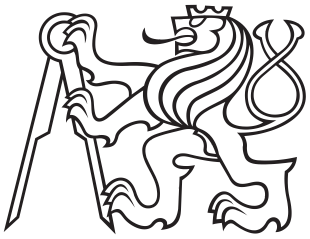
CZECH TECHNICAL UNIVERSITY IN PRAGUE

Faculty of Electrical Engineering

Department of Radioelectronics



Signal Processing in Wireless Physical Layer Coding Radio Networks with Weak System State Information

Doctoral Thesis

Ing. Petr Hron

Prague, 2023

Ph.D. Programme: **Electrical Engineering and Communications (P0714D060003)**
Field of study: **Radioelectronics**

Supervisor: **Prof. Ing. Jan Sýkora, CSc.**

Acknowledgements

I would like to thank my supervisor Doc. Ing. Jan Sýkora, CSc. for giving me the chance to perform research in digital communication theory under his valuable supervision.

Declaration

I hereby declare that this doctoral thesis *Signal Processing in Wireless Physical Layer Coding Radio Networks with Weak System State Information* is a result of my research activities carried out at the Faculty of Electrical Engineering at Czech Technical University in Prague based on the collaboration with my supervisor and other team members. It has not been previously included in a thesis or dissertation submitted to this or any other institution for a degree, diploma, or other qualifications. Information derived from published or unpublished work of others has been acknowledged in the text and a list of references is given in the bibliography.

Abstract

The growing demand for communication and connectivity within finite signal spectrum resources drives the need for novel and more efficient techniques of wireless network communications. One of the promising answers reflecting those requirements is the so-called WPNC paradigm which enables a more efficient way of exploiting available resources. The main difference between WPNC and classical multiuser techniques, used in today's wireless standards of cellular networks or wireless LAN for instance, is the perception of interferences. An interference is generally any signal which has disruptive effects on the reception. In dense networks, interferences formed by transmissions among other users are inevitable and account for one of the principal performance limitations. Plenty of methods to avoid interference exist. The majority of traditional methods reside on some form of signal space orthogonalization, which is wasteful w.r.t. the available resources. WPNC on the other hand, is designed in a particular such that it is possible to make use of the additional information carried by interference and to improve the overall network capacity and efficiency.

Because the WPNC concept does not separate signals from different users employing orthogonal division, one of its main components is a so-called hierarchical multiple access channel (H-MAC) formed between multiple source nodes and a relaying node. In an H-MAC, several source signals are allowed to interact, superimpose the receiver's antenna and form a joint constellation. In a dense network with numerous multi-hop paths, it is not always necessary to demodulate each source's message at the relay. Usually, it is sufficient to recover a many-to-one function that gets propagated further through the network. Those specifics render the H-MAC very different from a classical single-user channel and bring several new challenges. One important aspect is the channel parametrization in terms of attenuation, channel phase, and temporal alignment of the individual signals w.r.t. each other.

In this thesis, we focus on the H-MAC with weak system information. Most of the published research results on the properties of WPNC got derived under idealized assumptions about the system state knowledge, which is difficult or even impossible to assure in a real-world application. We analyze the situations in which the system knowledge is only partial or absent and develop techniques to deal with such a situation. In particular, we consider a transmission under unknown relative phase parametrization common time-base absence.

We give a detailed description and analysis of the considered problem together with a proposed solution in terms of theoretical derivation as well as a final algorithm. Where possible, we evaluate and verify the proposed solution in terms of a simulation and/or an over-the-air experiment and present the obtained results.

Keywords: Wireless physical layer network coding, Hierarchical MAC parametrization, asynchronous NOMA

Abstrakt

Stále rostoucí poptávka po bezdrátové komunikaci a konektivitě naráží na limity konečné šířky použitelného spektra čímž vzniká tlak na výzkum nových a efektivnějších způsobů komunikace. Jedním z možných řešení je takzvaný koncept WPNC, který umožňuj lepší využití dostupných prostředků. Hlavním rozdílem, kterým se koncept WPNC liší od klasických přístupů k mnohouživatelským komunikacím (využitých v dnešních buňkových sítích, wireless LAN apod.) je jeho přístup k interferenci. Jako interferenci označujeme jakýkoli signál, který má rušivý vliv na kvalitu příjmu. V hustých radiových sítích je interference tvořena přenosy mezi ostatními uživateli prakticky nevyhnutelný jev a představuje jeden z hlavních omezujících faktorů. Existuje mnoho postupů jak odstranit negativní dopad interference, většinou se jedná o ortogonální odděleni interference a užitečného signálu. Tyto způsoby vliv interference velmi efektivně potlačují, avšak nedovolují optimální využití přidělených prostředků. V případě konceptu WPNC je síť navržena takovým způsobem, že lze informaci nesenou interferencí využít ke zlepšení kapacity a efektivity přenosů v rámci celé sítě.

V síti s WPNC nejsou jednotlivé signály ortogonálně separovány a dochází tak k situacím, kdy je relay uzlem pozorována superpozice několika dílčích signálů. Takový kanál mezi několika zdroji a jedním přijímačem označujeme jako hierarchický MAC (H-MAC). V husté síti existuje zpravidla několik nezávislých cest (přes několik relay uzlů) kudy může informace mezi zdrojem a destinací putovat. Není tak vždy nutné aby byl relay uzel demoduloval všechny dílčí zprávy ale typicky stačí pokud demoduluje a přepošle nějakou many-to-one funkci jednotlivých zpráv. Právě H-MAC představuje zásadní odlišnost od klasických přístupů s ortogonálně separovatelnými kanály a přináší mnoho problémů specifických pro WPNC. Jeden z hlavních aspektů je relativní parametrizace jako je např. útlum, fáze kanálu nebo časová synchronizace mezi jednotlivými zdroji a přijímačem.

V této disertační práci se zaměřujeme na H-MAC se slabou znalostí systému. Většina publikovaných výsledků na téma sítí s WPNC vyžaduje idealizované předpoklady ohledně znalosti sytému v jednotlivých uzlech sítě. V praktických scénářích jsou takové předpoklady velmi obtížně splnitelné. Proto se zaměřujeme na analýzu situací kdy je stav systém znám pouze částečně či vůbec a navrhujeme postupy jak takové situace řešit. Konkrétně se jedná o přenos v H-MAC při neznámé relativní parametrizaci fáze nebo chybějící časové synchronizaci.

Výstupy prezentované v této práci obsahují detailní analýzu uvažovaného problému a navrhované řešení v podobě teoretického odvození a výsledného algoritmu. Navržené řešení je podrobeno verifikaci v podobě numerické simulace a/nebo reálného OTA přenosu.

Klíčová slova: Bezdrátové síťové kódování na fyzické vrstvě, Parametrizace hierarchického MAC, asynchronní NOMA

Contents

[1] All authors have contributed equally.

List of Figures

List of Tables

Chapter 1

Preface

1.1 Thesis Outline

1.1.1 Thesis Goals

This thesis "Signal Processing in Wireless Physical Layer Coding Radio Networks with Weak System State Information" summarizes my up-to-date research activities as a Ph.D. candidate at the Faculty of Electrical Engineering, at the Czech Technical University in Prague.

The research is motivated by wireless cloud networks with a high number of cooperating nodes that have only limited information about their radio environment. This includes the channel state, the processing/coding strategies of the other nodes as well as the overall connectivity. Each node is assumed to have only weak information about the system and the uncertainty is treated as an additional degree of freedom, fully integrated with the actual payload data. This poses a broad topic with a great number of new challenges.

The specific problems and contributions can be listed as follows:

- A solution to the problem of relative channel parametrization of an H-MAC without the need for additional orthogonal resources such as separate pilot or training sequences.
- Analysis and solution of imperfect synchronization of source nodes in an H-MAC stage. In particular, the design of modulation and coding methods, that would be resistant to mutual delays between individual source signals.

The studied problems are rigorously defined and mathematically described, analyzed, and solved in terms of proposed methods and algorithms. The obtained solutions are verified by a numerical simulation or an over-the-air experiment.

1.1.2 Thesis Contents

The rest of this thesis is organized as follows: The rest of this chapter contains an overview of publications and grants I participated in, and which made this work possible. The publications are divided into categories based on the publication type (journal or conference) and relation to the contents of this thesis. The rest of the thesis is divided into two parts.

Part I contains a general introduction to wireless communications and overviews the current research trends (Chapter 2). Chapter 3 is devoted to wireless multiuser communication methods, with a focus on the concept of WPNC as the dominant topic of this work.

The second Part (Contribution) is dedicated to the presentation of the results of our research. It is divided into three chapters based on the topic. The first chapter (Chapter 3.3.3: *Hierarchical*

Decision Aided H-MAC CSE) builds directly on the results of the master's thesis and contains the results of a hardware verification of the proposed hierarchical CSE. Chapter 5 (*Differential Hierarchical Demodulator of 2-source Hierarchical-MAC Channel*) presents our work on the application of a differential modulation in the context of an H-MAC and can be understood as an alternative to the hierarchical CSE. We propose several variants of demodulation and evaluate their performance utilizing computer simulations as well as an OTA evaluation. In the third chapter (Chapter 6: *Fractional Symbol Duration Asynchronous H-MAC*) we treat the topic of temporal asynchronicity in an H-MAC channel. We propose an OFDM-like approach and several receiver-side processing variants (symbol-spaced as well as fractionally-spaced processing). The proposed solutions get compared in terms of computer simulations. The thesis is concluded in Chapter 7, where we summarize our achievements and indicate viable future research directions.

1.2 Publications[1]

1.2.1 Publications Related to this Thesis

A list of publications divided into two parts based on the relation to the content of this thesis follows. Each part is subdivided into sections containing publications in journals with impact factor, conference publications indexed by Scopus, and other conference publications. The list reflects the state at the time of writing this thesis.

Publications in Journals with Impact Factor

- P. Hron, J. Sýkora, and L. Clavier. "Asynchronous WPNC with DFT Precoding and Symbol-Spaced Processing" *IEEE Communications Letters* **under review**.
- P. Hron and J. Sýkora. "Closed-Form Differential Hierarchical Demodulator of 2-source Hierarchical-MAC Channel" *IEEE Communications Letters* 25.11 (2021): 3704-3708.
- P. Hron, J. Lukáč, and J. Sýkora. "SDR Verification of Hierarchical Decision Aided 2-Source BPSK H-MAC CSE with Feed-Back Gradient Solver for WPNC Networks" *Radioengineering* 29.3 (2020): 573.

Conference Publications Indexed by Scopus

- P. Hron and J. Sykora. "Performance analysis of hierarchical decision aided 2-source BPSK H-MAC CSE with feed-back gradient solver for WPNC networks" *2019 IEEE Microwave Theory and Techniques in Wireless Communications (MTTW)*. Vol. 1. IEEE, 2019.

Other Conference Publications

- P. Hron, J. Sýkora and L. Clavier. "Linear precoding for asynchronous WPNC 2-source hierarchical multiple access channel" *4th MC and 4th Technical Meeting*. COST ACTION CA20120 INTERACT, 2023.
- P. Hron, J. Lukáč and J. Sýkora. "Noncoherent WPNC 2-source MAC channel demodulation methods" *2nd MC Meeting and 1st Technical Meeting*. COST ACTION CA20120 INTERACT, 2022.
- P. Hron, J. Lukáč and J. Sýkora. "SDR verification of noncoherent WPNC 2-source MAC channel reception" *3rd Post-IRACON Meeting*. COST ACTION CA15104 IRACON, 2021

[1] All authors have contributed equally.

- P. Hron and J. Sýkora. "Performance analysis of hierarchical decision aided 2-source BPSK H-MAC CSE with feed-back gradient solver for WPNC networks" *10th MC meeting and 10th Technical meeting*. COST ACTION CA15104 IRACON, 2019.

1.2.2 Publications Unrelated to this Thesis

Publications in Journals with Impact Factor

- J. Kolář, J. Sýkora and P. Hron. "Update-Based Machine Learning Classification of Hierarchical Symbols in a Slowly Varying Two-Way Relay Channel" *Mathematics* 8.11 (2020): 2007.

Conference Publications Indexed by Scopus

- J. Sýkora and P. Hron. "Achievable rates for HDF WPNC strategy with hierarchical bit-wise network coding maps for higher-order constellations in H-MAC channel with relative fading" *2018 European Conference on Networks and Communications* (EuCNC). IEEE, 2018.

1.3 Grants

The work presented in this thesis was supported by the following research projects and grants in which I participated.

Local Projects

- Ministry of Education, Youth and Sports (LTC17042): Cooperative Coding and Processing in Dense Radio Cloud Communication Networks, 2017-2020
- Czech Technical University in Prague (SGS22/063/OHK3/1T/13): Modulation and Physical Layer Distributed Algorithms for Dense Radio Networks, 2022
- Czech Technical University in Prague (SGS21/062/OHK3/1T/13): Modulation, Coding and Distributed Signal Processing Algorithms for Dense Radio Networks, 2021
- Czech Technical University in Prague (SGS20/068/OHK3/1T/13): Dense radio networks with wireless physical layer network coding, machine learning and distributed processing, 2020
- Czech Technical University in Prague (SGS19/069/OHK3/1T/13): Coding and Distributed Signal Processing in Radio Networks with Physical Layer Network Coding, 2019

International Projects

- EU COST (CA20120): Intelligence-Enabling Radio Communications for Seamless Inclusive Interactions (INTERACT), 2021-2025
- EU COST (CA15104): Inclusive Radio Communication Networks for 5G and Beyond (IRACON), 2016-2020

Part I

Wireless Physical Layer Network Coding Fundamentals:

Part I of this thesis contains an introduction to the research field of wireless communication covering the historical development and current state-of-the-art with a focus on multiuser access and WPNC. The introduction is intended for a broad audience and is written in a rather informal way. Definitions and principles which are necessary for an understanding of the contributions in part II are explained in terms of simple examples and accompanied by recommended references.

Chapter 2

Introduction

2.1 Historical Background

The idea to use electromagnetic phenomenons for fast communication dates back to the first electric telegraph by Samuel F. B. Morse patented in 1847 and the wireless telegraph by Guglielmo Marconi in 1896 [99]. Those were the two first practical implementations of long-distance communication systems based on the discoveries of electricity and radio waves respectively and gave rise to the development of today's omnipresent and highly sophisticated communication systems ranging from radio or TV broadcasts to cellular networks.

In the simple example of a telephone system, the transmitter is comprised of a microphone that transforms the acoustic pressure waves into an electric signal, an amplifier, and a subsequent filtration that adapts the signal for passing the channel. In this case, the channel is typically implemented by a metallic twin cable with a finite bandwidth, a finite attenuation, and additive noise. The job of the receiver is to recover the originally transmitted signal impaired by passing the channel. In the telephone system example, the received signal undergoes amplification, filtration, and possibly some kind of denoising before driving a speaker to recover the original acoustic pressure wave. A different example might be the storage of a file on a magnetic medium. In this example, the goal is to reproduce the saved data at some point in the future. The channel is realized by the magnetic medium with its characteristic properties such as zero gain for DC signal components. For a successful transmission through a magnetic tape, for example, the modulation process has to transform the source data into a continuous waveform with zero DC power.

The goal of a typical communication problem is to develop strategies for the transmitter and receiver that would allow reliable transmission given some particular form of the source data and the channel. Usually, there are additional restrictions on various parameters such as bit rate, latency, error rate, or energy efficiency to name a few. The objective of the research in the communication field is to develop an analytical model of the channel and design an applicable modulation/coding scheme meeting the requirements.

One of the fundamental tools for the design of communication system solutions is information theory with its foundations laid by Claude Shanon in his work *A mathematical theory of communication* [101] in 1984. Information theory introduced a novel perspective based on a probabilistic way of describing signals and the definition of entropy as a measure of information. The entropy is given as the expected value of the reciprocal of the probability density (or probability mass function for discrete-valued signals) and thus relates the information carried by the signal directly to its uncertainty. Let us imagine one discrete sample with two possible values 0 and 1. With a priori probabilities $\mathrm{p}(0) = 0$ and $\mathrm{p}(1) = 1$ there is no uncertainty and receiving such a sample does not carry any information, i.e. the entropy is zero. If, on the other hand, the probabilities are $\mathrm{p}(0) = \mathrm{p}(1)$ an observation of a sample carries exactly one bit of information. It can be shown that

a uniform distribution (in a discrete-valued case) maximizes the entropy as it introduces maximal uncertainty. It is possible to measure the amount of information between a pair of probabilistically described signals in terms of their so-called mutual information. This concept is very useful to characterize a wireless channel in terms of its capability of information transfer. In the case of a point-to-point channel, we define capacity as the maximal mutual information between the transmitted and received signal. The capacity is measured in bits and gives the maximal rate at which it is possible to communicate with a zero error probability. In the case of the Additive White Gaussian Noise (AWGN) channel, the capacity is given by the well-known Shannon-Hartley theorem [23]

$$C = \log_2 \left(1 + \frac{P}{N}\right) \text{ bits per channel use,} \tag{2.1}$$

where $\frac{P}{N}$ is the ratio of the signal power and the noise power, the so-called signal to noise ratio (SNR). There exist several such capacity theorems for various channel types and transmission techniques. They pose a theoretical limit but usually do not give a concrete procedure of how to achieve them. Shanon showed, for instance, that (2.1) can be achieved by utilizing random codes of infinite length. In a practical solution, such an approach would lead to an infinite delay and exponential complexity of the receiver which is prohibiting. This motivated a great number of works on the design of codes performing close to the capacity while having an internal structure allowing for decoding with low complexity at the same time [22]. Up to this time, the two main solutions to this problem are the so-called turbo codes [15] and low-density parity-check codes [39] (LDPC). Turbo codes are constructed as a set of convolution coders applied on bit-interleaved data at the transmitter. The decoding of turbocodes is based on iterative decoding algorithms with acceptable complexity. LDPC codes, on the other hand, are linear block codes with a long codeword and sparse parity check matrix. Although LDPC codes were first published in 1962, it was not until the invention of the sum-product algorithm [66] as an efficient way of decoding, that LDPC codes could be practically implemented. Even though it is known how to overcome the exponential decoding complexity, codes whose performance is approaching the channel capacity operate with very long codewords and are not suited for low latency applications.

A different topic, although connected, is the problem of channel effects such as fading and multipath propagation. Depending on the frequency band, environment, user mobility, etc. the wireless channel is difficult to model accurately enough while being mathematically tractable. A basic and widely used linear model is given as

$$x(t) = \int_0^\infty h(\tau)\, s(t-\tau)\, \mathrm{d}t + n(t), \tag{2.2}$$

where $n(t)$ is additive noise and $h(\tau)$ is the impulse response (typically the impulse response is time-dependent as $h(t,\tau)$). There have been a lot of different strategies how to limit the negative impacts of channel fading and noise on the quality of reception, such as the rake receiver for multipath channels or OFDM for frequency selective channels [112] to name a few. With the introduction of MIMO systems and in particular space-time coding, the perception of the wireless channel changed. It can be shown that by adding multiple transmit and receive antennas and by joint coding of the signals transmitted from individual antennas, the capacity can grow linearly with the number of antennas provided there is a rich scattering environment. In other settings, having multiple antennas might provide diversity gain or array gain, both improving the reliability of the transmission [50, 112, 84]. In the context of MIMO, the multipath propagation, and scattering in the channel is not perceived as a complication to be overcome, on the contrary, it allows to achieve better performance when exploited properly.

So far we addressed only situations with a single transmitter and a single receiver. Such a scenario rarely occurs in the real world where there are plenty of users and various wireless communication systems operating at the same time. This causes additional impacts on the wireless

channel in terms of interference. From the nice property of Gaussian noise, it follows that the likelihood of the observation $x[n]$ given the transmitted signal $s[n]$,

$$\mathrm{p}\left(x[n]\,|s[n]\right) \propto \exp\left(\frac{-1}{\sigma_w^2}\sum_n |x[n]-s[n]|^2\right) \tag{2.3}$$

depends on the energy of the signal difference. In turn, for two signals $s_1[n], s_2[n]$ with zero mutual energy $\sum s_1[n]\, s_2^*[n] = 0$ (orthogonal signals), the respective likelihoods become independent

$$\mathrm{p}\left(x[n]\,|s_1[n], s_2[n]\right) \propto \exp\left(\frac{-1}{\sigma_w^2}\sum_n |x[n]-s_1[n]|^2\right) \exp\left(\frac{-1}{\sigma_w^2}\sum_n |x[n]-s_2[n]|^2\right). \tag{2.4}$$

This fact provides a powerful tool to effectively separate signals from different users even though their respective electromagnetic (EM) waves interact freely. A heavily applied method is to divide the signal space into orthogonal subspaces, which get allocated to individual users. The signal space division can be done in plenty of different ways:

Space division multiple access (SDMA) The power of EM waves attenuates with the distance from the transmitter as $1/d^\alpha$, where d is the distance and α the path loss exponent. In free space $\alpha = 2$, in different environments (indoor, urban area) the exponent may be different [1, 78]. At a sufficient distance, the power of an interfering transmitter can be neglected w.r.t. the intended signal source and the signals do not interact. This way we get two independent channels. This is the fundamental enabler for cellular networks, where space separation is used to assures the orthogonality of signals within different cells.

Time division multiple access (TDMA) This is the simplest approach, where we ensure that the different signals exist only in non-overlapping time windows. The application of this method is very easy, still, a common time base of all participants and a controlling authority is needed. Additionally, this method of signal space sharing is the least optimal from the information-theoretic perspective. Nevertheless, it is a popular choice used even in today's 4G networks.

Frequency division multiple access (FDMA) Signals may be separated in frequency, such that their spectra do not overlap and thus their mutual energy is zero. In cellular networks, this method is commonly used together with SDMA and TDMA to partition the signal space into so-called resource blocks. The base station then allocates the resource blocks to individual user terminals. Also, two adjacent cells commonly operate at different frequencies to lower the interference.

Code division multiple access (CDMA) When using CDMA, each user modulates their data with the use of their personal code. The codes are chosen such as to have a low mutual correlation. The resulting signals are then transmitted through the same channel and are allowed to interact at the signal level. After the reception, the data can be separated based on the known code of the target user. Usually, CDMA users share a wide frequency band, which makes the signal more resistant to fading and also allows a more flexible resource allocation. On the other hand, it poses high demands on power control. The transmitting power of each user has to be set such that the received power level at the receiving antenna is equal for all the incoming signals. The most well-known usage of CDMA was its employment in the 3G networks.

For illustration, Table 2.1 shows the multiplexing methods employed by different network generations.

1G	2G	3G	4G	5G
FDMA	TDMA/CDMA	CDMA	CDMA/OFDM	OFDM/BDMA

Table 2.1: Multiplexing methods used in individual cellular network generations. OFDM (orthogonal frequency division multiplexing) falls into the category of FDMA. BDMA (beam division multiple access) uses narrow beams, aimed at different users and achieves spatial division [16, 7].

2.2 Current Trends

In today's world, the demand for communication is extremely high and is still growing, the wireless communication data rates double about every two years [9]. The number of actively used handheld devices is increasing as well as the traffic they generate. But the need for wireless communication does not originate only from mobile devices, in recent years a great number of stationary internet users have switched to wireless access. Another source of demand for communication comes from the so-called Machine-To-Machine (M2M) connections, which are estimated to represent one-half of the globally connected devices [20]. In the near future, both the number of communicating devices and the total throughput needed are likely to grow even more with the rise of a broad range of new application scenarios in new areas [113]. A highly discussed field is vehicular communication [46, 6]. The idea is to allow for real-time communication between nearby vehicles (V2V) and infrastructure (V2I) like traffic lights in an ad-hoc network. The ultimate goal is to develop a system, where autonomous cars could share their parameters such as speed, heading, or intentions and jointly sense the environment. Such a system promises to increase safety and efficiency by choosing optimal routes or avoiding traffic jams. A different emerging field is the smart grid [56, 30], which would connect the electric energy producers, consumers, and the transmission infrastructure. With a smart grid, it would be possible to collect a huge number of data on real-time power consumption and production. Together with the possibility of intelligent control of individual consumers, such as the charging of an electric vehicle or water heating, the consumption could be optimized w.r.t. current production from renewable sources for example. Other future areas which will benefit from wireless communication are smart homes and cities [10, 106, 59, 60, 65], smart health [2] and smart factories [89, 79] to name a few.

The next generation of mobile networks (6G) has the ambition to be a solution to the communication needs as described in the previous paragraph. Different applications have very different demands on the network. For example, a smart grid is expected to generate only small amounts of data with modest requirements for latency or reliability. With a potentially large amount of connected nodes, on the other hand, it may pose high demands on connectivity. Vehicular communications will require large data rates and ultra-low latency, medical applications such as remote surgeries are going to require high reliability and low latency and other applications will emphasize simple devices with long battery life for instance. In Fig. 2.1 we show the requirements on 5G and 6G networks in terms of a set of important parameters. It may be impractical, or even impossible, to assure all of those requirements at the same time. A viable design of the 6G network is a set of different communication strategies which trade off the various requirements and target a particular application scenario while being part of a common cross-layer architecture of a common unifying network. In the rest of this section, we briefly comment on the promising research directions and communication techniques which enable the rise of 6G[7, 123].

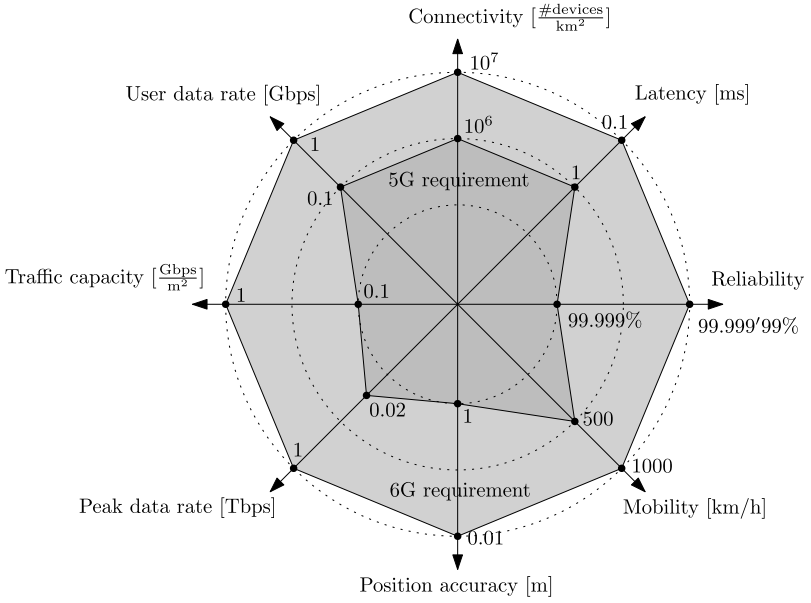


Figure 2.1: Qualitative comparison of the 5G and future 6G networks in terms of requirements [45].

2.2.1 New Spectrum Usage

To increase the capacity, there are two possible ways. Either to develop more resource-efficient methods or to increase the number of resources. One of the key radio resources is the utilized frequency spectrum. The traditional frequency bands used for communication in cellular networks (1G - 4G) as well as short-range standards such as WiFi or Bluetooth reach approximately 6 GHz with a wavelength from units to tens of centimeters. To meet future transmission rate and latency demands, higher frequency bands such as millimeter waves (mmWaves) with frequencies from 30-300 GHz or even terahertz communications in the range 300 GHz to 3 THz with a wavelength of tenths of millimeters are considered [123]. A further increase in frequency ultimately leads to optical wireless communications (OWC) in the infrared band or even visible light communication (VLC).

The application of mmWaves offers a huge available bandwidth and in turn large communication rates. At the same time, it brings several new challenges in contrast to microwave communications below 6 GHz. The wireless channel behaves differently for higher frequencies, the penetration loss is higher such that the waves get blocked by walls or moving people and it is difficult to cover an inside area with nodes deployed outside. Even in the case when a line-of-sight (LOS) is present, a higher path loss and stronger impact of atmospheric conditions, such as rain or fog, demand higher transmit power to achieve a comparable SNR [94]. A potential solution is the use of massive MIMO and the resulting array gain, where the radiated power gets focused into a narrow beam aimed at the user. The shorter wavelengths of mmWaves by contrast, allow for smaller antennas and facilitate the construction of small arrays with a great number of elements. There is however a difficulty with the aiming precision of such a narrow beam, especially in a scenario with moving users [122].

Terahertz communications offer theoretically even more bandwidth and in turn even higher data rates. At the same time the propagation properties are worse than for mmWaves. For this reason, the terahertz communication use cases are usually limited to short-range links of several units to tens of meters, such as in an office room or data center. A survey on this topic with numerous references can be found in [7, 4].

Optical wireless communications, while also based on electromagnetic waves, are significantly different from radio communications. Instead of complex and expensive RF devices, in the case of VLC, power-efficient light-emitting diodes (LEDs) or laser diodes can be used as a transmitter. Some techniques such as intensity modulation allow the same physical LEDs to combine the function of lighting and signal transmission at the same time. The main benefits are the potential throughput, low complexity, cost-efficient and low-energy equipment, low latency w.r.t. RF, and security due to very high penetration loss (LiFi). The highly directional property allows VLC to be used for sensing and relative positioning (LiDAR). More on this topic can be found in [90, 19].

2.2.2 Reconfigurable Intelligent Reflecting Surfaces

A comprehensible overview on this topic can be found in [12, 72, 120]. A reconfigurable intelligent reflecting surface (RIS) is a surface that acts as a controllable reflective mirror for radio waves. In contrast with ordinary "rough" surfaces which provide a diffused reflection with many different reflection directions and phase shifts, RISs are composed of patches of metamaterials and provide a specular reflection. A metamaterial, or more precisely metasurface in the case of a 2-dimensional metamaterial, is typically a periodic structure designed in a way to provide specific reflection properties such as the angle (generalized Snell's law [128]) or phase shift [81]. Depending on the wavelength and the required steering range and precision, RISs can be formed by simple antenna structures controlled by PIN diode RF switches as shown in [105, 57]. As such the RISs are intended as simple and passive devices.

From the communications point of view, the RIS can be perceived as a part of the channel which is controllable and gives additional degrees of freedom for optimizing the communication performance. Alternatively, the RIS can be understood as a fully passive relay node, which receives an incoming signal, performs some processing (phase shift), and re-radiates it to the receiver. The advantages over a simple amplify and forward (AF) relay are the low complexity, no need for a power source, the absence of noise amplification, and no half-duplex constraint. Another view is to interpret the RIS is a part of the receiver's antenna system, which reflects and focuses the incoming signal such that it adds up constructively at the antenna and effectively maximizes the receive antenna gain. Such a role is essential in the high-frequency bands (mmWaves) with the characteristic propagation directionality, where a RIS could be used for assuring connectivity even without a LOS.

One of the main open issues is channel estimation and feedback. Most of the theory developed on the application of RIS assumes some channel state knowledge at the RIS used for optimizing the reflection parameters [55]. It is not clear how to obtain the channel state information (CSI) at the RIS while keeping ti passive and without introducing its complexity extensively [12].

2.2.3 Multiple Antenna Technology

Multiple antenna technology is known to provide a capacity increase, diversity gain, or array gain based on the number of antennas at the transmitter and receiver (MIMO, MISO, SIMO) and the radio environment properties [50]. A natural extension is a multiuser MIMO system. A particularly suitable use case is a base station (BS) equipped with an antenna array communicating with multiple single-antenna users in its respective cell. Although expensive equipment is only necessary at the BS, which is stationary and is not subject to space or energy constraints, it is possible to achieve significant improvements in terms of spectral and energy efficiency [42, 76].

For 6G networks, massive MIMO with a very high number of antenna elements is considered as one of the main tools for efficiency improvement. Together with the exploitation of high-frequency bands such as mmWaves and terahertz communication [17], the size of individual antenna elements will shrink and thus allow the implementation of arrays even for smaller devices. On the other side,

(a) Half-duplex. (b) Full-duplex.

Figure 2.2: With a full-duplex transceiver at the relay node R, the spectral efficiency can be doubled.

massive MIMO and beam forming will become a necessity for overcoming propagation issues with short wavelength communications.

One of the main challenges is channel state tracking because of the high number of parameters and the short coherence time for higher frequency channels. Accurate channel state knowledge is a key assumption for attaining the advantages of massive MIMO such as the capacity gain or precise beam steering [7].

2.2.4 Full-Duplex Radio

The so-called full-duplex constraint is a hardware limitation that prohibits a device to transmit and receive at the same time (more precisely in the same signal subspace). A standard way is to allocate orthogonal subspaces for the TX and RX activity, usually in the time or frequency domain. While a temporal separation introduces additional delay, frequency domain separation requires an allocation of two separate bands for up-link and down-link and special care needs to be taken when designing the hardware in the case of close carrier frequencies. In the past, full duplex transceivers were used mainly in the continuous wave (CW) radar application, where the incident and reflected waves need to be emitted and observed simultaneously. In the field of communications, overcoming the full-duplex constraint would bring an increase in spectral efficiency for one thing and eliminate a fundamental limitation in the design of link layer protocols. A particular opportunity for the application of full-duplex transceivers are wireless networks with intermediate relaying nodes. A simple example is shown in Figure 2.2, where a message is sent from node A to node B employing a relay node R. With traditional half-duplex constrained transceivers we need two resource blocks (time slots, frequency bands, etc) for transmitting one message as indicated in Figure 2.2a. On the other hand, with a full-duplex relay as shown in Figure 2.2b, one-half of the resources can be saved.

The main difficulty with in-band full-duplex (IBFD) transceivers is the self-interference[96]. From the nature of IBFD, it follows that the incoming and transmitted signals cannot be easily separated by orthogonalization. Because the transmitted signal is known, at least in theory a perfect cancellation should be possible. In reality, there are many difficulties. One problem is that the transmitted signal passes the analog front-end and gets reflected from surrounding objects before being received as the self-interference. As a consequence, it cannot be canceled by simple subtraction. By contrast, it is necessary to take into account all the effects such as delay, nonlinear transformation by the power amplifier, and even the impact of the scattering by the environment. Another major issue is the huge power difference between the transmitted signal and the weak signal to be received which challenges the dynamical range of the front end and digital to analog converters (DACs).

In general, there are two types of IBFD nodes, those with a shared antenna and those with two (or more) separate antennas for reception and transmission. With separated antennas it is possible to apply a propagation-domain self-interference suppression, that is to design the antenna radiation patterns or antenna placement in a way that minimizes the signal transfer between the TX and RX antenna. This is the preferred way since it relaxes the dynamic range requirement on the front-end design. The downsides of this approach are its impracticality for small or mobile devices and the restrictions of the radiation patterns w.r.t. the receiving nodes. In the case of a shared antenna,

some part of the separation is achieved typically with a circulator or other types of waveguide devices. Further separation can be achieved in the analog-circuit-domain where it is possible to tap the transmit signal impacted by the hardware nonidealities directly from the antenna feed. In the analog-circuit-domain however, it is complicated to make the system adaptive such that it would be capable to track the changes in the environment for example. Another factor is the fact that analog-circuit-domain implementations are difficult and costly. The digital domain on the other hand is flexible and allows for the design of complex adaptive algorithms. The prohibiting problem for the digital domain self-interference cancellation is the limited dynamical range of the (DACs). Usually, all three domains are used, each removing some part of the interference.

2.2.5 Index modulation

Index modulation (IM) techniques promise high spectral- and energy-efficient communication while keeping the transceiver design relatively simple [13]. The central idea of index modulations is to use other parameters of the modulated signal than the phase or amplitude to convey some part of the transmitted information. It can be understood as a generalization of spatial modulation (SM) which was introduced around 2005 [83, 28].

In the most basic SM scenario, the transmitter is equipped with two TX antennas from which exactly one antenna is used at a time. The transmitted signal does not depend on the data and does not not carry any useful information. The message is translated to a sequence of indexes (1 or 2 in this case) which are used to select the TX antenna to use. Supposing the radio environment is rich enough such that the receiver can distinguish which antenna is active at a time, it is able to receive the message. In general both, the receiver and transmitter can have multiple antennas in which case the antenna index can carry more than one bit of information. Additionally, if we allow the signal itself to be modulated, the efficiency can be further increased although the transceiver design complexity increases. The approach where source data gets split into two parts, one for selecting the index and the rest for modulating the signal can be understood as a constrained case of general space-time coding where only one active antenna at a time is allowed. The main advantages are a simple design and energy efficiency. When compared to a full MIMO with general space-time coding, an SM transmitter is substantially simpler as it suffices to have only a single TX chain or even just a signal generator and some switches in the plain SM case.

In the general IM case, the parameters used for additional information transfer can be diverse [13]. Beyond TX antenna indexes it can be subchannels in an OFDM system, precoding matrices, modulation types, or RF mirrors. RF mirrors are devices mounted on or around a TX antenna, which can alter its far-field radiation pattern in a controlled manner [11]. This idea can be generalized to the so-called channel modulation (CH) where some information is carried by variations of the wireless channel that are distinguishable by the receiver. The source of the channel variations can be the transmitting node itself, as in the case of RF mirror IM, or some other entity (backscatter communications [88]). Such a concept considers passive nodes, i.e. small devices such as sensors that would transmit data by modulating the wireless channel of an ongoing transmission without the need of expending any energy.

2.2.6 Non-Orthogonal Multiple Access

The straightforward method of avoiding interference in a multi-node wireless network, used by previous network generations (1G - 4G), is orthogonal separation. The traditional approach to achieving this kind of separation is based on the partition of the available signal space into mutually orthogonal subspaces - resource blocks (RBs). A central authority (in the case of cellular networks it is the BS) allocates individual RBs to different users in an exclusive fashion which breaks the network into isolated point-to-point links. Based on multiuser information theory results, it is known that this approach is generally suboptimal in terms of spectral efficiency [112, 23].

There are additional drawbacks such as the communication overhead and latency introduced by the mechanism of managing and assigning free RBs to users on demand. There is a trade-off concerning the size of the RBs. Large RBs may be wasteful with low data-rate users, small blocks on the other hand increase the demands on the authority and thus generate more overhead. The most important issue, when considering the rising demands on connectivity, is the finite number of RBs available. If the number of active users is bigger than the number of RBs, the network gets overloaded and it is not possible to serve all the users, even in the case when the required sum rate is low [26].

In a non-orthogonal multiple access (NOMA) system, multiple users utilize the same RB, and controlled interference is admitted. The main two NOMA techniques are power domain NOMA and code domain NOMA.

In the first case, the users are differentiated based on their channel strength, and the capacity-achieving strategy for the AWGN MAC channel, successive decoding and interference cancellation (SIC) [112], is used. A typical example concerns two users, one close to the BS and the second at the cell edge (near-far problem). The BS transmits a non-orthogonal superposition of two signals with different power levels, a strong signal for the remote user and a weak signal for the close user such that the distant user is capable to decode its message. The close user applies the SIC strategy. Because its channel is stronger w.r.t. the distant user, it is possible to decode the distant user's data first. In the second step, based on the perfect knowledge of the interfering signal, the interference can be canceled by simple subtraction. The final step is then to decode its own data from an interference-free observation. In a general case with k users, the BS transmits a superposition of k signals with powers inversely proportional to the individual channel strengths. Each user needs to decode all the messages of the stronger users and cancel their interference before decoding its own message.

Code domain NOMA is based on CDMA access, which provides orthogonal subchannels based on orthogonal spreading sequences of individual users. In the NOMA case, overloading of such a system is allowed, that is non-orthogonal spreading sequences are allowed. Additionally to the SIC, joint decoding based on some kind of message passing algorithm can be used at the receivers. There exist a great number of techniques how to design the spreading sequences such as *sparse code multiple access* (SCMA), *pattern division multiple access* (PDMA), or l*ow density spreading* (LDS) [26, 31, 68, 129].

One of the drawbacks of the described methods is the receiver complexity and energy consumption increase with the growth of the number of users. This problem can be reduced by combining NOMA with the standard orthogonal division. Users are grouped based on some criterion and an RB (directional beam, OFDM subchannel, time slot, etc.) is assigned group-wise. This way, in each group the NOMA is applied only to a limited number of users. A different solution is a so-called *rate splitting multiple access* (RMSA) which is based on the idea of canceling only some part of the interference and treating the rest as noise [80].

More details and an overview of different NOMA techniques can be found in [75, 31].

Chapter 3

Wireless Networks

The majority of today's radio networks, carrying most of the traffic, are so-called cellular networks. One of the principles fundamental to classical cellular networks (1G - 4G) is orthogonal separation. The first level of separation is achieved by spatial division. The covered area is partitioned into individual cells as shown in Figure 3.1. Each cell is served by a dedicated transceiver, the BS. The BS acts as a central hub for data transmissions as well as the authority controlling the resource allocation within the cell. The individual BS are interconnected through a core network implemented typically by fiber optics. The second level of separation assures the mutual isolation of the users within the cell. Different techniques were utilized based on the cellular network generation to achieve this as listen in Table 2.1. This way the network gets transformed into a set of isolated point-to-point radio links (represented by dotted lines in Figure 3.1).

With the orthogonal multiple access paradigm, the only viable method to meet the growing demand for the number of active users and the throughput is the densification of the cellular network. That is, exploiting the SDMA and installing more base stations with reduced transmit power within shrinking cells. The densification allows for an approximately linear growth in achievable data rates. The use of TDMA does not enhance the total throughput and the currently used frequency spectrum is already very densely occupied. As cells get smaller, their size gets comparable to buildings and other city features. This brings the problem of providing adequate core network connectivity, which is often accomplished by additional wireless links. Also, adjusting the propagation model to a reduced distance between the transmitter and receiver is needed. The latter aspect is shown to have the potential to prevent further densification altogether [9].

A different approach to the problem of the rising number of users is the concept of the so-called mesh networks [5, 21]. Wireless mesh networks work without any central authority or access point. The information flow is relayed by individual users and routed through the network. This way, the physical distance between the transmitter and receiver is kept low and lower transmitting power can be used. In Figure 3.1 the user terminal ue7 has no connectivity to the network because it is located outside the range of all base stations. In a mesh network, however, user ue7 could use ue5 as a relaying node and get a multi-hop connection to the network. Another potential benefit can be seen when looking at users ue1 and ue2. They are located closely together but at the edge of the cell, far away from the base station. In the cellular setting, their mutual communication would have to be relayed over the base station BS1. In a mesh network, they could communicate directly, which would require less transmitting power and would greatly reduce interference to the surrounding users.

Giving up a central authority comes with a lot of problems such as distributed network topology estimation, the need for distributed routing algorithms, and opens security-related concerns just to name a few. A fundamental issue is the limited scalability of mesh networks. By adding nodes to the network, the capacity per node decreases. In [44] it was shown that in conventional networks, where

nodes communicate over point-to-point links while treating signals from other users as interference, the total capacity of the network scales as the square root of the number of users. From this, it seems inevitable that the capacity per user goes to zero for large ultra-dense networks.

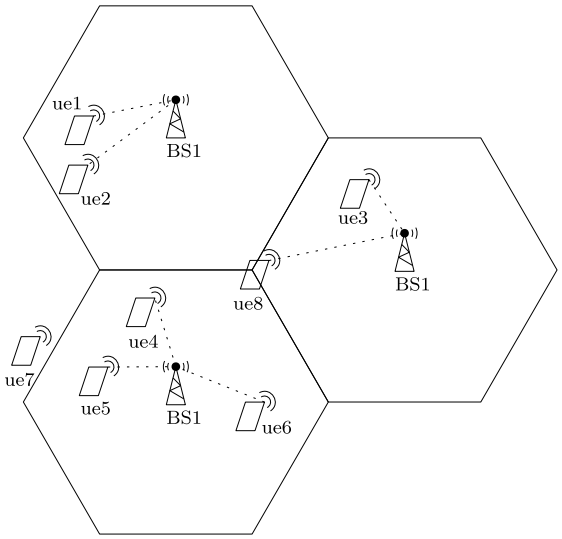


Figure 3.1: Classical cellular network topology.

3.1 Network Information Theory

In Section 2.1 we briefly touched the importance of information theory for understanding and characterizing the transmission of information as well as the design of practical communication methods. In the case of a point-to-point link, the channel's ability to carry information can be elegantly characterized by a single number, the channel capacity. In the case of a network with multiple nodes, the situation gets more complicated. In the point-to-point case, the capacity maximizes the information transfer between the two endpoints. With multiple nodes the optimization goal is not clear anymore, should we maximize the sum rate despite a potentially unfair solution or should we allocate the same capacity to all users even when it limits the overall throughput? In general, we cannot characterize a multi-node channel by a single number, instead, we use a so-called capacity (rate) region which contains all the vectors of individual rates achievable when communicating over a given channel. It can be shown that by proportional time sharing between two strategies, a corresponding convex combination of the respective rate vectors can be achieved. It directly follows, that the capacity region is convex. Besides the more complex characterization in terms of the rate regions, with multiple nodes there is a large number of degrees of freedom to optimize over coming from the broadcast nature of the wireless channel and the resulting interference, the various ways of routing and relaying in the network including potential feedback as well as the lack of the source-channel separation theorem. The source-channel separation theorem holds for a point-to-point link and states (in simple terms) that source and channel coding can be optimized independently. In a multi-node network, on the other hand, the encoding of correlated data sources at different nodes has to be part of a joint strategy to attain optimal performance. There exist special cases of node topologies and channel models (such as multiple access channel or degraded broadcast channel) where the capacity region is known, but a universal answer to the capacity of a general multi-node networks is not available [43]. A comprehensive treatment of the vast topic of network information theory can be found in the book [33].

cut	super TX	super RX	upper bound
a	X_1	X_1, Y	$R_1 \leq I(y; x_1) \leq B \log_2\left(1 + \frac{P_1}{2N_0 B}\right) = C_1$
b	X_2	X_2, Y	$R_2 \leq I(y; x_2) \leq B \log_2\left(1 + \frac{P_2}{2N_0 B}\right) = C_2$
c	X_1, X_2	Y	$R_1 + R_2 \leq I(y; x_1, x_2) \leq B \log_2\left(1 + \frac{P_1 + P_2}{2N_0 B}\right)$

Table 3.1: Upper bounds resulting from different cuts according to Figure 3.2.

3.1.1 Multiple Access Channel

To demonstrate some key features of network information theory and the capacity region we will investigate a simple three-node network, the well-known multiple access channel (MAC). It consists of two source nodes X_1, X_2, and one destination Y. For simplicity, we will assume a complex band-limited AWGN channel with noise power N, bandwidth B, and transmit powers $P_1 > P_2$ for sources X_1 and X_2 respectively. In the derivation of the capacity, we use the standard method of applying upper and lower bounds yielding the capacity in the case of intersection.

For the upper bound, we will use the so-called cut-set bound which is based on the idea of separating the set of nodes into two disjoint subsets and grouping the nodes at each side of the cut into a virtual super transmitter and super receiver respectively. The super transmitter and super receiver can be visualized by introducing infinite capacity links between all the nodes at the corresponding side of the cut. The information rate crossing the cut can be upper-bounded by the capacity of the point-to-point link between a super transmitter and a super transmitter. In Figure 3.2 we show the three possible cuts of the considered three-node MAC.

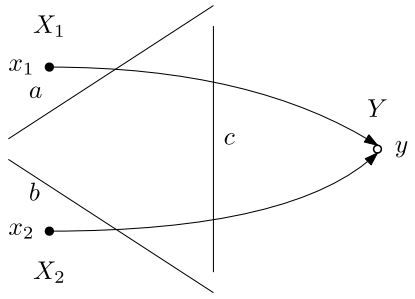


Figure 3.2: Two source MAC channel with three network cuts a, b, c.

The upper bounds on the rates R_1 and R_2 between nodes X_1, Y and X_2, Y respectively are summarized in Table 3.1. The mutual information expressions (maximized over the input distribution) are evaluated using the result for a point-to-point case given in (2.1) modified for a band-limited channel, giving the results in [bits/s/Hz]. In Figure 3.3 the inequalities obtained from individual cuts are presented graphically in the rate space. The rate vectors meeting all the inequalities are represented by the shaded area.

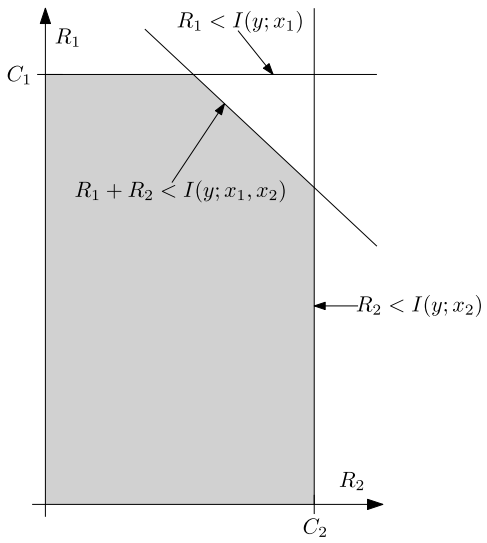


Figure 3.3: Cut-set bound resulting from the union of upper bounds given in Table 3.1.

Now we will investigate the lower bounds resulting from concrete communication strategies. As mentioned in the introduction to this section, proportional time sharing allows us to attain any convex combination of achievable rate pairs. In this case we consider the two extreme strategies where only a single source is transmitting at a time with the corresponding point-to-point capacities C_1, C_2. We note that in this case, the proportional time sharing is equivalent with TDMA. By varying the time proportion devoted to either user every point of the line segment between the rate points $[C_1, 0]$ and $[0, C_2]$ is achievable (Figure 3.4).

When employing FDMA instead, we allocate bandwidths $B_1 = \lambda B$ and $B_2 = (1 - \lambda) B$ to source X_1 and X_2 respectively with the proportionality factor $0 \leq \lambda \leq 1$. This approach yields two orthogonally separated links and we can apply the point-to-point capacity formulas as

$$B_1 \log_2 \left(1 + \frac{P_1}{2N_0 B_1}\right), \; B_2 \log_2 \left(1 + \frac{P_2}{2N_0 B_2}\right). \tag{3.1}$$

The resulting curve for varying λ can be found in Figure 3.4.

A more interesting strategy than orthogonal separation is the so-called superposition and interference cancellation (SIC). In this approach, both X_1 and X_2 are active at the same time and the receiver Y observes a superposition of both signals. In the first step of message decoding at R, one source is treated as pure interference. If, for instance, we first focus on decoding $X_1$the achievable capacity is given as

$$C_1' = B \log_2 \left(1 + \frac{P_1}{B\,(2N_0 + P_2)}\right). \tag{3.2}$$

Based on the knowledge of the message from X_1 it is possible to subtract its source signal from the observed signal and decode X_2 in the second step. The achievable rate of the second source results as

$$B \log_2 \left(1 + \frac{P_2}{2N_0 B}\right), \tag{3.3}$$

which is equal the full rate C_2 because the interference from X_1 was completely canceled. In Figure 3.4 this rate pair $[C_1', C_2]$ is denoted by b. Analogously it is possible to attain point a when first decoding X_2 and removing its interference. Using the convexity argument we obtain the achievability region as shown in Figure 3.4.

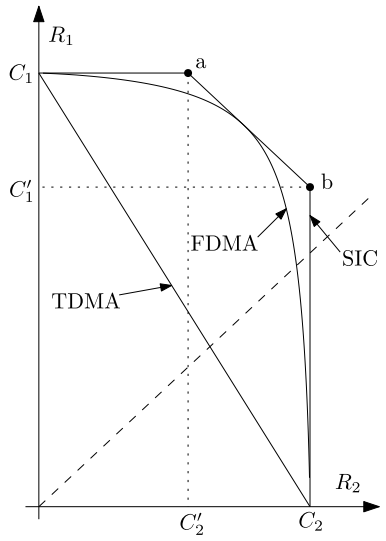


Figure 3.4: Lower capacity bounds achievable by different strategies.

When comparing the upper bound in Figure 3.3 and the lower bound achieved by SIC, we note that they coincide yielding the capacity region. The second observation is that both TDMA and FDMA are sub-optimal except for a set of isolated rate points. When requiring fairness, that is an equal rate for both users (dashed line in Figure 3.4), the SIC gives the optimal result whereas TDMA delivers noticeably bad performance.

3.2 Network Coding

In this section, we give a general overview of the Network Coding (NC) concept, which was introduced in 2000 by the "*Network information flow*" [3] and provoked a radical change in the perception of information propagating through a network. The authors considered a single-source multicast network and proved that the maximal flow of information is still given by the minimal cut as is the case in a single-source single-destination case, however, it cannot be achieved by simple routing. By routing, we understand a network communication strategy, where an information packet is treated as a commodity and is transported along individual links with intermediate nodes relaying or duplicating the incoming packets onto their outgoing links. The second major conclusion, regarding a multi-source multicast scenario, is that optimal throughput cannot be achieved by superposition coding in general. It has been shown in [3, 125] that even in a simple network with two independent sources superposition coding according to the source-separation theorem does not lead to an optimal solution.

