Z. ; RAAF, B. ; PAJUKOSKI, K. ; RINNE, M.J.: LTE-Advanced: The path towards gigabit/s in wireless mobile communications. In: *Wireless Communication, Vehicular Technology, Information Theory and Aerospace Electronic Systems Technology, 2009. Wireless VITAE 2009. 1st International Conference on*, 2009, S. 147–151

[MN97] MACKAY, D. J C. ; NEAL, R. M.: Near Shannon limit performance of low density parity check codes. In: *Electronics Letters* 33 (1997), Nr. 6, S. 457–458. http://dx.doi.org/10.1049/el:19970362. – DOI 10.1049/el:19970362. – ISSN 0013–5194

[PF01] PAPADIAS, C.B. ; FOSCHINI, G.J.: A space-time coding approach for systems employing four transmit antennas. In: *Acoustics, Speech, and Signal Processing, 2001. Proceedings. (ICASSP '01). 2001 IEEE International Conference on* Bd. 4, 2001. – ISSN 1520–6149, S. 2481 –2484 vol.4

[PSS05] PELEG, Michael ; SANDEROVICH, Amichai ; SHAMAI, Shlomo: On Extrinsic Information of Good Codes Operating Over Discrete Memoryless Channels. In: *CoRR* abs/cs/0504028 (2005). http://dblp.uni-trier.de/db/journals/corr/corr0504.html#abs-cs-0504028

[PSS07] PELEG, Michael ; SANDEROVICH, Amichai ; SHAMAI, Shlomo: Information Theory On extrinsic information of good binary codes operating over Gaussian channels. In: *European Transactions on Telecommunications* 18 (2007), Nr. 2, 133-139. http://dblp.uni-trier.de/db/journals/ett/ett18.html#PelegSS07

[Rei66] REIFFEN, B.: A per letter converse to the channel coding theorem. In: *Information Theory, IEEE Transactions on* 12 (1966), Oct, Nr. 4, S. 475–480. http://dx.doi.org/10.1109/TIT.1966.1053928. – DOI 10.1109/TIT.1966.1053928. – ISSN 0018–9448

[RFLT98] RASHID-FARROKHI, F. ; LIU, K.J.R. ; TASSIULAS, L.: Transmit beamforming and power control for cellular wireless systems. In: *Selected Areas in Communications, IEEE Journal on* 16 (1998), Oct, Nr. 8, S. 1437–1450. http://dx.doi.org/10.1109/49.730452. – DOI 10.1109/49.730452. – ISSN 0733–8716

[Rin08] RINNE, H.: *Taschenbuch der Statistik*. Deutsch Harri GmbH, 2008 http://books.google.de/books?id=_q5OzktC92AC. – ISBN 9783817118274

[Roh76] ROHATGI, V.K.: *An introduction to probability theory and mathematical statistics*. Wiley, 1976 (Wiley series in probability and mathematical statistics: Probability and mathematical statistics). http://books.google.de/books?id=YyXvAAAAMAAJ

[RT02] REGGIANI, L. ; TARTARA, G.: Probability density functions of soft information. In: *Communications Letters, IEEE* 6 (2002), Nr. 2, S. 52–54. http://dx.doi.org/10.1109/4234.984688. – DOI 10.1109/4234.984688. – ISSN 1089–7798

[RU08] RICHARDSON, Tom ; URBANKE, Ruediger: *Modern Coding Theory*. New York, NY, USA : Cambridge University Press, 2008. – ISBN 0521852293, 9780521852296

[Sha48] SHANNON, C. E.: A mathematical theory of communication. In: *Bell system technical journal* 27 (1948)

[Sha59] SHANNON, C. E.: Coding theorems for a discrete source with a fidelity criterion. In: *IRE Nat. Conv. Rec., Pt. 4.* 1959, S. 142–163

[SL05] SCHULZE, H. ; LUEDERS, C.: *Theory and Applications of OFDM and CDMA: Wideband Wireless Communications*. Wiley, 2005 http://books.google.de/books?id=PzWSwn9gHWEC. – ISBN 9780470017395

[SLV08] SNEESSENS, H.H. ; LOUVEAUX, J. ; VANDENDORPE, L.: Turbo-coded decode-and-forward strategy resilient to relay errors. In: *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*, 2008. – ISSN 1520–6149, S. 3213–3216

[TB00] TEN BRINK, S.: Rate one-half code for approaching the Shannon limit by 0.1 dB. In: *Electronics Letters* 36 (2000), Nr. 15, S. 1293–1294. http://dx.doi.org/10.1049/el:20000953. – DOI 10.1049/el:20000953. – ISSN 0013–5194

[Tho08] THOBABEN, R.: On distributed codes with noisy relays. In: *Signals, Systems and Computers, 2008 42nd Asilomar Conference on*, 2008. – ISSN 1058–6393, S. 1010–1014

[TSC98] TAROKH, Vahid ; SESHADRI, N. ; CALDERBANK, A.R.: Space-time codes for high data rate wireless communication: performance criterion and code construction. In: *Information Theory, IEEE Transactions on* 44 (1998), Mar, Nr. 2, S. 744–765. http://dx.doi.org/10.1109/18.661517. – DOI 10.1109/18.661517. – ISSN 0018–9448

[VA87] VAUGHAN, R.G. ; ANDERSEN, J.B.: Antenna diversity in mobile communications. In: *Vehicular Technology, IEEE Transactions on* 36 (1987), Nov, Nr. 4, S. 149–172. http://dx.doi.org/10.1109/T-VT.1987.24115. – DOI 10.1109/T–VT.1987.24115. – ISSN 0018–9545

[Vit67] VITERBI, A.J.: Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. In: *Information Theory, IEEE Transactions on* 13 (1967), Nr. 2, S. 260–269. http://dx.doi.org/10.1109/TIT.1967.1054010. – DOI 10.1109/TIT.1967.1054010. – ISSN 0018–9448

[Vor13] VORKÖPER, Sebastian: *Some notes on LLR calculation for practical relay networks*. Talk, April 2013. – Presentation at the Institute of Communications Engineering, University of Rostock

[Wei09] WEINSTEIN, S.B.: The history of orthogonal frequency-division multiplexing [History of Communications]. In: *Communications Magazine, IEEE* 47 (2009), November, Nr. 11, S. 26–35. http://dx.doi.org/10.1109/MCOM.2009.5307460. – DOI 10.1109/MCOM.2009.5307460. – ISSN 0163–6804

[Wei13] WEITKEMPER, Petra: *Signalverarbeitung für die Relay-gestützte Datenübertragung*. Bremen, Germany, University of Bremen, Diss., July 2013

[WFGV98] WOLNIANSKY, P.W. ; FOSCHINI, G.J. ; GOLDEN, G.D. ; VALENZUELA, R.: V-BLAST: an architecture for realizing very high data rates over the

rich-scattering wireless channel. In: *Signals, Systems, and Electronics, 1998. ISSSE 98. 1998 URSI International Symposium on*, 1998, S. 295–300

[Wol64] WOLFOWITZ, J.: *Coding theorems of information theory*. Springer, 1964 (Ergebnisse der Mathematik und ihrer Grenzgebiete). http://books.google.de/books?id=m0NgAAAAMAAJ

[Wol67] WOLFOWITZ, J.: Memory increases capacity. In: *Information and Control* 11 (1967), Nr. 4, 423 - 428. http://dx.doi.org/http://dx.doi.org/10.1016/S0019-9958(67)90640-7. – DOI http://dx.doi.org/10.1016/S0019–9958(67)90640–7. – ISSN 0019–9958

[WW96] WIBERG, Niclas ; WIBERG, Niclas: *Codes and Decoding on General Graphs*. 1996

