Channel Coding in Molecular Communication

Dissertation

zur Erlangung des akademischen Grades
Doktor der Ingenieurswissenschaften
(Dr.-Ing.)
der Technischen Fakultät
der Christian-Albrechts-Universität zu Kiel

vorgelegt von
Martin Damrath

Kiel, 2020

Tag der Einreichung: 17.04.2020
Tag der Disputation: 21.10.2020

Berichterstatter: Prof. Dr.-Ing. Peter Adam Höher
Prof. Dr.-Ing. Robert Schober
Prof. Dr. H. Birkan Yilmaz

Preface

The work presented in this dissertation is the outcome of my research conducted at the Chair of Information and Coding Theory at the Christian-Albrechts-Universität zu Kiel.

My special thanks belong to my supervisor Prof. Dr.-Ing. Peter Adam Hoeher, who has always supported, motivated, and inspired me through fruitful discussions. Under his supervision I had the opportunity to research freely and to let ideas come to life. Furthermore, I would like to thank my two co-referees Prof. Dr.-Ing. Robert Schober and Prof. Dr. H. Birkan Yilmaz for their interest in my work, their constructive feedback, and their valuable time.

I would also like to thank the members of the Chair of Information and Coding Theory, who made my PhD a special time due to the great scientific but also pleasant atmosphere inside and outside of the university: Abdullah, Adrian, Andrej, Gilbert, Julius, Kevin, Maurice, Max, Meelis, Niklas, Nils, Rebecca, Sami, and Sunasheer. My further thanks go to my supervised former students for their interest in my research, their efforts, and their support.

Additionally, I would like to thank Prof. H. Birkan Yilmaz, Prof. Chan-Byoung Chae, and Sascha Krüger for their fruitful collaborative research. Special thanks belong to Dr.-Ing. Sebastian Toxværd, Dr. rer. nat. Robert Felix Reischke, Sami Alkubti Almasri, and Sunasheer Bhattacharjee for conscientious proofreading of my work and their constructive comments.

My sincere thanks go to my parents Claudia and Norbert, who have always supported me without pressure and made my studies possible. Finally, I would like to thank my wife Yessi for her patience and her unconditional loving support during the final phase of my PhD. I am very much looking forward to our daughter, who will perhaps read these lines one day.

Kiel, October 2020 — Martin Damrath

Abstract

This dissertation establishes and analyzes a complete molecular transmission system from a communication engineering perspective. Its focus is on diffusion-based molecular communication in an unbounded three-dimensional fluid medium. As a basis for the investigation of transmission algorithms, an equivalent discrete-time channel model (EDTCM) is developed and the characterization of the channel is described by an analytical derivation, a random walk based simulation, a trained artificial neural network (ANN), and a proof of concept testbed setup. The investigated transmission algorithms cover modulation schemes at the transmitter side, as well as channel equalizers and detectors at the receiver side. In addition to the evaluation of state-of-the-art techniques and the introduction of orthogonal frequency-division multiplexing (OFDM), the novel variable concentration shift keying (VCSK) modulation adapted to the diffusion-based transmission channel, the low-complex adaptive threshold detector (ATD) working without explicit channel knowledge, the low-complex soft-output piecewise linear detector (PLD), and the optimal *a posteriori* probability (APP) detector are of particular importance and treated. To improve the error-prone information transmission, block codes, convolutional codes, line codes, spreading codes and spatial codes are investigated. The analysis is carried out under various approaches of normalization and gains or losses compared to the uncoded transmission are highlighted. In addition to state-of-the-art forward error correction (FEC) codes, novel line codes adapted to the error statistics of the diffusion-based channel are proposed. Moreover, the turbo principle is introduced into the field of molecular communication, where extrinsic information is exchanged iteratively between detector and decoder. By means of an extrinsic information transfer (EXIT) chart analysis, the potential of the iterative processing is shown and the communication channel capacity is computed, which represents the theoretical performance limit for the system under investigation. In addition, the construction of an irregular convolutional code (IRCC) using the EXIT chart is presented and its performance capability is demonstrated. For the evaluation of all considered transmission algorithms the bit error rate (BER) performance is chosen. The BER is determined by means of Monte Carlo simulations and for some algorithms by theoretical derivation.

Kurzfassung

In dieser Dissertation wird ein komplettes molekulares Übertragungssystem aus nachrichtentechnischer Sicht geschaffen und analysiert. Der Schwerpunkt liegt dabei auf einer diffusionsbasierten molekularen Kommunikation in einem unendlich ausgedehnten dreidimensionalen fluiden Medium. Als Grundlage zur Untersuchung von Übertragungsalgorithmen wird ein äquivalentes zeitdiskretes Kanalmodell des Systems entwickelt. Zur Charakterisierung des Kanals werden vier Methoden mit unterschiedlichen Ansätzen beschrieben. Diese basieren auf einer analytischen Herleitung, auf Random-Walk-basierten Simulationen, auf einem trainierten neuronalen Netzwerk und auf einem Demonstratoraufbau. Die untersuchten Übertragungsalgorithmen umfassen auf der Sendeseite Modulationsverfahren, während auf der Empfangsseite Kanalentzerrer und Detektoren betrachtet werden. Neben der Evaluation von aus der Literatur gängigen Techniken und der Einführung von *orthogonal frequency-division multiplexing* (OFDM) sind dabei vor allem die neuartige, an den Übertragungskanal angepasste variable Konzentrationsmodulation, der ohne Kanalkenntnis arbeitende aufwandsgünstige adaptive Schrankendetektor, der aufwandsgünstige abschnittsweise linear approximierende Softdetektor, sowie der optimale *a posteriori probability* (APP) Detektor hervorzuheben. Um die Fehlerwahrscheinlichkeit einer Informationsübertragung zu senken werden Blockcodes, Faltungscodes, Leitungscodes, Spreizcodes und räumliche Codierungen untersucht. Die Codierverfahren werden unter verschiedenen Normalisierungsansätzen analysiert. Dabei werden Gewinne und Verluste gegenüber einer uncodierten Übertragung herausgearbeitet. Neben aus der Literatur gängigen vorwärtsfehlerkorrigierenden Kanalcodes werden auch neuartige, an die Fehlerstatistik des diffusionsbasierten Kanals angepasste Leitungscodes vorgestellt. Darüber hinaus wird das Turbo-Prinzip in den Bereich der molekularen Kommunikation eingeführt, bei dem extrinsische Information iterativ zwischen Detektor und Decoder ausgetauscht wird. Mit Hilfe einer *extrinsic information transfer* (EXIT)-Chart-Analyse wird das Potential der iterativen Verarbeitung aufgezeigt und die Kanalkapazität des Kommunikationssystems berechnet, welche die theoretische Leistungsgrenze für das untersuchte System darstellt. Des Weiteren wird das Design eines irregulären Faltungscodes mit Hilfe der EXIT-Chart

vorgestellt und dessen Leistungsfähigkeit demonstriert. Für die Bewertung aller betrachteten Übertragungsalgorithmen wird die Bitfehlerrate herangezogen, die mit Hilfe von Monte-Carlo-Simulationen ermittelt wird. Zusätzlich wird die Bitfehlerrate für ausgewählte Algorithmen theoretisch hergeleitet.

Contents

1

Introduction

1.1 Overview on Molecular Communication

Arguably, the oldest naturally existing form of communication between two organisms is based on the exchange of molecules. It is used even by the simplest single-cell organisms to exchange information in the form of deoxyribonucleic acid (DNA) [vWW$^+$16]. This type of communication has been developed and optimized by nature over billions of years and is still used today in numerous forms within and between several organisms. In quorum sensing, bacteria release autoinductors to estimate their population density in the environment. If the density exceeds a certain threshold, specific genes of the bacteria are activated [Bas99]. In calcium signaling, calcium is used intracellularly to regulate cell activities in epithelial cells or intercellularly to transmit stimuli [Car02; NS$^+$05]. Neurotransmitters diffuse in chemical synapses through the synaptic cleft from a presynaptic cell to a postsynaptic cell in order to transfer action potentials between two neurons [NE$^+$13]. In fauna, pheromones are used for information transfer [BW63]. Ants, for example, use chemical signals for communication [HI$^+$15]. Even in flora, molecules are used for communication purposes. For instance, plants release volatiles in response to herbivore attacks to attract natural enemies and warn neighbouring plants [BK$^+$02].

In the research area of molecular communication [AB$^+$08; NE$^+$13; FY$^+$16], this naturally occurring form of communication is exploited or mimicked in order to realize information transmission. While electromagnetic waves (including micro-waves and light) and sound waves are used to transmit information in radio-based, optical, inductive, and acoustic communication, molecules are used as information carriers in molecular communication. Nevertheless, the transmission can be divided into the three main components: transmitter,

channel, and receiver. Molecular communication has been of great interest in recent times due to its strengths, which offers an advantage over classical transmissions in special application scenarios. Since it already occurs naturally in living organisms, biocompatible systems can be easily realized. In addition, the transmission is radiation-free. Molecular communication also enables transmission on very small scales where radio-based communication systems encounter problems due to the scaling of antenna size proportional to wavelength. Furthermore, the power supply in smallest scales is problematic. Molecular communication, instead, offers a very energy-efficient way of transmission. For example, if the molecules diffuse from the transmitter to the receiver, they use the environmental energy to move. Thus energy is needed only for the release, but no extra energy for the propagation of the molecules. Still, molecules propagate over long distances even in environments that are challenging for radio waves or light. This can be, for example, in water, in pipe and shaft systems or in heaps of rubble. But molecular communication also has disadvantages. Since the molecules have to be physically transported from the transmitter to the receiver, the transmission is slow, especially over long distances. If the molecules propagate by diffusion, this leads to slowly decreasing channel impulse responses and consequently to intersymbol interference (ISI). Furthermore, it is a stochastic propagation process which leads to an unreliable transmission. If there is no favorable drift, molecular communications systems are also limited in transmission range.

Many multi-scale application scenarios of molecular communication are conceivable. The main motivation for the initial molecular communication research is based on its significance as a potential key technology for the realization of a communication link between so-called nanomachines. Those nanomachines are defined as devices whose components scale in the nanometer range [AB+08]. They can, for example, be imagined as genetically modified bacteria or as synthetic cells engineered to perform a specific task [AF+12]. Due to their small dimensions, however, they are strongly limited in their energy and complexity. Consequently, they are only able to perform simple tasks [XY+03; SM+05]. In order to be able to handle even more complex tasks and to enable the autonomous use of nanomachines, a large number of machines need to work together as a swarm. For this purpose, they need a particularly energy-efficient way of communicating with each other, which could be provided by molecular communication. So far, nanomachines are more a theoretical concept, but partial functions have already been demonstrated [GA+13; VP+11]. One main application for these nanomachines is targeted drug delivery [CM+17]. As an example, nanomachines can be injected into the human body in the case of a tumor disease. Since tumor cells, in contrast to healthy cells, release a certain messenger substance, nanomachines can identify

the messenger substance and migrate towards the tumor cell. The nanomachines could then release the required amount of drugs locally at the tumor cell. This would considerably improve the effectiveness of pharmacology and significantly reduce potential adverse drug effects. In this example, the communication of the nanomachines is essential to locate the tumor cell, navigate to it and coordinate the drug release. Further micro-scale applications for molecular communication include, among others, lab-on-a-chip systems [Par10], tissue engineering [GN02], or monitoring and quality control [AB+08].

In addition to the mostly medically motivated microscopic applications, molecular communication can also find applications in the macroscopic field. There, it is often proposed as an alternative method in environments where wireless communication using conventional methods such as radio waves or light is not possible or challenging. For example, molecular communication enables communication in pipe and duct systems in the oil and gas industry [FY+16], in infrastructure monitoring applications [QG+14], for robots in underground tunnels, or for extracting embedded sensor data in cavities or machinery [GM+15]. The use of molecular communication for industrial applications is currently being investigated by the the German Federal Ministry of Education under the project entitled "Makroskopische Molekulare Kommunikation (MAMOKO)" [MAMOKO].

In general, molecular communication can be classified as active or passive depending on the way the molecules propagate. In active molecular communication, the information carriers are actively transported from the transmitter to the receiver. This can be achieved, for example, by bacterial motors carrying the information encoded in their DNA [CA10; GA10]. Alternatively, motor proteins can be used, which can travel on molecular rails [ME+06; HM+10; EM+11]. In passive molecular communication, information carriers propagate passively via diffusion through a fluid medium, thereby exploiting the energy available in the environment [Cus09].

1.2 Motivation and Contributions

Molecular communication is an interdisciplinary field of research that connects engineers, biologists, chemists, and physicians. As it is a comparably young research area, many fundamental questions remain unsolved. In this dissertation, molecular communication is examined from a communication engineering perspective. Unsolved issues in this area include a general information theoretical channel capacity expression and the related maximum rate of reliable transmissions [NE+13]. Furthermore, design and modulation of the three main components transmitter, channel, and receiver are under discussion. In the

context of this dissertation, a passive molecular communication scenario is considered, in which the information carrying molecules propagate only by free diffusion. This type is often referred to as diffusion-based molecular communication in the literature [PA11]. Compared to the classical wireless channel, the diffusion-based channel brings new interesting challenges to a transmission system. While, for example, radio waves propagate deterministically, the diffusion of molecules is described by stochastic propagation. This typically leads to slowly decreasing channel impulse responses and consequently a strong ISI effect. In addition, the random propagation causes an amplitude-dependent noise resulting in unreliable information transmission [PA11; SM$^+$14]. Transmitter and receiver should be adapted to the characteristics of the channel. In case of micro-scale applications, realistic transmitter and receiver algorithms are limited in their complexity by the low-energy assumption.

The diffusion-based channel with its new challenges is the main motivation for the contents covered in this dissertation. All three main components of a transmission system, from the transmitter over the channel to the receiver, are considered to give a complete overview of a communication system. The main goal is to identify transmission algorithms to provide the most reliable transmission of information. An equivalent discrete-time channel model (EDTCM) is proposed as a basis for the analysis and development of transmitter and receiver algorithms. Furthermore, the determination of the necessary channel coefficients is presented. In order to counteract the negative influence of the diffusion-based channel, appropriate transmission algorithms are analyzed and developed. On the transmitter side, the capabilities of existing modulation schemes are described and, in addition, a novel modulation scheme adapted to the channel is introduced. On the receiver side, the effect of channel equalization and detection is investigated. Besides the optimal but computational complex *a posteriori* probability (APP) detector, a low-complex biologically motivated detector and a low-complex soft-output detector are proposed. Furthermore, an extensive analysis of forward error correction (FEC) channel codes, line codes, spreading codes, and spatial codes is provided and their potential to enhance transmission reliability is evaluated. Novel line codes are presented which are adapted to the characteristics of the transmission channel. Moreover, modulation/detection and channel coding in the sense of the turbo principle are concatenated in order to further increase the reliability of the transmission by means of iterative processing. With the help of an extrinsic information transfer (EXIT) chart analysis, channel codes are designed that are matched to the channel. As theoretical performance limit the communication channel capacity and the influence of different system parameters on it are investigated. A macroscopic testbed design is presented and evaluated

as a proof of concept.

1.3 State-of-the-Art

The following provides an overview of the state-of-the-art in molecular communication scientific literature. The focus is on topics which are also covered in this dissertation. In detail, channel modeling, modulation, detection, ISI mitigation, channel coding, and multiple-input multiple-output (MIMO) are discussed.

Channel Modeling Depending on different system assumptions, the molecular communication literature contains numerous contributions dealing with channel modeling. As assumed in this dissertation, the focus in the following is on systems with absorbing receivers. For a more general comprehensive tutorial the reader is referred to [JA$^+$19]. The first analytical channel model was presented in [SE$^+$12]. It is a one-dimensional environment with a perfectly absorbing receiver. The propagation of the molecules is described by free diffusion, i. e. there is no interaction between the diffusing molecules. This channel model was extended in [KA$^+$12] by considering a constant drift. By solving the advection-diffusion equation, the hitting probability function of the molecules on an absorbing wall is derived. In [CT$^+$15], a one-dimensional environment is considered where the molecules propagate by anomalous diffusion which takes interactions and collisions between molecules into account. The three-dimensional case with a point transmitter and a perfectly absorbing spherical receiver is investigated in [YH$^+$14]. Assuming free diffusion, the first hitting probability function of the molecules on a spherical receiver is derived. This channel model is enhanced in [AY$^+$15] by considering receiver-side receptor effects. If the system assumptions change, for example due to the introduction of multiple absorbing spherical receiver or arbitrary drift, a closed analytical solution of the diffusion equation has often not been derived so far. As an alternative to computationally intensive Monte Carlo based random walk simulations, [LY$^+$17] suggests the use of an artificial neural network (ANN) to estimate the hitting probability function. In addition to the theoretical channel models, the literature also contains a few practical implementations of molecular communication systems, which often serve as proof of concept. The first microscopic experimental molecular communication system is presented in [HT$^+$08]. The authors use kinesin-driven microtubules to autonomously load and unload cargoes. In [OE12], cell to cell communication is considered where messages are exchanged in the form of DNA. Another demonstration of the key functionalities of molecular communication is given by wet laboratory experiments in [FS$^+$18], which is an

experimental testbed based on artificial cell systems. For message propagation, molecule replication is exploited by the cells. Another example of biological molecular communication testbeds includes [FF$^+$14; KA$^+$13; GK$^+$19]. In [FF$^+$14] the communication between manually stimulated platelets and endothelium is investigated experimentally. The experimental setups in [KA$^+$13; GK$^+$19] are based on *Escherichia coli* bacteria. Due to their fluorescent response to the injected stimulating molecules, they are used as receivers in [KA$^+$13]. In [GK$^+$19], however, they are used as transmitters where they release protons into their environment by light stimulation. The concentration of protons will finally be measured by a pH electrode. Besides the costly and implementation challenging microscopic testbeds, there are also a few macroscopic testbeds. The first macroscopic testbed is presented in [FG$^+$13]. It is based on an electric sprayer on the transmitter side, which sprays isopropyl alcohol. On the receiver side a metal-oxide sensor is used, which measures the alcohol concentration in its environment. With this setup, a simple data transmission link is realized. One of the biggest challenges of the setup is the nonlinearity, which is modeled in [FK$^+$14] by Gaussian noise. In [KF$^+$15], an end to end model is described, which takes the nonlinearity of the sensor into account. By introducing a second sprayer and metal-oxide sensor, the testbed is extended to MIMO in [KL$^+$16]. It is shown that the transmission rate can be further increased. The setup presented in [KS$^+$18] is also based on isopropyl alcohol, but considers a tube environment and a photon ionization gas detector. An alternative macroscopic testbed based on pH values is presented in [FP$^+$17]. Acid and base pulses are used to transmit binary data. It is shown that the new setup outperforms the old alcohol-based testbed. A testbed for in-vessel communication based on magnetic nanoparticles is introduced in [UK$^+$18]. On the transmitter side, the nanoparticles are pumped into a liquid stream. On the receiver side, the nanoparticles are detected using a susceptometer. In [SW$^+$18; WS$^+$18; SW$^+$19; WA$^+$19], theoretical channel models are developed for this setup. In addition, there are several other synthetic macroscopic molecular communication testbeds. Among others, these are based on the detection of odors in a gas stream network by mass spectroscopy [GM$^+$17], camera-detected microfluidic droplets [DD$^+$13] or fluorescent substances in liquid mediums, which are tracked either by cameras [AE$^+$18] or fluorescence detectors [TL$^+$18] at the receiver side.

Modulation The most popular modulation scheme in the literature is on-off keying (OOK). It is a binary modulation scheme in which a data bit is represented either by no release or by the release of a certain concentration of molecules at the transmitter. The first time it is considered in [MM$^+$10] and compared with a frequency modulation of the

transmission rate. In [KY+11], the release of the molecules is extended by several concentration levels to enable higher-order modulation schemes. Analogous to amplitude shift keying (ASK) in wireless communication, this modulation scheme is termed concentration shift keying (CSK). As an alternative to CSK, molecule shift keying (MoSK) in [KY+11] considers modulation by the type of released molecules. Consequently, each symbol is represented by the release of a certain type of molecule. It is shown that MoSK outperforms CSK and that it is more robust against ISI. The use of different molecule types is utilized in various modulation schemes. In [KC13], two new higher-order modulation schemes based on isomers are introduced. The proposed isomer CSK utilizes different isomers which are all CSK modulated. It is therefore a combination of CSK and MoSK and forms the basis for molecule concentration shift keying (MCSK). In isomer ratio shift keying, the information is represented by the concentration ratios between released isomers. In [AG+12], the information is represented by the order of the released molecule types. This leads to a modulation scheme which works without synchronization between transmitter and receiver. Often a second type of molecules is used to provide more robustness against ISI. In [AG+13], CSK is considered while changing the type of molecules in successive transmission intervals. This leads to a reduction of ISI. Another approach is considered in [TP+14] by adding a second type of molecules to OOK, which marks bit changes and thus reduces ISI at error-prone bit transitions. In [PS+14], zebra CSK is introduced. Similar to [AG+13], CSK with alternating molecule type is considered in successive transmission intervals. In addition to the information-carrying molecules, inhibitor molecules of alternated molecule type are released to reduce ISI. This leads to an increase in capacity and a decrease in the probability of symbol error. In addition to the concentration and type of molecules, information can also be modulated into the temporal dimension. To this aim, three different modulation schemes are proposed in [GL+11]. In pulse-position modulation (PPM), the information is represented by the time of release of a pulse during a transmission interval. In the sense of rate modulation, the information is represented by the release rate of the molecules within a transmission interval. In communication through silence (also known as differential PPM), the information is contained in the time interval between two consecutive pulses. The same principle was further developed in [KA+13] as smart time-elapsed communication by adding error differentiation and differential coding. A further degree of freedom for modulation can be created by several transmitters. Hence [GB+19a] presents molecular space shift keying, in which the information is modulated into the antenna index used for transmission. This idea can also be extended by including different types of molecules resulting in molecular spatial modulation. In [GB+19b], index

modulation is linked to PPM, which provides robustness against interlink interference (ILI). A combination of index modulation and CSK is discussed in [HW+19].

Detection The detection approaches in molecular communication can be divided into two classes: sampling-based/amplitude detection and strength/energy detection [MM+10; LC+13]. In the case of sampling-based/amplitude detection, the instantaneous number of molecules is measured at the receiver at a certain point of time. While for strength/energy detection, on the other hand, the number of molecules reaching the receiver is accumulated over a certain period of time. A comparison of the error performance of both classes is given in [AA15]. Probably the most common and one of the simplest detection algorithms in molecular communication literature is the fixed threshold detector (FTD). The idea is to compare the number of received molecules with a single or several fixed thresholds and thus to decide on the transmitted symbol. For example, FTD is considered in [MM+11]. In addition, a blind-reference threshold detector is proposed which adapts its threshold to the number of previously received molecules and is therefore also referred to as adaptive threshold detector (ATD). Nevertheless, a detailed performance analysis is missing in [MM+11]. The literature presents many approaches based on an adaptive threshold that make explicit channel knowledge unnecessary. In [HM+16], an ATD is presented that gradually adapts its threshold. While in [AA+17] the threshold is adapted on the basis of the previous receiver observation, in [SJ18], two previous symbol durations with optimal weighting are considered. Another alternative is the increase detection algorithm [ZL+18], in which the increase or non-rise of the molecule concentration is used as the basis for the symbol decision. To improve threshold detection, [NC+14b] presents the weighted sum detector with equal weights or weights adapted to the channel impulse response. This weighted sum detector is extended in [NC+14c] by a decision-feedback equalizer (DFE), which further increases its performance. Analogously, in [JA+17] a matched filter detector is proposed which maximizes the expected signal-to-interference-plus-noise ratio and exceeds the weighted sum detector. In addition to the simple threshold detectors, there are also more complex approaches. In [LL+12], the detection threshold is related to log-likelihood ratios (LLRs) and DFE is performed. A widely studied approach, for example in [MY+12b; MM+12], is based on an algorithm that generates test statistics from the receiver observations and compares them with a threshold in the sense of LLRs. This detector is stated as optimal. An optimal detector for MoSK and concentration encoded information is discussed in [SM+12]. In [KA13], maximum-likelihood sequence estimation (MLSE) is introduced for the first time into the field of molecular communication. Furthermore,

sequence detection based on the maximum *a posteriori* and maximum-likelihood criteria is presented. According to the reduced-state sequence estimation (RSSE) principle, the complexity of MLSE in [NC+14b; MY+14] is reduced by shortening the channel memory length considered in the Viterbi algorithm, while the non-considered channel memory is fed back with the detection results. A non-coherent detection approach exploiting the local convexity of the channel impulse response is presented in [LS+16]. In [JF+18], a detector is presented which is based on the statistical channel state information instead of the instantaneous channel state information. Two other notable detectors, operating without synchronization between transmitter and receiver, are introduced in [NE17].

Intersymbol Interference Mitigation One of the main challenges in molecular communication is the ISI which results from the slowly decreasing channel impulse responses of the diffusion-based channel. Many approaches for ISI mitigation are discussed in the molecular communication community. In classical wireless communication, equalizers are usually used to revert the ISI effect. In [KA13], different equalizers for molecular communication are investigated. Linear minimum mean square error equalization, nonlinear DFE, and sequence detection based on maximum *a posteriori* and maximum-likelihood criteria are analyzed. In addition to classical receiver-side equalization algorithms, other approaches are discussed as well. In [KE+14], the symbol intervals are optimized in order to reduce the ISI. That ISI can be reduced by shifting the receiver absorption intervals is demonstrated in [AP+18]. A simple threshold detector and a transmitter that adapts its transmission rate to ISI is used in [MS+16]. In [TP+15], a special signaling strategy based on two different molecule types is proposed. The idea is to transmit a delayed version of the transmit signal with the second type of molecule so that the ISI can be estimated at the receiver and subtracted from the received signal. In [YC+18], it is shown that the peak time of the channel impulse response derivative is less than the peak time of the channel impulse response itself. In addition, the tail vanishes faster. This is exploited in a derivative-based detector. A biologically inspired approach for ISI mitigation is given in [NC+14a]. In this approach, enzymes are used to reduce the remaining interfering molecules in the environment. Furthermore, there are specially proposed channel codes that are robust against ISI, which will be described in the next paragraph.

Channel Coding In order to counteract the effect of unreliable communication, many different channel codes are studied in the molecular communication literature. In general, a distinction can be made between codes adapted from classical wireless communication and

channel codes newly designed for molecular communication. Surprisingly, channel codes in connection with molecular communication are often used to counter the ISI effect, which in classical wireless communication is the task of the equalizer. In [LH12], Hamming codes are used for forward error correction and it is shown that a coding gain can be achieved. In addition, the critical distance at which the energy gain of the channel code exceeds the operational energy costs is analyzed. In [MM+13], convolutional codes are used to demonstrate that the transmission distance can be increased by applying channel codes. Minimum energy codes are codes that keep their desired code distance while minimizing the required energy. They are considered in [BL+14; Bai16] and compared to Hamming codes. It is shown that they outperform Hamming codes by their lower power consumption in bit error rate (BER) performance, although requiring longer code words. In addition, in [Bai16] Luby transform codes are analyzed, which exceed Hamming codes in scenarios with few molecules. In [LH+15b], it is demonstrated that self-orthogonal convolutional codes have a higher coding gain and a lower critical distance than Hamming codes. Further work in [LH+15a] analyzes the operation region of Hamming codes, Euclidean geometry low-density parity check (LDPC) codes and cyclic Reed-Muller codes. Reed-Solomon codes are investigated in [DD+17] and also compared to Hamming codes. It is indicated that Reed-Solomon codes have the potential to improve the BER performance. In [MP+18], molecular single parity check (SPC) codes are considered. The main focus is on their implementation in the biochemical domain for genetically engineered bacteria. An automatic repeat request system based on cyclic redundancy check (CRC) codes for bacterial quorum communications is addressed in [BL+15; BL+16b; BL+16a; Bai16]. With ISI-free codes [SL+12; YC+12; SL+13], novel codes for molecular communication are proposed. They have the property of being robust against several molecule crossovers. A crossover refers to the effect that molecules emitted later can reach the receiver earlier than previously emitted molecules. In [SL+12], it is shown that repetition codes and ISI-free codes outperform convolutional codes in the sense of a BER-complexity trade-off. Furthermore, it is proposed to normalize the information throughput for channel codes with different code rates for comparability. The ISI-free codes are modified in [AK+16; KJ+19] by introducing a time gap between consecutive code words. In addition, the possibility of variable codeword lengths is considered. It is demonstrated that the introduced modification leads to an increase in performance compared to the original ISI-free codes.

Multiple-Input Multiple-Output In classical wireless communications, MIMO techniques already appertain to the state-of-the-art. In molecular communication, however, they are

just rarely investigated. The first conjunction between molecular communication and MIMO is given in [MY+12a]. The authors introduce transmitter diversity, receiver-side diversity combining, and spatial multiplexing to the area of diffusion-based molecular communication. While focusing on multi-user interference, little attention is paid to the effect of ISI throughout the work. In [KL+16], several detection algorithms are proposed for spatial multiplexing scenarios in diffusion-based molecular communication and it is demonstrated, by a proof of concept testbed implementation, that utilizing a molecular MIMO concept has potential to increase the data rate. In contrast to [MY+12a], the authors in [KL+16] take both, ISI and ILI, in their channel model into account. Furthermore, they extend their tabletop molecular single-input single-output (SISO) testbed to a MIMO testbed. The authors in [LH+16] expand a molecular communication broadcast system by a second absorbing receiver and study the effect on the BER and the channel capacity. However, the authors consider asymptotic behavior (i. e., time goes towards infinity) without solving the time dependent capture probabilities, which is an open problem in the literature. A machine learning based MIMO channel modeling is presented in [LY+17]. In [MC+17], the authors interpret each individual transmission paths of an information particle as a single-input multiple-output channel. With focus on one-shot communication over a molecular timing channel, they study the system diversity gain depending on the number of released particles. Training-based maximum-likelihood and least square channel estimators for diffusion-based MIMO channels are addressed in [RS17]. A MIMO system based on molecular motors is investigated in [ML+18].

1.4 Author's Contributions

Referring to Section 1.2, this section links the author's peer reviewed scientific contributions to the objectives.

Channel Modeling In [DK+17a], analogous to classical wireless communication, the EDTCM for the diffusion-based molecular communication channel is established and analyzed. The focus is on a system with a point transmitter and a spherical absorbing receiver in an unbound three-dimensional fluid medium without drift. However, the channel model is generally applicable to any diffusion-based system as long as the arrival probabilities of the molecules at the receiver are known. It is applied in [KD+17] to a system with several transmitters, in [DY+17; DY+18] to a system with several transmitters and absorbing receivers, and in [DH18; DS+19] to a medium with drift. The EDTCM provides a simplified

representation of the complex diffusion-based molecular communication channel. By its analogy to the classical EDTCM of an additive white Gaussian noise channel with memory, it also enables easier access for the adaptation of classical communication algorithms into the area of molecular communication. Furthermore, it offers a significant reduction in simulation time and an increase in simulation accuracy compared to random walk based particle simulations. In addition to the EDTCM based on Poisson binomial distributions, approximations of this model based on binomial distribution, Gaussian distribution and Poisson distribution are proposed. As analyzed in [DK+17a], depending on the system parameters, the approximations lead only to minor deviations from the more complex Poisson binomial based EDTCM.