Network Coding is primarily focused on wireline networks comprised of a set of nodes interconnected by noiseless point-to-point links. Such networks are commonly modeled as a directed graph (V, E), with V the set of nodes, E the set of edges, and $c(e) > 0$ the capacity of each edge $e \in E$. In a single-source unicast scenario, there is one data source $s \in V$ and one destination node $d \in V$. The problem of finding the capacity of the network is to determine the maximal admissible information flow from the source to the destination. For an admissible flow, it must hold, that the transport load put on each individual edge is less or equal to its capacity. According to [34], the capacity of the maximal admissible flow is equal to the capacity of the minimal cut-set. A cut-set C is defined as a set of edges such that when removed from the network, the graph splits into multiple components, where s and d are elements of two different components. The capacity of the cut-set

is defined as the sum of its edge capacities as

$$c(C) = \sum_{e \in C} c(e). \tag{3.4}$$

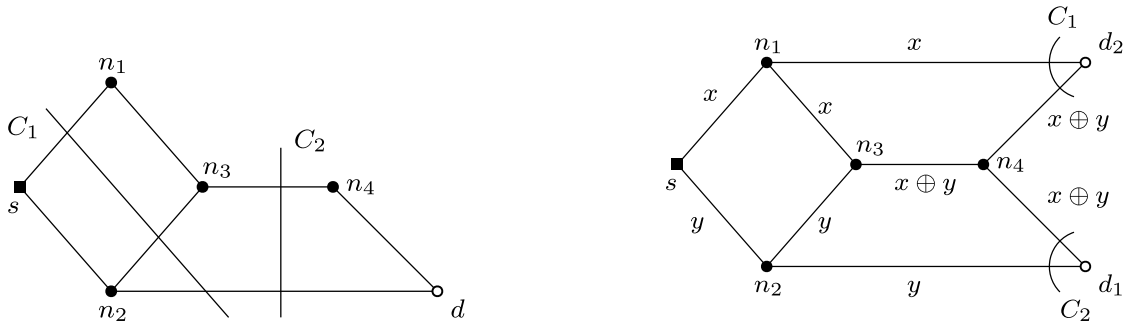
In the case of a single-source multicast network, there are multiple destination nodes $d_i \in V$. From [3] we know, that the broadcast rate h from the source s to all destinations d_i is admissible if and only if the values of all maximal flows from s to all individual nodes d_i are greater or equal to h. In consequence, the maximal broadcast rate h is given by the minimal cut-set. In contrast with the unicast case, the maximal rate h cannot be achieved by simple routing. Instead, nodes are assigned coding functions that compress the incoming packets to form the outgoing ones. As the original data messages pass through the network, they get progressively combined and compressed by multiple possibly different functions. The network has to be designed in a way that at the end, the incoming packets at each destination node allow to solve for the original messages.

In [69] it has been shown that optimal throughput can be achieved by linear coding. In a linear NC network, packets are formed as vectors over a finite field $\mathbb{F}_q$, and the condign functions are restricted to component-wise linear combinations with coefficients from the same field $\mathbb{F}_q$. This renders the destination node processing to a simple solution of a set of linear equations. There exist polynomial algorithms [97] for finding optimal coefficients of the linear maps at all nodes. The full full knowledge of the network topology is needed however. For practical applications where nodes do not know the whole network structure, or for dynamic networks where it is difficult to distribute the change of topology to all nodes, random linear coding was proposed [49]. It has been shown that even when coefficients at individual nodes are chosen randomly, there is a high probability of resulting in a good solution as long as the finite field $\mathbb{F}_q$ is big enough. With this approach, the destination cannot invert the network coding function as it does not know the coefficients. A straightforward solution is to prepend a special header to each packet, containing the coefficients [18]. This header would start as all 0 with a single 1 and get successively updated when passing different nodes. There is an overhead which can be a significant drawback for short packets. There are ways to reduce the overhead such as to utilize compressed coding vectors, where the number of packets accepted by the network coding function is limited or non-coherent NC, where the information about the coefficient is represented by linear subspace [61] that can be characterized by its basis and is invariant to linear transformations.

When multiple sources are allowed, things get more complicated. While a lot of cases still can be optimally solved by linear coding (a generalized version with full matrix multiplications at the nodes instead of simple component-wise combining might be necessary), there exist networks where non-linear coding has to be utilized [32] to achieve optimality. There are many more objectives of NC that differ from achieving the capacity, such as to provide correcting capacity [61] in networks with noisy links or malicious nodes, performing network anatomy, or simply providing a robust and reduced complexity method to the routing problem even in single-source unicast networks [29, 36].

For further reading, we suggest the introductory paper [37], a comprehensive tutorial [38], a historical perspective [127], survey papers [14, 98] containing references on the most impacting work, or several books [126, 48, 82] for in detail reading.

Example 3.1. Let us give an example network with one source node and one destination node as is shown in Figure 3.5a to demonstrate the fundamental principle of NC. The set of nodes is given as $V = \{s, n_1, n_2, n_3, n_4, d\}$, where s and d are the source and destination nodes respectively. An edge connecting two nodes represents an error-free communication link of unit capacity. The problem is to determine the maximal admissible information flow from the source to the destination. In the figure two example cut-sets $C_1 = \{(s, n_1), (n_2, n_3), (n_2, d)\}$ and $C_2 = \{(n_3, n_4), (n_2, d)\}$ are shown with respective capacities $c(C_1) = 3$ and $c(C_2) = 1$. Based on this, we can limit the value of the maximal flow to be smaller or equal to 2. In this simple network, it is easy to see, that the actual value of the maximal flow is equal to 2.



(a) Single source, single destination network.

(b) Single source, two destinations broadcast network.

Figure 3.5: Network for Example 3.1.

Now we slightly modify the example and add a second destination node d_2 together with its connecting edges. The goal is to transfer data from the source s to both destinations at the same time. As shown in Figure 3.5b, the source transmits a pair of source bits x and y. The optimal solution is given as follows. We send bit x from s to n_1, where we duplicate it and forward it to both d_1 and n_3. Symmetrically we deliver bit y to both n_3 and d_1. Next, we form the XOR of the bits x and y at node n_3 and transmit the result over the edge (n_3, n_4) to node n_4, which then forwards it to both destination nodes. At this point, node d_1 has received both y and $x \oplus y$ and can recover the original data as $y, x = (x \oplus y) \oplus y$. Analogously node d_2 obtains bits $x, y = (x \oplus y) \oplus x$. Using this strategy, we are able to transmit two bits using every edge only once. In this case, however, we needed to replicate the data at nodes n_1, n_2 and to compute the XOR at node n_3. It is no longer possible to treat the information simply as a fluid passing through a pipe network.

Example 3.2. This example is taken from [125] and demonstrates how by treating information flows from different sources separately it is not always possible to achieve the full capacity, not even with independent sources.

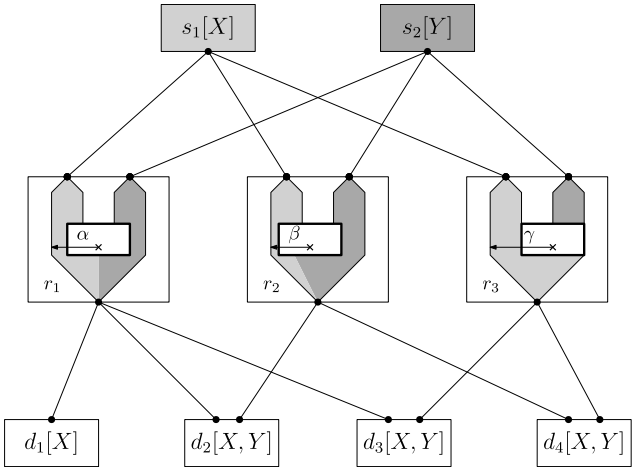
In Figure 3.6a we show a simple network with two sources s_1, s_2, three relays r_1, r_2, r_3, and four destination nodes d_1, d_2, d_3, d_4. All of the interconnecting links are error-free with a capacity of 1 bit/s. There are two independent information streams X and Y originating at s_1 and s_2 respectively. The goal is to reconstruct stream X at d_1 and both streams X and Y at the remaining destination nodes.

We first consider the superposition approach where the streams X, Y are routed independently through the network. As indicated in Figure 3.6a the relays treat the information flow as a liquid and allocate a fraction of their output link capacity to each stream without mixing them. The coefficient $0 \leq \alpha \leq 1$ at r_1 denotes the fraction allocated to stream X, and the rest of the capacity $(1 - \alpha)$ is allocated for stream Y (a practical implementation of such a fractional sharing would require appropriate coding with potentially long codewords). Similarly for coefficients β, γ at relays r_2, r_3 respectively. To decode an information stream at a destination, the sum of all its collected partial rates has to be at least 1. For example, at r_2 it must hold that

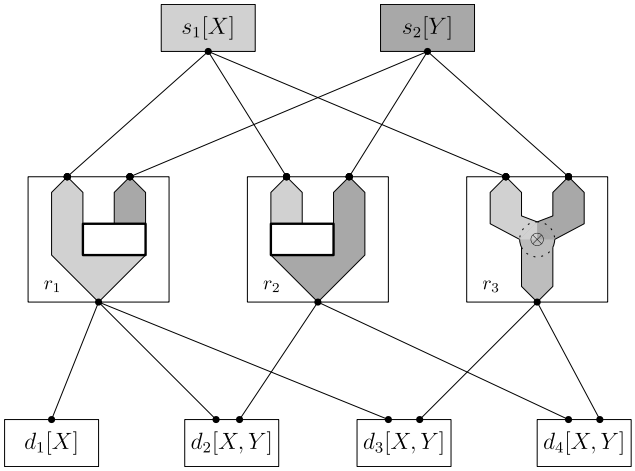
$$\alpha + \beta \geq 1 \tag{3.5}$$

$$(1 - \alpha) + (1 - \beta) = 2 - \alpha - \beta \geq 1 \tag{3.6}$$

for reconstructing stream X and Y respectively. In Table 3.2 we summarize all the conditions. It is easy to verify that there is no solution to meet all the inequalities and thus it is not possible to find a routing strategy to accomplish the goal. Such a conclusion can be made directly by observing



(a) Source separation approach.



(b) Application of NC.

Figure 3.6: Network for Example 3.2.

Figure 3.6a. Destination $d_1$forces $\alpha = 1$ and in turn $\beta = 0$ by the requirements of d_2, and no matter how we set γ it is not possible to deliver X and Y to both d_3 and d_4.

destination	stream X	stream Y
d_1	$\alpha \geq 1$	–
d_2	$\alpha + \beta \geq 1$	$2 - \alpha - \beta \geq 1$
d_3	$\alpha + \gamma \geq 1$	$2 - \alpha - \gamma \geq 1$
d_4	$\beta + \gamma \geq 1$	$2 - \beta - \gamma \geq 1$

Table 3.2: List of conditions for the requested streams at individual destinations to be decodable according to Figure 3.6a.

With NC however, it is possible to achieve the goal for all destinations. According to Figure 3.6b, where r_1, r_2 perform routing with $\alpha = 0$ and $\beta = 0$ and r_2 applies bit-wise XOR as the coding function. Table 3.3 summarizes the streams provided by the relays r_1, r_2, r_3 to individual destinations. We observe that all destinations can decode the required streams.

destination	r_1	r_2	r_3
$d_1[X]$	X	–	–
$d_2[X, Y]$	X	Y	–
$d_3[X, Y]$	X	–	$X \oplus Y$
$d_4[X, Y]$	–	Y	$X \oplus Y$

Table 3.3: Destination nodes observations according to Figure 3.6b.

3.3 Physical Layer Network Coding

Wireless Physical Layer Network (WPNC) coding is a communication concept for dense wireless networks whose origins date back to 2006 and the pioneering work [131, 92]. The goal of WPNC is to address the inefficiency of ordinary solutions based on the layered model and orthogonal user separation. The key idea is to build on the principles of NC and apply the results to wireless networks. The original NC target wireline networks, with separate noiseless links between the nodes. The wireless channel is very different from a wireline link, the differences essential for WPNC are the presence of interference and the broadcast nature. Interference is caused when electromagnetic (EM) waves crying signals from multiple transmitters add together and form a superposition at the receiver antenna. The broadcast nature of a wireless channel on the other hand allows multiple nodes to receive a signal emitted from one transmitter. The straightforward way of applying NC on wireless networks is by utilizing orthogonal separation and good channel coding, which eliminates interference and establishes isolated and error-free links as in a wireline network. Although even such a simple technique brings performance benefits w.r.t. standard routing, it is possible to exploit the wireless channel particularities to our advantage and perform even better.

A complex overview of WPNC and its different aspects can be found in [73, 111].

Example 3.3. Let us start with the simplest wireless network example, where the principles of WPNC can be nicely demonstrated, the so-called Two-Way Relay Channel (TWRC) as shown in Figure 3.7. It consists of two nodes A and B, that wish to interchange information. There does not exist a direct channel among them, so they have to use the intermediate relay node R. All three nodes are assumed to be subject to the half-duplex constraint. This constraint is implied by today's hardware capabilities and means that a node is not capable of transmitting and receiving signals from within the same signal subspace. For clarity, we will assume an orthogonal signal space division in the time domain only. With this assumption, the half-duplex constraint implies the inability of a node to transmit and receive during the same time period.

Figure 3.7: Two-way relay channel topology.

In Figure 3.8 we show three different strategies of communication over the TWRC. All three strategies ensure the delivery of packets $p_A$and p_B from A to B and B to A respectively. We assume the rates to be chosen such that each packet can be transmitted with a low probability of error.

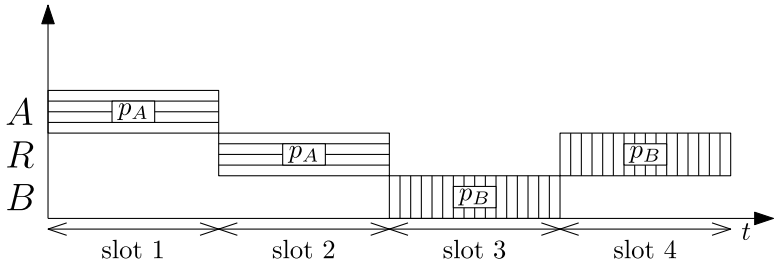
In Figure 3.8a we show the scheduling of a conventional bi-directional relaying. Each packet transmission uses its orthogonal subspace (time slot). In the first slot, node A transmits p_A to the relay. After receiving the whole packet, the relay R re-transmits it to its destination (node B) in the second slot. Similarly for the packet p_B from node B to node A. The whole transaction occupies 4 time slots.

In the perspective of Section 3.2, we know there is a better strategy as shown in Figure 3.8b. In this case, we apply network coding in the same fashion as it would be used in a wired network. Again, there is an orthogonal subspace allocated for each transmission. In the first slot node A transmits p_A to the relay and similarly in the second slot for node B and packet p_B. The relay first buffers both packets and then applies the network coding (in this simple case it again uses a bit-wise XOR map) after that it broadcasts the resulting packet to both nodes A and B. Knowing their respective outgoing packets, nodes A and B are then able to reconstruct their respective target packets p_B and p_A. This approach does save one time slot and makes use of the wireless channel broadcast nature ($p_A \oplus p_B$ is transmitted simultaneously to both nodes A and B).

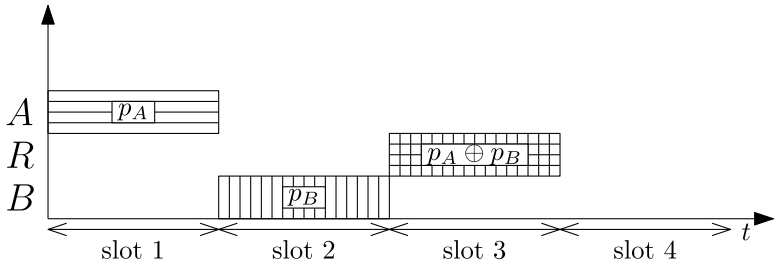
Still, in many cases, there is a better alternative. The main idea of WPNC is to exploit the intrinsic superposition property of the wireless channel and transfer the network coding from the

network layer to the physical layer (Table 3.4). In the third scenario (Figure 3.8c), nodes A and B transmit in the first slot simultaneously. The relay R observes a superposition of the two signals and extracts only the information related to its target packet, namely the XOR function $p_A \oplus p_B$. The broadcast in the second slot is identical to slot 3 in the former case.

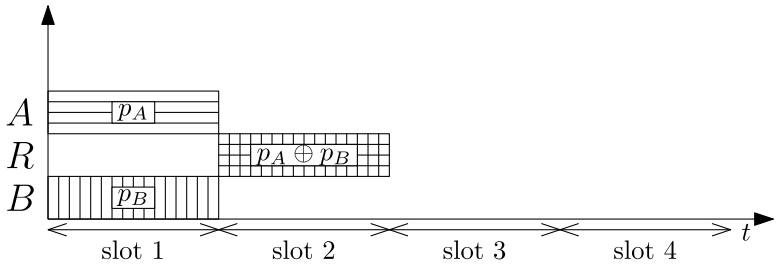
Based on this qualitative comparison we see that the number of time slots needed can be reduced by 25% or 50% w.r.t. the conventional bi-directional relaying strategy.



(a) Conventional bi-directional relaying.



(b) Straightforward network coding scheme.



(c) Physical layer network coding.

Figure 3.8: Two-way relay channel relaying strategies.

3.3.1 Relaying Techniques

There exist various relay strategies referred to as *relaying techniques*, which specify how the observed superposition is processed before being relayed. Different techniques address various criteria such as relay node complexity, overall delay, or throughput maximization. We list the principal approaches together with a short description ordered by the level of the decision made by the relay. The terminology used is based on the monograph [111].

Amplify and Forward (AF)

This technique assumes any linear operation with the observed signal, mostly applied in the analog domain. In the simplest, one-dimensional case, the received signal gets multiplied by a complex

constant. That is, the received signal gets amplified and phase-shifted before being broadcasted. In the case of TWRC, the receiving nodes A and B subtract their own transmitted signal (assuming the knowledge of the channel coefficients) and demodulate their target data [91, 58]. The benefits of this technique are the easy processing at the relay and its independence from the alphabet and received signal structure. The main problem of the AF approach is that together with the useful signals the additive noise gets amplified as well, which makes this method impractical for low SNR scenarios.

Soft Forward (SF)

Sometimes also called Estimate and Forward. In this case, the relay demodulates the signal but does not perform any hard decisions. The resulting entity is a soft measure of a given target data, extracted from the observation. An example is the set of continuous-valued likelihoods for individual target symbols. The final hard decision is left to the receiving nodes, which combine the obtained soft metric with their knowledge of their transmitted signal.

Quantize and Forward (QF)

QF is a strategy very similar to SF, but the soft metric gets additionally quantized and only the resulting discrete metric is transmitted.

Decode and Forward (DF)

In this approach, the relaying node performs a complete demodulation with a hard decision on the target function of the original transmitted data packets. After the demodulation, the relay re-encodes the result and broadcasts it to both nodes.

In wireline NC networks, the network coding functions are typically linear functions over a finite field with coefficients chosen such that the resulting overall network code function can be inverted at all the destinations. In the case of a wireless channel, the signals combine naturally and the relay directly observes a linear combination. This is an advantage as the combination comes for free but poses a problem at the same time because the coefficients of the observed combination are complex and determined by the channel state. There exist various levels of exploiting the signal combination performed by the channel.

Joint Decode and Forward (JDF) The relay first decodes all the incoming messages from the observed superposition and then applies the NC function separately. This allows the implementation of an arbitrary NC function over a finite field and simplifies the design as it is completely isolated from the linear combination formed by the channel. On the other hand, it is restricted by the classical MAC capacity as shown in Figure 3.4.

Hierarchical Decode and Forward (HDF) In contrast to joint decode and forward, the relay does not decode the individual messages to apply the NC function to them but extracts the result directly from the observation. The mapping of the channel-combined observation onto the requested NC linear combination over a finite field introduces difficulties such as an additional layer of ambiguity and strong dependence on the channel parametrization which can make it even impossible in some cases. The main benefit is a capacity gain over the JDF method [109, 110, 111].

Compute and Forward Here, the combinations resulting from signals space superposition are directly mapped onto the results of a finite field linear function with the use of lattice codes [130] which establish the link between the recurrent nature of a finite field used for NC and nonfinie fields

7	Application Layer
6	Presentation Layer
5	Session Layer
4	Transport Layer
3	Network Layer
2	Data Link Layer
1	Physical Layer

Table 3.4: Communication layers of the OSI model.

such as the real or complex numbers which are a natural way of describing the observation. High-dimensional lattice codes allow us to avoid the limitations of HDF and offer higher achievable rates.

Additional details on the topic of relaying techniques and their information-theoretic aspects can be found among others in [93, 110]. In [93], the authors compare different 2-stage protocols (i.e. AF and DF with different target functions) in terms of their respective achievable rates and state the upper bound of the TWRC as the AGWN capacity of the weaker user $C = \log_2(1 + \gamma_{\min})$. With the use of structured lattice codes [133, 87, 119], it has been shown that the DF strategy achieves a rate of $\log_2\left(\frac{1}{2} + \gamma_{\min}\right)$ that is close to the upper bound.

3.3.2 Networking Aspects of WPNC

In this section, we focus on WPNC from the perspective of the entire network. The traditional approach to networking is based on a layered model, which divides the problem of transferring information from a source to its destination through the network into smaller, mutually independent, sub-problems. The well-known OSI model contains seven layers. The first layer interacts directly with the physical channel (i.e. voltage, current, EM waves, etc.) and the seventh layer serves as the interface to the user application. The intermediate layers interact only with adjacent layers such that they serve the layer above and use the services from the layer below. Table 3.4 lists the individual layers of the OSI model. For our purposes, we are interested mainly in the lower three layers, i.e.

Physical Layer This layer is responsible for transmitting individual bits from one network node to another, making direct use of the physical channel. In the case of wireless transmission, it performs the functions of modulation and coding.

Data Link Layer The second layer is responsible for logical node-to-node communication. It receives data frames from the network layer, transforms them into a suitable bit stream, and hands it over to the physical layer. It additionally takes care of multiple access and provides error control.

Network Layer The network Layer ensures the transfer of packets of data over the whole network, that is from the actual data source node to its final destination. The main function of this layer is to route the individual packets, often over multiple node-to-node links through the network. This layer handles the addressing of individual nodes and tracks the topology of the whole network.

The separation of network-related operations (routing, topology estimation) and point-to-point transmissions is intrinsic to wired networks, where independent point-to-point links physically exist. The same approach is commonly used in wireless networks as well, although it is not optimal. It is sub-optimal even in the wired case as was explained in Section 3.2. In its nature, the wireless channel does not offer point-to-point connections. They have to be actively enforced by one of the

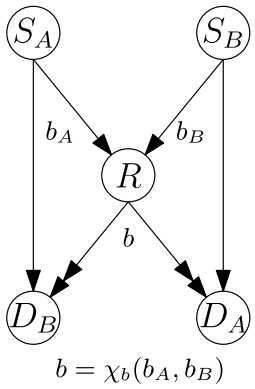


Figure 3.9: Butterfly network.

multiple access techniques. But often, the broadcast nature of the wireless channel can be exploited as well as the unavoidable interference which can be harnessed and used to our advantage.

An example can be found in Figure 3.1. The user terminal ue8 is assigned to BS1 and located at the very edge of the cell. Because it is distant from its base station, it has to transmit with high power and acts as a source of strong interference for the remaining base stations. Alternatively, the signal from ue8 could be received by all three base stations and the independent observations could be jointly processed by the network. This way a better reception in terms of lower SNR and higher diversity could be achieved. Similarly for the up-link. When connected only to BS1, ue8 receives a low-power signal and experiences poor connectivity. Better results (in terms of information rate or outage probability) could be achieved in the case when all three base stations would transmit a useful signal for ue8. In [67], the authors developed and analyzed cooperative diversity protocols which take advantage of the broadcast nature of the wireless channel.

In the WPNC paradigm, the physical layer is network aware. This effectively merges the lower three layers into one complex entity. The benefit is the ability to jointly optimize the local signal processing at individual nodes together with the routing and information flow through the entire network.

Example 3.4. We demonstrate the central idea in the example of the so-called butterfly network [3], as shown in Figure 3.9. It consists of a total of five nodes, two sources S_A, S_B a relay node R, and two destination nodes D_A and D_B. All nodes are subject to the half-duplex constraint. The goal is to convey the two symbols b_A, b_B to their respective destinations D_A, D_B.

To describe the WPNC solution of this network, two orthogonal signal sub-spaces are utilized which we call stages. In this example, the orthogonality will be achieved by a partition in the time domain only. We assign nodes S_A, S_B to transmit in the first stage and node R to transmit in the second stage. The simple and double arrows in the figure represent the transmissions during the first and the second stage respectively. Such a schedule meets the half-duplex constraint for all nodes in the network. The first stage induces two point-to-point channels from S_A to D_B and from S_B to D_A and a MAC from S_A and S_B to the relay R. We call this stage the MAC stage accordingly. In the second stage (the broadcast stage (BC)), the relay broadcasts b to both destination nodes D_A and D_B. The definition of stages in a network determines the connectivity among the nodes.

At the relay, the HDF strategy is applied and the data processing is defined by the Hierarchical Network Code map (HNC map). In general, it can be an arbitrary discrete function of the received messages. In the butterfly network example, the HNC map χ_b is a function of the source symbols b_A, b_B and yields the hierarchical target symbol b. The HNC maps in the network need to be chosen in a way that ensures the solvability of the respective target messages at all destination nodes. In our particular example, a suitable HNC map is again the bit-wise XOR function, i.e. $\chi_b\left(b_A, b_B\right) =$

$b_A \oplus b_B$. After the second stage, D_A knows the symbol b_B and also the hierarchical symbol b. Together they allow the target symbol b_A to be determined without ambiguity as $b_A = b_B \oplus b$. Analogously for the destination D_B.

The butterfly network (a generalization of the TWRC channel (Example 3.3)) has gained much attention. For example in [115], the authors suggest an approach of splitting the information flow and using superposition coding to distribute it optimally over multiple direct and hierarchical paths through the network. This way it is possible to avoid performance degradation in case of unequal rates of the individual links. A general treatment of the networking aspects, together with formal conditions for solvability and a polynomial formalism for analyzing the flow of information in a WPNC network can be found in [111].

3.3.3 Hierarchical Multiple Access Channel

In this section, we focus on the Hierarchical Multiple Access Channel (H-MAC) which is commonly found as a part of a WPNC network. For clarity, we will assume a two-source H-MAC with the HDF strategy, similarly as in Example 3.4 shown in Figure 3.9. Under the HDF strategy, the relay only aims to decode the hierarchical symbol, the result of the corresponding HNC map which can be classified based on its output alphabet cardinality as follows.

Full HNC map A full HNC map is a one-to-one map from the Cartesian product of the source alphabets $\mathcal{A}_{b_i}$ to the output alphabet $\mathcal{A}_b$ of equal cardinality as the input space ($\Pi_i |\mathcal{A}_{b_i}|$). In the case of a two-source H-MAC with $b_A, b_B \in \{0, 1\}$, the hierarchical symbol b would be from an alphabet with 4 elements. Such an HNC map allows all sources to be decoded independently and corresponds to a classical MAC channel.

Minimal HNC map The cardinality of the output alphabet $\mathcal{A}_b$ of a minimal HNC map is equal to the maximal cardinality of the individual source alphabets. It allows an arbitrary source to be determined, given the perfect knowledge of all the other sources. This is the case of the XOR function used in the Examples 3.3, 3.4. In both cases, based resulting hierarchical symbol b it is possible to distinguish between all the values of b_A given b_B and vice versa.

Extended HNC map For an extended HNC map, the output alphabet cardinality is between the full and minimal HNC map. For the decoding at the destination, some independent side information is still needed but it does not have to be perfect as in the case of the minimal map.

Lossy HNC map The output alphabet is smaller than the maximal input alphabet and as such even the full knowledge of one message does not resolve the ambiguity and additional side information is needed.

Besides the question of how to transform the observed signal combination in the signal space into a particular NC function, there is another issue when applying NC from wireline networks to a wireless scenario connected to presence of noise. Generally, we need a mapping that maps the messages at the source nodes into the signal space in such a way, that after passing a noisy wireless MAC, the relay is be able to decode the desired NC function with a low probability of error. In [108, 111] this mapping is termed a network coded modulation (NCM) because it fulfills three tasks, the modulation as it converts the discrete input bits into a signal to be transmitted by the antenna, channel coding to provide protection against noise and the NC which has to respect the network structure. A convenient design of such an NCM is based on a layered structure [110], with a separation of channel coding and the NC. This idea is straightforward in the case of a linear NC and a linear channel code.

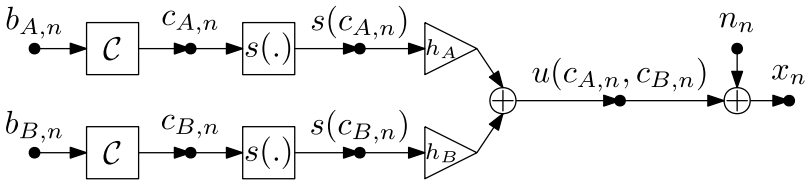


Figure 3.10: Two source H-MAC with isomorphic layered NCM.

Given a linear finite field $\mathbb{F}_q$ coder $\mathcal{C}$, the source messages $\boldsymbol{b}_A, \boldsymbol{b}_B$ get first encoded as $\boldsymbol{c}_A = \mathcal{C}(\boldsymbol{b}_A) = \boldsymbol{G}\boldsymbol{b}_A$ and $\boldsymbol{c}_B = \mathcal{C}(\boldsymbol{b}_B) = \boldsymbol{G}\boldsymbol{b}_B$. Because the channel code is one-to-one, the NC at the message level $\chi_b(b_{A,n}, b_{B,n}) = b_n$ directly implies a corresponding NC at the codeword level $\chi_c(c_{A,n}, c_{B,n}) = c_n$. Assuming a linear $\chi_b(b_{A,n}, b_{B,n}) = \alpha b_{A,n} + \beta b_{B,n}$ with α, β being nonzero elements from the same field $\mathbb{F}_q$, we have

$$\boldsymbol{c} = \mathcal{C}(\boldsymbol{b}) = \boldsymbol{G}\boldsymbol{b} = \boldsymbol{G}(\alpha\boldsymbol{b}_A + \beta\boldsymbol{b}_B) = \alpha\boldsymbol{G}\boldsymbol{b}_A + \beta\boldsymbol{G}\boldsymbol{b}_B = \alpha\boldsymbol{c}_A + \beta\boldsymbol{c}_B = \chi_c(\boldsymbol{c}_A, \boldsymbol{c}_B). \tag{3.7}$$

An NCM fulfilling the property $(\mathcal{C}(\chi_b(b_{A,n}, b_{B,n})) = \chi_c(c_{A,n}, c_{B,n}))$ is called an *isomorphic layered NCM* and will be assumed for the rest of this work. For the expansion part, a standard linear modulation can be applied.

The diagram in Figure 3.10 shows a two-source flat-fading AWGN H-MAC with the source node processing. At both sources A, B, their respective messages $\boldsymbol{b}_A, \boldsymbol{b}_B$ get encoded by a common linear code $\mathcal{C}$ resulting in codewords $\boldsymbol{c}_A, \boldsymbol{c}_B$ respectively. A one-to-one mapper $s(.)$ maps the codewords symbol-wise onto the constellation space. The channel combination forms a hierarchically channel-combined symbol

$$h(c_{A,n}, c_{B,n}) = h_A s(c_{A,n}) + h_B s(c_{B,n}), \tag{3.8}$$

which constitutes the observed hierarchical constellation. The last step is the addition of noise.

The hierarchical constellation (H-constellation) is defined as a map $\mathcal{U}$ from the set of hierarchical symbols $\mathcal{A}_c$ to the subsets of the hierarchical alphabet (H-alphabet) $\mathcal{A}_u$ given by all the hierarchical channel-combined symbols u. The receiver aims to only distinguish different hierarchical target symbols $c_n = \chi_c(c_{A,n}, c_{B,n})$ rather than individual $c_{A,n}, c_{B,n}$. Because χ_c is typically a many-to-one function, the observed constellation has a larger cardinality than the target symbol alphabet $\mathcal{A}_c$.

Example 3.5. We explain this effect in an example shown in Figure 3.11.

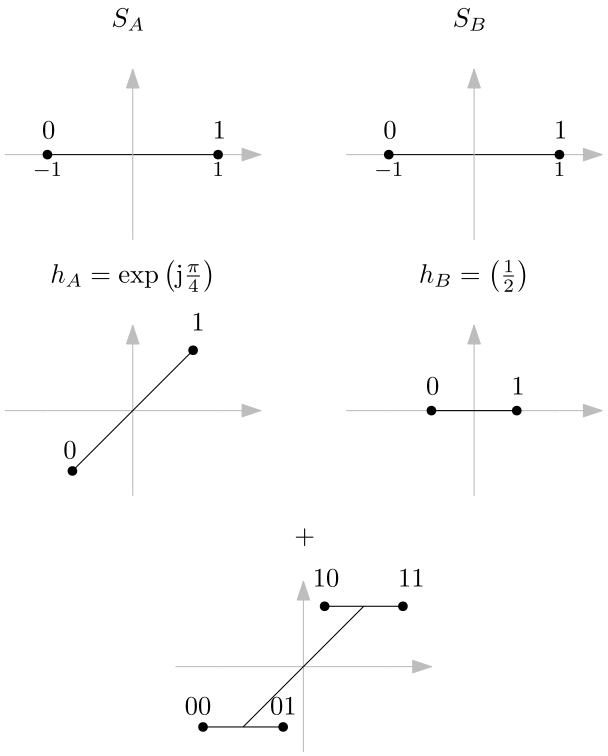


Figure 3.11: Example of an H-constellation.

We assume two sources, each transmitting a single binary symbol b_A, b_B respectively. In both cases, the source symbols get mapped onto BPSK alphabets such that $s_A(c_A) = 2c_A - 1$ and $s_B(c_B) = 2c_B - 1$. The transmission of the useful signal through the MAC channel is modeled as $u = h_A s_A(c_A) + h_B s_B(c_B)$, i.e. the source signals get multiplied by their corresponding channel coefficients and summed at the receiver antenna. In this particular example, the resulting H-alphabet is given as $\mathcal{A}_u = \{h_A + h_B, h_A - h_B, -h_A + h_B, -h_A - h_B\}$. Assuming a minimal XOR HNC map, the H-constellation mapping is given as follows

$$\mathcal{U}(c) = \begin{cases} \{h_A + h_B, -h_A - h_B\} & c = 0 \\ \{h_A - h_B, -h_A + h_B\} & c = 1. \end{cases} \tag{3.9}$$

This effect is called hierarchical dispersion and can be regarded as a nuisance parameter that requires an additional step of marginalization. The observation x_n is typically processed by a soft demodulator (SODEM) according to

$$\mathrm{p}(x_n|c_n) = \frac{\sum_{\chi_c(c_A,c_B)=c_n} \mathrm{p}_w(x_n - u(c_A,c_B))\,\mathrm{p}(c_A,c_B)}{\sum_{\chi_c(c_A,c_B)=c_n} \mathrm{p}(c_A,c_B)}, \tag{3.10}$$

where we sum over all pairs of c_A, c_B for which the HNC map χ_c result is c_n. Even in a simple AWGN channel, the marginalization over the hierarchical dispersion noticeably complicates the error probability analysis [107].

In Figure 3.12 an equivalent channel model is presented, with the target symbol b_n as the input. The model can be divided into three serially connected parts. The H-Encoder performs is a

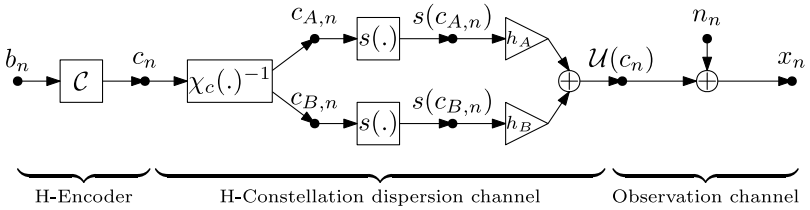


Figure 3.12: Equivalent channel model of a 2-source MAC with isomorphic layered NCM (Figure 3.10). The many-to-one function inversion $\chi_c(.)^{-1}$ is implemented as random sampling from an appropriate probability mass function $\mathrm{p}(c_A, c_B|c_n)$.

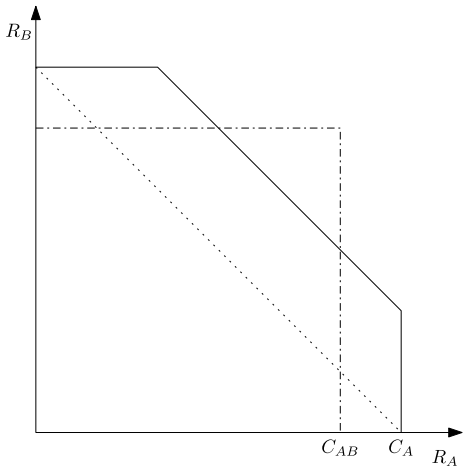


Figure 3.13: Illustrative rate regions for a symmetrical two-source H-MAC. Dotted line: achievable rate for a full HNC map with simple time sharing. Full line: capacity for a full HNC map equals a standard H-MAC with SIC. Dash-dotted line: capacity for a minimal HNC map.

one-to-one mapping from the H-message $\boldsymbol{b}$ onto the H-codeword $\boldsymbol{c}$, given by codebook $\mathcal{C}$ thanks to the isomorphic layered NCM. The second part is the so-called H-Constellation dispersion channel which transforms the H-codeword into the H-constellation. In the real channel, the H-constellation results from the channel combination of multiple signals. From the communication point of view, the only important thing is its structure and stochastic properties and as such it is possible to use the model shown in Figure 3.12, where $\mathcal{U}(c_n)$ is modeled by the inversion of the HNC map $\chi_c(.)^{-1}$. Formally an inversion of a many-to-one function does not exist. In our model, $\chi_c(.)^{-1}$ is implemented as a random sample from the image set according to an appropriate probability mass function $\mathrm{p}(c_A, c_B|c_n)$. If, for example, $\chi_c(.)$ is the bit-wise XOR and c_A, c_B are uniformly distributed on $\{0, 1\}$, $\chi_c(1)^{-1}$ results either in $[c_A = 1, c_b = 0]$ or in $[c_A = 0, c_b = 1]$, each with probability $\frac{1}{2}$. The last part is a standard Gaussian observation channel.

In [110, 111] the rate region of the H-MAC channel was studied. It was shown that the throughput rate region is rectangular and given by the mutual information $I(c_n; x_n)$. As is apparent from the equivalent channel model shown in Figure 3.12, the H-Constellation dispersion channel introduces additional randomness when compared to a single-user channel with the same SNR. Unless the H-Constellation is self-folded, that is all the constellation points of $[c_A, c_B] : \chi_c(c_A, c_B) = c$ fall into the same point, $C_A > C_{A,B}$ as indicated in Figure 3.13.

3.3.3.1 Channel Parametrization

The information-theoretic analysis of the H-MAC usually assumes channel state knowledge at all involved nodes, perfect time and frequency synchronization, and an unconstrained source alphabet [93, 95]. From a practical point of view, all of those conditions are difficult to achieve and pose additional challenges.

The relative parametrization $h = \frac{h_B}{h_A}$ of the channel combined symbol (3.8) is a phenomenon that is not present in the single-user case and that heavily affects the resulting H-constellation. A complex coefficient in a single-user channel has the impact of channel rotation and change of amplitude. Both of those effects can be equalized at the receiver provided accurate channel state information (CSI) is available. In the H-MAC, the relative parametrization completely changes the shape of the constellation. An example of decision regions resulting from a hard decision on (3.10) for a two-source BPSK H-MAC with XOR HNC map is given in Figure 3.14. We notice that the decision region shape changes completely for different relative parametrization values (Subfigures 3.14a, 3.14b).

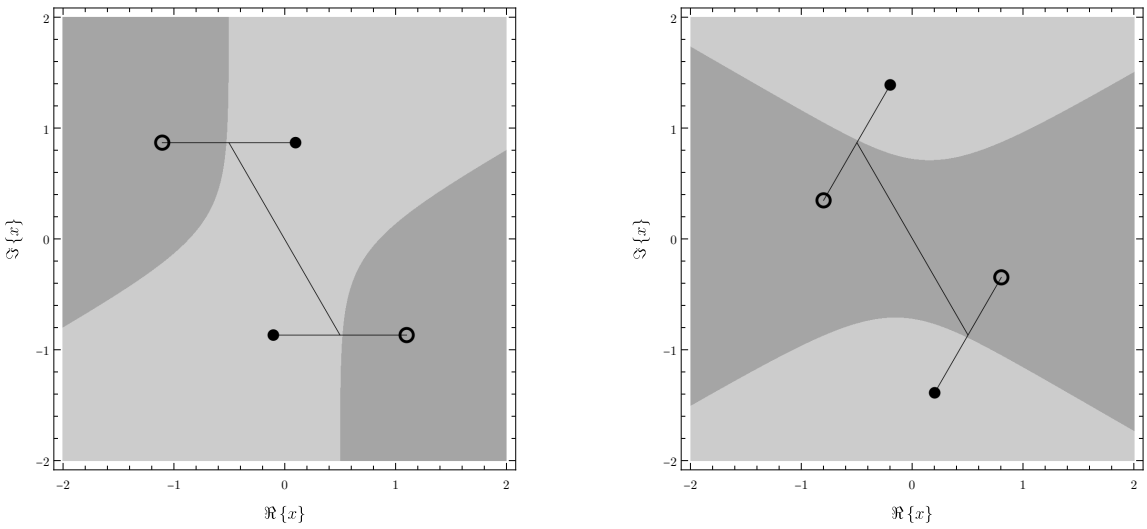


Figure 3.14: Constellation space decision regions for BPSK and XOR HNC map. For convenience, we included the constellation with useful signal points designated by two different marker styles according to the corresponding hierarchical target symbol.

Besides the nonlinear decision region change, the relative parametrization has the potential to completely prevent successful reception. This effect is called singular fading and arises when different combinations of input symbols result in an identical constellation point

$$h\left(c_{A,1}, c_{B,1}\right) = h\left(c_{A,2}, c_{B,2}\right) \;\; c_{A,1} \neq c_{A,2} \vee c_{B,1} \neq c_{B,2}. \tag{3.11}$$

We differentiate between two cases, based on the impact on the H-symbol decision. The resolved singular fading when

$$\chi_c\left(c_{A,1}, c_{B,1}\right) = \chi_c\left(c_{A,2}, c_{B,2}\right) \tag{3.12}$$

or unresolved singular fading if

$$\chi_c\left(c_{A,1}, c_{B,1}\right) \neq \chi_c\left(c_{A,2}, c_{B,2}\right). \tag{3.13}$$

Resolved fading is favorable as it reduces the hierarchical dispersion. In contrast, unresolved fading prevents the receiver to distinguish between different H-symbols $c_1 = \chi_c(c_{A,1}, c_{B,1})$ and $c_2 = \chi_c(c_{A,2}, c_{B,2})$.

Example 3.6. In this example, we demonstrate the effect of singular fading on a two-source H-MAC with QPSK source alphabets with a Gray mapping and the XOR HNC map. In Figure 3.15 the two source constellation get combined with relative fading coefficient h. The resulting H-constellation is composed of 16 points with markers according to the H-symbol $c = \mathrm{XOR}(c_A, c_B) \in \{00, 01, 10, 11\} \equiv \{\bullet, \blacksquare, \square, \circ\}$. The shaded areas mark the H-constellation points which are close and difficult to distinguish by the receiver. In the singular case when $|h| \to 1$ they coincide perfectly which results in singular fading.

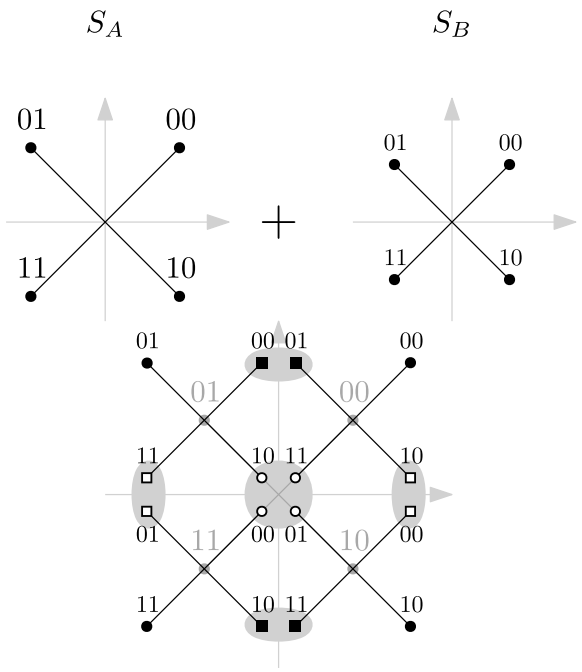
In Figure 3.15a, $h = 0.8$ and the close points belong to the same H-symbol which is apparent as they share the same markers. This case is called resolved singular fading w.r.t. the particular HNC map and does not have a negative effect.

In Figure 3.15b, $h = 0.8\mathrm{e}^{-\frac{\mathrm{j}\pi}{2}}$ and different points within the shaded areas belong to a different H-symbol. For $h = \mathrm{e}^{-\frac{\mathrm{j}\pi}{2}}$ the shaded areas collapse into a single point which results in a 9-point H-constellation. Because it is not possible to distinguish between different values of the H-symbol, this situation is called unresolved singular fading and causes a severe performance degradation.

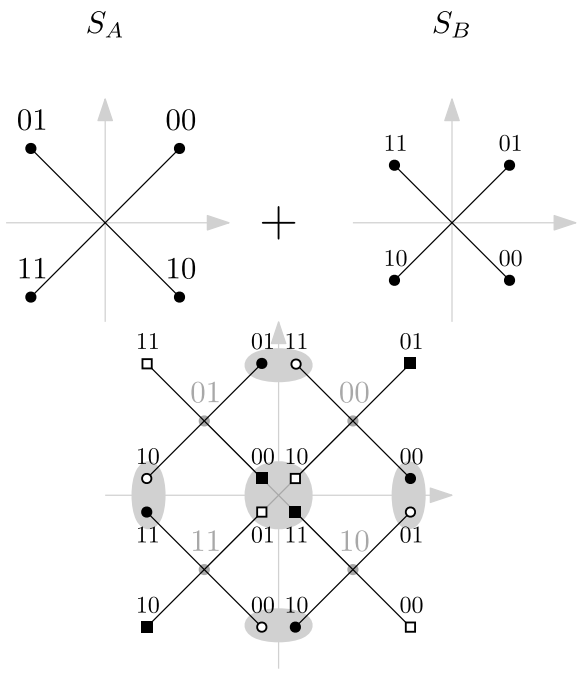
The issue of unresolved singular fading has attracted a lot of attention [62, 47, 114, 35, 74]. Besides the straightforward solution of transmitter-side pre-rotation which is impractical as it requires a feedback channel, suggested solutions rely on particular designs of source constellations with increased cardinality, application of different alphabets on individual sources, or adaptive selection of the HNC map or constellation based on actual channel state. A different approach is to accept the presence of occasional unresolved fading and decrease the negative impact thereof by exploiting receiver diversity [63].

In the case of the AF strategy, the relay does not need any information about the relative parametrization as it only amplifies the received signal. The processing is done at the target nodes, in [41] the channel reciprocity in the particular case of a TWRC was used. In contrast, for an H-MAC and with DF relay strategy, an accurate CSE at the relay is important not only for the implementation of (3.10) but also for more complex adaptive receiver strategies [124]. A detailed treatment of estimation methods and training signal design for various channel models in the context of WPNC can be bound in [40, 70]. In most of the cases, a block-fading assumption is made and the estimate is obtained based on known pilot signals. In Chapter 4 we introduce and verify a H-data aided CSI estimator [54, 53, 52] with relaxed assumptions on block-fading - only the gains are assumed to stay constant and the phase is free to drift within the frame.

In Chapter 5 we propose an alternative approach of demodulation in a phase-varying channel. We apply a differential phase keying to a two-source H-MAC and construct a noncoherent receiver w.r.t. the channel phase parametrization. We demonstrate its good BER performance for the BPSK alphabet and XOR HNC map combination. A noncoherent reception in an H-MAC is not a new idea, in [103, 134] the authors applied a differential modulation to a TWRC with AF strategy and show that a 3 dB loss compared to a coherent detection scheme can be achieved. In [25] a DF scenario is considered and a MAP detector for Gaussian fading coefficients is derived. Noncoherent reception in a WPNC MAC setting with the use of nonlinear BFSK and CPSK was studied in [104] and [116] respectively. In our work [54] we propose multiple differential PSK demodulation methods and show that a 3 dB loss w.r.t. a coherent case can be achieved in the case of the HDF strategy.



(a) Resolved singular fading.



(b) Unresolved singular fading.

Figure 3.15: Singular fading example.

3.3.3.2 Synchronization

Besides channel parametrization, an important facet of WPNC networks is the synchronization of the timebase and/or frequency among the nodes. There are two opposing aspects. On one hand, interfering signals from multiple source nodes are processed and demodulated jointly which poses high demands on accurate synchronization [132, 85]. On the other hand, the lack of a central authority that would provide a reference, adequate synchronization is not trivial to achieve [102, 64, 8]. In Example 3.7 we give a simple example network to demonstrate, that even with a perfect synchronization it is not always possible to assure synchronicity among all the interacting signals.

Example 3.7. To illustrate the importance of the synchronization or an appropriate solution to the asynchronous reception in an H-MAC channel, we present a simple scenario according to Figure 3.16. There are two source nodes S_A and S_B which participate in two H-MAC stages w.r.t. relays R_1 and R_2. Given the topology, the path delays are uneven such that source S_B is unable to adjust its time base to assure a synchronous reception w.r.t. the signal from S_A at both R_1 and R_2.

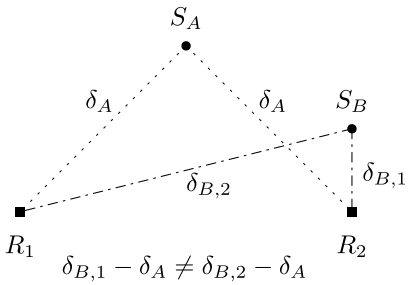


Figure 3.16: WPNC scenario with two H-MACs: (S_A, S_B) $\rightarrow R_1$, (S_A, S_B) $\rightarrow R_2$ such that it is impossible for sources S_A and S_B to synchronize w.r.t. both R_1 and R_2.

The problem of delay between arriving signals in the same MAC stage is generally divided into two cases which are treated separately. An integer multiple of the symbol time duration [70] and a fractional delay. The first case is usually approached by appropriate coding techniques such as cyclic or constitutional codes [117, 121, 118]. The fractional delay is addressed in multiple ways, that differ by the used modulation pulse, the number of samples taken per symbol duration, and how the arising inter-symbol-interference (ISI) is eliminated. In [85, 77, 121] the signal is oversampled (typically there is one matched filter (MF) for each delayed signal), and subsequently a maximum likelihood sequence detector algorithm is applied with various approximations to reduce its complexity. In [86], the authors oversample and apply a linear whitening transformation which results in a synchronous observation although with smaller gain. An alternative with symbol-rate processing is discussed in [71], where mid-point sampling is performed and iterative ISI cancellation is proposed. In Chapter 6 we consider a two-source H-MAC with fractional symbol time duration and analyze the ISI from the H-symbol perspective. We show that by application of an OFDM-like precoding, the ISI can be eliminated. We propose a symbol-spaced sampling and fractionally-spaced sampling receiver side processing and compare the resulting performance for the root-raised-cosine (RRC) modulation pulse.

Part II

Contribution:

In the contributions part, we present the results of our research. The works were published in [53, 52, 54]. In contrast with the publications, this thesis is not conditioned by a page limit, which allows us to present our results in more detail and even to include some additional parts which were not published due to length constraints.
Based on the topics, we divide our contributions into three chapters, *Hierarchical Decision Aided H-MAC CSE, Differential Hierarchical Demodulator of 2-source Hierarchical-MAC Channel,* and *Fractional Symbol Duration Asynchronous H-MAC*. In a closing chapter, we give a conclusion of our work and discuss potential directions of future research and possible practical implementations.
We decided to write individual sections as self-contained units with a proper introduction, result, and conclusion. We believe that this makes it easier for the reader to navigate and allows them to read the chapters independently. A minor drawback of this approach is the inevitable presence of redundancy. We made an effort to keep the amount of redundancy minimal and to use identical notation among different chapters.

Chapter 4

Hierarchical Decision Aided H-MAC CSE

4.1 Introduction

This chapter is devoted to channel state estimation (CSE) in a hierarchical multiple access channel (H-MAC) which was introduced in Section 3.3.3.1. The motivation comes from the non-linear dependence of the hierarchical constellation on the channel coefficients, especially on the relative phase parametrization. For the hierarchical decoding at a relay, it is of great importance to have an accurate estimate of the channel state.

In this chapter, we build on the results of my master's thesis [51]. Therefore, we do not present all the details and refer the reader to the thesis [51] and related publications [108, 53, 52]. To keep this work self-contained, we give a summary of the main assumptions, ideas, and solutions in Sections 4.2 and 4.3, to render the contribution in Section 4.4 comprehensible without the need to study the references in detail.

As mentioned in the previous paragraph, the contribution of this chapter is the verification of the proposed channel state estimator on a real channel using a network of experimental transceivers. To verify that the estimator performs well in a real-world scenario, we compared its performance w.r.t. the results obtained by numerical simulation in terms of various versions of MSE, as well as in terms of the achievable hierarchical BER (H-BER) when used together with an ML demodulator.

The work presented in this chapter is based on my published as well as unpublished research, especially [53, 52].