[WWKK08] WEITKEMPER, P. ; WÜBBEN, D. ; KÜHN, V. ; KAMMEYER, K.-D.: Soft Information Relaying for Wireless Networks with Error-Prone Source-Relay Link. In: *7th International ITG Conference on Source and Channel Coding*. Ulm, Germany, Jan 2008

[Wyn73] WYNER, A.D.: A theorem on the entropy of certain binary sequences and applications–II. In: *Information Theory, IEEE Transactions on* 19 (1973), Nov, Nr. 6, S. 772–777. http://dx.doi.org/10.1109/TIT.1973.1055108. – DOI 10.1109/TIT.1973.1055108. – ISSN 0018–9448

[WZ73] WYNER, A.D. ; ZIV, J.: A theorem on the entropy of certain binary sequences and applications–I. In: *Information Theory, IEEE Transactions on* 19 (1973), Nov, Nr. 6, S. 769–772. http://dx.doi.org/10.1109/TIT.1973.1055107. – DOI 10.1109/TIT.1973.1055107. – ISSN 0018–9448

[WZ76] WYNER, Aaron D. ; ZIV, Jacob: The rate-distortion function for source coding with side information at the decoder. In: *IEEE Trans. Inform. Theory* 22 (1976), S. 1–10

Publications

[HHV+12] HINKFOTH, M. ; HEINRICH, E. ; VORKÖPER, S. ; KÜHN, V. ; SALOMON, R.: X-ORCA: FPGA-based wireless localization in the sub-millimeter range. In: *20th ACM/SIGDA International Symposium on Field-Programmable Gate Arrays (FPGA'12)*. Monterey, California, USA, february 2012, S. 29–32. – ISBN: 978-1-4503-1155-7, DOI: 10.1145/2145694.2145700

[KKV09] KHAN, A. ; KÜHN, V. ; VORKÖPER, S.: Information Combining Based ARQ in Relay Networks. In: *15th European Wireless Conference (EW 2009)*. Aalborg, Denmark, may 2009

[KV09a] KÜHN, V. ; VORKÖPER, S.: Application of Information Combining to Relay Networks. In: *IEEE Vehicular Technology Conference (VTC 2009 Spring)*. Barcelona, Spain, 2009

[KV09b] KÜHN, V. ; VORKÖPER, S.: Information Combining for Relay Networks. In: *IEEE International Conference on Communications (ICC 2009)*. Dresden, Germany, 2009

[KVK14] KERN, D. ; VORKÖPER, S. ; KÜHN, V.: A new code construction for polar codes using min-sum density. In: *8th International Symposium on Turbo Codes and Iterative Information Processing (ISTC), 2014*. Bremen, Germany, August 2014, S. 228–232

[VK08] VORKÖPER, S. ; KÜHN, V.: Information Combining for OFDM-based Relay Networks. In: *13th International OFDM-Workshop*. Hamburg, Germany, august 2008

[VK09a] VORKÖPER, S. ; KÜHN, V.: Noise Variance Estimation in OFDM Cyclic Delay Diversity Relay Networks. In: *14th International OFDM-Workshop (InOWo 2009)*. Hamburg, Germany, september 2009

[VK09b] VORKÖPER, S. ; KÜHN, V.: Reducing the Channel Estimation Error for Relay Networks with Cyclic Delay Diversity. In: *15th European Wireless Conference (EW 2009)*. Aalborg, Denmark, may 2009, S. 53–57

[VK11a] VORKÖPER, S. ; KÜHN, V.: Analyzing the Performance of Error-Prone Relay Networks. In: *Proc. IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP'11)*. Prague, may 2011

[VK11b] VORKÖPER, S. ; KÜHN, V.: CFO Estimation and Compensation in LTE Compliant DF Relay Networks. In: *Proc. of the 16th International OFDM-Workshop 2011 (InOWo'11)*. Hamburg, Germany, august 2011, S. 76–80

[VK11c] VORKÖPER, S. ; KÜHN, V.: Error-Prone Relay Networks with Soft- and Hard-Decision Re-Encoding. In: *International ITG Workshop on Smart Antennas (WSA'11)*. Aachen, Germany, february 2011

[VK12a] VORKÖPER, S. ; KÜHN, V.: Performance Comparison of Distributed IDM-STC versus Cooperative OFDM for practical Decode-and-Forward Relay-Networks. In: *International ITG Workshop on Smart Antennas (WSA'12)*. Dresden, march 2012, S. 145–151

[VK12b] VORKÖPER, S. ; KÜHN, V.: Power and Subcarrier Allocation Scheme for Uplink OFDMA. In: *Proceedings of OFDM 2012, 17th International OFDM Workshop 2012 (InOWo'12)*, 2012, S. 1–6

[VLK+13] VORKÖPER, S. ; LENKEIT, F. ; KÜHN, V. ; WÜBBEN, D. ; DEKORSY, A.: Distributed IDM-STC versus Cooperative OFDM for the Two-Hop Decode-and-Forward Multiple-Access Relay Channel. In: *International ITG Workshop on Smart Antennas (WSA'13)*. Stuttgart, march 2013

[WVW+09] WEITKEMPER, P. ; VORKÖPER, S. ; WÜBBEN, D. ; KAMMEYER, K.-D. ; KÜHN, V.: Distributed IDM-STC versus OFDM-CDD in Two-Hop Relay-Networks. In: *14th International OFDM-Workshop (InOWo 2009)*. Hamburg, Germany, september 2009

Appendix

A

A.1 Proof of the ideal coding scheme

In [HHJF01] and [HHFJ02], Huettinger, Huber, Johannesson and Fischer introduced and proofed the concept of ideal coding schemes, employing the data processing theorem [CT06], the rate distortion function [Sha59] and Fano's inequality [Fan61].

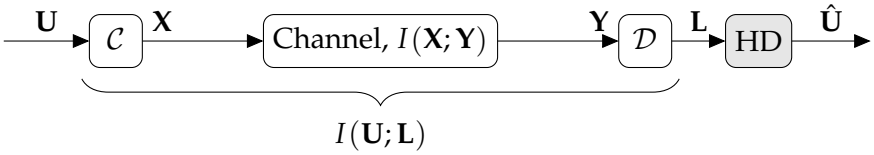


Fig. A.1.: Simplified communication model

Depicted in **Fig. A.1** is a simplified communication model, including the i.i.d. binary input sequence $\mathbf{U}$, encoder $\mathcal{C}$ with binary output sequence $\mathbf{X}$, the underlying channel with mutual information $I(\mathbf{X};\mathbf{Y})$, the decoder $\mathcal{D}$ and hard-decision HD delivering soft- and hard-estimates $\mathbf{L}$ and $\hat{\mathbf{U}}$, respectively. The mutual informations $I(\mathbf{U};\mathbf{L})$ and $I(\mathbf{U};\hat{\mathbf{U}})$ of the soft- and hard-estimate end-to-end channel is taken from the encoder input and decoder or hard-decision output.

Relating to **Fig. A.1**, the proof for the ideal coding scheme starts with the chain of inequalities

$$\begin{aligned}\sum_{k=1}^{K} I(U_k;L_k) &\overset{a}{\leq} I(\mathbf{U};\mathbf{L}) \\ &\overset{b}{\leq} I(\mathbf{X};\mathbf{Y}) \\ &\overset{c}{\leq} NC \ , \end{aligned} \tag{A.1}$$

where the inequality in **a** comes from the fact that memory increases mutual information [Wol67] and, thus, symbol-wise mutual information cannot exceed vector-wise (sequence-wise) MI [HH05]. Data processing theorem is utilized in **b**, where the MI of the end-to-end channel cannot be larger than that of the physical one and in **c**, the channel capacity is maximized mutual information. Please note that, K denotes the information- and N the code-sequence length, respectively.