Modulation and Detection In [BD+19], a novel single-carrier modulation scheme called variable concentration shift keying (VCSK) is introduced. Symbols are represented by different concentrations of released molecules. In contrast to CSK, which is known from literature, the concentration levels are not equidistant but have variable distances to each other. Thus, the concentration levels can be adapted to the stochastics of the amplitude-dependent noise of the diffusion-based molecular communication channel. As shown in [BD+19], VCSK outperforms classical CSK due to its adaptation to the properties of the channel. In [DK+17b], the multi-carrier modulation scheme orthogonal frequency-division multiplexing (OFDM) is introduced into the field of molecular communication for the first time. Analogous to optical communication, three different variants, namely direct current biased optical OFDM (DCO-OFDM), asymmetrically clipped optical OFDM (ACO-OFDM), and pulse amplitude modulation discrete multitone transmission (PAM-DMT), are presented and compared. A major advantage of OFDM is its comparatively low-complexity equalization. By adding a cyclic prefix, only a simple single-tap equalizer is required at the receiver side. With the help of BER simulations it is shown that the OFDM parameters should be chosen carefully since they have a big impact on the BER performance. The performance for the considered OFDM systems lies between a standard OOK transmission with FTD and MLSE. Furthermore, in [DK+17b] a simple bit-loading algorithm is proposed which can further improve the BER performance of the OFDM systems. To evaluate the performance of modulation schemes in iterative receiver systems, in [BD+19] different modulation schemes are compared using an EXIT chart analysis. The theoretical performance limit of modulation schemes is determined in [DH18; DS+19] using the example of OOK modulation. For this purpose, the communication channel capacity of the transmission system is determined exploiting the EXIT chart.

In [DH16], a low-complexity ATD is proposed and analyzed. Motivated by the effect of synaptic plasticity, this ATD, in contrast to an FTD, varies its threshold value based on the last received sample. Therefore, no channel knowledge is necessary. Its low complexity makes it particularly suitable for the use in nanomachines. In [DH16], the theoretical BER of the ATD is derived and compared with the performance of FTD, RSSE and MLSE in BER simulations. It is demonstrated that the ATD, despite its non-required channel knowledge and its low complexity, outperforms the FTD in environments with ISI. In [DK$^+$17a], the optimal APP detector based on the Bahl-Cocke-Jelinek-Raviv (BCJR) algorithm is introduced for the first time considering the introduced channel models. It is shown that it can outperform the MLSE algorithm based on an Euclidean distance metric in the sense of BER. For complexity limited systems, it can serve as a benchmark for other detection algorithms. In addition, the APP provides LLRs, which are used for iterative detection in [DH18; DS$^+$19] to analyze and further increase the performance of the system. Although originally designed for extremely higher-order modulation schemes, the idea of piecewise linear detector (PLD) [DH$^+$17; DH$^+$18] can also be applied to the area of molecular communication as demonstrated in this dissertation. The idea is based on the observation that LLRs plotted over all possible receiver samples can be approximated by several straight lines. Thus a low-complex detector can be realized, which is able to provide soft-outputs.

System Design with Multiple Transmitters/Receivers In [KD$^+$17], a system with an absorbing spherical receiver and four point transmitters is considered. To organize the reception of all transmitter signals, multiple access channel techniques are analyzed. As a main contribution the focus is on molecular code-division multiple access (CDMA), which is applied and analyzed in [KD$^+$17] for the first time in the area of molecular communication. For detection, a low-complexity adaptive threshold detector is proposed. In order to evaluate the performance of molecular CDMA, it is compared with molecular time-division multiple access and a single-transmitter transmission using BER simulations. It is shown that molecular CDMA can outperform the other two systems. Furthermore, it is pointed out that the spreading sequences have to be chosen carefully, since the amplitude-dependent noise of the diffusion-based channel causes different error probabilities. In [DY$^+$17; DY$^+$18], a symmetrical 2×2 molecular MIMO scenario consisting of two point transmitters and two absorbing spherical receivers is considered. For this scenario, an ANN is trained to determine the channel coefficients and its estimates are shown to be sufficiently accurate. The novel contribution of [DY$^+$17; DY$^+$18] is the first investigation of spatial diversity

algorithms in a molecular MIMO scenario under consideration of ISI. For this purpose, transmitter-side repetition MIMO is compared with Alamouti-type coding. At the receiver side, equal gain combining (EGC) which is equal to maximum ratio combining (MRC) in the investigated symmetric scenario is analyzed in combination with ATD and MLSE detection. This analysis is extended in [DY+18] by selection diversity (SD) and FTD. With the help of BER simulations it is shown that the simpler repetition MIMO exceeds the Alamouti-type coding and a SISO transmission system. Furthermore, [DY+18] shows that compared to the SISO transmission an array gain is achieved rather than a spatial diversity gain in the diffusion-based channel.

Advanced Channel Coding In the molecular communication literature, channel equalization and channel coding are often considered separately, but both with the aim of providing robustness against ISI. The main contribution of [DH18] is to consider channel equalization and channel coding simultaneously and to introduce the turbo principle into the area of molecular communication. Analogous to classical wireless communication systems, channel equalization is used to reverse the ISI effect and channel coding is utilized for robustness against the amplitude-dependent noise of the diffusion-based channel. For this purpose, the bit-interleaved coded modulation with iterative detection (BICM-ID) structure is applied, which enables iterative detection/decoding with exchange of extrinsic information according to the turbo principle. With the help of EXIT chart analysis, the performance of the investigated systems is predicted. While in [DS+19] state-of-the-art LDPC codes are analyzed, in [DH18] an irregular convolutional code (IRCC) is matched to the EXIT function of the detector. BER simulations demonstrate the performance of the BICM-ID structure based on the maximum transmission distance. While the simultaneous consideration of channel equalization and channel coding already leads to a significant increase of the maximum transmission distance, this can be further increased by iterative detection (ID). In [DH18], it is additionally demonstrated that a system with an IRCC adapted to the detector gets close to the communication channel capacity.

Author's Publication List

[DH16] M. Damrath and P. A. Hoeher, "Low-complexity adaptive threshold detection for molecular communication," *IEEE Trans. Nanobiosci.*, vol. 15, no. 3, pp. 200–208, Apr. 2016.

Abstract: This paper studies detection algorithms for diffusion-based molecular communication systems, where molecules freely diffuse as information carrier from a transmitter to a receiver in a fluid medium. The main limitations are strong intersymbol interference due to the random propagation of the molecules, and the low-energy/low-complexity assumption regarding future implementations in so-called nanomachines. In this contribution, a new biologically inspired detection algorithm suitable for binary signaling, named adaptive threshold detection, is proposed, which deals with these limitations. The proposed detector is of low complexity, does not require explicit channel knowledge, and seems to be biologically reasonable. Numerical results demonstrate that the proposed detector can outperform the common low-complexity fixed threshold detector under certain conditions. As a benchmark, maximum-likelihood sequence estimation (MLSE) and reduced-state sequence estimation (RSSE) are also analyzed by means of numerical simulations. In addition, the effect of molecular denaturation on the detection performances is studied. It is shown that denaturation generally improves the detection performances, while RSSE is able to outperform MLSE in the case of no denaturation.

[DK+17a] M. Damrath, S. Korte, and P. A. Hoeher, "Equivalent discrete-time channel modeling for molecular communication with emphasize on an absorbing receiver," *IEEE Trans. Nanobiosci.*, vol. 16, no. 1, pp. 60–68, Jan. 2017.

Abstract: This paper introduces the equivalent discrete-time channel model (EDTCM) to the area of diffusion-based molecular communication (DBMC). Emphasis is on an absorbing receiver, which is based on the so-called first passage time concept. In the wireless communications community the EDTCM is well known. Therefore, it is anticipated that the EDTCM improves the accessibility of DBMC and supports the adaptation of classical wireless communication algorithms to the area of DBMC. Furthermore, the EDTCM has the capability to provide a remarkable reduction of computational complexity compared to random walk based DBMC simulators. Besides the exact EDTCM, three approximations thereof based on binomial, Gaussian, and Poisson approximation are proposed and analyzed in order to further reduce computational complexity. In addition, the Bahl-Cocke-Jelinek-Raviv (BCJR) algorithm is adapted to all four channel models. Numerical results show the performance of the exact EDTCM, illustrate the performance of the adapted BCJR algorithm, and demonstrate the accuracy of the approximations.

[DK+17b] M. Damrath, J. J. Koshy, and P. A. Hoeher, "Application of OFDM in diffusion-based molecular communication," *IEEE Trans. Mol. Biol. Multi-Scale Commun.*, vol. 3, no. 4, pp. 254–258, Dec. 2017.

Abstract: In this letter, the application of orthogonal frequency division multiplexing (OFDM) in the area of diffusion-based molecular communication is studied. The focus is on direct current biased optical OFDM, asymmetrically clipped optical OFDM, and pulse amplitude modulation discrete multi-tone transmission known from optical communications. Simulation results show that the OFDM parameters have to be chosen carefully and that the bit error rate performance is in-between that of standard on-off keying transmission with maximum-likelihood sequence estimation and fixed threshold detection. In addition, it is demonstrated that a proposed bit-loading algorithm can further improve the OFDM performance.

[DY+17] M. Damrath, H. B. Yilmaz, C. Chae, and P. A. Hoeher, "Spatial coding techniques for molecular MIMO," in *Proc. IEEE Information Theory Workshop (ITW)*, Kaohsiung, Taiwan, Nov. 2017, pp. 324–328.

Abstract: This paper presents spatial diversity techniques applied to multiple-input multiple-output (MIMO) diffusion-based molecular communications (DBMC). Two types of spatial coding techniques, namely Alamouti-type coding and repetition MIMO coding are suggested and analyzed. In addition, we consider receiver-side equal-gain combining, which is equivalent to maximum-ratio combining in symmetrical scenarios. For numerical analysis, the channel impulse responses of a symmetrical 2x2 MIMO-DBMC system are acquired by a trained artificial neural network. It is demonstrated that spatial diversity has the potential to improve the system performance and that repetition MIMO coding outperforms Alamouti-type coding.

[DH+17] M. Damrath, P. A. Hoeher, and G. J. Forkel, "Symbol detection based on Voronoi surfaces with emphasis on superposition modulation," *Digital Communications and Networks*, vol. 3, no. 3, pp. 141–149, Aug. 2017.

Abstract: A challenging task when applying high-order digital modulation schemes is the complexity of the detector. Particularly, the complexity of the optimal a posteriori probability (APP) detector increases exponentially with respect to the number of bits per data symbol. This statement is also true for the Max-Log-APP detector, which is a common simplification of the APP detector. Thus it is important to design new detection algorithms which combine a sufficient per-

formance with low complexity. In this contribution, a detection algorithm for two-dimensional digital modulation schemes which cannot be split-up into real and imaginary parts (like phase shift keying and phase-shifted superposition modulation (PSM)) is proposed with emphasis on PSM with equal power allocation. This algorithm exploits the relationship between Max-Log-APP detection and a Voronoi diagram to determine planar surfaces of the soft outputs over the entire range of detector input values. As opposed to state-of-the-art detectors based on Voronoi surfaces, a priori information is taken into account, enabling iterative processing. Since the algorithm achieves Max-Log-APP performance, even in the presence of a priori information, this implies a great potential for complexity reduction compared to the classical APP detection.

[KD+17] S. Korte, M. Damrath, and P. A. Hoeher, "Multiple channel access techniques for diffusion-based molecular communications," in *Proc. Int. ITG-Conf. on Systems, Communications and Coding (SCC)*, Hamburg, Germany, Feb. 2017, pp. 1–6.

Abstract: In this work, a transmission link between an arbitrary number of point sources and one spherical absorbing receiver is established considering diffusion-based molecular propagation in a three-dimensional environment filled with a fluid medium without flow. Therefore an equivalent discrete time channel model for multiuser communication is derived, molecule shift keying is used as modulation technique and an appropriate low-complexity adaptive threshold detector is implemented. For multiple channel access, molecular time division multiple access (MoTDMA) and as a main contribution molecular code division multiple access (MoCDMA) are introduced. The performances of the developed communication systems are analyzed in terms of bit-error-rate running numerical simulations.

[DH18] M. Damrath and P. A. Hoeher, "EXIT-chart-aided code matching in molecular communications," in *Proc. Int. Symp. on Turbo Codes & Iterative Information Processing (ISTC)*, Hong Kong, Dec. 2018, pp. 1–5.

Abstract: Bit-interleaved coded modulation (BICM) systems are introduced to the area of diffusion-based molecular communication (DBMC). To improve receiver-side iterative detection, irregular convolutional codes (IRCC) are matched to the extrinsic information transfer (EXIT) function of the detector. Numerical bit error ratio (BER) simulations confirm that a channel code adapted to the EXIT function of the detector increases the performance of the DBMC system in terms of BER and distance.

[DY+18] M. Damrath, H. B. Yilmaz, C. Chae, and P. A. Hoeher, "Array gain analysis in molecular MIMO communications," *IEEE Access*, vol. 6, pp. 61 091–61 102, Oct. 2018.

Abstract: In this paper, spatial transmission techniques in the area of multiple-input multipleoutput (MIMO) diffusion-based molecular communications (DBMC) are investigated. For transmitter-side spatial coding, Alamouti-type coding and repetition MIMO coding are analyzed. At the receiver-side, selection diversity and equal-gain combining are studied as combining strategies. Throughout the numerical analysis, a symmetrical 2x2 MIMO-DBMC system is assumed. Furthermore, a trained artificial neural network is utilized to acquire the channel impulse responses. The numerical analysis demonstrates that there is no spatial diversity gain in the DBMC system under investigation, but that it is possible to achieve an array gain instead. In addition, it is shown that for MIMO-DBMC systems repetition MIMO coding is superior to Alamouti-type coding.

[DH+18] M. Damrath, P. A. Hoeher, and G. J. Forkel, "Piecewise linear detection for direct superposition modulation," *Digital Communications and Networks*, vol. 4, no. 2, pp. 98–105, Apr. 2018.

Abstract: Considering high-order digital modulation schemes, the bottleneck in consumer products is the detector rather than the modulator. The complexity of the optimal a posteriori probability (APP) detector increases exponentially with respect to the number of modulated bits per data symbol. Thus, it is necessary to develop low-complexity detection algorithms with an APP-like performance, especially when performing iterative detection, for example in conjunction with bit interleaved coded modulation. We show that a special case of superposition modulation, dubbed Direct Superposition Modulation (DSM), is particularly suitable for complexity reduction at the receiver side. As opposed to square QAM, DSM achieves capacity without active signal shaping. The main contribution is a low-cost detection algorithm for DSM, which enables iterative detection by taking a priori information into account. This algorithm exploits the approximate piecewise linear behavior of the soft outputs of an APP detector over the entire range of detector input values. A theoretical analysis and simulation results demonstrate that at least max-log APP performance can be reached, while the complexity is significantly reduced compared to classical APP detection.

[BD+19] S. Bhattacharjee, M. Damrath, and P. A. Hoeher, "EXIT chart analysis of higher order modulation schemes in molecular communications,"

in *Proc. ACM Int. Conf. on Nanoscale Computing and Communication (NANOCOM)*, Dublin, Ireland, Sep. 2019, pp. 1–6.

Abstract: This paper discusses and compares the performance of existing modulation schemes in diffusion-based molecular communication systems by means of extrinsic information transfer chart and bit error rate analysis. Additionally, a new modulation scheme called variable concentration shift keying is introduced with proven advantage over conventional concentration shift keying in terms of bit error rate performance.

[DS+19] M. Damrath, M. Schurwanz, and P. A. Hoeher, "The turbo principle in molecular communications," in *Proc. Int. ITG-Conf. on Systems, Communications and Coding (SCC)*, Rostock, Germany, Feb. 2019, pp. 1–6.

Abstract: This work introduces the turbo principle into the area of diffusion-based molecular communication (DBMC). By means of bit-interleaved coded modulation (BICM) at the transmitter side, channel encoder and modulator are serially concatenated with a random interleaver in-between. At the receiver side, iterative processing between detector and channel decoder is performed. The convergence of the iterative processing is investigated by an extrinsic information transfer (EXIT) chart analysis. In addition, the EXIT chart is used to predict the system performance and to determine the communication channel capacity including the DBMC channel. Bit-error-rate (BER) simulations are performed for advanced low-density parity-check (LDPC) codes. These numerical results confirm the benefit of turbo processing over non-iterative processing for DBMC applications.

1.5 Structure of Dissertation

The content of the dissertation is structured as follows. Chapter 2 deals with the modeling of the diffusion-based transmission channel and thus represents the basis for the analyses presented in the subsequent chapters. First, the system model and its assumptions are described. Three different methods for determining the channel impulse response are then presented and the equivalent discrete-time channel model is established including different approximations. Finally, a macroscopic alcohol-based testbed setup, for which a channel impulse response is determined, is described as proof of concept. This channel impulse response is used for prediction of a sequence transmission and compared with testbed measurements. Chapter 3 deals with transmitter and receiver algorithms based

on the channel model presented in Chapter 2. The focus is on modulation schemes at the transmitter side and on detectors and channel equalizers at the receiver side. Furthermore, orthogonal frequency-division multiplexing is regarded as a multi-carrier method which elegantly combines modulation and channel equalization/detection. Chapter 4 examines the improvement of transmission by utilizing channel codes. In addition to classical block and convolutional FEC codes, line codes, spreading codes and spatial codes are analyzed as well. The analysis of the codes is performed under different aspects of normalization and shows corresponding gains and losses. Chapter 5 deals with advanced channel coding applying the turbo principle, in which the decoder and detector are connected by an interleaver and exchange extrinsic information. Using an extrinsic information transfer chart analysis, the communication channel capacity is determined, the iterative detection performance is predicted and an adapted irregular convolutional code is designed. The performance is evaluated using bit error rate simulations. Chapter 6 concludes the dissertation with a summary of the contents. In addition, an outlook gives inspiration for further research.

The notation throughout the dissertation, which comprises amongst others acronyms and mathematical notations, is listed in Appendix A. Appendix B collects supporting mathematical derivations. Furthermore, Appendix C summarizes the simulation parameters of all simulations presented in this dissertation. The listed values are intended as default values, which are fixed if the parameters are not varied throughout the simulation.

2

Diffusion-Based Channel Modeling

This chapter forms the basis for the transmission algorithms examined in the following chapters. In addition to a description of the system model and its assumptions, various methods for channel characterization are presented. Based on these, a channel model is introduced. Finally, a testbed setup is presented as a proof of concept.

2.1 System Model and Assumptions

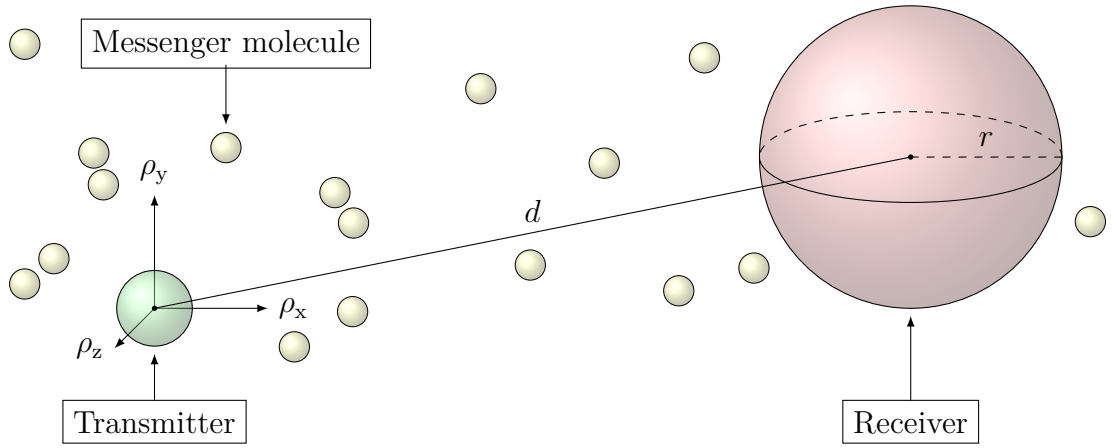


Figure 2.1: Visualization of the diffusion-based molecular communication scenario under investigation.

In this dissertation, a three-dimensional communication scenario is assumed, which is shown in Fig. 2.1 [DH16]. The unbounded environment is filled with a fluid medium, which

together with the diffusing particle is defined by its diffusion coefficient [NE+13]

$$D = \frac{k_{\mathrm{B}}\vartheta}{6\pi\eta r_{\mathrm{mol}}}, \tag{2.1}$$

where $k_{\mathrm{B}} \approx 1.38 \times 10^{-23}\,\mathrm{J/K}$ is the Boltzmann constant, ϑ and η are the temperature and viscosity of the medium respectively, and r_{mol} is the hydrodynamic radius of the diffusing particle. It is assumed that the medium is stationary and homogeneous, i. e., it does not change its temperature and viscosity over time or location. Furthermore, the medium is free of any form of drift/flow.

As shown in Fig. 2.1, the communication scenario consists of a transmitter and a receiver. It is assumed that the transmitter is a point source with an infinitesimally small volume. In the sense of a point source, it is able to release an impulse of messenger molecules at its location at any time. Without loss of generality, it is assumed that the transmitter is in the coordinates origin of the three-dimensional environment. The receiver is assumed to have a spherical shape defined by the radius r. It is located at a distance $d > r$ away from the transmitter, where d is defined as the distance between the centers of the receiver and the transmitter, respectively. The surface of the receiver consists of ideal receptors which are sensitive to the messenger molecules. Whenever a messenger molecule hits a receptor, it is immediately counted and removed from the environment. The removal of message molecules after reaching the receiver is a natural phenomenon. In nature, receiving cells often modify or degrade signal molecules [LB+00]. For example, acetylcholines are degraded by acetylcholinesterase in neuromuscular junctions [KY+13]. The spatial expansion of the receptors is assumed to be infinitesimally small. The density of receptors on the surface is assumed to be high enough to count messenger molecules every time they hit the surface of the receiver. Consequently, the receiver is modeled as an ideal absorbing receiver. Throughout this dissertation, it is assumed that the receiver performs energy detection, i. e., it is capable to sum up the number of molecules received over a symbol duration T. It is also assumed that the transmitter and receiver are perfectly synchronized in time. Time synchronization can be achieved, for example, by an external signal such as the human heartbeat or, as described in [MN13], by the release of inhibitory molecules.

The messenger molecules, emitted by the transmitter and detected by the receiver, act as information carriers in the communication system. It is assumed that they are all of the same type and are the only molecules in the environment. They propagate independently of each other, do not collide, and can be in the same position at the same time. Consequently, the messenger molecules follow the principle of Brownian motion. The stochastic propagation

of the messenger molecules by diffusion leads to a diffusion noise, which describes the deviation from the expected propagation. In this dissertation, it is assumed that this is the only noise source in the system.

This system model modulates a diffusion-based intercellular transmission for example between two bacteria or two nanomachines. It is a rather simple model, which has its advantages in its comparatively simple description. Moreover, the model focuses on the diffusion propagation of the molecules without considering other effects such as additional molecules in the environment, a heterogeneous fluid medium, receptor binding processes, or interactions between molecules. This is consistent with the focus of the dissertation on the effect of diffusion.

Table 2.1: Assumed default parameters for the scenarios theoretically examined in this dissertation. While the results are presented only for microscopic parameters, the macroscopic parameters lead to equivalent results.

Parameter		Value	
		Microscopic	Macroscopic
Diffusion coefficient	D	$4.367\,64 \times 10^{-10}\,\mathrm{m^2/s}$	$4.367\,64 \times 10^{-4}\,\mathrm{m^2/s}$
Receiver radius	r	$4.5\,\mu\mathrm{m}$	$4.5\,\mathrm{mm}$
Transmission distance	d	$10\,\mu\mathrm{m} \ldots 150\,\mu\mathrm{m}$	$10\,\mathrm{mm} \ldots 150\,\mathrm{mm}$

Tab. 2.1 presents the default parameters for the scenario which is the basis for the theoretical investigations carried out in this dissertation. While results for the microscopic parameters are presented in the following, the macroscopic parameters lead to equivalent results. Thus, the presented results can also be interpreted for macroscopic molecular communication.

The microscopic parameters are biologically motivated and inspired by "System 1" applied in [NC+14a]. Water is considered as a transmission medium at a temperature of $\vartheta = 25\,^{\circ}\mathrm{C}$ with a viscosity of $\eta = 1 \times 10^{-3}\,\mathrm{kg/m/s}$. The radius of the messenger molecules is assumed to be $r_{\mathrm{mol}} = 0.5\,\mathrm{nm}$, which corresponds to the size of common small organic molecules such as glucose, amino acids, and nucleotides [NC+14a]. These values result in the diffusion coefficient D according to equation (2.1). In contrast to [NC+14a], the receiver radius is set to $r = 4.5\,\mu\mathrm{m}$ instead of 45 nm, which is a more realistic value for a cell. The size of prokaryotic cells ranges typically between $1\,\mu\mathrm{m}$ and $10\,\mu\mathrm{m}$ [NE+13].

The macroscopic parameters are obtained by upscaling the microscopic parameters under the constraint of providing equivalent results. Nevertheless, they represent realistic values for a macroscopic scenario. The receiver radius of $r = 4.5\,\mathrm{mm}$ is feasible as the size of the sensitive layer of a sensor (see Section 2.4). The diffusion coefficient D is in the order of

magnitude of diffusing substances in gases. Typical values ranges between $0.1 \times 10^{-4}\,\mathrm{m}^2/\mathrm{s}$ and $1 \times 10^{-4}\,\mathrm{m}^2/\mathrm{s}$ at a pressure of 1 atm and near room temperature [Cus09].

The resulting transmission distances in the interval of seconds by pure diffusion are in micrometers in the microscopic scenario, while millimeters are reached in the macroscopic scenario. It should be noted that a significant increase in range can be achieved by a favorable drift in the medium. Thus the transmission distances can be increased as shown for example in [DH18] to millimeters in the microscopic scenario and to meters in the macroscopic scenario. The macroscopic testbed presented in Section 2.4 also demonstrates transmission in the meter range.

2.2 Channel Characterization

The channel describes the diffusion propagation of a molecule in the environment after its release at the transmitter. Diffusion is a non-directional random motion of particles due to their thermal energy and the collisions with their vicinity. Since the movement is based only on the thermal energy of the environment, it is a passive type of propagation. Even if individual particles perform a non-directional random motion, diffusion counteracts the concentration gradient in a fluid. This can be illustrated by the following thought experiment: Imagine a container with a fluid. Initially, the concentration of a substance on the right side of the container is much higher than on the left side of the container. All particles of the substance perform a non-directional random motion, so the probability of a specific particle changing sides is the same on both sides. Since there are more particles on the right side, in total more particles will diffuse from the right to the left in the beginning. This will continue until the concentrations on the left and right side are equal. In this balanced state of concentration, on average the same amount of particles diffuse from right to left and from left to right. Thus the concentration is balanced and no longer changes significantly.

The concentration change of a substance in a medium can be described by Fick's second law [Ber93]

$$\frac{\partial \phi(\boldsymbol{\rho}, t)}{\partial t} = D \nabla^2 \phi(\boldsymbol{\rho}, t), \tag{2.2}$$

where $\phi(\boldsymbol{\rho}, t)$ is the concentration of the substance at location $\boldsymbol{\rho} = [\rho_\mathrm{x}, \rho_\mathrm{y}, \rho_\mathrm{z}]$ at time t and ∇^2 is the Laplace operator.

Assuming linearity in a communication system, the channel can be precisely described by its impulse response $h(t)$ [PS07]. In case of the molecular communication scenario under

consideration, the impulse response is proportional to the first passage time density [FY+16; DK+17a], which describes the probability that a molecule will reach the receiver per unit of time t. The integral over the first passage time density defines the probability that a molecule reaches the receiver. In the unbounded three-dimensional case, the probability that a molecule will not reach the receiver within infinite time is non-zero, so the integral over the first passage time density is less than one. Since a system based on time slots with a summing receiver is assumed, the probability that a molecule will reach the receiver during the ℓth symbol duration after its release is proportional to the ℓth channel coefficient

$$h_\ell = \int_{\ell T}^{(\ell+1)T} h(\tau)\,\mathrm{d}\tau. \tag{2.3}$$

Due to the proportionality, the channel impulse response is set equal to the first passage time density and the channel coefficients equal to the arrival probabilities of a molecule throughout the remainder of this dissertation. Thus, the terms will be used interchangeably. There are different methods to determine $h(t)$ and h_ℓ. In the following, three ways are introduced. The first way is based on the numerical simulation of propagating molecules, the second on the analytical solution of Fick's second law, and the third on an artificial neural network (ANN). While the analytical solution can only be derived for certain scenarios and assumptions, the numerical simulation and training of an ANN is always possible. The latter two methods are always associated with a small inaccuracy in a practical implementation. A trained ANN offers a computational complexity advantage over numerical simulations, if the channel is to be determined for many varying system parameters.

2.2.1 Random Walk Characterization

Since the message molecules collide with the typically smaller and faster particles of the fluid medium, their propagation can be described by Brownian motion [NE+13]. Brownian motion is derived by solving (2.2). Separate derivation with respect to time and spatial directions (see Appendix B.1) shows that

$$\phi(\boldsymbol{\rho} = [\rho_\mathrm{x}, \rho_\mathrm{y}, \rho_\mathrm{z}], t) = \frac{A}{\sqrt{t^3}} \exp\left(-\frac{\rho_\mathrm{x}^2 + \rho_\mathrm{y}^2 + \rho_\mathrm{z}^2}{4Dt}\right) \tag{2.4}$$

is a valid solution of (2.2), where A is an arbitrary constant. By normalizing the concentration of molecules in the entire space to unity (just a single molecule), the constant A

can be determined:

$$
\begin{aligned}
1 &\stackrel{!}{=} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \phi(\boldsymbol{\rho}, t)\,\mathrm{d}\rho_\mathrm{x}\,\mathrm{d}\rho_\mathrm{y}\,\mathrm{d}\rho_\mathrm{z} \\
&= \frac{A}{\sqrt{t^3}}\left[\int_{-\infty}^{\infty} \exp\left(-\frac{\rho_\mathrm{x}^2}{4Dt}\right)\mathrm{d}\rho_\mathrm{x} \cdot \int_{-\infty}^{\infty} \exp\left(-\frac{\rho_\mathrm{y}^2}{4Dt}\right)\mathrm{d}\rho_\mathrm{y} \cdot \int_{-\infty}^{\infty} \exp\left(-\frac{\rho_\mathrm{z}^2}{4Dt}\right)\mathrm{d}\rho_\mathrm{z}\right] \qquad (2.5) \\
&= \frac{A(4\pi Dt)^{\frac{3}{2}}}{\sqrt{t^3}}, \\
A &= \frac{\sqrt{t^3}}{(4\pi Dt)^{\frac{3}{2}}}. \qquad (2.6)
\end{aligned}
$$

Consequently, entering the result of (2.6) into (2.4), the concentration of a molecule is distributed as

$$
\phi(\boldsymbol{\rho} = [\rho_\mathrm{x}, \rho_\mathrm{y}, \rho_\mathrm{z}], t) = \frac{1}{(4\pi Dt)^{\frac{3}{2}}} \exp\left(-\frac{\rho_\mathrm{x}^2 + \rho_\mathrm{y}^2 + \rho_\mathrm{z}^2}{4Dt}\right). \qquad (2.7)
$$

From (2.7), it is seen that Brownian motion can be described by a Wiener process $W(t)$, with the increments $W(t+\Delta t) - W(t)$ in all spatial directions following the distribution $\mathcal{N}(0, 2D\Delta t)$, where $\mathcal{N}\left(\mu, \sigma^2\right)$ is the Gaussian distribution with mean μ and variance σ^2. The Wiener process can be approximated by a three-dimensional random walk for small step size Δt, which converges to the Wiener process for a step size $\Delta t \to 0$ [NE+13]. The update formulas of the corresponding random walk are given as

$$
\begin{aligned}
\rho_\mathrm{x}[i+1] &= \rho_\mathrm{x}[i] + \Delta\rho_\mathrm{x}, \qquad &\text{with} \quad \Delta\rho_\mathrm{x} \sim \mathcal{N}(0, 2D\Delta t)\,, \\
\rho_\mathrm{y}[i+1] &= \rho_\mathrm{y}[i] + \Delta\rho_\mathrm{y}, \qquad &\text{with} \quad \Delta\rho_\mathrm{y} \sim \mathcal{N}(0, 2D\Delta t)\,, \\
\rho_\mathrm{z}[i+1] &= \rho_\mathrm{z}[i] + \Delta\rho_\mathrm{z}, \qquad &\text{with} \quad \Delta\rho_\mathrm{z} \sim \mathcal{N}(0, 2D\Delta t)\,,
\end{aligned} \qquad (2.8)
$$

where $i \in \mathbb{N}_0$ denotes the simulation step, $\left[\rho_\mathrm{x}[i], \rho_\mathrm{y}[i], \rho_\mathrm{z}[i]\right]$ is the position of the molecule at time instant $i\Delta t$, and $\left[\rho_\mathrm{x}[0], \rho_\mathrm{y}[0], \rho_\mathrm{z}[0]\right]$ defines the initial release position of the molecule at $t = 0$ [FY+16].