4.2 System Model

We assume a three-node network with two source nodes (S_A, S_B) and one relay R. Further we assume a perfect symbol-timing synchronization among all three nodes. The network operates in an H-MAC stage. The hierarchical target message at the relay is formed as a many-to-one function (Hierarchical Network Code (HNC) map) of the source messages $b = \chi(b_A, b_B)$. At both source nodes, the messages are encoded by a common linear codebook $\mathcal{C}$ ($\mathbf{c}_A = \mathcal{C}(b_A)$, $\mathbf{c}_B = \mathcal{C}(b_B)$) and symbol-wise mapped on the constellation points $s_{A,n} = \mathrm{s}(c_{A,n})$, $s_{B,n} = s(c_{B,n})$ using a one-dimensional alphabet $\mathcal{A}_s$ of cardinality M_s. The target H-code symbols are denoted as $\mathbf{c} = \mathcal{C}(b)$. We assume an isomorphic layered NCM which implies that $c_n = \chi_c(c_{A,n}, c_{B,n})$. Further, we choose a minimal HNC map χ_c, such that given any two elements from $\{c_n, c_{A,n}, c_{B,n}\}$, the remaining element can be uniquely determined.

We assume a memoryless 2-source hierarchical MAC (H-MAC) AWGN channel

$$\mathbf{x} = \mathbf{u}\left(\mathbf{c}_A, \mathbf{c}_B\right) + \boldsymbol{w} = h_A \boldsymbol{s}\left(\mathbf{c}_A\right) + h_B \boldsymbol{s}\left(\mathbf{c}_B\right) + \mathbf{w}. \tag{4.1}$$

The fading coefficients are decomposed into magnitude and true phase as $h_A = \mathrm{e}^{\mathrm{j}\phi_A}$, $h_B = \eta \mathrm{e}^{\mathrm{j}\phi_B}$, $\eta \in \mathbb{R}^+$ and $\mathbf{w}$ is the AWGN with σ_w^2 variance per dimension. The SNR will be related w.r.t. S_A. The observed frame is of length N.

In this chapter, we consider the special case of a BPSK alphabet $\mathcal{A}_s = \{\pm 1\}$ with the natural H-constellation mapper $s(c) = 2c - 1$ and binary coded symbols $c_A, c_B \in \{0, 1\}$ at both source nodes. In this special case, it follows that the only minimal HNC map is accomplished by the XOR function $c = c_A \oplus c_B$, $c \in \{0, 1\}$.

4.2.1 H-MAC Channel Phase Invariance

Because the target symbol at R is a many-to-one function, multiple source symbol combinations $\mathrm{s}\left(c_A\right), \mathrm{s}\left(c_B\right)$ correspond to one H-code symbol c. This phenomenon is called hierarchical dispersion and, in our observation model, it demonstrates through phase ambiguity. For the useful signal corresponding to the target symbol c $u\left(\boldsymbol{\phi}, c = \chi_c\left(c_A, c_B\right)\right) = \mathrm{e}^{\mathrm{j}\Phi_A}\mathrm{s}\left(c_A\right) + \eta \mathrm{e}^{\mathrm{j}\Phi_B}\mathrm{s}\left(c_B\right)$, it holds that

$$u\left(\boldsymbol{\varphi}, c\right) = u\left(\boldsymbol{\Phi} + \left[\left(2k_1 + 1\right)\pi, \left(2k_2 + 1\right)\right]^{\mathrm{T}}, c\right), \; k_1, k_2 \in \mathbb{Z}. \tag{4.2}$$

The shift in either component of $\boldsymbol{\Phi}$ results in a change of sign of both source constellation points. That is equivalent to a change of sign of both source symbols which does not affect the target symbol because $\chi_c\left(c_A, c_B\right) = \chi_c\left(1 - c_A, 1 - c_B\right)$.

In other words, it means that from the perspective of the relay, we can define a rectangle $\mathbb{P} = \left\{[\varphi_A, \varphi_B] : -\frac{\pi}{2} \leq \varphi_A < \frac{\pi}{2}, -\pi < \varphi_B \leq \pi\right\}$ such that $\boldsymbol{\varphi} \in \mathbb{P}$ are unambiguous w.r.t. c. All other solutions are equivalent from the perspective of H-data-aided CSE. Notice that it does not generally hold for a classical full data c_A, c_B aided estimator.

4.3 H-MAC Phase Estimator

4.3.1 Estimator H-data Aided Metric

Having a memoryless AWGN channel and symbol-wise constellation mappers, the observation likelihood can be factorized as

$$\mathrm{p}_N\left(\mathbf{x}|\boldsymbol{\varphi}, \mathbf{c}_A, \mathbf{c}_B\right)) = \prod_n \mathrm{p}\left(x_n|\boldsymbol{\varphi}, c_{A,n}, c_{B,n}\right). \tag{4.3}$$

With the assumption of known η and the hierarchical H-code symbols $\mathbf{c}$ at R, the final metric is obtained by marginalization over the H-dispersion. Because we assume a minimal HNC map we obtain

$$\mathrm{p}\left(x|\boldsymbol{\varphi}, c\right) = \frac{1}{\mathrm{p}\left(c\right)} \sum_{s_A} \mathrm{p}\left(x|\boldsymbol{\varphi}, s_A, s_B\left(s_A, c\right)\right) \mathrm{p}\left(s_A\right) \mathrm{p}\left(s_B\right). \tag{4.4}$$

Having a BPSK alphabet and i.i.d. source symbols, after some manipulations, we can express the logarithm of scaled (dropping scaling independent on φ_A, φ_B) $\mathrm{p}\left(x|\varphi_A, \varphi_B, c\right)$ as

$$\rho = -\frac{2}{\sigma_w^2}\eta s\left(c\right)\cos\left(\varphi_A - \varphi_B\right) + \ln\cosh\left(\frac{2}{\sigma_w^2}\left(\Re\left\{x\mathrm{e}^{-\mathrm{j}\varphi_A}\right\} + \eta s\left(c\right)\Re\left\{x\mathrm{e}^{-\mathrm{j}\varphi_B}\right\}\right)\right), \tag{4.5}$$

where $s(c) = s_A s_B^* = 1 - 2c$ directly follows from the minimal XOR HNC map. In turn, the corresponding ML estimator is given as

$$\hat{\boldsymbol{\varphi}} = [\hat{\varphi}_A, \hat{\varphi}_B] = \arg\max_{\boldsymbol{\varphi}} \left\{ \rho_N(\boldsymbol{\varphi}, \mathbf{c}) = \sum_{n=1}^{N} \rho(\varphi, c_n) \right\}. \tag{4.6}$$

When assuming a very long frame $N \to \infty$ and equiprobable target symbols $\Pr(c_n = 1) = \Pr(c_n = 0)$ (those assumptions are reasonable, since frame lengths of today's used LDPC codes are in the order of 64800, and a minimal HNC map has indeed equiprobable c), we observe that

$$-\frac{2}{\sigma_w^2}\eta \cos(\varphi_A - \varphi_B) \sum_{n=1}^{N} s(c_n) \to 0, \tag{4.7}$$

while the second term of (4.5) is a monotonically increasing function of N because $\ln(\cosh(x)) \geq 0$, $\forall x$. For the high-SNR regime, $\frac{2}{\sigma_w^2} \gg 1$ we can use the approximation $\ln(\cosh(x)) \approx \ln\left(\frac{1}{2}\right) + |x|$ for $x \gg 1$ and write

$$\rho_N(\boldsymbol{\varphi}) \propto \sum_{n=1}^{N} \left| \left(\Re\left\{ x_n \mathrm{e}^{-\mathrm{j}\varphi_A} \right\} + \eta s(c_n) \Re\left\{ x_n \mathrm{e}^{-\mathrm{j}\varphi_B} \right\} \right) \right|. \tag{4.8}$$

When analyzing (4.8), it can be shown that the metric attains a maximum at the true phase vector $\boldsymbol{\varphi}_0 = [\phi_A, \phi_B]$. If, in addition, the condition

$$\left(\eta^2 + 1\right) \cos(\phi_A - \phi_B) \pm 2\eta \gtreqless 0 \tag{4.9}$$

is fulfilled, there exists a second false maximum at $\boldsymbol{\varphi}_1 = [\phi_B, \phi_A]$. Further, it can be shown, that

$$\rho_N(\boldsymbol{\varphi}_0) > \rho_N(\boldsymbol{\varphi}_1) \text{ for } \eta < 1 \tag{4.10}$$

$$\rho_N(\boldsymbol{\varphi}_0) = \rho_N(\boldsymbol{\varphi}_1) \text{ for } \eta = 1. \tag{4.11}$$

In the singular case when $\eta = 1$, it holds that

$$u\left([\phi_A, \phi_B]^{\mathrm{T}}, \chi_c(c_A, c_B) \right) = u\left([\phi_B, \phi_A]^{\mathrm{T}}, \chi_c(1 - c_A, 1 - c_B) \right) \tag{4.12}$$

and in turn

$$u(\boldsymbol{\varphi}_0, c) = u(\boldsymbol{\varphi}_1, c). \tag{4.13}$$

In summary, it is possible to distinguish the true maximum from the false one based on the value of the metric unless $\eta = 1$ in which case both solutions are equivalent w.r.t. the hierarchical target symbol c.

4.3.2 Feed-Back Gradient Solver

The extreme of ρ_N can be found using a feed-back iterative solver with additive updates according to

$$\hat{\boldsymbol{\varphi}}(i+1) = \hat{\boldsymbol{\varphi}}(i) + K\boldsymbol{\mu}_N, \tag{4.14}$$

where K is the properly chosen step-size and $\boldsymbol{\mu}_N = \nabla_{\boldsymbol{\varphi}} \rho_N(\hat{\boldsymbol{\varphi}}(i), \boldsymbol{c})$ is the gradient with symbol-wise components $\mu_{N,A} = \frac{\partial \rho_N}{\partial \varphi_A}$ and $\mu_{N,B} = \frac{\partial \rho_N}{\partial \varphi_B}$. After carrying out the partial derivatives, we get (see details in [108])

$$\begin{aligned} \mu_{N,A} = \sum_{n=1}^{N} & \frac{2}{\sigma_w^2} \eta s(c_n) \sin(\varphi_A - \varphi_B) + \frac{2}{\sigma_w^2} \Im\left\{ x_n \mathrm{e}^{-\mathrm{j}\varphi_A} \right\} \\ & \times \tanh\left(\frac{2}{\sigma_w^2} \left(\Re\left\{ x_n \mathrm{e}^{-\mathrm{j}\varphi_A} \right\} + \eta s(c_n) \Re\left\{ x_n \mathrm{e}^{-\mathrm{j}\varphi_B} \right\} \right) \right), \end{aligned} \tag{4.15}$$

$$\begin{aligned}\mu_{N,B} = \sum_{n=1}^{N} \frac{2}{\sigma_w^2}\eta s\left(c_n\right)\sin\left(\varphi_B - \varphi_A\right) + \frac{2}{\sigma_w^2}\eta\Im\left\{x_n \mathrm{e}^{-\mathrm{j}\varphi_B}\right\} \\ \times \tanh\left(\frac{2}{\sigma_w^2}\left(\eta\Re\left\{x_n\mathrm{e}^{-\mathrm{j}\varphi_B}\right\} + s\left(c_n\right)\Re\left\{x_n\mathrm{e}^{-\mathrm{j}\varphi_A}\right\}\right)\right).\end{aligned} \tag{4.16}$$

Alternatively, we can consider a more practical scenario, where the phase $\boldsymbol{\varphi}_n = [\varphi_{A,n}, \varphi_{B,n}]$ varies during the frame. For such a model, we modify the full-frame gradient descent search and iterate over $\hat{\boldsymbol{\varphi}}_n$, in each step applying a correction based on the gradient of a local neighborhood given by a window length W.

$$\hat{\boldsymbol{\varphi}}_{i+1} = \hat{\boldsymbol{\varphi}}_i + K\nabla_{\tilde{\boldsymbol{\varphi}}}\left(\sum_{n=i-W-1}^{i-1}\rho\left(\varphi_{A,n}, \varphi_{B,n}, c_n, x_n\right)\right)\Bigg|_{\tilde{\boldsymbol{\varphi}}=\hat{\boldsymbol{\varphi}}_i}. \tag{4.17}$$

The window length should reflect the rate of phase change and the knowledge of the initial phase i.e. $\hat{\boldsymbol{\varphi}}_1 = \boldsymbol{\varphi}_1$ is assumed.

4.4 Over-the-Air Verification

In this section, we evaluate the performance of the proposed estimator based on gradient tracking (4.17) using an over-the-air transmission and compare it with simulation results. For this purpose, we considered multiple scenarios and evaluated the mean squared error (MSE) of the phase estimates and H-BER of the hierarchical message $\mathbf{c}$ for different values of SNR. The final goal of the relay processing is an estimate of the hierarchical $\mathbf{c}$. Under such conditions, the classical MSE of the form $\mathrm{E}\left[\left(\boldsymbol{\varphi} - \hat{\boldsymbol{\varphi}}\right)^2\right]$ is not suitable because the many-to-one χ function introduces a many-to-one relationship between the channel phase and observation as indicated in (4.2). To account for this, we evaluate a modified MSE using

$$\min_{k_1,k_2}\mathrm{E}\left[\left(\boldsymbol{\varphi} - \hat{\boldsymbol{\varphi}} + \left[\left(2k_1+1\right)\pi, \left(2k_2+1\right)\pi\right]\right)^2\right]. \tag{4.18}$$

4.4.1 Considered Scenarios

4.4.1.1 True Hierarchical Data-Aided Estimation (scenario A)

In the first setting, the estimator has full access to the hierarchical target message $\mathbf{c}$. This scenario, although it can not be practically implemented, should produce the best result and set a performance limit.

4.4.1.2 Hierarchical Data Estimation Based on Hard Decision (scenario B)

The second case resembles a real scenario, where the hierarchical target message is not available at the relay and the phase estimator uses only an estimate $\hat{\mathbf{c}}$. The estimate was obtained by a hard decision

$$\hat{c}_n = \begin{cases} 1 & \log\left(\frac{\mathrm{p}(x_n|c_n=0)}{\mathrm{p}(x_n|c_n=1)}\right) \geq 0 \\ 0 & \log\left(\frac{\mathrm{p}(x_n|c_n=0)}{\mathrm{p}(x_n|c_n=1)}\right) < 0. \end{cases} \tag{4.19}$$

The values of c_n are equiprobable which allows us to marginalize over the hierarchical dispersion as

$$\begin{aligned} \mathrm{p}\left(x_n|c\right) &\propto \sum_{(c_A,c_B):\chi(c_A,c_B)=c} \mathrm{p}\left(x_n|c_A,c_B\right)\mathrm{p}\left(c_A,c_B\right) \\ &\propto \sum_{(c_A,c_B):\chi(c_A,c_B)=c} \exp\left(\left|x_n - u_n\left(c_A,c_B\right)\right|^2\right). \end{aligned} \tag{4.20}$$

For details see [111].

For this scenario, the gradient tracking of the channel phase given in (4.17) needs to be modified to

$$\hat{\boldsymbol{\varphi}}_{i+1} = \hat{\boldsymbol{\varphi}}_i + K\nabla_{\check{\boldsymbol{\varphi}}}\left(\sum_{n=i-W-1}^{i-1} \rho\left(\varphi_{A,n},\varphi_{B,n},\hat{c}_n,x_n\right)\right)\Bigg|_{\check{\boldsymbol{\varphi}}=\hat{\boldsymbol{\varphi}}_i}, \tag{4.21}$$

where estimates $[\hat{c}_n]_{n=i-W-1}^{i-1}$ are obtained by (4.19) and (4.20) using the past phase estimates $[\hat{\varphi}_n]_{n=i-W-1}^{i-1}$ to form the useful signal values $[u_n\left(c_A,c_B\right)]_{n=i-W-1}^{i-1}$.

4.4.2 Over-the-air implementation

For the evaluation of the estimator performance in a real-world setting, we used three software-defined radio transceivers. Each for one node S_A, S_B, and R. The particular type used in our experiments was the *Ettus Research N210* transceiver. The experimental setting was made up of three stationary antennas placed in a time-invariant office environment according to Figure 4.1. For carrier frequency and symbol timing synchronization, the *OctoClock* clock distribution module was employed. The parameter setup used for our transmissions is summarized in Table 4.1.

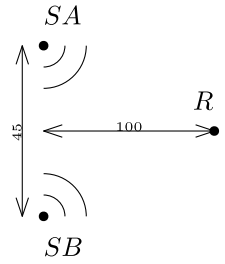


Figure 4.1: Antenna arrangement.

parameter	value
carrier frequency	2.4 GHz
sampling frequency	10 MHz
modulation pulse	RRC
roll-off factor	0.35
samples per symbol	8

Table 4.1: Settings used for over-the-air transmissions.

Since the phase estimator is unbiased, the MSE measurement of the phase estimate could be implemented by computing the sample variance of the estimation error. In our measurement, we used an invariant environment and a perfect synchronization of all three nodes (synchronization of the symbol timing as well as the carrier frequency) such that the channel phases could be assumed constant during one measurement realization. To make the channel phase variable while allowing

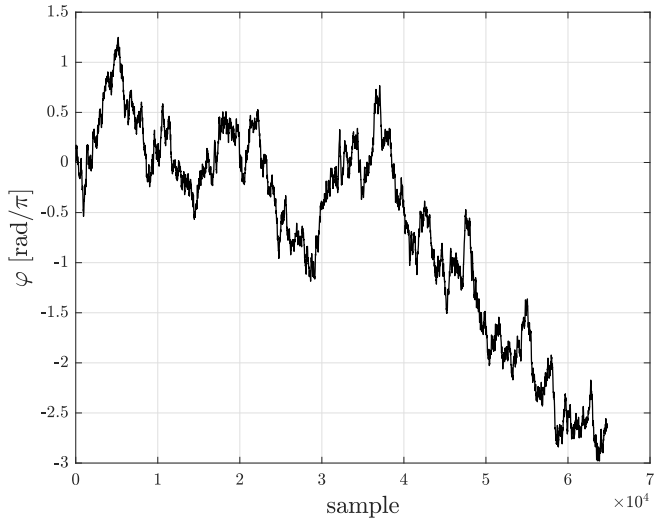


Figure 4.2: Example of generated phase time dependence.

us to compute the estimation error, we modeled the variable phase $\varphi'_{A,n}, \varphi'_{B,n}$ as a random walk process and embedded it artificially in the generated signals such that

$$\boldsymbol{\varphi}_n = \left[\angle h_A + \varphi'_{A,n}, \angle h_B + \varphi'_{B,n}\right]. \tag{4.22}$$

In Figure 4.2 we show a typical realization of the generated phase. Coefficients h_A and h_B are constant channel coefficients of the real OTA channel which were estimated using orthogonal pilot sequences $\boldsymbol{p}_A, \boldsymbol{p}_B$ and a standard ML estimator

$$\begin{aligned} \angle h_A &= \angle \langle \boldsymbol{x}, \boldsymbol{p}_A \rangle \\ |h_A| &= \frac{|\langle \boldsymbol{x}, \boldsymbol{p}_A \rangle|}{\|\boldsymbol{p}_A\|} \end{aligned} \tag{4.23}$$

$$\hat{\sigma}_w^2 = \operatorname{var}\left(\boldsymbol{x} - h_A \boldsymbol{p}_A\right) \tag{4.24}$$

at the start of every measurement realization. A suitable normalization was applied to comply with the channel model (4.1) ($|h_A| = 1$). Desired values of η and SNR of the normalized channel were achieved by an appropriate selection of transmit gains at both sources S_A and S_B.

4.5 Results

In this section, we present the performance evaluation of the phase estimator (4.17) in terms of its estimate accuracy as well as in terms of the resulting H-BER when used together with a ML demodulator (4.20). We present two types of results. One method involves a numerical simulation of the channel model (4.1), while the second is based on a real OTA transmission according to the description in Section 4.4.2. The observations of those two sources are passed to the phase estimator under two different scenarios. In the first scenario (scenario A, described in Section 4.4.1.1), the estimator is aided by artificially supplied hierarchical target symbols c_n. In the second, and more practical scenario (scenario B, described in Section 4.4.1.2), the phase estimator is aided only by estimates $\hat{c}_n$ according to (4.19).

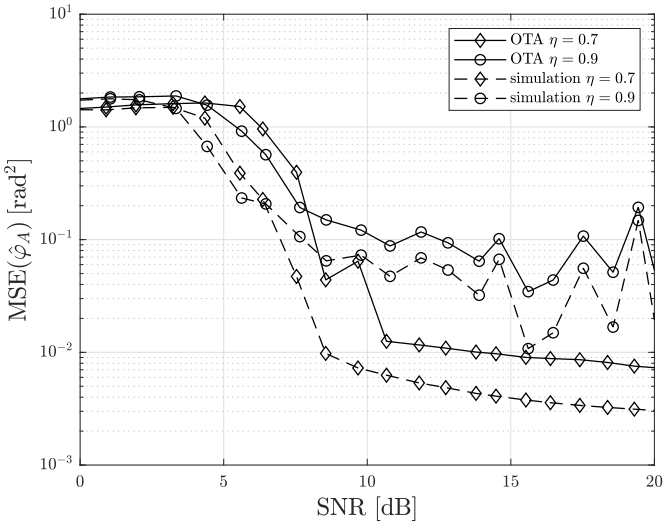


Figure 4.3: Results of simulation and OTA under scenario B (Section 4.4.1.2). We plot the SNR dependence of the $\hat{\varphi}_A$ MSE according to (4.18) for $\eta \in \{0.7, 0.9\}$.

In Figure 4.3, we present the evaluation of the MSE according to (4.18) for different values of SNR. For clarity, we plot only results for the estimate $\hat{\varphi}_A$. We observe, that the simulation performs reasonably close to the OTA. It may be surprising, that from 10dB, the OTA MSE exhibits a plateau and does not decrease with rising SNR. The explanation for this behavior can be found in Section 4.3.1, namely the presence of a false maximum of the metric. To eliminate this effect, we define a modified version of the MSE such that we compute two versions of the square error, the direct $(\hat{\varphi}_A - \varphi_A)^2, (\hat{\varphi}_B - \varphi_B)^2$ and the flipped $(\hat{\varphi}_B - \varphi_A)^2, (\hat{\varphi}_A - \varphi_B)^2$, and select the smaller one. This modified MSE will be denoted with a star as MSE*.

In Figure 4.4 we show the results of the same experiment as in Figure 4.3, now employing MSE*. It can be seen, that the plateau of the OTA MSE dissapeared.

For completeness, Figure 4.5 shows the MSE* for both phases estimates under $\eta = 0.5$. We observe a better result for the stronger signal in both cases, the simulation and OTA which is in accordance with the Cramér–Rao bound analysis presented in [52].

We also evaluated the performance difference between scenarios A and B, that is the two cases where we supplied correct hierarchical data to the estimator and where we used hard-decision estimates only. The results are presented in Figure 4.6, which shows that the aided case performs better, as anticipated. For high SNR however, both curves merge as the hard-decision error probability drops.

Finally, in Figure 4.7 we present the attained H-BER under scenario B. We conclude that there is no noticeable performance gap between the simulation and the OTA and thus, the proposed method is reasonably robust w.r.t. violations to the observation model due to hardware imperfections.

4.6 Conclusion

This chapter was motivated by the sensitivity of the hierarchical demodulation on the knowledge of the channel state parameters and the need for an accurate channel state estimate. In Section 4.2 we described our system model and stated our assumptions. Section 4.3 contains a summary of our previous work, published in [51, 53, 52, 108], devoted to a hierarchical data-aided channel

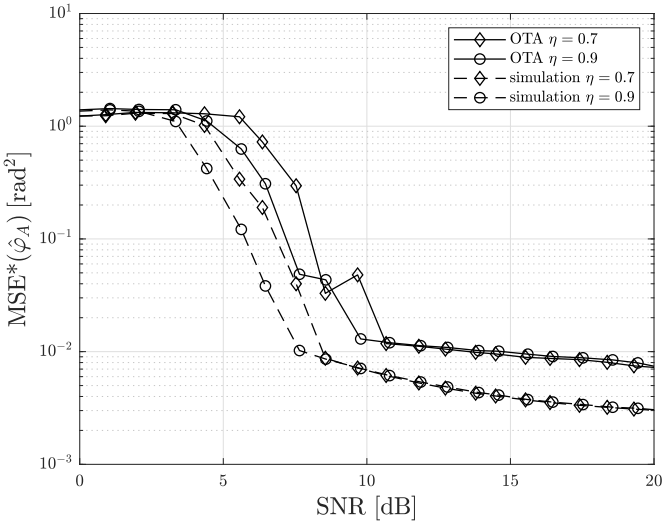


Figure 4.4: Results of simulation and OTA under scenario B (Section 4.4.1.2). We plot the SNR dependence of the $\hat{\varphi}_A$ MSE* for $\eta \in \{0.7, 0.9\}$.

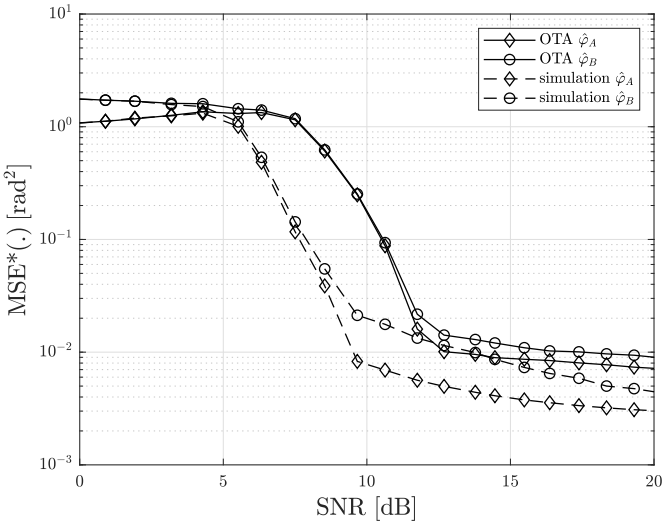


Figure 4.5: Comparison of simulation and OTA experiment in terms of MSE* of both $\hat{\varphi}_A, \hat{\varphi}_B$, and $\eta = 0.5$ under scenario B (Section 4.4.1.2).

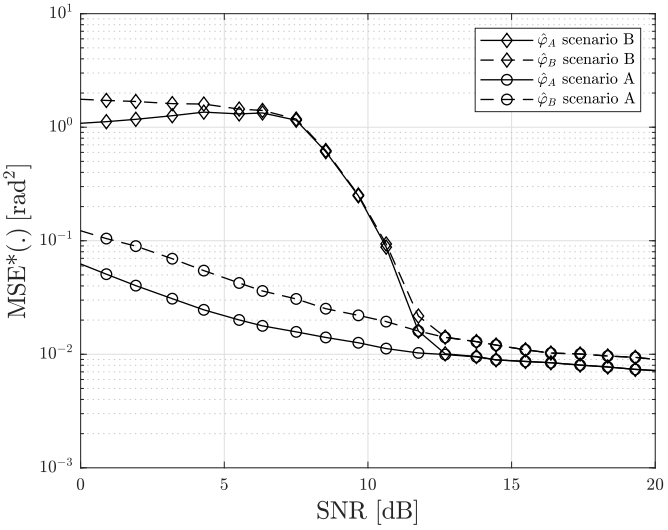


Figure 4.6: Comparison of the MSE* performance of the OTA experiment under scenarios A (Section 4.4.1.1) and B (Section 4.4.1.2) for $\eta = 0.5$.

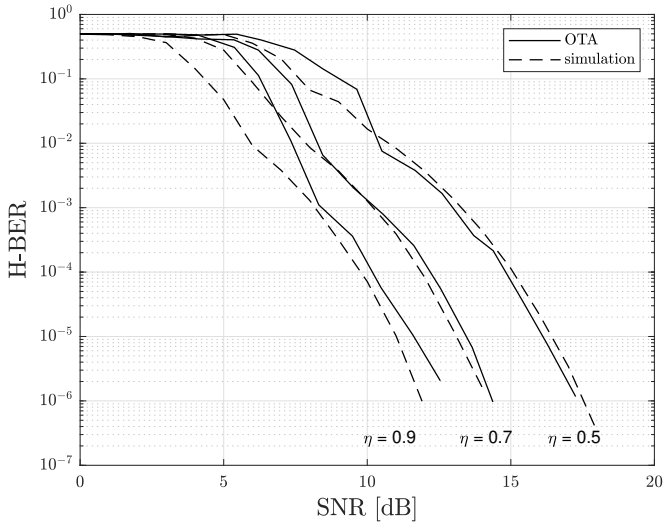


Figure 4.7: Comparison of H-BER performance attained by the simulation and OTA under scenario B (Section 4.4.1.2) for different values of $\eta \in \{0.5, 0.7, 0.9\}$.

phase estimator and a gradient-search based solver together with the analysis of the corresponding metric.

The contribution was presented in Section 4.4. First, we described the setup of the experimental network as well as the methodology. In particular, we defined a genie-aided reference scenario, with perfect hierarchical data aid, and a practical case without any external help. To measure the performance of the CSE, we defined two variants of MSE which respect the many-to-one nature of the HNC map and thus give meaningful results in the H-MAC context. To verify the CSE under real-life conditions, we run multiple simulations and OTA experiments and compared their respective performance in terms of the previously mentioned MSE measures and achievable H-BER rates.

The main observation is that while there is a significant performance gap between the simulation and OTA experiment in terms of the MSE, the resulting H-BER rates performance is comparable. Based on this observation we conclude, that a receiver implementation based on our CSE is robust w.r.t. violations of the observation model due to hardware imperfections.

Chapter 5

Differential Hierarchical Demodulator of 2-source Hierarchical-MAC Channel

5.1 Introduction

We focus on the application of the differential modulation principle to a non-orthogonal H-MAC scenario as a way of dealing with the effects of relative parametrization as discussed in Section 3.3.3.1. The major difference between a single-user case and the H-MAC is the phenomenon of relative channel parametrization. There is a nonlinear impact on the hierarchical constellation in the H-MAC case, while in the single-user case, it results in simple rotation. In the particular case of differential modulation, there is a difference in the number of degrees of freedom. While in the single-user case, there is only one phase parameter that gets completely eliminated by the differential encoding. In the H-MAC, on the other hand, there is a phase parameter connected with each user. In the case of a 2-user H-MAC, there are two independent phases that can not be unambiguously eliminated by the application of differential encoding.

In this chapter we propose several demodulation methods for a two-user H-MAC and differential modulation which form the main contribution. For evaluating their performance we run several numerical simulations as well as an over-the-air transmission. We show that for the BPSK alphabet, the proposed closed-form Sequential ML demodulator performs roughly 3 dB away from the reference case of a fully coherent setting and thus performs similarly as a single-user channel with differential modulation.

The work presented in this chapter is based on my published as well as unpublished research, especially [54].

5.1.1 Motivation & Related Work

Most of the results on fundamental limits, system design, or performance analysis of WPNC networks assume relatively idealized assumptions related to relative channel parametrization and NCM codebooks used at source nodes. The maximum likelihood (ML) decision regions depend heavily on the relative channel phase in a nonlinear way. Because of the highly interacting signals, however, the channel state estimation is complex and we cannot simply apply methods used in single-user cases [108, 40]. This motivates the research on a non-coherent reception for an H-MAC. One approach is to use a nonlinear FSK modulation which we can find in [116], where the authors studied multiple scenarios with various assumptions on the channel state information (CSI) knowledge at

the receiver. A particular demodulation procedure of a hierarchical BFSK modulation on a relay with the HDF strategy can be found in [104]. Another approach is to use a linear PSK modulation with a differential precoder. In [134, 103, 25] the authors assumed a two-way relay channel (TWRC) with a differential PSK modulation and the AF strategy at the relay. The authors of [25] additionally considered the HDF strategy and derived a hierarchical MAP (maximum a posteriori probability) demodulator for a Rayleigh fading channel.

5.2 System Model

We consider a single MAC stage with two source nodes S_A, S_B, and one relay node R. All three nodes are assumed to have perfect symbol-timing synchronization. The observation model at R is a memoryless constellation space AWGN channel given as

$$x_n = h_{A,n} s_{A,n}\left(q\left(\boldsymbol{c}_{A,n}\right)\right) + h_{B,n} s_{B,n}\left(q\left(\boldsymbol{c}_{B,n}\right)\right) + w_n \tag{5.1}$$

as shown in Figure 5.1. The input to each transmitter node is an encoded binary sequence $\boldsymbol{c}$, which can be treated as a vector of i.i.d. uniformly distributed random variables. The source node processing consists of a one-to-one channel symbol mapper $q\left(.\right)$ which takes the nth M tuple from the codeword $\boldsymbol{c}$ and maps it to a channel symbol from the alphabet $\mathcal{A}_q$, and constellation point mapper $s\left(.\right)$ which maps the channel symbols to constellation points in the complex plain. We will consider only PSK channel symbol alphabets given as $\mathcal{A}_q = \exp\left\{\left(\frac{2\pi \mathrm{j}}{M} i\right)\right\}_{i=0}^{M-1}$. We assume a flat fading channel with complex fading coefficients $h_{A,n} = \eta_A \mathrm{e}^{\mathrm{j}\angle h_{A,n}}$, $h_{B,n} = \eta_B \mathrm{e}^{\mathrm{j}\angle h_{B,n}}, \eta_A, \eta_B \in \mathbb{R}^+$. Note that the fading channel magnitudes η_A, η_B are assumed constant w.r.t. a frame duration and known to the relay. Finally, white Gaussian noise w with a variance σ_w^2 per dimension is added to form the observation x. The SNR will be related w.r.t. S_A as $\mathrm{E}\left[|s_A|^2\right]/\sigma_w^2$.

We conceive this setting as a part of a bigger radio network with an *isomorphic layered* NCM (for example a TWRC or a butterfly network) [111] and thus, at node R, we only aim to decode the hierarchical target codeword symbol $\boldsymbol{c}_n$ which is given as a many-to-one function (Hierarchical Network Code (HNC) map) $\boldsymbol{c}_n = \chi_c\left(\boldsymbol{c}_{A,n}, \boldsymbol{c}_{B,n}\right)$. In particular, we assume a *minimal* HNC map χ_c given by the bit-wise XOR function.

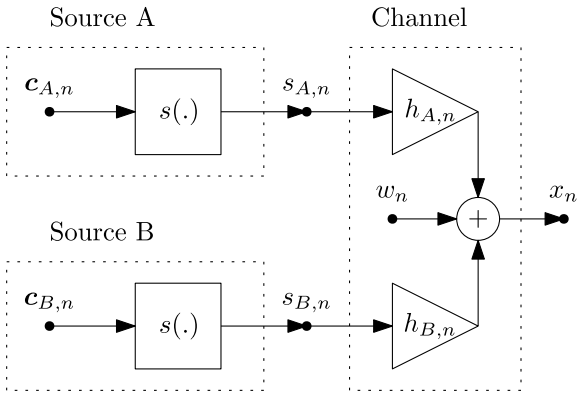


Figure 5.1: Schematic representation of constellation space H-MAC.

5.2.1 Differential Encoding

We propose to apply a differential encoding at the source nodes according to $s_{n,A} = s_{A,n-1}q_{,A,n}$ and $s_{n,B} = s_{B,n-1}q_{,B,n}$. A schematic representation can be found in Figure 5.2. Because the channel symbol mapper is one-to-one, $\boldsymbol{c}_n$ are equivalent to q_n. As such we will perform our analysis on the channel symbol level. The one-to-one relationship allows us to define a hierarchical target channel symbol and a channel symbol hierarchical map

$$q_n = q\left(\boldsymbol{c}_n\right) = \chi_q\left(q\left(\boldsymbol{c}_{A,n}\right), q\left(\boldsymbol{c}_{B,n}\right)\right). \tag{5.2}$$

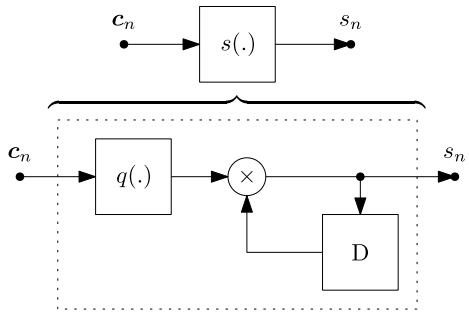


Figure 5.2: Schematic representation of a differential encoder. The block denoted by D performs a delay of one symbol.

5.2.2 Two-point Observation

At R, the observation consists of two superimposed differentially modulated signals according to (5.1). For the proposed demodulation methods, however, it is sufficient to consider a pair of two consecutive observations which we call the *two-point observation*

$$\begin{aligned} x_0 &= \underbrace{\eta_A \mathrm{e}^{\mathrm{j}\varphi_{A,0}} + \eta_B \mathrm{e}^{\mathrm{j}\varphi_{B,0}}}_{u_0} + w_0 \\ x_1 &= \underbrace{\eta_A \mathrm{e}^{\mathrm{j}\varphi_{A,1}} q_A + \eta_B \mathrm{e}^{\mathrm{j}\varphi_{B,1}} q_B}_{u_1} + w_1, \end{aligned} \tag{5.3}$$

where u_0, u_1 are the useful signal of observation 0 and 1 respectively. Phases $\varphi_{A,n}$ and $\varphi_{B,n}$ are given as the sum of the channel coefficient angles $\angle h_{A,n}, \angle h_{B,n}$ and all the preceding symbol phases. Their difference within one two-point observation $(\varphi_{A,0} - \varphi_{A,1}, \varphi_{B,0} - \varphi_{B,1})$ is given only by the change of the channel phases from one symbol to the next one. We expect this change to be negligible and continue with the approximation

$$\begin{aligned} \varphi_{A,0} &= \varphi_{A,1} \\ \varphi_{B,0} &= \varphi_{B,1}. \end{aligned} \tag{5.4}$$

5.3 Noisy Reference

Our first approach is motivated by the single-user differential PSK demodulation, where differential precoding at the transmitter allows for non-coherent demodulation (that is the signal is demodulated without the need for any channel phase estimation). The demodulation of one symbol requires two observations and the symbol is estimated from their phase difference. From a different point of view one observation can be understood as a reference and the second as a standard PSK demodulation w.r.t. the reference. This other way of understanding is the base of the extension to the two-user case.

5.3.1 Phase Reference

For the matter of the derivation we consider a noiseless situation (that is we work only with the useful signal of the two-point observation u_0, u_1). After separating the magnitude and phase we have

$$|u_0|\,\mathrm{e}^{\mathrm{j}\angle u_0} = \eta_A \mathrm{e}^{\mathrm{j}\varphi_A} + \eta_B \mathrm{e}^{\mathrm{j}\varphi_B} \tag{5.5}$$

$$|u_1|\,\mathrm{e}^{\mathrm{j}\angle u_1} = \eta_A \mathrm{e}^{\mathrm{j}(\varphi_A+\Delta_A)} + \eta_B \mathrm{e}^{\mathrm{j}(\varphi_B+\Delta_B)}, \tag{5.6}$$

where $\Delta_A = \angle q_A, \Delta_B = \angle q_B$. Both q_A and q_B are elements of a PSK alphabet and as such, it is sufficient to work with their phases Δ_A, Δ_B.

Now we analyze the number of degrees of freedom of this set of equations. The 2-point observation gives us 4 real-valued equations and 6 unknowns - $\eta_A, \eta_B, \varphi_A, \varphi_B, \Delta_A$ and Δ_B. Because we assume the channel coefficient magnitudes η_A, η_B to be constant and known at the receiver, we are left with only four unknowns. Having the same number of equations as degrees of freedom allows us to solve the system. In practice, the amplitudes would be typically estimated once and remained fairly constant over the rest of the frame. This cannot be assumed about the phases $\angle h_{A,n}, \angle h_{B,n}$ as they change even in a static environment due to imperfections of local oscillators.

The first step is to solve (5.5) for φ_A, φ_B. For this, we multiply it by $\mathrm{e}^{-\mathrm{j}\angle u_0}$which results in

$$|u_0| = \eta_A \mathrm{e}^{\mathrm{j}(\varphi_A - \angle u_0)} + \eta_B \mathrm{e}^{\mathrm{j}(\varphi_B - \angle u_0)} = \eta_A \mathrm{e}^{\mathrm{j}\alpha} + \eta_B \mathrm{e}^{\mathrm{j}\beta}. \tag{5.7}$$

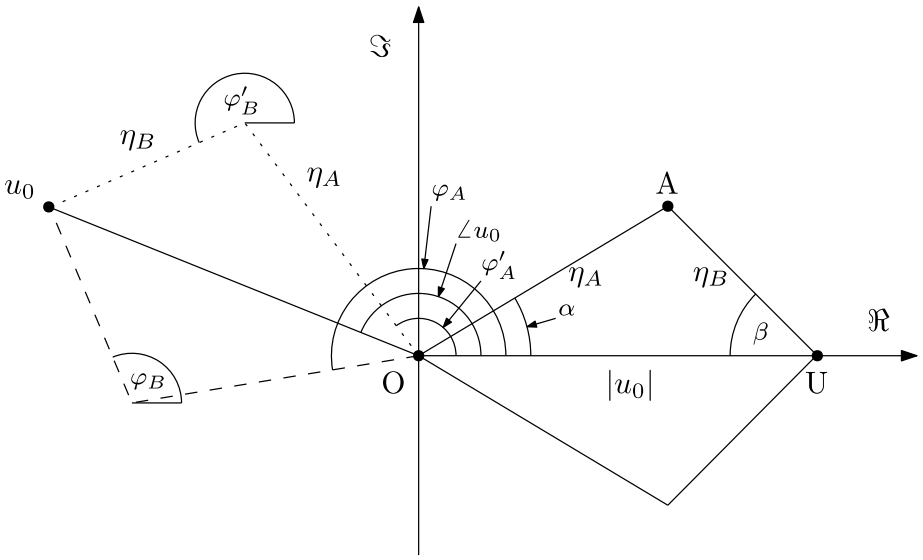


Figure 5.3: Decomposition of the observation u_0 into two components of known magnitudes (η_A, η_B) and its rotated copy.

Equation (5.7) is geometrically expressed in Figure 5.3. In Figure 5.3 we plot the useful signal u_0 in the complex plane. Since we assume to know η_A and η_B, there are generally two possible ways how to decompose u_0 into a sum of two contributions of given absolute values. In the figure, the two solutions are shown with a dotted and dashed line respectively. We note that there is a pair of phases φ_A, φ_B associated to each solution (the latter being denoted with a prime). After a rotation (multiplication by $\mathrm{e}^{-\mathrm{j}\angle u_0}$), equation (5.7) can be represented by the quadrangle on the right side of the figure. With the help of auxiliary angles α, β we can use a geometric argument

and the law of cosine on the triangle OAU to get

$$\begin{aligned}\cos(\alpha) &= \frac{|u_0|^2 + \eta_A^2 - \eta_B^2}{2\eta_A |u_0|} \\ \cos(\beta) &= \frac{|u_0|^2 + \eta_B^2 - \eta_A^2}{2\eta_B |u_0|},\ |u_0| \neq 0\end{aligned} \tag{5.8}$$

and in turn

$$\begin{aligned}\alpha(|u_0|) &= \arccos\left(\frac{|u_0|^2 + \eta_A^2 - \eta_B^2}{2\eta_A |u_0|}\right) \\ \beta(|u_0|) &= \arccos\left(\frac{|u_0|^2 + \eta_B^2 - \eta_A^2}{2\eta_B |u_0|}\right),\ |u_0| \neq 0.\end{aligned} \tag{5.9}$$

The solution for α and β is simple because we know all the distances η_A, η_B and $|u_0|$. The two possible solutions are given as

$$\begin{aligned}[\varphi_A, \varphi'_A] &= [\angle u_0 + \alpha(|u_0|), \angle u_0 - \alpha(|u_0|)] \\ [\varphi_B, \varphi'_B] &= [\angle u_0 - \beta(|u_0|), \angle u_0 + \beta(|u_0|)].\end{aligned} \tag{5.10}$$

The only case when we get a single, unambiguous solution ($\varphi'_A = \varphi_A, \varphi'_B = \varphi_B$) occurs when

$$\alpha = 0 \mod \pi \tag{5.11}$$

$$\beta = 0 \mod \pi. \tag{5.12}$$

As can be seen from Figure 5.3, (5.11) and (5.12) are equivalent and thus we either get a single solution for both phases or for neither of them. In most cases, the reference observation u_0 will give us a pair of correct and a pair of false channel phases which we need to reflect in the subsequent processing.

Example 5.1. Let us analyze an example with a BPSK alphabet and the XOR HNC map χ_c. Given the observation u_0, we get a pair of phases φ_A, φ_B which correspond to the dashed constellation in Figure 5.4, and the primed pair φ'_A, φ'_B with the corresponding constellation shown with dotted lines. We do not know which constellation was used by the transmitters but from the point of view of the hierarchical target symbol c_n (or equivalently the target channel symbol q_n) the relay needs only to distinguish between the points corresponding to $c_n = 1$ (denoted by crosses) and points corresponding to $c_n = 0$ (denoted by squares). In the example in Figure 5.4, we see that points corresponding to $c_n = 0$ coincide for both constellation variants. In general, this is not always the case but we expect them to be close as it is with the points corresponding to $c_n = 1$. The shaded area indicates a possible decision region for c_n assuming both constellation variants are equally likely.

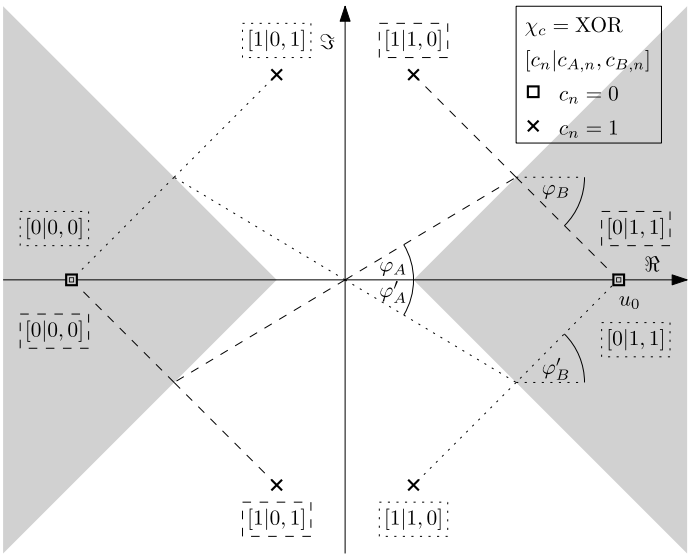


Figure 5.4: Figure for Example 5.1.

So far we ignored the presence of the noise term w_0. To make this approach practical and still retain the closed-form solution, we simply substitute $u_0 = x_0$ and understand the results as approximate estimates $\hat{\varphi}_A, \hat{\varphi}_B$. We cannot make the simple substitution of $u_0 = x_0$ because equation (5.7) might not have a solution. The reason is that if for some noise realizations $|x_0| > \eta_A + \eta_B$ or $|x_0| < |\eta_A - \eta_B|$, it cannot be decomposed as a sum of two vectors of magnitudes η_A, η_B. That is why we first need to map the noisy observation x_0 to $x'_0 \in \mathbb{A}$, where $\mathbb{A}$ is an annulus given as

$$\mathbb{A} = \{x \in \mathbb{C} | \, |\eta_A - \eta_B| = r_1 \leq |x| \leq |\eta_A + \eta_B| = r_2\}. \tag{5.13}$$

An example of $\mathbb{A}$ is shown in Figure 5.5. Because $u_0 \in \mathbb{A}$ and $x_0 = u_0 + w_0$, the most likely estimate of u_0 given x_0 is the projection of x_0 onto $\mathbb{A}$. Such a projection (Figure 5.5) x'_0 is given as

$$\begin{aligned} \angle x'_0 &= \angle x_0 \\ |x_0| &= \begin{cases} |\eta_A - \eta_B| & |x_0| < |\eta_A - \eta_B| \\ |x_0| & |\eta_A - \eta_B| \leq |x_0| \leq |\eta_A + \eta_B| \\ |\eta_A + \eta_B| & |\eta_A + \eta_B| < |x_0|. \end{cases} \end{aligned} \tag{5.14}$$

From (5.10) and (5.9) we can observe the dependence of the phase estimates on the magnitude and the phase of u_0. The phase $\angle u_0$ appears in the estimate directly (5.10) and the magnitude $|u_0|$

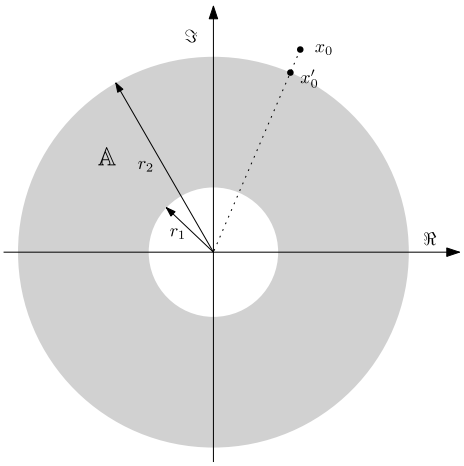


Figure 5.5: Visualization of the annulus $\mathbb{A}$ (5.13) and the projection (5.14).

undergoes a nonlinear transformation and defines the angles α and β. After the substitution of the noisy projection x_0', we have

$$\begin{aligned} |x_0'| &= |u_0| + \xi\left(u_0, w_0\right) \\ \angle x_0' &= \angle u_0 + \psi\left(u_0, w_0\right). \end{aligned} \tag{5.15}$$

To evaluate the impact of $\xi\left(u_0, w_0\right)$ and $\psi\left(u_0, w_0\right)$ on the estimates we can consider a Taylor expansion of (5.9) around $\alpha\left(|u_0|\right)$ and $\beta\left(|u_0|\right)$ and derive the first derivatives

$$\begin{aligned} \frac{\partial \varphi_A}{\partial \angle u_0} &= 1 \\ \frac{\partial \varphi_A}{\partial \left|u_0\right|} &= \frac{-1}{\sqrt{1-\left(\frac{|u_0|^2+\eta_A^2-\eta_B^2}{2\eta_A|u_0|}\right)^2}} \frac{\left|u_0\right|^2-\eta_A^2+\eta_B^2}{2\eta_A\left|u_0\right|^2} \end{aligned} \tag{5.16}$$

$$\begin{aligned} \frac{\partial \varphi_B}{\partial \angle u_0} &= 1 \\ \frac{\partial \varphi_B}{\partial \left|u_0\right|} &= \frac{1}{\sqrt{1-\left(\frac{|u_0|^2+\eta_B^2-\eta_A^2}{2\eta_B|u_0|}\right)^2}} \frac{\left|u_0\right|^2-\eta_B^2+\eta_A^2}{2\eta_B\left|u_0\right|^2}. \end{aligned} \tag{5.17}$$

We see that the derivatives w.r.t. $\angle u_0$ are constant and do not depend on the actual useful signal u_0. The derivatives w.r.t. $|u_0|$ are dependent on $|u_0|$ as shown in Figure 5.6. We note that in the limiting case where $|u_0|$ goes to $\eta_A + \eta_B$ or $|\eta_A - \eta_B|$ the absolute value of the derivatives goes to infinity and we expect the resulting phase estimates to be poor.

Remark. It is interesting to note the correspondence between $\beta = \frac{\pi}{2}$ in Figure 5.3 and $\frac{\partial \varphi_A}{\partial |u_0|} = 0$.