Since, for the information processing characteristic of ideal codes one is interested in the symbol-wise MI, the chain of inequalities from (A.1) results in

$$I\left(U;\hat{U}\right) \leq I(U;L) \leq \frac{1}{K} I(\mathbf{U};\mathbf{L}) \leq \frac{1}{K} I(\mathbf{X};\mathbf{Y}) \leq \frac{C}{R_c} \ , \tag{A.2}$$

with the definition of mean MI $I(U;L) = \frac{1}{K} \sum_{k=1}^{K} I(U_k;L_k)$ and the obvious fact that a hard-decision (retaining $\hat{U}$) on soft-symbols L cannot increase information.

Given Fano's inequality[1] $H_b(\epsilon_k) \geq H\left(U_k|\hat{U}_k\right)$ and Jensen's inequality $\mathrm{E}_x\left\{f(x)\right\} \leq f\left(\mathrm{E}_x\left\{x\right\}\right)$ for convex-$\cap$ functions $f(x)$, one can show that

$$\frac{1}{K} \sum_{k=1}^{K} H\left(U_k|\hat{U}_k\right) \leq \frac{1}{K} \sum_{k=1}^{K} H_b(\epsilon_k) = \mathrm{E}_k\left\{H_b(\epsilon_k)\right\} \leq H_b\left(\mathrm{E}_k\left\{\epsilon_k\right\}\right) = H_b(\epsilon) \ , \tag{A.3}$$

the conditional entropy is bounded by the binary entropy function $H_b(\epsilon)$ with crossover (error) probability ϵ. For further explanation and since the binary entropy function is symmetric, the crossover probabilities are restricted to $\epsilon,\ \epsilon_k \in \left[0, \frac{1}{2}\right]$.

Inserting (A.3) in $I\left(U;\hat{U}\right)$ delivers

$$I\left(U;\hat{U}\right) = \frac{1}{K} \sum_{k=1}^{K} \left(H(U_k) - H\left(U_k|\hat{U}_k\right)\right) \geq 1 - H_b(\epsilon) \ , \tag{A.4}$$

i.e. the bit error rate of any coding scheme providing an end-to-end mutual information $I\left(U;\hat{U}\right)$ has to be larger or equal to the BER $\epsilon = H_b^{-1}\left(1 - I\left(U;\hat{U}\right)\right)$ of the hard-decision output $\hat{U}$ of a memoryless channel.

Following rate-distortion theory [Sha59], a code with rate R_c and appropriate decoding rule exists which achieves an average bit error rate $\epsilon \leq \epsilon_{\mathrm{T}}$ as long as

$$R_c \leq \frac{C}{1 - H_b(\epsilon_{\mathrm{T}})} \tag{A.5}$$

and the code word length N tends to infinity [HHJF01] [Mac02, p. 162]. Thus $\frac{C}{R_c} \geq 1 - H_b(\epsilon_{\mathrm{T}}) = C_{\mathrm{BSC}}$ holds.

In (A.5) $\epsilon_{\mathrm{T}} \in \left[0, \frac{1}{2}\right]$ is the tolerated BER which is greater or equal to the BER ϵ of an ideal coding scheme, i.e. $\epsilon \leq \epsilon_{\mathrm{T}}$ holds. Since (A.2) and (A.3) requires $\epsilon \geq \epsilon_{\mathrm{T}}$, an

[1] The inequality $H_b(\epsilon_k) \geq H\left(U_k|\hat{U}_k\right)$ is the stronger form [CT06, p. 39] of Fano's inequality with a binary alphabet.

ideal coding scheme only exists in the equality of $\epsilon = \epsilon_{\mathrm{T}}$ and (A.2) restricts to the chain of equalities

$$1 - H_b(\epsilon) = I\left(U;\hat{U}\right) = \frac{1}{K} I\left(\mathbf{U};\hat{\mathbf{U}}\right) = 1 - H_b(\epsilon_{\mathrm{T}}) \ . \tag{A.6}$$

Hence, for an ideal coding scheme, the end-to-end channel is a BSC, achieving minimum possible BER. Therefore, rewriting (A.5) by inserting equality (A.6) results in the ideal information processing characteristic for memoryless symmetric channels

$$\mathrm{IPC}^{\mathrm{ideal}}(C) = I\left(U;\hat{U}\right) \leq \begin{cases} C/R_c & \text{for } C < R_c \\ 1 & \text{for } C \geq R_c \end{cases} \ . \tag{A.7}$$

A.2 Weak converse versus ideal coding scheme

The weak converse of the coding theorem states that the decoding error probability is bounded away from zero if $R_c > C$, i.e. having a lower bound on the error probability strictly greater than zero, as the code block-length goes to infinity. In fact, in [Fan61, p. 189] and [CT06, p. 208] the average probability of decoding error (word error probability) is lower bounded by

$$\mathrm{Pr}_{\mathrm{err}} = \mathrm{Pr}\left(\mathbf{U} \neq \hat{\mathbf{U}}\right) \geq \begin{cases} 1 - \frac{C}{R_c} & \text{for } C > R_c \\ 0 & \text{for } C \leq R_c \end{cases} \ , \tag{A.8}$$

given the i.i.d. binary sequence of random variable $\mathbf{U} = [U_k]_{Kx1}$, $U_k \in \{0,1\}$. The prove [Höh13, p. 70] starts with (consider **Fig. A.1** and Appendix A.1 for the meaning of the different variables)

$$\begin{aligned} H\left(\mathbf{U}|\hat{\mathbf{U}}\right) &= H(\mathbf{U}) - I\left(\mathbf{U};\hat{\mathbf{U}}\right) \\ &\overset{a}{\geq} H(\mathbf{U}) - I\left(\mathbf{X};\hat{\mathbf{Y}}\right) \\ &\overset{b}{\geq} H(\mathbf{U}) - NC = K - NC \end{aligned} \tag{A.9}$$

a the data-processing inequality and **b** the definition of the channel capacity for a memoryless channel. Given Fano's inequality

$$H\left(\mathbf{U}|\hat{\mathbf{U}}\right) \leq H_b(\mathrm{Pr}_{\mathrm{err}}) + \mathrm{Pr}_{\mathrm{err}} \cdot \log_2\left(|\mathbb{M}| - 1\right) \tag{A.10}$$

and the alphabet size $|\mathbb{M}| = 2^K$ of all possible code-words, the decoding error probability is calculated by inserting (A.10) into (A.9)

$$H_b(\mathrm{Pr}_{\mathrm{err}}) + \mathrm{Pr}_{\mathrm{err}} \cdot \log_2\left(|\mathbb{M}| - 1\right) \geq K - NC \tag{A.11}$$

and resolving for $\mathrm{Pr}_{\mathrm{err}}$ via $\log_2\left(2^K - 1\right) \approx K$ and $H_b(\mathrm{Pr}_{\mathrm{err}}) = 1$ (the largest possible binary entropy) will deliver

$$\begin{aligned} \mathrm{Pr}_{\mathrm{err}} &\geq 1 - \frac{C}{R_c} - \frac{1}{NR_c} \\ &\overset{a}{\approx} 1 - \frac{C}{R_c} \ , \end{aligned} \tag{A.12}$$

given the code rate $R_c = K/N$ and in **a** letting K and N approach infinity.

Now, to show the equality in the decoding error probability of the ideal coding scheme and the weak converse, the alphabet size of all possible code-words is restricted to $|\mathbb{M}| = 2$, since only one binary information symbol ($K = 1$) is considered. Then Fano's inequality reduces to $H\left(U|\hat{U}\right) \leq H_b(\mathrm{Pr}_{\mathrm{err}})$ (see the analogy to Appendix A.1). Thus, inserting $H\left(U|\hat{U}\right)$ in (A.9) delivers ($K = 1$ and $N = K/R_c$)

$$H_b(\mathrm{Pr}_{\mathrm{err}}) \geq K - NC = 1 - \frac{C}{R_c} \ , \tag{A.13}$$

which by resolving for $I\left(U;\hat{U}\right) = 1 - H_b(\mathrm{Pr}_{\mathrm{err}}) = C/R_c$ is equivalent to (A.7).