Channel Impulse Response

Due to its proportionality, the channel impulse response $h(t)$ in the scenario under investigation is set equal to the first passage time density, which describes the probability that a

molecule will reach the spherical absorbing receiver per unit of time t. This density can be determined using a Monte Carlo simulation in which the propagation of the molecules is simulated as a three-dimensional random walk [NE+13]. For this purpose, in each simulation step i it is checked whether a molecule is inside the receiver. If this is the case, the molecule is counted as received at time $i\Delta t$ and removed from the simulation. If $\boldsymbol{y}$ is the vector containing the number of received molecules $y[i]$ in each simulation step i and the propagation of N molecules has been simulated, then the following relationship applies

$$h(i\Delta t) \approx \frac{y[i]}{N\Delta t}. \tag{2.9}$$

Since (2.9) is executed by a Monte Carlo simulation, this approximation becomes more accurate when the number of simulated propagations N is increased. In addition, the accuracy increases with smaller time steps Δt.

Channel Coefficients

According to (2.3), the channel coefficients h_ℓ result from the integration of the channel impulse response $h(t)$ over the symbol duration T. Due to proportionality, this results in the probability that a molecule is absorbed by the receiver in the ℓth symbol duration after its release. This probability can be obtained directly from the random walk based Monte Carlo simulation. Assuming that T is a multiple of simulation time step Δt, the channel coefficients can be approximated as

$$h_\ell \approx \frac{1}{N} \sum_{i=\ell\frac{T}{\Delta t}+1}^{(\ell+1)\frac{T}{\Delta t}} y[i]. \tag{2.10}$$

As with the channel impulse response, the accuracy of the approximation can be increased by smaller simulation steps Δt and a larger number N of simulated molecules.

2.2.2 Analytical Characterization

As an alternative to the simulative approximation of the channel impulse response by the random walk, it can also be determined analytically for the scenario under investigation. For this, (2.2) must be solved taking the absorbing receiver sphere into account [SK00; YH+14]. If, in contrast to the assumption in Section 2.1, the origin of the coordinate system is chosen in the center of the absorbing sphere, the problem shows a spherical symmetry. In the spherical coordinate system defined by $[\rho, \theta, \varphi]$, the probability that a molecule will

be absorbed by the receiver depends only on the radial component ρ, i. e., the distance d to the emission point, but not on the orientation θ and φ. Thus $\phi(\rho, t)$ is considered in the following, where the origin of the spherical coordinate system lies in the center of the receiving sphere.

Consequently, by applying Laplace operator for spherical coordinates to (2.2), Fick's second law results in [SK00]

$$\begin{aligned} \frac{\partial \phi(\rho,t)}{\partial t} &= \frac{D}{\rho^2}\frac{\partial}{\partial \rho}\left[\rho^2 \frac{\partial \phi(\rho,t)}{\partial \rho}\right] \\ &= \frac{D}{\rho}\frac{\partial^2}{\partial \rho^2}\Big[\rho\,\phi(\rho,t)\Big], \end{aligned} \tag{2.11}$$

$$\frac{\partial}{\partial t}\Big[\rho\,\phi(\rho,t)\Big] = D\frac{\partial^2}{\partial \rho^2}\Big[\rho\,\phi(\rho,t)\Big], \tag{2.12}$$

where the identity $\frac{1}{\rho^2}\frac{\partial}{\partial \rho}\left[\rho^2 \frac{\partial f(\rho)}{\partial \rho}\right] = \frac{1}{\rho}\frac{\partial^2}{\partial \rho^2}\Big[\rho\, f\,(\rho)\Big]$, proved in Appendix B.2, is applied. In order to solve (2.12), the following initial condition and two boundary conditions must be satisfied [SK00; YH+14]:

$$\text{Initial condition:} \qquad \lim_{t\to 0}\phi(\rho,t) = \frac{1}{4\pi d^2}\delta(\rho - d), \tag{2.13}$$

$$\text{1. Boundary condition:} \qquad \lim_{\rho\to\infty}\phi(\rho,t) = 0, \tag{2.14}$$

$$\text{2. Boundary condition:} \qquad \phi(\rho = r, t) = 0. \tag{2.15}$$

The initial condition (2.13) results from the fact that the molecules are located at $\rho = d$ when they are released at $t = 0$. Hence $\lim_{t\to 0}\phi(\rho,t) = B\delta(\rho - d)$ results, where B is an arbitrary constant. The constant B can be determined by normalizing the concentration in the whole environment to unity (just a single molecule):

$$\begin{aligned} 1 &\overset{!}{=} \int_0^{2\pi}\int_0^{\pi}\int_0^{\infty} B\delta(\rho - d)\rho^2 \sin\theta \,\mathrm{d}\rho\,\mathrm{d}\theta\,\mathrm{d}\varphi \\ &= 4\pi \int_0^{\infty} B\delta(\rho - d)\rho^2 \,\mathrm{d}\rho \\ &= 4\pi B d^2, \end{aligned} \tag{2.16}$$

$$B = \frac{1}{4\pi d^2}. \tag{2.17}$$

The first boundary condition (2.14) ensures that the concentration of molecules vanishes for locations far away from their release point. The second boundary condition (2.15) ensures the absorption of the molecules on the surface of the spherical receiver. To solve equation (2.12), $\phi(\rho, t)$ is partitioned into

$$\phi(\rho, t) = \xi(\rho, t) + \upsilon(\rho, t), \tag{2.18}$$

where $\xi(\rho, t)$ solves (2.12) with respect to (2.13) and (2.14) without an absorbing sphere, and $\upsilon(\rho, t)$ solves (2.12) with respect to (2.14) and ensures that $\phi(\rho, t)$ satisfies (2.15) [SK00].

Determination of $\xi(\rho, t)$

The following derivation is based on the procedure given in [SK00]. A general structure of a valid solution of (2.12) is not straightforward. In order to solve (2.12) for $\xi(\rho, t)$, the Fourier transform of $\rho \cdot \xi(\rho, t)$ defined as

$$\Xi(\zeta, t) = \mathcal{F}\{\rho \cdot \xi(\rho, t)\} = \int_{-\infty}^{\infty} \rho \cdot \xi(\rho, t) \cdot \exp(-j\zeta\rho)\, \mathrm{d}\rho \tag{2.19}$$

is applied. Combining (2.12) and (2.19) yields

$$\frac{\partial}{\partial t}\Xi(\zeta, t) = -D\zeta^2 \Xi(\zeta, t), \tag{2.20}$$

which has the straightforward solution

$$\Xi(\zeta, t) = A_\xi \exp\left(-D\zeta^2 t\right), \tag{2.21}$$

where A_ξ is an arbitrary time-independent constant which can be determined by the initial condition (2.13). Modifying (2.13) and applying Fourier transform leads to

$$\lim_{t\to 0} \rho \cdot \xi(\rho, t) = \frac{1}{4\pi d}\delta(\rho - d) \tag{2.22}$$

and

$$\mathcal{F}\left\{\frac{1}{4\pi d}\delta(\rho - d)\right\} = \frac{1}{4\pi d}\exp(-j\zeta d) \stackrel{!}{=} \lim_{t\to 0} \Xi(\zeta, t) = A_\xi. \tag{2.23}$$

By applying the inverse Fourier transform

$$\rho \cdot \xi(\rho, t) = \mathcal{F}^{-1}\{\Xi(\zeta, t)\} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Xi(\zeta, t) \cdot \exp(j\zeta\rho)\, \mathrm{d}\zeta \tag{2.24}$$

and exploiting identity [GR+14]

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \exp\left(-a\zeta^2\right) \cdot \exp(jb\zeta)\, \mathrm{d}\zeta = \frac{1}{\sqrt{4\pi a}} \exp\left(-\frac{b^2}{4a}\right) \tag{2.25}$$

the final solution for $\rho \cdot \xi(\rho, t)$ is obtained as

$$\rho \cdot \xi(\rho, t) = \frac{1}{4\pi d} \frac{1}{\sqrt{4\pi D t}} \exp\left(-\frac{(\rho - d)^2}{4Dt}\right). \tag{2.26}$$

Determination of $\upsilon(\rho, t)$

The following derivation is based on the procedure given in [SK00]. If equation (2.12) is solved for $\upsilon(\rho, t)$, the initial condition is

$$\lim_{t \to 0} \rho \cdot \upsilon(\rho, t) = 0 \tag{2.27}$$

in order not to violate the initial condition (2.13) of $\phi(\rho, t)$. To find a solution of (2.12), the Laplace transform

$$\Upsilon(\rho, s) = \mathcal{L}\{\upsilon(\rho, t)\} = \int_{0}^{\infty} \upsilon(\rho, t) \cdot \exp(-st)\, \mathrm{d}t \tag{2.28}$$

is applied. Combining (2.12) and (2.28) yields

$$s\rho\Upsilon(\rho, s) - \rho \cdot \upsilon(\rho, 0) = s\rho\Upsilon(\rho, s) = D \frac{\partial^2}{\partial \rho^2} \Big[\rho\Upsilon(\rho, s)\Big], \tag{2.29}$$

which, considering (2.14), has the solution

$$\rho\Upsilon(\rho, s) = A_{\upsilon} \exp\left(-\sqrt{\frac{s}{D}}\rho\right), \tag{2.30}$$

where A_υ is a constant which should be determined with respect to (2.15). By using the Laplace transform identity [GR+14]

$$\mathcal{L}\left\{\frac{1}{\sqrt{\pi t}}\exp\left(-\frac{a^2}{4t}\right)\right\} = \frac{1}{\sqrt{s}}\exp\left(-|a|\sqrt{s}\right), \tag{2.31}$$

$\rho \cdot \xi(\rho, t)$ is Laplace transformed to

$$\mathcal{L}\left\{\rho \cdot \xi(\rho, t)\right\} = \frac{1}{4\pi d}\frac{1}{\sqrt{4Ds}}\exp\left(-\sqrt{\frac{s}{D}}|\rho - d|\right). \tag{2.32}$$

Applying the second boundary condition (2.15) in Laplace domain to $\Phi(\rho, s) = \mathcal{L}\left\{\rho \cdot \phi(\rho, t)\right\}$ yields a solution for constant A_υ:

$$0 \stackrel{!}{=} \frac{1}{4\pi d}\frac{1}{\sqrt{4Ds}}\exp\left(-\sqrt{\frac{s}{D}}(d - r)\right) + A_\upsilon \exp\left(-\sqrt{\frac{s}{D}}r\right), \tag{2.33}$$

$$A_\upsilon = -\frac{1}{4\pi d}\frac{1}{\sqrt{4Ds}}\exp\left(-\sqrt{\frac{s}{D}}(d - 2r)\right). \tag{2.34}$$

Performing the inverse Laplace transform utilizing identity (2.31), the final solution for $\rho \cdot \upsilon(\rho, t)$ for $\rho \geq r$ can be obtained as

$$\rho \cdot \upsilon(\rho, t) = -\frac{1}{4\pi d}\frac{1}{\sqrt{4\pi Dt}}\exp\left(-\frac{(\rho + d - 2r)^2}{4Dt}\right). \tag{2.35}$$

Channel Impulse Response

Due to its proportionality, the channel impulse response $h(t)$ in the scenario under investigation is set equal to the first passage time density, which describes the probability that a molecule will reach the spherical absorbing receiver per unit of time t. In other words, it represents the expected flux of molecules at the surface of the spherical receiver at time t. The flux of molecules at a point ρ is described by Fick's first law [SK00]

$$\begin{aligned} D\frac{\partial\phi(\rho, t)}{\partial\rho} &= D\frac{\partial}{\partial\rho}\left\{\frac{1}{4\pi d\rho}\frac{1}{\sqrt{4\pi Dt}}\left[\exp\left(-\frac{(\rho - d)^2}{4Dt}\right) - \exp\left(-\frac{(\rho + d - 2r)^2}{4Dt}\right)\right]\right\} \\ &= \frac{D}{4\pi d\sqrt{4\pi Dt}}\left\{\left[\frac{1}{\rho^2} + \frac{2\rho + 2d - 4r}{4Dt\rho}\right]\exp\left(-\frac{(\rho + d - 2r)^2}{4Dt}\right)\right. \\ &\quad \left. - \left[\frac{1}{\rho^2} + \frac{2\rho - 2d}{4Dt\rho}\right]\exp\left(-\frac{(\rho - d)^2}{4Dt}\right)\right\}, \end{aligned} \tag{2.36}$$

where $\phi(\rho, t)$ is observed from combining (2.18), (2.26), and (2.35). The channel impulse response is obtained from evaluating (2.36) for $\rho = r$ and integrating over the spherical surface

$$\begin{aligned} h(t) &= \int_0^{\pi} \int_0^{2\pi} \left[D \frac{\partial \phi(\rho, t)}{\partial \rho} \right] \rho^2 \sin\theta \, \mathrm{d}\varphi \, \mathrm{d}\theta \Bigg|_{\rho=r} \\ &= 4\pi\rho^2 D \frac{\partial \phi(\rho, t)}{\partial \rho} \Bigg|_{\rho=r} \\ &= \frac{r^2 D}{d\sqrt{4\pi D t}} \frac{4d - 4r}{4Dtr} \exp\left(-\frac{(d-r)^2}{4Dt} \right) \\ &= \frac{r}{d} \frac{d-r}{\sqrt{4\pi D t^3}} \exp\left(-\frac{(d-r)^2}{4Dt} \right). \end{aligned} \tag{2.37}$$

Channel Coefficients

The channel coefficients h_ℓ can be determined according to (2.3) by integrating (2.37). This results in the channel coefficients [DK+17a]

$$\begin{aligned} h_\ell &= \int_{\ell T}^{(\ell+1)T} \frac{r}{d} \frac{d-r}{\sqrt{4\pi D t^3}} \exp\left(-\frac{(d-r)^2}{4Dt} \right) \mathrm{d}\tau \\ &= \int_0^{(\ell+1)T} \frac{r}{d} \frac{d-r}{\sqrt{4\pi D t^3}} \exp\left(-\frac{(d-r)^2}{4Dt} \right) \mathrm{d}\tau - \int_0^{\ell T} \frac{r}{d} \frac{d-r}{\sqrt{4\pi D t^3}} \exp\left(-\frac{(d-r)^2}{4Dt} \right) \mathrm{d}\tau \\ &= \frac{r}{d} \operatorname{erfc}\left(\frac{d-r}{\sqrt{4D(\ell+1)T}} \right) - \frac{r}{d} \operatorname{erfc}\left(\frac{d-r}{\sqrt{4D\ell T}} \right), \end{aligned} \tag{2.38}$$

where

$$\operatorname{erfc}(t) = \frac{2}{\sqrt{\pi}} \int_t^{\infty} \exp\left(-\tau^2 \right) \mathrm{d}\tau \tag{2.39}$$

is the complementary error function. From (2.38) for the case $\ell = 0$ and $T \to \infty$, it can be seen that in the unbounded three-dimensional scenario under investigation the probability of a molecule not reaching the receiver is non-zero, i. e. $h_0 = r/d < 1$.

2.2.3 Artificial Neural Network Characterization

While the random walk based characterization of the channel can be performed for arbitrary scenarios, an analytical solution only exists for a few special scenarios. For multiple absorbing receivers or for arbitrary drift in the three-dimensional environment, for example, an closed analytical solution does not exist so far. An individual random walk based characterization of the channel for each parameter set often requires a very high computational complexity. In these cases it is advisable to train an ANN for reference parameter sets in order to estimate the channel impulse response for new parameter sets. Since only a few reference parameter sets have to be determined with the help of random walk simulations, the computational effort for the other parameter sets can be avoided. The design, training and optimization of ANNs is an independent and extensive field of research that would go beyond the scope of this dissertation. Therefore, only a fully connected feed-forward neural network with three layers structured and trained as in [LY+17; DY+17; DY+18] is considered here. In order to explain the general procedure and the network structure, an ANN is exemplarily designed, which estimates the cumulative channel impulse response of the point-to-point transmission under investigation.

Model Function

For this task, a model function of the cumulative channel impulse response is constructed. Inspired by the results of Section 2.2.2, the model function is chosen as

$$f_{\mathrm{m}}\left(t, b_1, b_2, b_3\right) = b_1 \frac{r}{d} \operatorname{erfc}\left(\frac{d-r}{(4D)^{b_2} t^{b_3}}\right), \tag{2.40}$$

where the fitting parameters b_i provide the freedom to adapt the model function to simulated and estimated cumulative channel impulse responses. The fitting parameters are to be estimated by the ANN for different system parameters. In this trivial example, the ideal fitting parameters are constant and independent of the system parameters r, d, and D. In other scenarios, such as environments with two absorbing receivers as in [DY+18], the fitting parameters vary with the system parameters.

ANN Structure

The ANN under consideration is a fully connected feed-forward neural network with three layers as depicted in Fig. 2.2. Generally, it consists of an input layer, a hidden layer, and an output layer. While the number of nodes in the input layer is equal to the number of

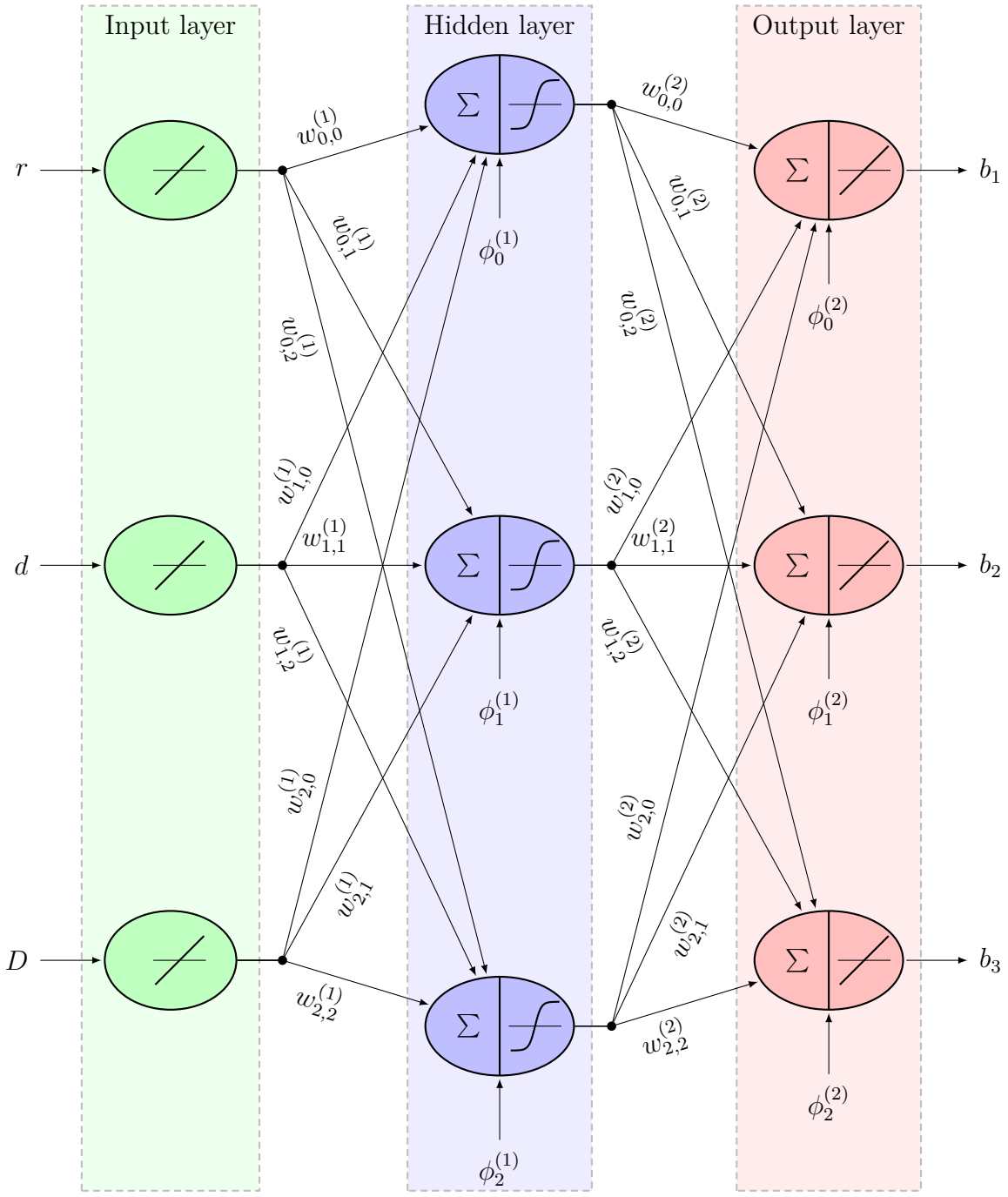


Figure 2.2: Fully connected feed-forward neural network with three layers.

system parameters, the number of nodes in the output layer is equal to the number of fitting parameters. In this trivial example, the number of nodes in the hidden layer is set to three. In practical realizations, however, more nodes in the hidden layer would be desirable to improve the estimation accuracy at the expense of an increasing computational complexity. Too many nodes, however, will degrade the estimation performance due to over-fitting [WC99]. In the sense of a fully connected feed-forward neural network, the output of each node is restricted to be a weighted input at all nodes of the following layer. As shown in Fig. 2.2, the weighting factor from node i of the previous layer to node j of the νth layer is denoted by $w_{i,j}^{(\nu)}$. The input nodes are simple neuron models with a linear transfer function which can be used to directly scale the inputs. The hidden layer nodes represent a more complex neuron model, which consists of an input and a transfer function. By forming the weighted sum of the previous node outputs and a bias $\phi_i^{(\nu)}$, the input function models the signal transmission at the synaptic clefts between axons and dendrites of biological neurons. The transfer function transforms the result of the sum following a hyperbolic tangent function. The output nodes are based on the same neuron model as the hidden nodes, with the difference that a linear transfer function is used. The weights $w_{i,j}^{(\nu)}$ and biases $\phi_i^{(\nu)}$ can be adjusted by training to change the input/output behavior of the ANN.

Training Data

To train the ANN, input/output pairs for reference parameter sets must be provided. Since, in contrast to the trivial example considered here, an analytical description of the channel is typically not possible, the random walk based channel characterization according to Section 2.2.1 is used to generate the training data. In a first step, the cumulative channel impulse response for reference parameters is determined using Monte Carlo simulation. During a second step, the corresponding output parameters $b_{\mathrm{ref},i}$ are calculated using curve fitting. In the method under investigation, nonlinear least square curve fitting [HP+13] is applied:

$$\left\{b_{\mathrm{ref},1}, b_{\mathrm{ref},2}, b_{\mathrm{ref},3}\right\} = \underset{b_1,b_2,b_3}{\arg\min} \sum_{j=1}^{N_{\Delta t}} \left(f_{\mathrm{m}}\left(j\Delta t, b_1, b_2, b_3\right) - \frac{1}{N} \sum_{i=1}^{j} y[i] \right)^2, \tag{2.41}$$

where $N_{\Delta t}$ is the number of observations in the random walk simulation. With these two steps, input/output pairs can be determined for training, which in the example consist of the fitting parameters $\left\{b_{\mathrm{ref},1}, b_{\mathrm{ref},2}, b_{\mathrm{ref},3}\right\}$ and the corresponding reference parameters r, d, and D.

Training

In the training phase, the outputs b_i of the ANN are calculated for the training parameter sets. Using the reference outputs $b_{\mathrm{ref},i}$, the quadratic error $\sum_i |b_{\mathrm{ref},i} - b_i|^2$ between the training data and the ANN outputs can be calculated, which should be minimized by adjusting the weights. It is thus an optimization problem. In the sense of Bayesian regularization [DH97; Mac92], the square of the weights $|w_{i,j}^{(\nu)}|^2$ and biases $|\phi_i^{(\nu)}|^2$ is minimized in addition to the square error to avoid over-fitting. In the ANN under investigation, the optimization is solved according to the Levenberg-Marquardt algorithm [Mar63; HM94]. The required gradients of the weights and biases are calculated by back propagation exploiting the chain rule [Dre62].

Channel Coefficients

After the ANN has been trained, it can be used to estimate the channel coefficients of unknown parameter sets. For this purpose, the ANN output b_i is calculated for the parameter set under investigation. Finally, using the model function (2.40), the channel coefficients h_ℓ can be calculated analogous to (2.38):

$$h_\ell = b_1 \frac{r}{d} \operatorname{erfc}\left(\frac{d-r}{(4D)^{b_2}((\ell+1)T)^{b_3}}\right) - b_1 \frac{r}{d} \operatorname{erfc}\left(\frac{d-r}{(4D)^{b_2}(\ell T)^{b_3}}\right). \tag{2.42}$$

2.3 Equivalent Discrete-Time Channel Model

The equivalent discrete-time channel model (EDTCM) is an exact description of the input/output behavior of a transmission system in the discrete-time domain. Similar to the transmit pulse, the physical channel, the receive filter, as well as sampling in radio-based communication, the release of molecules at the transmitter, the propagation of molecules through the environment, and the absorption with summation of molecules at the receiver can be described by an EDTCM [DK$^+$17a]. This simplifies the understanding and the adaptation of classical transmission algorithms to the diffusion-based transmission channel. It also allows a significant reduction of computational complexity in contrast to random walk based propagation simulation, especially with long channel memory lengths, large number of molecules, and long info sequences. The molecular EDTCM differs from the classical EDTCM in the sense that the molecules propagate stochastically in contrast to the deterministically propagating radio waves. In addition to the exact EDTCM, which is based on Poisson binomial distributions, approximations are also considered, which simplify

the understanding and description of the diffusion-based channel. The following derivations of the channel models as well as their analysis and evaluation are based on the results presented in [DK+17a].

2.3.1 Poisson Binomial EDTCM

The absorption of a molecule at the receiver in the ℓth transmission interval after its release can be described by a Bernoulli trial. This consists of two possible events, either the molecule hits the receiver and gets absorbed, or it does not hit the receiver and diffuses further through the environment. As shown in the previous sections, the channel coefficient h_ℓ is set equal to the probability that a molecule will be absorbed by the receiver in the ℓth transmission interval. Consequently, a molecule does not hit the receiver with a probability of $1 - h_\ell$. If not only a single molecule but a number of N independently propagating molecules with the same absorption probability are released at the transmitter at the same time, the N Bernoulli trials can be described by the binomial distribution $\mathcal{B}(N, h_\ell)$, where N is the number of trials and h_ℓ is the success probability. Consequently, the total number of molecules absorbed by the receiver at discrete-time step k is the sum of binomial distributions, which is described by a Poisson binomial distribution:

$$y[k] = \sum_{\ell=0}^{L} z_\ell, \qquad \text{with} \quad z_\ell \sim \mathcal{B}(x[k-\ell], h_\ell)\,, \tag{2.43}$$

where L is the channel memory length and $x[k]$ is the transmitted number of molecules in the kth transmission interval. If $x[k]$ is interpreted as a transmit symbol, (2.43) describes

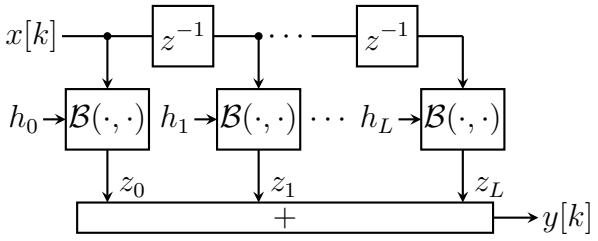


Figure 2.3: Block diagram of the Poisson binomial equivalent discrete-time channel model.

the EDTCM shown in Fig. 2.3. The blocks with binomial distributions represent random number generators that follow the corresponding distribution.

2.3.2 Binomial Approximated EDTCM

Especially for simulations with long channel memory length L it is very time-consuming to generate binomial distributed random values for each branch in Fig 2.3. Furthermore, arithmetic expressions with Poisson binomial distributions quickly become very complicated. Therefore, different approximations of the Poisson binomial distributed EDTCM are presented in the following, whose approximations are suitable under different conditions. The sum of binomial distributions with different success probabilities cannot be described exactly with a single binomial distribution. However, if the individual success probabilities h_ℓ are small enough and the number of attempts N is high enough, the EDTCM shown in Fig 2.3 can be approximated by a single binomial distribution:

$$y[k] \sim \mathcal{B}\left(N, \sum_{\ell=0}^{L} u[k-\ell]h_\ell\right), \tag{2.44}$$

where $u[k]$ is the kth info bit. Note that this approximation is only suitable for binary on-off keying (OOK) modulation where $u[k] = 1$ is represented by releasing N molecules and $u[k] = 0$ is represented by releasing no molecules. The following two approximations in the next sections are suitable for any modulation scheme. The resulting binomial approximated EDTCM is shown in Fig 2.4.

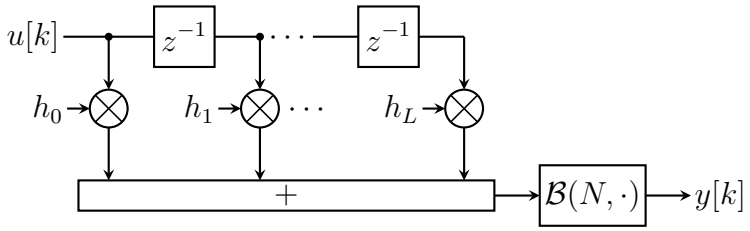


Figure 2.4: Block diagram of the binomial approximated equivalent discrete-time channel model.

2.3.3 Gaussian Approximated EDTCM

If the number of trials is sufficiently large and the probability of success is not close to zero or one, then the binomial distribution can be well approximated by a Gaussian distribution [KY+10]. The relationship of the approximation is given by

$$\mathcal{B}(N, h_\ell) \approx \mathcal{N}\left(\mu_\ell = Nh_\ell, \sigma_\ell^2 = Nh_\ell(1-h_\ell)\right), \tag{2.45}$$

where μ_ℓ is the mean and σ_ℓ^2 is the variance of the Gaussian distribution. Since the sum of independent Gaussian distributions can be described by a single Gaussian distribution, $y[k]$ can be approximated as

$$y[k] \approx z[k] + \omega[k], \tag{2.46}$$

with

$$z[k] = \sum_{\ell=0}^{L} x[k-\ell]h_\ell, \tag{2.47}$$

$$\omega[k] \sim \mathcal{N}\left(0, \sigma^2[k] = \sum_{\ell=0}^{L} x[k-\ell]h_\ell(1-h_\ell)\right). \tag{2.48}$$

Thereby, $z[k]$ represents the expected number of molecules at the receiver. The zero-mean additive amplitude-dependent noise $\omega[k]$ models the uncertainty of the number of molecules arriving at the receiver, which is known as Brownian noise in the diffusion-based channel [SM+12]. The resulting Gaussian approximated EDTCM is shown in Fig. 2.5. It should be

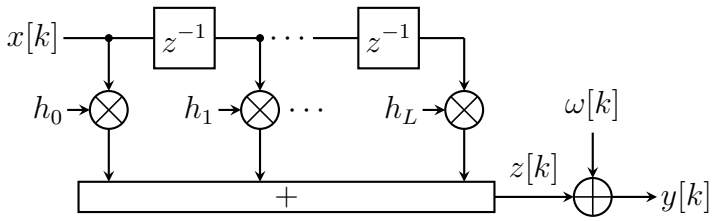


Figure 2.5: Block diagram of the Gaussian approximated equivalent discrete-time channel model.

noted that the structure is the same as that of the classical additive white Gaussian noise channel with memory. The only difference is that the additive noise in the diffusion-based molecular channel is amplitude-dependent. The amplitude-dependent variance $\sigma^2[k]$ of the noise can be efficiently calculated by two convolutions as shown in Fig. 2.6. Since the number of molecules received is a non-negative integer, but the Gaussian distribution is defined as a continuous distribution over all real numbers, $y[k]$ has to be post-processed. By rounding to nearest integer and limiting the minimum number to zero,

$$y'[k] = \lfloor \max\{0, y[k]\} \rceil, \tag{2.49}$$

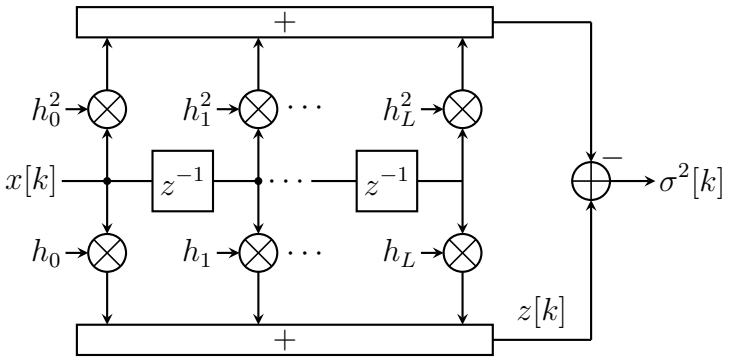


Figure 2.6: Variance calculation for the Gaussian approximated equivalent discrete-time channel model.