5.3.2 Demodulation

Now we use the noisy estimates $\hat{\varphi}_A$ and $\hat{\varphi}_B$ obtained from the first observation x_0 to form a hierarchical maximum likelihood (ML) estimator. The phase estimate ambiguity gives us one extra degree of freedom compared to the standard H-ML situation. Because both variants of the phase estimates are equally likely conditioned by x_0 we eliminate it by an additional marginalization with

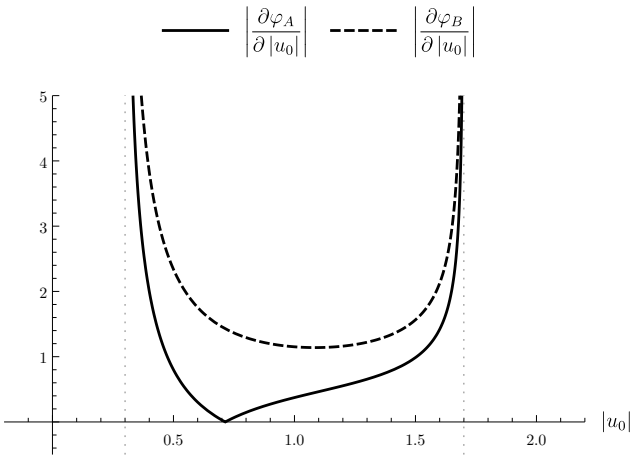


Figure 5.6: Evaluation of the absolute value of (5.16) and (5.17) for $\eta_A = 1, \eta_B = 0.7$.

a uniform prior. For the hierarchical target channel symbol we write

$$\hat{q} = \underset{q}{\operatorname{argmax}} \sum_{q_A,q_B:\chi_q(q_A,q_B)=q} \mathrm{p}\left(x_1|q_A,q_B,\hat{\varphi}_A,\hat{\varphi}_B\right) + \sum_{q_A,q_B:\chi_q(q_A,q_B)=q} \mathrm{p}\left(x_1|q_A,q_B,\hat{\varphi}'_A,\hat{\varphi}'_B\right), \quad (5.18)$$

where $\mathrm{p}\left(.\right)$ is the likelihood function given up to unimportant scaling as

$$\mathrm{p}\left(x_1|q_A,q_B,\hat{\varphi}_A,\hat{\varphi}_B\right) = \exp\left(\frac{-1}{\sigma_w^2}\left|x_1 - \eta_A \mathrm{e}^{\mathrm{j}(\hat{\varphi}_A+\Delta_A)} - \eta_B \mathrm{e}^{\mathrm{j}(\hat{\varphi}_B+\Delta_B)}\right|^2\right). \quad (5.19)$$

An exchange of x_0 and x_1 results in a sign flip of Δ_A, Δ_B or equivalently performs a complex conjugate of q_A and q_B. This means the reference observation used for the phase estimation can be selected arbitrarily. To distinguish which observation was used as the reference we add additional indexes such that $\left[\hat{\varphi}_{0,A},\hat{\varphi}_{0,B}\right], \left[\hat{\varphi}'_{0,A},\hat{\varphi}'_{0,B}\right]$ result from x_0 and $\left[\hat{\varphi}_{1,A},\hat{\varphi}_{1,B}\right], \left[\hat{\varphi}'_{1,A},\hat{\varphi}'_{1,B}\right]$ from x_1. As we saw in the previous section, the quality of the estimates $\hat{\varphi}_A$ and $\hat{\varphi}_B$ varies largely based on the actual $|u_0|$. As a consequence, the quality of the estimates generally depends on whether x_0 or x_1 is used as the reference. There exist a lot of ways to pick the reference observation based on different optimization criteria. A simple and efficient approach is to consider this as an additional degree of freedom over which we perform the likelihood marginalization. The resulting ML is results as

$$\begin{aligned}\hat{q} = \underset{q}{\operatorname{argmax}} & \sum_{q_A,q_B:\chi_q(q_A,q_B)=q} \mathrm{p}\left(x'_1|q_A,q_B,\hat{\varphi}_{0,A},\hat{\varphi}_{0,B}\right) \\ & + \sum_{q_A,q_B:\chi_q(q_A,q_B)=q} \mathrm{p}\left(x'_1|q_A,q_B,\hat{\varphi}'_{0,A},\hat{\varphi}'_{0,B}\right) \\ & + \sum_{q_A,q_B:\chi_q(q_A,q_B)=q} \mathrm{p}\left(x'_0|q^*_A,q^*_B,\hat{\varphi}_{1,A},\hat{\varphi}_{1,B}\right) \\ & + \sum_{q_A,q_B:\chi_q(q_A,q_B)=q} \mathrm{p}\left(x'_0|q^*_A,q^*_B,\hat{\varphi}'_{1,A},\hat{\varphi}'_{1,B}\right).\end{aligned} \quad (5.20)$$

5.4 Sequential ML

In contrast to the previous section, where the two-point observation got split into two parts - the reference for estimating channel phases and the observation for symbol estimation, in this section we consider the joint likelihood $\mathrm{p}\left(x'_0, x'_1 | q_A, q_B, \varphi_A, \varphi_B\right)$ of the whole two-point observation. The memoryless channel allows us to factorize the likelihood as

$$\begin{aligned}\mathrm{p}\left(x'_0, x'_1 | q_A, q_B, \varphi_A, \varphi_B\right) &= \mathrm{p}\left(x'_0, x'_1 | \Delta_A, \Delta_B, \varphi_A, \varphi_B\right) \\ &= \exp\left(\frac{-1}{\sigma_w^2}\left|x'_0 - \eta_A \mathrm{e}^{\mathrm{j}\varphi_A} - \eta_B \mathrm{e}^{\mathrm{j}\varphi_B}\right|^2\right) \\ &\times \exp\left(\frac{-1}{\sigma_w^2}\left|x'_1 - \eta_A \mathrm{e}^{\mathrm{j}(\varphi_A+\Delta_A)} - \eta_B \mathrm{e}^{\mathrm{j}(\varphi_B+\Delta_B)}\right|^2\right).\end{aligned} \tag{5.21}$$

Because the joint maximization of (5.21) w.r.t. all the conditioning variables would be hard to solve, we propose a sequential maximization. In the first step we maximize (5.21) only w.r.t. the phases φ_A, φ_B resulting in

$$\mathrm{m}\left(x'_0, x'_1 | \Delta_A, \Delta_B\right) = \max_{\varphi_A, \varphi_B} \mathrm{p}\left(x'_0, x'_1 | \Delta_A, \Delta_B, \varphi_A, \varphi_B\right). \tag{5.22}$$

In the second step, we marginalize the obtained metric over the hierarchical dispersion and perform a second maximization over the target hierarchical channel symbol

$$\hat{q} = \underset{q}{\operatorname{argmax}} \sum_{q=\chi_q\left(\mathrm{e}^{\mathrm{j}\Delta_A}, \mathrm{e}^{\mathrm{j}\Delta_B}\right)} \mathrm{m}\left(x'_0, x'_1 | \Delta_A, \Delta_B\right). \tag{5.23}$$

Equation (5.22) is hard to solve in general. It is possible to find an analytical solution for two special cases. Namely when $\Delta_A = \Delta_B + \pi$ and when $\Delta_A = \Delta_B$. Based on those two cases a closed-form solution for the case of a BPSK alphabet $\mathcal{A}_q = \{\pm 1\}$ can be constructed.

5.4.1 Special Case for BPSK

For the case of an XOR HNC map and a BPSK alphabet, (5.23) will reduce to

$$\mathrm{m}\left(x'_0, x'_1 | 0, 0\right) + \mathrm{m}\left(x'_0, x'_1 | \pi, \pi\right) \underset{\hat{q}=-1}{\overset{\hat{q}=1}{\lessgtr}} \mathrm{m}\left(x'_0, x'_1 | \pi, 0\right) + \mathrm{m}\left(x'_0, x'_1 | 0, \pi\right). \tag{5.24}$$

In the rest of this section, we derive closed-form solutions of the m metrics (5.24).

Special case for $\Delta_B = \Delta_A + \pi$

After merging the exponential functions and expanding the square magnitudes in (5.21) we obtain

$$\begin{aligned}\mathrm{p}\left(x'_0, x'_1 | \Delta_A, \Delta_B, \varphi_A, \varphi_B\right) = \exp\Bigg(\frac{-1}{\sigma_w^2}\Big[&\left|x'_0\right|^2 + \left|x'_1\right|^2 + 2\eta_A^2 + 2\eta_B^2 \\ &+ 2\eta_A\eta_B \Re\left\{\mathrm{e}^{\mathrm{j}(\varphi_A-\varphi_B)}\left(1 + \mathrm{e}^{\mathrm{j}(\Delta_A-\Delta_B)}\right)\right\} \\ &- 2\Re\left\{x'_0\left(\eta_A \mathrm{e}^{-\mathrm{j}\varphi_A} + \eta_B \mathrm{e}^{-\mathrm{j}\varphi_B}\right)\right\} \\ &- 2\Re\left\{x'_1\left(\eta_A \mathrm{e}^{-\mathrm{j}(\varphi_A+\Delta_A)} + \eta_B \mathrm{e}^{-\mathrm{j}(\varphi_B+\Delta_B)}\right)\right\}\Big]\Bigg).\end{aligned} \tag{5.25}$$

We focus only on the real parts because the rest does not depend on φ_A, φ_B. After substituting $\Delta_B = \Delta_A + \pi$ the first term disappears completely

$$2\eta_A\eta_B\Re\left\{\mathrm{e}^{\mathrm{j}(\varphi_A-\varphi_B)}\left(1+\mathrm{e}^{\mathrm{j}(\Delta_A-\Delta_B)}\right)\right\}\Big|_{\Delta_B=\Delta_A+\pi} = 0 \tag{5.26}$$

and decouples the two phase variables. We are left with the following two terms which can be maximized independently.

$$\Re\left\{\eta_A\mathrm{e}^{-\mathrm{j}\varphi_A}\left(x_0' + x_1'\mathrm{e}^{-\mathrm{j}\Delta_A}\right)\right\},\ \ \Re\left\{\eta_B\mathrm{e}^{-\mathrm{j}\varphi_B}\left(x_0' - x_1'\mathrm{e}^{-\mathrm{j}\Delta_A}\right)\right\} \tag{5.27}$$

We are not interested in the particular values of φ_A and φ_B which maximize the above terms, only in their respective maximum values for which we have

$$\begin{aligned}\max_{\varphi_A}\Re\left\{\eta_A\mathrm{e}^{-\mathrm{j}\varphi_A}\left(x_0' + x_1'\mathrm{e}^{-\mathrm{j}\Delta_A}\right)\right\} &= \eta_A\left|x_0' + x_1'\mathrm{e}^{-\mathrm{j}\Delta_A}\right| \\ \max_{\varphi_B}\Re\left\{\eta_B\mathrm{e}^{-\mathrm{j}\varphi_B}\left(x_0' - x_1'\mathrm{e}^{-\mathrm{j}\Delta_A}\right)\right\} &= \eta_B\left|x_0' - x_1'\mathrm{e}^{-\mathrm{j}\Delta_A}\right|.\end{aligned} \tag{5.28}$$

After substituting (5.28) into (5.25) we get the solution

$$\begin{aligned}\mathrm{m}\left(x_0', x_1'|\Delta_A, \Delta_A+\pi\right) = \exp\Bigg(\frac{-1}{\sigma_w^2}\Big[&\left|x_0'\right|^2 + \left|x_1'\right|^2 + 2\eta_A^2 + 2\eta_B^2 \\ &- 2\left(\eta_A\left|x_0' + x_1'\mathrm{e}^{-\mathrm{j}\Delta_A}\right| + \eta_B\left|x_0' - x_1'\mathrm{e}^{-\mathrm{j}\Delta_A}\right|\right)\Big]\Bigg).\end{aligned} \tag{5.29}$$

For the particular cases in (5.24) we have

$$\mathrm{m}\left(x_0', x_1'|0, \pi\right) = \exp\left(\frac{-1}{\sigma_w^2}\left[\left|x_0'\right|^2 + \left|x_1'\right|^2 + 2\eta_A^2 + 2\eta_B^2 - 2\left(\eta_A\left|x_0' + x_1'\right| + \eta_B\left|x_0' - x_1'\right|\right)\right]\right) \tag{5.30}$$

$$\mathrm{m}\left(x_0', x_1'|\pi, 0\right) = \exp\left(\frac{-1}{\sigma_w^2}\left[\left|x_0'\right|^2 + \left|x_1'\right|^2 + 2\eta_A^2 + 2\eta_B^2 - 2\left(\eta_A\left|x_0' - x_1'\right| + \eta_B\left|x_0' + x_1'\right|\right)\right]\right). \tag{5.31}$$

Special case for $\Delta_A = \Delta_B$

For the second special case when $\Delta_A = \Delta_B = \Delta$, we start again from the likelihood expression (5.21) and substitute $\Delta_A = \Delta$ and $\Delta_B = \Delta$ to get

$$\begin{aligned}\mathrm{p}\left(x_0', x_1'|\Delta, \Delta, \varphi_A, \varphi_B\right) &= \exp\left(\frac{-1}{\sigma_w^2}\left|x_0' - \eta_A\mathrm{e}^{\mathrm{j}\varphi_A} - \eta_B\mathrm{e}^{\mathrm{j}\varphi_B}\right|^2\right) \\ &\quad\times\exp\left(\frac{-1}{\sigma_w^2}\left|x_1' - \eta_A\mathrm{e}^{\mathrm{j}(\varphi_A+\Delta)} - \eta_B\mathrm{e}^{\mathrm{j}(\varphi_B+\Delta)}\right|^2\right) \\ &= \exp\left(\frac{-1}{\sigma_w^2}\left[\left|x_0' - B\right|^2 + \left|x_1' - \mathrm{e}^{\mathrm{j}\Delta}B\right|^2\right]\right) \\ &= \exp\left(\frac{-1}{\sigma_w^2}\left[\left|x_0' - B\right|^2 + \left|x_1'\mathrm{e}^{-\mathrm{j}\Delta} - B\right|^2\right]\right),\end{aligned} \tag{5.32}$$

where $B = \eta_A\mathrm{e}^{\mathrm{j}\varphi_A} + \eta_B\mathrm{e}^{\mathrm{j}\varphi_B}$. The problem can be understood as a search for a complex point B which minimizes the sum of squared distances to the given points x_0' and $x_1'\mathrm{e}^{-\mathrm{j}\Delta}$ subject to the constraint of $x_0', x_1'\mathrm{e}^{-\mathrm{j}\Delta}, B \in \mathbb{A}$.

Let us consider a general case where we minimize

$$f(x) = \left|a - x\right|^2 + \left|b - x\right|^2,\ a \neq b,\ a, b, x \in \mathbb{A}. \tag{5.33}$$

We start by computing the generalized derivatives [100] w.r.t. x and put them equal to zero

$$\frac{\tilde{\partial} f(x)}{\tilde{\partial} x} = -a^* - b^* + 2x^* = 0 \;\Rightarrow x^* = \frac{a^* + b^*}{2} \tag{5.34}$$

$$\frac{\tilde{\partial} f(x)}{\tilde{\partial} x^*} = -a - b + 2x = 0 \;\Rightarrow x = \frac{a+b}{2}, \tag{5.35}$$

which gives us the solution as the midpoint between a and b. Because $\mathbb{A}$ is not convex, it may or may not contain the midpoint (an example of the situation when $\mathbb{A}$ contains the midpoint is shown in Figure 5.7a). Since $a, b \in \mathbb{A}$ it follows that

$$\left|\frac{a+b}{2}\right| \leq \frac{1}{2}\left[|a| + |b|\right] \leq \frac{1}{2}\left[r_2 + r_2\right] = |r_2| \tag{5.36}$$

and we need to only address the case when $\left|\frac{a+b}{2}\right| < r_1$. In such a case the arguments of the extremes will lie on the borders of $\mathbb{A}$. The outer border can be excluded because

$$f(a) \leq f(x), |x| = r_2 \tag{5.37}$$

which is a consequence of the triage inequality. We are left with an optimization problem subject to an equality constraint

$$f(x) = |a - x|^2 + |b - x|^2 \tag{5.38}$$

$$g(x) = |x|^2 - r_1^2 = 0. \tag{5.39}$$

We form the Lagrangian function as $\mathcal{L}(x, \lambda) = f(x) + \lambda g(x)$ and put its gradient equal to the zero vector [100]

$$2x^* - a^* - b^* + \lambda x^* = 0 \;\Rightarrow x^* = \frac{a^* + b^*}{2 + \lambda} \tag{5.40}$$

$$2x - a - b + \lambda x = 0 \;\Rightarrow x = \frac{a+b}{2+\lambda} \tag{5.41}$$

$$|x|^2 - r_1^2 = 0 \;\Rightarrow \frac{a^* + b^*}{2+\lambda}\frac{a+b}{2+\lambda} - r_1^2 = 0. \tag{5.42}$$

This way we obtain a quadratic equation in λ

$$\lambda^2 r_1^2 + \lambda\left(4r_1^2\right) + 4r_1^2 - |a+b|^2 \tag{5.43}$$

with the following two roots (assuming $r_1 \neq 0$)

$$\lambda_{1,2} = \frac{-2r_1 \pm |a+b|}{r_1}. \tag{5.44}$$

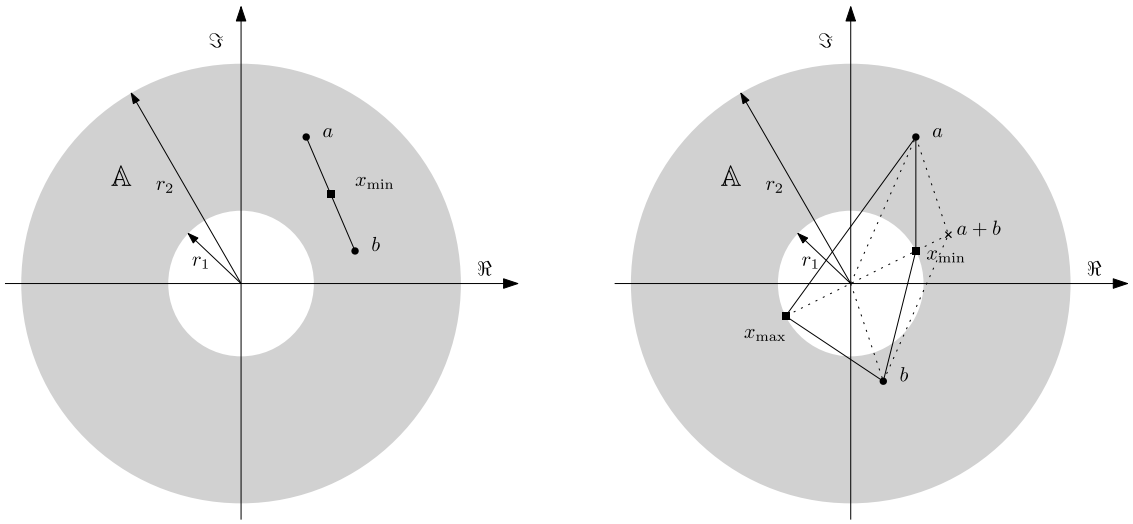
After a substitution, we get the solution

$$x_{1,2} = \pm\frac{r_1(a+b)}{|a+b|}, \tag{5.45}$$

where one solution corresponds to the minimum and the other to the maximum. An example of this situation is shown in Figure 5.7b.

Now we apply the general result to (5.32) substituting $a = x_0', b = x_1' \mathrm{e}^{-\mathrm{j}\Delta}$ and get

$$\begin{aligned} \mathrm{m}(x_0', x_1' | q_A, q_B) &= \exp\left(\frac{-2}{\sigma_w^2}\left|x_0' - \frac{x_0' + x_1'\mathrm{e}^{-\mathrm{j}\Delta}}{2}\right|^2\right) \\ &= \exp\left(\frac{-1}{2\sigma_w^2}\left|x_0' - x_1'\mathrm{e}^{-\mathrm{j}\Delta}\right|^2\right) \end{aligned} \tag{5.46}$$



(a) $\frac{a+b}{2} \in \mathbb{A}$ (b) $\frac{a+b}{2} \notin \mathbb{A}$

Figure 5.7: Geometric interpretation of the maximization of (5.33).

for the case when $\frac{x'_0+x'_1 e^{-j\Delta}}{2} \in \mathbb{A}$. In the other case, we substitute the solutions (5.45) into (5.32), get

$$\begin{aligned}
&\exp\left(\frac{-1}{\sigma_w^2}\left[\left|x'_0 \pm \frac{r_1\left(x'_0+x'_1 e^{-j\Delta}\right)}{\left|x'_0+x'_1 e^{-j\Delta}\right|}\right|^2 + \left|x'_1 e^{-j\Delta} \pm \frac{r_1\left(x'_0+x'_1 e^{-j\Delta}\right)}{\left|x'_0+x'_1 e^{-j\Delta}\right|}\right|^2\right]\right) \\
=&\exp\left(\frac{-1}{\sigma_w^2}\left[\left|x'_0\right|^2+\left|x'_1\right|^2+2r_1^2 \pm \frac{2r_1}{\left|x'_0+x'_1 e^{-j\Delta}\right|}\Re\left\{\left(x'_0+x'_1 e^{-j\Delta}\right)^*\left(x'_0+x'_1 e^{-j\Delta}\right)\right\}\right]\right) \\
=&\exp\left(\frac{-1}{\sigma_w^2}\left[\left|x'_0\right|^2+\left|x'_1\right|^2+2r_1^2 \pm 2r_1\left|x'_0+x'_1 e^{-j\Delta}\right|\right]\right)
\end{aligned} \tag{5.47}$$

and select the negative sign to obtain the maximum.

The final solution is given as

$$\mathrm{m}\left(x'_0, x'_1|\Delta,\Delta\right) = \begin{cases} e^{\frac{-1}{2\sigma_w^2}\left|x'_0-x'_1 e^{-j\Delta}\right|^2} & r_1 \le |M| \\ e^{\frac{-1}{\sigma_w^2}\left[\left|x'_0\right|^2+\left|x'_1\right|^2+2r_1^2-2r_1\left|x'_0+x'_1 e^{-j\Delta}\right|\right]} & r_1 > |M| \end{cases}, \tag{5.48}$$

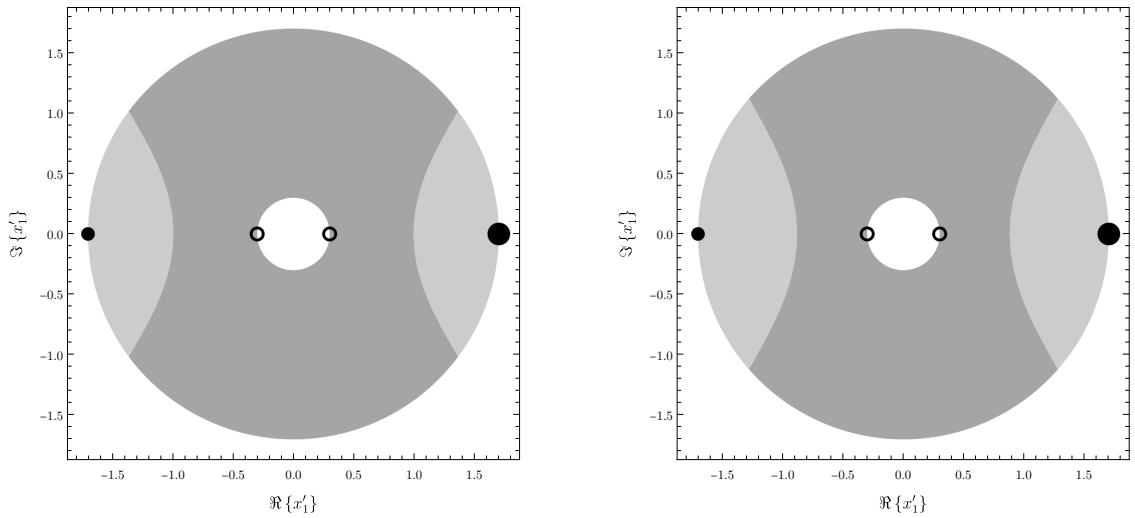
where $M = \frac{x'_0+x'_1 e^{-j\Delta}}{2}$ is the midpoint. For the particular cases in (5.24) we have

$$\mathrm{m}\left(x'_0, x'_1|0,0\right) = \begin{cases} e^{\frac{-1}{2\sigma_w^2}\left|x'_0-x'_1\right|^2} & r_1 \le \left|\frac{x'_0+x'_1}{2}\right| \\ e^{\frac{-1}{\sigma_w^2}\left[\left|x'_0\right|^2+\left|x'_1\right|^2+2r_1^2-2r_1\left|\left(x'_0+x'_1\right)\right|\right]} & r_1 > \left|\frac{x'_0+x'_1}{2}\right| \end{cases} \tag{5.49}$$

$$\mathrm{m}\left(x'_0, x'_1|\pi,\pi\right) = \begin{cases} e^{\frac{-1}{2\sigma_w^2}\left|x'_0+x'_1\right|^2} & r_1 \le \left|\frac{x'_0-x'_1}{2}\right| \\ e^{\frac{-1}{\sigma_w^2}\left[\left|x'_0\right|^2+\left|x'_1\right|^2+2r_1^2-2r_1\left|x'_0-x'_1\right|\right]} & r_1 > \left|\frac{x'_0-x'_1}{2}\right| \end{cases}. \tag{5.50}$$

5.4.2 Decision Regions

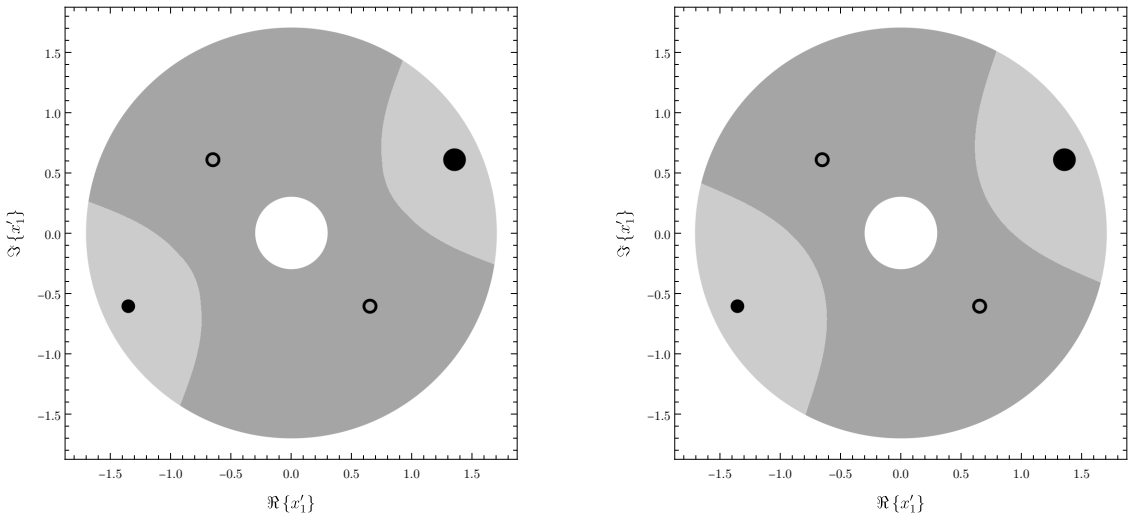
We use the results from Section 5.4.1 and visualize (5.24) using decision region plots shown in Figures 5.8, 5.9, 5.10, 5.11, and 5.12. We consider a two-point observation, where $u_0 = 1 + 0.7\mathrm{e}^{\mathrm{j}\varphi_B}$ is fixed and decide on $\hat{q}$ based on the location of $x'_1 \in \mathbb{A}$ (in a real case we would not know u_0 but only the noisy observation x'_0, as a result, the decision regions would be shifted by $x'_0 - u_0$). The dark region corresponds to the set of x'_1 based on which the ML estimator (5.24) selects $\hat{q} = -1$ and the light region corresponds to $\hat{q} = 1$. For reference, we include the four possible values of u_1: u_0 (big black point), $-u_0$ (small black point), and $\pm 1 \mp 0.7\mathrm{e}^{\mathrm{j}\varphi_B}$ (black rings). As can be observed, the individual values of u_1 fall into the appropriate decision regions (the points into the light region and the circles into the dark region).



(a) SNR = 0 dB

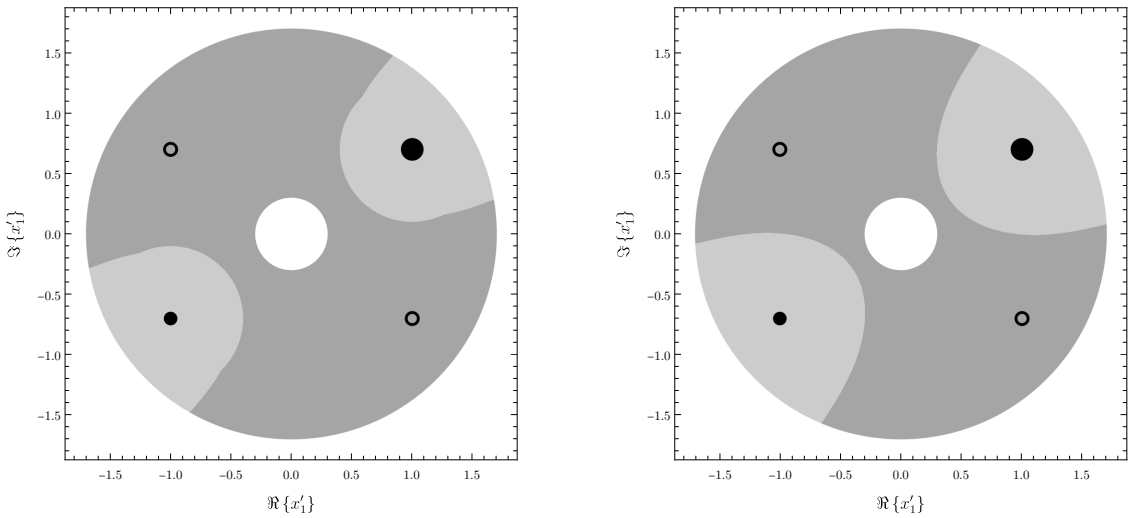
(b) SNR = 15 dB

Figure 5.8: Decision regions of (5.24) for $h_A = 1, h_B = 0.7\angle 0$.



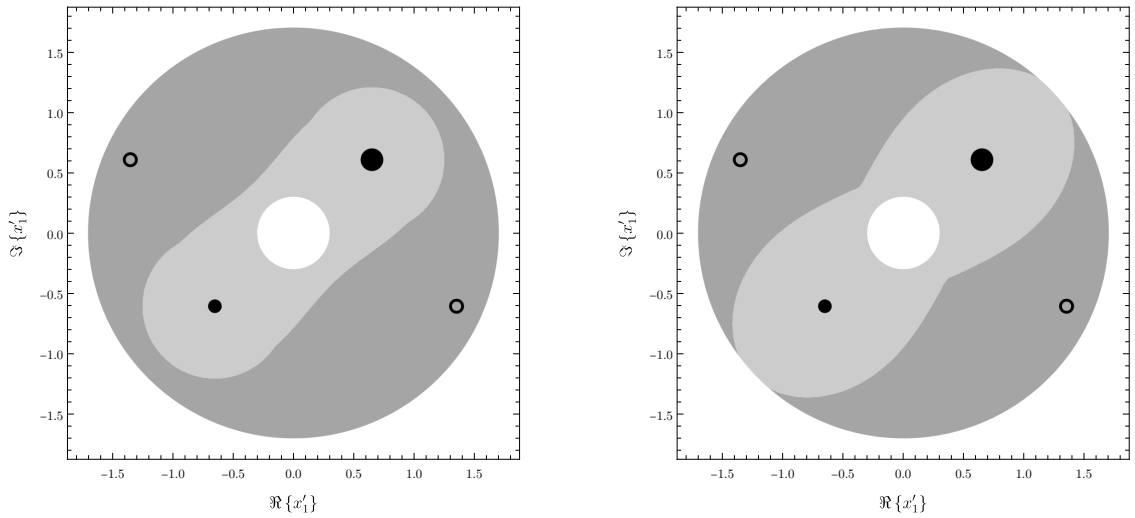
(a) SNR = 0 dB (b) SNR = 15 dB

Figure 5.9: Decision regions of (5.24) for $h_A = 1, h_B = 0.7\angle\frac{\pi}{3}$.



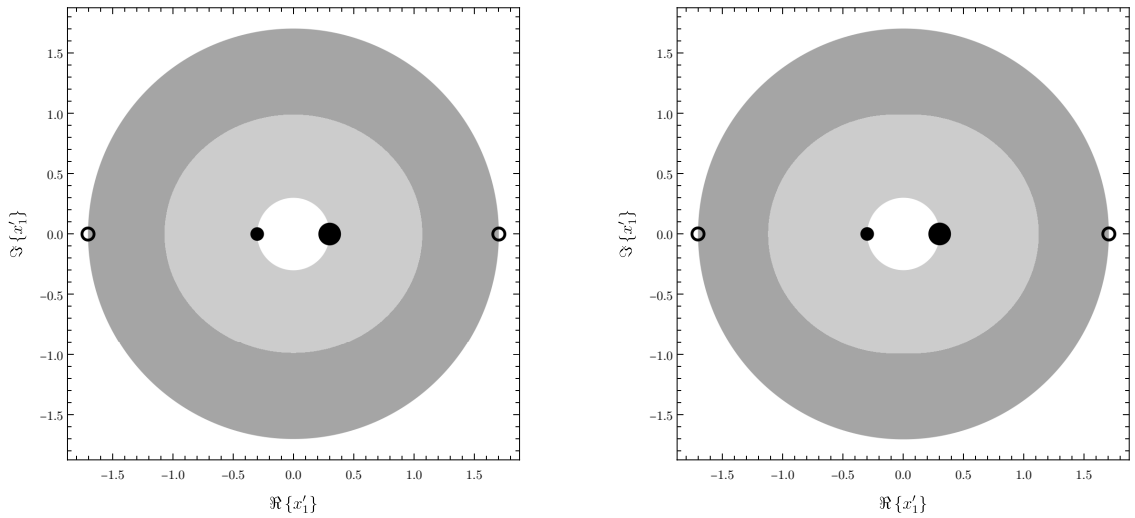
(a) SNR = 0 dB (b) SNR = 15 dB

Figure 5.10: Decision regions of (5.24) for $h_A = 1, h_B = 0.7\angle\frac{\pi}{2}$.



(a) SNR = 0 dB (b) SNR = 15 dB

Figure 5.11: Decision regions of (5.24) for $h_A = 1, h_B = 0.7\angle\frac{2\pi}{3}$.



(a) SNR = 0 dB (b) SNR = 15 dB

Figure 5.12: Decision regions of (5.24) for $h_A = 1, h_B = 0.7\angle\pi$.

5.5 Ad-Hoc Methods

In this section, we present additional ad-hoc demodulation methods for the BPSK and QPSK alphabets, both with the bitwise XOR as the HNC map.

5.5.1 BPSK

We start by considering only the useful signal. Given $u_0 = \eta_A \mathrm{e}^{\varphi_A} + \eta_B \mathrm{e}^{\varphi_B}$, we have

$$u_1 \in \{\eta_A \mathrm{e}^{\varphi_A} + \eta_B \mathrm{e}^{\varphi_B}, -\eta_A \mathrm{e}^{\varphi_A} - \eta_B \mathrm{e}^{\varphi_B}\} = \{u_0, -u_0\} = \mathbb{U}_1 \tag{5.51}$$

for $q = 1$ and

$$u_1 \in \{\eta_A \mathrm{e}^{\varphi_A} - \eta_B \mathrm{e}^{\varphi_B}, \eta_B \mathrm{e}^{\varphi_B} - \eta_A \mathrm{e}^{\varphi_A}\} = \mathbb{U}_{-1} \tag{5.52}$$

for $q = -1$. The idea is to take the noisy observation x'_0 and construct a noisy variant of (5.51) as

$$\hat{\mathbb{U}}_1 = \{x'_0, -x'_0\}\,. \tag{5.53}$$

A decision on $\hat{q}$ can be made based on the Euclidean distance of the second observation x'_1 and $\hat{\mathbb{U}}_1$

$$\mathrm{d}\left(x'_1, \hat{\mathbb{U}}_0\right) = \min\left\{\left|x'_1 - x'_0\right|, \left|x'_1 + x'_0\right|\right\}. \tag{5.54}$$

The decision is made such that $\hat{q} = 1$ if

$$\mathrm{d}\left(x'_1, \hat{\mathbb{U}}_0\right) < \frac{1}{2}\,\mathrm{d}\left(\mathbb{U}_1, \mathbb{U}_0\right) = \frac{1}{2}\min\left\{2\eta_A, 2\eta_B\right\}. \tag{5.55}$$

The resulting estimate can be computed very easily as

$$\min\left\{\left|x'_1 - x'_0\right|, \left|x'_1 + x'_0\right|\right\} \underset{\hat{q}=-1}{\overset{\hat{q}=1}{\lessgtr}} \min\left\{\eta_A, \eta_B\right\}. \tag{5.56}$$

A graphical illustration of the decision regions of (5.56) is shown in Figure 5.13. The regions are constructed based on u_0. The set of x'_1 for which (5.56) decides as $\hat{q} = -1$ is marked with a gray background and the set for which $\hat{q} = 1$ is given as the complement (in a real case we would not know u_0 but only the noisy observation x'_0, as a result, the decision regions would be shifted by $x'_0 - u_0$).

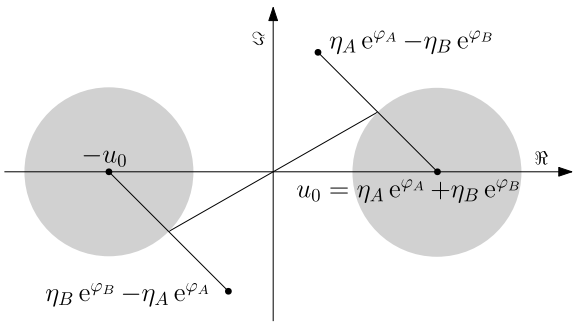


Figure 5.13: Example illustration of the resulting decision regions produced by the ad-hoc BPSK demodulation procedure. If x'_1 falls inside the gray area, $\hat{q} = 1$, otherwise $\hat{q} = -1$.

5.5.2 QPSK

In this section, we extend the previous ad-hoc demodulator for BPSK to a higher order alphabet. In (5.56) we compare the magnitudes of the sum and difference of the two observations. In the QPSK case, we consider two additional combinations $|x'_0 + \mathrm{j}x'_1|$ and $|x'_0 - \mathrm{j}x'_1|$ resulting in four combinations in total. The combinations are designated

$$x'^{(i)} = x'_0 + k_i x'_1, k_i \in \{\pm 1, \pm \mathrm{j}\} \tag{5.57}$$

and

$$u^{(i)} = u_0 + k_i u_1, k_i \in \{\pm 1, \pm \mathrm{j}\} \tag{5.58}$$

in case of the useful signal only. Motivated by (5.56) we evaluate the magnitudes of $u^{(i)}$ conditioned by q_A, q_B and assemble Table 5.1. In some cases, the magnitude can not be solved without the knowledge of φ_A, φ_B. In such cases we specify the set of possible values in terms of an interval.

The values $x'^{(i)}$ follow a Gaussian distribution with variance $2\sigma_w^2$ per dimension and their magnitude is distributed according to a folded normal distribution. For instance, the PDF of $\left|x'^{(1)}\right| = |x'_0 + x'_1|$ is proportional to

$$\exp\left(-\frac{(x'-\mu)^2}{2\sigma_w^2}\right) + \exp\left(-\frac{(x'+\mu)^2}{2\sigma_w^2}\right), x' \geq 0, \tag{5.59}$$

where μ is one of the values from Table 5.1. In this initial approach, we assume σ_w^2 to be small w.r.t. η_A, η_B such that we can approximate the PDF of $\left|x'^{(i)}\right|$ by a Gaussian one with mean $\left|u^{(i)}\right|$ and variance $2\sigma_w^2$ in the vicinity of $\left|u^{(i)}\right|$. The hierarchical target symbol estimate is decided as follows.

We start by computing the values of $\left|x'^{(i)}\right|$ and calculate the squared distances w.r.t. the expected values (highlighted in Table 5.1) conditioned by various combinations of q_A and q_B. Next, we sum the obtained squared differences over the rows of the table and get a distance metric ρ for each combination of q_A and q_B. In the end, we perform a marginalization over the hierarchical dispersion

$$\rho\left(x'_0, x'_1 | q\right) = \min_{q_A, q_B : \chi_q(q_A, q_B) = q} \rho\left(x'_0, x'_1 | q_A, q_B\right) \tag{5.60}$$

and get the final estimate as

$$\hat{q} = \operatorname*{argmin}_q \rho\left(x_0, x_1 | q\right). \tag{5.61}$$

Example 5.2. An example evaluation of $\rho\left(x'_0, x'_1 | q_A, q_B\right)$ for $q_A = 1$ and $q_B = \mathrm{j}$ is done as follows. In Table 5.1 we find the row given the condition $q_A = 1, q_B = \mathrm{j}$, which in this case is the fifth row. In the second step, we look for the highlighted magnitudes, which we find to be $\left|u^{(3)}\right|$ and $\left|u^{(4)}\right|$. To get $\rho\left(x'_0, x'_1 | 1, \mathrm{j}\right)$ we need to sum the squared distances of the observations and expected values for x'_0, x'_1. As such we have

$$\begin{aligned} \rho\left(x'_0, x'_1 | 1, \mathrm{j}\right) &= \left(\left|x'^{(3)}\right| - \sqrt{2}\eta_B\right)^2 + \left(\left|x'^{(4)}\right| - \sqrt{2}\eta_A\right)^2 \\ &= \left(|x'_0 - x'_1| - \sqrt{2}\eta_B\right)^2 + \left(|x'_0 - \mathrm{j}x'_1| - \sqrt{2}\eta_A\right)^2. \end{aligned} \tag{5.62}$$

q_A	q_B	$\lvert u_0+u_1\rvert$	$\lvert u_0+\mathrm{j}u_1\rvert$	$\lvert u_0-u_1\rvert$	$\lvert u_0-\mathrm{j}u_1\rvert$
$q=1$					
1	1	$2\lvert u_0\rvert$	$\sqrt{2}\lvert u_0\rvert$	0	$\sqrt{2}\lvert u_0\rvert$
j	j	$\sqrt{2}\lvert u_0\rvert$	0	$\sqrt{2}\lvert u_0\rvert$	$2\lvert u_0\rvert$
-1	-1	0	$\sqrt{2}\lvert u_0\rvert$	$2\lvert u_0\rvert$	$\sqrt{2}\lvert u_0\rvert$
$-\mathrm{j}$	$-\mathrm{j}$	$\sqrt{2}\lvert u_0\rvert$	$2\lvert u_0\rvert$	$\sqrt{2}\lvert u_0\rvert$	0
$q=\mathrm{j}$					
1	j	①	②	$\sqrt{2}\eta_B$	$\sqrt{2}\eta_A$
j	1	②	$\sqrt{2}\eta_B$	$\sqrt{2}\eta_A$	①
-1	$-\mathrm{j}$	$\sqrt{2}\eta_B$	$\sqrt{2}\eta_A$	①	②
$-\mathrm{j}$	-1	$\sqrt{2}\eta_A$	①	②	$\sqrt{2}\eta_B$
$q=-1$					
1	-1	$2\eta_A$	③	$2\eta_B$	③
j	$-\mathrm{j}$	③	$2\eta_B$	③	$2\eta_A$
-1	1	$2\eta_B$	③	$2\eta_A$	③
$-\mathrm{j}$	j	③	$2\eta_A$	③	$2\eta_B$
$q=-\mathrm{j}$					
1	$-\mathrm{j}$	①	$\sqrt{2}\eta_A$	$\sqrt{2}\eta_B$	②
j	-1	$\sqrt{2}\eta_A$	$\sqrt{2}\eta_B$	②	①
-1	j	$\sqrt{2}\eta_B$	②	①	$\sqrt{2}\eta_A$
$-\mathrm{j}$	1	②	①	$\sqrt{2}\eta_A$	$\sqrt{2}\eta_B$

(a)

①	$\in \left\langle \lvert 2\eta_A-\sqrt{2}\eta_B\rvert ; 2\eta_A-\sqrt{2}\eta_B \right\rangle$
②	$\in \left\langle \lvert \sqrt{2}\eta_A-2\eta_B\rvert ; \sqrt{2}\eta_A-2\eta_B \right\rangle$
③	$\in \left\langle \sqrt{2}\lvert \eta_A-\eta_B\rvert ; \sqrt{2}(\eta_A-\eta_B) \right\rangle$

(b)

Table 5.1: Conditions for the magnitudes of various combinations of the useful signals u_0, u_1 and unknown channel phase values φ_A, φ_B. The table is grouped into four segments, based on the corresponding hierarchical target symbol q.

5.6 Error Performance Evaluation

5.6.1 Numerical Simulation

In this section, we numerically evaluate the H-BER, i.e. the BER of the target hierarchical symbol q. The simulations are performed as follows. First, we generate an i.i.d. sequence of 10^4 channel symbols for each source. For every channel symbol tuple $(q_{A,n}, q_{B,n})$ we compute a pair of useful signal constellation space points

$$\begin{aligned} u_{0,n} &= \eta_A \mathrm{e}^{\mathrm{j}\varphi_{A,n}} + \eta_B \mathrm{e}^{\mathrm{j}\varphi_{B,n}} \\ u_{1,n} &= \eta_A \mathrm{e}^{\mathrm{j}\varphi_{A,n}} q_{A,n} + \eta_B \mathrm{e}^{\mathrm{j}\varphi_{B,n}} q_{B,n}, \end{aligned} \tag{5.63}$$

with phases $\varphi_{A,n}, \varphi_{B,n}$ taken as i.i.d. random samples from a uniform distribution. Each pair $u_{0,n}, u_{1,n}$ gets added together with the AWGN forming the resulting two-point observation $x_{0,n}, x_{1,n}$. The obtained two-point observation $x_{0,n}, x_{1,n}$ is then used as the input to the hierarchical demodulator. In the, end the achieved H-BER rates are evaluated.

For reference, we include a coherent demodulation with artificially supplied channel phase values φ_A, φ_B. The demodulator for the coherent case is given as a simple marginalization over the hierarchical dispersion followed by a ML estimator

$$\hat{q} = \underset{q}{\operatorname{argmax}} \sum_{q_A, q_B : \chi_q(q_A, q_B) = q} \mathrm{p}\left(x_1 | q_A, q_B, \varphi_A, \varphi_B\right). \tag{5.64}$$

The resulting H-BER rates can be found in Figures 5.14 and 5.15 for the BPSK and QPSK alphabets respectively. In both cases, we fix η_A (a common channel gain can be equivalently represented by a shift in SNR) and run the simulations for two values of η_B. We present the performance of several demodulation scenarios:

NR single ref Noisy reference demodulation given by (5.18), where only the first observation is used to estimate φ_A, φ_B while the second observation is used to estimate q_n.

NR single ref resolved Same as the case above with the difference that in this case we artificially resolve the phase estimate ambiguity.

NR Noisy reference demodulation given by (5.20), where we average over the two possibilities of which observation to use for the reference when estimating the phases.

NR resolved Same as the case above, with the difference that we artificially resolve the phase estimate ambiguity. That is we reject the false solution.

Coherent A reference scenario given by (5.64) with complete knowledge of the channel state.

Sequential ML Demodulation given by (5.23) implemented using the closed form solution for the BPSK alphabet.

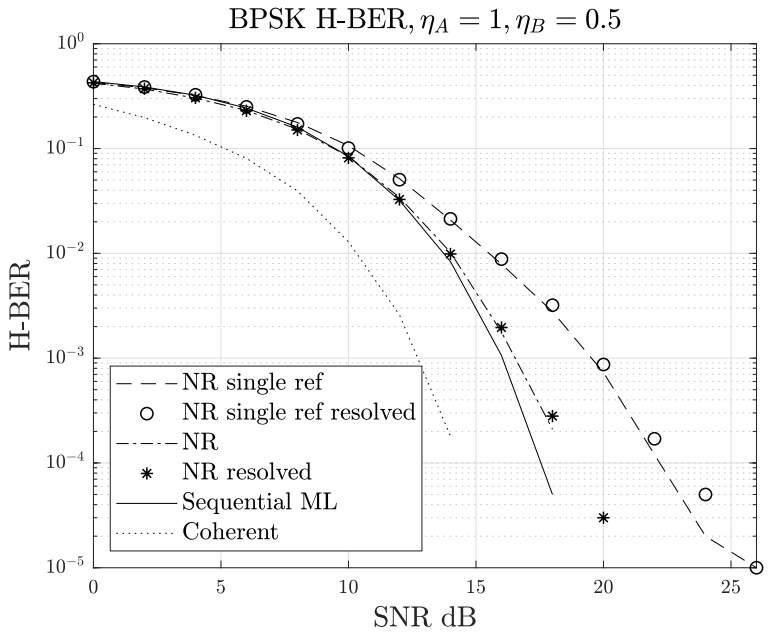
Based on Figure 5.14 we make several observations:

- The performance gap between *NR single* and *NR* is significant and more pronounced for similar values of η_A, η_B which nicely illustrates the dependence of the channel phase estimation quality on the magnitude of the reference observation as shown in Figure 5.6.

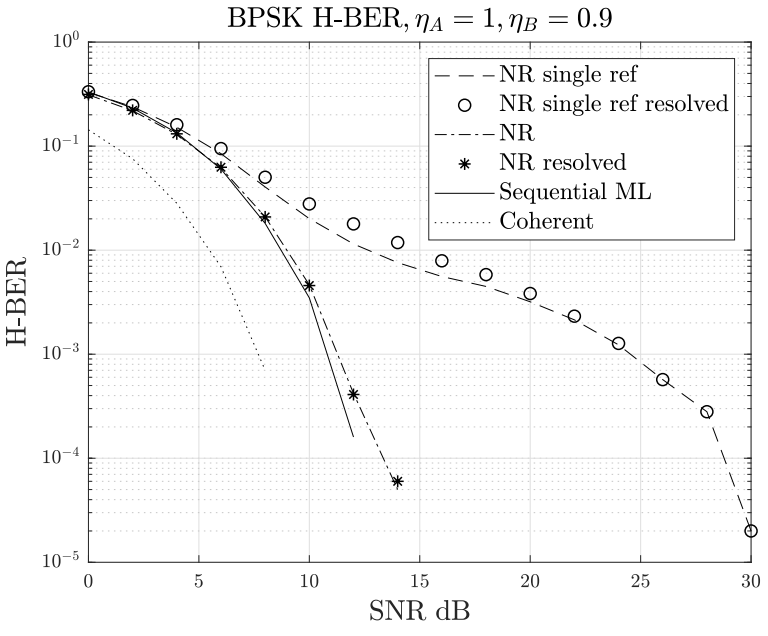
- The artificially resolved variants (*NR single ref resolved, NR resolved*) perform close to the cases where we average over the phase estimate ambiguity (*NR single ref, NR*). This is something that we illustrated in Example 5.1.
- The *Sequential ML* method gives the best performance although it is very close to *NR*.
- We observe a slight reduction of the performance drop w.r.t. the *Coherent* case for the higher value of η_B.
- Interestingly the H-MAC differential modulation is roughly 3dB away from the *Coherent* case and as such, gives a similar performance as a differentially modulated single-user case.

For the QPSK ($\mathcal{A}_q = \{1, \mathrm{j}, -1, -\mathrm{j}\}$) case (Figure 5.15) we observe:

- For this alphabet, we do not have a closed solution for the *Sequential ML* approach.
- The major performance drop is caused by channel phase estimate ambiguity.
- Similarly as in the BPSK case, *NR* performs better than *NR single ref* but only for the resolved cases which are not applicable to a practical situation.
- When focusing on the resolved cases we see a performance improvement w.r.t. the *Coherent* case for higher η_B.

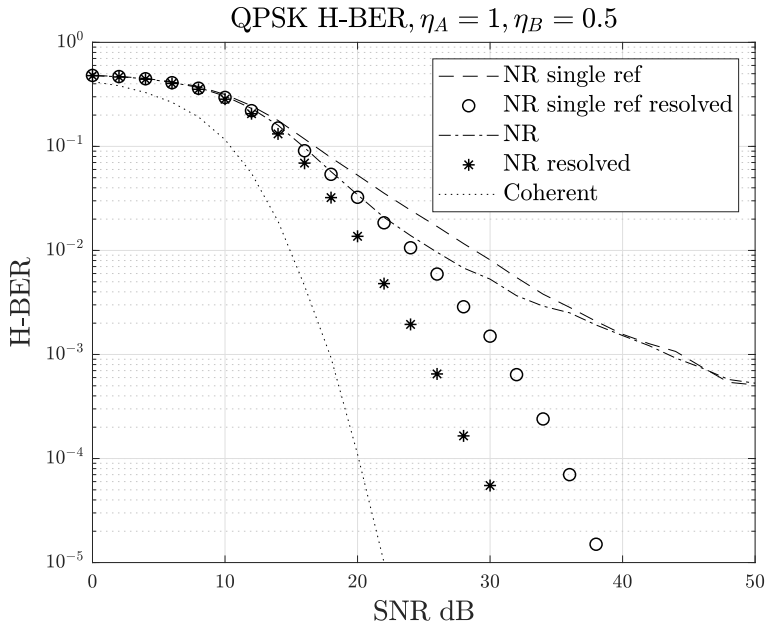


(a) $\eta_B = 0.5$

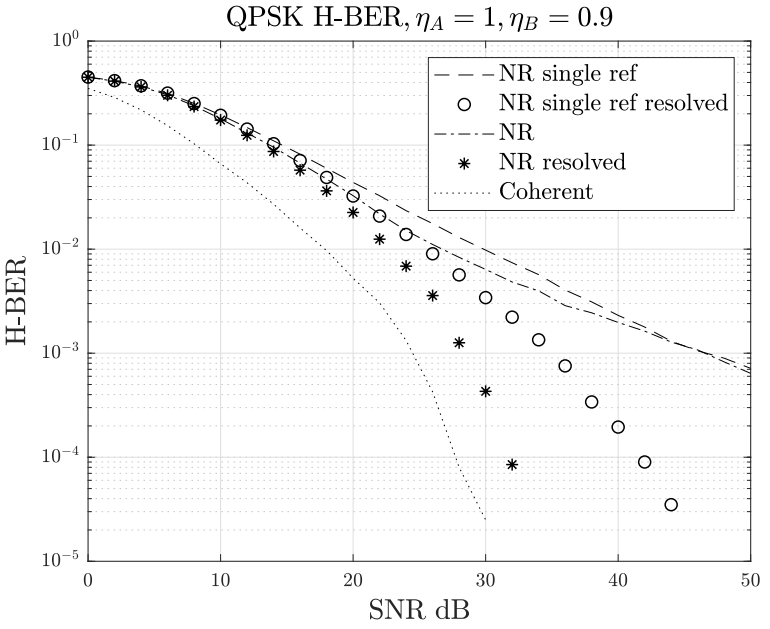


(b) $\eta_B = 0.9$

Figure 5.14: Simulated H-BER results for different differential demodulation variants for the BPSK alphabet and XOR HNC map.



(a) $\eta_B = 0.5$



(b) $\eta_B = 0.9$

Figure 5.15: Simulated H-BER results for different differential demodulation variants for the BPSK alphabet and XOR HNC map.