A.3 Random variables, probabilities and probability functions

Throughout this dissertation, a random variable is denoted by capital letters, like X. Numerical values or outcomes of the random variable X, which are real valued numbers, are denoted by small letters x. Blackboard bold letters denote the set of outcomes $\mathbf{x} \in \mathbb{X} \subset \mathbb{R}$ of the random variable X, i.e. the possible values of X.

If X is a discrete random variable (rv), then the set of outcomes is at most countably infinite and the corresponding probability distribution is given by the probability mass function (pmf) $p_X(x) = \mathrm{Pr}(X = x)$. Simply, $p_X(x)$ delivers the probability of the event $\{X = x\}$. Furthermore, the sum over the pmf for all possible outcomes of

X yields the property $\sum_{x \in \mathbb{X}} p_X(x) = 1$. The sum over the probability mass function $p_X(x)$ in case of $X \leq x$

$$F_X(x) = \Pr(X \leq x) = \sum_{k \leq x} p_X(k) \ ,$$

will lead to the cumulative distribution function (cdf) of a discrete distribution.

Please note, that throughout this dissertation mostly discrete random variables are utilized. In the rare cases, there continuous random variables are needed, they are specially emphasized.

A.4 Some properties and distributions of random variables

In this section, some important properties, distributions and transformations of random variables are given. Most of them are well known and commonly used in publications without references. For others, which might be unusual, references will be given.

Besides many good and interesting books in the field of probability theory, there exist two of them, which should be specially recommended. The first book [BT08] is from Dimitri P. Bertsekas and John N. Tsitsiklis. This book not only delivers a good introduction into probability theory, but also has a good and clean notation, which is sometimes missing in other academic books. The second book [Roh76] from V K. Rohatgi is not recommended for beginners, but delivers a (almost) complete set of different formulas in the scientific field of probability.

A.4.1 Gaussian distribution, $\mathcal{N}(\mu_X, \sigma_X^2)$

Let the random variable $X \sim \mathcal{N}(\mu_X, \sigma_X^2)$ be Gaussian (or normal) distributed with mean $\mu_X \in \mathbb{R}$ and variance $\sigma_X^2 \geq 0$, then the probability mass function (pmf) is calculated via

$$p_X(x) = \frac{1}{\sqrt{2\pi\sigma_X^2}} e^{-\frac{(x-\mu_X)^2}{2\sigma_X^2}} \ , \tag{A.14}$$

with corresponding cumulative distribution function (cdf)

$$F_X(x) = \frac{1}{2} \cdot \left[1 + \operatorname{erf}\left(\frac{x - \mu_X}{\sqrt{2\sigma_X^2}} \right) \right] \quad . \tag{A.15}$$

A.4.2 Bernoulli distribution, $\mathcal{B}(\epsilon)$

Let $U \sim \mathcal{B}(\epsilon)$ be a Bernoulli distributed binary random variable taking on values $u \in \{0, 1\}$ and corresponding probability of occurrence $\Pr(U = 0) = 1 - \epsilon$ and $\Pr(U = 1) = \epsilon$, then the probability mass function for U is calculated via

$$p_U(u) = \begin{cases} 1 - \epsilon & \text{for } u = 0 \\ \epsilon & \text{for } u = 1 \end{cases} , \tag{A.16}$$

with corresponding cumulative distribution function (cdf)

$$F_U(u) = \begin{cases} 0 & \text{for } u < 0 \\ 1 - \epsilon & \text{for } 0 \leq u < 1 \\ 1 & \text{for } u \geq 1 \end{cases} . \tag{A.17}$$

A.4.3 Derived distributions, $Y = g(X)$

Let X be a discrete random variable with probability mass function $p_X(x)$. Suppose, that the discrete rv $Y = g(X)$ is a function of X. Then, the pmf $p_Y(x)$ of Y is given by [BT08, p. 80]

$$p_Y(y) = \sum_{\{x | g(x) = y\}} p_X(x) \quad . \tag{A.18}$$

In case of continuous X, calculating the probability density function $f_Y(y)$ of Y is a two step process [BT08, p. 202]:

1. Calculate the cumulative distribution function $F_Y(x)$ of Y via

$$F_Y(y) = \Pr(g(X) \leq y) = \int_{\{x | g(x) \leq y\}} f_X(x) \, \mathrm{d}x \tag{A.19}$$

2. Take the derivative of the cdf to retain the pdf $f_Y(x)$ of Y

$$f_Y(y) = \frac{d}{dy} F_Y(y) \tag{A.20}$$

If $g(x)$ is a strictly monotonic function, with its inverse $x = g^{-1}(y)$ being differentiable, then the pdf of $f_Y(y)$ can be calculated directly from

$$f_Y(y) = f_X\left(g^{-1}(y)\right) \left| \frac{d}{dy} g^{-1}(y) \right| \ , \tag{A.21}$$

given $f_X(x)$.

A.4.4 Sum of two random variables, $Z = X + Y$

Let X and Y to be i.i.d. distributed discrete random variables with probability mass function $p_X(x)$ and $p_Y(y)$ then the pmf of the rv $Z = X + Y$ is given by [BT08, p. 213]

$$p_Z(z) = \sum_{x \in \mathbb{X}} p_X(x) p_Y(z - x) \ , \tag{A.22}$$

with corresponding shorthand notation $p_Z(z) = \mathcal{F}^+ \left(p_X(x), p_Y(y)\right)$.

A.4.5 Product of two random variables, $Z = X \cdot Y$

Let X and Y be to i.i.d. distributed random variables with probability mass function $p_X(x)$ and $p_Y(x)$ than the pmf of the rv $Z = X \cdot Y$ is given by [Roh76, p. 141]

$$p_Z(z) = \sum_{x \in \mathbb{X}} \frac{1}{|x|} p_X(x) p_Y\left(\frac{z}{x}\right) \ , \tag{A.23}$$

with corresponding shorthand notation $p_Z(z) = \mathcal{F}^\times \left(p_X(x), p_Y(y)\right)$.

A.4.6 Linear function of a random variable, $Y = aX + b$

If the random variable $Y = aX + b$ is a linear function of the rv X with scaling and shifting factor $a \neq 0$ and b, respectively, then the probability mass function $p_Y(y)$ of Y is given by [BT08, p. 205]

$$p_Y(y) = \frac{1}{|a|} p_X\left(\frac{y - b}{a}\right) \ , \tag{A.24}$$

with $p_X(x)$ being the probability mass function of X. The shorthand notation is $p_Y(y) = \mathcal{F}^{\triangleright}(p_X(x), a, b)$.

Since scaling a random variable is often used in this dissertation, the pmf of this function is denoted by the shorthand notation of $\mathcal{F}^{\triangleright}(p_X(x), a) = \mathcal{F}^{\triangleright}(p_X(x), a, 0)$.

A.4.7 Derived distribution for $Y = \tanh(X/2)$

Let X be a discrete random variable with probability mass function $p_X(x)$, than the pmf of the rv $Y = \tanh(X/2)$ is given by (A.21)

$$p_Y(y) = p_X\left(2 \cdot \tanh^{-1}(y)\right) \left| \frac{2}{1-y^2} \right| \quad , \tag{A.25}$$

with $y = \tanh(x/2)$, its inverse function $x = 2 \cdot \tanh^{-1}(y)$ and corresponding derivative

$$\left| \frac{d}{dy} g^{-1}(y) \right| = \left| \frac{2}{1-y^2} \right| \quad . \tag{A.26}$$

The shorthand notation is $p_Y(y) = \mathcal{F}^{\curlywedge}(p_X(x))$.