$y'[k]$ becomes a non-negative integer. This post-processing step leads to an increasing approximation error when the mean value of $y[k]$ is close to zero.

2.3.4 Poisson Approximated EDTCM

Especially for unlikely events, the binomial distribution can be well approximated by the Poisson distribution. The approximated relationship of the distributions is given as

$$\mathcal{B}(N, h_\ell) \approx \mathcal{P}(\lambda_\ell = N h_\ell)\,, \tag{2.50}$$

where λ_ℓ is the event rate describing the mean and the variance of the Poisson distribution. Similar to Gaussian distributions, the sum of independent Poisson distributions can be described by a single Poisson distribution whose event rate is the sum of the individual event rates. Therefore, $y[k]$ is approximated by

$$y[k] \sim \mathcal{P}\left(\lambda[k] = z[k] = \sum_{\ell=0}^{L} x[k-\ell] h_\ell\right). \tag{2.51}$$

Since the Poisson distribution is a discrete probability distribution which is only defined for non-negative integer values, no post-processing is necessary in contrast to the Gaussian approximated EDTCM. The resulting Poisson approximated EDTCM is shown in Fig. 2.7.

2.3.5 Complexity Discussion

For the realization of the different EDTCMs, from Section 2.3.1 to Section 2.3.4 parameters such as mean or variance have to be calculated from the channel coefficients h_ℓ and the

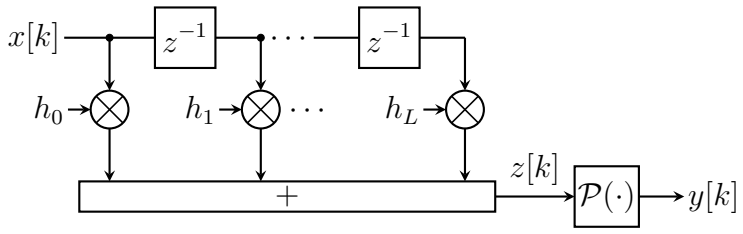


Figure 2.7: Block diagram of the Poisson approximated equivalent discrete-time channel model.

number of released molecules $x[k]$. These parameters, however, can be efficiently calculated by convolution or can be stored in lookup tables. Therefore, the effort for calculating these parameters can be neglected when considering the overall computational complexity of the models. The main contribution to computational complexity arises from the random number generators (RNGs). There exist efficient algorithms for Gaussian RNG, Poisson RNG [AD82] and binomial RNG [KS88] with a constant computational complexity $\mathcal{O}(1)$ independent of N and h_ℓ. As a result, the computational complexity of all approximated EDTCMs is $\mathcal{O}(1)$, whereas the Poisson binomial EDTCM has a computational complexity of $\mathcal{O}(L+1)$. Thus, the approximated channel models reduce the computational complexity by a factor proportional to the $L+1$ channel coefficients.

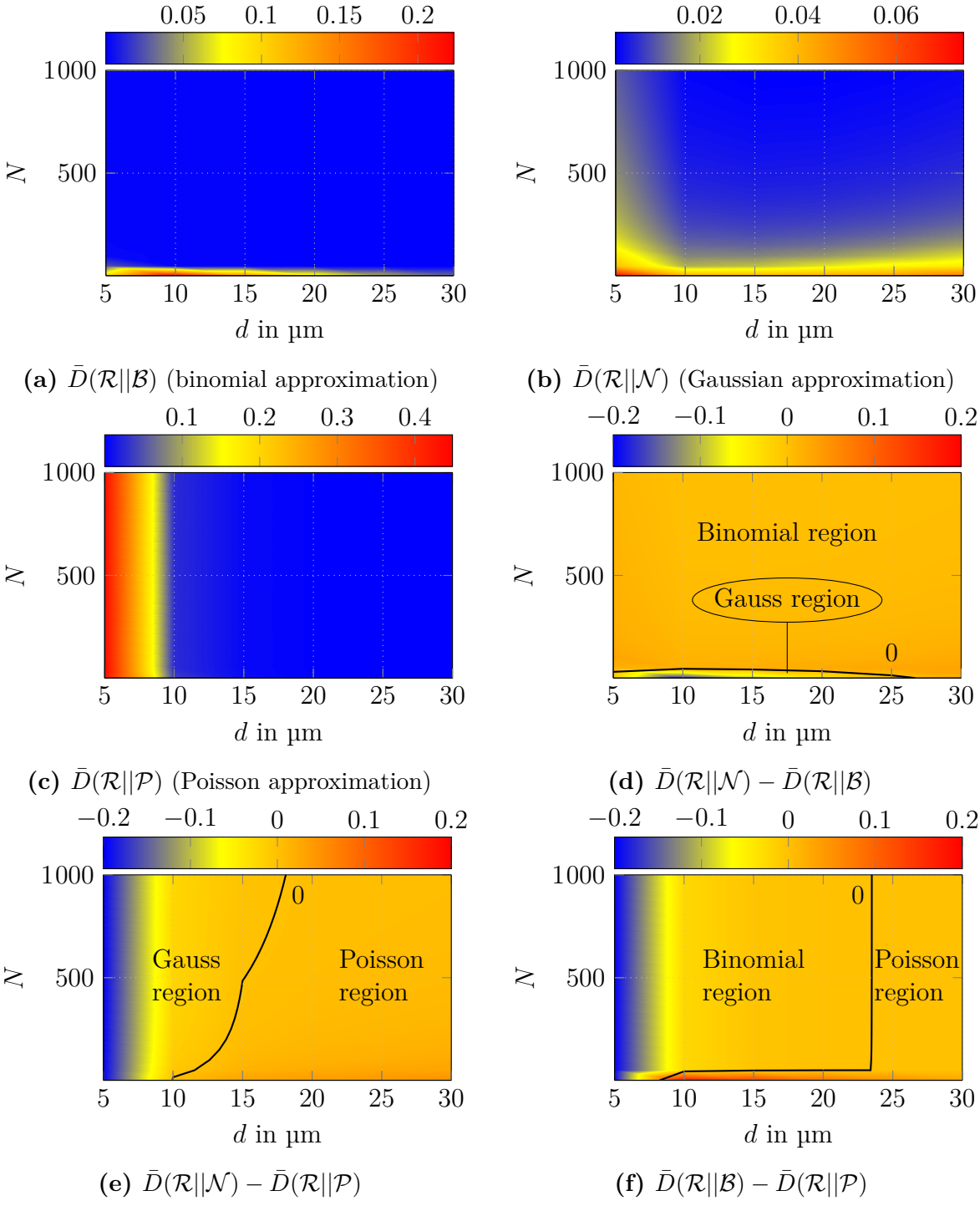
2.3.6 Approximation Accuracy Analysis

The channel models introduced in Section 2.3.2, Section 2.3.3, and Section 2.3.4 approximate the sum of binomial random variables defined for the exact Poisson binomial channel model introduced in Section 2.3.1. The accuracy of approximation can be analyzed by investigating the mean Kullback-Leibler divergence (MKLD) [KL51] between the Poisson binomial probability mass function $P_{\mathcal{R}}(n)$ and the proposed approximations $P_{\mathcal{B}/\mathcal{N}/\mathcal{P}}(n)$ over all possible input sequences $\boldsymbol{s}$:

$$\bar{D}(\mathcal{R}||\mathcal{B}/\mathcal{N}/\mathcal{P}) = \frac{1}{2^{L+1}} \sum_{\boldsymbol{s}\in\{0,1\}^{L+1}} \left[\sum_{n=0}^{w_{\mathrm{H}}(\boldsymbol{s})N} P_{\mathcal{R}}(n) \log_2 \left(\frac{P_{\mathcal{R}}(n)}{P_{\mathcal{B}/\mathcal{N}/\mathcal{P}}(n)} \right) \right], \tag{2.52}$$

where $w_{\mathrm{H}}(\cdot)$ is the Hamming weight. It should be noted that, for the sake of simplicity, only binary OOK modulation is considered in which $u[k]=1$ is represented by the release of $x[k]=N$ molecules and $u[k]=0$ by the release of $x[k]=0$ molecules.

In Fig. 2.8(a)-2.8(c) the MKLDs for the system parameters given in Sec. C.1 are shown.



(a) $\bar{D}(\mathcal{R}||\mathcal{B})$ (binomial approximation)

(b) $\bar{D}(\mathcal{R}||\mathcal{N})$ (Gaussian approximation)

(c) $\bar{D}(\mathcal{R}||\mathcal{P})$ (Poisson approximation)

(d) $\bar{D}(\mathcal{R}||\mathcal{N}) - \bar{D}(\mathcal{R}||\mathcal{B})$

(e) $\bar{D}(\mathcal{R}||\mathcal{N}) - \bar{D}(\mathcal{R}||\mathcal{P})$

(f) $\bar{D}(\mathcal{R}||\mathcal{B}) - \bar{D}(\mathcal{R}||\mathcal{P})$

Figure 2.8: Mean Kullback-Leibler divergence between Poisson binomial distribution and approximations. Parameters according to Appendix C.1.

For binomial approximation, the MKLD decreases for an increasing distance, which is equivalent to a decreasing absorption probability. Furthermore, it decreases with an increasing number of released molecules, which also applies to the Gaussian approximation. An increasing number of released molecules leads to a decreasing uncertainty of the channel outputs, since the variance becomes smaller relatively to the mean value. It can be seen that Gaussian approximation is worse if the absorption probabilities are close to zero (large distances) or one (small distances). For the Poisson approximation, the MKLD decreases for an increasing distance. This confirms the statement that the Poisson approximation fits best for unlikely events.

In Fig. 2.8(d)-2.8(f), the differences between MKLDs of the different approximations are shown. Consequently, it is possible to define regions in which one approximation is more accurate than another. In the scenario under investigation, however, the differences between the MKLDs are often only small. For a small number of released molecules it can be stated that binomial approximation leads to higher approximation errors. In addition, Poisson approximation should be avoided for short distances. Without these two exceptions, however, all three approximations lead to fair approximations of the Poisson binomial channel model. A practical comparison by means of bit error simulations confirms this result in [DK$^+$17a].

2.4 Testbed Verification

In addition to theoretical models, the channel can also be characterized by real-world experiments. For this purpose, a testbed based on alcohol molecules is constructed as a proof of concept. This setup is based on the setup presented in [KF$^+$15]. On the transmitter side there is an electronically controllable sprayer filled with an alcoholic liquid. In the setup presented here, isopropanol is used which is diluted with water to a desired concentration. By opening and closing the valve, the information to be transmitted can be modulated by the sprayer. In order to accelerate the transmission process and increase the transmission range, an optional fan is placed behind the sprayer. On the receiver side, an alcohol sensor is used to measure the change in concentration in its environment. With the help of an *Analog Discovery* the sprayer is controlled and the alcohol sensor is read out at the same time. The transmission medium between transmitter and receiver is air. Spatial limits are only given by the room in which the measurements are taken. In the following, the main components of the testbed are explained in detail.

Sprayer

Figure 2.9: *Gloria Type 89* pressure sprayer connected with solenoid valve.

The *Gloria Type 89* pressure sprayer, shown in Fig. 2.9, is used for emitting the alcohol molecules on the transmitter side. It provides a maximum filling capacity of 1 L. With the integrated pump a maximum pressure of 3 bar can be generated inside the container. The liquid is sprayed through a pivoting hollow-cone nozzle with a spraying angle of 60°. At a pressure level of 3 bar, the output rate is 0.54 L/min and droplets with an average size of 3.3 µm are generated. In order to be able to electronically control the on and off switching of the sprayer, a 12 V solenoid valve is placed in front of the nozzle, which is closed at idle state. This valve tolerates a maximum pressure of 10 bar and is controlled by an *Analog Discovery* and a control circuit. The minimum duration of a switching cycle including complete opening and closing of the valve is in the order of 50 ms.

Optional Fan

To speed up the transmission process from the transmitter to the receiver and to increase the transmission range, an optional fan is positioned behind the sprayer. The fan, shown in Fig. 2.10, is a tabletop fan of type *ChillOut GF601E* with rotor blades. It has a maximum power consumption of 31 W and has two control levels for the rotation speed of the blades. In the results presented here, the fan is only operated at its highest speed.

Figure 2.10: *ChillOut GF601E* tabletop fan with rotor blades.

Alcohol Sensor

An *MQ-3* gas sensor [MQ3] is used to measure the alcohol concentration in the environment of the receiver. The sensor is shown in Fig. 2.11. It is a metal-oxide sensor based on a

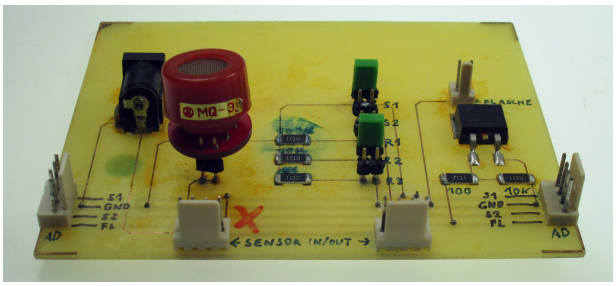


Figure 2.11: *MQ-3* gas sensor with control board.

resistive measuring principle. The sensitive layer consists of tin dioxide and has a diameter of 16.8 mm. By heating with the help of a 5 V direct or alternating current operated heating coil, this layer becomes semiconducting. If alcohol molecules encounter the sensitive layer, an oxidization reaction takes place. This results in a change of the resistance. If the alcohol concentration increases, the sensitive layer becomes more conductive and the resistance decreases. A decreasing alcohol concentration leads to an increase of the resistance. Sensitivity depends on temperature and humidity [MQ3]. In addition to alcohol, the sensor is

sensitive to other volatile liquids and gases as well. Fig. 2.12 shows the *MQ-3* gas sensor sensitivity characteristics for several gases.

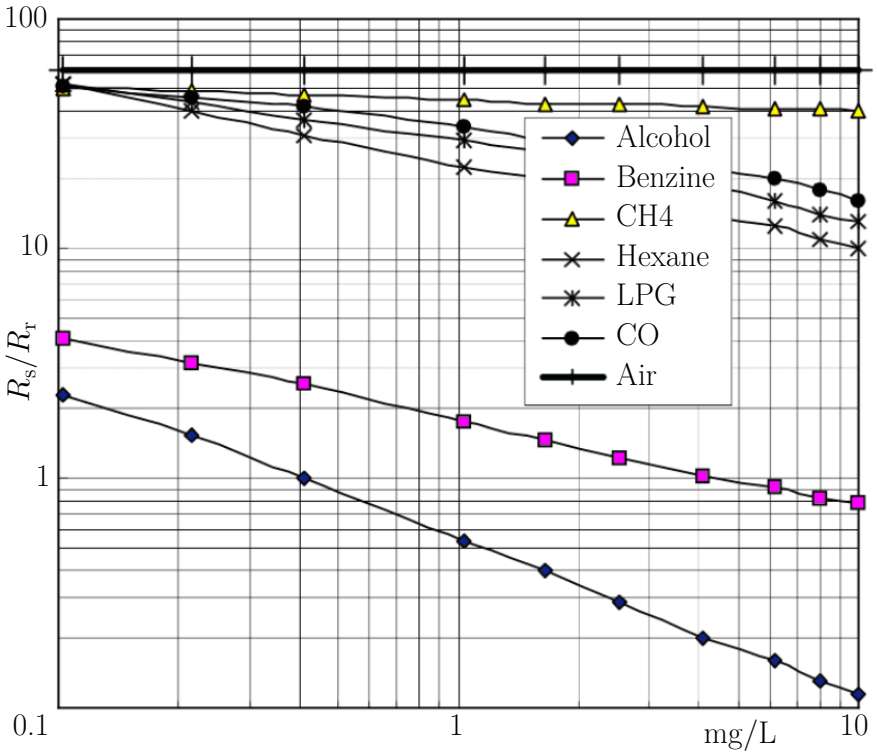


Figure 2.12: *MQ-3* gas sensor sensitivity characteristics for several gases, where R_s is the sensor resistance at various concentration of gases and R_r is the sensor resistance at 0.4 mg/L of alcohol in clean air [MQ3].

The simplified relationship between alcohol concentration ϕ and the resistance R_s of the sensitive layer can be generally described by

$$R_s(t) \propto R_r \phi(t)^{b_0}, \tag{2.53}$$

$$\log\left(\frac{R_s(t)}{R_r}\right) \propto b_0 \log\left(\phi(t)\right), \tag{2.54}$$

where R_r is the sensor resistance at a reference point of concentration. The approximately constant parameter b_0 is according to (2.54) equal to the gradient of the curve in Fig. 2.12 and thus $b_0 \approx -0.65$. To determine the alcohol concentration ϕ, the voltage U_L is measured over a constant load resistance. This voltage is inversely proportional to the resistance R_s

of the sensor. Consequently, the relationship between U_L and ϕ is

$$U_\mathrm{L}(t) \propto \phi(t)^{-b_0}, \tag{2.55}$$

$$U_\mathrm{L}(t)^{-\frac{1}{b_0}} \propto \phi(t) \approx \frac{b_1 d}{\sqrt{4\pi b_2 t}} \exp\left(-\frac{(d-b_3 t)^2}{4 b_2 t}\right), \tag{2.56}$$

where the general shape is given by the solution of the one-dimensional unbounded advection-diffusion equation [FY+16] and b_1, b_2, b_3 are unknown parameters that can be determined by curve fitting on measurement data. Nevertheless, all fitting parameters describe a major effect on the channel impulse response. The number of molecules released at the transmitter is described by b_1. The unknown diffusion coefficient and the influence of turbulence can be modulated by b_2. The drift caused by the initial velocity of the alcohol molecules at the outlet of the sprayer and by the application of a fan is covered by b_3. The spatial dimensions of the receiver in relation to the transmission distance are assumed to be negligibly small and are therefore not considered in (2.56).

Analog Discovery

Figure 2.13: *Analog Discovery* device.

The *Analog Discovery* manufactured by *Digilent* and depicted in Fig. 2.13, is a compact universal serial bus based device that combines many useful laboratory instruments. Besides power supplies and digital inputs and outputs it offers a waveform generator and a digital oscilloscope. The latter two handle the control of the sprayer as well as the readout of the sensor in the testbed. The solenoid valve of the sprayer is controlled by a

metal–oxide–semiconductor field-effect transistor through an analog output of the *Analog Discovery*. The analog output has a resolution of 14 bit and offers up to ±5 V at a sampling rate of 100 megasamples per second at 5 MHz bandwidth. The sensor voltage is recorded via oscilloscope input. The input offers ±25 V at a resolution of 14 bit with a sampling rate of up to 100 megasamples per second at 5 MHz bandwidth.

2.4.1 Experimental Determination of the Channel Impulse Response

In the following, the experimental determination of a channel impulse response is performed. The distances between the individual devices are specified as shown in Fig. 2.14. A 3.5 %

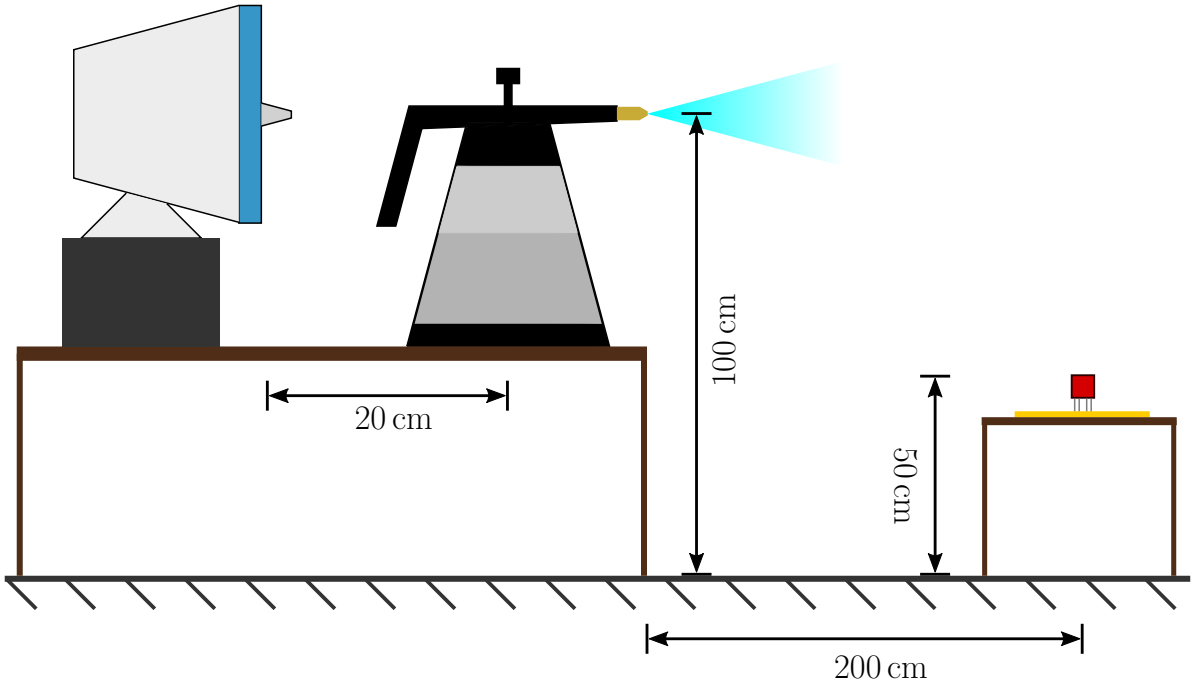


Figure 2.14: Experimental testbed setup.

alcohol solution is used inside the sprayer. The spraying time for a pulse representing the info bit $u = 1$ is fixed at 500 ms. The sampling frequency of the *Analog Discovery* is fixed at 100 Hz. A total of 10 impulse responses are measured and averaged to determine the average channel impulse response. It should be noted that the channel impulse response in the testbed is interpreted proportional to the alcohol concentration at the receiver and not the first passage time density. The result is shown in Fig. 2.15. It can be observed that the measured channel impulse responses follow the typical characteristic of a diffusion-based molecular communication channel impulse response. After a comparably steep rise and reaching its maximum, a slowly decaying tail follows. In addition, significant deviations can

be seen between the individual measurement runs, which could be described by a stochastic channel impulse response. There are many possible reasons for these deviations. Besides the pure diffusion process, macroscopic effects are rather conceivable in the testbed. These include inaccuracies during the spraying process, turbulent air flow caused by the fan, air circulation in the room, as well as temperature and humidity changes.

To determine a theoretical channel impulse response, the fitting parameters of (2.56) are obtained by nonlinear least square curve fitting applying the principle of (2.41). The

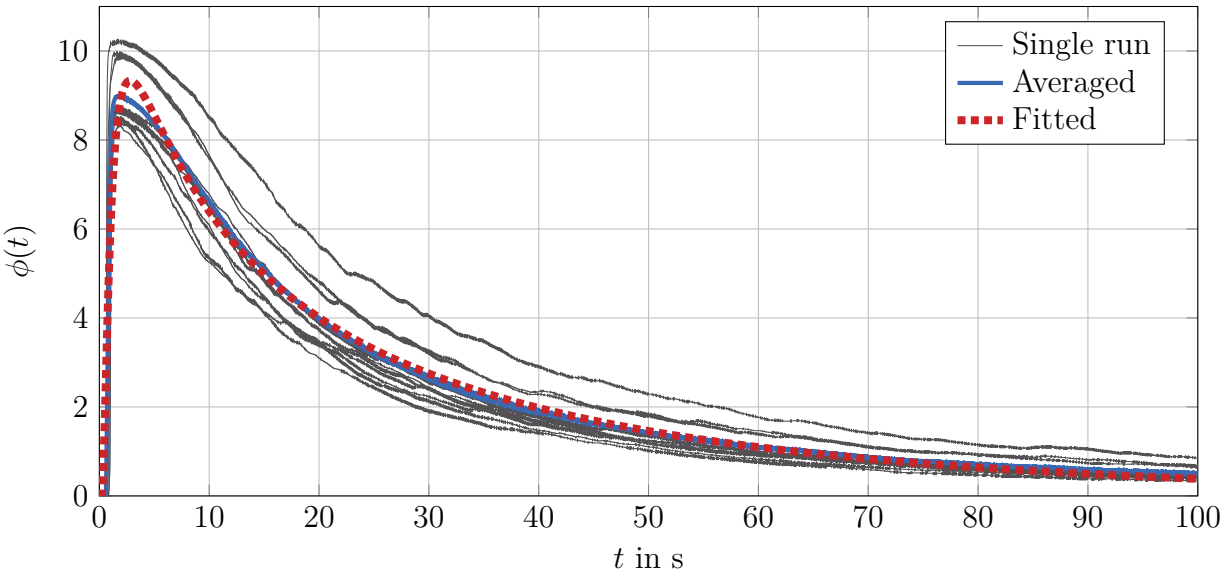


Figure 2.15: Measured and fitted channel impulse response of testbed (Fitting parameters: $b_1 = 28.73$, $b_2 = 0.64$, $b_3 = 0.23$).

resulting channel impulse response with its fitting parameters is shown in Fig. 2.15 as well. The fitted curve particularly describes the area of the rise and the decaying tail well. Only the region around the maximum is less well modeled and could be improved by introducing more fitting parameters. However, the root-mean-square error (RMSE) over the shown time range is around 0.196 and is assumed to be sufficiently good.

2.4.2 Experimental Sequence Transmission

To check the accuracy of the determined channel impulse response, the example sequence $\boldsymbol{u} = [1\,0\,0\,1\,0\,0\,1\,1\,0\,0\,0\,0\,1\,1\,1\,0\,1\,0\,1\,0\,0]$ is transmitted, which represents the message "ICT" according to the american standard code for information interchange code. The modulation follows the OOK described in Section 3.1.1. The symbol duration is set to $T = 4\,\mathrm{s}$.

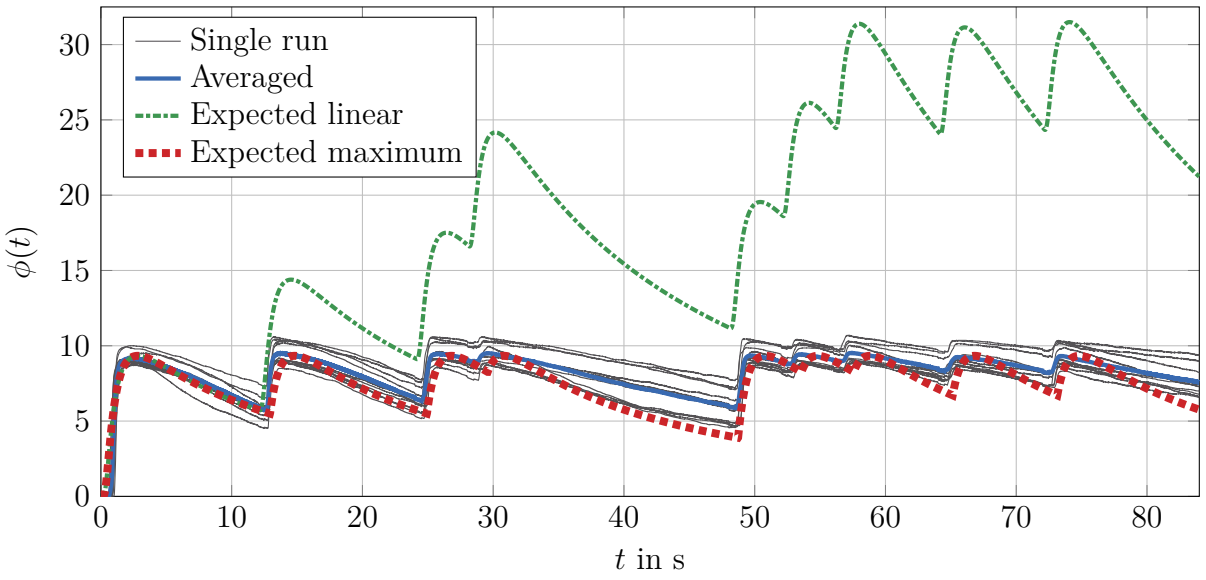


Figure 2.16: Transmission of text message "ICT".

This transmission is also repeated 10 times. The individually recorded and the averaged measurement curves are shown in Fig. 2.16. In addition, the expected received sequence from the theoretically determined channel impulse response is shown. As for the measured channel impulse responses, the individual measurement curves deviate from each other, but generally agree well. The transmitted bit sequence can also be clearly detected in all measurements by the nine local maxima in the curve. Looking at the theoretically expected received sequence (dash-dotted line), it becomes clear that the general shape is also well in agreement with its maxima, but a certain offset can be observed after the first pulse. This offset indicates that the channel impulse responses cannot be superimposed linearly to reconstruct the measured values. There are several possible reasons for this nonlinear superposition. On the one hand, the sensor saturation and the nonlinear sensor characteristic can cause this effect. On the other hand, it is possible that the decaying tail is caused rather by a sensor resume time than by the residual alcohol particles in the environment. The sensor resume time describes the time that a sensor needs to get back to its original state. In addition, a concentration and not the absolute number of particles is measured at the sensor. To compensate for the nonlinear superposition effect, the maximum values instead of the superposition of the individual impulse responses is selected in the red dashed curve. This results in a better match between the measured and expected signals. However, this also leads to a slight temporal shift of the local maxima, as it becomes visible after the first impulse. The RMSE between the expected and averaged

measured curve is around 1.0324.

In conclusion, the presented testbed is suitable as proof of concept, since the measured average impulse response can be described by a modified impulse response of the diffusion-based channel. Nevertheless, there are many additional factors that influence the experiment and the obtained measurements. These are, for example, the nonlinear superposition at the sensor, temperature and humidity fluctuations in the room, turbulences, or the influence of gravity.

3

Transmitter and Receiver Algorithms

On the basis of the system and channel model introduced in Section 2.1, various transmitter and receiver algorithms are presented and analyzed in detail in this chapter. On the transmitter side, different modulation schemes are considered. On the receiver side, the focus is on channel equalization and detection algorithms. Furthermore, orthogonal frequency-division multiplexing (OFDM) is used to adapt a multi-carrier approach for molecular communication.

3.1 Modulation Schemes

3.1.1 On-Off Keying

In wireless communication, on-off keying (OOK) is a special case of binary amplitude shift keying (ASK) modulation, where the information is represented by switching the signal on or off. As it does not consume any transmission energy when switched off, it is particularly suitable for energy-limited battery-operated devices such as remote controls. It also produces a unipolar output signal, which is needed, for instance, in optical transmission. In nature, such kind of information modulation can be observed at the synaptic cleft of chemical synapses. Only when the presynaptic cell is excited by an action potential, it releases neurotransmitters into the synaptic cleft, which diffuse as information carriers to the postsynaptic cell. Motivated by this, in the sense of OOK in molecular communication, binary information is represented by the presence or absence of molecule release at the transmitter [MM$^+$10]. In this dissertation, it is assumed that the transmitter releases either N or no molecules at the beginning of a transmission interval to represent an info bit.