So far the two-point observations were generated according to the assumption (5.4) of no channel phase change between the first and second observation x_0, x_1. To get more realistic results and evaluate the robustness of the suggested demodulation algorithms we include additional phase noise in our simulations. The updated equation for the two-point observation generation is given as follows

$$\begin{aligned} x_{0,n} &= \eta_A \mathrm{e}^{\mathrm{j}\varphi_{A,n}} + \eta_B \mathrm{e}^{\mathrm{j}\varphi_{B,n}} + w_0 \\ x_{1,n} &= \eta_A \mathrm{e}^{\mathrm{j}(\varphi_{A,n}+\Delta\varphi_{A,n})} q_{A,n} + \eta_B \mathrm{e}^{\mathrm{j}(\varphi_{B,n}+\Delta\varphi_{B,n})} q_{B,n} + w_1, \end{aligned} \tag{5.65}$$

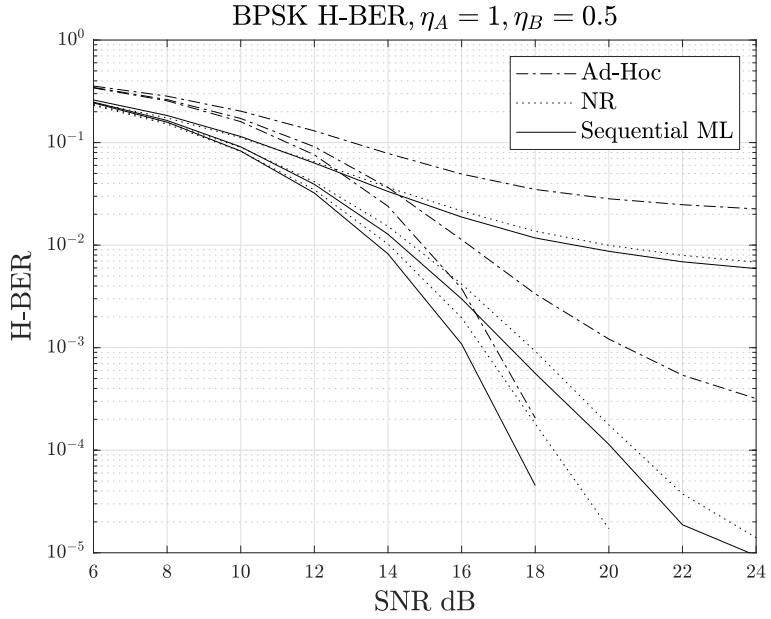
where $\Delta\varphi_{A,n}, \Delta\varphi_{B,n}$ are i.i.d. distributed according to $\mathcal{N}\left(0, \sigma_\varphi^2\right)$ while the rest stays identical to (5.63). The results are presented in Figures 5.16 and 5.17 again for the BPSK and QPSK alphabets respectively. In addition to the *NR* and *Sequential ML* methods, we also include the *Ad-Hoc* algorithms as described in Sections 5.5.1 and 5.5.2. In each graph, we show the results for three different values of the phase noise variance σ_φ^2. For each of the methods, there are three curves with rising H-BER performance corresponding to $\sigma_\varphi^2 = 0, 0.01$ and 0.04 respectively.

Based on Figure 5.16 we observe:

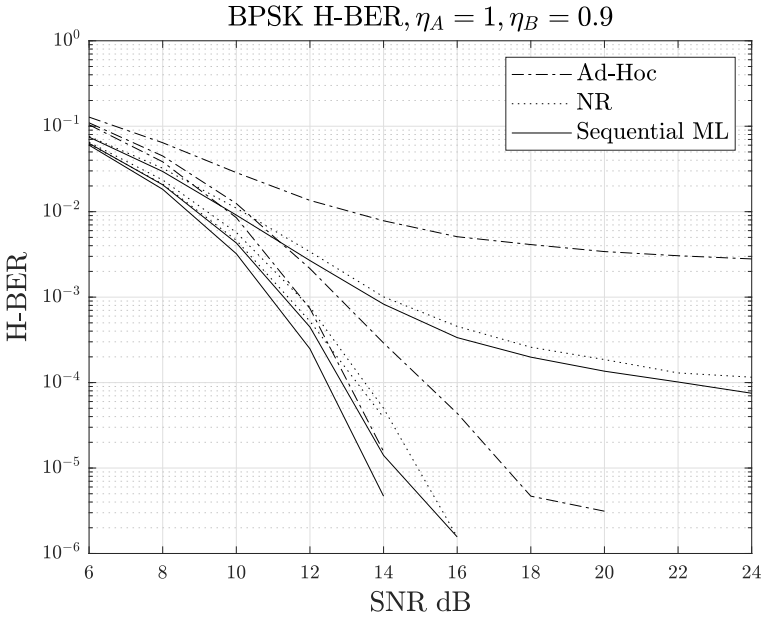
- The performance is negatively affected by the additional channel phase noise. The performance penalty depends on the degree to which (5.4) is violated (the magnitude of σ_φ^2). In a practical scenario, we expect only a minor violation and thus only a slight increase of the H-BER.
- The *Ad-Hoc* approach is inferior to *NR* and the *Sequential ML*.
- The *Ad-Hoc* approach is more sensitive to the violation of (5.4).

For the QPSK ($\mathcal{A}_q = \{1, \mathrm{j}, -1, -\mathrm{j}\}$) case (Figure 5.17) we observe:

- As in the previous case, we observe a performance degradation with rising σ_φ^2 although it is not as strong.
- Similarly, the *Ad-Hoc* method gets outperformed by *NR* although the performance difference is not as pronounced.
- The main differences in comparison to the BPSK case seem to be related to the fact that the performance with QPSK is heavily limited by the channel phase ambiguity. It is interesting to note that although the ambiguity problem does not appear in the *Ad-Hoc* approach, the final performance is very similar.

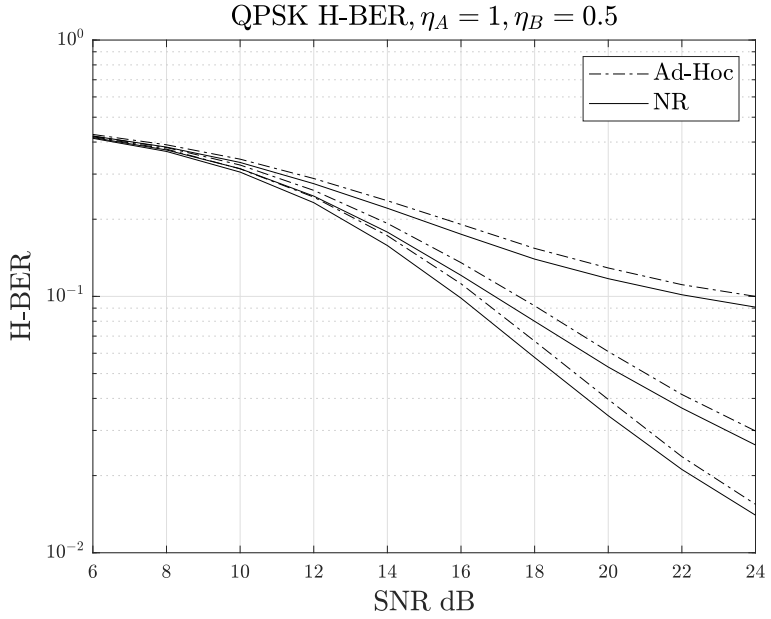


(a) $\eta_B = 0.5$

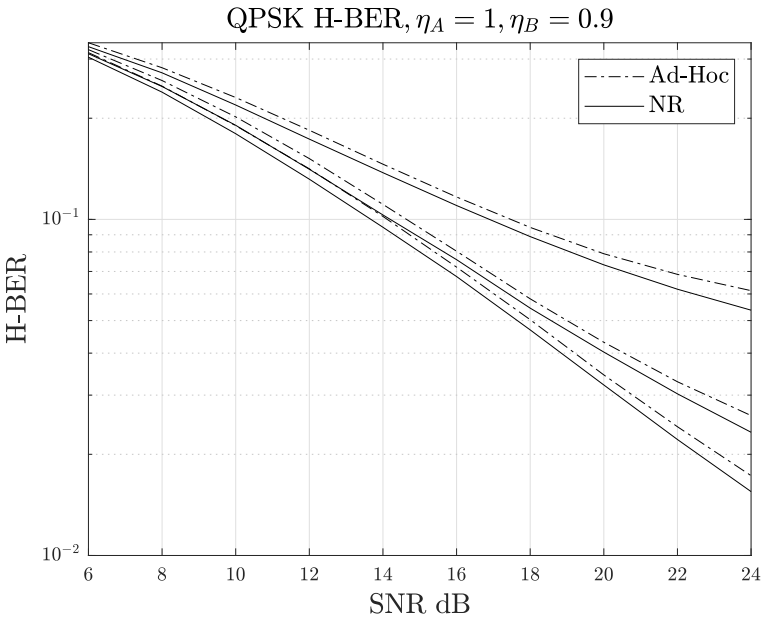


(b) $\eta_B = 0.9$

Figure 5.16: Simulated H-BER for the BPSK alphabet with a XOR HNC map. The performance of three demodulation methods and a set of phase variances $\sigma_\varphi^2 \in \{0, 0.01, 0.04\}$ is shown.



(a) $\eta_B = 0.5$



(b) $\eta_B = 0.9$

Figure 5.17: Simulated H-BER for the QPSK alphabet with a XOR HNC map. The performance of two demodulation methods and a set of phase variances $\sigma_\varphi^2 \in \{0, 0.01, 0.04\}$ is shown.

5.6.2 Over-the-air Verification

For the over-the-air (OTA) verification, we use the same mechanism of generating the useful signal for individual two-point observations as was described in the previous section (5.63). The channel phases $\varphi_{A,n}, \varphi_{B,n}$ are still generated artificially to get a uniform coverage of the full range. The phase noise $\Delta\varphi_{A,n}, \Delta\varphi_{B,n}$ which violates the assumption (5.4) and is an inherent part of a real live scenario arises naturally due to hardware imperfections and the real channel. Our setup allows to use a carrier frequency and symbol timing synchronization between all the nodes using the *OctoClock* clock distribution module which eliminates $\Delta\varphi_{A,n}, \Delta\varphi_{B,n}$. In our setup, we used three *Ettus Research N210* SDR transceivers for nodes S_A, S_B, and R. The experimental setting was made up of three stationary antennas placed in a time-invariant office environment according to Figure 5.18. The parameter setup used for the transmissions is summarized in Table 5.2. For each link S_A-R, S_B-R we used an $N_p = 5 \times 10^5$ symbols long random noise-like pilot sequence and employed a standard ML estimator for estimating the actual values of η_A, η_B and σ_w^2 which are assumed to be known at R.

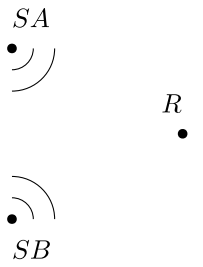
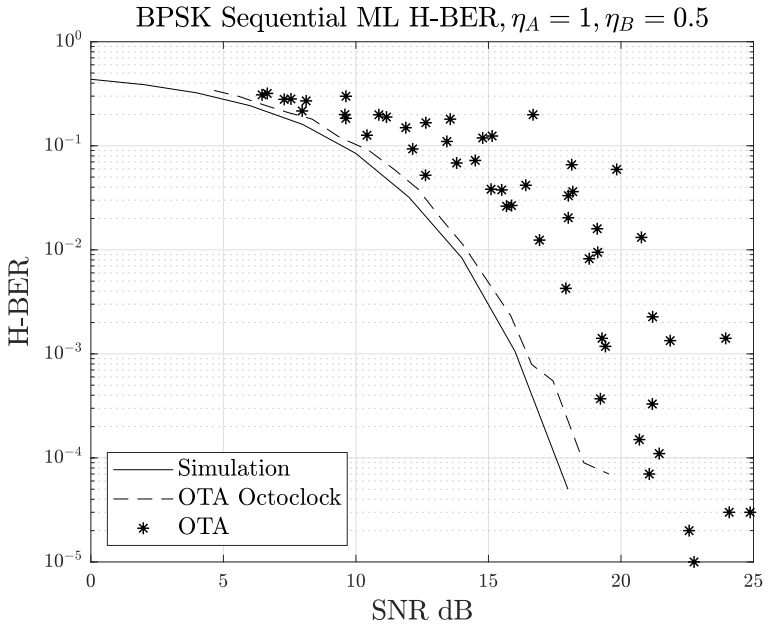


Figure 5.18: Antenna arrangement.

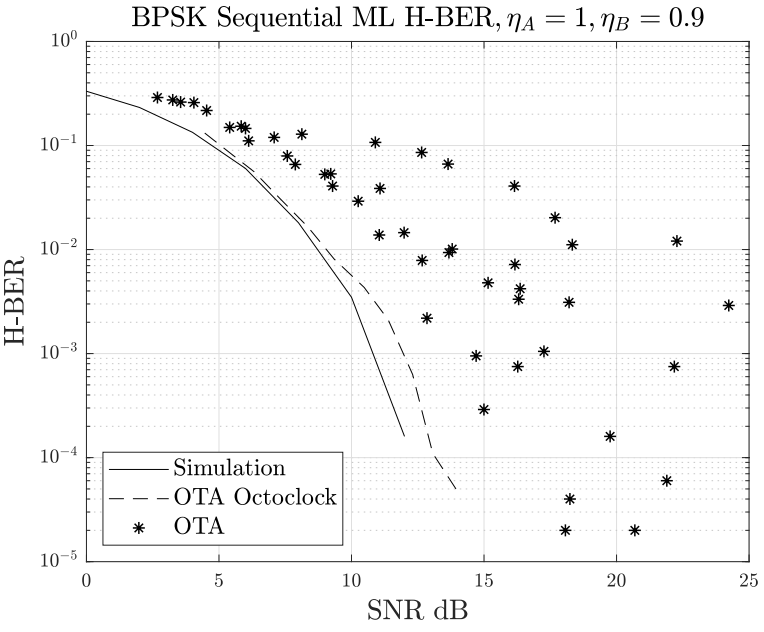
Parameter	Value
carrier frequency	2.4 GHz
sampling frequency	10 MHz
modulation pulse	RRC
roll-off factor	0.35
samples per symbol	8

Table 5.2: Settings used for OTA transmission.

The measured results are presented in Figure 5.19. We plot the H-BER for a BPSK alphabet and the *Sequential ML* method for three cases. The first case is an OTA transmission using the artificial *OctoClock* synchronization (shown by a dashed line), the second case is a fully asynchronous, shown by a set of individual measurements. For reference, we include the simulated performance shown by a solid line. We observe that the synchronous OTA performance is almost identical to the simulation which shows, that the approach is rather robust w.r.t. channel model imperfections. In the non-synchronized case, the H-BER is worse and exhibits a large variance which is caused by the stochastic nature of the phase noise. Such a result is not surprising when compared to the H-BER dependence on σ_φ^2 shown in Figure 5.16. It is important to state that in the OTA experiment without the application of *OctoClock* synchronization, we didn't apply any frequency offset compensation.



(a) $\eta_B = 0.5$



(b) $\eta_B = 0.9$

Figure 5.19: OTA H-BER results for BPSK alphabet with an XOR HNC map.

5.7 Conclusion

This chapter was motivated by the severe impacts of inaccurate CSI on the performance of an H-MAC stage in the context of a WPNC network and the resulting need for effective estimation methods. The main idea was to apply differential encoding to a PSK modulation which would allow for a non-coherent reception and thus substantially reduce the receiver complexity. While the differential encoding at the transmitter is straightforward, the receiver-side processing is very different from a standard single-user case. The main contribution of this chapter was the design of several demodulation methods for an H-MAC with differential PSK modulation.

We considered a two-source H-MAC and focused on a so-called two-point observation. It consists of two consecutive observations which share the same channel coefficients. At the receiver, we considered an a priori knowledge of the channel coefficient magnitudes and no knowledge about their phases. In a practical scenario, the channel phase may change much faster than the magnitude and as such poses higher demands on the CSE and tracking.

The first proposed method (Noisy Reference) is based on the single-user case and extends it naturally to the 2-source H-MAC case. The main idea is to use one observation as a reference to estimate the actual channel phase and the second for estimating the transmitted symbol. We show that unlike in the single-user case, the reference observation leaves certain ambiguity in the channel phase estimate. This in turn introduces an additional degree of freedom which has to be accounted for in the demodulation phase. We show that for the particular case of an XOR HNC map and BPSK alphabets, the geometrical situation of the resulting constellation is favorable and the impact of the phase estimate ambiguity is minimal.

A different method is based on the likelihood maximization of the two-point observation as a whole (Sequential ML). To make this optimization problem solvable we suggest performing it in two successive steps. First, we find the argument of the maximum w.r.t. the unknown channel phases, and in the second step we do the ML w.r.t. the hierarchical target symbol. Even with such a simplification, it was possible to find a closed-form solution only for the BPSK alphabet. In numerical simulations, this method proved to have a slight performance benefit w.r.t. the noisy reference technique in terms of the achievable H-BER rate. When compared with a reference fully coherent scenario the penalty was around 3dB which is comparable with a differential single-user case.

In the rest of the chapter, we presented several ad-hoc methods for the BPSK and QPSK alphabets with simpler processing needed for demodulation. An achievable H-BER comparison showed that the ad-hoc methods perform slightly worse than the systematic ones and are less robust w.r.t. the violation of the assumption of constant channel coefficients during one two-point observation.

In the end, we presented the results of an over-the-air verification of the Sequential ML method for a BPSK alphabet. A comparison with the simulation results shows that the method is rather robust w.r.t. channel model imperfections as long as the carrier frequency and symbol timing synchronization among the sources is maintained.

Chapter 6

Fractional Symbol Duration Asynchronous H-MAC

6.1 Introduction

In this chapter, we address the topic introduced in Section 3.3.3.2 and study a hierarchical MAC (H-MAC) with asynchronous sources. In particular, we consider a situation with two signals which are mutually shifted by a fraction of the symbol duration. We analyze the impacts of such asynchronicity, specifically the intersymbol interference and attenuation, and present numerical results for the case of a root-raised-cosine (RRC) modulation pulse.

Further, we suggest several techniques to mitigate the negative impacts of asynchronicity on the goal of hierarchical demodulation at the relay. First, develop a stochastic model of the interference and use it for the elimination of the interference as a nuisance parameter. As this approach does not prove to be effective in the context of the hierarchical nature of the channel, we suggest a DFT precoding with symbol spaced and double rate processing at the receiver. For each of those variants, we develop an equivalent channel and analyze its properties. In the end, we present numerical evaluations of the achievable hierarchical BER (H-BER) rates of the suggested methods and discuss the results.

6.1.1 Motivation & Related Work

Most of the results on fundamental limits, the system design and performance analysis in basic topologies and scenarios [111], [131, 63, 87, 110] assume relatively idealized assumptions on the synchronization of signals which are received during one H-MAC stage. Usually, a perfect synchronization among all source nodes and the relay is assumed. A violation of this assumption has a strong negative impact on the resulting performance of the system [132, 131, 85]. In a real-world scenario, perfect synchronization is difficult to achieve when not impossible. In the target application of WPNC, we expect to have a large number of small and relatively cheap devices and synchronization on the frame level only can be assumed. The synchronization of multiple nodes is not the only difficulty that makes this problem much more complex than in a single-user case. The possibilities to compensate for a delay are much more restricted as well. First, because the observation in an H-MAC is formed by the superposition of individually delayed signals, even perfect knowledge of the delays at the receiver does not permit a simple compensation w.r.t. all the component signals. For some network topologies uneven path delays w.r.t. different relays, even the knowledge of channel parameters at individual sources does not allow to solve the situation by a simple time base adjustment at the sources as illustrated on Example 3.7.

The delay Δ is usually split into an integer multiple of the symbol duration part and a fractional symbol duration remainder as $\Delta = lT_s + \delta$. In most cases, the first part lT_s is approached by a suitable codebook design (cyclic codes) while the fractional remainder δ is dealt with at the signal processing level. In this chapter, we focus only on the fractional delay δ. Codebook design and techniques to combat the integer multiple of symbol duration delay can be found among others in [70, 117]. Further work on the estimation of the delay parameters and asynchronous reception method considering integer (lT_s) as well as fractional (δ) delays in a WPNC or NOMA context can be found in [71, 77, 118, 132, 24] and in references therein.

6.2 System Model

In this Section, we define the system model used in this chapter. Generally, we aim to solve the situation where there are multiple transmitting sources that are forced to communicate in the same signal subspace without a common time base.

6.2.1 2-source asynchronous H-MAC Channel

We assume two sources S_A, S_B, and one receiver node R. The observation model at R is a memoryless AWGN channel given as

$$x(t) = h_A x_A(t - \delta_A) + h_B x_B(t - \delta_B) + w(t), \tag{6.1}$$

where $h_A, h_B \in \mathbb{C}$ are known channel coefficients, $w(t)$ is a white Gaussian noise process with variance σ_w^2, and $x_A(t)$, $x_B(t)$ are the transmitted signals. Without loss of generality $0 \leq \delta_A \leq \delta_B$ are constant channel delays. We assume the delay parameters δ_A, δ_B to be measured and known to the relay. Because we are interested in fractional symbol time asynchronicity only, we modify the observation model (6.1) and remove any common delay and adjust the signals mutually by a whole multiple of T_s such that

$$x(t) = h_A x_A(t) + h_B x_B(t - \delta) + w(t), \tag{6.2}$$

where $0 \leq \delta \leq \frac{T_s}{2}$.

Both transmitters use a linear modulation given as

$$\begin{aligned} x_A(t) &= \sum_n s(c_{A,n})\, g(t - nT_S) \\ x_B(t) &= \sum_n s(c_{B,n})\, g(t - nT_S), \end{aligned} \tag{6.3}$$

where $s(.)$ is the alphabet mapper, $c_{A,n}, c_{B,n}$ are i.i.d. binary source data and $g(t)$ is a Nyquist modulation pulse w.r.t. the symbol duration period T_s. At R, we apply a matched filter (MF) shifted by $\epsilon\delta, \epsilon \in \langle 0, 1\rangle$ and sample according to Figure (6.1) to obtain

$$x_n(\epsilon) = \int_{-\infty}^{\infty} x(t)\, g^*(t - nT_s - \epsilon\delta)\, \mathrm{d}t. \tag{6.4}$$

It follows, that the two extreme cases $\epsilon = 0$ and $\epsilon = 1$ correspond with synchronous sampling with source S_A and S_B respectively. After substitution of (6.2) for $x(t)$ into (6.4) we get

$$\begin{aligned} x_n(\epsilon) = {} & h_A \sum_{n'} s(c_{A,n'})\, R_g((n' - n)T_s - \epsilon\delta) \\ & + h_B \sum_{n'} s(c_{B,n'})\, R_g((n' - n)T_s + (1 - \epsilon)\delta) \\ & + w_n, \end{aligned} \tag{6.5}$$

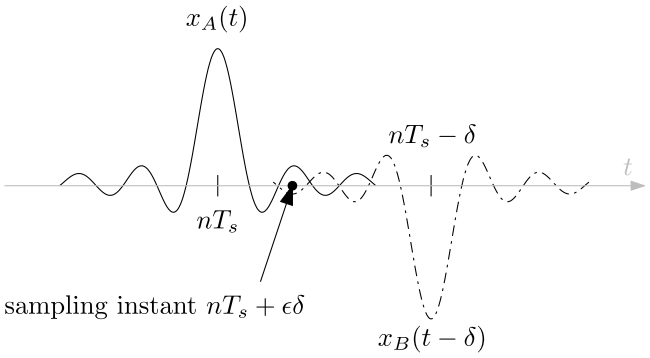


Figure 6.1: Illustration of the sampling instant corresponding to ϵ.

where $w_n \sim CN\left(0, \sigma_w^2\right)$ and

$$R_g(t) = \int_{-\infty}^{\infty} g(\tau)\, g^*(\tau + t)\, \mathrm{d}\tau \tag{6.6}$$

is the autocorrelation function (ACF) of the modulation pulse.

Example 6.1. In Figure 6.2 we show a schematic representation of the interference generation during misaligned MF sampling. Two modulation pulses pass through a MF and a subsequent sampler. The sampler is synchronized with the first pulse (drawn with a dash-dotted curve) such that its whole energy gets collected into a single sample for $n = 0$. The second pulse (drawn with a solid line) is delayed by δ w.r.t. the sampler. As a consequence, part of its energy gets spread over the neighboring samples and gives rise to intersymbol interference.

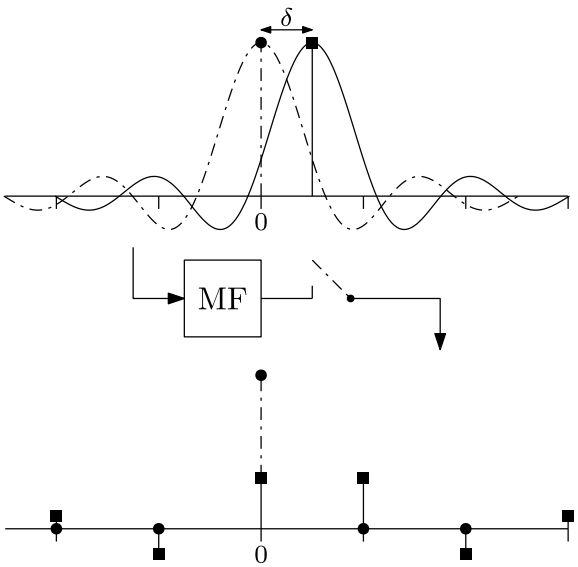


Figure 6.2: Schematic representation of the interference generation during misaligned MF sampling.

Equation (6.5) can be rewritten as

$$\begin{aligned} x_n(\epsilon) = & \underbrace{R_g(\epsilon\delta)\, h_A s(c_{A,n}) + R_g((1-\epsilon)\delta)\, h_B s(c_{B,n})}_{\text{attenuated usefull signal}} \\ & + \underbrace{h_A \sum_{l\neq 0} s(c_{A,n+l})\, R_g(lT_s - \epsilon\delta)}_{\text{ISI from SA}} \\ & + \underbrace{h_B \sum_{l\neq 0} s(c_{B,n+l})\, R_g(lT_s + (1-\epsilon)\delta)}_{\text{ISI from SB}} \\ & + \underbrace{w_n}_{\text{noise}} \end{aligned} \tag{6.7}$$

to stress the two ways in which the observation is impacted by asynchronous reception, namely the overall attenuation of the useful signal and the presence of intersymbol interference.

Practically useful modulation pulses have to be finite or approximately finite signals, which translates to their ACF as well. This motivates the introduction of an approximation where only the central elements of the ACF are considered. The rate of the approximation is given by the parameter m according to $R_g(t) = 0$ for $|t| \geq (m+1)T_s$ which gives us

$$\begin{aligned} x_n(\epsilon) \approx & R_g(\epsilon\delta)\, h_A s(c_{A,n}) + R_g((1-\epsilon)\delta)\, h_B s(c_{B,n}) \\ & + h_A \sum_{l\in\{-m,m\}\setminus\{0\}} s(c_{A,n+l})\, R_g(lT_s - \epsilon\delta) \\ & + h_B \sum_{l\in\{-m,m\}\setminus\{0\}} s(c_{B,n+l})\, R_g(lT_s + (1-\epsilon)\delta) \\ & + w_n. \end{aligned} \tag{6.8}$$

The goal at the relay node R is to decode the hierarchical target codeword symbol c_n which is given by a many-to-one function (Hierarchical Network Code (HNC) map)

$$c_n = \chi_c(c_{A,n}, c_{B,n}). \tag{6.9}$$

In particular, we assume a *minimal* HNC map [111] χ_c given by a bitwise XOR function.

6.2.2 Matrix Description

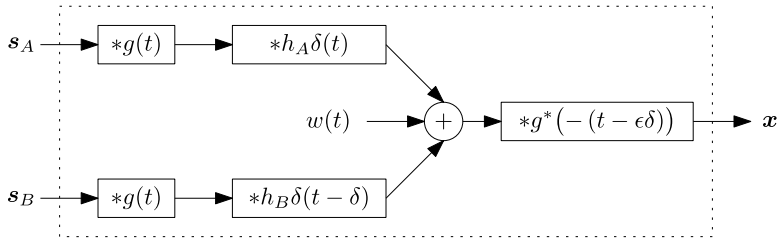
We consider a single frame consisting of $N = 2l-1, l \in \mathbb{N}$ symbols and describe the 2-source asynchronous MAC, introduced in the previous section, using matrix notation. Please note that we work on the level of discrete-time samples, effectively ignoring the physical continuous-time reality of the channel as emphasized in Figure (6.3).

We start by defining the signal column vectors as $\boldsymbol{x}(\epsilon) = [x_n(\epsilon)]_{n=0}^{N-1}$, $\boldsymbol{s}_A = [s(c_{A,n})]_{n=0}^{N-1}$, $\boldsymbol{s}_B = [s(c_{B,n})]_{n=0}^{N-1}$ and the AWGN vector $\boldsymbol{w} = [w_n]_{n=0}^{N-1}$. Equation (6.7) can be rewritten as

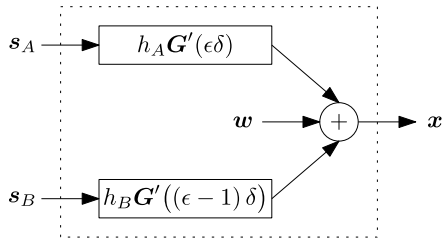
$$\boldsymbol{x}(\epsilon) = h_A \boldsymbol{G}'(\epsilon\delta)\, \boldsymbol{s}_A + h_B \boldsymbol{G}'((\epsilon-1)\delta)\, \boldsymbol{s}_B + \boldsymbol{w}, \tag{6.10}$$

where $\boldsymbol{G}'(.)$ is the $N \times N$ delay matrix containing samples of the ACF $R_g(t)$ according to

$$\boldsymbol{G}'(\tau) = \begin{bmatrix} R_g(-\tau) & R_g(T_s-\tau) & R_g(2T_s-\tau) & R_g(3T_s-\tau) & \dots \\ R_g(-T_s-\tau) & R_g(-\tau) & R_g(T_s-\tau) & R_g(2T_s-\tau) & \dots \\ R_g(-2T_s-\tau) & R_g(-T_s-\tau) & R_g(-\tau) & R_g(T_s-\tau) & \dots \\ R_g(-3T_s-\tau) & R_g(-2T_s-\tau) & R_g(-T_s-\tau) & R_g(-\tau) & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \tag{6.11}$$



(a) Actual continuous-time channel structure.



(b) Discrete-time model.

Figure 6.3: 2-source asynchronous discrete-time channel.

Similarly to the previous section, we consider the same approximation by neglecting small values of the ACF resulting in an approximated delay matrix $\boldsymbol{G}'^{(m)}(\tau)$. After that, (6.8) can be rewritten as

$$\boldsymbol{x}(\epsilon) \approx h_A \boldsymbol{G}'^{(m)}(\epsilon\delta)\,\boldsymbol{s}_A + h_B \boldsymbol{G}'^{(m)}((\epsilon-1)\,\delta)\,\boldsymbol{s}_B + \boldsymbol{w}. \tag{6.12}$$

Example 6.2. For an example with $m = 1$, the approximated delay matrix is of the form

$$\boldsymbol{G}'^{(1)} = \begin{bmatrix} R_g(-\tau) & R_g(T_s-\tau) & 0 & 0 & \dots \\ R_g(-T_s-\tau) & R_g(-\tau) & R_g(T_s-\tau) & 0 & \dots \\ 0 & R_g(-T_s-\tau) & R_g(-\tau) & R_g(T_s-\tau) & \dots \\ 0 & 0 & R_g(-T_s-\tau) & R_g(-\tau) & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \tag{6.13}$$

with all elements equal to zero except the k-diagonals where $k \in \{-m,\dots,m\} = \{-1,0,1\}$.

6.3 Interference Suppression

In this section, we present an approach where we sample at the mid-point [71] between the two transmitted signals ($\epsilon = \frac{1}{2}$) and try to suppress the negative impacts of asynchronous reception. Although it does not lead to an effective method solving for performance degradation, it gives some insight into the problem.

We start with the assumption of a symmetric ACF, that is $R_g(t) = R_g(-t)$ and rearrange (6.8)

$$\begin{aligned} x_n\left(\frac{1}{2}\right) &\approx R_g\left(\frac{\delta}{2}\right)\left[h_A s\left(c_{A,n}\right) + h_B s\left(c_{B,n}\right)\right] \\ &\quad + h_A \sum_{l\in\{-m,m\}\setminus\{0\}} s\left(c_{A,n-l}\right) R_g\left(-lT_s - \frac{\delta}{2}\right) \\ &\quad + h_B \sum_{l\in\{-m,m\}\setminus\{0\}} s\left(c_{B,n+l}\right) R_g\left(lT_s + \frac{\delta}{2}\right) \\ &\quad + w_n \\ &= R_g\left(\frac{\delta}{2}\right)\left[h_A s\left(c_{A,n}\right) + h_B s\left(c_{B,n}\right)\right] \\ &\quad + \sum_{l\in\{-m,m\}\setminus\{0\}} R_g\left(lT_s + \frac{\delta}{2}\right)\left[h_A s\left(c_{A,n-l}\right) + h_B s\left(c_{B,n+l}\right)\right] \\ &\quad + w_n. \end{aligned} \tag{6.14}$$

For the rest of this section, we drop the explicit notation of $\epsilon = \frac{1}{2}$. Equation (6.14) can be rewritten as

$$x_n = K\left[h_A s\left(c_{A,n}\right) + h_B s\left(c_{B,n}\right)\right] + I\left(\boldsymbol{c}_A, \boldsymbol{c}_B\right) + w_n, \tag{6.15}$$

where $K = R_g\left(\frac{\delta}{2}\right)$ is attenuation of the useful signal, $I\left(\boldsymbol{c}_A, \boldsymbol{c}_B\right)$ the interference term, and w_n the noise. In Figure (6.4) we show a plot of the dependence of $|K|$on the delay δ for the case of a RRC pulse with different values of the roll-off factor β. It holds that $|K| \leq 1$ with equality for $\delta = 0$. We observe that the negative impact of the delay gets amplified for larger roll-off factor values. As mentioned in the previous section, the performance of the asynchronous reception is impacted by two major factors, the nonzero intersymbol interference $I\left(\boldsymbol{c}_A, \boldsymbol{c}_B\right)$ and the reduced energy of the useful signal. The approach of interference suppression can not avoid the drop in K. In turn, Figure (6.4) gives us the residual degradation even after a complete elimination of the interference term.

For the signal demodulation, we use a standard likelihood maximization according to the general equation

$$\hat{c}_n = \mathrm{argmax}_{\check{c}} \sum_{\chi_c(c_A,c_B)=\check{c}} \rho\left(x_n - K\left[h_A s\left(c_A\right) + h_B s\left(c_B\right)\right]\right), \tag{6.16}$$

where $\rho\left(x_n - K\left[h_A s\left(c_A\right) + h_B s\left(c_B\right)\right]\right)$ is proportional to the modeled probability of the observation x_n, given the useful signal (fully determined by c_A, c_B).

6.3.1 Random Interference Modeling

The first attempt is to consider symbols $s\left(c_{A,i\neq n}\right)$ and $s\left(c_{B,i\neq n}\right)$ as purely random i.i.d. uniform distributed during the demodulation of the nth symbol. Additionally, we apply the ACF assumption as shown in (6.8) and obtain

$$I^{(m)} = \sum_{l\in\{-m,m\}\setminus\{0\}} \left(\boldsymbol{h}\boldsymbol{u}_l\right) R_g\left(lT_s - \frac{\delta}{2}\right), \tag{6.17}$$

where $\boldsymbol{h} = \begin{bmatrix} h_A & , h_B \end{bmatrix}$ and $\boldsymbol{u}_l$ are i.i.d. random vectors $\boldsymbol{u} = \left[u_1, u_2\right]^{\mathrm{T}}$ with u_1, u_2 independent random variables which take values $\{\pm 1\}$ with equal probability (in case of a BPSK alphabet which is assumed for the rest of this section). The PMF of I in in turn uniform, with a fractal-like structure as shown in Figure (6.5).

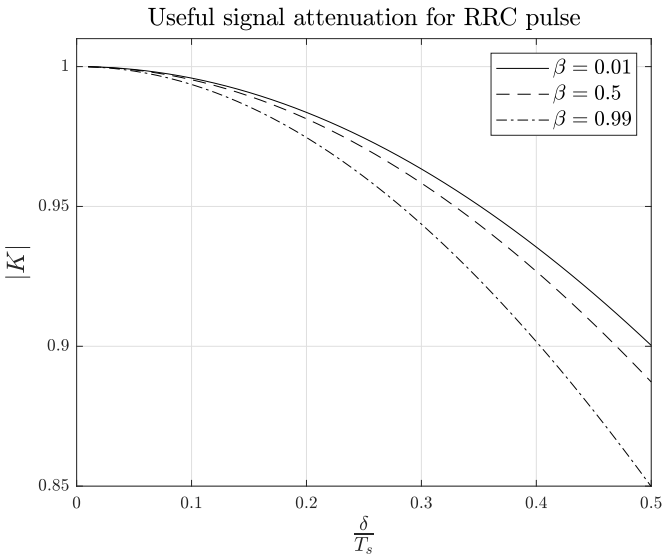


Figure 6.4: Numerical evaluation of $|K|$as a function of the delay δ for a RRC modulation pulse.

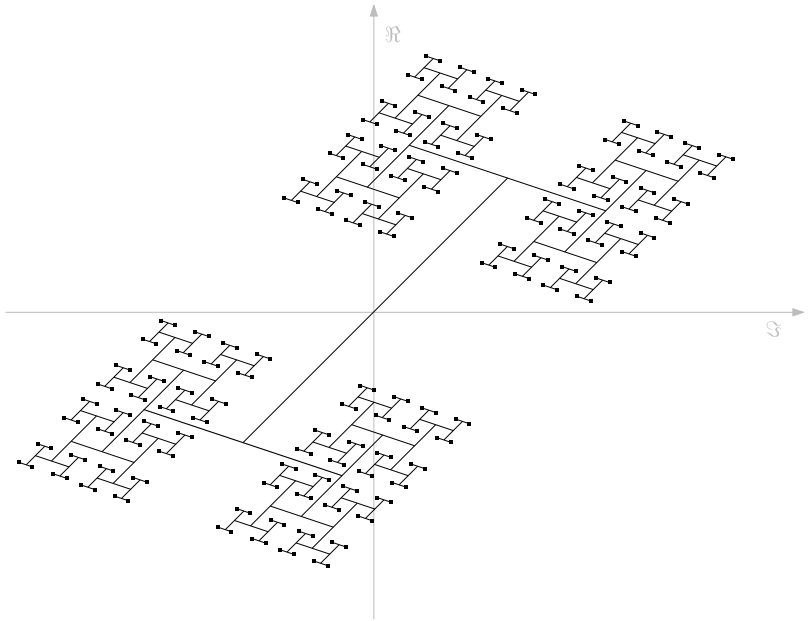


Figure 6.5: Schematic representation of the support of I defined by (6.17) for $m = 2$.

For the whole interference and noise part $I + w_n$, the PDF is given as the convolution of the discrete uniform PMF of I and the Gaussian density. The resulting PDF is a uniform Gaussian mixture given as

$$\rho\left(r_n\right) = \sum_{U_i \in \mathbb{U}} \exp\left(\frac{-1}{\sigma_w^2}\left|r_n - \boldsymbol{h} U_i \boldsymbol{R}_g\right|^2\right), \tag{6.18}$$

where

$$\boldsymbol{R}_g = \left[R_g\left(-mT_s - \frac{\delta}{2}\right), \ldots, R_g\left(-T_s - \frac{\delta}{2}\right), R_g\left(T_s - \frac{\delta}{2}\right), \ldots, R_g\left(mT_s - \frac{\delta}{2}\right)\right]^{\mathrm{T}} \tag{6.19}$$

and $\mathbb{U}$ is the set of $2 \times 2m$ matrices with elements from $\{\pm 1\}$. Since for each $U_i \in \mathbb{U}$ there is also an $-U_i \in \mathbb{U}$, the expression (6.18) can be simplified by joining together the corresponding pairs as

$$\begin{aligned}
&\exp\left(\frac{-1}{\sigma_w^2}\left|r - \boldsymbol{h} U \boldsymbol{R}_g\right|^2\right) + \exp\left(\frac{-1}{\sigma_w^2}\left|r - \boldsymbol{h}\left(-U\right)\boldsymbol{R}_g\right|^2\right) \\
=&\exp\left(\frac{-1}{\sigma_w^2}\left(|r|^2 + |\boldsymbol{h} U \boldsymbol{R}_g|^2\right)\right)\left[\exp\left(\frac{-2}{\sigma_w^2}\Re\left\{r \boldsymbol{h}^* U \boldsymbol{R}_g\right\}\right) + \exp\left(\frac{2}{\sigma_w^2}\Re\left\{r \boldsymbol{h}^* U \boldsymbol{R}_g\right\}\right)\right] \\
=&2\exp\left(\frac{-1}{\sigma_w^2}\left(|r|^2 + |\boldsymbol{h} U \boldsymbol{R}_g|^2\right)\right)\cosh\left(\frac{2}{\sigma_w^2}\Re\left\{r \boldsymbol{h}^*\right\} U \boldsymbol{R}_g\right).
\end{aligned} \tag{6.20}$$

The resulting PDF can be rewritten as

$$\rho\left(r_n\right) = \exp\left(\frac{-1}{\sigma_w^2}\left|r_n\right|^2\right) \sum_{U_i \in \bar{\mathbb{U}}} \exp\left(\frac{-1}{\sigma_w^2}\left|\boldsymbol{h} U_i \boldsymbol{R}_g\right|^2\right) \cosh\left(\frac{2}{\sigma_w^2}\Re\left\{r_n \boldsymbol{h}^*\right\} U_i \boldsymbol{R}_g\right), \tag{6.21}$$

where $\bar{\mathbb{U}}$ is the quotient set of $\mathbb{U}$ by the relation $\sim$ where $U_i \sim U_j$ iff $U_i = -U_j$. After plugging (6.21) into (6.16) we get the final maximum likelihood (ML) demodulation equation as

$$\begin{aligned}
\hat{c}_n = \operatorname{argmax}_{\check{c}} &\sum_{\chi_c(c_A, c_B) = \check{c}} \exp\left(\frac{-1}{\sigma_w^2}\left|x_n - K\left[h_A s\left(c_A\right) - h_B s\left(c_B\right)\right]\right|^2\right) \\
&\times \sum_{U_i \in \bar{\mathbb{U}}} \exp\left(\frac{-1}{\sigma_w^2}\left|\boldsymbol{h} U_i \boldsymbol{R}_g\right|^2\right) \cosh\left(\frac{2}{\sigma_w^2}\Re\left\{\left(x_n - K\left[h_A s\left(c_A\right) - h_B s\left(c_B\right)\right]\right) \boldsymbol{h}^*\right\} U_i \boldsymbol{R}_g\right).
\end{aligned} \tag{6.22}$$

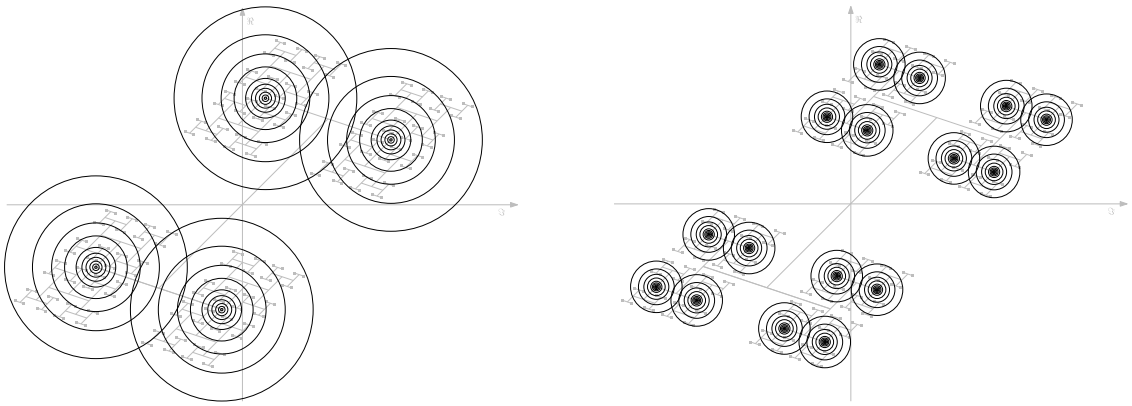
Motivated by Figure (6.6), it seems appropriate to set the approximation parameter m according on the additive noise variance.

6.3.2 Full Data-Aided Interference Compensation

A more promising approach is a data-aided solution. The idea is to respect the dependence of the interference on the neighboring symbols. In the first approach we consider an imaginary, genie-aided scenario, where at the time of the nth symbol demodulation the m preceding and m subsequent symbols are expected to be known. We can imagine a practical extension of such a scenario based on an iterative technique, where a soft decision metric on the neighboring symbols could be used to enhance the estimate of the current symbol.

In this approach, the interference term $I\left(\boldsymbol{c}_A, \boldsymbol{c}_B\right)$ can be approximated by

$$\sum_{l \in \{-m, m\} \setminus \{0\}} R_g\left(lT_s - \frac{\delta}{2}\right)\left[h_A s\left(c_{A,n-l}\right) + h_B s\left(c_{B,n+l}\right)\right]. \tag{6.23}$$



(a) Case with large σ_w^2. (b) Case with small σ_w^2.

Figure 6.6: Graphical representation of the Gaussian mixture approximation for the PDF (6.21) of $I + w_n$.

The general metric from (6.16) is given as

$$\rho\left(r_n\right) = \exp\left(\frac{-1}{\sigma_w^2}\left|r_n - \sum_{l\in\{-m,m\}\setminus\{0\}} R_g\left(lT_s - \frac{\delta}{2}\right)\left[h_A s\left(c_{A,n-l}\right) + h_B s\left(c_{B,n+l}\right)\right]\right|^2\right) \tag{6.24}$$

and in turn

$$\begin{aligned}\hat{c}_n = \operatorname{argmax}_{\check{c}} \sum_{\chi_c(c_A,c_B)=\check{c}} \exp\Bigg(\frac{-1}{\sigma_w^2}\Bigg|x_n - K\left[h_A s\left(c_A\right) - h_B s\left(c_B\right)\right] - \\ \sum_{l\in\{-m,m\}\setminus\{0\}} R_g\left(lT_s - \frac{\delta}{2}\right)\left[h_A s\left(c_{A,n-l}\right) + h_B s\left(c_{B,n+l}\right)\right]\Bigg|^2\Bigg).\end{aligned} \tag{6.25}$$

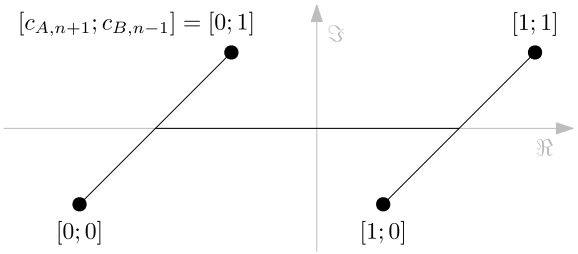
6.3.3 Hierarchical Data-Aided Interference Compensation

Because a WPNC setting is considered, we can not expect to have the full data aid nor an information metric on it. At the relay, the goal is to demodulate the hierarchical target symbol only. Hence we need to update the previous section only considering a hierarchical target data aid, in particular the knowledge of the XOR function of the neighboring symbols. The ML demodulator expression stays the same except for the degree to which we are able to model the interference. Since the HNC map is many-to-one we experience the hierarchical dispersion not only in the final marginalization of the likelihood but also while modeling the interference. This case is somewhere between the fully random interference model and the full data-aided case.

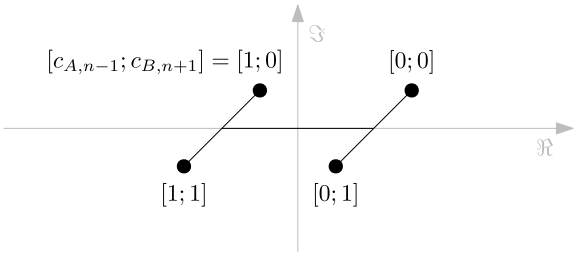
The demodulator equation is the same as (6.22) with a reduction of the set $\mathbb{U}$ to only contain elements that are consistent with the values of c_{n+l}, $l \in \{-m, m\} \setminus \{0\}$.

Example 6.3. We explain the statement "to reduce the set $\mathbb{U}$ to only contain elements that are consistent with the values of c_{n+l}, $l \in \{-m, m\} \setminus \{0\}$" in terms of an example with $m = 1$. In equation (6.14) it was shown, that the interference term $I\left(\left[c_{n-1}, c_{n+1}\right]^T\right)$ is equal to a linear combination of $\left[h_A s\left(c_{A,n-l}\right) + h_B s\left(c_{B,n+l}\right)\right]$ weighted by the ACF samples. The elements $\left[h_A s\left(c_{A,n-l}\right) + h_B s\left(c_{B,n+l}\right)\right]$ look almost like the useful signal, except for the index mismatch between $s\left(c_{A,n-l}\right)$ and $s\left(c_{B,n+l}\right)$. Still, it is useful to perceive $\left[h_A s\left(c_{A,n-l}\right) + h_B s\left(c_{B,n+l}\right)\right]$ as a

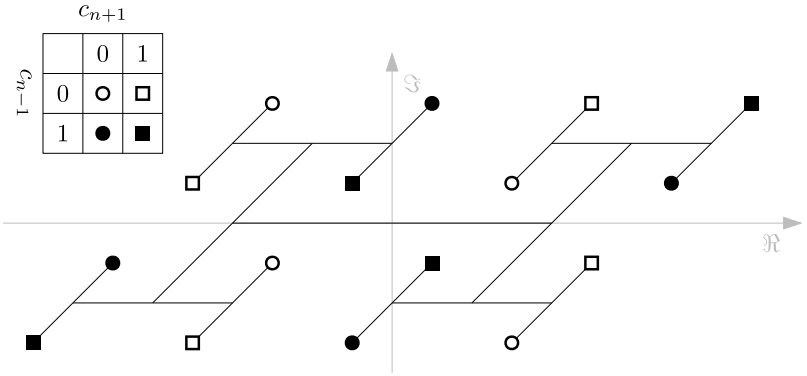
virtual constellation as depicted in Figures 6.7a and 6.7b. The possible values of the resulting interference are shown in Figure 6.7c. With the knowledge of the hierarchical symbol help, the set of possible interference values can be narrowed down. In Figure 6.7c the 4 possible subsets according to the values of $[c_{n-1}, c_{n+1}]^{\mathrm{T}}$ are labeled by various markers.



(a) $R_g\left(-T_s - \frac{\delta}{2}\right)\left[h_A s\left(c_{A,n-l}\right) + h_B s\left(c_{B,n+l}\right)\right]$



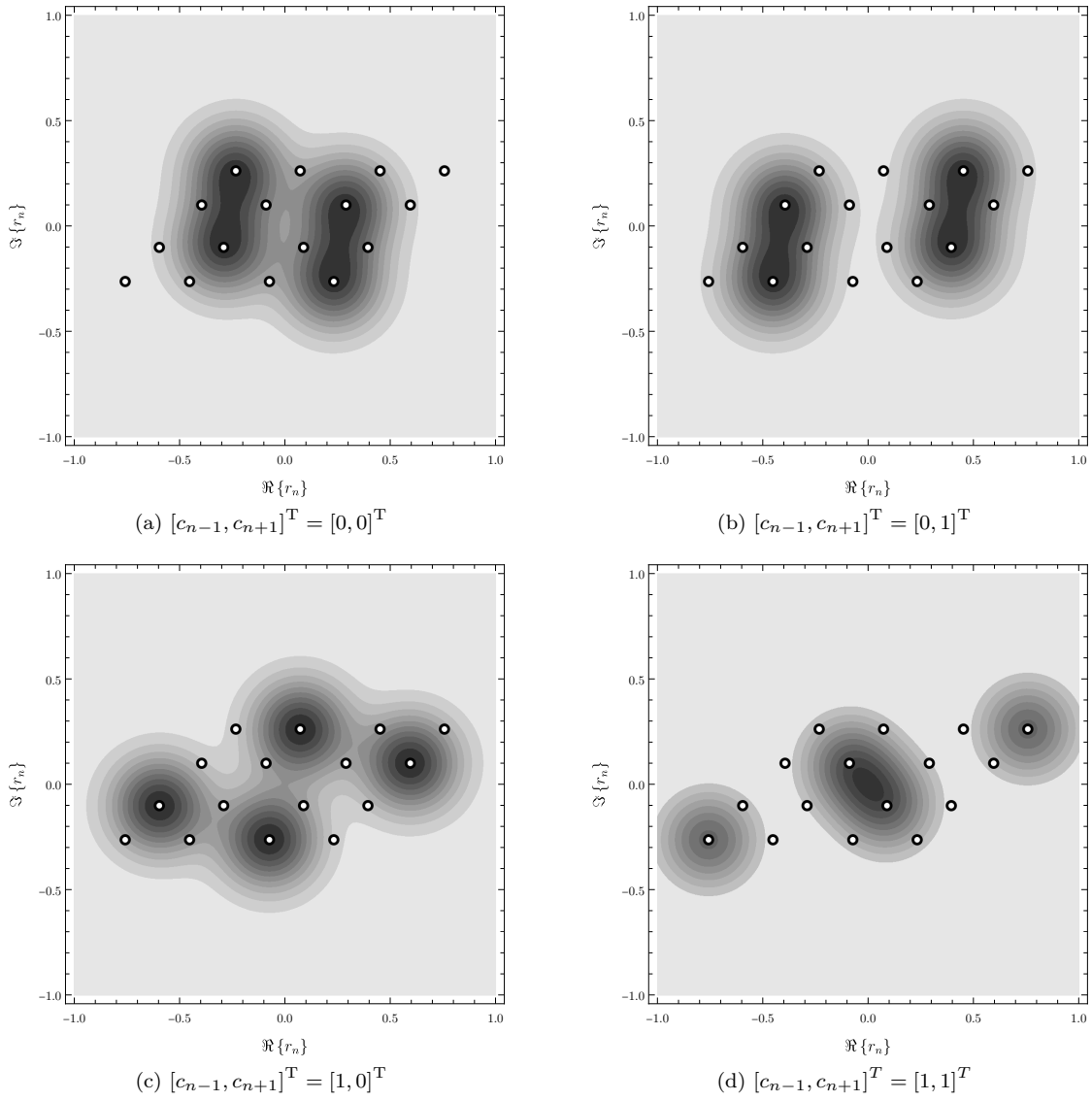
(b) $R_g\left(T_s - \frac{\delta}{2}\right)\left[h_A s\left(c_{A,n+l}\right) + h_B s\left(c_{B,n-l}\right)\right]$



(c) Resulting interference pattern $I^{(m=1)}$ based on the hierarchical target symbol aid.