A.4.8 Derived distribution for $Y = |X|$

Let X be a discrete random variable with probability mass function $p_X(x)$, than the pmf of the rv $Y = |X|$ is calculated by [Rin08, p. 214]

$$p_Y(y) = \begin{cases} 0 & \text{for } y < 0 \\ p_X(y) + p_X(-y) & \text{for } y \geq 0 \end{cases} \quad , \tag{A.27}$$

with corresponding cumulative distribution function

$$F_Y(y) = \begin{cases} 0 & \text{for } y < 0 \\ F_X(y) - F_X(-y) & \text{for } y \geq 0 \end{cases} \quad . \tag{A.28}$$

The shorthand notation is given by $p_Y(y) = \mathcal{F}^{|X|}(p_X(x))$.

A.4.9 Modulo-2 sum density, $U_Z = U_X \oplus U_Y$

Let $U_X \sim \mathcal{B}(\epsilon_X)$ and $U_Y \sim \mathcal{B}(\epsilon_Y)$ be i.i.d. binary random variables with pmfs $p_{U_X}(u_x)$, $p_{U_Y}(u_y)$ and probabilities of occurrence ϵ_X and ϵ_Y. The pdf for the modulo-2 sum $U_Z = U_X \oplus U_Y$ (*xor* operation) is given by the probabilities

$$\begin{aligned}
\Pr(U_Z = 0) &= \Pr(U_X = 0 \cap U_Y = 0) + \Pr(U_X = 1 \cap U_Y = 1) \\
&= (1-\epsilon_X)(1-\epsilon_Y) + \epsilon_X\epsilon_Y \\
\Pr(U_Z = 0) &= \Pr(U_X = 0 \cap U_Y = 0) + \Pr(U_X = 1 \cap U_Y = 1) \\
&= (1-\epsilon_X)\epsilon_Y + \epsilon_X(1-\epsilon_Y)
\end{aligned}$$

and thus

$$p_{U_Z}(u_z) = \begin{cases} (1-\epsilon_X)(1-\epsilon_Y) + \epsilon_X\epsilon_Y & \text{for } u_z = 0 \\ (1-\epsilon_X)\epsilon_Y + \epsilon_X(1-\epsilon_Y) & \text{for } u_z = 1 \end{cases} \tag{A.29}$$

holds. The shorthand notation is $p_{U_Z}(u_z) = \mathcal{F}^{\oplus}\left(p_{U_X}(u_x), p_{U_Y}(u_y)\right)$.

A.4.10 Boxplus and min-sum density

Let L_X and L_Y be i.i.d. random variables with probability mass functions $p_{L_X}(l_x)$, $p_{L_Y}(l_y)$ and corresponding cumulative distribution functions $F_{L_X}(l_x)$, $F_{L_Y}(l_y)$. Let the random variable L_Z be the outcome from the *boxplus* operation [Hag94] of L_X and L_Y given by

$$\begin{aligned}
L_Z &= L_X \boxplus L_Y \\
&= 2 \cdot \tanh^{-1}\left(\tanh\left(\frac{L_X}{2}\right) \cdot \tanh\left(\frac{L_Y}{2}\right)\right) \quad .
\end{aligned} \tag{A.30}$$

In order to calculate the probability mass function $p_{L_Z}(l_z)$ for the rv L_Z, the second line in (A.30) is utilized. Here, the steps to retain $p_{L_Z}(l_z)$ includes:

1. variable transformation for $X = \tanh(L_X/2)$ and $Y = \tanh(L_Y/2)$
2. pmf for the the product of two i.i.d. random variables $Z = X \cdot Y$
3. (inverse) variable transformation $L_Z = 2 \cdot \tanh^{-1}(Z)$.

Suppose, $L_X \sim \mathcal{N}(\mu_X, \sigma_X^2)$ and $L_Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$ are discrete Gaussian distributed, then the pmfs for $X = \tanh(L_X/2)$ and $Y = \tanh(L_Y/2)$ are given by (A.25)

$$p_X(x) = \frac{2}{(1-x^2)\sqrt{2\pi\sigma_X^2}} e^{-\frac{(2\tanh^{-1}(x)-\mu_X)^2}{2\sigma_X^2}} \tag{A.31}$$

and $p_Y(y)$, respectively.

Since L_X and L_Y are i.i.d. random variables, the same holds for X and Y. Therefore, the pmf of the product $Z = X \cdot Y$ is given by (A.23)

$$p_Z(z) = \sum_{x \in \mathbb{X}} \frac{1}{|x|} p_X(x) p_Y\left(\frac{z}{x}\right) \ . \tag{A.32}$$

Sadly, this sum can not be solved analytically. Thus, the derived mass $p_{L_Z}(l_z)$ from $p_Z(z)$, given the variable transformation $L_Z = 2 \cdot \tanh^{-1}(Z)$, is also not expressed in an analytically way. Still, the shorthand notation $p_{L_Z}(l_z) = \mathcal{F}^{\boxplus}\left(p_{L_X}(l_x), p_{L_Y}(l_y)\right)$ is used for the numerical evaluation of the boxplus density.

To overcome this limitation, one can utilize the approximation of the *boxplus* operation, which is given by [Hag94]

$$\begin{aligned} L_Z &= L_X \boxplus L_Y \\ &\approx \operatorname{sign}(L_X) \cdot \operatorname{sign}(L_Y) \cdot \min(|L_X|, |L_Y|) \ . \end{aligned} \tag{A.33}$$

The pmf for the discrete random variable L_Z is obtained by [WW96][Ana01]

$$p_{L_Z}(l_z) = \begin{cases} p_{L_X}(l_z)\left[1 - F_{L_Y}(l_z)\right] + p_{L_Y}(l_z)\left[1 - F_{L_X}(l_z)\right] & \\ \quad + p_{L_X}(-l_z) \cdot F_{L_Y}(-l_z) + p_{L_Y}(-l_z) \cdot F_{L_X}(-l_z) & \text{for } l_z \geq 0 \\ p_{L_X}(l_z)\left[1 - F_{L_Y}(-l_z)\right] + p_{L_Y}(l_z)\left[1 - F_{L_X}(-l_z)\right] & \\ \quad + p_{L_X}(-l_z) \cdot F_{L_Y}(l_z) + p_{L_Y}(-l_z) \cdot F_{L_X}(l_z) & \text{for } l_z < 0 \end{cases} , \tag{A.34}$$

with the shorthand notation $p_{L_Z}(l_z) = \mathcal{F}^{\widetilde{\boxplus}}\left(p_{L_X}(l_x), p_{L_Y}(l_y)\right)$.

A.5 Log-likelihood ratios and soft bits

Let A be an event and $\Pr(A)$ the probability that the event A occurs. Then, the *log-likelihood ratio* (LLR) of this event is defined as [SL05, p. 105]

$$L(A) = \log\left(\frac{\Pr(A)}{1 - \Pr(A)}\right) \ , \tag{A.35}$$

i.e. the logarithmic ratio between the probability that the event A occurs and the probability that event A did not happen.

Given a random variable X with the binary set of possible outcomes $x \in \{-1, +1\}$, then the LLR for $L(x)$ is calculated as [Hag94]

$$l = L(x) = \log\left(\frac{\Pr(X = +1)}{\Pr(X = -1)}\right) = \log\left(\frac{p_X(x = +1)}{p_X(x = -1)}\right) \quad . \tag{A.36}$$

Taking the sign of (A.36), i.e. $\text{sign}(L(x))$ corresponds to the hard decision, whereas the absolute value $|L(x)|$ delivers the reliability of that decision.