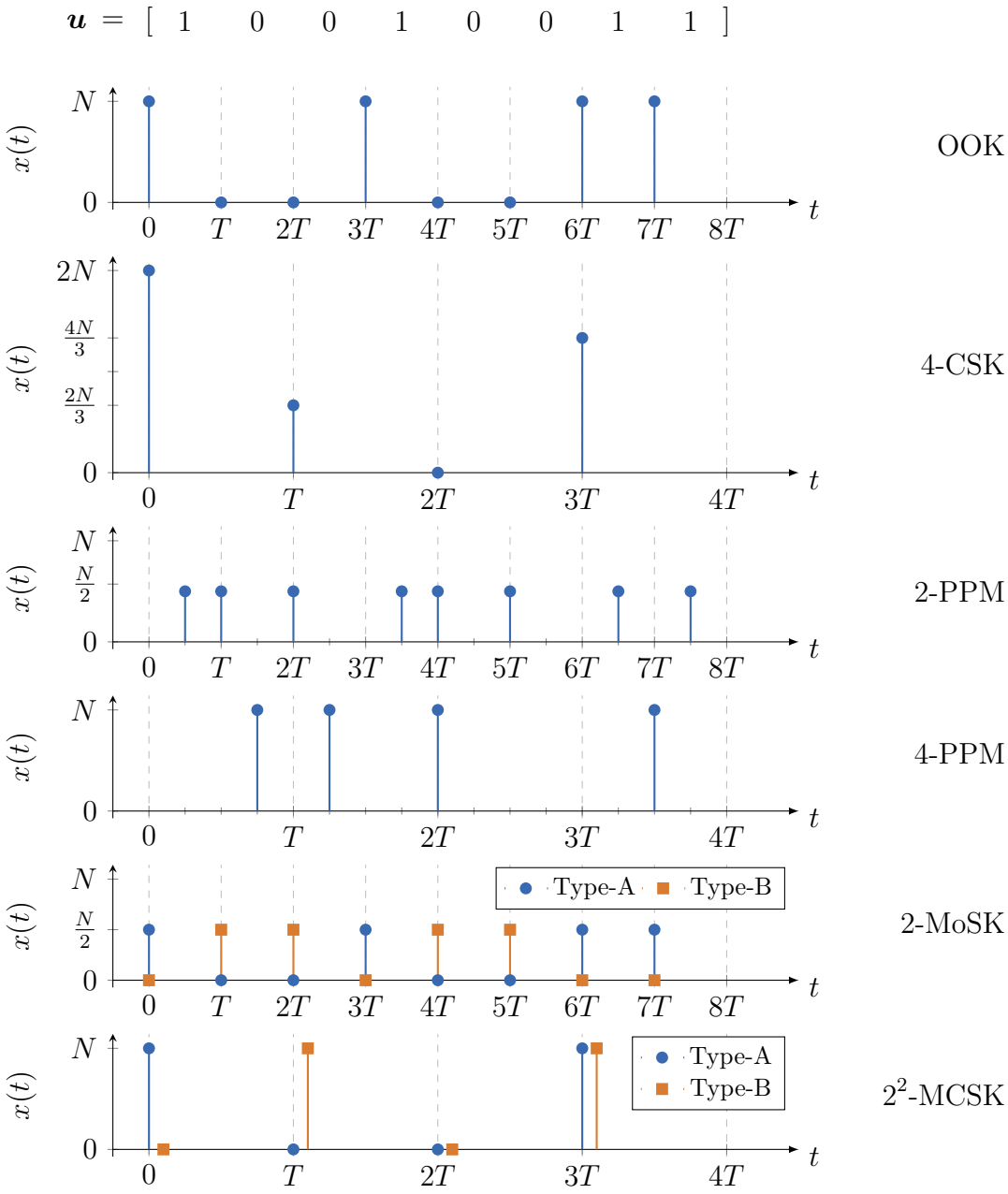


Figure 3.1: Overview of different modulation schemes.

Hence, the discrete-time representation of the transmitted OOK modulated symbol is

$$x[k] = \begin{cases} 0 & \text{if } u[k] = 0, \\ N & \text{if } u[k] = 1. \end{cases} \tag{3.1}$$

Since it is a binary modulation scheme, the bit duration is equal to the symbol duration $T_\mathrm{b} = T$. For the transmission of bit sequence $\boldsymbol{u} = [1\,0\;1\,0\,0\;1\,1]$, the corresponding OOK-modulated transmit signal is depicted in Fig. 3.1. In molecular communication, the number of molecules emitted is proportional to the energy required [GD+18]. Thus the average number of released molecules per bit is used as bit energy E_b, and the averaged number of released molecules per symbol is used as symbol energy E_s throughout this dissertation. Under the assumption of equally probable info bits, OOK results in

$$E_\mathrm{s} = \sum_{x \in \mathcal{X}} p(x)x = \frac{1}{2}(0 + N) = \frac{N}{2} = E_\mathrm{b}, \tag{3.2}$$

where $\mathcal{X}$ is the symbol alphabet.

Theoretical Symbol Error Rate and Bit Error Rate

For the sake of simplicity, the case without intersymbol interference (ISI) is considered for the derivation of the theoretical bit error rate (BER) and symbol error rate (SER). Consequently, there is only a single channel coefficient h, which is equivalent to the absorption probability of a molecule at the receiver during the transmission interval of its release. Fig. 3.2 shows an example of the probability distribution functions $P(y|x)$ of the two OOK symbols at the receiver. In this simple case, $\hat{u}[k] = 1$ is detected when a molecule reaches the receiver within the transmission interval, i. e. $y[k] > 0$. If none is absorbed, $y[k] = 0$, $\hat{u}[k] = 0$ is decided. This results in the theoretical BER

$$P_\mathrm{b} = \sum_{x \in \mathcal{X}} P(x)P(\hat{x} \neq x|x) = \frac{1}{2}\left[(1 - \mathcal{B}(y \leq 0\,|0,1)) + \mathcal{B}(y \leq 0\,|N,h\,)\right] = \frac{(1-h)^N}{2}, \tag{3.3}$$

where $\mathcal{B}(y \leq Y\,|N,p) = \sum_{y=0}^{Y} \binom{N}{y} p^y (1-p)^{N-y}$ is the cumulative binomial distribution and $\mathcal{B}(0,1)$ is a unit impulse. Since OOK is a binary modulation scheme, the BER is equal to the SER, i. e. $P_\mathrm{b} = P_\mathrm{s}$. Fig. 3.3 shows numerically simulated and theoretically calculated BER for different h, where the results of both match well.

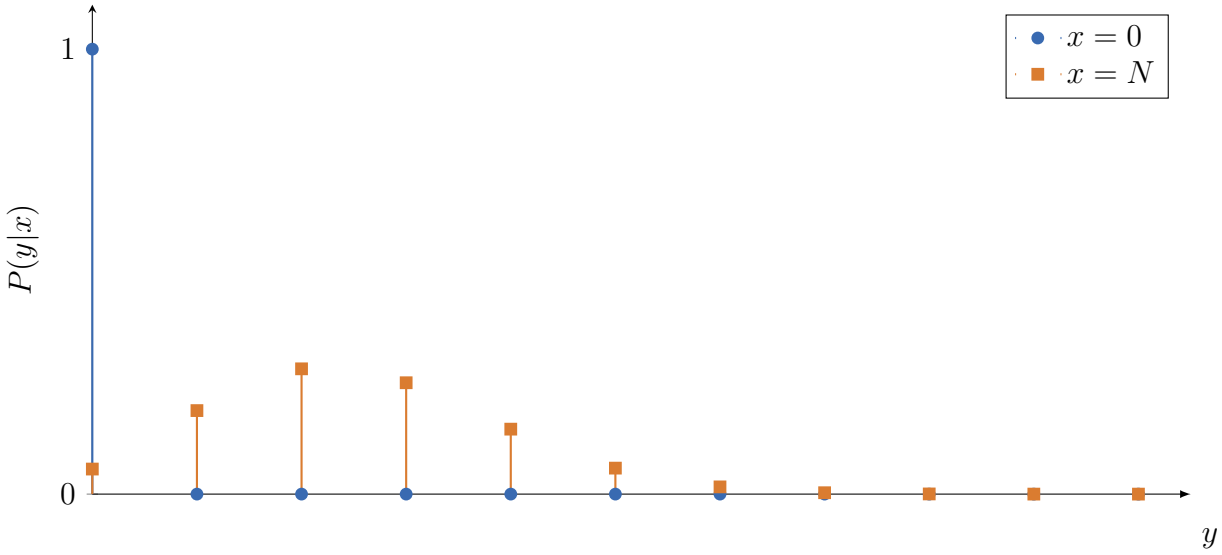


Figure 3.2: Symbol probability distribution for OOK.

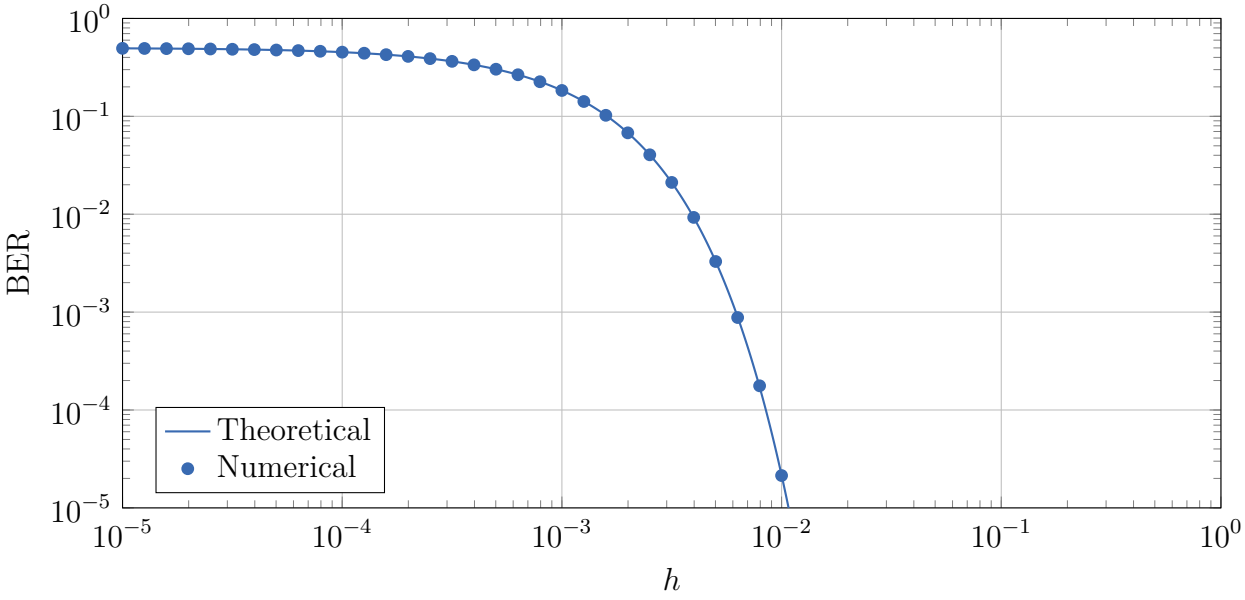


Figure 3.3: Bit error rate of OOK. Parameters according to Appendix C.2.

3.1.2 Concentration Shift Keying

For higher-order modulation, OOK can be extended by several concentration levels. This leads to the so-called concentration shift keying (CSK) [KY+11], which is similar to ASK in classical communication. The individual symbols are arranged equidistantly, which leads to a minimum BER for the additive white Gaussian noise channel with equiprobable symbols. The symbols can also be arranged not equidistantly, which is discussed in Section 3.1.3. In order to make different modulation schemes comparable, OOK is used as a reference. In the sense of a fixed bit rate, the number of bits per second are kept constant in all modulation schemes. In addition, an energy normalization on E_b is performed. Thus the average number of released molecules per bit is set constant to $E_\mathrm{b} = N/2$ for all modulation schemes.

In M-CSK, there are M different symbols which represent $m = \log_2(M)$ bits, where M is the modulation order. Assuming that the lowest concentration value is zero, the distance α between two neighboring symbols in M-CSK can be obtained by normalizing E_b:

$$E_\mathrm{b} = \frac{E_\mathrm{s}}{m} = \frac{1}{m}\sum_{x\in\mathcal{X}} p(x)x \tag{3.4}$$

$$= \frac{1}{mM}\sum_{i=0}^{M-1} \alpha i = \frac{\alpha}{mM}\left(\frac{(M-1)M}{2}\right) \tag{3.5}$$

$$= \frac{\alpha(M-1)}{2m} \overset{!}{=} \frac{N}{2} \tag{3.6}$$

$$\alpha = \frac{mN}{(M-1)}. \tag{3.7}$$

Consequently, the symbol alphabet for M-CSK is defined as $\mathcal{X} = \{x^{(0)}, x^{(1)}, \ldots, x^{(M-1)}\} = \{0, \alpha, \ldots, (M-1)\alpha\}$. Note that α is not necessarily an integer. In this case rounding must be applied. The deviation in E_b that can occur as a consequence of rounding is assumed to be negligible. The bit to symbol mapping in this dissertation is done following Gray labeling, which minimizes the BER in the uncoded case. An example of the resulting transmit signal for 4-CSK is shown in Fig. 3.1.

Theoretical Symbol Error Rate and Bit Error Rate

The theoretical BER for 4-CSK is demonstratively derived. For higher-order CSK, the derivation can be extended accordingly. Fig 3.4 shows an example of the symbol probability distributions $P(y|x)$ at the receiver, which consists of a unit pulse $\mathcal{B}(0,1)$ at zero and three binomial distributions $\mathcal{B}(2N/3,h)$, $\mathcal{B}(4N/3,h)$, and $\mathcal{B}(2N,h)$. In total, there are

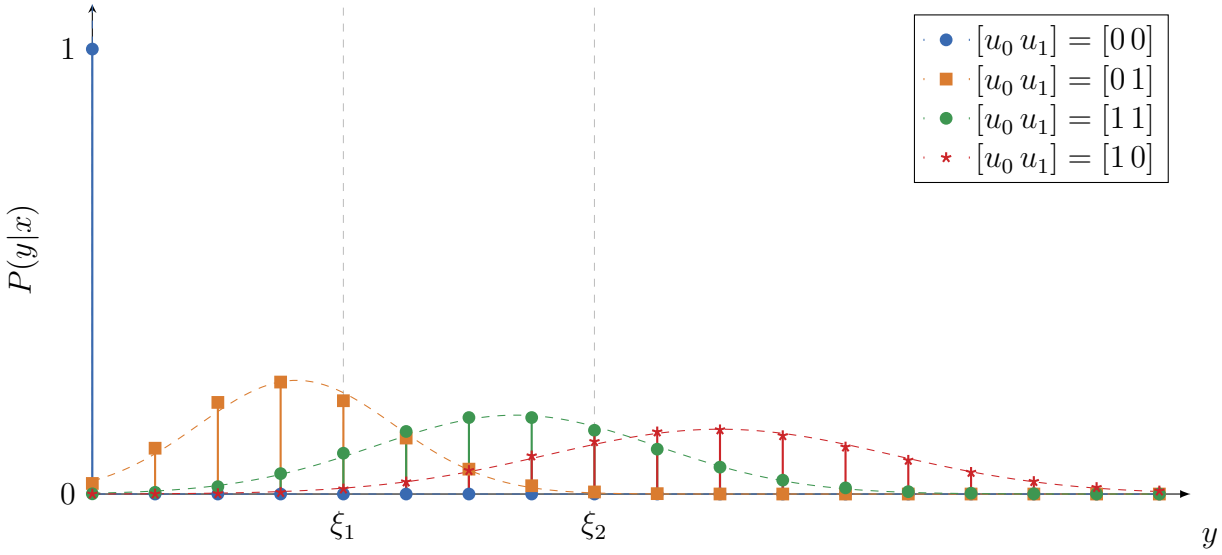


Figure 3.4: Symbol probability distribution for CSK.

three threshold values, 0, ξ_1, and ξ_2, to distinguish the four symbols that result from the intersections of the individual distributions. Due to the binomial coefficient, the intersection of two binomial distributions cannot be calculated easily, but can be determined numerically. If the binomial distributions are approximated by Gaussian distributions (as shown in Fig 3.4 by dashed lines), the intersection between Gaussian distributions can be calculated as shown in Appendix B.3. The approximation by a Gaussian distribution is particularly precise for large sample sizes and when the success probabilities are not close to zero or one [KY+11]. This result can be used to simplify the numerical determination of the intersection between binomial distributions. Taking the threshold values into account, the theoretical SER results in

$$P_{\mathrm{s}}^{[0\,0]} = 0, \tag{3.8}$$

$$P_{\mathrm{s}}^{[0\,1]} = \mathcal{B}\left(y \leq 0 \,\middle|\, \frac{2N}{3}, h\right) + 1 - \mathcal{B}\left(y \leq \xi_1 \,\middle|\, \frac{2N}{3}, h\right), \tag{3.9}$$

$$P_{\mathrm{s}}^{[1\,1]} = \mathcal{B}\left(y \leq \xi_1 \,\middle|\, \frac{4N}{3}, h\right) + 1 - \mathcal{B}\left(y \leq \xi_2 \,\middle|\, \frac{4N}{3}, h\right), \tag{3.10}$$

$$P_{\mathrm{s}}^{[1\,0]} = \mathcal{B}(y \leq \xi_2 \,|2N, h)\,, \tag{3.11}$$

$$P_{\mathrm{s}} = \frac{1}{4}\left(P_{\mathrm{s}}^{[0\,0]} + P_{\mathrm{s}}^{[0\,1]} + P_{\mathrm{s}}^{[1\,1]} + P_{\mathrm{s}}^{[1\,0]}\right). \tag{3.12}$$

To calculate the BER, the error probabilities must be calculated bit by bit for each symbol:

$$P_{\mathrm{b}}^{[0\,0]} = 0, \tag{3.13}$$

$$P_{\mathrm{b}}^{[0\,1]} = 1 - \mathcal{B}\left(y \leq \xi_1 \left| \frac{2N}{3}, h \right.\right) + \mathcal{B}\left(y \leq 0 \left| \frac{2N}{3}, h \right.\right) + 1 - \mathcal{B}\left(y \leq \xi_2 \left| \frac{2N}{3}, h \right.\right), \tag{3.14}$$

$$P_{\mathrm{b}}^{[1\,1]} = \mathcal{B}\left(y \leq \xi_1 \left| \frac{4N}{3}, h \right.\right) + \mathcal{B}\left(y \leq 0 \left| \frac{4N}{3}, h \right.\right) + 1 - \mathcal{B}\left(y \leq \xi_2 \left| \frac{4N}{3}, h \right.\right), \tag{3.15}$$

$$P_{\mathrm{b}}^{[1\,0]} = \mathcal{B}(y \leq \xi_1 \,|2N, h) + \mathcal{B}(y \leq \xi_2 \,|2N, h) - \mathcal{B}(y \leq 0 \,|2N, h), \tag{3.16}$$

$$P_{\mathrm{b}} = \frac{1}{2} \cdot \frac{1}{4} \left(P_{\mathrm{b}}^{[0\,0]} + P_{\mathrm{b}}^{[0\,1]} + P_{\mathrm{b}}^{[1\,1]} + P_{\mathrm{b}}^{[1\,0]} \right). \tag{3.17}$$

Fig. 3.5 shows numerically simulated and theoretically calculated BER and SER for 4-CSK

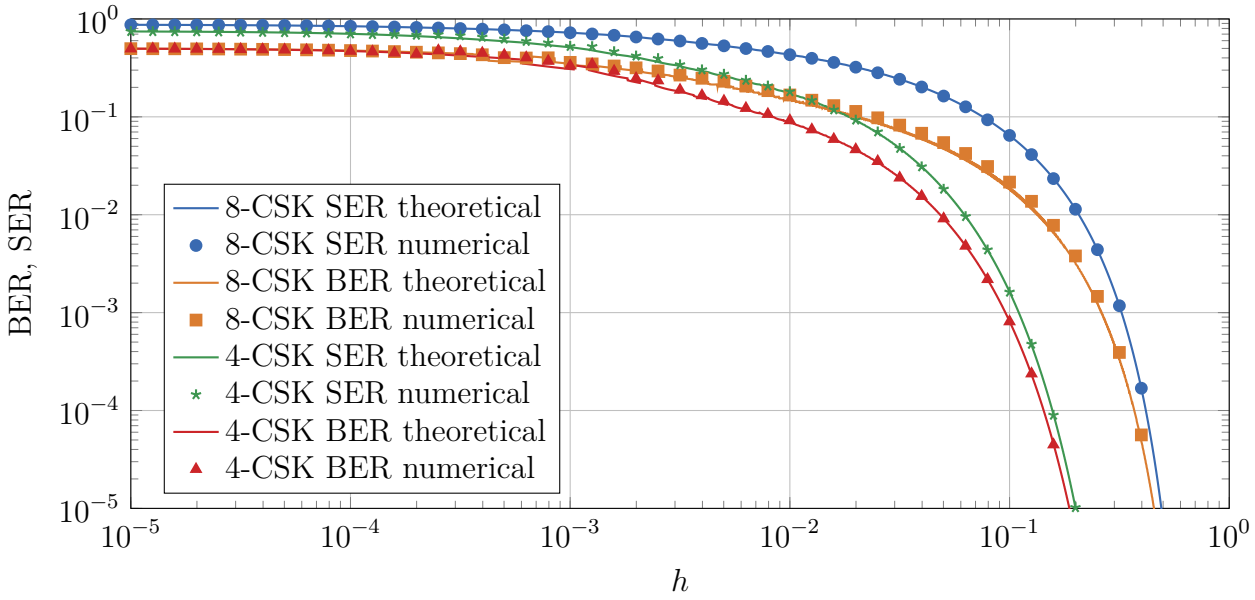


Figure 3.5: Bit error rate and symbol error rate of CSK. Parameters according to Appendix C.2.

and 8-CSK with respect to different h, where the numerical results match well to the theoretical results. The slight deviations from the smooth shape of the curves are due to the discrete nature of the binomial distribution. When determining the threshold values, a quantization error occurs, since the theoretical continuous intersection of two binomial distributions does not necessarily have to be a discrete value. This effect leads to slight jumps in the BER and SER performance. The performance loss at higher-order CSK is based on symbol energy normalization. This leads to a smaller distance between adjacent symbols when CSK is of higher order. As a result, symbol detection is more error-prone than for lower-order CSK.

3.1.3 Variable Concentration Shift Keying

Since the diffusion-based channel is affected by amplitude-dependent noise, equidistant spaces between adjacent symbols are no longer optimal. This limitation of CSK is eliminated by variable concentration shift keying (VCSK) through the introduction of variable distances [BD+19]. For this purpose, factors $0 \leq \alpha_i \leq 1$ with $1 \leq i \leq M-2$ are introduced, which define the distance of a symbol to its adjacent neighbor that has either the same or a higher concentration level. Thus the M symbols of M-VCSK result in $x^{(0)} = 0$ and $x^{(i)} = \alpha_i x^{(i+1)}$ which leads to the concentration-increasing property $0 \leq x^{(i)} \leq x^{(i+1)}$. In the sense of a fixed bitrate, the individual concentration levels can be determined considering the following condition:

$$E_\mathrm{b} = \frac{E_\mathrm{s}}{m} = \frac{1}{m} \sum_{x \in \mathcal{X}} p(x) x \tag{3.18}$$

$$= \frac{1}{mM} \left(\sum_{i=1}^{M-2} \alpha_i x^{(i+1)} + x^{(M-1)} \right) \stackrel{!}{=} \frac{N}{2}. \tag{3.19}$$

In the following, 4-VCSK is considered as an example. Consequently, (3.19) results in

$$E_\mathrm{b} = \frac{1}{2} \cdot \frac{1}{4} \left(\alpha_1 x^{(2)} + \alpha_2 x^{(3)} + x^{(3)} \right) = \frac{1}{2} \cdot \frac{1}{4} \left(\alpha_1 \alpha_2 x^{(3)} + \alpha_2 x^{(3)} + x^{(3)} \right) \stackrel{!}{=} \frac{N}{2}. \tag{3.20}$$

Thus, the concentration levels are obtained as [BD+19]

$$x^{(1)} = \frac{\alpha_1 \alpha_2 4N}{\alpha_1 \alpha_2 + \alpha_2 + 1} \tag{3.21}$$

$$x^{(2)} = \frac{\alpha_2 4N}{\alpha_1 \alpha_2 + \alpha_2 + 1} \tag{3.22}$$

$$x^{(3)} = \frac{4N}{\alpha_1 \alpha_2 + \alpha_2 + 1}. \tag{3.23}$$

As in CSK, $x^{(i)}$ does not necessarily have to result in an integer. In this case rounding must be applied. The deviation in E_b that can occur as a consequence of rounding is assumed to be negligible. The bit to symbol mapping in this dissertation is done following Gray labeling, which minimizes the BER.

Theoretical Symbol Error Rate and Bit Error Rate

In the following, the theoretical BER and SER for 4-VCSK is exemplarily derived. This derivation can be extended to higher-order VCSK modulation schemes. Fig. 3.6 sketches

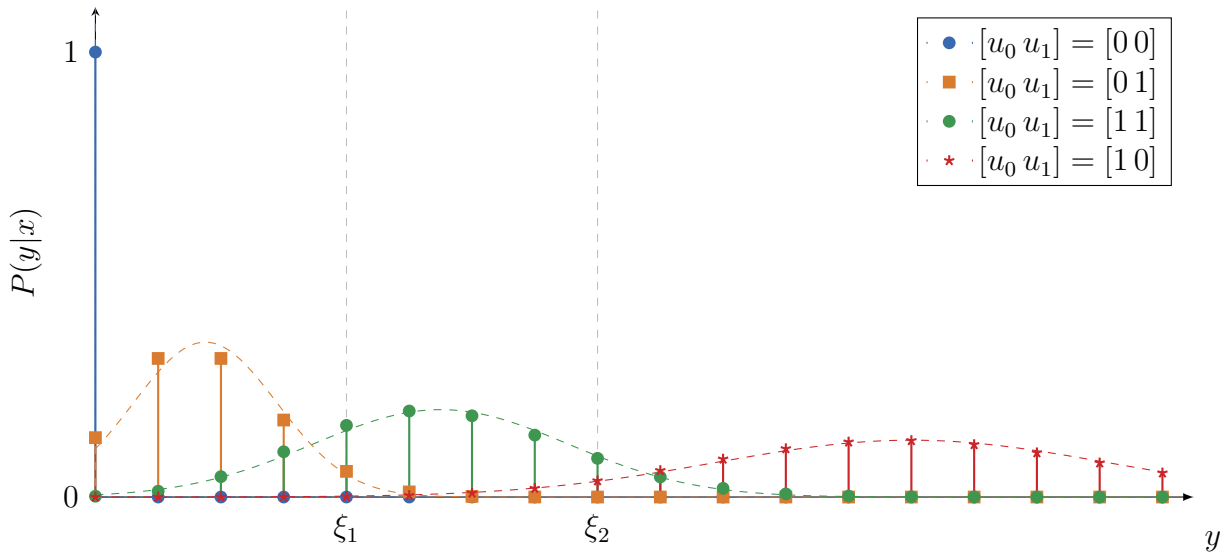


Figure 3.6: Symbol probability distribution for VCSK.

the symbol probability distribution at the receiver for a 4-VCSK transmission, which consists of a unit pulse $\mathcal{B}(0,1)$ at zero and three binomial distributions $\mathcal{B}\left(x^{(1)} = \frac{\alpha_1\alpha_2 4N}{\alpha_1\alpha_2+\alpha_2+1}, h\right)$, $\mathcal{B}\left(x^{(2)} = \frac{\alpha_2 4N}{\alpha_1\alpha_2+\alpha_2+1}, h\right)$, and $\mathcal{B}\left(x^{(3)} = \frac{4N}{\alpha_1\alpha_2+\alpha_2+1}, h\right)$. There are three different threshold values, namely 0, ξ_1, and ξ_2, which can be used to distinguish between the four possible transmitted symbols at the receiver side. The threshold values result from the intersections of the individual distributions. As discussed for CSK, they can be determined either numerically or analytically by Gaussian approximation. The Gaussian approximation of the distributions is represented in Fig. 3.6 by dashed lines. Taking the threshold values into account, the theoretical SER results in

$$P_{\mathrm{s}}^{[0\,0]} = 0, \tag{3.24}$$

$$P_{\mathrm{s}}^{[0\,1]} = \mathcal{B}\left(y \leq 0 \,\middle|\, x^{(1)}, h\right) + 1 - \mathcal{B}\left(y \leq \xi_1 \,\middle|\, x^{(1)}, h\right), \tag{3.25}$$

$$P_{\mathrm{s}}^{[1\,1]} = \mathcal{B}\left(y \leq \xi_1 \,\middle|\, x^{(2)}, h\right) + 1 - \mathcal{B}\left(y \leq \xi_2 \,\middle|\, x^{(2)}, h\right), \tag{3.26}$$

$$P_{\mathrm{s}}^{[1\,0]} = \mathcal{B}\left(y \leq \xi_2 \,\middle|\, x^{(3)}, h\right), \tag{3.27}$$

$$P_{\mathrm{s}} = \frac{1}{4}\left(P_{\mathrm{s}}^{[0\,0]} + P_{\mathrm{s}}^{[0\,1]} + P_{\mathrm{s}}^{[1\,1]} + P_{\mathrm{s}}^{[1\,0]}\right). \tag{3.28}$$

To calculate the BER, the error probabilities must be calculated bit by bit for each symbol:

$$P_{\mathrm{b}}^{[0\,0]} = 0, \tag{3.29}$$

$$P_{\mathrm{b}}^{[0\,1]} = 1 - \mathcal{B}\left(y \leq \xi_1 \left| x^{(1)}, h\right.\right) + \mathcal{B}\left(y \leq 0 \left| x^{(1)}, h\right.\right) + 1 - \mathcal{B}\left(y \leq \xi_2 \left| x^{(1)}, h\right.\right), \tag{3.30}$$

$$P_{\mathrm{b}}^{[1\,1]} = \mathcal{B}\left(y \leq \xi_1 \left| x^{(2)}, h\right.\right) + \mathcal{B}\left(y \leq 0 \left| x^{(2)}, h\right.\right) + 1 - \mathcal{B}\left(y \leq \xi_2 \left| x^{(2)}, h\right.\right), \tag{3.31}$$

$$P_{\mathrm{b}}^{[1\,0]} = \mathcal{B}\left(y \leq \xi_1 \left| x^{(3)}, h\right.\right) + \mathcal{B}\left(y \leq \xi_2 \left| x^{(3)}, h\right.\right) - \mathcal{B}(y \leq 0 \,|2N, h), \tag{3.32}$$

$$P_{\mathrm{b}} = \frac{1}{2} \cdot \frac{1}{4} \left(P_{\mathrm{b}}^{[0\,0]} + P_{\mathrm{b}}^{[0\,1]} + P_{\mathrm{b}}^{[1\,1]} + P_{\mathrm{b}}^{[1\,0]}\right). \tag{3.33}$$

If the distances between adjacent symbols are appropriately selected through α_i, the BER and SER can be minimized. The gain through variable concentration levels compared to CSK can be observed in different visualizations. If the distributions in Fig. 3.4 are compared with those in Fig. 3.6, it becomes clear that the error areas of 4-VCSK are smaller compared to 4-CSK. Fig. 3.7 shows the theoretical SER over α_1 and α_2 for $h = 0.01$. A 4-CSK is

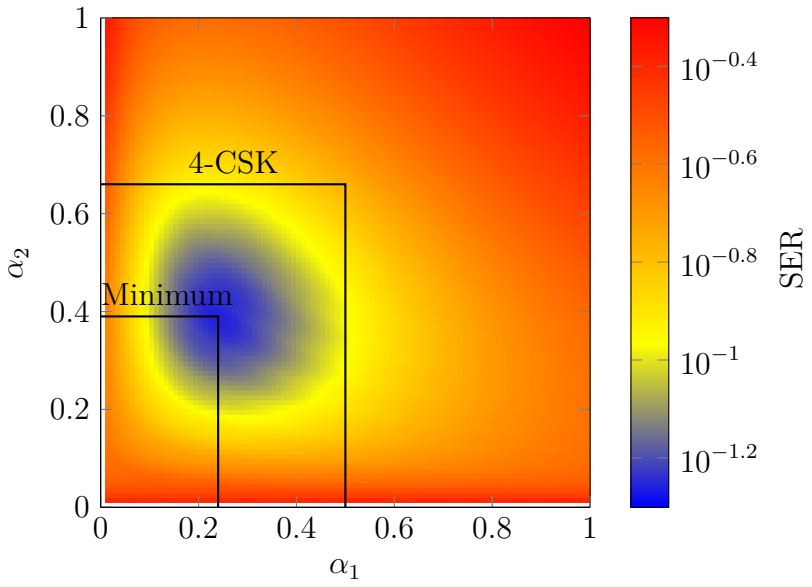


Figure 3.7: Symbol error rate of 4-VCSK with respect to α_1 and α_2 ($h = 0.01$). Parameters according to Appendix C.3.

equivalent to a 4-VCSK with $\alpha_1 = 1/2$ and $\alpha_2 = 2/3$. However, a minimum SER is achieved with $\alpha_1 \approx 0.24$ and $\alpha_2 \approx 0.39$. Fig. 3.8 compares numerically and theoretically determined BER and SER for 4-VCSK under the influence of different h. The optimal α_1 and α_2 values for 4-VCSK are selected for each h, which lead to a minimum BER and SER. The example of 4-VCSK demonstrates that the variable concentration levels lead to a significant

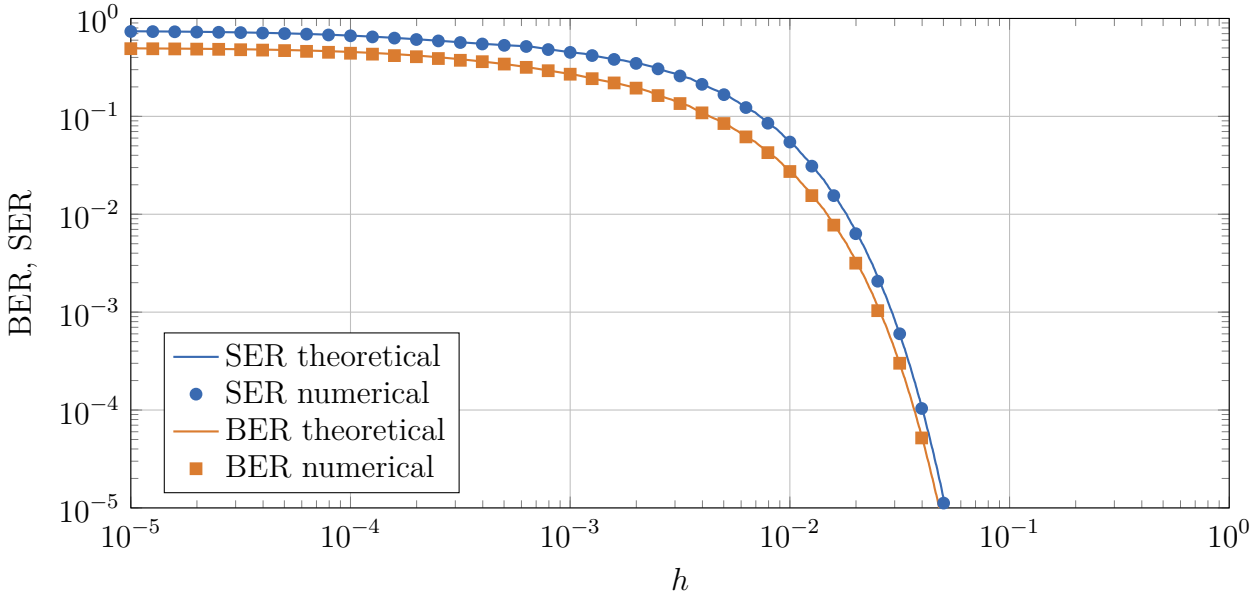


Figure 3.8: Bit error rate and symbol error rate of 4-VCSK. Parameters according to Appendix C.2.

increase in performance compared to 4-CSK in Fig. 3.5. Furthermore, it is confirmed that the theoretical results agree well with the numerical results.