Figure 6.7: Example of hierarchical data aided interference model with two neighboring symbol aid.

To obtain the interference and noise PDF, we simply convolve the uniform PMF of $I\left(\left[c_{n-1}, c_{n+1}\right]^{\mathrm{T}}\right)$ with the Gaussian density. The resulting densities are shown in Figure 6.8 for the four possible combinations of $\left[c_{n-1}, c_{n+1}\right]^{T}$. The included points represent the support of the discrete distribution of I. Please note the correspondence between Figure 6.7c and Figure 6.8.



(a) $[c_{n-1}, c_{n+1}]^{\mathrm{T}} = [0,0]^{\mathrm{T}}$

(b) $[c_{n-1}, c_{n+1}]^{\mathrm{T}} = [0,1]^{\mathrm{T}}$

(c) $[c_{n-1}, c_{n+1}]^{\mathrm{T}} = [1,0]^{\mathrm{T}}$

(d) $[c_{n-1}, c_{n+1}]^{T} = [1,1]^{T}$

Figure 6.8: Interference + noise PDF given different values of the XOR aid with $h_A = 1, h_B = 0.75\mathrm{e}^{\frac{\mathrm{j}\pi}{4}}, \delta = 0.3T_s$ and $m = 1$.

6.3.4 Numerical Results

In this section, we present the dependence of the achievable H-BER on the relative delay δ. The results are plotted in Figure 6.9 which contains the evaluation of the three introduced suppression methods. The solid lines indicate the results of a full data aid (Section 6.3.2) for various values of m. We observe that there is a huge performance gain between $m = 0$ and $m = 1$ which signifies a big impact of the interference coming from the two neighboring symbols. From the curve corresponding to $m = \max$, we observe that even though the interference term $I(\boldsymbol{c}_A, \boldsymbol{c}_B)$ gets completely removed, some performance drop is still present for $\delta \neq 0$. This is related to the behavior of $|K|$ as shown in Figure 6.4. The dotted and dashed lines indicate the results of a XOR-aided approach (Section 6.3.3) and fully random modeling (Section 6.3.1) respectively. We conclude, that those techniques

are not efficient and result only in a minor improvement. An intuitive understanding of why this is the case can be found in Figure 6.8. We observe that the restriction of the possible values of I based on the XOR-aid does not limit the variance of the resulting PDF significantly and does not help much with the ML decision.

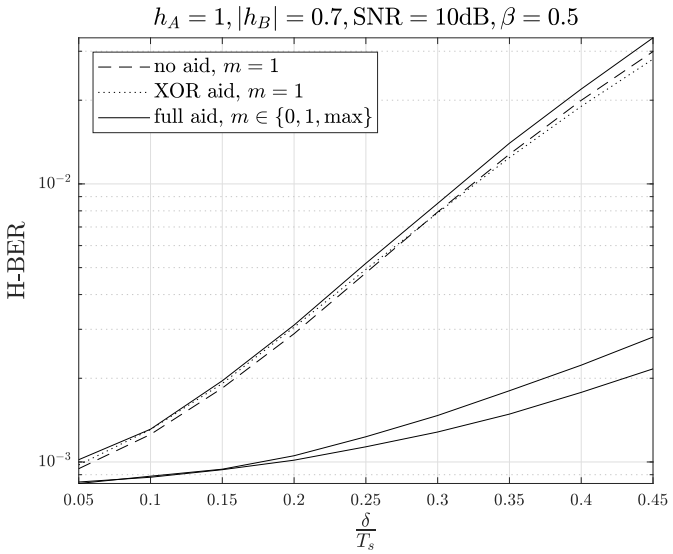


Figure 6.9: Results of a simulation using a RRC pulse with roll-off factor $\beta = 0.5$. The H-BER is averaged over a uniformly distributed relative channel phase $\angle h_B$. The solid lines represent the H-BER with a full data aid for different values of the approximation parameter: $m = 0$ - no interference suppression gives the worst performance, $m = 1$ - interference suppression given the full knowledge of one neighboring symbol, $m = \max$ - fully removing the interference gives the best performance limited only by the dropping magnitude of K.

6.4 Linear Precoding

In this section, we make use of the finite frame matrix description introduced in Section 6.2.2 and suggest a linear precoding at the transmitters which makes it easy to compensate for the delay δ at the receiver.

6.4.1 Circular Extension

We start by looking at equation (6.12) and focus on the delay matrix. The original delay matrix $\boldsymbol{G}'$ is Toeplitz which means that each row is a shifted version of the row above, where a new element gets added on the left side and one element gets dropped on the right side of the row.

Example 6.4. The general structure of a square $N \times N$ Toeplitz matrix is given as follows:

$$\begin{bmatrix} a_0 & a_1 & a_2 & a_3 & \dots & a_{N-2} & a_{N-1} \\ a_{-1} & a_0 & a_1 & a_2 & \dots & a_{N-3} & a_{N-2} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{-(N-1)} & a_{-(N-2)} & a_{-(N-3)} & a_{-(N-4)} & \dots & a_{-1} & a_0 \end{bmatrix}. \tag{6.26}$$

A special type of Toeplitz matrices are the so-called circulant matrices, whose rows are related in terms of a circular shift operation. A circulant matrix is fully defined by its first row, as the

remaining rows can be obtained as cyclically shifted versions thereof. As will be explained in the next section, it would be favorable if the delay matrix would be circulant.

Example 6.5. The general structure of a square $N \times N$ circulant matrix is given as follows:

$$\begin{bmatrix} a_0 & a_1 & a_2 & a_3 & \dots & a_{N-2} & a_{N-1} \\ a_{N-1} & a_0 & a_1 & a_2 & \dots & a_{N-3} & a_{N-2} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_1 & a_2 & a_3 & a_4 & \dots & a_{N-1} & a_0 \end{bmatrix}. \tag{6.27}$$

We recall that after the ACF approximation, the general delay matrix changes to $\boldsymbol{G}'^{(m)}$. We notice that (depending on the value of m) the matrix $\boldsymbol{G}'^{(m)}$ is almost circulant with only a few rows not respecting the circulant property.

Example 6.6. Example of $\boldsymbol{G}'^{(1)}$ with elements not respecting the circulant property indicated by a cross:

$$\begin{bmatrix} R_g(-\tau) & R_g(T_s-\tau) & 0 & 0 & \cancel{0} \\ R_g(-T_s-\tau) & R_g(-\tau) & R_g(T_s-\tau) & 0 & 0 \\ 0 & R_g(-T_s-\tau) & R_g(-\tau) & R_g(T_s-\tau) & 0 \\ 0 & 0 & R_g(-T_s-\tau) & R_g(-\tau) & R_g(T_s-\tau) \\ \cancel{0} & 0 & 0 & R_g(-T_s-\tau) & R_g(-\tau) \end{bmatrix}. \tag{6.28}$$

We modify matrix $\boldsymbol{G}'^{(m)}$ to become circulant by changing the small number of elements which do not fulfill the circulant property and denote the resulting matrix as $\boldsymbol{G}^{(m)}$.

Example 6.7. Example of $\boldsymbol{G}^{(1)}$ resulting from a modification of $\boldsymbol{G}'^{(1)}$ such that $\boldsymbol{G}^{(1)}$ is a circulant matrix:

$$\begin{bmatrix} R_g(-\tau) & R_g(T_s-\tau) & 0 & 0 & R_g(-T_s-\tau) \\ R_g(-T_s-\tau) & R_g(-\tau) & R_g(T_s-\tau) & 0 & 0 \\ 0 & R_g(-T_s-\tau) & R_g(-\tau) & R_g(T_s-\tau) & 0 \\ 0 & 0 & R_g(-T_s-\tau) & R_g(-\tau) & R_g(T_s-\tau) \\ R_g(T_s-\tau) & 0 & 0 & R_g(-T_s-\tau) & R_g(-\tau) \end{bmatrix}. \tag{6.29}$$

It can be shown that when altering the transmitted vector $\boldsymbol{s}_A$ by prepending m last elements $s_{A,N-m}, \dots, s_{A,N-1}$ at the beginning (as elements $s_{A,-m}, \dots, s_{A,-1}$ respectively) and appending the first m elements $s_{A,0}, \dots, s_{A,m-1}$ at the end (as elements $s_{A,N}, \dots, s_{A,N+m-1}$ respectively), the output of the H-MAC becomes equal to $\boldsymbol{G}^{(m)}\boldsymbol{s}_A$ (Example 6.8). The same holds for the signal from the second source $\boldsymbol{s}_B$. We call this modification the *cyclic extension.* From now on we will assume such an cyclic extension to be applied, and model the physical channel by the corresponding circulant delay matrices. For ease of notation, we will drop the explicit specification of the approximation parameter m and write simply $\boldsymbol{G}$ instead of $\boldsymbol{G}^{(m)}$.

Example 6.8. Example of the cyclic extension:

$$\overbrace{\begin{bmatrix} X(-T_s) & X(0) & X(T_s) & 0 & 0 & 0 \\ 0 & X(-T_s) & X(0) & X(T_s) & 0 & 0 \\ 0 & 0 & X(-T_s) & X(0) & X(T_s) & 0 \\ 0 & 0 & 0 & X(-T_s) & X(0) & X(T_s) \end{bmatrix}}^{\boldsymbol{G}'^{(1)}} \begin{bmatrix} s_4 \\ s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_1 \end{bmatrix} = \tag{6.30}$$

$$\underbrace{\begin{bmatrix} X(0) & X(T_s) & 0 & X(-T_s) \\ X(-T_s) & X(0) & X(T_s) & 0 \\ 0 & X(-T_s) & X(0) & X(T_s) \\ X(T_s) & 0 & X(-T_s) & X(0) \end{bmatrix}}_{\boldsymbol{G}^{(1)}} \underbrace{\begin{bmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \end{bmatrix}}_{\boldsymbol{s}},$$

where $X(t) = R_g(t - \tau)$ to make the notation simpler.

6.4.2 Eigenvalue Decomposition

Having a circulant delay matrix, we can make use of the eigenvector property of circulant matrices. Namely, it holds that eigenvectors of circulant matrices do not depend on the matrix element values [27] and are given solely by the circulant structure of the matrix. It can be shown that the set of normalized eigenvectors of a circulant $N \times N$ matrix is given as the rows of the DFT matrix $\boldsymbol{U}$.

Example 6.9. Example of the DFT matrix $\boldsymbol{U}$:

$$\frac{1}{\sqrt{N}} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & \dots & 1 \\ 1 & \omega & \omega^2 & \omega^3 & \omega^4 & \dots & \omega^{N-1} \\ 1 & \omega^2 & \omega^4 & \omega^6 & \omega^8 & \dots & \omega^{2(N-1)} \\ 1 & \omega^3 & \omega^6 & \omega^9 & \omega^{12} & \dots & \omega^{3(N-1)} \\ 1 & \omega^4 & \omega^8 & \omega^{12} & \omega^{16} & \dots & \omega^{4(N-1)} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{N-1} & \omega^{2(N-1)} & \omega^{3(N-1)} & \omega^{4(N-1)} & \dots & \omega^{(N-1)(N-1)} \end{bmatrix}, \tag{6.31}$$

where $\omega = \mathrm{e}^{-\frac{2\pi \mathrm{j}}{N}}$. We notice that $\boldsymbol{U}$ is symmetric ($\boldsymbol{U}^{\mathrm{T}} = \boldsymbol{U}$) and unitary ($\boldsymbol{U}^{\mathrm{H}} = \boldsymbol{U}^{-1}$). As consequence, it holds that $\boldsymbol{U}^* = \boldsymbol{U}^{-1}$.

The eigenvalue decomposition of the delay matrix follows as

$$\boldsymbol{G}(\tau) = \boldsymbol{U}\boldsymbol{\Sigma}(\tau)\boldsymbol{U}^*, \tag{6.32}$$

where $\boldsymbol{\Sigma}(\tau)$ is the diagonal matrix containing the eigenvalues $\{\lambda_0(\tau), \dots, \lambda_{N-1}(\tau)\}$. To understand the meaning of the eigenvalues better, we can multiply (6.32) by $\boldsymbol{U}^*$ from the left and by $\boldsymbol{U}$ from the right obtaining

$$\boldsymbol{\Sigma}(\tau) = \boldsymbol{U}^*\boldsymbol{G}(\tau)\boldsymbol{U}. \tag{6.33}$$

The RHS multiplication by $\boldsymbol{U}$ performs the DFT to the rows of $\boldsymbol{G}(\tau)$. Because the rows are circularly shifted w.r.t. each other, they contain modulated versions of the first row's DFT. After the multiplication by $\boldsymbol{U}^*$ from the left, all except the diagonal elements cancel out and the result is the diagonal matrix $\boldsymbol{\Sigma}(\tau)$. This way it is possible to relate the diagonal elements of $\boldsymbol{\Sigma}(\tau)$, the eigenvalues, to the DFT of the first row of $\boldsymbol{G}(\tau)$. Because the delay matrix $\boldsymbol{G}(\tau)$ is real-valued, the eigenvalues exhibit even symmetry

$$\lambda_i = \lambda^*_{(N-i) \mod N}. \tag{6.34}$$

Which in turn implies that λ_0 is real-valued and equal to the sum of the first row of $\boldsymbol{G}(\tau)$.

Example 6.10. In this example for $N = 4$, we show the correspondence between the eigenvalues and the DFT of the first row of $\boldsymbol{G}$. We start with

$$\boldsymbol{GU} = \begin{bmatrix} g_0 & g_1 & g_2 & g_3 \\ g_3 & g_0 & g_1 & g_2 \\ g_2 & g_3 & g_0 & g_1 \\ g_1 & g_2 & g_3 & g_0 \end{bmatrix} \boldsymbol{U} = \frac{1}{\sqrt{4}} \begin{bmatrix} d_0 & d_1 & d_2 & d_3 \\ d_0 & \omega d_1 & \omega^2 d_2 & \omega^3 d_3 \\ d_0 & \omega^2 d_1 & \omega^4 d_2 & \omega^6 d_3 \\ d_0 & \omega^3 d_1 & \omega^6 d_2 & \omega^9 d_3 \end{bmatrix}, \tag{6.35}$$

where $\boldsymbol{d} = \mathrm{DFT}\{\boldsymbol{g}\}$. After the multiplication by $\boldsymbol{U}^*$ from the left we have

$$\begin{aligned} \boldsymbol{U}^*\boldsymbol{GU} &= \frac{1}{\sqrt{4}} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & \omega^{-1} & \omega^{-2} & \omega^{-3} \\ 1 & \omega^{-2} & \omega^{-4} & \omega^{-6} \\ 1 & \omega^{-3} & \omega^{-6} & \omega^{-9} \end{bmatrix} \frac{1}{\sqrt{4}} \begin{bmatrix} d_0 & d_1 & d_2 & d_3 \\ d_0 & \omega d_1 & \omega^2 d_2 & \omega^3 d_3 \\ d_0 & \omega^2 d_1 & \omega^4 d_2 & \omega^6 d_3 \\ d_0 & \omega^3 d_1 & \omega^6 d_2 & \omega^9 d_3 \end{bmatrix} \\ &= \frac{1}{4} \begin{bmatrix} 4d_0 & 0 & 0 & 0 \\ 0 & 4d_1 & 0 & 0 \\ 0 & 0 & 4d_2 & 0 \\ 0 & 0 & 0 & 4d_3 \end{bmatrix} = \begin{bmatrix} \lambda_0 & 0 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_2 & 0 \\ 0 & 0 & 0 & \lambda_3 \end{bmatrix}. \end{aligned} \tag{6.36}$$

When applied to the delay matrix $\boldsymbol{G}$, the first row is given as

$$\begin{aligned} \boldsymbol{g} = \big[&R_g(-\tau), R_g(T_s - \tau), R_g(2T_s - \tau), \ldots, R_g(mT_s - \tau), 0, \ldots, \\ &0, R_g(-mT_s - \tau), \ldots, R_g(-2T_s - \tau), R_g(-T_s - \tau)\big]. \end{aligned} \tag{6.37}$$

As such, the λ_i are directly related to the energy spectrum of the modulation pulse $g(t)$. Let us denote the energy spectrum as

$$C(f) = \mathcal{F}\{R_g(t)\} \tag{6.38}$$

and the τ modulated energy spectrum follows as

$$C^\tau(f) = \mathcal{F}\{R_g(t+\tau)\} = C(f)\,\mathrm{e}^{2\pi \mathrm{j} f\tau}. \tag{6.39}$$

The central row of $\boldsymbol{G}(\tau)$ is a circularly shifted version of $\boldsymbol{g}$ such that it contains the samples of $R_g(t-\tau)$ with the sampling frequency $f_s = \frac{1}{T_s}$. The spectrum of the sampled ACF $C(F)$ is related to $C^\tau(f)$ according to

$$\begin{aligned} C^\tau(F) &= \mathrm{DtFT}\left\{R_g^\tau(nT_s)\right\} \\ &= \frac{1}{T_s}\sum_{k=-\infty}^{-\infty} C^\tau\left(\frac{F-k}{T_s}\right) \\ &= \frac{1}{T_s}\sum_{k=-\infty}^{-\infty} C\left(\frac{F-k}{T_s}\right)\mathrm{e}^{2\pi \mathrm{j}\tau\frac{F-k}{T_s}}. \end{aligned} \tag{6.40}$$

In the approximation (6.12), we put $R_g(nT_s - \tau) = 0$ for all $|n| > m$, that is we multiply the sampled ACF by a rectangular window with the width of $2m+1$, centered around 0. This results in a convolution with the DtFT of the window

$$S_W^{(m)}(F) = \sum_{k=-m}^{m} \mathrm{e}^{-2\pi \mathrm{j} Fk} = \frac{\sin\left(2\pi F\left(m+\frac{1}{2}\right)\right)}{\sin(\pi F)} \tag{6.41}$$

as

$$S^{\tau,(m)}(F) = C^\tau(F) * S_W^{(m)}(F). \tag{6.42}$$

Alternatively, the final spectrum can be obtained according to

$$S^{\tau,(m)}(F) = \sum_{k=-m}^{m} R_g(mT_s + \tau)\, \mathrm{e}^{2\pi kF}. \tag{6.43}$$

It follows that the DFS of the central row of matrix $\boldsymbol{G}(\tau)$ is equal to the samples of $S^{\tau,(m)}(F)$. The final step is to relate those samples to the eigenvalues

$$\begin{aligned} \lambda_i &= d_i \\ &= \frac{1}{\sqrt{N}} \mathrm{e}^{\frac{2\pi \mathrm{j}}{N}\left(\frac{N-1}{2}+1\right)i} S^{-\tau,(m)}\left(\frac{1}{N}i\right) \\ &= \frac{1}{\sqrt{N}} \mathrm{e}^{\pi \mathrm{j} \frac{N+1}{N} i} S^{-\tau,(m)}\left(\frac{1}{N}i\right). \end{aligned} \tag{6.44}$$

6.4.3 Eigenvalues for RRC pulse

In this section, we consider the class of Nyquist RRC pulses with various roll-off factors and examine the energy spectrum and the resulting eigenvalues as described in the previous section. The delay matrix $\boldsymbol{G}(\tau)$ and its set of eigenvalues depend on the ACF, or equivalently on the energy spectrum of the pulse. That is why it is sufficient to only consider the second-order characteristics, given as

$$R_g(t) = T_s^3 \sin\left(\frac{\pi t}{T_s}\right) \frac{\cos\left(\beta \frac{\pi t}{T_s}\right)}{\pi t \left(T_s^2 - 4\beta^2 t^2\right)} \tag{6.45}$$

$$C(f) = \begin{cases} T_s & |f| < \frac{1-\beta}{2T_s} \\ T_s \cos^2\left(\frac{\pi T_s}{2\beta}\left(|f| - \frac{1-\beta}{2T_s}\right)\right) & \frac{1-\beta}{2T_s} \le |f| < \frac{1+\beta}{2T_s} \\ 0 & \frac{1+\beta}{2T_s} \le |f| \end{cases} \tag{6.46}$$

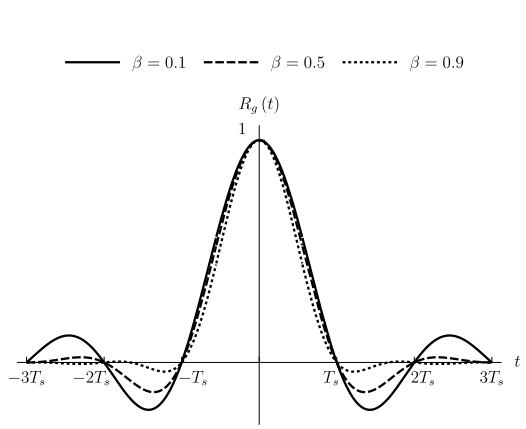
and depicted in Figure 6.10.

When sampling the ACF with a sampling frequency of $f_s = \frac{1}{T_s}$, the resulting energy spectrum gets periodic according to (6.40). Because the RRC pulse is Nyquist w.r.t. T_s, the overlapping spectral replicas add up to a constant spectrum (Figure 6.11a). This is true only for the case with synchronous sampling ($\tau = 0$). If there is a misalignment, individual spectral replicas get multiplied by a complex exponential and the resulting spectrum $C^{\tau}(F)$ becomes complex. An example of several complex replicas $\frac{1}{T_s} C^{\tau}\left(\frac{F-k}{T_s}\right) = \frac{1}{T_s} C\left(\frac{F-k}{T_s}\right) \mathrm{e}^{2\pi \mathrm{j} \tau \frac{F-k}{T_s}}$ is depicted in Figure 6.11b. After summing the shifted spectral replicas, we obtain the final spectrum $C^{\tau}(F)$ which equals to 1 for $\tau = 0$ and is a complex function otherwise ($\tau \neq 0$) as shown in Figures 6.12 and 6.13. We notice that the magnitude of $C^{\tau}(F)$ falls below 1 only on intervals, for which there are multiple nonzero contributions in the sum (6.40). As a consequence, for larger values of β, the magnitude of $C^{\tau}(F)$ gets affected more by the misalignment ($\tau > 0$).

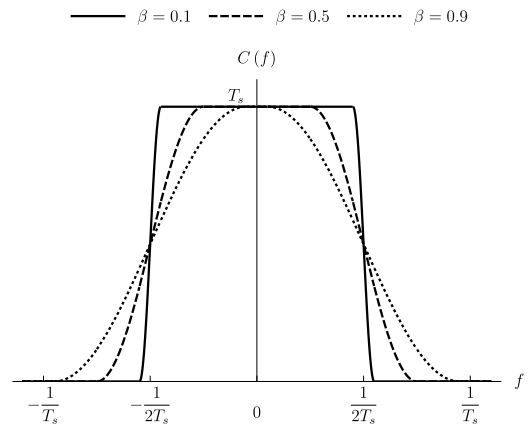
The reason why we are interested in the ACF spectrum is its direct link to the eigenvalues of the delay matrix. Without the approximation, that is when $m = \infty$, the eigenvalues can by obtained by uniform sampling of one period of $C^{\tau}(F)$. In the case where $m < \infty$, a convolution with the spectrum of a window needs to be applied first, according to (6.42).

In Figures 6.14a and 6.14b we show a parametric plot of $S^{\tau,(m)}(F)$ for various combinations of τ and β. In addition, we also include the corresponding eigenvalues $\lambda_i, 0 \le i < N = 11$. Observations:

- For τ close to 0, the curve tightens around 1, this is expected as for $\tau = 0$ $C^{\tau}(F)$ becomes equal to 1.
- There are more eigenvalues with small magnitude for bigger τ.

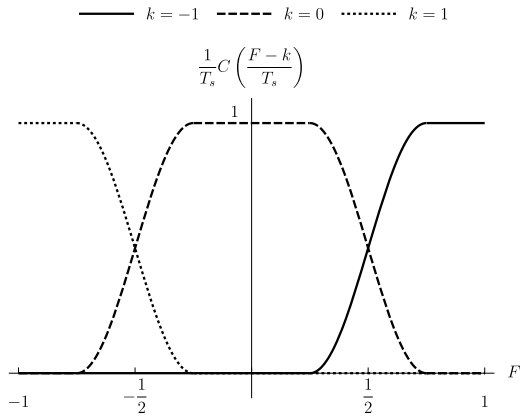


(a) RRC ACF (6.45) for various values of the roll-off factor β.

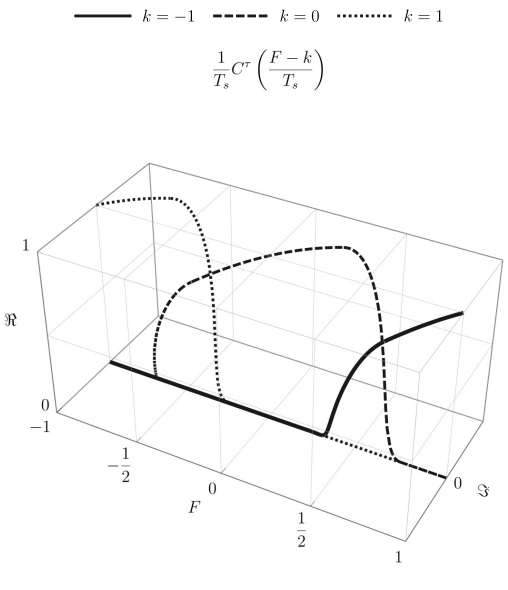


(b) Energy spectrum (6.46) for various values of the roll-off factor β.

Figure 6.10: Second-order characteristics of the RRC modulation pulse $g(t)$.



(a) Synchronous sampling ($\tau = 0$).



(b) Asynchronous sampling ($\tau \neq 0$).

Figure 6.11: Overlapping energy spectrum contributions according to (6.40) for $\beta = \frac{1}{2}$.

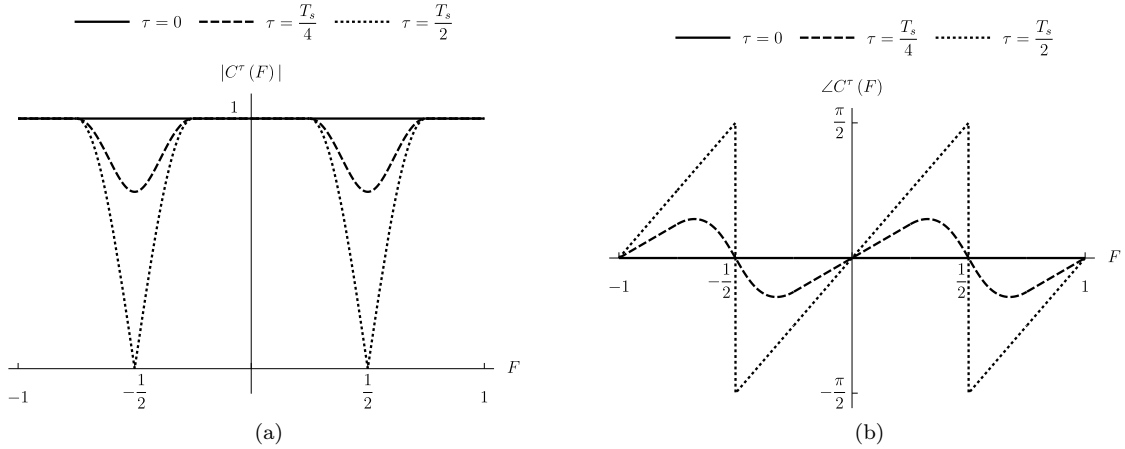


Figure 6.12: Energy spectrum of sampled $g(t)$ with $f_s = \frac{1}{T_s}$ for various values of the temporal offset τ and for $\beta = \frac{1}{2}$.

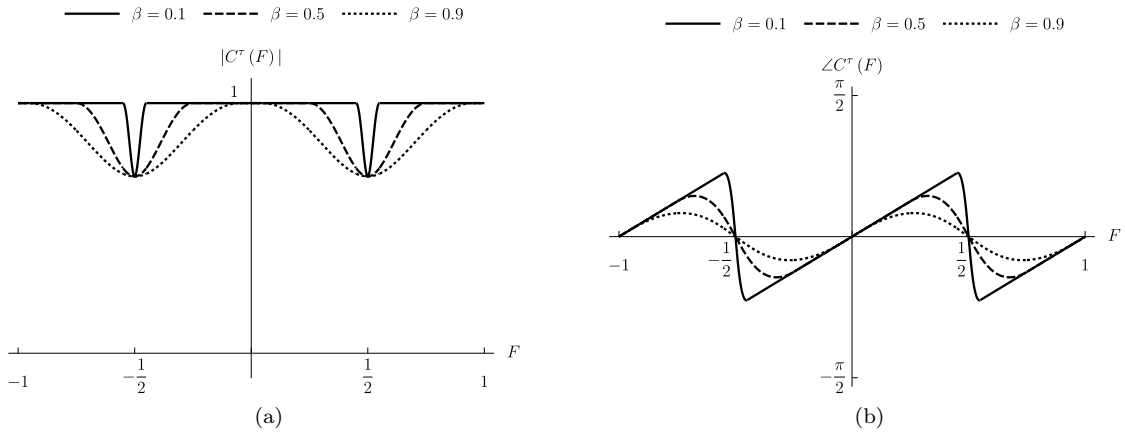
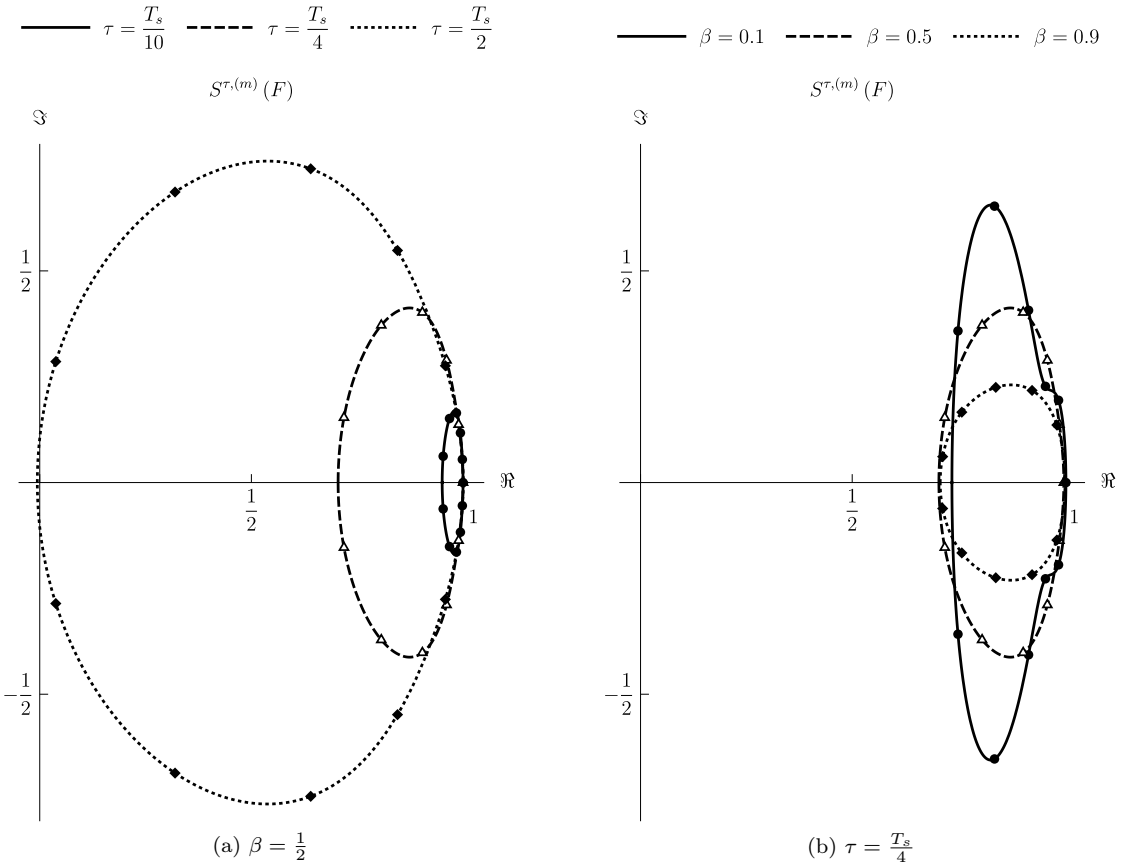


Figure 6.13: Energy spectrum of sampled $g(t)$ with $f_s = \frac{1}{T_s}$ for various values of the roll-off factor β and $\tau = \frac{1}{4}$.



(a) $\beta = \frac{1}{2}$ (b) $\tau = \frac{T_s}{4}$

Figure 6.14: A parametric plot of $S^{\tau,(m)}(F)$, where $F \in \left[-\frac{1}{2}, \frac{1}{2}\right]$ and $m = 3$. The markers indicate corresponding eigenvalues as uniform samples of one period of $S^{\tau,(m)}(F)$.

- For smaller values of β, the absolute values of the eigenvalues are less affected by τ.
- For small values of β, the approximation has a larger impact on$|\lambda_i|$.

6.4.4 Equivalent Channel

The essential observation is that the transformation matrix $\boldsymbol{U}$ is not affected by τ and as such, it is possible to apply the transformation at the transmitters without the knowledge of the delay. To make this more precise, we compute a transformed source signal at transmitter S_A as $\boldsymbol{s}'_A = \boldsymbol{U}\boldsymbol{s}_A$. Similarly for S_B and $\boldsymbol{s}'_B = \boldsymbol{U}\boldsymbol{s}_B$. Next, we substitute the transformed source signals together with (6.32) into (6.12) and multiply the equation with $\boldsymbol{U}^*$ from the left:

$$\begin{aligned}\boldsymbol{x}'(\epsilon) &\approx h_A \boldsymbol{U}^* \boldsymbol{G}(\epsilon\delta)\, \boldsymbol{s}'_A + h_B \boldsymbol{U}^* \boldsymbol{G}((\epsilon-1)\delta)\, \boldsymbol{s}'_B + \boldsymbol{U}^* \boldsymbol{w} \\ &= h_A \boldsymbol{U}^* \boldsymbol{U} \boldsymbol{\Sigma}(\epsilon\delta)\, \boldsymbol{U}^* \boldsymbol{U} \boldsymbol{s}_A + h_B \boldsymbol{U}^* \boldsymbol{U} \boldsymbol{\Sigma}((\epsilon-1)\delta)\, \boldsymbol{U}^* \boldsymbol{U} \boldsymbol{s}_B + \boldsymbol{U}^* \boldsymbol{w} \\ &= h_A \boldsymbol{\Sigma}(\epsilon\delta)\, \boldsymbol{s}_A + h_B \boldsymbol{\Sigma}((\epsilon-1)\delta)\, \boldsymbol{s}_B + \boldsymbol{w}',\end{aligned} \tag{6.47}$$

where $\boldsymbol{x}' = \boldsymbol{U}^* \boldsymbol{x}$ is a transformation of the received signal and $\boldsymbol{w}' = \boldsymbol{U}^* \boldsymbol{w}$ is transformed noise. First, we note that $\boldsymbol{w}'$ is still i.i.d. Gaussian with the same variance as $\boldsymbol{w}$ because we applied a unitary transformation. Second, as stated before, since $\boldsymbol{U}$ does not depend on τ, the source symbol vector transformation can be performed at each individual transmitter without the knowledge of

the channel state. Finally because $\boldsymbol{\Sigma}$ is diagonal, (6.47) can be viewed as a set of individual independent AWGN subchannels, with gains given by the eigenvalues of $\boldsymbol{G}(\epsilon\delta)$ and $\boldsymbol{G}((\epsilon-1)\delta)$. The nth subchannel follows as

$$x'_n(\epsilon) \approx \lambda_n(\epsilon\delta) h_A s(c_{A,n}) + \lambda_n((\epsilon-1)\delta) h_B s(c_{B,n}) + w'_n. \tag{6.48}$$

It is interesting to compare this approach with the case without precoding (6.8) from an energetic point of view. If we focus on an arbitrary transmitted symbol, its total energy available after the MF is given as $\sum_{n=-\infty}^{\infty} |R_g(nT_S-\tau)|^2$ which is spread over all the samples and mixed up with the rest of the symbols. Using Parseval's theorem we have

$$\sum_{n=-\infty}^{\infty} |R_g(nT_S-\tau)|^2 \approx \sum_{n=-m}^{m} |R_g(nT_S-\tau)|^2 = \frac{1}{N}\sum |\lambda_i|^2. \tag{6.49}$$

The LHS of (6.49) corresponds to the energy of one symbol passing through the asynchronous channel. The RHS represents the energy per symbol averaged over all the subchannels (6.59) in the case with precoding. Hence the use of precoding allows to effectively collect all the available energy spread over the whole frame due to misalignment and concentrate it into a set of independent channels with individual gains given by λ_n. If we compare this with (6.8), we conclude that this approach allows for the removal of the interference but is not able to capture all of the transmitted energy. In Figure 6.15 we show the dependence of the symbol energy on the delay τ and compare the case without precoding with the maximal, minimal, and mean over all subchannels in the coded case.

6.4.5 Impact of Channel Parametrization

We examine the impact of time-varying channel coefficients on the asynchronous channel reception with precoding. To support the matrix notation, we represent the channel coefficients in terms of a diagonal matrix as

$$\boldsymbol{H}_A = \begin{bmatrix} h_{A,1} & 0 & 0 & 0 \\ 0 & h_{A,2} & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & h_{A,N} \end{bmatrix}. \tag{6.50}$$

As such, the observation is

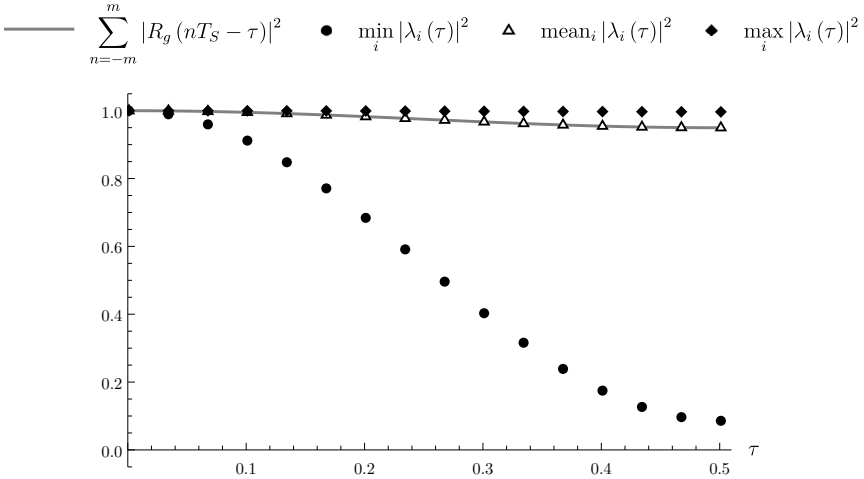
$$\begin{aligned} \boldsymbol{x}'(\epsilon) &\approx \boldsymbol{U}^*\boldsymbol{H}_A\boldsymbol{G}(\delta)\boldsymbol{s}'_A + \boldsymbol{U}^*\boldsymbol{H}_B\boldsymbol{G}((\epsilon-1)\delta)\boldsymbol{s}'_B + \boldsymbol{U}^*\boldsymbol{w} \\ &= \boldsymbol{U}^*\boldsymbol{H}_A\boldsymbol{U}\boldsymbol{\Sigma}(\epsilon\delta)\boldsymbol{s}_A + \boldsymbol{U}^*\boldsymbol{H}_B\boldsymbol{U}\boldsymbol{\Sigma}((\epsilon-1)\delta)\boldsymbol{s}_B + \boldsymbol{w}'. \end{aligned} \tag{6.51}$$

We see that the unitary transform does not chancel unless the channel matrices $\boldsymbol{H}_A$ and $\boldsymbol{H}_B$ commute with $\boldsymbol{U}^*$. Because $\boldsymbol{U}^*$ is not diagonal, the necessary and sufficient condition for $\boldsymbol{H}_A$ and $\boldsymbol{U}^*$ to commute is that $\boldsymbol{H}_A = \alpha\boldsymbol{I}$. In other words, when the channel coefficient does not change during the frame. If we assume the observation matrix $\boldsymbol{H}_A$ to be known at the receiver we can remove it and get

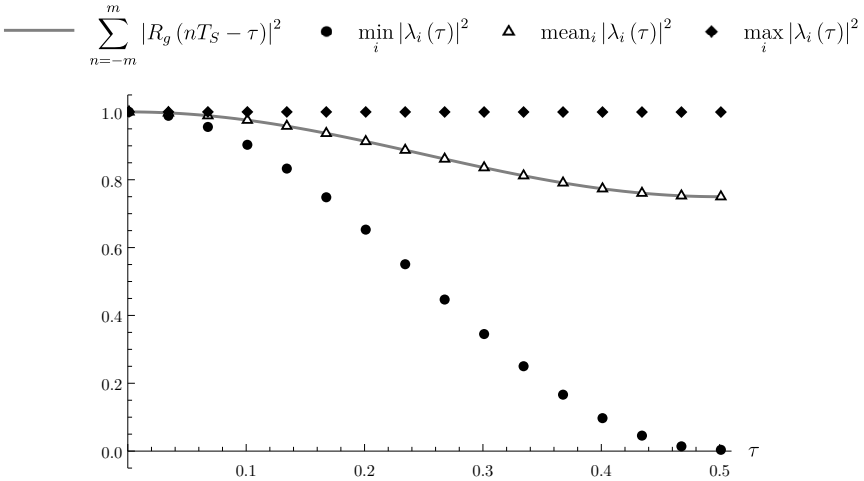
$$\boldsymbol{U}^*\boldsymbol{H}_A^{-1}\boldsymbol{x}(\epsilon) \approx \boldsymbol{\Sigma}(\epsilon)\boldsymbol{s}_A + \boldsymbol{U}^*\boldsymbol{H}\boldsymbol{U}\boldsymbol{\Sigma}((\epsilon-1)\delta)\boldsymbol{s}_B + \boldsymbol{U}^*\boldsymbol{H}_A^{-1}\boldsymbol{w}, \tag{6.52}$$

where

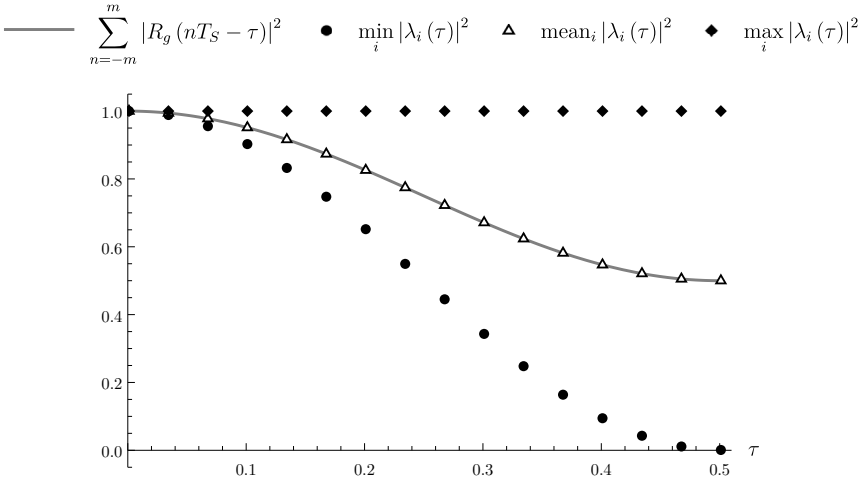
$$\boldsymbol{H} = \boldsymbol{H}_A^{-1}\boldsymbol{H}_B = \begin{bmatrix} \frac{h_{B,0}}{h_{A,0}} & 0 & 0 & 0 \\ 0 & \frac{h_{B,1}}{h_{A,1}} & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \frac{h_{B,(N-1)}}{h_{A,(N-1)}} \end{bmatrix} \tag{6.53}$$



(a) $\beta = \frac{1}{10}$



(b) $\beta = \frac{1}{2}$



(c) $\beta = 1$

Figure 6.15: Symbol energy comparison according to (6.49), where $N = 51$ and $m = 10$.

is the relative fading matrix. Because $\boldsymbol{H}$ is diagonal, $\boldsymbol{U}^*\boldsymbol{H}\boldsymbol{U} = \boldsymbol{C}$ is circular and

$$\boldsymbol{U}^*\boldsymbol{H}_A^{-1}\boldsymbol{x}\left(\epsilon\right) \approx \boldsymbol{\Sigma}\left(\epsilon\right)\boldsymbol{s}_A + \boldsymbol{C}\boldsymbol{\Sigma}\left(\left(\epsilon-1\right)\delta\right)\boldsymbol{s}_B + \boldsymbol{U}^*\boldsymbol{H}_A^{-1}\boldsymbol{w}. \tag{6.54}$$

We conclude that with known $\boldsymbol{H}_A$ and $\boldsymbol{H}_B$, the equivalent channel turns into a sum of $\boldsymbol{\Sigma}\left(\epsilon\right)\boldsymbol{s}_A$ and $\boldsymbol{\Sigma}\left(\left(\epsilon-1\right)\delta\right)\boldsymbol{s}_B$ filtered by a FIR with a known response. Based on the channel dynamics, a viable approach may be to split the frame into smaller sub-frames such that we can approximate $\boldsymbol{H}_{A,i} \approx h_{A,i}\boldsymbol{I}$ and $\boldsymbol{H}_{B,i} \approx h_{B,i}\boldsymbol{I}$, where i is the index of the sub-frame. The precoding would be performed on a sub-frame basis turning $\boldsymbol{G}$ and $\boldsymbol{U}$ into block diagonal matrices.

6.4.6 Midpoint Sampling (A)

In this section, we present a strategy of sampling once every symbol duration at an instant that is in between the two time-shifted signals, i.e. $\epsilon = \frac{1}{2}$. The assumption of an even ACF ($R_g\left(t\right) = R_g\left(-t\right)$) results in

$$\boldsymbol{G}\left(\frac{-\delta}{2}\right) = \boldsymbol{G}^T\left(\frac{\delta}{2}\right). \tag{6.55}$$

When using (6.55) and following the steps in (6.47) we obtain

$$\begin{aligned}\boldsymbol{x}'\left(\frac{1}{2}\right) &\approx h_A\boldsymbol{U}^*\boldsymbol{G}\left(\frac{\delta}{2}\right)\boldsymbol{s}'_A + h_B\boldsymbol{U}^*\boldsymbol{G}^{\mathrm{T}}\left(\frac{\delta}{2}\right)\boldsymbol{s}'_B + \boldsymbol{U}^*\boldsymbol{w}\\ &= h_A\boldsymbol{U}^*\boldsymbol{U}\boldsymbol{\Sigma}\left(\frac{\delta}{2}\right)\boldsymbol{U}^*\boldsymbol{U}\boldsymbol{s}_A + h_B\boldsymbol{U}^*\boldsymbol{U}^*\boldsymbol{\Sigma}^{\mathrm{T}}\left(\frac{\delta}{2}\right)\boldsymbol{U}\boldsymbol{U}\boldsymbol{s}_B + \boldsymbol{w}'\\ &= h_A\boldsymbol{\Sigma}\left(\frac{\delta}{2}\right)\boldsymbol{s}_A + h_B\boldsymbol{U}^*\boldsymbol{U}^*\boldsymbol{\Sigma}^{\mathrm{T}}\left(\frac{\delta}{2}\right)\boldsymbol{U}\boldsymbol{U}\boldsymbol{s}_B + \boldsymbol{w}'.\end{aligned} \tag{6.56}$$

It can be shown that $\boldsymbol{U}^2$ is a permutation matrix with a 1 in the upper left corner and on the -1th anti-diagonal as shown in Example 6.11.

Example 6.11. Example of U^2 for $N = 5$.

$$\boldsymbol{U}^2 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 1\\ 0 & 0 & 0 & 1 & 0\\ 0 & 0 & 1 & 0 & 0\\ 0 & 1 & 0 & 0 & 0 \end{bmatrix} \tag{6.57}$$

Because $\boldsymbol{U}^2$ is a real matrix, in turn, it holds $\boldsymbol{U}^*\boldsymbol{U}^* = \left(\boldsymbol{U}\boldsymbol{U}\right)^* = \boldsymbol{U}^2$. Further, it can be shown that the multiplication of a diagonal matrix by $\boldsymbol{U}^2$ from both sides has the effect of reversing the elements on the main diagonal besides the first element. When considering this, together with the symmetry of the eigenvalues (6.34) we conclude that $\boldsymbol{U}^2\boldsymbol{\Sigma}\boldsymbol{U}^2 = \boldsymbol{\Sigma}^*$ (Example 6.12) and thus

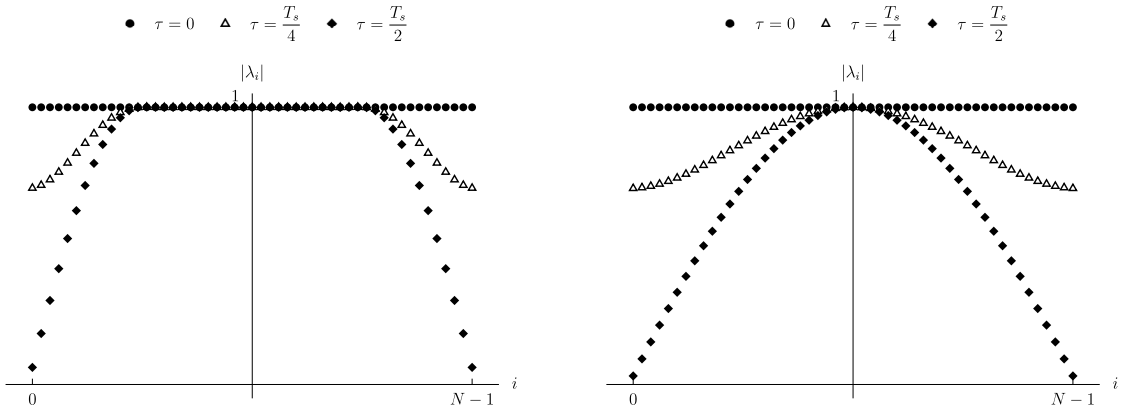
$$\boldsymbol{x}' \approx h_A\boldsymbol{\Sigma}\boldsymbol{s}_A + h_B\boldsymbol{\Sigma}^*\boldsymbol{s}_B + \boldsymbol{w}', \tag{6.58}$$

where we dropped the explicit dependence on δ. The resulting independent AWGN subchannels follow as

$$x'_n \approx \lambda_n h_A s\left(c_{A,n}\right) + \lambda_n^* h_B s\left(c_{B,n}\right) + w'_n. \tag{6.59}$$

The corresponding ML demodulator is given simply as

$$\hat{c}_n = \mathrm{argmax}_{\check{c}} \sum_{\chi_c(c_A,c_B)=\check{c}} \exp\left(\frac{-1}{\sigma_w^2}\left|x'_n - \lambda_n h_A s\left(c_A\right) - \lambda_n^* h_B s\left(c_B\right)\right|^2\right). \tag{6.60}$$



(a) $\beta = \frac{1}{2}$, $m = 10$ (b) $\beta = 1$, $m = 10$

Figure 6.16: The eigenvalues magnitudes resulting from a frame of length $N = 51$ and RRC modulation pulse.