In many cases, especially in wireless communication, one is often interested in the probability of a transmitted symbol, denoted by the rv X, conditioned on a received symbol, denoted by the rv Y. Here, the input rv X, $x \in \{-1, +1\}$ is supposed to be binary, whereas the output Y, $y \in \mathbb{R}$ is often a discretized random variable. The so called a-posteriori probability with corresponding LLR is now calculated via the conditional probability mass function

$$l = L(x|y) = \log\left(\frac{p_{X|Y}(x = +1|y)}{p_{X|Y}(x = -1|y)}\right) \quad , \tag{A.37}$$

which includes the extrinsic and the a-priori LLR values

$$\begin{aligned} &= \log\left(\frac{p_X(x = +1) \cdot p_{X|Y}(x = +1|y)}{p_X(x = -1) \cdot p_{X|Y}(x = -1|y)}\right) \\ &= \log\left(\frac{p_X(x = +1)}{p_X(x = -1)}\right) + \log\left(\frac{p_{Y|X}(y|x = +1)}{p_{Y|X}(y|x = -1)}\right) \quad . \end{aligned} \tag{A.38}$$

Under the assumption of i.i.d. BPSK modulated transmit symbols, the a-posteriori log-likelihood ratios for an BIAWGN channel with

$$Y = X + N \tag{A.39}$$

and $N \sim \mathcal{N}(0, \sigma_N^2)$, is only depending on the extrinsic log-likelihood ratio. Given the channel transition probability

$$p_{Y|X}(y|x = \pm 1) = \frac{1}{\sqrt{2\pi\sigma_N^2}} e^{-\frac{(y \mp 1)^2}{2\sigma_N^2}} \quad , \tag{A.40}$$

the second term in (A.38) results to

$$l = L(y|x) = \log\left(\frac{p_{Y|X}(y|x=+1)}{p_{Y|X}(y|x=-1)}\right) = \log\left(\frac{\frac{1}{\sqrt{2\pi\sigma_N^2}}e^{-\frac{(y-1)^2}{2\sigma_N^2}}}{\frac{1}{\sqrt{2\pi\sigma_N^2}}e^{-\frac{(y+1)^2}{2\sigma_N^2}}}\right)$$
$$= \frac{2}{\sigma_N^2}\cdot y \ . \tag{A.41}$$

In case of a multiplicative fading channel, the log-likelihood ratios for the channel output $Y = A \cdot X + N$, with an attenuation factor A, $a \in \mathbb{R}$, is obtained by

$$l = L(y|x,a) = \frac{2}{\sigma_N^2}\cdot a \cdot y \ . \tag{A.42}$$

Please note, that time indices have been omitted in the preceding paragraph and throughout this dissertation, parentheses will be omitted, i.e. $l = L(x)$.

Nomenclature

B

AF	Amplify-and-Forward
AlF	Alpha Function
ARQ	Automatic Repeat Request
AWGN	Additive White Gaussian Noise
BCJR	Bahl, Coke, Jelinek, Raviv
BEC	Binary Erasure Channel
BER	Bit Error Rate
BIAWGN	Binary Input Additive White Gaussian Noise
BISMC	Binary Input Symmetric Memoryless Channel
BPSK	Binary Phase Shift Keying
BSC	Binary Symmetric Channel
CC	Convolutional Code
CF	Compress-and-Forward
DF	Decode-and-Forward
DFIR	Decode-and-Forward with Incremental Redundancy
DFRC	Decode-and-Forward with Repetition Coding
EXIT	Extrinsic Information Transfer
FEC	Forward Error Correction
GIAWGN	Gaussian Input Additive White Gaussian Noise
GISMC	Discrete Symmetric Memoryless Channel

GISMC	Gaussian Input Symmetric Memoryless Channel
HIHO	Hard-Input Hard-Output
ICT	Information Combining Technique
IPC	Information Processing Characteristic
ITC	Information Transfer Characteristic
LDPC	Low-Density Parity-Check
LF	Limiter Function
LLR	Log-Likelihood Ratio
LTE	Long Term Evolution
MAP	Maximum-a-Posteriori
MFMC	Max-Flow Min-Cut
MHT	Multi-Hop Transmission
MI	Mutual Information
MIMO	Multiple Input Multiple Output
ML	Maximum-Likelihood
MLSE	Maximum-Likelihood Sequence Estimator
MRC	Maximum Ratio Combining
NSC	Non-Systematic Convolutional Code
OFDM	Orthogonal Frequency Division Multiplexing
OSI	Open Systems Interconnection
PDF	Probability Density Function
PICF	Parallel Information Combining Function
PMF	Probability Mass Function
QoS	Quality of Service

RAC	Repeat-Accumulate Code
RSC	Recursive Systematic Convolutional Code
SICF	Serial Information Combining Function
SISO	Soft-Input Soft-Output
SMC	Symmetric Memoryless Channel
SNR	Signal-to-Noise Ratio
SPC	Single Parity-Check
TC	Turbo Code
UMTS	Universal Mobile Telecommunications System
WiMAX	Worldwide interoperability for Microwave Access
WLAN	Wireless Local Area Network

C Notation

Throughout this dissertation, the following symbols and notations are used:

Mathematical symbols and expressions	
a, A	Scalars, i.e. variables for a real or complex value
$\mathbf{a} = \left[a_m\right]_{M\times 1}$	Column vector of size M, i.e. $m = 1, \ldots, M$ rows and one column
$\mathbf{A} = \left[A_{m,n}\right]_{M\times N}$	Matrix of size $M \times N$, i.e. $m = 1, \ldots, M$ rows and $n = 1, \ldots, N$ columns
$\mathbf{a}^T, \mathbf{A}^T$	Transpose of a vector or matrix
$\mathbf{A}^H$	Herimitian or conjugate transpose of a matrix
$\mathbb{A} = \{\ldots\}$	Set, e.g. $\mathbb{R}$ as the set of real numbers
$\lvert\mathbb{A}\rvert$	Cardinality or size of the set $\mathbb{A}$

Entropy, mutual information and capacity	
$H(X)$	Entropy or Shannon entropy of X, measured in bit
$H(X,Y)$	Joint entropy between X and Y
$H(X\|Y)$	Conditional entropy between X and Y
$H_b(\epsilon)$	Binary entropy function with crossover prob. ϵ
$I(X;Y)$	Mutual information between X and Y
$I(X;Y\|Z)$	Conditional mutual information between X and Y, conditioned on Z
C	Capacity of a channel
C_{BSC}	Capacity of a binary symmetric channel
C_{BEC}	Capacity of a binary erasure channel
C_{BIAWGN}	Capacity of a binary input additive white Gaussian noise channel
C_{GIAWGN}	Capacity of a Gaussian input additive white Gaussian noise channel
R	Rate or throughput
$R_c = K/N$	Code rate for code with K information symbols and N correspodning code symbols

MIs for IPC and ICT	
$\text{IPC}^{\text{ideal}}(C)$	IPC for ideal coding scheme
$\text{IPC}^{\text{good}}(C)$	IPC for coding scheme with good codes
$\text{IPC}^{\text{info}}_{\text{SD}}(C)$	Measured IPC for soft inf. symbols
$\text{IPC}^{\text{info}}_{\text{HD}}(C)$	Measured IPC for hard inf. symbols
$\text{IPC}^{\text{extr}}(C)$	Measured IPC for extrinsic symbols
$\text{IPC}^{\text{info}}_{\text{ext}}(C)$	Measured extended IPC
$\mathcal{I}^{\text{ser}}$	Serial information combining function
$\mathcal{I}^{\text{ser,up}}$	Serial information combining function, upper bound
$\mathcal{I}^{\text{ser,low}}$	Serial information combining function, lower bound
$\mathcal{I}^{\text{par}}$	Parallel information combining function
$\mathcal{I}^{\text{par,up}}$	Parallel information combining function, upper bound
$\mathcal{I}^{\text{par,low}}$	Parallel information combining function, lower bound