3.1.4 Pulse-Position Modulation

As an alternative to modulating the released molecule concentration, the time of molecule release can also be modulated. A popular example is pulse-position modulation (PPM) [GL+11] which is named in analogy to the modulation frequently used in optical communication and remote controls. In M-PPM, $m = \log_2(M)$ bits are combined into a PPM symbol. The transmission interval with symbol duration $T = mT_\mathrm{b}$ is divided into M equally sized time slots. Depending on the m bits, $(mN)/2$ molecules are released at the beginning of the uniquely assigned time slot. For the transmission of bit sequence $\boldsymbol{u} = [1\,0\;\;1\,0\,0\;\;1\,1]$, the corresponding 2-PPM and 4-PPM-modulated transmit signals are depicted in Fig. 3.1. As for CSK, Gray labeling is assumed to minimize the BER.

If the info bits are mapped as shown exemplarily for 4-PPM in Tab. 3.1, PPM transmission can be modeled as OOK transmission. The corresponding new symbol duration

Table 3.1: Mapping from PPM modulated info bits to OOK modulated info bits.

$\boldsymbol{u}$	$\boldsymbol{u}'$
[0 0]	[1 0 0 0]
[0 1]	[0 1 0 0]
[1 1]	[0 0 1 0]
[1 0]	[0 0 0 1]

results in $T' = T'_{\mathrm{b}} = T/M$ and the OOK symbols in

$$x'[k] = \begin{cases} 0 & \text{if } u'[k] = 0, \\ \frac{m}{2}N & \text{if } u'[k] = 1. \end{cases} \tag{3.34}$$

Theoretical Symbol Error Rate and Bit Error Rate

For the theoretical derivation, PPM is interpreted as OOK with symbol duration T'. Since no ISI is considered, the symbols at the receiver have the probability distribution shown in Fig. 3.2 in each transmission interval. In the sense of PPM, exactly one bit equal to one is transmitted each M transmission intervals. Consequently, an error occurs if no molecules are absorbed at the receiver during those intervals. Thus, the theoretical BER is equal to

$$P_{\mathrm{b}} = \frac{1}{2}\left[\mathcal{B}\left(y \leq 0 \middle| \frac{m}{2}N, h\right)\right] = \frac{1}{2}(1-h)^{\frac{m}{2}N}. \tag{3.35}$$

The factor 1/2 arises from the fact that the transmitted info bits have to be selected randomly with a probability of 1/2 if no molecule reaches the transmitter. A symbol error also occurs when no molecules reach the receiver during the transmission interval. Since in this case the symbol has to be decided randomly, the error probability of the decision process is $(M-1)/M$. Thus, the theoretical SER results in

$$P_{\mathrm{s}} = \frac{M-1}{M}\left[\mathcal{B}\left(y \leq 0 \middle| \frac{m}{2}N, h\right)\right] = \frac{M-1}{M}(1-h)^{\frac{m}{2}N}. \tag{3.36}$$

Fig. 3.9 shows numerically simulated and theoretically calculated BER and SER for 2-PPM, 4-PPM, and 8-PPM with respect to different h, where the numerical results match well to the theoretical results. It can be observed that a higher-order PPM leads to a better performance with constant h. The gain is due to the energy normalization. With higher-order PPM, several bits are combined to form a symbol, so that more molecules can

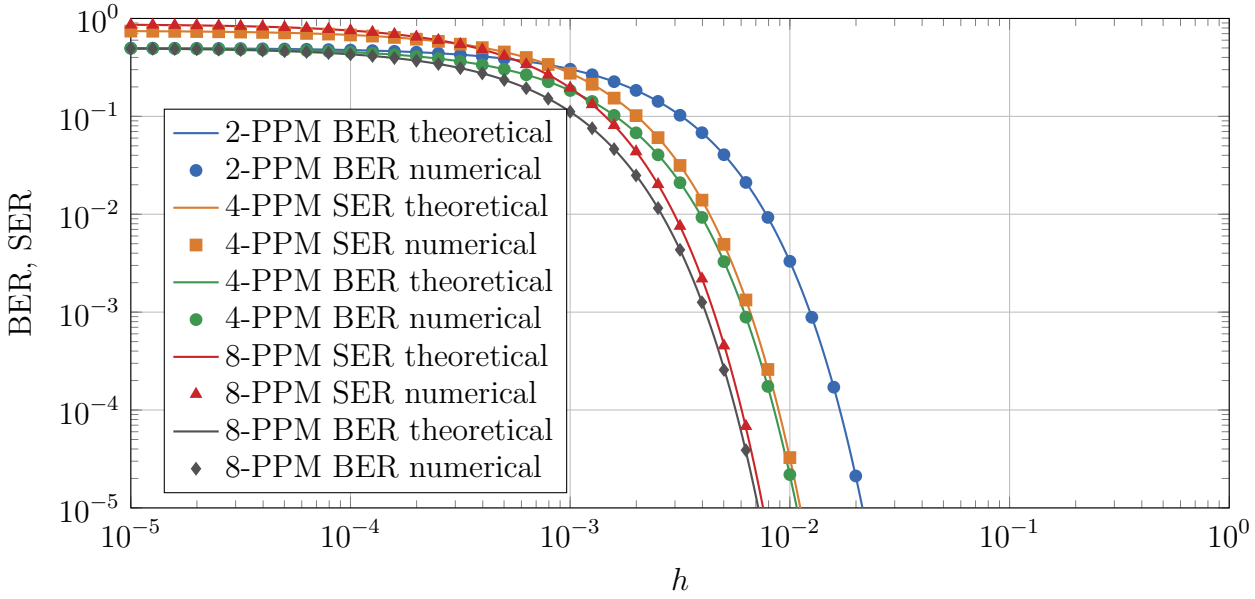


Figure 3.9: Bit error rate and symbol error rate of PPM. Parameters according to Appendix C.2.

be released for the single impulse. It should be noted that an equal h in a diffusion-based channel is typically achieved by equal transmission intervals with duration T'. However, the total PPM symbol duration results in $T = MT'$, so the information rate decreases with higher-order PPM. In addition, ISI can significantly influence the performance of PPM. A comparison with normalized information rate and ISI is given in Section 3.1.7.

3.1.5 Molecule Shift Keying

In contrast to radio-based wireless communication, molecular communication offers a further degree of freedom in modulation. If the receptors of the receiver are sensitive to different types of molecules, then information can also be represented by the type of molecule. The resulting modulation format is called molecule shift keying (MoSK) [KY+11]. It is assumed that a type of molecule is emitted at the transmitter in each symbol interval. The modulation order M is therefore equal to the number of different molecule types. Consequently, $m = \log_2(M)$ bits can be represented by an MoSK symbol. The energy normalization leads to

$$E_\mathrm{b} = \frac{E_\mathrm{s}}{m} \stackrel{!}{=} \frac{N}{2}. \tag{3.37}$$

Thus, for example, the mapping for 2-MoSK is defined as

$$x[k] = \begin{cases} \frac{N_\mathrm{A}}{2} & \text{if } u[k] = 0, \\ \frac{N_\mathrm{B}}{2} & \text{if } u[k] = 1, \end{cases} \tag{3.38}$$

where $N_\mathrm{A}/2$ and $N_\mathrm{B}/2$ are the number of released molecules of type A and type B, respectively. Fig. 3.1 shows the resulting transmit signal for 2-MoSK of an info sequence. The M-MoSK transmission can be described by M equivalent discrete-time channel models (EDTCMs), whereby the transmission of each molecule type can be represented by its own OOK transmission.

Theoretical Symbol Error Rate and Bit Error Rate

In each transmission interval, a pulse of molecules of exactly one type is released. Therefore, the probability distribution of the symbols in Fig. 3.2 results at the receiver. Since no ISI is assumed, an error can only occur if no molecules are absorbed at the receiver during a transmission interval. The theoretical BER is therefore

$$P_\mathrm{b} = \frac{1}{2}\left[\mathcal{B}\left(y \leq 0 \middle| \frac{m}{2}N, h\right)\right] = \frac{1}{2}(1-h)^{\frac{m}{2}N}. \tag{3.39}$$

The factor 1/2 describes the random selection of the transmitted bits in the event of an error. If symbols are considered, the case that no molecules reach the receiver also leads to errors. The random selection of the transmitted symbol in case of error has an error probability of $(M-1)/M$. Therefore, the theoretical SER results in

$$P_\mathrm{s} = \frac{M-1}{M}\left[\mathcal{B}\left(y \leq 0 \middle| \frac{m}{2}N, h\right)\right] = \frac{M-1}{M}(1-h)^{\frac{m}{2}N}. \tag{3.40}$$

Fig. 3.10 shows the numerically simulated and theoretically calculated BER and SER for 2-MoSK, 4-MoSK, and 8-MoSK with respect to different h. It can be seen that the theoretical and numerical results agree well. In higher-order MoSK, more bits per symbol can be transmitted, thus more molecules are available for a symbol. This improves BER and SER performance for higher-order MoSK. The performance without ISI of M-MoSK is the same as M-PPM. However, it should be noted that with PPM, the transmission intervals for higher-order PPM are shorter and h is typically smaller. The opposite is true for MoSK. Due to the gained degree of freedom of the additional molecule types, higher-order MoSK leads to longer transmission intervals and thus typically to higher h. Thus, MoSK is also more robust against ISI than PPM.

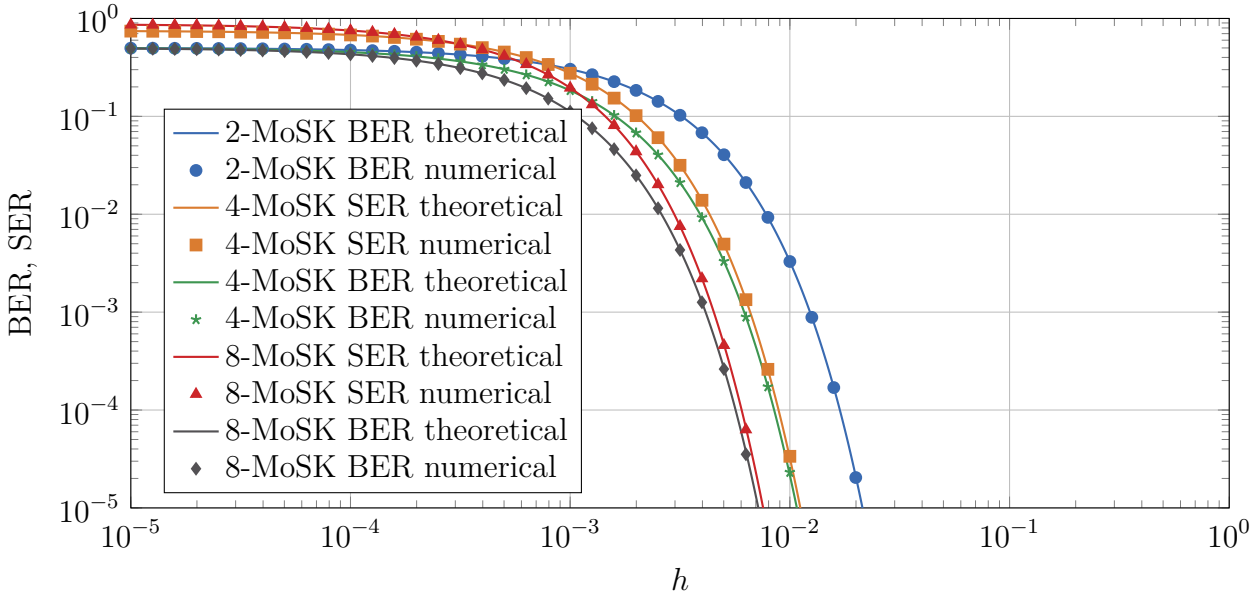


Figure 3.10: Bit error rate and symbol error rate of MoSK. Parameters according to Appendix C.2.

3.1.6 Molecule Concentration Shift Keying

In U^Z-molecule concentration shift keying (MCSK) [KC13; BD+19], CSK from Section 3.1.2 is combined with the degrees of freedom obtained by MoSK using several molecule types from Section 3.1.5. The info bits are represented by U possible concentration levels of each Z molecule type. U^2-MCSK can be regarded as an equivalent to U^2ary quadrature amplitude modulation (QAM) in classical wireless communication, which is constructed of two independent U-ASK schemes, one in the real domain and one in the imaginary domain. While QAM is limited to the complex plane, MCSK allows to exploit further degrees of freedom through $Z > 2$. The $M = U^Z$ symbols can be represented by $m = \log_2\left(U^Z\right)$ bits, where m/z bits are represented by a U-CSK, respectively. Thus, a U^Z-MCSK can also be interpreted as Z parallel U-CSK schemes. Consequently, the distance between adjacent symbols of each U-CSK is analogous to (3.4)-(3.7)

$$\alpha = \frac{\frac{m}{Z}N}{(U-1)}. \tag{3.41}$$

Similar to CSK, α does not necessarily have to be an integer and rounding might be necessary. The resulting deviation in E_b is assumed to be negligibly small. In addition, Gray labeling is applied for each U-CSK scheme. An example of the resulting transmit signal for 2^2-MCSK is shown in Fig. 3.1.

Theoretical Symbol Error Rate and Bit Error Rate

Since U^Z-MCSK can be interpreted as Z parallel U-CSK schemes, the result of the theoretical SER and BER of a U^Z-MCSK corresponds for the ISI-free case exactly to that of U-CSK in Section 3.1.2. The results in Fig 3.5 are also equivalent. However, this behaviour changes in a system affected by ISI. Analogous to MoSK, the additional degree of freedom of Z different molecule types in U^Z-MCSK leads to an improvement in performance compared to U-CSK. At a fixed information rate, the use of different molecule types typically results in longer symbol durations and thus larger h. An additional increase in performance can be achieved by using variable concentration levels according to VCSK in Section 3.1.3.

3.1.7 Bit Error Rate Performance

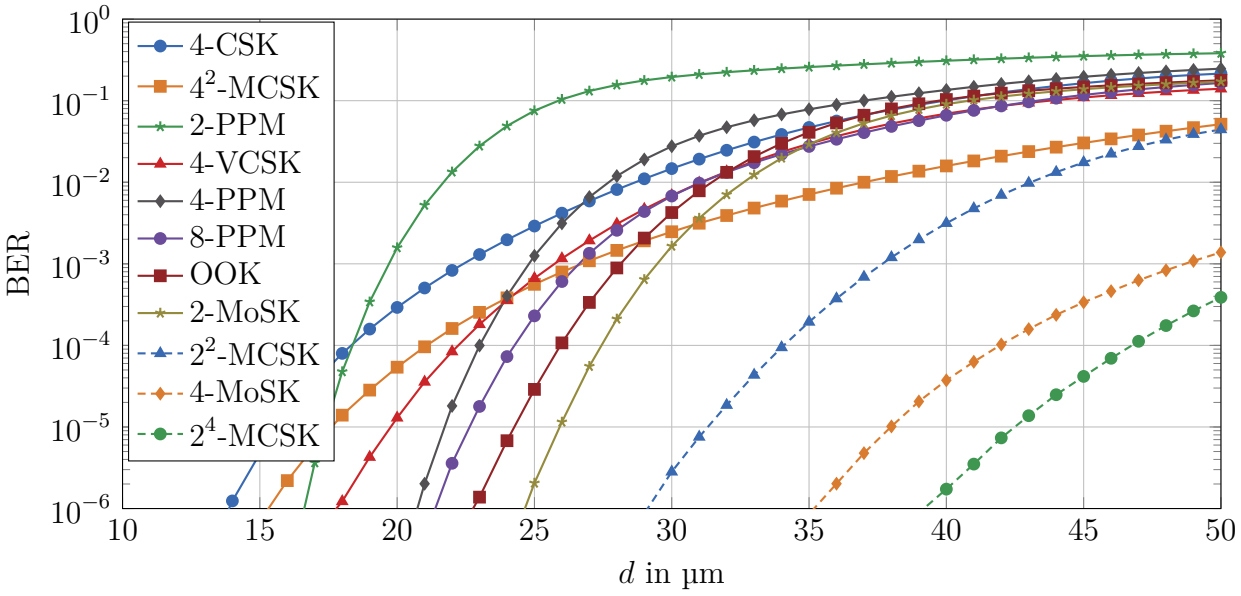


Figure 3.11: Bit error rate comparison between different single-carrier modulation schemes for varying transmission distances. Parameters according to Appendix C.4.

In contrast to the previous derived theoretical symbol error rates and bit error rates, Fig. 3.11 compares all introduced modulation schemes in an ISI-affected scenario. For

this purpose the bit error rate is determined with the help of Monte Carlo simulations. Detection is performed applying the *a posteriori* probability (APP) detector introduced in Section 3.2.8. In order to allow a fair comparison, the following three parameters are chosen to be the same for all modulation schemes:

1. The average number of released molecules per info bit E_b,
2. the bit duration T_b,
3. the absolute channel memory duration $(L+1)T$.

Fig. 3.11 demonstrates the strength of OOK compared to higher-order modulation schemes which only consider a single type of molecule. While the performance of CSK with higher level suffers from the smaller distances between adjacent symbols, PPM with higher order benefits from the increased number of molecules per pulse. If 4-VCSK is compared with 4-CSK, the variable concentration levels bring a significant improvement. Nevertheless, the BER performance is worse than that of OOK. If modulation formats with 2-ary concentration levels and more than a single molecule type are considered, they bring a significant performance gain over OOK. The example of U^Z-MCSK demonstrates this observation. While an increase of the concentration levels U has a negative influence on the BER performance, an increase of the molecule types Z leads to an improvement. At a target BER of 10^{-3}, expanding OOK to two molecule types results in an increase of the maximum transmission distance from 28 µm to approximately 38 µm. If it is further extended to four molecule types, the maximum transmission distance can be increased to 53 µm.

3.2 Channel Equalization and Detection

3.2.1 Fixed Threshold Detection

The fixed threshold detector (FTD) is a very simple and low-complexity detector, which often occurs in nature. One example is the quorum sensing of bacteria. Bacteria release autoinducers to determine their population size. If a certain concentration of autoinducers is reached, the bacteria change their behaviour. Another example are neurons that become active when they are excited beyond a certain threshold potential. In molecular communication, the FTD is used for symbol detection. If the number of absorbed molecules $y[k]$ in a transmission interval exceeds or falls below a certain threshold value ξ, the decision $\hat{x}[k]$ is

made for the corresponding transmitted symbol [MM+11]. Therefore, it is only suitable for concentration-based modulations. These include OOK, CSK, VCSK, and partial detection of MCSK. For binary OOK modulation, the FTD rule is

$$\hat{x}[k] = \begin{cases} 0 & \text{if } y[k] \leq \xi, \\ N & \text{if } y[k] > \xi. \end{cases} \tag{3.42}$$

If M concentration levels have to be distinguished, $M-1$ threshold values result for the FTD. Despite the favorable low-complexity realization, there are also a few disadvantages of FTD. On the one hand, it cannot inherently counteract ISI. On the other hand, the algorithm is not able to adapt to varying channel conditions due to the fixed threshold. As a result, the threshold might only be well-chosen for a specific scenario.

Theoretical Symbol Error Rate and Bit Error Rate

The theoretical BER and SER for OOK and a channel with memory length $L = 1$ is determined. The equations here are used as an example, but can also be easily extended to higher-order modulation schemes such as CSK and longer channel memory lengths. Given

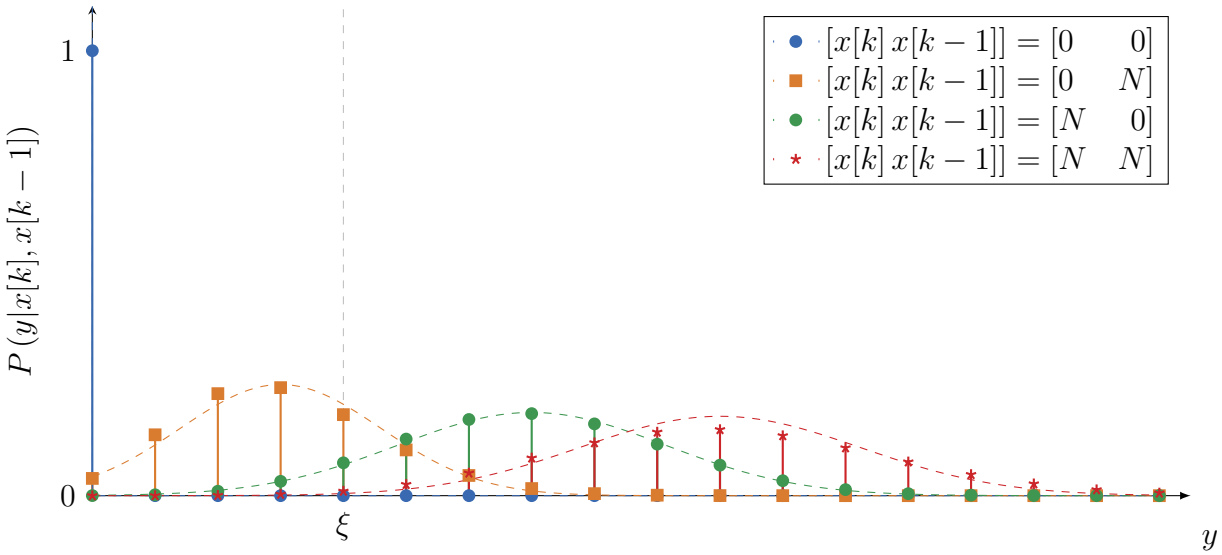


Figure 3.12: Symbol probability distribution for OOK with channel memory length $L = 1$.

$L = 1$, four different probability distributions can occur at the receiver, depending on the transmitted sequence, as shown in Fig. 3.12. Since the distribution for $[x[k]\, x[k-1]] = [N\, N]$

has a Poisson binomial distribution which is not easy to handle in mathematical derivations, the distributions are approximated by Gaussian distributions as described in Section 2.3.3. The approximating Gaussian distributions are represented by dashed lines in Fig 3.12. The bit error probabilities $P_{\mathrm{b}}^{[u[k]\,u[k-1]]}$ of the four different possibilities can be determined as a function of the fixed threshold:

$$P_{\mathrm{b}}^{[0\,0]} = 1 - \mathcal{N}\Big(y \leq \frac{1}{2}\,|0,0\Big) = 0, \tag{3.43}$$

$$P_{\mathrm{b}}^{[0\,1]} \approx 1 - \mathcal{N}(y \leq \xi\,|Nh_1, Nh_1(1-h_1))\,, \tag{3.44}$$

$$P_{\mathrm{b}}^{[1\,0]} \approx \mathcal{N}(y \leq \xi\,|Nh_0, Nh_0(1-h_0))\,, \tag{3.45}$$

$$P_{\mathrm{b}}^{[1\,1]} \approx \mathcal{N}(y \leq \xi\,|N(h_0+h_1), N(h_0+h_1)(1-(h_0+h_1)))\,, \tag{3.46}$$

where $\mathcal{N}\Big(y \leq \frac{1}{2}\,|0,0\Big)$ is the cumulative distribution function of a Dirac delta function. The integration threshold $y \leq 1/2$ is used to reflect the rounding of the continuous Gaussian distribution values. The total BER is the average of the four individual bit error probabilities:

$$P_{\mathrm{b}} = \frac{1}{4}\left(P_{\mathrm{b}}^{[0\,0]} + P_{\mathrm{b}}^{[0\,1]} + P_{\mathrm{b}}^{[1\,1]} + P_{\mathrm{b}}^{[1\,0]}\right) = P_{\mathrm{s}}. \tag{3.47}$$

Since OOK is considered here, the BER is equal to the SER. From the theoretical BER it can be seen that certain sequences are more prone to errors than others. Fig 3.13 shows a comparison of the theoretical BER with the numerical BER. The result shows that the theoretically derived BER agrees well with the simulatively determined BER.

3.2.2 Adaptive Threshold Detection

The adaptive threshold detector (ATD) is a biologically inspired low-complexity detector for OOK which, in contrast to the FTD, changes its threshold value depending on the received number of molecules [DH16]. The idea of an adaptive threshold is motivated by the effect of short-term synaptic plasticity [ZR02] in human neural networks. Depending on previous stimuli, the synapses increase (enhancement) or decrease (depression) their activity for a stimulus. On the postsynaptic side, the activity can be adjusted, for example, by modifying receptors or by releasing enzymes into the synaptic cleft. Therefore, a detector with a threshold that adapts to environmental stimuli is a very realistic approach especially for biological microscopic molecular communication. Since it already occurs in nature, it can be assumed that such a detector can also be implemented in small energy-limited receivers. The detection rule, which is illustrated in Fig. 3.14 by a block diagram, is as

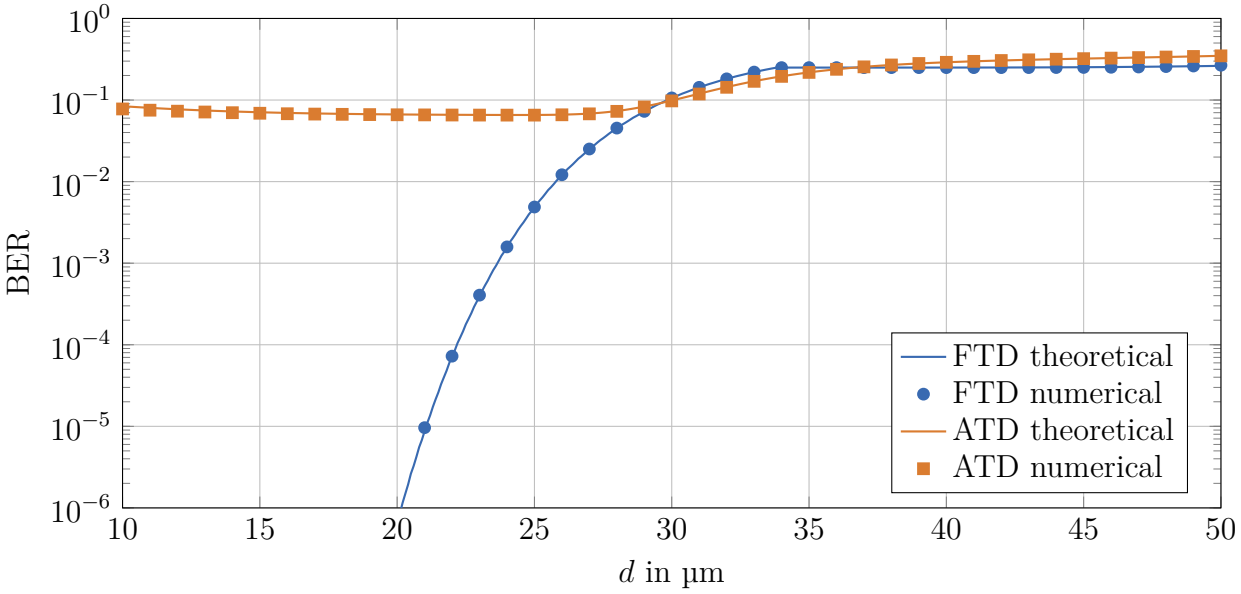


Figure 3.13: Bit error rate of FTD and ATD. Parameters according to Appendix C.5.

follows [DH16]:

$$\hat{x}[k] = \begin{cases} 0 & \text{if } y[k] \leq \xi[k] = y[k-1], \\ N & \text{if } y[k] > \xi[k] = y[k-1]. \end{cases} \tag{3.48}$$

Only if the number of received molecules in the current transmission interval exceeds the

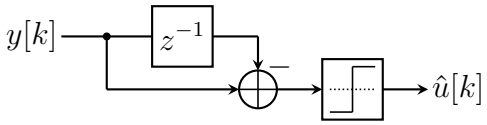


Figure 3.14: Block diagram of the adaptive threshold detector.

number of received molecules in the previous transmission interval, the detector decides for $\hat{u}[k] = 1$. Unlike the FTD, the ATD inherently needs ISI in the channel. Without ISI, especially the detection of a sequence of consecutive ones leads to an increased probability of error (due to the need to exceed the number of molecules received from the previous transmission interval). A major advantage is that the ATD is defined independently of the channel coefficients h_ℓ. Thus it works without the knowledge of h_ℓ for varying channels and the computational complexity and overhead of channel estimation can be saved.