Example 6.12. Here we show that indeed $\boldsymbol{U}^2\boldsymbol{\Sigma}\boldsymbol{U}^2 = \boldsymbol{\Sigma}^*$, in an example for $N = 4$.

$$\boldsymbol{U}^2\boldsymbol{\Sigma} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \lambda_0 & 0 & 0 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 & 0 \\ 0 & 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & 0 & \lambda_3 & 0 \\ 0 & 0 & 0 & 0 & \lambda_4 \end{bmatrix} = \begin{bmatrix} \lambda_0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \lambda_1 \\ 0 & 0 & 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 & 0 & 0 \\ 0 & \lambda_4 & 0 & 0 & 0 \end{bmatrix} \tag{6.61}$$

$$\begin{aligned} \boldsymbol{U}^2\boldsymbol{\Sigma}\boldsymbol{U}^2 &= \begin{bmatrix} \lambda_0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \lambda_1 \\ 0 & 0 & 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 & 0 & 0 \\ 0 & \lambda_4 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} \lambda_0 & 0 & 0 & 0 & 0 \\ 0 & \lambda_4 & 0 & 0 & 0 \\ 0 & 0 & \lambda_3 & 0 & 0 \\ 0 & 0 & 0 & \lambda_2 & 0 \\ 0 & 0 & 0 & 0 & \lambda_1 \end{bmatrix} = \begin{bmatrix} \lambda_0 & 0 & 0 & 0 & 0 \\ 0 & \lambda_1^* & 0 & 0 & 0 \\ 0 & 0 & \lambda_2^* & 0 & 0 \\ 0 & 0 & 0 & \lambda_3^* & 0 \\ 0 & 0 & 0 & 0 & \lambda_4^* \end{bmatrix} \end{aligned} \tag{6.62}$$

From (6.59) it follows that the nth subchannel experiences an SNR of

$$\mathrm{SNR}_n = \frac{|\lambda_n|^2}{\sigma_w^2} \tag{6.63}$$

and a relative parametrization

$$\frac{\lambda_n^* h_B}{\lambda_n h_A} = \frac{|\lambda_n|\,|h_B|\,\mathrm{e}^{\mathrm{j}(\angle h_B - \angle \lambda_n)}}{|\lambda_n|\,|h_A|\,\mathrm{e}^{\mathrm{j}(\angle h_A \angle \lambda_n)}} = \mathrm{e}^{-2\mathrm{j}\angle\lambda_n} h. \tag{6.64}$$

In Figure 6.16 we see an example distribution of $|\lambda_n|$ for a RRC pulse and various values of β and τ. It is not surprising to see that for a bigger offset τ, there is a greater number of weaker subchannels. Also, for a bigger roll-off factor the eigenvalue magnitudes are affected more by the asynchronicity τ.

6.4.7 Double Rate Sampling (B)

In the previous section, we considered mid-point sampling at the receiver and showed that the application of DFT precoding allows to resolve all the interference arising from asynchronous reception. In addition, we observed that even after the interference is removed, some part of the performance degradation remains present. This is the case because with one sample per T_s, the MF is not able to collect the whole energy of the modulation pulse $g(t)$ whenever $\tau \neq 0$ or $\beta \neq 0$ and the resulting samples do not represent a sufficient statistic.

In this section, we consider taking two samples per T_s at the receiver. In particular, we take one sample synchronously with each source ($\epsilon_1 = 0, \epsilon_2 = 1$) resulting in $\boldsymbol{x}_1 = \boldsymbol{x}(0)$ and $\boldsymbol{x}_2 = \boldsymbol{x}(1)$ respectively. The observations are given as

$$\begin{aligned} \boldsymbol{x}_1 &= h_A \boldsymbol{G}(0)\, \boldsymbol{s}'_A + h_B \boldsymbol{G}(-\delta)\, \boldsymbol{s}'_B + \boldsymbol{w}_1 \\ \boldsymbol{x}_2 &= h_A \boldsymbol{G}(\delta)\, \boldsymbol{s}'_A + h_B \boldsymbol{G}(0)\, \boldsymbol{s}'_B + \boldsymbol{w}_2, \end{aligned} \tag{6.65}$$

with $\boldsymbol{w}_1$ and $\boldsymbol{w}_2$ correlated. The (k,l)th element of the cross-covariance matrix $\boldsymbol{K}_{\boldsymbol{w}_1,\boldsymbol{w}_2} = \mathrm{E}\left[\boldsymbol{w}_1 \boldsymbol{w}_2^{\mathrm{H}}\right]$ can be written down as

$$\begin{aligned} \sigma_{k,l} &= \mathrm{E}\left[\int_{-\infty}^{\infty} n(t)\, g^*(t - kT_S)\, \mathrm{d}t \int_{\infty}^{\infty} n^*(t')\, g(t' - lT_S - \delta)\, \mathrm{d}t'\right] \\ &= \mathrm{E}\left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} n(t)\, g^*(t - kT_S)\, n(t')\, g(t' - lT_S - \delta)\, \mathrm{d}t'\mathrm{d}t\right] \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g^*(t - kT_S)\, \mathrm{E}\left[n(t)\, n^*(t')\right] g(t' - lT_S - \delta)\, \mathrm{d}t'\mathrm{d}t. \end{aligned} \tag{6.66}$$

Assuming the noise process $n(t)$ to be uncorrelated, i.e. $K_n(t,t') = \mathrm{E}\left[n^*(t)\, n(t')\right] = \delta(t - t')\, \sigma_w^2$ we get

$$\begin{aligned} \sigma_{k,l} &= \sigma_w^2 \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g^*(t - kT_S)\, \delta(t - t')\, g(t' - lT_S - \delta)\, \mathrm{d}t'\mathrm{d}t \\ &= \sigma_w^2 \int_{-\infty}^{\infty} g^*(t - kT_S)\, g(t - lT_S - \delta)\, \mathrm{d}t \\ &= \sigma_w^2 R_g((k - l)\, T_S - \delta) \\ &= \sigma_w^2 \boldsymbol{G}'(\delta)_{l,k} \end{aligned} \tag{6.67}$$

and

$$\boldsymbol{K}_{\boldsymbol{w}_1,\boldsymbol{w}_2} = \sigma_w^2 \boldsymbol{G}'(\delta)^{\mathrm{T}} = \sigma_w^2 \boldsymbol{G}'(-\delta) \tag{6.68}$$

for an even ACF ($R_g(t) = R_g(-t)$).

For the (k,l)th element of the pseudo cross-covariance matrix $\boldsymbol{M}_{\boldsymbol{w}_1,\boldsymbol{w}_2} = \mathrm{E}\left[\boldsymbol{w}_1 \boldsymbol{w}_2^{\mathrm{T}}\right]$, we can write

$$m_{k,l} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(t - kT_S)\, \mathrm{E}\left[n(t)\, n(t')\right] g(t' - lT_S - \epsilon_2)\, \mathrm{d}t'\mathrm{d}t \tag{6.69}$$

and from the assumption of a circularly symmetric noise process, i.e. $\mathrm{E}\left[n(t)\, n(t')\right] = 0$ it follows that $m_{k,l} = 0$ and in turn

$$\boldsymbol{M}_{\boldsymbol{w}_1,\boldsymbol{w}_2} = \boldsymbol{0}. \tag{6.70}$$

After the inverse DFT, the observation (6.65) transforms into

$$\begin{aligned} \boldsymbol{x}'_1 &\approx h_A \boldsymbol{U}^* \boldsymbol{G}(0)\, \boldsymbol{s}'_A + h_B \boldsymbol{U}^* \boldsymbol{G}(-\delta)\, \boldsymbol{s}'_B + \boldsymbol{U}^* \boldsymbol{w}_1 \\ &= h_A \boldsymbol{U}^* \boldsymbol{U} \boldsymbol{s}_A + h_B \boldsymbol{U}^2 \boldsymbol{\Sigma}(\delta)\, \boldsymbol{U}^2 \boldsymbol{s}_B + \boldsymbol{w}'_1 \\ &= h_A \boldsymbol{s}_A + h_B \boldsymbol{\Sigma}^*(\delta)\, \boldsymbol{s}_B + \boldsymbol{w}'_1 \\ \boldsymbol{x}'_2 &\approx h_A \boldsymbol{U}^* \boldsymbol{G}(\delta)\, \boldsymbol{s}'_A + h_B \boldsymbol{U}^* \boldsymbol{G}(0)\, \boldsymbol{s}'_B + \boldsymbol{U}^* \boldsymbol{w}_2 \\ &= h_A \boldsymbol{\Sigma}(\delta)\, \boldsymbol{s}_A + h_B \boldsymbol{s}_B + \boldsymbol{w}'_2, \end{aligned} \tag{6.71}$$

where $\boldsymbol{G}(0) = \boldsymbol{I}$ follows from (6.13). First, we need to examine the stochastic properties of the noise vectors $\boldsymbol{w}'_1 = \boldsymbol{U}^* \boldsymbol{w}_1$ and $\boldsymbol{w}'_2 = \boldsymbol{U}^* \boldsymbol{w}_2$. Knowing that $\boldsymbol{w}_1$ and $\boldsymbol{w}_2$ are jointly Gaussian and circularly symmetric, the cross-covariance matrix of $\boldsymbol{w}'_1$ and $\boldsymbol{w}'_2$ can be written as

$$\begin{aligned} K_{\boldsymbol{w}'_1, \boldsymbol{w}'_2} &= \mathrm{E}\left[\boldsymbol{U}^* \boldsymbol{w}_1 \boldsymbol{w}_2^{\mathrm{H}} \boldsymbol{U}\right] \\ &= \sigma_w^2 \boldsymbol{U}^* \boldsymbol{G}'(-\delta)\, \boldsymbol{U} \\ &= \sigma_w^2 \boldsymbol{U}^* \left(\boldsymbol{G}(-\delta) + \boldsymbol{\Delta}(-\delta)\right) \boldsymbol{U} \\ &= \sigma_w^2 \left(\boldsymbol{\Sigma}(-\delta) + \boldsymbol{U}^* \boldsymbol{\Delta}(-\delta)\, \boldsymbol{U}\right), \end{aligned} \tag{6.72}$$

where $\boldsymbol{\Delta} = \boldsymbol{G}' - \boldsymbol{G}$. Matrix $\boldsymbol{\Delta}$ cannot be removed by a cyclic extension as we did with the useful signal. Based on numerical evaluation however, the term $\boldsymbol{U}^* \boldsymbol{\Delta}(-\delta)\, \boldsymbol{U}$ is small w.r.t. $\boldsymbol{\Sigma}(-\delta)$ and as such, we can approximate $\boldsymbol{K}_{\boldsymbol{w}'_1, \boldsymbol{w}'_2}$ by the diagonal matrix $\sigma_w^2 \boldsymbol{\Sigma}(-\delta) = \sigma_w^2 \boldsymbol{\Sigma}^*(\delta)$.

Remark. The equality $\boldsymbol{\Sigma}(-\tau) = \boldsymbol{\Sigma}^*(\tau)$ can be shown as follows

$$\boldsymbol{\Sigma}(-\tau) = \boldsymbol{U}^* \boldsymbol{G}(-\tau)\, \boldsymbol{U} = \boldsymbol{U}^* \boldsymbol{G}^{\mathrm{T}}(\tau)\, \boldsymbol{U} = \left(\boldsymbol{U} \boldsymbol{G}(\tau)\, \boldsymbol{U}^*\right)^{\mathrm{T}} = \left(\boldsymbol{U}^* \boldsymbol{G}(\tau)\, \boldsymbol{U}\right)^{\mathrm{H}} = \left(\boldsymbol{\Sigma}(\tau)\right)^{\mathrm{H}} = \boldsymbol{\Sigma}^*(\tau). \tag{6.73}$$

We observe that when assuming $\boldsymbol{K}_{\boldsymbol{w}'_1, \boldsymbol{w}'_2}$ diagonal, the equivalent channel (6.71) falls apart into a set of independent two-dimensional subchannels given by

$$\begin{bmatrix} x'_{1,n} \\ x'_{2,n} \end{bmatrix} = \boldsymbol{S}_n \begin{bmatrix} h_A s(c_{A,n}) \\ h_B s(c_{B,n}) \end{bmatrix} + \begin{bmatrix} w'_{1,n} \\ w'_{2,n} \end{bmatrix} \tag{6.74}$$

with

$$\boldsymbol{S}_n = \begin{bmatrix} 1 & \lambda_n^* \\ \lambda_n & 1 \end{bmatrix} \tag{6.75}$$

$$\begin{bmatrix} w'_{1,n} \\ w'_{2,n} \end{bmatrix} \sim CN\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \sigma_w^2 \boldsymbol{S}_n\right). \tag{6.76}$$

In turn it is possible to apply a two-dimensional ML demodulator according to

$$\hat{c}_n = \mathrm{argmax}_{\check{c}} \sum_{\chi_c(c_A, c_B) = \check{c}} \mathrm{e}^{-\boldsymbol{q}_n(c_A, c_B)^{\mathrm{H}} \boldsymbol{K}^{-1} \boldsymbol{q}_n(c_A, c_B)} \tag{6.77}$$

$$\boldsymbol{q}_n = \begin{bmatrix} x'_{1,n} - u_A - \lambda_n^* u_B \\ x'_{2,n} - \lambda_n u_A - u_B \end{bmatrix}, \tag{6.78}$$

$$\boldsymbol{K}^{-1} = \frac{1}{\sigma_w^2 \left(1 - |\lambda|^2\right)} \begin{bmatrix} 1 & -\lambda_n^* \\ -\lambda_n & 1 \end{bmatrix}, \tag{6.79}$$

where we $u_A = h_A s(c_A)$ and $u_B = h_B s(c_B)$. After substituting for $\boldsymbol{q}$ and $\boldsymbol{K}^{-1}$ into (6.77) and assuming a PSK alphabet, we get

$$\sigma_w^2 \boldsymbol{q}_n(c_A, c_B)^{\mathrm{H}} \boldsymbol{K}^{-1} \boldsymbol{q}_n(c_A, c_B) = \frac{\left|x'_{1,n}\right|^2 + \left|x'_{2,n}\right|^2 - 2\Re\left\{\lambda_n x'_{1,n} \left(x'_{2,n}\right)^*\right\}}{\left(1 - |\lambda_n|^2\right)} + |h_A|^2 + |h_B|^2 + 2\Re\left\{\lambda_n u_A u_B^* - x'_{1,n} u_A^* - x'_{2,n} u_B^*\right\}, \tag{6.80}$$

for $|\lambda_n| \neq 1$. After removing terms which do not depend on c_A or c_B, the final ML demodulator results in

$$\hat{c}_n = \mathrm{argmax}_{\check{c}} \sum_{\chi_c(c_A, c_B) = \check{c}} \mathrm{e}^{-\frac{2}{\sigma_w^2}\Re\left\{\lambda_n u_A u_B^* - x'_{1,n} u_A^* - x'_{2,n} u_B^*\right\}} \tag{6.81}$$

which is defined even for the case when $|\lambda_n| = 1$.

Remark. It is interesting to note what happens for the synchronous case when $\delta = 0$:

$$\boldsymbol{G}(0) = \boldsymbol{I} \tag{6.82}$$

$$\boldsymbol{\Sigma}(0) = \boldsymbol{U}^* I \boldsymbol{U} = \boldsymbol{I} \tag{6.83}$$

$$\lambda_n = 1, \forall n. \tag{6.84}$$

From this follows that $x_1 = x_2$ and (6.81) changes to

$$\hat{c}_n = \mathrm{argmax}_{\check{c}} \sum_{(c_A, c_B):\ \check{c}} \mathrm{e}^{-\frac{2}{\sigma_w^2}\Re\left\{u_A u_B^* - x'_{1,n}(u_A + u_B)^*\right\}} \tag{6.85}$$

which is the standard hierarchical ML for the synchronous case. We can see that the double rate sampling with precoding naturally coincides with the standard situation in the case when $\delta = 0$.

As $u_A = h_A s(c_A)$ and $u_B = h_B s(c_B)$, in (6.81) it can be observed that $\bar{x}_{1,n} = \Re\left\{x'_{1,n} h_A^*\right\}$ and $\bar{x}_{2,n} = \Re\left\{x'_{2,n} h_B^*\right\}$ are sufficient statistics for the observations $x'_{1,n}$ and $x'_{2,n}$ respectively. If we substitute the observation model (6.74) for $x'_{1,n}$ and $x'_{2,n}$, we get

$$\begin{aligned} \bar{x}_{1,n} &= \Re\left\{\left(h_A s(c_{A,n}) + \lambda_n h_B s(c_{B,n}) + w'_{1,n}\right) h_A^*\right\} \\ &= \Re\left\{|h_A|^2 s(c_{A,n}) + \lambda_n h_B h_A^* s(c_{B,n})\right\} + \Re\left\{h_A^* w'_{1,n}\right\} \\ &= \bar{u}_{1,n} + \bar{w}_{1,n} \end{aligned} \tag{6.86}$$

$$\begin{aligned} \bar{x}_{2,n} &= \Re\left\{\left(\lambda_n^* h_A s(c_{A,n}) + h_B s(c_{B,n}) + w'_{2,n}\right) h_B^*\right\} \\ &= \Re\left\{\lambda_n^* h_A h_B^* s(c_{A,n}) + |h_B|^2 s(c_{B,n})\right\} + \Re\left\{h_B^* w'_{2,n}\right\} \\ &= \bar{u}_{2,n} + \bar{w}_{2,n}. \end{aligned} \tag{6.87}$$

The noise cross-covariance matrix gets transformed to

$$\boldsymbol{K}\begin{bmatrix} h_A^* w'_{1,n} \\ h_B^* w'_{2,n} \end{bmatrix} = \sigma_w^2 \begin{bmatrix} h_A^* & 0 \\ 0 & h_B^* \end{bmatrix} \boldsymbol{S}_n \begin{bmatrix} h_A & 0 \\ 0 & h_B \end{bmatrix} = \sigma_w^2 \begin{bmatrix} |h_A|^2 & h_A^* h_B \lambda_n \\ h_A h_B^* \lambda_n^* & |h_B|^2 \end{bmatrix} \tag{6.88}$$

and because $\left[h_A^* w'_{1,n}, h_B^* w'_{2,n}\right]^{\mathrm{T}}$ is circularly symmetric it follows that

$$\boldsymbol{K}\begin{bmatrix} \bar{w}_{1,n} \\ \bar{w}_{2,n} \end{bmatrix} = \frac{1}{2}\Re\left\{\boldsymbol{K}\begin{bmatrix} h_A^* w'_{1,n} \\ h_B^* w'_{2,n} \end{bmatrix}\right\} = \frac{\sigma_w^2}{2}\begin{bmatrix} |h_A|^2 & \Re\left\{h_A^* h_B \lambda_n\right\} \\ \Re\left\{h_A h_B^* \lambda_n^*\right\} & |h_B|^2 \end{bmatrix}. \tag{6.89}$$

Now, λ_n only appears multiplied by h_B, and as such $\angle\lambda_n$ is effectively added to the relative phase $\angle h_B - \angle h_A$. The magnitude $|\lambda_n|$ strengthens the dependence between the two observations (in the useful signal part as well as in the noise part). Again it holds that for $\lambda_n = 1$ both observations coincide and for $\lambda_n = 0$ we get fully independent observations of individual sources.

BPSK Alphabet and XOR HNC Map

Here we focus on (6.77) for the special case of the BPSK alphabet and XOR HNC map. We express the likelihoods given the two possible hierarchical target symbol values as

$$\begin{aligned}\mathrm{p}\left(\bar{x}_{1,n}, \bar{x}_{2,n} | c_n = 0\right) &\propto \exp\left(\frac{-1}{\sigma_w^2}\Re\left\{\lambda_n^* h_A h_B^*\right\} - \bar{x}_{1,n} - \bar{x}_{2,n}\right) \\ &\quad + \exp\left(\frac{-1}{\sigma_w^2}\Re\left\{\lambda_n^* h_A h_B^*\right\} + \bar{x}_{1,n} + \bar{x}_{2,n}\right) \\ &= 2\mathrm{e}^{-\frac{1}{\sigma_w^2}\Re\{\lambda_n^* h_A h_B^*\}} \cosh\left(\frac{1}{\sigma_w^2}\left(\bar{x}_{1,n} + \bar{x}_{2,n}\right)\right)\end{aligned} \tag{6.90}$$

$$\begin{aligned}\mathrm{p}\left(\bar{x}_{1,n}, \bar{x}_{2,n} | c_n = 1\right) &\propto \exp\left(\frac{-1}{\sigma_w^2}\Re\left\{\lambda_n^* h_A h_B^*\right\} - \bar{x}_{1,n} + \bar{x}_{2,n}\right) \\ &\quad + \exp\left(\frac{-1}{\sigma_w^2}\Re\left\{\lambda_n^* h_A h_B^*\right\} + \bar{x}_{1,n} - \bar{x}_{2,n}\right) \\ &= 2\mathrm{e}^{-\frac{1}{\sigma_w^2}\Re\{\lambda_n^* h_A h_B^*\}} \cosh\left(\frac{1}{\sigma_w^2}\left(\bar{x}_{1,n} - \bar{x}_{2,n}\right)\right).\end{aligned} \tag{6.91}$$

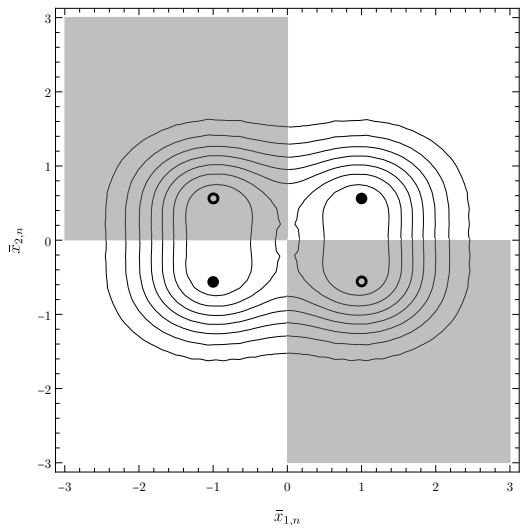
Then we can form the log-likelihood ratio (LLR) as

$$\Lambda\left(c_n\right) = \frac{\mathcal{L}\left(\bar{x}_{1,n}, \bar{x}_{2,n} | c_n = 1\right)}{\mathcal{L}\left(\bar{x}_{1,n}, \bar{x}_{2,n} | c_n = 0\right)} = \frac{\log\left(\cosh\left(\frac{1}{\sigma_w^2}\left(\bar{x}_{1,n} - \bar{x}_{2,n}\right)\right)\right) + \frac{1}{\sigma_w^2}\Re\left\{\lambda_n^* h_A h_B^*\right\}}{\log\left(\cosh\left(\frac{1}{\sigma_w^2}\left(\bar{x}_{1,n} + \bar{x}_{2,n}\right)\right)\right) - \frac{1}{\sigma_w^2}\Re\left\{\lambda_n^* h_A h_B^*\right\}}. \tag{6.92}$$

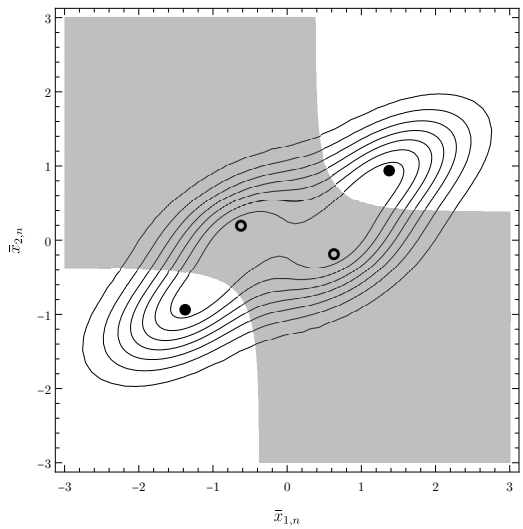
In Figures 6.17, 6.18, 6.19 we present the resulting decision regions in the "sufficient statistics space". That is the two axes represents the possible values of $\bar{x}_{1,n} = \Re\left\{x'_{1,n} h_A^*\right\}$ and $\bar{x}_{2,n} = \Re\left\{x'_{2,n} h_B^*\right\}$ respectively. For convenience, we include the useful signal points $[\bar{u}_{1,n}, \bar{u}_{2,n}]$ designated by two different marker styles according to the corresponding hierarchical target symbol c_n. In addition, we represent the additive noise $[\bar{w}_{1,n}, \bar{w}_{2,n}]$ distribution by means of several contour lines. We make the following observations:

- For $\lambda = 0$ the two observations $\bar{x}_{1,n}$ and $\bar{x}_{2,n}$ are independent because $\bar{u}_{1,n}$ and $\bar{u}_{2,n}$ are function of only $s\left(c_{A,n}\right)$ and $s\left(c_{A,n}\right)$ respectively and the noise samples $\bar{w}_{1,n}$ and $\bar{w}_{2,n}$ are independent. Corresponding Figures are 6.17a, 6.18a, and 6.19a.
- The same effect can be observed when $\angle h_A$ and $\angle h_B + \angle\lambda$ differ by a an odd multiple of π, in which case $\Re\left\{\lambda_n^* h_A h_B^*\right\} = 0$. Corresponding Figures are 6.19a, 6.19b and 6.19c.
- If $\lambda = 1$ this case coincides with a standard synchronous situation and the dependence of $\bar{x}_{1,n}$ and $\bar{x}_{2,n}$ is fully specified by the channel parameters h_A, h_B. Corresponding Figures are 6.17c, 6.18c and 6.19c.

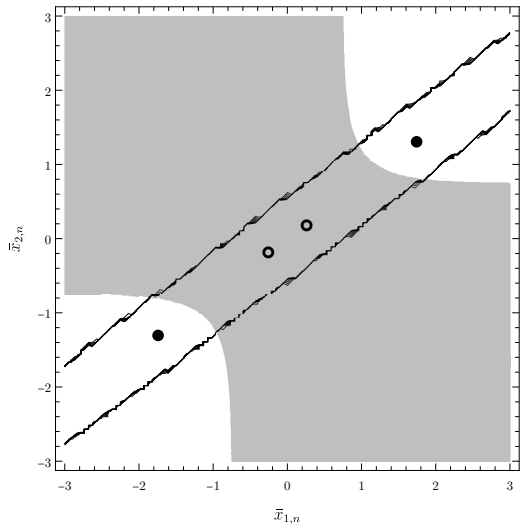
Remark. In Figures 6.17, 6.18, and 6.19 we used real-valued λ only because the phase $\angle\lambda$ directly contributes to the relative channel phase $\angle h_B - \angle h_A$. As such it is equivalent to a varying $\angle h_B$ while $\angle h_A = 0$.



(a) $\lambda = 0$

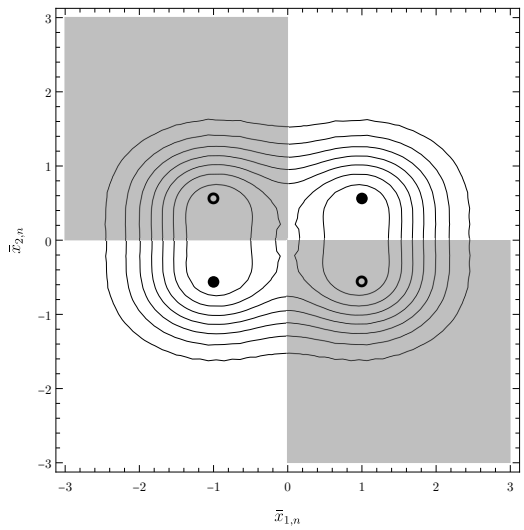


(b) $\lambda = \frac{1}{2}$

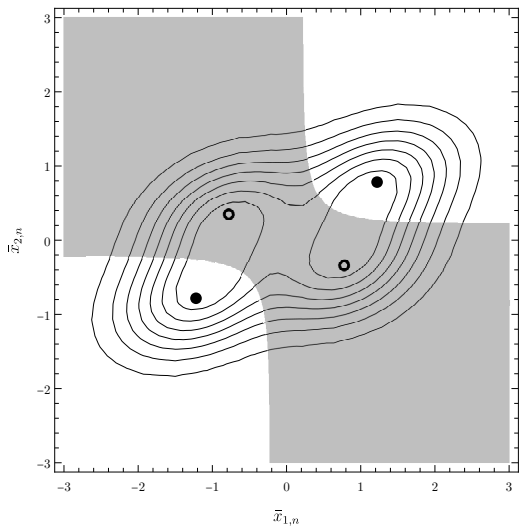


(c) $\lambda = 1$

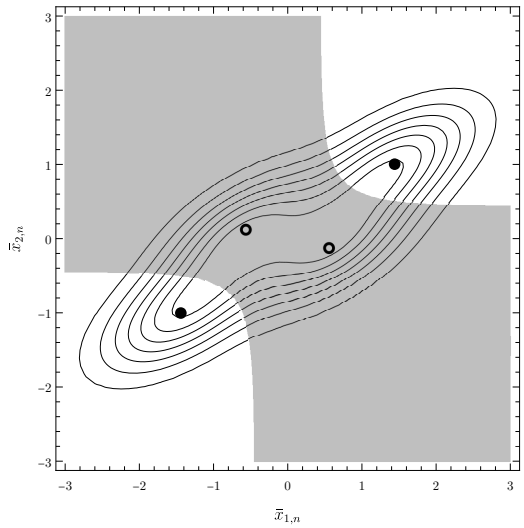
Figure 6.17: Illustration of the useful signal $[\bar{u}_{1,n}, \bar{u}_{2,n}]$ and the noise PDF in the sufficient statistics space. Decision regions are given by (6.92). Chanel parameters are given as $h_A = 1$ and $h_B = 0.75$.



(a) $\lambda = 0$

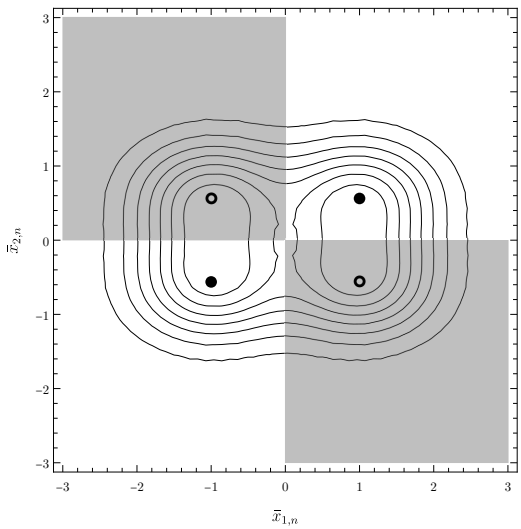


(b) $\lambda = \frac{1}{2}$

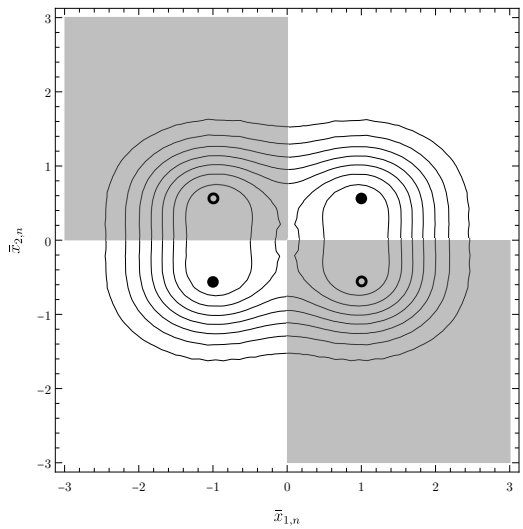


(c) $\lambda = 1$

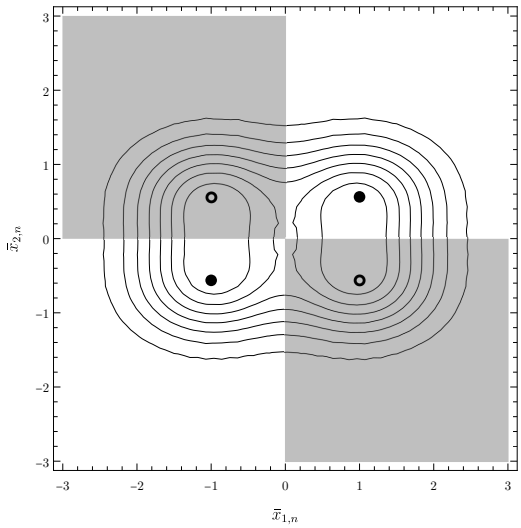
Figure 6.18: Illustration of the useful signal $\bar{u}_{1,n}, \bar{u}_{2,n}$ and the noise PDF in the sufficient statistics space. Decision regions are given by (6.92). Chanel parameters are given as $h_A = 1$ and $h_B = \frac{3}{4}\angle\frac{\pi}{3}$.



(a) $\lambda = 0$



(b) $\lambda = \frac{1}{2}$



(c) $\lambda = 1$

Figure 6.19: Illustration of the useful signal $\bar{u}_{1,n}, \bar{u}_{2,n}$ and the noise PDF in the sufficient statistics space. Decision regions are given by (6.92). Chanel parameters are given as $h_A = 1$ and $h_B = \frac{3}{4}\angle\frac{\pi}{2}$.

Two-Dimensional Sub-Channels

Let us focus on the two-dimensional subchannel given by (6.74) and evaluate its properties. For the channel matrix $\boldsymbol{S}_n$ we have

$$\det\left(\boldsymbol{S}_n\right) = 1 - \left|\lambda_n\right|^2 \tag{6.93}$$

which is nonzero as long as $|\lambda_n| \neq 1$. With this assumption, $\boldsymbol{S}_n$ is invertible and we can apply an inverse transform according to

$$\begin{aligned} \boldsymbol{S}_n^{-1} \begin{bmatrix} x'_{1,n} \\ x'_{2,n} \end{bmatrix} &= \begin{bmatrix} h_A s_A \\ h_B s_B \end{bmatrix} + \boldsymbol{S}_n^{-1} \begin{bmatrix} w'_{1,n} \\ w'_{2,n} \end{bmatrix} \\ \boldsymbol{S}_n^{-1} \begin{bmatrix} w'_{1,n} \\ w'_{2,n} \end{bmatrix} &\sim CN\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \boldsymbol{S}_n^{-1} \sigma_w^2 \boldsymbol{S}_n \left(\boldsymbol{S}_n^{-1}\right)^{\mathrm{H}} = \sigma_w^2 \boldsymbol{S}_n^{-1} \right), \end{aligned} \tag{6.94}$$

where

$$\boldsymbol{S}_n^{-1} = \frac{1}{1 - \left|\lambda_n\right|^2} \begin{bmatrix} 1 & -\lambda_n \\ -\lambda_n^* & 1 \end{bmatrix}. \tag{6.95}$$

Every sub-channel with $|\lambda| \neq 1$ (*orthogonal sub-channel*) allows for a separate observation of each source. The two observations are not independent though, because of the correlated additive noise. Assuming a RRC pulse and considering the eigenvalues λ_n as equidistant samples of the modulated energy spectrum, a direct connection between the number of *orthogonal sub-channels* and the roll-off factor can be made. The amplitude of the modulated energy spectrum is equal to one on the interval $\frac{-1}{2}(1-\beta) < F < \frac{1}{2}(1-\beta)$ and less than one on the rest of the interval $\left\langle \frac{-1}{2}; \frac{1}{2} \right\rangle$. As consequence, the number of *non-orthogonal sub-channels* is given as

$$\#\left(|\lambda_n| = 1\right) = 1 + 2\left\lfloor \frac{(1-\beta)\frac{1}{2}}{\frac{1}{N}} \right\rfloor = 1 + 2\left\lfloor \frac{(1-\beta)N}{2} \right\rfloor = \left\lfloor 1 + (1-\beta)N \right\rfloor. \tag{6.96}$$

And in turn, the number of *orthogonal sub-channels* equals to $N - \lfloor 1 + (1-\beta)N \rfloor$. It follows that for $\beta = 0$, no extra spectrum is used thus no *orthogonal sub-channels* are available.

In the rest of this section, we evaluate the mutual information (MI) of such a two-dimensional subchannel. Generally, we are interested in the hierarchical target symbol, i.e. the result of the many-to-one HNC map. Therefore we are interested in the overall hierarchical information [111]

$$\mathrm{I}\left(C; X\right) = \mathcal{H}\left[\boldsymbol{X}\right] - \mathcal{H}\left[\boldsymbol{X}|C\right], \tag{6.97}$$

where $\boldsymbol{X} = \left[X_1, X_2\right]^{\mathrm{T}}$. For the computation of (6.97) we express

$$\mathrm{p}_{\boldsymbol{X}}\left(\left[x'_{1,n}, x'_{2,n}\right]^{\mathrm{T}}\right) = \frac{1}{M^2} \sum_{(c_A, c_B)} \mathrm{e}^{-\boldsymbol{q}_n(c_A,c_B)^{\mathrm{H}} \boldsymbol{K}^{-1} \boldsymbol{q}_n(c_A,c_B)} \tag{6.98}$$

and

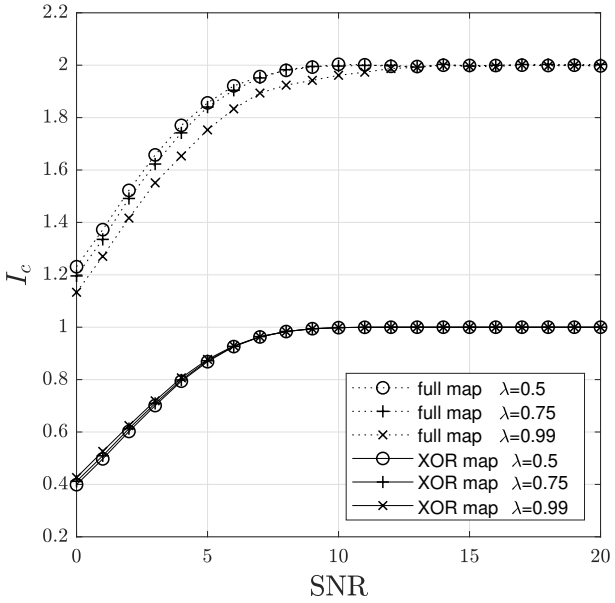
$$\mathrm{p}_{\boldsymbol{X}|C}\left(\left[x'_{1,n}, x'_{2,n}\right]^{\mathrm{T}} | c_n\right) = \frac{1}{M} \sum_{\chi_c(c_A,c_B) = \ c_n} \mathrm{e}^{-\boldsymbol{q}_n(c_A,c_B)^{\mathrm{H}} \boldsymbol{K}^{-1} \boldsymbol{q}_n(c_A,c_B)}, \tag{6.99}$$

where M is the alphabet cardinality, and $\boldsymbol{q}, \boldsymbol{K}$ are defined according to (6.78), (6.79) respectively. The entropies were obtained by numerical Monte-Carlo evaluation as

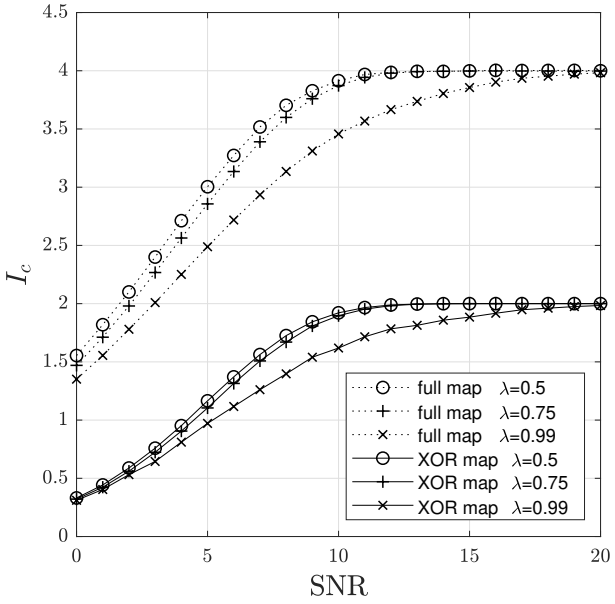
$$\mathcal{H}\left[\boldsymbol{X}\right] \approx -\mathrm{E}_{\left[x'_{1,n}, x'_{2,n}\right]^{\mathrm{T}} \sim \mathrm{p}_{\boldsymbol{X}}} \left[\lg\left(\mathrm{p}_{\boldsymbol{X}}\left(\left[x'_{1,n}, x'_{2,n}\right]^{\mathrm{T}}\right)\right)\right] \tag{6.100}$$

$$\mathcal{H}\left[\boldsymbol{X}|C\right] \approx -\mathrm{E}_{\left(\left[x'_{1,n}, x'_{2,n}\right]^{\mathrm{T}}, c_n\right) \sim \mathrm{p}_{\boldsymbol{X},C}} \left[\lg\left(\mathrm{p}_{\boldsymbol{X}|C}\left(\left[x'_{1,n}, x'_{2,n}\right]^{\mathrm{T}} | c_n\right)\right)\right]. \tag{6.101}$$

The results are shown in Figures 6.20a and 6.20b for BPSK and QPSK alphabets respectively. For reference, we included the MI of a full map, i.e. the MI between $\boldsymbol{X}$ and $[C_A, C_B]$. The plots show the dependence of MI on the channel SNR for various values of λ while $h_A = 1$ and $|h_B| = 0.75$ are fixed and $\angle h_B$ is sampled from a uniform distribution. We considered only real-valued λ because $\angle\lambda$ gets combined with the relative phase $\angle h_B - \angle h_A$ randomization. We observe an increase of the MI for decreasing values of $|\lambda|$. While for the BPSK case and XOR map the results seem almost independent on $|\lambda|$, for the QPSK case this dependence is significant and more pronounced in the case of a full map. This behavior can be related to an increase in the number of degrees of freedom formed by the channel. In a noiseless situation, $|\lambda| = 1$ would result in $x'_{1,n} = \left(x'_{2,n}\right)^*$ with only one complex degree of freedom while $|\lambda| \leq 1$ on the other hand would lead to an invertible channel (6.94) with two complex degrees of freedom. With the noise included, the transition gets "smooth" in the sense that smaller $|\lambda|$ give better results.



(a) BPSK alphabet



(b) QPSK alphabet

Figure 6.20: Numerical evaluation of the overall hierarchical mutual information $\mathrm{I}(C;X)$ of a subchannel given by (6.74). The channel coefficient $h_A = 1$ was kept constant and $|h_B| = 0.75$ while $\angle h_B$ was considered random with uniform distribution.

6.4.8 Numerical Results

In this section, we present numerical results obtained by computer simulations. The simulations were run with fixed channel coefficient amplitudes $|h_A| = 1, |h_B| = 0.7$ and random, uniformly distributed channel coefficient phases $\angle h_A, \angle h_B$. Individual frames with $N = 101$ symbols were passed through an appropriate channel model with SNR = 10dB and various demodulation techniques were applied. In the flowing figures we present the H-BER dependence on the temporal misalignment δ obtained using a RRC pulse with various roll-off values β.

BPSK Alphabet with XOR HNC Map

Observations:

- In general, we see that for smaller β a larger m is needed and vice versa, for $\beta = \frac{1}{2}$ the choice of $m = 3$ seems to be sufficient. This is not surprising as with the increase of β, the pulse widens in the frequency spectrum and gets more focused in the time domain.
- The reference method with artificial interference suppression (Figure 6.21) does not permit to fully overcome the performance degradation with rising δ for neither roll-off.
- Mid-point sampling with precoding (method (A) in Figure 6.21) achieves an H-BER independent on the delay δ only for $\beta = 0$. Still it performs better than the reference method.
- The precoding method (B) (double rate sampling in Figure 6.21) is able to achieve an H-BER independent on δ for all the cases with different β.
- In Figure 6.21 we observe that method (A) performs almost equally as the reference scenario with the maximal $m = 50$. For the selected roll-off $\beta = 0.5$, only method (B) is able to fully remove the performance penalty resulting from a nonzero delay δ.

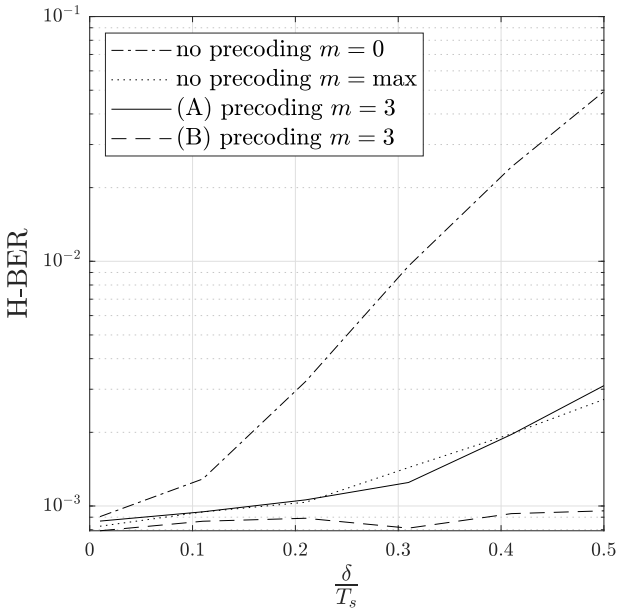
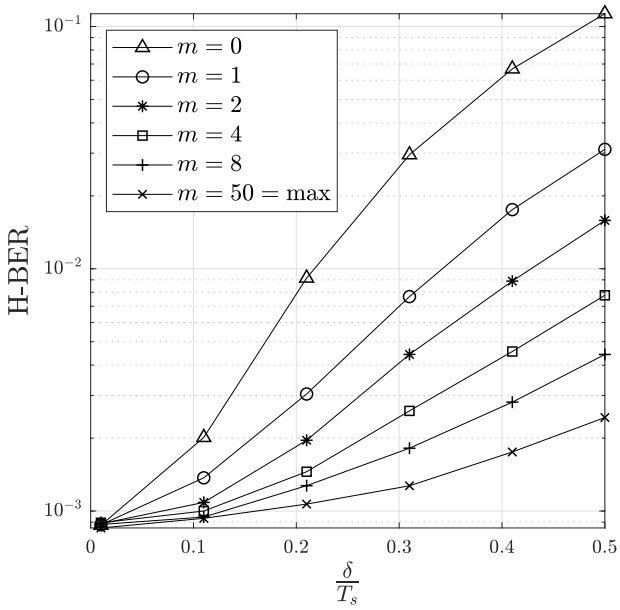
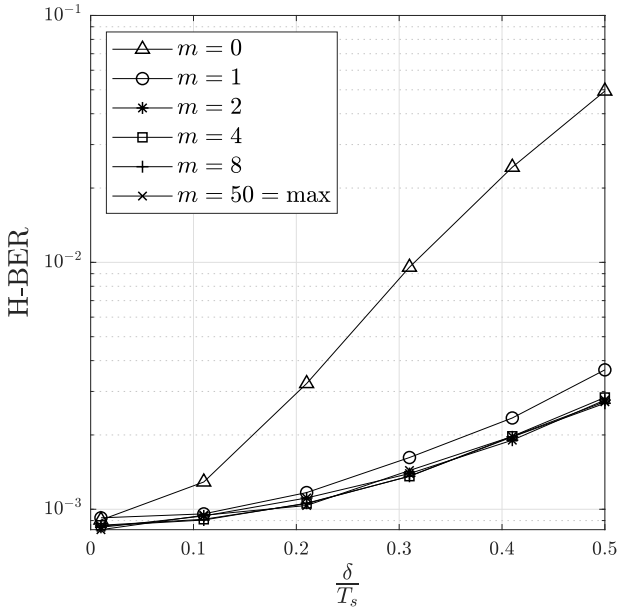


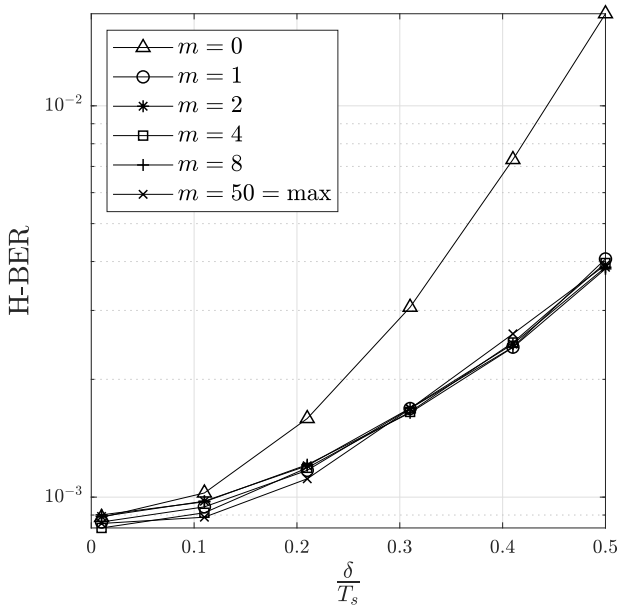
Figure 6.21: Comparison of the performance of all three methods with the BPSK alphabet and $\beta = \frac{1}{2}$.



(a) $\beta = 0$

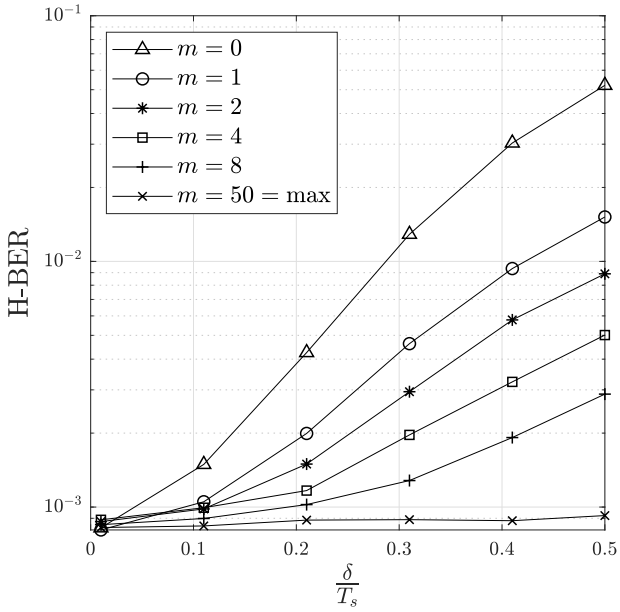


(b) $\beta = \frac{1}{2}$

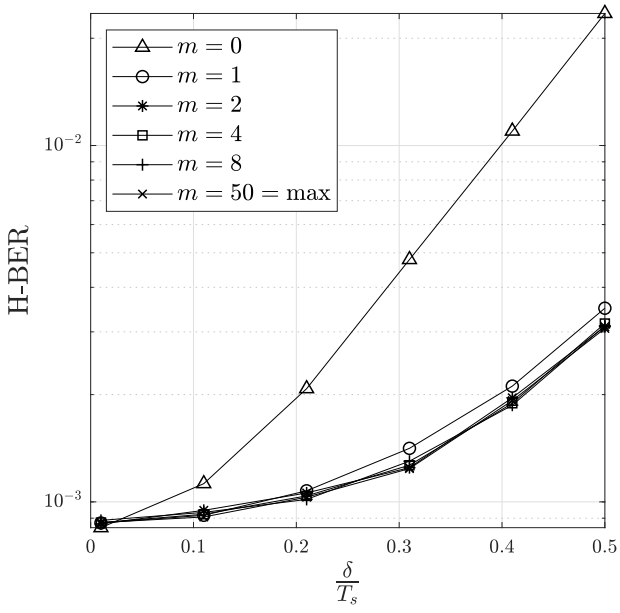


(c) $\beta = 1$

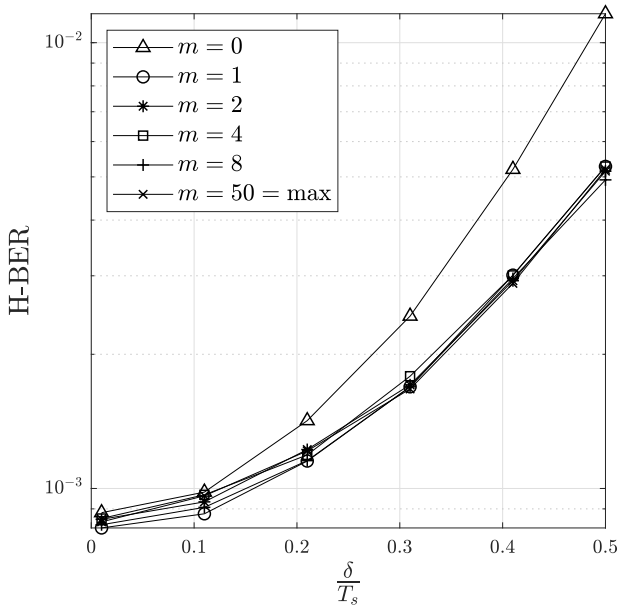
Figure 6.21: H-BER results for a reference scenario with full data aid and mid-point sampling without precoding according to (6.25).