Notations for nodes, channel in- and outputs and channel parameters	
$\mathcal{S}$	Source
$\mathcal{R}_m$	Relay m, $m \in \{1,\ldots,M\}$
$\mathcal{D}$	Destination
$\mathbf{x}_T$	Transmit sequence of the transmitter, $T \in \{\mathcal{S},\mathcal{R}\}$
$\mathbf{y}_{TR}$	Receive sequence or channel output of the transmitter T and receiver R link, $R \in \{\mathcal{R},\mathcal{D}\}$
γ_{TR}	Signal-to-noise ratio
$\sigma_N^2,\ \sigma_X^2$	Noise or signal variance
$a_{TR} = \sqrt{d_{TR}^{-\alpha}}$	Path-loss attenuation factor a_{TR} with distance d_{TR} and path-loss exponent α
$T_1,\ T_2$	Time slot indicator with normalized time slots lengths $t_1 + t_2 = 1$

Probability		
Ω		Sample space (set of all possible outcomes or events ω of the random experiment)
ω		Event (subset of Ω)
X	r.v.	Random variable
$\mathbb{X}$		Alphabet of X
$X : \Omega \to \mathbb{X}$		Mapping from the sample space to the alphabet
$\omega \to X(\omega)$		Function, that the r.v. takes on value $X(\omega)$, if event ω occurs
X		Stochastic process, with $\mathrm{X} = \{X_l : l = 1, \ldots, L\}$ over a collection of L random variables X
$\Pr(X)$		probability of X
$\Pr(X = x)$		probability of the event that X takes on the value x, i.e. $\{\omega : X(\omega) = x\}$ or simple $\{X = x\}$
$p_X(x)$	pmf	Probability mass function, with $\sum_{x \in \mathbb{X}} p_X(x) = 1$ or
	pdf	Probability density function, with $\int_{-\infty}^{+\infty} f_X(x)dx = 1$
$F_X(x)$	cdf	Cumulative distribution function with $F_X(x) = \Pr(X \leq x) = \begin{cases} \sum_{k \leq x} p_X(k) & \text{if } X \text{ is discrete} \\ \int_{-\infty}^{x} p_X(t)dt & \text{if } X \text{ is continuous} \end{cases}$
$p_{X,Y}(x,y)$		Joint probability mass function, with $p_{XY}(x,y) = \Pr(X = x, Y = y)$
$p_{X\|Y}(x\|y)$		Conditional probability mass function, with $p_{X\|Y}(x\|y) = \Pr(X = x\|Y = y)$
ϵ		Crossover or erasure probability with either $\epsilon = \Pr(y = \pm 1 \| x = \mp 1)$ or $\epsilon = \Pr(y = 0 \| x = \pm 1)$

Shorthand notations	
$p_Z(z) = \mathcal{F}^{+}\left(p_X(x), p_Y(y)\right)$	Pdf or pmf of $p_Z(z)$ for $Z = X + Y$
$p_Z(z) = \mathcal{F}^{\times}\left(p_X(x), p_Y(y)\right)$	Pdf or pmf of $p_Z(z)$ for $Z = X \cdot Y$
$p_Y(y) = \mathcal{F}^{\trianglerighteq}\left(p_X(x), a, b\right)$	Pdf or pmf of $p_Y(y)$ for $Y = aX + b$
$p_Y(y) = \mathcal{F}^{\triangleright}\left(p_X(x), a\right)$	Pdf or pmf of $p_Y(y)$ for $Y = aX$
$p_Y(y) = \mathcal{F}^{\curlywedge}\left(p_X(x)\right)$	Pdf or pmf of $p_Y(y)$ for $Y = \tanh(X/2)$
$p_Y(y) = \mathcal{F}^{\lvert X\rvert}\left(p_X(x)\right)$	Pdf or pmf of $p_Y(y)$ for $Y = \lvert X\rvert$
$p_{U_Z}(u_z) = \mathcal{F}^{\oplus}\left(p_{U_X}(u_x), p_{U_Y}(u_y)\right)$	Pdf or pmf of $p_{U_Z}(u_z)$ for $U_Z = U_X \oplus U_Y$
$p_{L_Z}(l_z) = \mathcal{F}^{\boxplus}\left(p_{L_X}(l_x), p_{L_Y}(l_y)\right)$	Pdf or pmf of $p_{L_Z}(l_Z)$ for $L_Z = L_X \boxplus L_Y$
$p_{L_Z}(l_z) = \mathcal{F}^{\widetilde{\boxplus}}\left(p_{L_X}(l_x), p_{L_Y}(l_y)\right)$	Pdf or pmf of $p_{L_Z}(l_Z)$ for $L_Z = \operatorname{sign}(L_X) \cdot \operatorname{sign}(L_Y) \cdot \min\left(\lvert L_X\rvert, \lvert L_Y\rvert\right)$

Theses

D

1. Rate prediction for a three-node relay network with the information combining technique is possible.

2. Orthogonal relay networks with error-prone relays can be semi-analytically predicted by means of mutual information.

3. The information combining technique for relay networks will reduce the computational complexity in predicting the mutual information and/or bit error rate of such networks.

4. Soft re-encoding will not always achieve better bit error rate performance in a relay network compared to hard re-encoding, depending on the combining/decoding strategies.

5. Channel log-likelihood ratios at the destination's decoder input can be calculated.

6. Maximum ratio combining of the information symbols at the destination can increase the mutual information.

7. Advantages of maximum ratio combining or turbo decoding at the destination depend on the position and, therewith, the channel qualities of the receiving nodes.

Lebenslauf

Sebastian Vorköper

Beruf

07/2016– **Polizeivollzugsbeamter, Abteilung 2, Dezernat 24 "Besondere Dienste"**, *Landeskriminalamt Mecklenburg-Vorpommern.*

08/2015–06/2016 **Sachbearbeiter, Abteilung 2, Dezernat 24 "Besondere Dienste"**, *Landeskriminalamt Mecklenburg-Vorpommern.*

05/2015–07/2015 **Projektingenieur**, *SII Deutschland GmbH.*

03/2014–12/2014 **Wissenschaftlicher Mitarbeiter am Institut für Nachrichtentechnik, Forschungsgruppe Nachrichtenübertragungstechnik**, *Universität Rostock.*

08/2013–12/2013 **Forschungsingenieur am Schifffahrtsinstitut Warnemünde e.V.**, *Institut an der Hochschule Wismar.*

07/2007–07/2013 **Wissenschaftlicher Mitarbeiter am Institut für Nachrichtentechnik, Forschungsgruppe Nachrichtenübertragungstechnik**, *Universität Rostock.*

Studium, Schule

10/2001–05/2007 **Diplom-Studium im Studiengang Informationstechnik / technische Informatik**, *Universität Rostock*, Gesamtnote: "sehr gut".

11/2005 **Bachelor im Studiengang Informationstechnik / technische Informatik**, *Universität Rostock*, Gesamtnote: "gut".

08/2000–06/2001 **Zivildienst im Bereich Maschinen- und Schiffswartung**, *"MS Stubnitz"*, Rostock.

09/1997–07/2000 **Fachabitur in Elektrotechnik**, *Berufliche Schule für Technik und Handwerk*, Stralsund, Note: 2,0.

08/1996–07/1997 **Schüleraustausch in den USA (Central High School)**, *La Crosse*, Wisconsin, USA.

09/1991–08/1996 **Besuch des Goethe-Gymnasiums**, Stralsund.

09/1986–08/1991 **Besuch der Grundschule Andershof**, Stralsund.