Theoretical Symbol Error Rate and Bit Error Rate

For the derivation of the theoretical BER and SER a channel memory length of $L = 1$ is assumed. In contrast to the derivation of FTD in Section 3.2.1, here the binomial distributions are approximated by Poisson distributions as discussed in Section 2.3.4 to avoid complicated expressions with Poisson binomial distributions. The four possible probability distributions at the receiver are shown in Fig 3.12. Since the adaptive threshold $\xi[k]$ is determined by the previous number of received molecules $y[k]$, the BER depends on the current transmitted bit $u[k]$ and the two previous ones $u[k-1]$ and $u[k-2]$. This results in eight different sequences, each with their own bit error probabilities $P_\mathrm{b}^{[u[k]\,u[k-1]\,u[k-2]]}$. For a threshold value $\xi[k]$, each of these bit error probabilities results from the occurrence probability of this threshold value multiplied by the probability that an error occurs in the current transmission interval. In order to consider all threshold values that may occur, the sum over all possible threshold values must be calculated. This results in

$$P_\mathrm{b}^{[0\,0\,0]} = \sum_{\xi=0}^{\infty} \mathcal{P}(y = \xi\,|0)\,(1 - \mathcal{P}(y \leq \xi\,|0)) = 0, \tag{3.49}$$

$$P_\mathrm{b}^{[1\,0\,0]} \approx \sum_{\xi=0}^{\infty} \mathcal{P}(y = \xi\,|0)\,\mathcal{P}(y \leq \xi\,|Nh_0) = \mathcal{P}(y \leq 0\,|Nh_0)\,, \tag{3.50}$$

$$P_\mathrm{b}^{[0\,1\,0]} \approx \sum_{\xi=0}^{\infty} \mathcal{P}(y = \xi\,|Nh_0)\,(1 - \mathcal{P}(y \leq \xi\,|Nh_1))\,, \tag{3.51}$$

$$P_\mathrm{b}^{[1\,1\,0]} \approx \sum_{\xi=0}^{\infty} \mathcal{P}(y = \xi\,|Nh_0)\,\mathcal{P}(y \leq \xi\,|N(h_0 + h_1))\,, \tag{3.52}$$

$$P_\mathrm{b}^{[0\,0\,1]} = \sum_{\xi=0}^{\infty} \mathcal{P}(y = \xi\,|Nh_1)\,(1 - \mathcal{P}(y \leq \xi\,|0)) = 0, \tag{3.53}$$

$$P_\mathrm{b}^{[1\,0\,1]} \approx \sum_{\xi=0}^{\infty} \mathcal{P}(y = \xi\,|Nh_1)\,\mathcal{P}(y \leq \xi\,|Nh_0)\,, \tag{3.54}$$

$$P_\mathrm{b}^{[0\,1\,1]} \approx \sum_{\xi=0}^{\infty} \mathcal{P}(y = \xi\,|N(h_0 + h_1))\,(1 - \mathcal{P}(y \leq \xi\,|Nh_1))\,, \tag{3.55}$$

$$P_\mathrm{b}^{[1\,1\,1]} \approx \sum_{\xi=0}^{\infty} \mathcal{P}(y = \xi\,|N(h_0 + h_1))\,\mathcal{P}(y \leq \xi\,|N(h_0 + h_1))\,, \tag{3.56}$$

where $\mathcal{P}(0)$ is a unit impulse. The total BER is the average of the individual bit error probabilities:

$$P_\mathrm{b} = \frac{1}{8}\left(P_\mathrm{b}^{[0\,0\,0]} + P_\mathrm{b}^{[1\,0\,0]} + P_\mathrm{b}^{[0\,1\,0]} + P_\mathrm{b}^{[1\,1\,0]} + P_\mathrm{b}^{[0\,0\,1]} + P_\mathrm{b}^{[1\,0\,1]} + P_\mathrm{b}^{[0\,1\,1]} + P_\mathrm{b}^{[1\,1\,1]}\right). \tag{3.57}$$

Since OOK is considered here, the BER is equal to the SER, i. e. $P_\mathrm{b} = P_\mathrm{s}$. From the theoretical BER it can be seen that certain sequences are more prone to errors than others. Fig 3.13 shows a comparison of the theoretical BER with the numerical BER. The result indicates that the theoretically derived BER agrees well with the BER determined by simulations.

3.2.3 Piecewise Linear Detection

In [DH+18; DH+17], the piecewise linear detector (PLD) with emphasis on superposition modulation was introduced. It is based on the observation that the extrinsic log-likelihood ratios (LLRs) depicted over the received samples can be approximated by piecewise straight lines. This idea can also be applied to diffusion-based molecular communication as shown in Fig. 3.15. As in [DH+18; DH+17], symbol-wise detection rather than sequence detection is considered. Consequently, only the current received sample $y[k]$ instead of the whole sequence $\boldsymbol{y}$ of received samples is taken into account. In the following, binary OOK transmission is assumed, so the terms symbol and bit can be used interchangeably. The symbol-wise maximum APP detection is defined by

$$\hat{u}[k] = \arg\max_{\tilde{u}[k]} p(\tilde{u}[k]|y[k]). \tag{3.58}$$

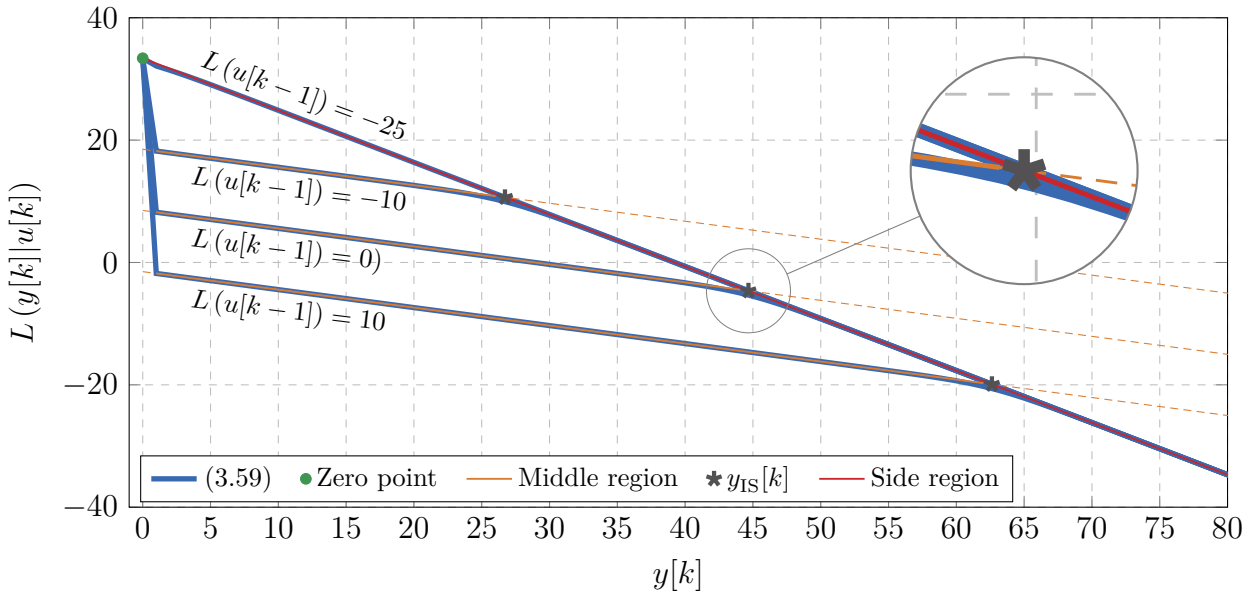
In the sense of a symbol-wise APP detection, LLRs are determined as

$$L\left(u[k]|y[k]\right) \doteq \log \frac{P\left(u[k]=0|y[k]\right)}{P\left(u[k]=1|y[k]\right)} = \underbrace{\log \frac{P\left(y[k]|u[k]=0\right)}{P\left(y[k]|u[k]=1\right)}}_{L(y[k]|u[k])} + \underbrace{\log \frac{P\left(u[k]=0\right)}{P\left(u[k]=1\right)}}_{L(u[k])}, \tag{3.59}$$

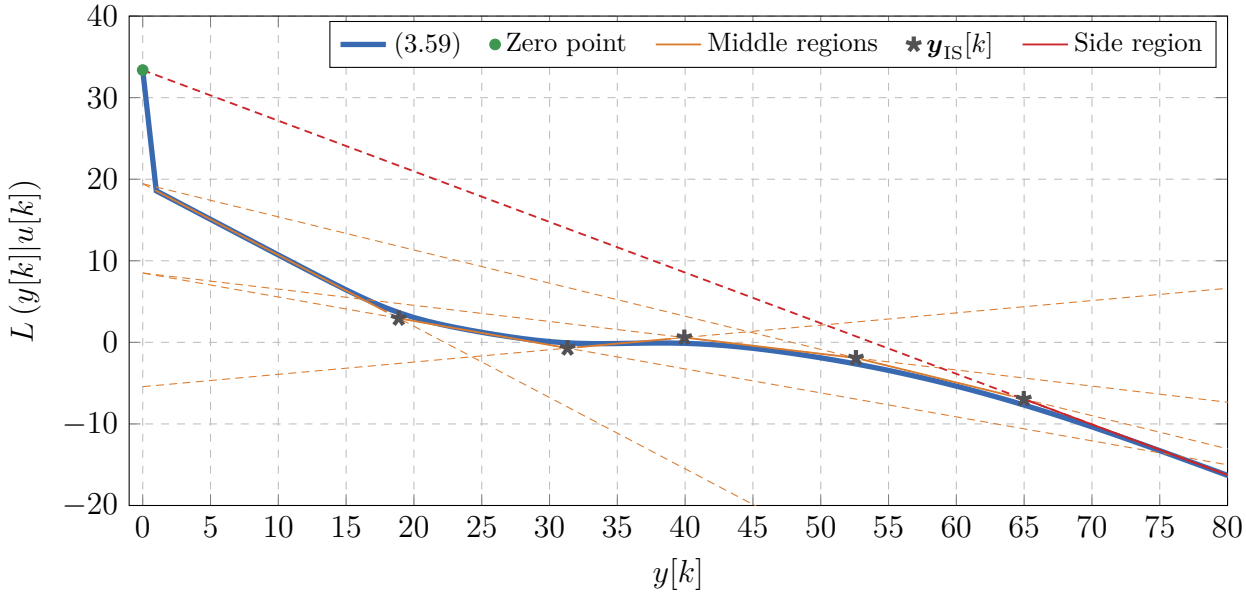
where $L\left(u[k]|y[k]\right)$, $L\left(y[k]|u[k]\right)$, $L\left(u[k]\right)$ are the *a posteriori* , extrinsic, and *a priori* LLRs, respectively. The LLR represents soft information about the info bit u. The sign of an LLR can be used for a hard decision, while the magnitude represents information about the reliability of the decision.

To derive the PLD for diffusion-based molecular communication, a channel with memory length $L = 1$ is assumed for simplicity. Nevertheless, the PLD can be extended to any $L > 1$ by following the same procedure (see Fig. 3.15). Using the Poisson approximation from Section 2.3.4, the extrinsic LLR is approximated as

$$L\left(y[k]|u[k]\right) \approx \log \frac{P\left(u[k-1]=0\right)\mathcal{P}(y[k]\,|0\,) + P\left(u[k-1]=1\right)\mathcal{P}(y[k]\,|Nh_1\,)}{P\left(u[k-1]=0\right)\mathcal{P}(y[k]\,|Nh_0\,) + P\left(u[k-1]=1\right)\mathcal{P}(y[k]\,|N(h_0+h_1)\,)}, \tag{3.60}$$



(a) $L = 1$.



(b) $L = 2$, no *a priori* information.

Figure 3.15: Extrinsic LLR of symbol-wise APP detector and PLD depicted over the received samples for different channel memory lengths. The discrete functions are represented as continuous in favor of visibility. Parameters according to Appendix C.6.

where $\mathcal{P}(y\,|\lambda) = \lambda^y/y! \cdot \exp(-\lambda)$ for $\lambda > 0$, $\mathcal{P}(y = 0\,|\lambda = 0) = 1$, and $\mathcal{P}(y > 0\,|\lambda = 0) = 0$, respectively.

As in [DH+18; DH+17], not all terms in (3.60) are taken into account for the PLD. Instead, only the respective dominant terms in the numerator and denominator are considered. For the diffusion-based channel with $L = 1$ the following three regions exist:

Zero point $y[k] = 0$**:** At this point the impulse-type distribution $\mathcal{P}(\lambda = 0)$ dominates in the numerator and the distribution $\mathcal{P}(Nh_0)$ in the denominator, since $h_0 \leq (h_0 + h_1)$ applies. The extrinsic LLR can thus be approximated at the point $y[k] = 0$ by

$$L\left(y[k]|u[k]\right) \approx \log \frac{P\left(u[k-1]=0\right)\mathcal{P}(y[k]=0\,|0\,)}{P\left(u[k-1]=0\right)\mathcal{P}(y[k]=0\,|Nh_0\,)} = \log \frac{1}{\exp(-Nh_0)} = Nh_0. \tag{3.61}$$

Interestingly, this point can be calculated very easily and is independent of the *a priori* information of the previous bit (see Fig. 3.15(a)).

Middle region $0 < y[k] < y_{\mathrm{IS}}[k]$**:** In this region, the dominant term in the numerator changes compared to the zero point through $y[k] > 0$, since $\mathcal{P}(y > 0\,|\lambda = 0) = 0$ applies. The dominant term in the denominator does not change until $y_{\mathrm{IS}}[k]$, which describes the point of intersection where the two terms in the denominator are equal. The approximated extrinsic LLR is simplified in this region as

$$\begin{aligned} L\left(y[k]|u[k]\right) &\approx \log \frac{P\left(u[k-1]=1\right)\mathcal{P}(y[k]\,|Nh_1\,)}{P\left(u[k-1]=0\right)\mathcal{P}(y[k]\,|Nh_0\,)} \\ &= \log \frac{P\left(u[k-1]=1\right)(Nh_1)^{y[k]}\exp(-Nh_1)}{P\left(u[k-1]=0\right)(Nh_0)^{y[k]}\exp(-Nh_0)} \\ &= \left(\log(h_1) - \log(h_0)\right) y[k] + N(h_0 - h_1) - L\left(u[k-1]\right). \end{aligned} \tag{3.62}$$

From (3.62) it can be observed that the approximation results in a linear function, which depends only on the channel coefficients, the number of released molecules, and the *a priori* LLR of the previous bit. If *a priori* LLR is present, the linear function of the middle region is shifted by this amount as shown in Fig. 3.15(a).

Side region $y[k] \geq y_{\mathrm{IS}}[k]$**:** When the intersection point $y_{\mathrm{IS}}[k]$ is reached, the dominant term in the denominator changes in favor of $\mathcal{P}(\lambda = N(h_0 + h_1))$ in contrast to the middle region. The dominant term in the numerator, on the other hand, does not change.

The approximated extrinsic LLR results in

$$\begin{aligned} L\left(y[k]|u[k]\right) &\approx \log \frac{P\left(u[k-1]=1\right)\mathcal{P}(y[k]\,|Nh_1\,)}{P\left(u[k-1]=1\right)\mathcal{P}(y[k]\,|N(h_0+h_1)\,)} \\ &= \log \frac{(Nh_1)^{y[k]}\exp(-Nh_1)}{(N(h_0+h_1))^{y[k]}\exp(-N(h_0+h_1))} \\ &= \left(\log(h_1)-\log(h_0+h_1)\right)y[k]+Nh_0. \end{aligned} \tag{3.63}$$

As in the middle region, the approximation is given by a linear function, which depends on the channel coefficients and the number of released molecules. Interestingly, similar to the zero point, the result is independent of *a priori* information of the previous bit as shown in Fig. 3.15(a). It should be noted that the side region result evaluated at $y[k]=0$ is equal to the extrinsic LLR of the zero point.

To separate the middle region and side region, the intersection point $y_{\mathrm{IS}}[k]$ must be determined. This point is obtained by equating their corresponding extrinsic LLR and solving it with respect to $y[k]$:

$$y_{\mathrm{IS}}[k] = \frac{Nh_1 + L\left(u[k-1]\right)}{\log(h_0+h_1)-\log(h_0)}. \tag{3.64}$$

Alternatively, $y_{\mathrm{IS}}[k]$ can be determined by equating the two terms in the denominator of (3.60). Depending on channel coefficients, number of released molecules and *a priori* LLR, the size of the middle region changes. For $y_{\mathrm{IS}}[k]<1$ it vanishes completely and the PLD is completely described by the side region. This behavior is demonstrated in Fig. 3.15(a).

When comparing the extrinsic LLRs according to (3.59) and the approximated LLR of the PLD in Fig. 3.15, it can be observed that the approximation for the depicted cases shows only small negligible deviations. The deviation is maximum at the intersections between the regions. Overall, the PLD provides low-complexity detection, since only linear functions have to be evaluated. Alternatively, the zero point can be stored together with all intersection points and the extrinsic LLR can be determined by linear interpolation between the respective points. A major advantage over FTD and ATD is that PLD provides soft-output values in the form of LLRs.

3.2.4 Zero-Forcing Linear Equalization

In classical wireless communication, equalization methods are typically used to counteract the ISI of the transmission channel [PS07]. One method is the so-called zero-forcing (ZF)

detector, which is a linear equalizer. The idea is to interpret the EDTCM with its memory and channel coefficients as a finite impulse response (FIR) filter. The ZF equalizer with its filter coefficients g_j is designed in such a way that it acts as an inversion of the channel. Thus, it will cancel the ISI in the noise-free case completely. In the discrete-time domain, the convolution of the channel coefficients with the ZF filter coefficients should result in a unit impulse [PS07]:

$$\sum_{j=-\infty}^{\infty} g_j h_{\ell-j} \stackrel{!}{=} \begin{cases} 1 & \text{if } \ell = 0, \\ 0 & \text{if } \ell \neq 0. \end{cases} \tag{3.65}$$

In the z-domain this corresponds to

$$1 \stackrel{!}{=} H(z) \cdot G(z), \qquad G(z) \stackrel{!}{=} \frac{1}{H(z)}. \tag{3.66}$$

Since the channel behaves like an FIR filter, the ZF equalizer theoretically results in an infinite impulse response (IIR) filter. In practice, however, it is difficult to realize an IIR

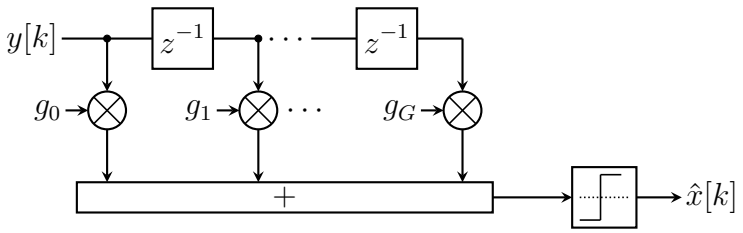


Figure 3.16: Block diagram of zero-forcing equalization.

filter with infinite filter coefficients, so the ZF equalizer is realized as an FIR filter with G filter coefficients, as shown in Fig. 3.16. This restriction leads to the fact that the ISI is not completely eliminated. However, the diffusion channel considered in this dissertation is causal, stable, and minimum-phase, therefore the ZF filter is also causal, stable, and minimum-phase. Accordingly, the influence of the non-considered IIR filter coefficients can be estimated as small. It should be noted that other channels may also lead to non-causal or anti-causal ZF equalizers, whereas Fig. 3.16 only depicts the causal case. After equalization, a detector without ISI consideration can be used. As shown in Fig. 3.16, in this dissertation the FTD presented in Section 3.2.1 is used to detect the transmitted symbols $\hat{x}[k]$.

3.2.5 Decision Feedback Equalization

The decision-feedback equalizer (DFE) is a nonlinear equalizer. It consists of an optional, often linear, feedforward filter and a nonlinear feedback filter [PS07]. The task of the feedforward filter is to shorten the length of the channel impulse response, while the feedback filter is designed to cancel out the remaining ISI. In this dissertation, a DFE

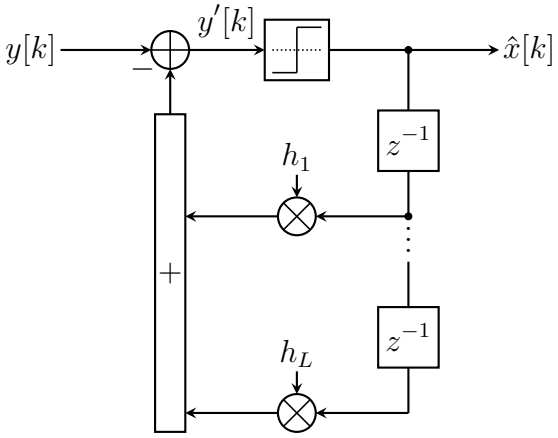


Figure 3.17: Block diagram of decision-feedback equalizer without feedforward filter.

without a feedforward filter is considered as shown in Fig. 3.17. In this special case, the filter coefficients are equal to the coefficients of the channel memory and the equalized received sample is [LL+12]

$$y'[k] = y[k] - \sum_{\ell=1}^{L} \hat{x}[k-\ell]h_\ell. \tag{3.67}$$

The expected ISI $\hat{x}[k-\ell]h_\ell$ is estimated and subtracted from the received sample $y[k]$ by the feedback of previous decided symbols $\hat{x}[k-\ell]$. The equalized receive sample $y'[k]$ can be detected by a detector that does not take ISI into account. In this dissertation, the FTD according to Section 3.2.1 is used for detection. If the previous symbols have been correctly detected, the DFE does not amplify the noise, which can occur in a ZF equalizer. However, if there is an error in the symbol detection, an error propagation occurs, which limits the performance of the DFE [PS07].

3.2.6 Maximum-Likelihood Sequence Estimation

The optimal detector in terms of minimizing the sequence error of the detected sequence follows the maximum-likelihood sequence estimation (MLSE) approach. According to the

MLSE approach, the most likely transmitted sequence is determined taking into account all possible sequences [PS07],

$$\hat{\boldsymbol{u}} = \arg\max_{\tilde{\boldsymbol{u}}} p(\boldsymbol{y}|\tilde{\boldsymbol{u}}), \tag{3.68}$$

whereby $\tilde{\boldsymbol{u}}$ is an info sequence hypothesis. A major advantage of the MLSE is that, in contrast to ZF, it does not lead to noise amplification and, in contrast to the DFE, it does not lead to error propagation. An elegant realization of the MLSE can be achieved by the Viterbi algorithm [For73], which searches for the most probable path in a trellis diagram. A trellis diagram consists of K trellis segments, where K is the number of symbols transmitted. This results in one trellis segment per transmission interval. Fig. 3.18

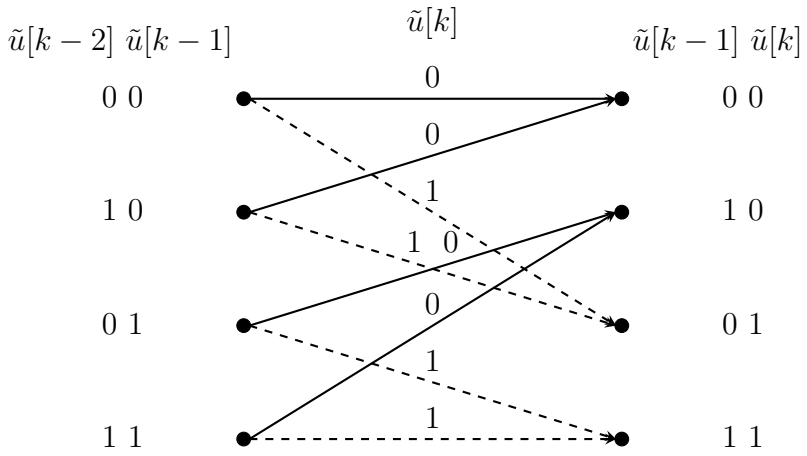


Figure 3.18: Trellis segment of a binary transmission with channel memory length $L = 2$.

shows an example of a trellis segment for a binary transmission with channel memory length $L = 2$. In general, a trellis segment consists of M^L states, where M corresponds to the modulation order of the transmitted symbols. The states describe all possible symbol combinations in the memory elements of the EDTCM. Accordingly, in the example in Fig. 3.18 there are $2^2 = 4$ states. A connection between two states is called a branch. Each state can be reached by M branches and from it leave M branches. The branches represent the symbol transmitted in the current time step, which is present in a memory element of the EDTCM in the next time step. There are only two possible symbols for binary transmission in Fig. 3.18. The solid branches represent the symbol which corresponds to $\tilde{u}[k] = 0$ and the dashed branches represent the symbol which corresponds to $\tilde{u}[k] = 1$. Costs are assigned to each valid state transition, which are called branch metrics. The branch metric from state $\boldsymbol{s}_i$ to $\boldsymbol{s}_j$ at time step k is denoted by $\lambda_{\boldsymbol{s}_i,\boldsymbol{s}_j}[k]$. The

sum of all branch metrics that lead to state $\boldsymbol{s}_j$ at time step k is called path metric and denoted by $\Gamma_{\boldsymbol{s}_j}[k]$. In general, the Viterbi algorithm consists of the following three steps:

1. *Initialization:* In this step all branch metrics are calculated. Additionally, assuming that the state memory is empty at the beginning of the transmission, the path metric of the first all-zero state is set to zero, while the path metrics of the remaining states are initialized to infinite.
2. *Add, compare, select:* This step is executed in chronological order for each new state $\boldsymbol{s}_j$ in the respective trellis segment. First, the new path metrics are determined according to $\Gamma_{\boldsymbol{s}_j}[k] = \Gamma_{\boldsymbol{s}_i}[k-1] + \lambda_{\boldsymbol{s}_i,\boldsymbol{s}_j}[k]$ (*add*). Then the different path metrics of the state $\boldsymbol{s}_j$ are compared with each other (*compare*). The path with the minimum cost is retained as the surviving path, while all other paths are discarded due to their higher cost (*select*).
3. *Trace-back:* At the end of the trellis diagram, the minimum path metrics of all final states are compared. The state with the lowest path metric represents the path of the most probable sequence. Starting from it, the path can be traced backwards through the trellis diagram. Each selected branch is directly assigned to a certain symbol and the transmitted sequence can be detected.

Calculation of Branch Metrics

Depending on the selected modulation scheme, there are slight differences in the branch metrics. In the following, the branch metrics for concentration-based modulation schemes such as OOK, CSK, and VCSK will be determined exemplarily. For other modulation schemes, the branch metrics can be determined in a similar fashion. According to the MLSE criterion in (3.68), the conditional probability that $y[k]$ molecules arrive must be calculated for each branch in the kth time step, under the assumption that symbol hypothesis $\tilde{\boldsymbol{x}}[k] = [\tilde{x}[k-L] \ \dots \ \tilde{x}[k]]$ was transmitted. This probability is referred to as the state transition probability $\gamma_{\boldsymbol{s}_i,\boldsymbol{s}_j}[k]$ and describes the probability of a transition from state $\boldsymbol{s}_i$ to state $\boldsymbol{s}_j$ in the trellis diagram. The corresponding symbol hypothesis $\tilde{\boldsymbol{x}}[k]$ results from state $\boldsymbol{s}_i$ and $\boldsymbol{s}_j$. According to the channel model in Section 2.3.1, the transition probability follows a Poisson binomial distribution, which can be elegantly represented using the discrete Fourier transform (DFT). The transition probability for $\nu_{\boldsymbol{s}_i,\boldsymbol{s}_j} > 0$ is given by [DK+17a]

$$\gamma_{\boldsymbol{s}_i,\boldsymbol{s}_j}[k] = P(\tilde{x}[k]) \frac{1}{1+\nu_{\boldsymbol{s}_i,\boldsymbol{s}_j}} \sum_{n=0}^{\nu_{\boldsymbol{s}_i,\boldsymbol{s}_j}} \exp\left(-j\frac{2\pi n y[k]}{1+\nu_{\boldsymbol{s}_i,\boldsymbol{s}_j}}\right) \Xi_{\boldsymbol{s}_i,\boldsymbol{s}_j}[n], \tag{3.69}$$

where

$$\Xi_{\boldsymbol{s}_i,\boldsymbol{s}_j}[n] = \prod_{\ell=0}^{L} \left(1 + \left(\exp\left(j\frac{2\pi n}{1+\nu_{\boldsymbol{s}_i,\boldsymbol{s}_j}}\right) - 1\right) h_\ell\right)^{\tilde{x}[k-\ell]}, \tag{3.70}$$

and for $\nu_{\boldsymbol{s}_i,\boldsymbol{s}_j} = 0$ by

$$\gamma_{\boldsymbol{s}_i,\boldsymbol{s}_j}[k] = \begin{cases} P(\tilde{x}[k]) & \text{if } y[k] = 0, \\ 0 & \text{if } y[k] > 0. \end{cases} \tag{3.71}$$

The variable $\nu_{\boldsymbol{s}_i,\boldsymbol{s}_j} = \sum_{\ell=0}^{L} \tilde{x}[k-\ell]$ describes the total number of molecules released into the channel. It should be noted that in case of maximum likelihood detection, as is the case of MLSE, no *a priori* information is considered. In this case, $P(\tilde{x}[k])$ should be chosen as a constant factor. The calculation of the transition probability is relatively complex despite the exploitation of the DFT. However, since the Poisson binomial distribution is a discrete distribution with a finite number of points, the transition probability distribution for each symbol hypothesis $\tilde{\boldsymbol{x}}[k]$ can be calculated in advance and stored in lookup tables. The branch metrics in the Viterbi algorithm are typically associated with costs that need to be minimized. Therefore, the error probability is used as branch metric instead of the transition probability

$$\lambda_{\boldsymbol{s}_i,\boldsymbol{s}_j}[k] = 1 - \gamma_{\boldsymbol{s}_i,\boldsymbol{s}_j}[k]. \tag{3.72}$$

For the binomial approximated [DK+17a], Gauss approximated [KA13; MY+14; DK+17a], and Poisson approximated [NC+14b; DK+17a] EDTCM, the corresponding transition probabilities are

$$\gamma^{\mathcal{B}}_{\boldsymbol{s}_i,\boldsymbol{s}_j}[k] = P(\tilde{x}[k]) \binom{N}{y[k]} p^{y[k]}_{\boldsymbol{s}_i,\boldsymbol{s}_j} \left(1 - p_{\boldsymbol{s}_i,\boldsymbol{s}_j}\right)^{N-y[k]}, \tag{3.73}$$

$$\gamma^{\mathcal{N}}_{\boldsymbol{s}_i,\boldsymbol{s}_j}[k] = \begin{cases} P(\tilde{x}[k]) \frac{1}{2}\left(1 + \operatorname{erf}\left(-\frac{\mu_{\boldsymbol{s}_i,\boldsymbol{s}_j}}{\sqrt{2\sigma^2_{\boldsymbol{s}_i,\boldsymbol{s}_j}}}\right)\right) & \text{if } y[k] = 0, \\ P(\tilde{x}[k]) \frac{1}{\sqrt{2\pi\sigma^2_{\boldsymbol{s}_i,\boldsymbol{s}_j}}} \exp\left(-\frac{(y[k]-\mu_{\boldsymbol{s}_i,\boldsymbol{s}_j})^2}{2\sigma^2_{\boldsymbol{s}_i,\boldsymbol{s}_j}}\right) & \text{if } y[k] > 0, \end{cases} \tag{3.74}$$

$$\gamma^{\mathcal{P}}_{\boldsymbol{s}_i,\boldsymbol{s}_j}[k] = P(\tilde{x}[k]) \frac{\lambda^{y[k]}_{\boldsymbol{s}_i,\boldsymbol{s}_j}}{y[k]!} \exp\left(-\lambda_{\boldsymbol{s}_i,\boldsymbol{s}_j}\right), \tag{3.75}$$

with

$$p_{\boldsymbol{s}_i,\boldsymbol{s}_j} = \sum_{\ell=0}^{L} \tilde{u}[k-\ell]h_\ell, \tag{3.76}$$

$$\mu_{\boldsymbol{s}_i,\boldsymbol{s}_j} = \lambda_{\boldsymbol{s}_i,\boldsymbol{s}_j} = \sum_{\ell=0}^{L} \tilde{x}[k-\ell]h_\ell, \tag{3.77}$$

$$\sigma^2_{\boldsymbol{s}_i,\boldsymbol{s}_j} = \mu_{\boldsymbol{s}_i,\boldsymbol{s}_j} - \sum_{\ell=0}^{L} \tilde{x}[k-\ell]h_\ell^2. \tag{3.78}$$

It should be noted that the binomial approximated EDTCM is only intended to be used for OOK transmission since it is based on the assumption that binomial distributions are merged with the same number of trials.