(a) $\beta = 0$

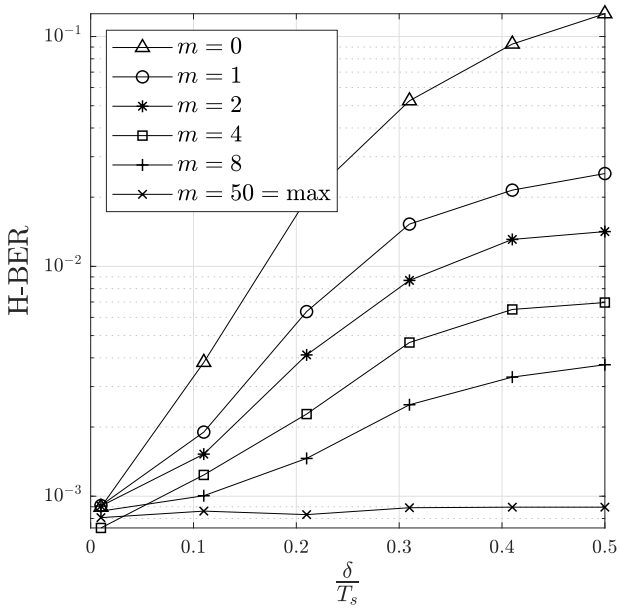


(b) $\beta = \frac{1}{2}$

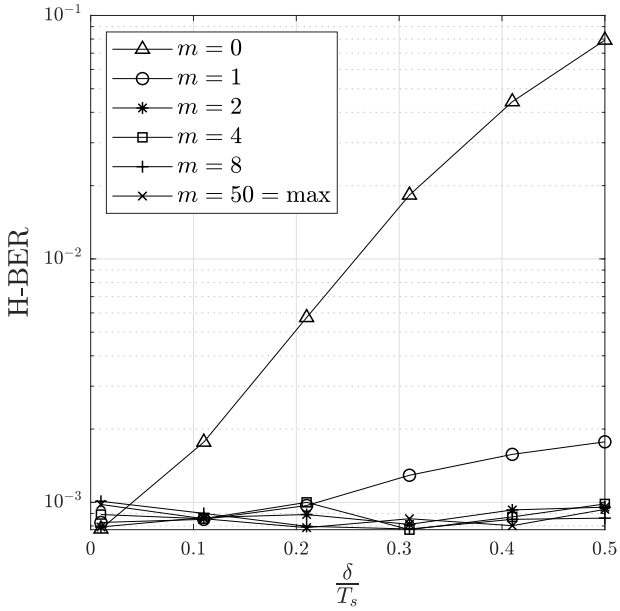


(c) $\beta = 1$

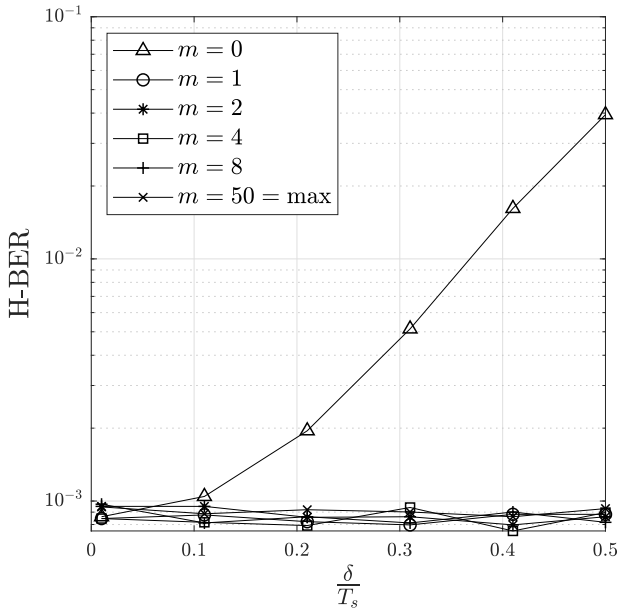
Figure 6.21: H-BER results for the precoding approach method (A), i.e. sampling once per T_s in between both sources according to (6.60).



(a) $\beta = 0$



(b) $\beta = \frac{1}{2}$



(c) $\beta = 1$

Figure 6.21: H-BER results for the precoding method (B), i.e. sampling twice per T_s once synchronously with each of the sources according to (6.81).

QPSK Alphabet with XOR HNC Map

The observations stated in the previous section for the BPSK alphabet case apply here as well. The plots for QPSK are shown in Figures 6.22 (reference scenario with artificial interference suppression), 6.22 (precoding with symbol spaced sampling), 6.22 (precoding with double rate sampling), and 6.22 (comparison of all the methods with $\beta = 0.5$). The only difference is that for method (B) we observe (Figures 6.22 and 6.22) a performance increase with rising δ for $\beta > 0$.

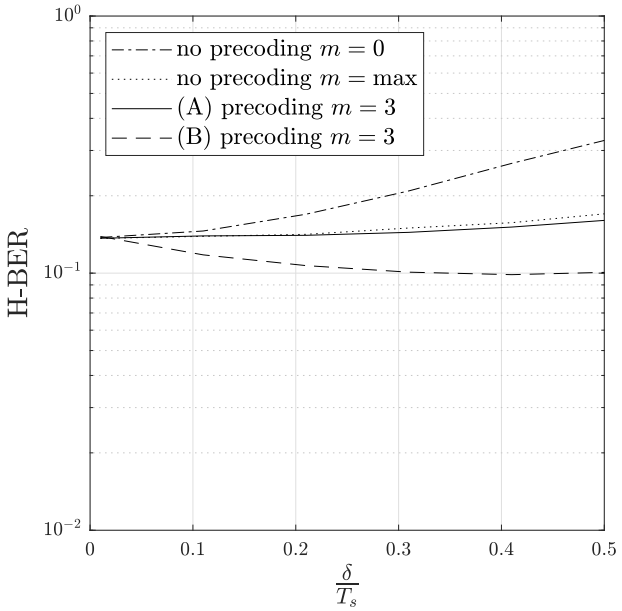
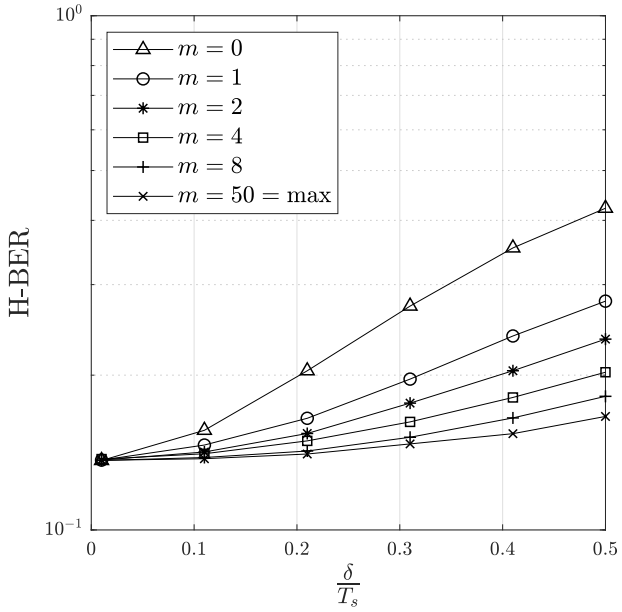
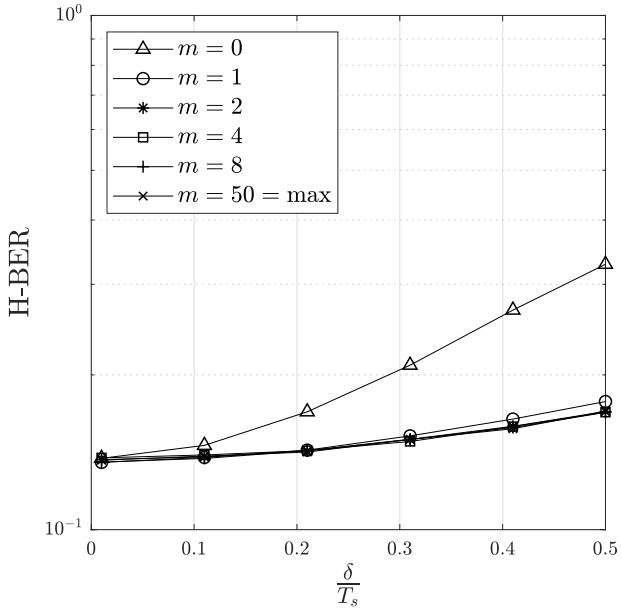


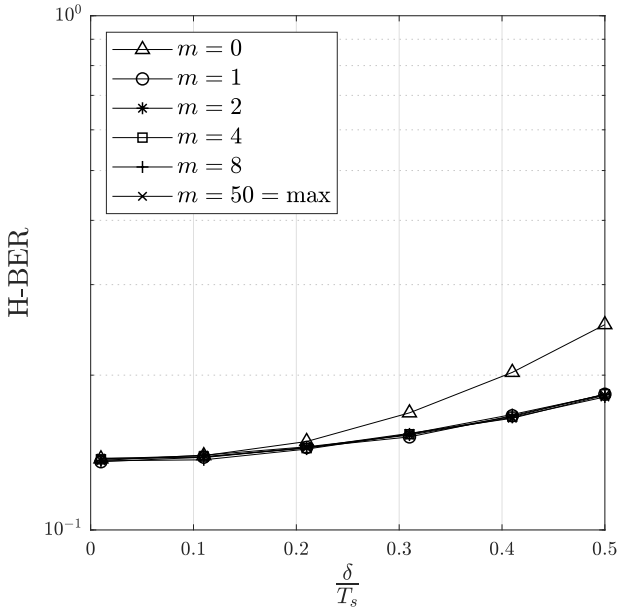
Figure 6.22: Comparison of the performance of all three methods with QPSK alphabet and $\beta = \frac{1}{2}$.



(a) $\beta = 0$

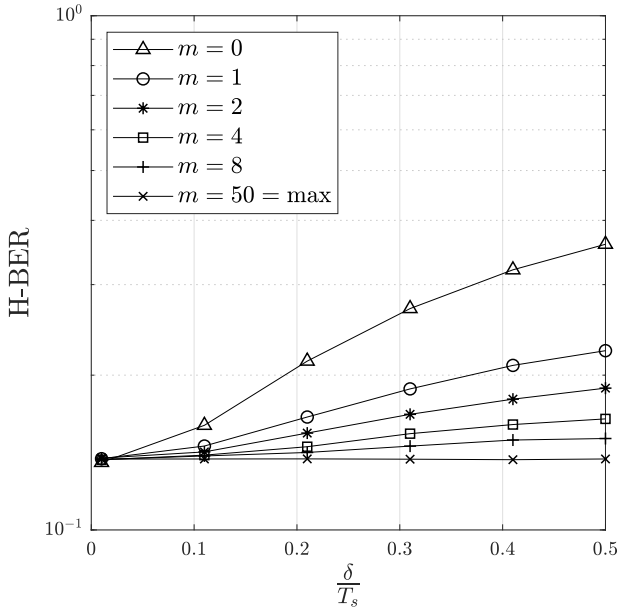


(b) $\beta = \frac{1}{2}$

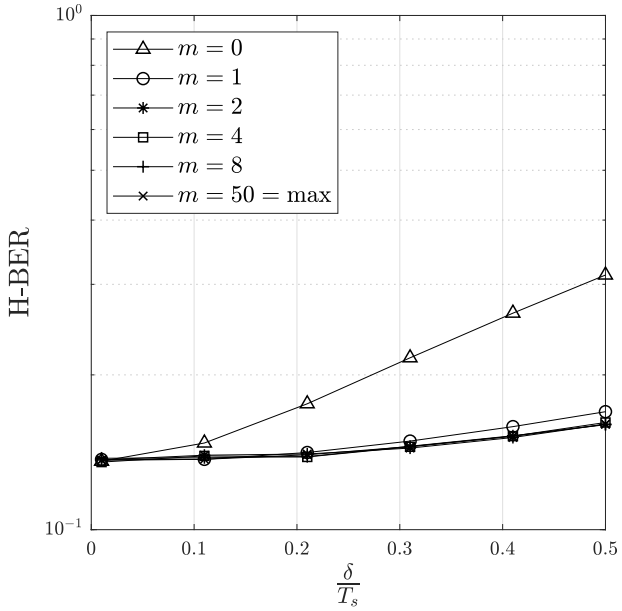


(c) $\beta = 1$

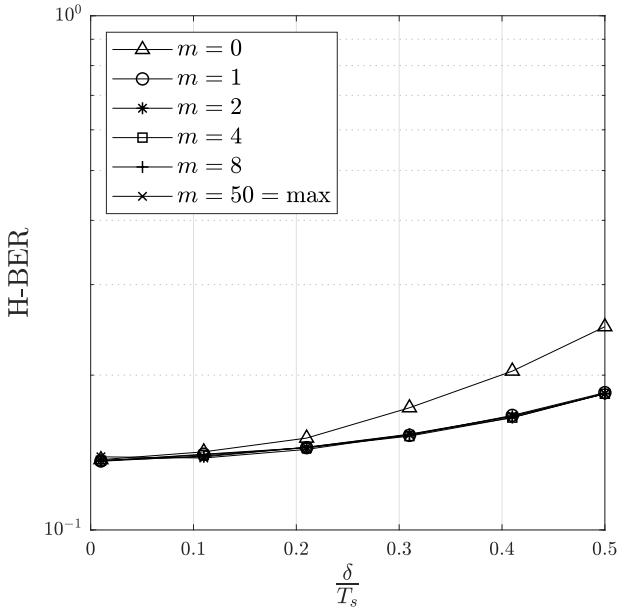
Figure 6.22: H-BER results for the reference scenario with full data aid and uncoded mid-point sampling according to (6.25).



(a) $\beta = 0$

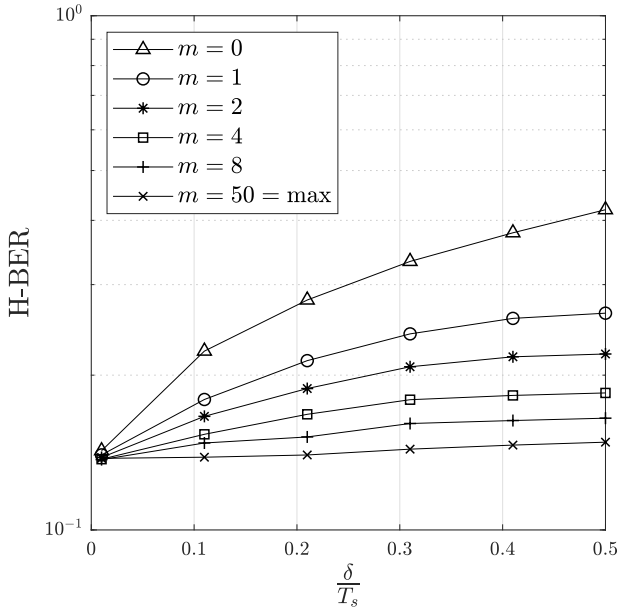


(b) $\beta = \frac{1}{2}$

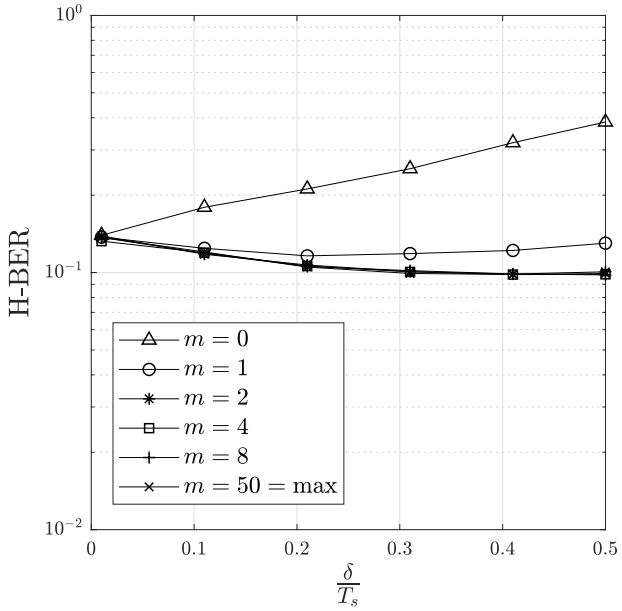


(c) $\beta = 1$

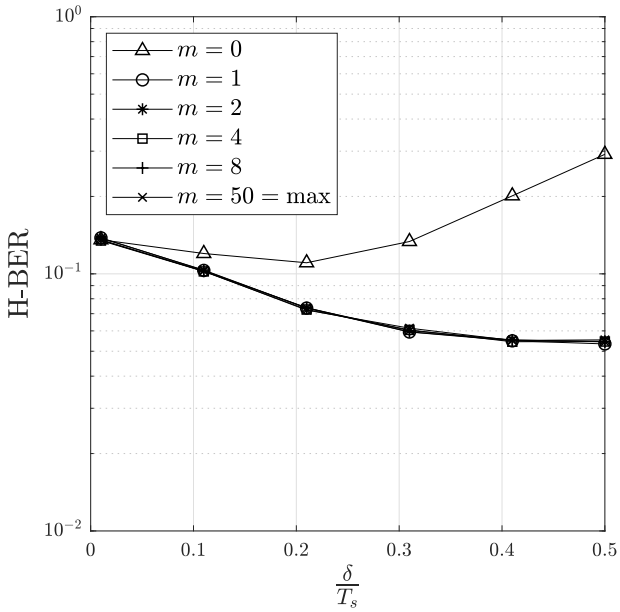
Figure 6.22: H-BER results for the precoding approach method (A), that is sampling once per T_s in between both sources according to (6.60).



(a) $\beta = 0$



(b) $\beta = \frac{1}{2}$



(c) $\beta = 1$

Figure 6.22: H-BER results for the precoding method (B), i.e. sampling twice per T_s once synchronously with each of the sources according to (6.81).

6.4.9 Conclusion

For clarity, we restate the considered scenario, suggested solutions and conclude the obtained results.

In this chapter, we considered the problem of asynchronous reception in a two-source H-MAC channel with constant channel coefficients and AWGN. At the transmitters, we considered a linear modulation with a general Nyquist modulation pulse $g(t)$ and a standard MF demodulation at the receiver. Based on those assumptions, we analyzed the impact of the temporal misalignment between the two source signals on the received samples. To do so, we considered a compound discrete channel by concatenating the transmitter side expansion part, the continuous time channel, and the MF sampling at the receiver as shown in Figure 6.23.

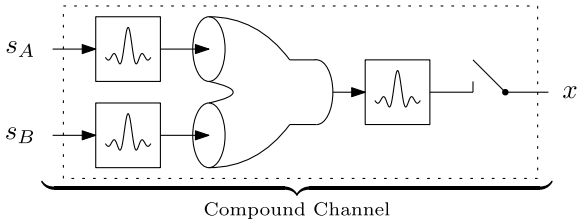


Figure 6.23: Unison of the modulation pulse filtering, continuous time channel, and MF sampling into one discrete time channel.

Further, we used a vector representation and showed how the discrete-time channel can be modeled by a simple matrix multiplication according to Figure 6.24.

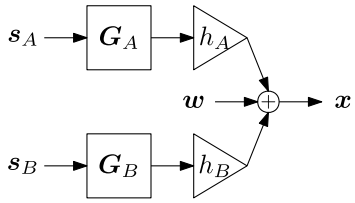


Figure 6.24: Compound channel vector representation.

The asynchronicity of the continuous-time channel translates to a gain loss in the corresponding discrete-time compound channel in two fundamental ways:

Problem 1 As long as the considered modulation pulse $g(t)$ is synchronized with the corresponding MF $g^*(t)$, the Nyquist property assures that the whole energy of the pulse gets projected on the corresponding MF. In turn, the entire energy of the symbol s_n is captured by the corresponding sample x_n. On the other hand, if the pulse is delayed by a $\delta \neq lT_s$, the Nyquist property does not hold anymore and we observe nonzero mutual energy between $g(t-\delta)$ and the MFs $g^*(t-nT_s)$ with $n \neq 0$. As a consequence, the energy of the symbol s_n gets spread over multiple samples x_l and *intersymbol interference* is introduced. A schematic representation is shown in Figure 6.25).

Problem 2 Based on the bandwidth of the modulation pulse $g(t)$ and the sampling rate at the receiver, some energy of the delayed pulse $g(t-\delta)$ may get lost when passing through the channel. This happens in the case when $g(t-\delta) \notin \operatorname{span}\langle\{g^*(t-nT_s)\}_n\rangle$, where only the energy associated to the projection of $g(t-\delta)$ onto $\operatorname{span}\langle\{g^*(t-nT_s)\}_n\rangle$ is received.

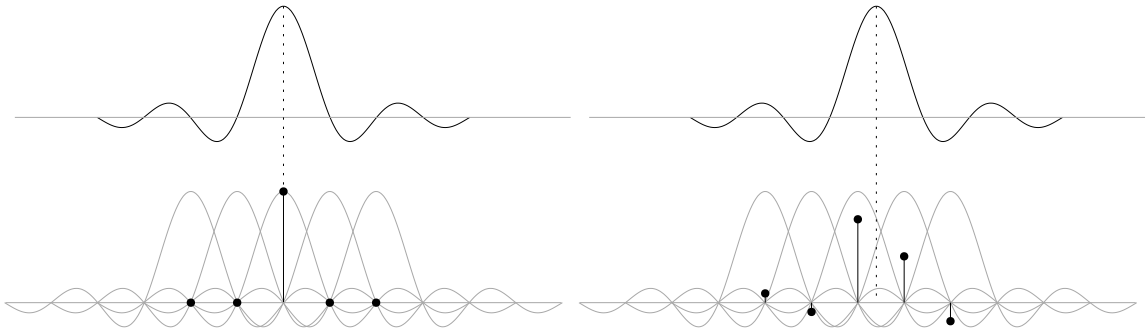


Figure 6.25: Intersymbol interference due to temporal misalignment.

The first strategy we studied was based on intersymbol interference suppression. The idea of the interference suppression approach was to perform a mid-point sampling and to remove the symbol interference from the individual samples before passing them through a standard ML demodulator. In the context of the energy transfer stated above, this method does not have the potential to solve either of the two problems. By a simple removal of the interference, the observation becomes memoryless and allows for a simple symbol by symbol demodulation. At the same time, the energy carried by the interference is wasted. The difficulty with this method is to accurately model the interference, which depends on a large number of surrounding symbols. The first approach was to model it as purely random with a Gaussian mixture density. Because it did not provide significant improvement, we considered a data-aided method. The idea was to use some a priori knowledge of the surrounding symbols based on which we could model the interference more accurately. The simulations have shown that a full data aid does help and reduces the H-BER of the subsequent ML demodulation. In the context of WPNC however, we do not have access to the full data at the relay. An attempt of using only a hierarchical data aid, a XOR function in our case, did not prove to be effective. An analysis of the resulting interference distribution showed that the XOR aid does not help much to lower the variance w.r.t. the fully random model.

A different method arose from utilizing an approximation of the modulation pulse ACF and cyclic extension to render the channel matrix $\boldsymbol{G}$ circulant. As a consequence, it was possible to apply a DFT to the transmitted signals and convert the matrix to a diagonal form as is the case in a regular OFDM. We showed that working with a circulant matrix and DFT precoding together with symbol-spaced midpoint sampling allows for solving the first problem. In contrast with the interference suppression method, in this case it is possible to collect the interference coherently and make use of the energy it contains. DFT precoding does not solve the second problem because the symbol-spaced sampling does not form a sufficient statistics in general and some part of the energy pulse may get lost.

To combat the second problem, we continued with the DFT precoding and applied double rate receiver-side processing. In particular, we used to sets of MFs, each synchronized with one source, and obtained two samples per symbol time duration. Based on numerical evaluation in the case of a RRC pulse, we argued that the double-rate sample pairs can be considered almost independent. With this assumption, we defined a corresponding two-dimensional channel and analyzed its properties such as invertibility, mutual information, and H-BER. This method is able to solve both of the problems. Thanks to the DFT precoding, it is able to collect the energy from the interference and the double rate sampling allows to capture all the dimensions of the useful signal. Subsequent mutual information analysis and H-BER evaluations showed that for a higher alphabet, the 2-dimensional channel performs better with a nonzero delay δ.

In Table 6.1 we give a brief summary of the relations between the discussed methods and problems 1 and 2.

method	problem 1	problem 2
interference suppression	Can be solved partially, the removal of interference facilitates a symbol by symbol demodulation but the energy carried by the interference is lost.	Does not prevent the energy from leaking into hidden dimensions. For a RRC pulse with roll-off $\beta = 0$, this problem does not occur.
DFT precoding, symbol spaced sampling	Allows to solve the interference problem and exploit the interference energy.	Does not prevent the energy from leaking into hidden dimensions. For a RRC pulse with roll-off $\beta = 0$, this problem does not occur.
DFT precoding, double rate sampling	Allows to solve the interference problem and exploit the interference energy.	Allows to observe all the dimensions of the delayed signal and capture all the energy for a RRC pulse with an arbitrary roll-off β.

Table 6.1: Summary of individual methods characteristics.

Chapter 7

Conclusions

7.1 Summary of Contribution

We presented several communication techniques for a parametrical H-MAC with partial channel state knowledge. We cover two main cases, the problem of unknown channel phase and reception with temporary misaligned source signals. Under the conditions of a WPNC network with nonorthogonal multiple access, the channel parametrization is very different from a classical single-user case and requires a different approach and new modulation/demodulation methods.

Our first contribution consists of an analysis of the relative channel phase parametrization of an H-MAC and the proposition of several methods how to cope with it in the case of a two-source setup. One method is based on a hierarchical data-aided estimation procedure, which allows for precise channel phase tracking during the payload transmission without the need for separate orthogonal resources for pilot signals. A different approach is based on differential modulation and several demodulation algorithms which operate in a noncoherent fashion and allow a simple and robust implementation without the need for precise channel phase estimation.

The second contribution is devoted to an asynchronous H-MAC. In a real H-MAC, it is essentially impossible to assure perfect synchronization among all the source nodes and the corresponding relay. We analyze the impact of a fractional symbol time duration delay in a two-source H-MAC and propose several techniques to minimize the negative impact thereof. The solution is based on OFDM-like precoding at the transmitters and corresponding demodulation at the receiver. We suggest different variants, employing symbol-spaced as well as fractionally-spaced processing at the relay.

More detailed conclusions can be found at the end of each respective chapter.

7.2 Future Research

The ultimate goal of wireless cloud networks with weak system information is to integrate the radio environment sensing and the payload data transmission in a joint process. The uncertainty about the network state (such as individual channel states, processing/coding strategies of the constituent nodes, and overall connectivity) would be treated as an additional degree of freedom in a similar way as the payload data itself. This could give a rise to so-called super-relay clouds, that is networks consisting of many autonomous nodes acting as relays and forming a compact could network. The hierarchical information would pass through many, mutually interacting paths within the network, without revealing the internal cloud structure to the source and/or destination terminals. The overall end-to-end connectivity as well as the intrinsic structure of the cloud would be achieved by a distributed and cooperative effort of the relay nodes.

Bibliography

[1] MN Abdallah, W Dyab, TK Sarkar, MVSN Prasad, CS Misra, A Lamparez, M Salazar-Palma, and SW Ting. Further validation of an electromagnetic macro model for analysis of propagation path loss in cellular networks using measured driving-test data. *IEEE Antennas and Propagation Magazine*, 56(4):108–129, 2014.

[2] Giuseppe Aceto, Valerio Persico, and Antonio Pescapé. Industry 4.0 and health: Internet of things, big data, and cloud computing for healthcare 4.0. *Journal of Industrial Information Integration*, 18:100129, 2020.

[3] Rudolf Ahlswede, Ning Cai, S-YR Li, and Raymond W Yeung. Network information flow. *IEEE Transactions on information theory*, 46(4):1204–1216, 2000.

[4] Ian F Akyildiz, Ahan Kak, and Shuai Nie. 6g and beyond: The future of wireless communications systems. *IEEE access*, 8:133995–134030, 2020.

[5] Ian F Akyildiz, Xudong Wang, and Weilin Wang. Wireless mesh networks: a survey. *Computer networks*, 47(4):445–487, 2005.

[6] Saif Al-Sultan, Moath M Al-Doori, Ali H Al-Bayatti, and Hussien Zedan. A comprehensive survey on vehicular ad hoc network. *Journal of network and computer applications*, 37:380–392, 2014.

[7] Muntadher Alsabah, Marwah Abdulrazzaq Naser, Basheera M Mahmmod, Sadiq H Abdulhussain, Mohammad R Eissa, Ahmed Al-Baidhani, Nor K Noordin, Sadiq M Sait, Khaled A Al-Utaibi, and Fazirul Hashim. 6g wireless communications networks: A comprehensive survey. *IEEE Access*, 9:148191–148243, 2021.

[8] Abdollah Amirkhani and Amir Hossein Barshooi. Consensus in multi-agent systems: a review. *Artificial Intelligence Review*, 55(5):3897–3935, 2022.

[9] Jeffrey G Andrews, Xinchen Zhang, Gregory D Durgin, and Abhishek K Gupta. Are we approaching the fundamental limits of wireless network densification? *IEEE Communications Magazine*, 54(10):184–190, 2016.

[10] Hamidreza Arasteh, Vahid Hosseinnezhad, Vincenzo Loia, Aurelio Tommasetti, Orlando Troisi, Miadreza Shafie-khah, and Pierluigi Siano. Iot-based smart cities: A survey. In *2016 IEEE 16th international conference on environment and electrical engineering (EEEIC)*, pages 1–6. IEEE, 2016.

[11] Ertugrul Basar and Ibrahim Altunbas. Space-time channel modulation. *IEEE Transactions on Vehicular Technology*, 66(8):7609–7614, 2017.

[12] Ertugrul Basar, Marco Di Renzo, Julien De Rosny, Merouane Debbah, Mohamed-Slim Alouini, and Rui Zhang. Wireless communications through reconfigurable intelligent surfaces. *IEEE access*, 7:116753–116773, 2019.

[13] Ertugrul Basar, Miaowen Wen, Raed Mesleh, Marco Di Renzo, Yue Xiao, and Harald Haas. Index modulation techniques for next-generation wireless networks. *IEEE access*, 5:16693–16746, 2017.

[14] Riccardo Bassoli, Hugo Marques, Jonathan Rodriguez, Kenneth W Shum, and Rahim Tafazolli. Network coding theory: A survey. *IEEE Communications Surveys & Tutorials*, 15(4):1950–1978, 2013.

[15] Claude Berrou, Ramesh Pyndiah, Patrick Adde, Catherine Douillard, and Raphaël Le Bidan. An overview of turbo codes and their applications. In *The European Conference on Wireless Technology, 2005.*, pages 1–9. IEEE, 2005.

[16] Mudit Ratana Bhalla, Anand Vardhan Bhalla, et al. Generations of mobile wireless technology: A survey. *International Journal of Computer Applications*, 5(4):26–32, 2010.

[17] Sherif Adeshina Busari, Kazi Mohammed Saidul Huq, Shahid Mumtaz, and Jonathan Rodriguez. Terahertz massive mimo for beyond-5g wireless communication. In *ICC 2019-2019 IEEE International Conference on Communications (ICC)*, pages 1–6. IEEE, 2019.

[18] Philip A Chou, Yunnan Wu, and Kamal Jain. Practical network coding. In *Proceedings of the annual Allerton conference on communication control and computing*, volume 41, pages 40–49. The University; 1998, 2003.

[19] Mostafa Zaman Chowdhury, Md Tanvir Hossan, Amirul Islam, and Yeong Min Jang. A comparative survey of optical wireless technologies: Architectures and applications. *ieee Access*, 6:9819–9840, 2018.

[20] U Cisco. Cisco annual internet report (2018–2023) white paper. 2020. *Acessado em*, 10(01):1–35, 2021.

[21] Marco Conti and Silvia Giordano. Multihop ad hoc networking: The theory. *IEEE communications Magazine*, 45(4):78–86, 2007.

[22] Daniel J Costello and G David Forney. Channel coding: The road to channel capacity. *Proceedings of the IEEE*, 95(6):1150–1177, 2007.

[23] T. M. Cover and Joy A. Thomas. *Elements of information theory.* Wiley, Hoboken, 2nd edition, 2006.

[24] Jianxiong Cui, Guangliang Dong, Shengli Zhang, Haitao Li, and Guinian Feng. Asynchronous noma for downlink transmissions. *IEEE Communications Letters*, 21(2):402–405, 2016.

[25] Tao Cui, FeiFei Gao, and Chintha Tellambura. Physical layer differential network coding for two-way relay channels. In *IEEE GLOBECOM 2008-2008 IEEE Global Telecommunications Conference*, December 2008.

[26] Linglong Dai, Bichai Wang, Yifei Yuan, Shuangfeng Han, I Chih-Lin, and Zhaocheng Wang. Non-orthogonal multiple access for 5g: solutions, challenges, opportunities, and future research trends. *IEEE Communications Magazine*, 53(9):74–81, 2015.

[27] Philip J Davis. Circulant matrices wiley. *New york*, 1979.

[28] Marco Di Renzo, Harald Haas, and Peter M Grant. Spatial modulation for multiple-antenna wireless systems: A survey. *IEEE Communications Magazine*, 49(12):182–191, 2011.

[29] Marco Di Renzo, Michela Iezzi, and Fabio Graziosi. Beyond routing via network coding: An overview of fundamental information-theoretic results. In *21st Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, pages 2745–2750. IEEE, 2010.

[30] G Dileep. A survey on smart grid technologies and applications. *Renewable energy*, 146:2589–2625, 2020.

[31] Zhiguo Ding, Xianfu Lei, George K Karagiannidis, Robert Schober, Jinhong Yuan, and Vijay K Bhargava. A survey on non-orthogonal multiple access for 5g networks: Research challenges and future trends. *IEEE Journal on Selected Areas in Communications*, 35(10):2181–2195, 2017.

[32] Randall Dougherty, Christopher Freiling, and Kenneth Zeger. Insufficiency of linear coding in network information flow. *IEEE transactions on information theory*, 51(8):2745–2759, 2005.

[33] Abbas El Gamal and Young-Han Kim. *Network information theory*. Cambridge university press, 2011.

[34] Peter Elias, Amiel Feinstein, and Claude Shannon. A note on the maximum flow through a network. *IRE Transactions on Information Theory*, 2(4):117–119, 1956.

[35] Dong Fang and Alister Burr. Rotationally invariant coded modulation for physical layer network coding in two-way relay fading channel. In *European Wireless 2012; 18th European Wireless Conference 2012*, pages 1–6. VDE, 2012.

[36] Christina Fragouli. Network coding: Beyond throughput benefits. *Proceedings of the IEEE*, 99(3):461–475, 2011.

[37] Christina Fragouli, Jean-Yves Le Boudec, and Jörg Widmer. Network coding: an instant primer. *ACM SIGCOMM Computer Communication Review*, 36(1):63–68, 2006.

[38] Christina Fragouli, Emina Soljanin, et al. Network coding fundamentals. *Foundations and Trends® in Networking*, 2(1):1–133, 2007.

[39] Robert Gallager. Low-density parity-check codes. *IRE Transactions on information theory*, 8(1):21–28, 1962.

[40] Feifei Gao, Chengwen Xing, and Gongpu Wang. *Channel Estimation for Physical Layer Network Coding Systems*. Springer, 2014.

[41] Feifei Gao, Rui Zhang, and Y-C Liang. Optimal channel estimation and training design for two-way relay networks. *IEEE Transactions on Communications*, 57(10):3024–3033, 2009.

[42] David Gesbert, Marios Kountouris, Robert W Heath, Chan-Byoung Chae, and Thomas Salzer. Shifting the mimo paradigm. *IEEE signal processing magazine*, 24(5):36–46, 2007.

[43] Andrea Goldsmith, Michelle Effros, Ralf Koetter, Muriel Médard, Asu Ozdaglar, and Lizhong Zheng. Beyond shannon: the quest for fundamental performance limits of wireless ad hoc networks. *IEEE Communications Magazine*, 49(5):195–205, 2011.

[44] Piyush Gupta and Panganmala R Kumar. The capacity of wireless networks. *IEEE Transactions on information theory*, 46(2):388–404, 2000.

[45] Shimaa A Abdel Hakeem, Hanan H Hussein, and HyungWon Kim. Vision and research directions of 6g technologies and applications. *Journal of King Saud University-Computer and Information Sciences*, 2022.

[46] Hannes Hartenstein and LP Laberteaux. A tutorial survey on vehicular ad hoc networks. *IEEE Communications magazine*, 46(6):164–171, 2008.

[47] Miroslav Hekrdla and Jan Sykora. Hexagonal constellations for adaptive physical-layer network coding 2-way relaying. *IEEE Communications Letters*, 18(2):217–220, 2014.

[48] Tracey Ho and Desmond Lun. *Network coding: an introduction*. Cambridge University Press, 2008.

[49] Tracey Ho, Muriel Médard, Ralf Koetter, David R Karger, Michelle Effros, Jun Shi, and Ben Leong. A random linear network coding approach to multicast. *IEEE Transactions on information theory*, 52(10):4413–4430, 2006.

[50] Bengt Holter. On the capacity of the mimo channel: A tutorial introduction. In *Proc. IEEE Norwegian Symposium on Signal Processing*, volume 2, pages 167–172. Citeseer, 2001.

[51] Petr Hron. Channel estimation and network coded modulation for parametric h-mac channel in wpnc radio networks. Master's thesis, Prague CZ, 2019.

[52] Petr Hron, Jozef Lukac, and Jan Sykora. Sdr verification of hierarchical decision aided 2-source bpsk h-mac cse with feed-back gradient solver for wpnc networks. *Radioengineering*, 29(3):573, 2020.

[53] Petr Hron and Jan Sykora. Performance analysis of hierarchical decision aided 2-source bpsk h-mac cse with feed-back gradient solver for wpnc networks. In *2019 IEEE Microwave Theory and Techniques in Wireless Communications (MTTW)*, volume 1, pages 72–75. IEEE, 2019.

[54] Petr Hron and Jan Sykora. Closed-form differential hierarchical demodulator of 2-source hierarchical-mac channel. *IEEE Communications Letters*, 25(11):3704–3708, 2021.

[55] Chongwen Huang, Alessio Zappone, George C Alexandropoulos, Mérouane Debbah, and Chau Yuen. Reconfigurable intelligent surfaces for energy efficiency in wireless communication. *IEEE transactions on wireless communications*, 18(8):4157–4170, 2019.

[56] Yasin Kabalci. A survey on smart metering and smart grid communication. *Renewable and Sustainable Energy Reviews*, 57:302–318, 2016.

[57] Nadège Kaina, Matthieu Dupré, Geoffroy Lerosey, and Mathias Fink. Shaping complex microwave fields in reverberating media with binary tunable metasurfaces. *Scientific reports*, 4(1):6693, 2014.

[58] Sachin Katti, Shyamnath Gollakota, and Dina Katabi. Embracing wireless interference: Analog network coding. *ACM SIGCOMM Computer Communication Review*, 37(4):397–408, 2007.

[59] Latif U Khan, Ibrar Yaqoob, Nguyen H Tran, SM Ahsan Kazmi, Tri Nguyen Dang, and Choong Seon Hong. Edge-computing-enabled smart cities: A comprehensive survey. *IEEE Internet of Things Journal*, 7(10):10200–10232, 2020.

[60] Hakpyeong Kim, Heeju Choi, Hyuna Kang, Jongbaek An, Seungkeun Yeom, and Taehoon Hong. A systematic review of the smart energy conservation system: From smart homes to sustainable smart cities. *Renewable and sustainable energy reviews*, 140:110755, 2021.

[61] Ralf Koetter and Frank R Kschischang. Coding for errors and erasures in random network coding. *IEEE Transactions on Information theory*, 54(8):3579–3591, 2008.

[62] Toshiaki Koike-Akino, Petar Popovski, and Vahid Tarokh. Denoising maps and constellations for wireless network coding in two-way relaying systems. In *IEEE GLOBECOM 2008-2008 IEEE Global Telecommunications Conference*, pages 1–5. IEEE, 2008.

[63] Toshiaki Koike-Akino, Petar Popovski, and Vahid Tarokh. Optimized constellations for two-way wireless relaying with physical network coding. *IEEE journal on Selected Areas in Communications*, June 2009.

[64] Jakub Kolar, Jan Sykora, and Umberto Spagnolini. Distributed network tomography applied to stochastic delay profile estimation. *Radioengineering*, 29(1), 2020.

[65] Nikos Komninos, Eleni Philippou, and Andreas Pitsillides. Survey in smart grid and smart home security: Issues, challenges and countermeasures. *IEEE Communications Surveys & Tutorials*, 16(4):1933–1954, 2014.

[66] Frank R Kschischang, Brendan J Frey, and H-A Loeliger. Factor graphs and the sum-product algorithm. *IEEE Transactions on information theory*, 47(2):498–519, 2001.

[67] J Nicholas Laneman, David NC Tse, and Gregory W Wornell. Cooperative diversity in wireless networks: Efficient protocols and outage behavior. *IEEE Transactions on Information theory*, 50(12):3062–3080, 2004.

[68] Mai TP Le, Guido Carlo Ferrante, Giuseppe Caso, Luca De Nardis, and Maria-Gabriella Di Benedetto. On information-theoretic limits of code-domain noma for 5g. *IET Communications*, 12(15):1864–1871, 2018.

[69] S-YR Li, Raymond W Yeung, and Ning Cai. Linear network coding. *IEEE transactions on information theory*, 49(2):371–381, 2003.

[70] Yixin Li, Fu-Chun Zheng, and Michael Fitch. Physical layer network coding with channel and delay estimation. *IET Communications*, 7(11):1109–1116, 2013.

[71] Yixin Li, Fu-Chun Zheng, and Yanxiang Jiang. Inter-symbol-interference cancelation in time-domain physical-layer network coding with fractional delay. In *2014 IEEE 25th Annual International Symposium on Personal, Indoor, and Mobile Radio Communication (PIMRC)*, pages 232–236. IEEE, 2014.

[72] Christos Liaskos, Shuai Nie, Ageliki Tsioliaridou, Andreas Pitsillides, Sotiris Ioannidis, and Ian Akyildiz. A new wireless communication paradigm through software-controlled metasurfaces. *IEEE Communications Magazine*, 56(9):162–169, 2018.

[73] Soung Chang Liew, Shengli Zhang, and Lu Lu. Physical-layer network coding: Tutorial, survey, and beyond. *Physical Communication*, 6:4–42, 2013.

[74] Jianquan Liu, Meixia Tao, and Youyun Xu. Pseudo exclusive-or for ldpc coded two-way relay block fading channels. In *2011 IEEE International Conference on Communications (ICC)*, pages 1–5. IEEE, 2011.

[75] Yuanwei Liu, Zhijin Qin, Maged Elkashlan, Zhiguo Ding, Arumugam Nallanathan, and Lajos Hanzo. Non-orthogonal multiple access for 5g and beyond. *Proceedings of the IEEE*, 105(12):2347–2381, 2017.

[76] Lu Lu, Geoffrey Ye Li, A Lee Swindlehurst, Alexei Ashikhmin, and Rui Zhang. An overview of massive mimo: Benefits and challenges. *IEEE journal of selected topics in signal processing*, 8(5):742–758, 2014.

[77] Lu Lu and Soung Chang Liew. Asynchronous physical-layer network coding. *IEEE Transactions on Wireless Communications*, 11(2):819–831, 2011.

[78] John MacLellan, Steve Lam, and Xiaoyang Lee. Residential indoor rf channel characterization. In *IEEE 43rd Vehicular Technology Conference*, pages 210–213. IEEE, 1993.

[79] Praveen Kumar Malik, Rohit Sharma, Rajesh Singh, Anita Gehlot, Suresh Chandra Satapathy, Waleed S Alnumay, Danilo Pelusi, Uttam Ghosh, and Janmenjoy Nayak. Industrial internet of things and its applications in industry 4.0: State of the art. *Computer Communications*, 166:125–139, 2021.

[80] Yijie Mao, Bruno Clerckx, and Victor OK Li. Rate-splitting multiple access for downlink communication systems: bridging, generalizing, and outperforming sdma and noma. *EURASIP journal on wireless communications and networking*, 2018:1–54, 2018.

[81] E Martini and S Maci. Theory, analysis, and design of metasurfaces for smart radio environments. *Proceedings of the IEEE*, 110(9):1227–1243, 2022.

[82] Muriel Médard and Alex Sprintson. *Network coding: Fundamentals and applications*. Academic Press, 2011.

[83] Raed Y Mesleh, Harald Haas, Sinan Sinanovic, Chang Wook Ahn, and Sangboh Yun. Spatial modulation. *IEEE Transactions on vehicular technology*, 57(4):2228–2241, 2008.

[84] Jan Mietzner, Robert Schober, Lutz Lampe, Wolfgang H Gerstacker, and Peter A Hoeher. Multiple-antenna techniques for wireless communications-a comprehensive literature survey. *IEEE communications surveys & tutorials*, 11(2):87–105, 2009.

[85] Scott L Miller. Detection of imperfectly synchronized data streams in physical layer network coding. In *2014 IEEE Global Communications Conference*, pages 1577–1582. IEEE, 2014.

[86] Hossein Najafi, Mohamed Oussama Damen, and Are Hjorungnes. Asynchronous compute-and-forward. *IEEE transactions on communications*, 61(7):2704–2712, 2013.

[87] Bobak Nazer and Michael Gastpar. Compute-and-forward: Harnessing interference through structured codes. *IEEE Transactions on Information Theory*, 57(10):6463–6486, 2011.

[88] Jin-Ping Niu and Geoffrey Ye Li. An overview on backscatter communications. *Journal of Communications and Information Networks*, 4(2):1–14, 2019.

[89] Philipp Osterrieder, Lukas Budde, and Thomas Friedli. The smart factory as a key construct of industry 4.0: A systematic literature review. *International Journal of Production Economics*, 221:107476, 2020.

[90] Parth H Pathak, Xiaotao Feng, Pengfei Hu, and Prasant Mohapatra. Visible light communication, networking, and sensing: A survey, potential and challenges. *IEEE communications surveys & tutorials*, 17(4):2047–2077, 2015.

[91] P. Popovski and H. Yomo. Bi-directional amplification of throughput in a wireless multi-hop network. In *2006 IEEE 63rd Vehicular Technology Conference*, volume 2, pages 588–593, 2006.

[92] Petar Popovski and Hiroyuki Yomo. The anti-packets can increase the achievable throughput of a wireless multi-hop network. In *2006 IEEE international conference on communications*, volume 9, pages 3885–3890. IEEE, 2006.

[93] Petar Popovski and Hiroyuki Yomo. Physical network coding in two-way wireless relay channels. In *2007 IEEE international conference on communications*, pages 707–712. IEEE, 2007.

[94] Vasanthan Raghavan, Andrzej Partyka, Lida Akhoondzadeh-Asl, Mohammad Ali Tassoudji, Ozge Hizir Koymen, and John Sanelli. Millimeter wave channel measurements and implications for phy layer design. *IEEE Transactions on Antennas and Propagation*, 65(12):6521–6533, 2017.

[95] Boris Rankov and Armin Wittneben. Achievable rate regions for the two-way relay channel. In *2006 IEEE international symposium on information theory*, pages 1668–1672. IEEE, 2006.

[96] Ashutosh Sabharwal, Philip Schniter, Dongning Guo, Daniel W Bliss, Sampath Rangarajan, and Risto Wichman. In-band full-duplex wireless: Challenges and opportunities. *IEEE Journal on selected areas in communications*, 32(9):1637–1652, 2014.

[97] Peter Sanders, Sebastian Egner, and Ludo Tolhuizen. Polynomial time algorithms for network information flow. In *Proceedings of the fifteenth annual ACM symposium on Parallel algorithms and architectures*, pages 286–294, 2003.

[98] Michele Sanna and Ebroul Izquierdo. A survey of linear network coding and network error correction code constructions and algorithms. *International Journal of Digital Multimedia Broadcasting*, 2011, 2011.

[99] Tapan K Sarkar, Robert Mailloux, Arthur A Oliner, Magdalena Salazar-Palma, and Dipak L Sengupta. *History of wireless*. John Wiley & Sons, 2006.

[100] Peter J Schreier and Louis L Scharf. *Statistical signal processing of complex-valued data: the theory of improper and noncircular signals*. Cambridge university press, 2010.

[101] Claude E Shannon. A mathematical theory of communication. *The Bell system technical journal*, 27(3):379–423, 1948.

[102] Osvaldo Simeone, Umberto Spagnolini, Yeheskel Bar-Ness, and Steven H Strogatz. Distributed synchronization in wireless networks. *IEEE Signal Processing Magazine*, 25(5):81–97, 2008.

[103] Lingyang Song, Yonghui Li, Anpeng Huang, Bingli Jiao, and Athanasios V Vasilakos. Differential modulation for bidirectional relaying with analog network coding. *IEEE Transactions on Signal Processing*, August 2010.

[104] Jesper H Sørensen, Rasmus Krigslund, Petar Popovski, Toshiaki Koike Akino, and Torben Larsen. Physical layer network coding for fsk systems. *IEEE Communications Letters*, 13(8):597–599, 2009.

[105] Ludek Subrt and Pavel Pechac. Controlling propagation environments using intelligent walls. In *2012 6th European Conference on Antennas and Propagation (EUCAP)*, pages 1–5. IEEE, 2012.

[106] Abbas Shah Syed, Daniel Sierra-Sosa, Anup Kumar, and Adel Elmaghraby. Iot in smart cities: A survey of technologies, practices and challenges. *Smart Cities*, 4(2):429–475, 2021.

[107] Jan Sykora. Hierarchical pairwise error probability for hierarchical decode and forward strategy in plnc. *IEEE Communications Letters*, 20(1):49–52, 2016.

[108] Jan Sykora. Hierarchical data decision aided 2-source bpsk h-mac channel phase estimator with feed-back gradient solver for wpnc networks. In *2018 14th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, pages 89–96. IEEE, 2018.

[109] Jan Sykora and Alister Burr. Network coded modulation with partial side-information and hierarchical decode and forward relay sharing in multi-source wireless network. In *2010 European Wireless Conference (EW)*, pages 639–645. IEEE, 2010.

[110] Jan Sykora and Alister Burr. Layered design of hierarchical exclusive codebook and its capacity regions for hdf strategy in parametric wireless 2-wrc. *IEEE Transactions on Vehicular Technology*, June 2011.

[111] Jan Sykora and Alister Burr. *Wireless Physical Layer Network Coding*. Cambridge University Press, January 2018.

[112] David Tse and Pramod Viswanath. *Fundamentals of wireless communication*. Cambridge university press, 2005.

[113] IIMT Union. Imt traffic estimates for the years 2020 to 2030. *Report ITU*, 2370, 2015.

[114] Tomas Uricar and Jan Sykora. Non-uniform 2-slot constellations for bidirectional relaying in fading channels. *IEEE Communications Letters*, 15(8):795–797, 2011.

[115] Tomas Uricar, Jan Sykora, Bin Qian, and Wai Ho Mow. Superposition coding for wireless butterfly network with partial network side-information. In *2013 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*, pages 149–154. IEEE, 2013.

[116] Matthew C Valenti, Don Torrieri, and Terry Ferrett. Noncoherent physical-layer network coding with fsk modulation: Relay receiver design issues. *IEEE Transactions on Communications*, July 2011.

[117] Dong Wang, Shengli Fu, and Kejie Lu. Channel coding design to support asynchronous physical layer network coding. In *GLOBECOM 2009-2009 IEEE Global Telecommunications Conference*, pages 1–6. IEEE, 2009.

[118] Ping-Chung Wang, Yu-Chih Huang, and Krishna R Narayanan. Asynchronous physical-layer network coding with quasi-cyclic codes. *IEEE Journal on Selected Areas in Communications*, 33(2):309–322, 2014.

[119] Makesh Pravin Wilson, Krishna Narayanan, Henry D Pfister, and Alex Sprintson. Joint physical layer coding and network coding for bidirectional relaying. *IEEE Transactions on Information Theory*, 56(11):5641–5654, 2010.

[120] Qingqing Wu and Rui Zhang. Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network. *IEEE communications magazine*, 58(1):106–112, 2019.

[121] Xiaofu Wu, Chunming Zhao, and Xiaohu You. Joint ldpc and physical-layer network coding for asynchronous bi-directional relaying. *IEEE Journal on Selected Areas in Communications*, 31(8):1446–1454, 2012.

[122] Ming Xiao, Shahid Mumtaz, Yongming Huang, Linglong Dai, Yonghui Li, Michail Matthaiou, George K Karagiannidis, Emil Björnson, Kai Yang, I Chih-Lin, et al. Millimeter wave communications for future mobile networks. *IEEE Journal on Selected Areas in Communications*, 35(9):1909–1935, 2017.

[123] Ping Yang, Yue Xiao, Ming Xiao, and Shaoqian Li. 6g wireless communications: Vision and potential techniques. *IEEE network*, 33(4):70–75, 2019.

[124] Keyvan Yasami, Abolfazl Razi, and Ali Abedi. Analysis of channel estimation error in physical layer network coding. *IEEE Communications Letters*, 15(10):1029–1031, 2011.

[125] Raymond W Yeung. Multilevel diversity coding with distortion. *IEEE Transactions on Information Theory*, 41(2):412–422, 1995.

[126] Raymond W Yeung. *Information theory and network coding*. Springer Science & Business Media, 2008.

[127] Raymond W Yeung. Network coding: A historical perspective. *Proceedings of the IEEE*, 99(3):366–371, 2011.

[128] Nanfang Yu, Patrice Genevet, Mikhail A Kats, Francesco Aieta, Jean-Philippe Tetienne, Federico Capasso, and Zeno Gaburro. Light propagation with phase discontinuities: generalized laws of reflection and refraction. *science*, 334(6054):333–337, 2011.

[129] Zhifeng Yuan, Guanghui Yu, Weimin Li, Yifei Yuan, Xinhui Wang, and Jun Xu. Multi-user shared access for internet of things. In *2016 IEEE 83rd Vehicular Technology Conference (VTC Spring)*, pages 1–5. IEEE, 2016.

[130] Ram Zamir. *Lattice Coding for Signals and Networks: A Structured Coding Approach to Quantization, Modulation, and Multiuser Information Theory*. Cambridge University Press, 2014.

[131] Shengli Zhang, Soung Chang Liew, and Patrick P Lam. Hot topic: Physical-layer network coding. In *Proceedings of the 12th annual international conference on Mobile computing and networking*, pages 358–365, 2006.

[132] Shengli Zhang, Soung-Chang Liew, and Patrick P Lam. On the synchronization of physical-layer network coding. In *2006 IEEE Information Theory Workshop-ITW'06 Chengdu*, pages 404–408. IEEE, 2006.

[133] Shengli Zhang, Soung Chang Liew, Hui Wang, and Xiaohui Lin. Capacity of two-way relay channel. In *International Conference on Access Networks*, pages 219–231. Springer, 2009.

[134] Wu Zhuo, Li Guangxiang, and Wang Tao. Two-way amplify-and-forward relaying with carrier offsets in the absence of csi: differential modulation-based schemes. *EURASIP Journal on Wireless Communications and Networking*, July 2014.