Familie

Familienstand:	Verheiratet
Kinder:	Jessica Anazie, geb. am 15.01.1999
	Finn Vorköper, geb. am 16.09.2009
	Hanna Vorköper, geb. am 14.10.2011

Projekte und persönliches Engagement

- Ständiges Mitglied der Linux- und Unix-Usergroup Rostock
- Software- und Hardwareprojekte, wie z.B. dem *GNOME Photo Printer*, *plot2gp* oder dem Aufbau von Mikrocontroller- bzw. FPGA-Boards inklusive Layouten, Ätzen und Bestücken
- Interesse am Amateurfunk (SWL) und der Elektrotechnik
- Gutachter für die *IEEE Transactions on Communications*, *European Transactions on Telecommunications*, u.a.

Computerkenntnisse

Betriebssysteme:	Linux/Unix, Windows, Mac OSX u.a.
Progr.-sprachen:	C/C++, Matlab/Simulink, Python, Javascript, VHDL, u.a.
Datenbanken:	SQL
Hardware:	Aufbau und Wartung von Rechner- und Netzwerkanlagen
Textsatz:	LaTeX, Word, PowerPoint, OpenOffice, u.a.

Sprachkenntnisse

Deutsch:	Muttersprache
Englisch:	Fließend in Wort und Schrift,
	Verfassen von wissenschaftlichen Publikationen
Französisch:	Schulkenntnisse

Betreute Diplom-/Masterarbeiten

- D. Kern: *Implementierung und Evaluierung von Polar Codes.* Universität Rostock 2012
- M. Rogge: *Implementierung eines drahtlosen MIMO-OFDM-Systems auf einer FPGA/DSP-Hardware.* Universität Rostock 2011
- M. Schultz: *Implementierung von Detektionsalgorithmen für ein V-BLAST System.* Universität Rostock 2011
- R. Al-Salim: *Untersuchung des Trägerfrequenzversatzes in OFDM Relay-Netzwerken.* Universität Rostock 2010
- S. Simanowski: *Implementierung eines VBLAST-Systems auf einer MIMO-Demonstrationshardware.* Universität Rostock 2010
- M-K. Ali: *Implementation of an OFDM-Transceiver in Matlab/Simulink for the Lyrtech MIMO-ADS.* Universität Rostock 2009
- R. Amling: *Analyse der Leistungsfähigkeit von OFDM in Relay-Netzwerken.* Universität Rostock 2008

Betreute Bachelorarbeiten

- P. Bartmann: *Implementierung eines Algorithmus zur Routenfindung in vermaschten Netzen bei Minimierung der Energie.* Universität Rostock 2011
- D. Kern: *Implementierung eines Algorithmus zur Routenfindung in vermaschten Netzen bei Maximierung des Datendurchsatzes.* Universität Rostock 2011
- R. Yasin: *Implementierung eines AM-Senders.* Universität Rostock 2011
- M. Sabaana: *Implementierung eines AM-Empfängers.* Universität Rostock 2011

Vorträge im Bereich "Angewandte Informationstheorie"

- **S. Vorköper**, M. Woltering: *Hardware implementation of a MIMO-OFDM System.* 19. ITG-Fachgruppe 2012, Bremen
- **S. Vorköper**: *Analysis of Cooperative Coding Relay Networks.* 16. ITG-Fachgruppe 2010, Rostock
- **S. Vorköper**: *Cyclic Delay Diversity in OFDM-based Relay Networks.* 13. ITG-Fachgruppe 2009, Rostock
- **S. Vorköper**, V. Kühn: *Application of Information Combining to Relay Networks.* 12. ITG-Fachgruppe 2008, Berlin

Wissenschaftliche Publikationen

- D. Kern, **S. Vorköper**: *A New Code Construction for Polar Codes Using Min-Sum Density.* 8th International Symposium on Turbo Codes and Iterative Information Processing (ISTC'14), Bremen, Deutschland.
- **S. Vorköper**, F. Lenkeit, V. Kühn, D. Wübben, A. Dekorsy: *Distributed IDM-STC versus Cooperative OFDM for the Two-Hop Decode-and-Forward Multiple-Access Relay Channel.* International ITG Workshop on Smart Antennas (WSA'13), Dresden, Deutschland.
- **S. Vorköper**, V. Kühn: *Power and Subcarrier Allocation Scheme for Uplink OFDMA.* 17th International OFDM Workshop 2012 (InOWo'12), Essen, Deutschland.
- **S. Vorköper**, V. Kühn: *Performance Comparison of Distributed IDM-STC versus Cooperative OFDM for practical Decode-and-Forward Relay-Networks.* International ITG Workshop on Smart Antennas (WSA'12), Dresden, Deutschland.
- M. Hinkfoth, E. Heinrich, **S. Vorköper**, V. Kühn, R. Salomon: *X-ORCA: FPGA-based wireless localization in the sub-millimeter range.* 20th ACM/SIGDA International Symposium on Field-Programmable Gate Arrays (FPGA'12), Monterey, CA, USA.
- **S. Vorköper**, V. Kühn: *CFO Estimation and Compensation in LTE Compliant DF Relay Networks.* 16th International OFDM Workshop 2011 (InOWo'11), Hamburg, Deutschland.
- **S. Vorköper**, V. Kühn: *Analyzing the Performance of Error-Prone Relay Networks.* IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP'11), Prag, Tschechische Republik.

- **S. Vorköper**, V. Kühn: *Error-Prone Relay Networks with Soft- and Hard-Decision Re-Encoding.* International ITG Workshop on Smart Antennas (WSA'11), Aachen, Deutschland
- **S. Vorköper**, V. Kühn: *Noise Variance Estimation in OFDM Cyclic Delay Diversity Relay Networks.* 14th International OFDM Workshop 2009 (InOWo'09), Hamburg, Deutschland.
- P. Weitkemper, **S. Vorköper**, D. Wübben, K.-D. Kammeyer, V. Kühn: *Distributed IDM-STC versus OFDM-CDD in Two-Hop Relay-Networks.* 14th International OFDM Workshop 2009 (InOWo'09), Hamburg, Deutschland.
- A. Khan, **S. Vorköper**, V. Kühn: *Information Combining Based ARQ in Relay Networks.* 15th European Wireless Conference (EW 2009), Aalborg, Dänemark.
- **S. Vorköper**, V. Kühn: *Reducing the Channel Estimation Error for Relay Networks with Cyclic Delay Diversity.* 15th European Wireless Conference (EW 2009), Aalborg, Dänemark.
- V. Kühn, **S. Vorköper**: *Information Combining for Relay Networks.* IEEE International Conference on Communications (ICC 2009), Dresden, Deutschland.
- V. Kühn, **S. Vorköper**: *Application of Information Combining to Relay Networks.* IEEE Vehicular Technology Conference (VTC 2009 Spring), Barcelona, Spanien.
- **S. Vorköper**, V. Kühn: *Information Combining for OFDM-based Relay Networks.* 13th International OFDM Workshop 2008 (InOWo'08), Hamburg, Deutschland.

Statement of Authentication

I hereby certify that I have written the submitted dissertation independently, without assistance from external parties and without use of other resources than those indicated. Ideas taken directly or indirectly from other resources are acknowledged in the text.

This dissertation was neither in full, nor in similar form, or in part previously submitted for grading at any academic institution.

Signed:

December 26, 2016, Sebastian Vorköper

Eigenständigkeitserklärung

Hiermit erkläre ich, dass ich die eingereichte Dissertation selbständig und ohne fremde Hilfe verfasst, andere als die von mir angegebenen Quellen und Hilfsmittel nicht benutzt und die den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Diese Dissertation wurde weder in gleicher, noch ähnlicher Form, noch auszugsweise im Rahmen einer anderen Prüfung vorgelegt.

Unterschrift:

December 26, 2016, Sebastian Vorköper