3.2.7 Reduced-State Sequence Estimation

The computational complexity of the Viterbi algorithm is proportional to its M^L states. Especially with higher-order modulation schemes and channels with a long memory length, MLSE becomes computationally complex. The reduced-state sequence estimation (RSSE) lowers the computational complexity by reducing the number of states. The idea is to shorten the memory length to $J < L$, which results in only M^J states inside the trellis. The residual ISI is taken into account by hard decisions according to the DFE principle [EQ88]. The main energy should therefore be in the first J channel coefficients to reduce the effect of error propagation. Consequently, a minimum phase channel impulse response, which is also typical in the diffusion-based channel under investigation, is desirable. The RSSE works in basically the same way as the MLSE in Section 3.2.6. There are only small differences in the calculation of transition probabilities through state reduction and consideration of hard decisions. For the Poisson binomial channel model, $\gamma_{\boldsymbol{s}_i,\boldsymbol{s}_j}[k]$ can be obtained according to (3.69) with the following difference:

$$\begin{aligned}\Xi_{\boldsymbol{s}_i,\boldsymbol{s}_j}[n] = &\prod_{\ell=0}^{J}\left(1+\left(\exp\left(j\frac{2\pi n}{1+\nu_{\boldsymbol{s}_i,\boldsymbol{s}_j}}\right)-1\right)h_\ell\right)^{\tilde{x}[k-\ell]} \\ &\cdot\prod_{\ell=J+1}^{L}\left(1+\left(\exp\left(j\frac{2\pi n}{1+\nu_{\boldsymbol{s}_i,\boldsymbol{s}_j}}\right)-1\right)h_\ell\right)^{\hat{x}[k-\ell]},\end{aligned} \tag{3.79}$$

$$\nu_{\boldsymbol{s}_i,\boldsymbol{s}_j} = \sum_{\ell=0}^{J}\tilde{x}[k-\ell] + \sum_{\ell=J+1}^{L}\hat{x}[k-\ell]. \tag{3.80}$$

For the binomial approximated, Gauss approximated [MY+14], and Poisson approximated [NC+14b] channel model, (3.73), (3.74), and (3.75) can be used with the small modification

$$p_{\boldsymbol{s}_i,\boldsymbol{s}_j} = \sum_{\ell=0}^{J} \tilde{u}[k-\ell]h_\ell + \sum_{\ell=J+1}^{L} \hat{u}[k-\ell]h_\ell, \tag{3.81}$$

$$\mu_{\boldsymbol{s}_i,\boldsymbol{s}_j} = \lambda_{\boldsymbol{s}_i,\boldsymbol{s}_j} = \sum_{\ell=0}^{J} \tilde{x}[k-\ell]h_\ell + \sum_{\ell=J+1}^{L} \hat{x}[k-\ell]h_\ell, \tag{3.82}$$

$$\sigma^2_{\boldsymbol{s}_i,\boldsymbol{s}_j} = \mu_{\boldsymbol{s}_i,\boldsymbol{s}_j} - \sum_{\ell=0}^{J} \tilde{x}[k-\ell]h_\ell^2 - \sum_{\ell=J+1}^{L} \hat{x}[k-\ell]h_\ell^2. \tag{3.83}$$

3.2.8 A Posteriori Probability Detection

In contrast to MLSE, the purpose of maximum APP detection is to minimize the bit error probability [PS07]

$$\hat{u}[k] = \arg\max_{\tilde{u}[k]} p(\tilde{u}[k]|\boldsymbol{y}). \tag{3.84}$$

To achieve this goal, the LLR for each info bit is determined in an APP detector as

$$L\left(u[k]|\boldsymbol{y}\right) \doteq \log \frac{P\left(u[k]=0|\boldsymbol{y}\right)}{P\left(u[k]=1|\boldsymbol{y}\right)} = \underbrace{\log \frac{P\left(\boldsymbol{y}|u[k]=0\right)}{P\left(\boldsymbol{y}|u[k]=1\right)}}_{L(\boldsymbol{y}|u[k])} + \underbrace{\log \frac{P\left(u[k]=0\right)}{P\left(u[k]=1\right)}}_{L(u[k])}, \tag{3.85}$$

where $L\left(u[k]|\boldsymbol{y}\right)$, $L\left(\boldsymbol{y}|u[k]\right)$, and $L\left(u[k]\right)$ are the *a posteriori*, extrinsic, and *a priori* LLRs, respectively. An elegant realization of the APP detector can be achieved by the Bahl-Cocke-Jelinek-Raviv (BCJR) algorithm [BC+74], which works similar to the MLSE on a trellis diagram. While the MLSE runs only in one direction through the trellis, the BCJR runs in both directions, replacing the final trace-back operation. The *a posteriori* LLR can be represented for the BCJR algorithm as follows:

$$L\left(u[k]|\boldsymbol{y}\right) = \log \frac{\sum\limits_{\mathcal{S}_0} P\left(\boldsymbol{s}_i, \boldsymbol{s}_j, \boldsymbol{y}\right)}{\sum\limits_{\mathcal{S}_1} P\left(\boldsymbol{s}_i, \boldsymbol{s}_j, \boldsymbol{y}\right)} = \log \frac{\sum\limits_{\mathcal{S}_0} \alpha_{\boldsymbol{s}_i}\left[k-1\right] \gamma_{\boldsymbol{s}_i,\boldsymbol{s}_j}\left[k\right] \beta_{\boldsymbol{s}_j}\left[k\right]}{\sum\limits_{\mathcal{S}_1} \alpha_{\boldsymbol{s}_i}\left[k-1\right] \gamma_{\boldsymbol{s}_i,\boldsymbol{s}_j}\left[k\right] \beta_{\boldsymbol{s}_j}\left[k\right]}. \tag{3.86}$$

Thereby $\mathcal{S}_0$ and $\mathcal{S}_1$ denote the set of all existing trellis transmissions from state $\boldsymbol{s}_i$ to $\boldsymbol{s}_j$ with hyothesis $\tilde{u}[k] = 0$ and $\tilde{u}[k] = 1$, respectively. The forward recursion probability is defined as

$$\alpha_{\boldsymbol{s}_j}\left[k\right] = \sum_{\boldsymbol{s}_i} \alpha_{\boldsymbol{s}_i}\left[k-1\right] \gamma_{\boldsymbol{s}_i,\boldsymbol{s}_j}\left[k\right] \tag{3.87}$$

and the backward recursion probability is defined as

$$\beta_{\boldsymbol{s}_i}[k-1] = \sum_{\boldsymbol{s}_j} \beta_{\boldsymbol{s}_j}[k]\, \gamma_{\boldsymbol{s}_i,\boldsymbol{s}_j}[k]\,. \tag{3.88}$$

Depending on the termination of the trellis diagram, $\alpha_{\boldsymbol{s}_i}[0]$ and $\beta_{\boldsymbol{s}_j}[K]$ can be initialized differently. If not stated otherwise, it is assumed that the trellis diagram is only terminated at the beginning of the transmission in the all-zero state. Accordingly, this results in

$$\alpha_{\boldsymbol{s}_i}[0] = \begin{cases} 1 & \text{if } \boldsymbol{s}_i = [0 \ldots 0]\,, \\ 0 & \text{otherwise,} \end{cases} \tag{3.89}$$

and

$$\beta_{\boldsymbol{s}_j}[K] = \frac{1}{M^L}. \tag{3.90}$$

Depending on the channel model, the transition probability $\gamma_{\boldsymbol{s}_i,\boldsymbol{s}_j}[k]$ can be determined according to (3.69), (3.73), (3.74), or (3.75).

3.2.9 Bit Error Rate Performance

To analyze the performance of the different algorithms, the BERs of the detectors are compared. A channel with memory length $L = 5$ and OOK modulation is assumed. Only for the analysis of the PLD, as in its derivation, a channel memory length of $L = 1$ is considered. It should be noted that the principle of the PLD can also be extended to longer channel memory lengths. Furthermore, it is assumed that all detectors have perfect channel knowledge. In order to investigate the influence of different system parameters, the BER for varying transmission distances, number of released molecules, and symbol durations is determined numerically.

Fig. 3.19 shows the BER for varying transmission distances. As the transmission distance increases, the channel impulse response becomes weaker and more dispersive. Consequently, the interference energy increases relative to the signal energy. As a result, the error probability of all detection algorithms, with the exception of ATD at short distances, increases with increasing transmission distance. The ATD is inherently based on ISI. For short transmission distances, however, the channel impulse response is at least nearly ISI-free, which reduces the probability of exceeding the adaptive threshold. Therefore, the ATD performs worse than the FTD at short transmission distances. In Fig. 3.19, the ATD shows

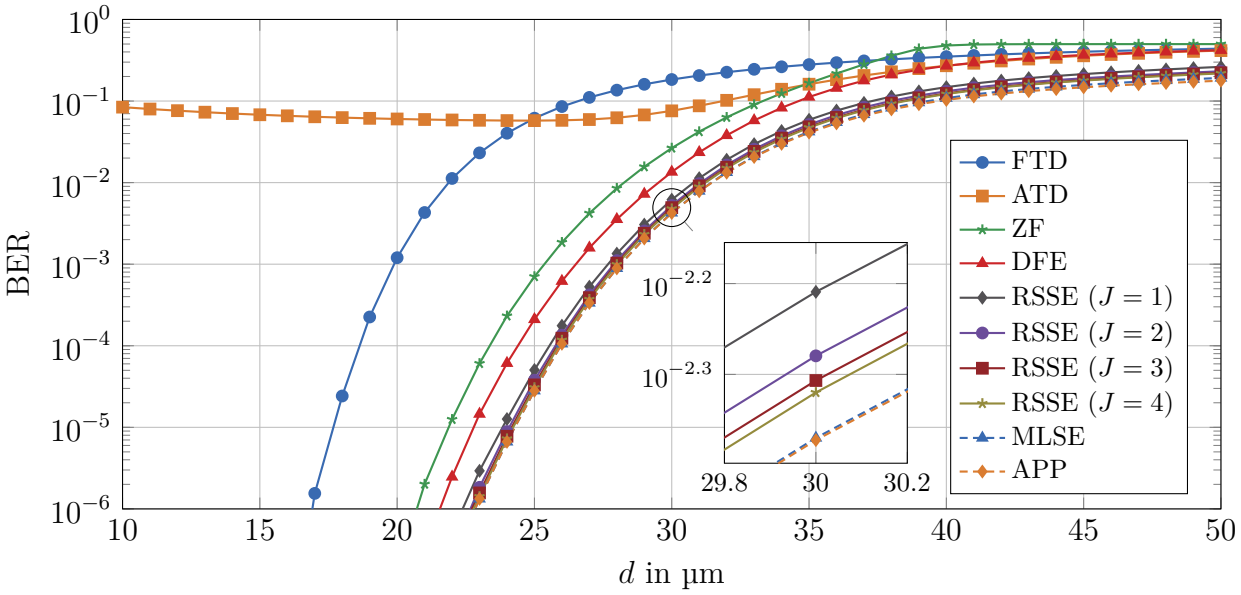


Figure 3.19: Bit error rate of different detection schemes for varying transmission distances. Parameters according to Appendix C.6.

a lower BER than the FTD from a distance of 25 µm. While the FTD, which does not take ISI into account, is increasingly less effective in detection, the ATD can benefit from the dispersive channel. It should be noted that FTD and ATD are low-complexity detectors. While FTD has to know the channel coefficients to determine the optimal threshold, ATD works without channel knowledge. If enhanced receiver structures with equalization are considered, performance increases with computational complexity. While the ZF and DFE essentially have the same computational complexity, the ZF is more error-prone than the DFE. This is mainly due to the inversion of the channel impulse response when determining the ZF filter coefficients. Especially with small channel coefficients, noise amplification can occur. The DFE, on the other hand, only feeds back hard decisions and thus does not amplify the noise. However, error propagation can occur if a decision is made incorrectly. The trellis based equalizers show the highest performance. Even the RSSE with $J = 1$ suffers from a small loss of performance compared to the MLSE. With increasing J, the RSSE approaches the computational complexity and BER performance of the MLSE. The MLSE achieves almost the same performance as the APP detector. The difference between the two algorithms is that the APP detector minimizes the bit error, while the MLSE minimizes the sequence error rate. Considering the BER, the APP detector is the optimal detector.

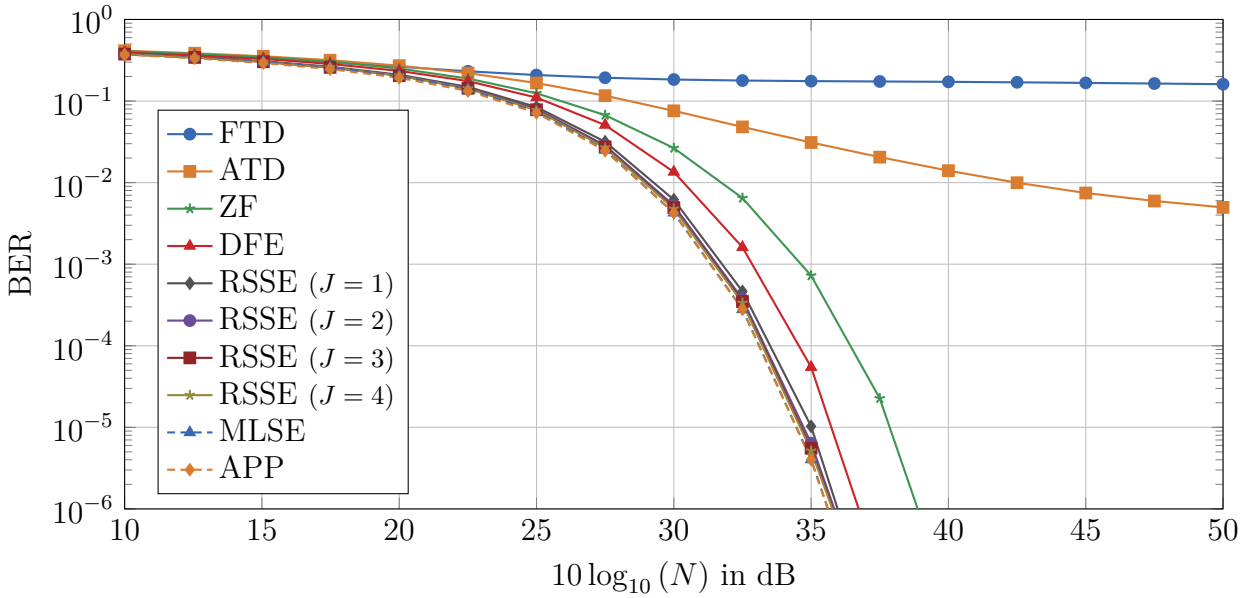


Figure 3.20: Bit error rate of different detection schemes for varying number of molecules. Parameters according to Appendix C.6.

Similar results can also be observed when changing the number of molecules and the symbol duration. Since the number of released molecules can be interpreted as signal power [GD+18], the BER over $10\log_{10}(N)$ in dB is shown in Fig. 3.20 analogously to classical wireless communication. If N and thus the signal energy is increased, the noise decreases relative to the signal power. As a result, the channel becomes more deterministic and the BER of all detectors decreases. If the symbol duration is increased, the interference power is reduced relative to the signal power. In addition, the relative noise power decreases because molecules are absorbed over a longer period of time. Consequently, the bit error probability decreases for all detectors with increasing symbol duration in Fig. 3.21. Only the ATD, which inherently is based on ISI, degrades when the symbol duration in the system under consideration exceeds $T = 0.7\,\mathrm{s}$.

To evaluate the detection performance of the PLD, the BER is analyzed at varying transmission distances for a channel with memory length $L = 1$. The numerically determined result is shown in Fig. 3.22. For comparison, the BER of the FTD, ATD, APP detector and the Poisson approximated symbol-wise APP detector according to (3.60) are shown as well. It can be concluded that the PLD has almost the same performance as the symbol-wise APP detector and thus represents a reasonable approximation. Both detectors have the same performance as the FTD for $d < 34\,\mu\mathrm{m}$. Due to the symbol-wise processing, $h_0 > h_1$,

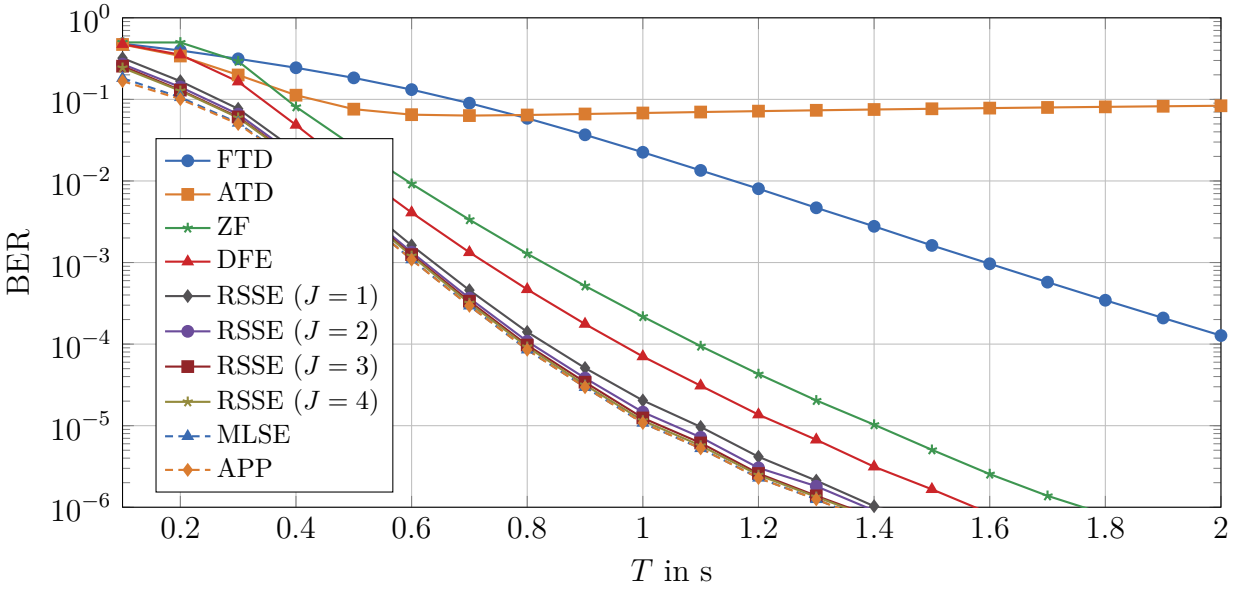


Figure 3.21: Bit error rate of different detection schemes for varying symbol duration. Parameters according to Appendix C.6.

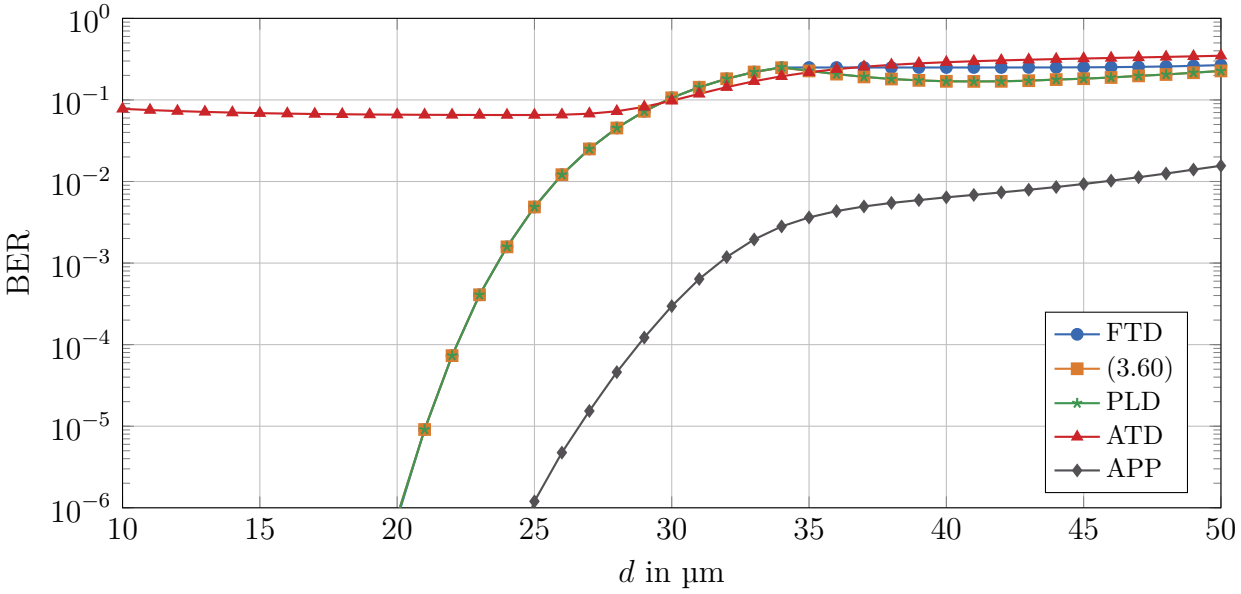


Figure 3.22: Bit error rate of different detection schemes for varying transmission distances and $L = 1$. Parameters according to Appendix C.6.

and the hard decision due to the sign of the LLRs, in principle only a single threshold detection takes place. The threshold is given by the zero-crossing in Fig. 3.15(a). Only beyond $d > 34\,\mu\text{m}$ the PLD and symbol-wise APP detector outperform the FTD. In this region $h_0 < h_1$ holds and thus two zero-crossings occur, which is not covered by the FTD with a single threshold. Furthermore, the PLD and symbol-wise APP detector provide soft information at their output, which can be used in combination with channel coding to achieve a significant increase in performance as demonstrated in Chapter 5. When compared with the (sequence-wise) APP detector, a significant performance difference is observed. It is evident that, if possible, sequence-wise processing is preferable to symbol-wise processing. However, this also involves a much higher computational complexity.

3.3 Orthogonal Frequency-Division Multiplexing

The modulation schemes discussed in Section 3.1 are suitable for single-carrier transmission. By contrast, OFDM can be used to modulate several subcarriers in parallel, which is known as multi-carrier modulation. Each subcarrier refers to a modulated frequency component of the transmitted OFDM signal. OFDM can be considered as a special form of frequency-division multiplexing, where the subcarriers are orthogonal to each other. Due to orthogonality the inter-carrier interference can be avoided. If the dashed boxes are neglected, Fig. 3.23 shows the block diagram of a classic OFDM system [PS07]. In the OFDM modulator (upper part in Fig. 3.23), the incoming bit sequence $\boldsymbol{u}$ is multiplexed into several parallel data streams. In each data stream, the bits are mapped to symbols $\boldsymbol{X}$ using any digital modulation scheme. With the help of an N_{DFT}-point inverse DFT (IDFT), the symbols are modulated onto N_{DFT} orthogonal subcarriers. Typically, each IDFT output block is extended by a cyclic repetition of L_{CP} symbols, called cyclic prefix (CP). Finally, the OFDM signal is parallel to serial converted and transmitted via the channel. In the OFDM demodulator (lower part in Fig. 3.23) all steps of the modulator are inverted and equalization is performed. If $L_{\text{CP}} \geq L$, the convolution of the OFDM signal with the channel becomes a cyclic convolution. In the frequency domain this implies an element-wise multiplication of the subcarrier signal $\boldsymbol{X}$ with the Fourier transformed channel coefficient $\boldsymbol{H} = \mathcal{F}_{N_{\text{DFT}}}\{\boldsymbol{h}\}$, where $\boldsymbol{h} = [h_0 \ldots h_L]$ is the vector of the channel coefficients in the time domain. Consequently, OFDM transforms a frequency-selective channel into parallel non-dispersive channels. Thus, ISI is avoided completely and equalization reduces to a low-complexity single-tap equalizer at the receiver side. If $N_{\text{DFT}} \gg L_{\text{CP}}$, the losses of power efficiency and bandwidth efficiency are negligibly small [PS07]. This is a key motivation of

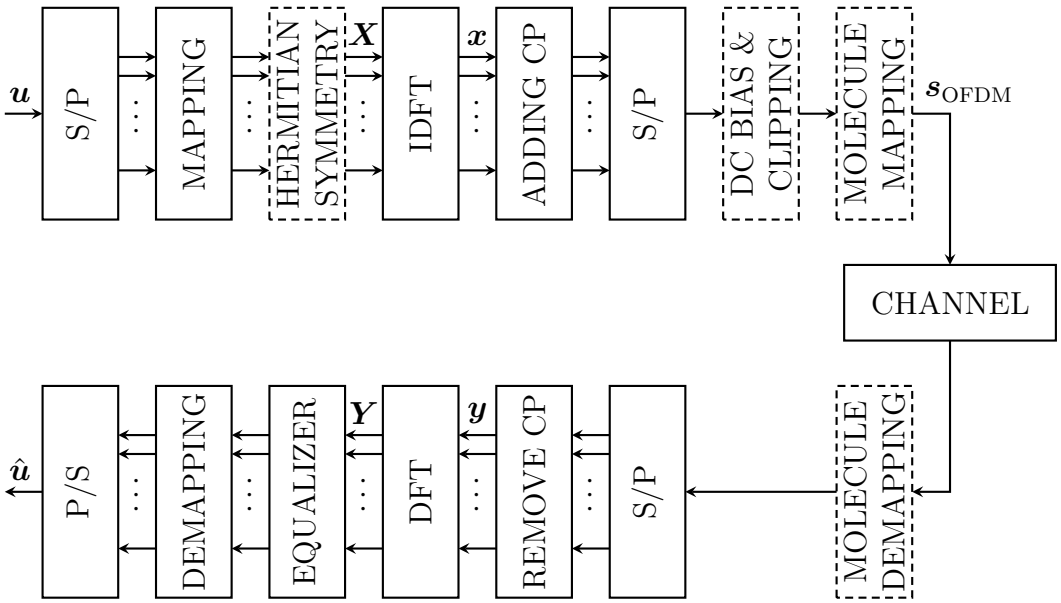


Figure 3.23: Block diagram of an OFDM system for molecular communication. Dashed boxes highlight additional blocks compared to a standard OFDM system.

using OFDM in dispersive channels. Furthermore, OFDM offers an elegant way to support arbitrary digital modulation schemes. By means of bit-loading [GA04], the data-to-carrier mapping can be adapted to the frequency characteristic of the channel to further improve the system performance.

The output signal of the OFDM modulator is typically complex-valued and bipolar. In molecular communication, however, it is only possible to release a non-negative integer number of molecules. A prerequisite for OFDM in molecular communication [DK$^+$17b] is that the OFDM signal must be positive and real-valued. The same requirements hold for intensity modulated direct detection in optical communication, where the intensity of light is a real-valued non-negative information carrier. The requirements are met by restricting the IDFT input to Hermitian symmetry and by clipping (limiting all negative values to zero) before transmission over the channel. Optionally, a direct current (DC) bias can be applied to reduce the effect of clipping. These steps are depicted by the dashed boxes in Fig. 3.23. Note that the OFDM demodulator does not differ from a classical OFDM demodulator. In addition, the signal has to be mapped to molecules and rounded before transmission via the channel to ensure integer values as shown in Fig. 3.24.

In the following, three OFDM modifications known from the optical community are introduced, which are adapted for molecular communication systems [DK$^+$17b]. Furthermore,

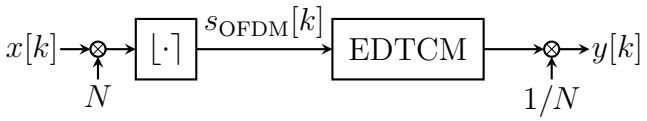


Figure 3.24: Block diagram of the considered molecule mapping, channel model, and molecule demapping.

a bit-loading algorithm suitable for the diffusion-based molecular communication channel is proposed.

3.3.1 Direct Current Biased Optical OFDM

In DC biased optical OFDM (DCO-OFDM) [CK96], the data symbols are mapped to the N_{DFT}-point IDFT input $\boldsymbol{X} = [X_0 \ \dots \ X_f \ \dots \ X_{N_{\mathrm{DFT}}-1}]$ with subject to Hermitian symmetry

$$X_f = X^*_{N_{\mathrm{DFT}}-f}, \qquad f \in \{1, \dots, N_{\mathrm{DFT}}/2 - 1\}, \tag{3.91}$$

where X^* denotes the complex conjugate of X. The DC subcarrier and the Nyquist tone subcarrier are set to zero:

$$X_0 = X_{N_{\mathrm{DFT}}/2} = 0. \tag{3.92}$$

Thus, the IDFT output block $\boldsymbol{x} = [x_0 \ \dots \ x_n \ \dots \ x_{N_{\mathrm{DFT}}-1}]$ consists of real numbers. To reduce the clipping noise (impact of clipping negative values), a DC bias is added to the signal after CP extension. Note that a DC bias leads to power inefficiency, since the power of the DC bias is not used for data modulation. Furthermore, the amplitude-dependent diffusion noise will increase by the DC bias. The DC bias can be either fixed or adaptive. The optimal fixed DC bias is for example addressed in [ZZ14] and varies for different modulation schemes. If the smallest negative value of each IDFT output block is chosen as an adaptive DC bias, clipping noise is completely avoided.

3.3.2 Asymmetrically Clipped Optical OFDM

The concept of asymmetrically clipped optical OFDM (ACO-OFDM) [AL06] differs from DCO-OFDM in the missing DC bias and in the organization of the IDFT input $\boldsymbol{X}$. While (3.91) still holds, only the odd subcarriers are used for data transmission. The even subcarriers are set to zero. Consequently, the IDFT output $\boldsymbol{x}$ will have an antisymmetric

property:

$$x_n = -x_{n+N_{\mathrm{DFT}}/2}, \qquad n \in \{0, \ldots, N_{\mathrm{DFT}}/2 - 1\}. \tag{3.93}$$

Thus, clipping noise only occurs in the even unmodulated subcarriers. In the odd subcarriers, clipping reduces the signal amplitude by a factor of two.

3.3.3 Pulse Amplitude Modulation Discrete Multitone Transmission

In pulse amplitude modulation discrete multitone transmission (PAM-DMT) [LR+09] equations (3.91) and (3.92) hold. Only the imaginary part of each subcarrier is used to transmit bipolar pulse amplitude modulation (PAM) symbols. The IDFT output block $\boldsymbol{x}$ will have the antisymmetric property:

$$x_n = -x_{N_{\mathrm{DFT}}-n}, \qquad n \in \{1, \ldots, N_{\mathrm{DFT}}/2 - 1\}. \tag{3.94}$$

Clipping noise will only occur in the real part of subcarriers. The imaginary part will be scaled by a factor of 1/2 after clipping.

3.3.4 Bit-loading

In conventional OFDM, the number of bits per subcarrier is typically equally allocated. In frequency-selective channels, however, the channel quality might differ remarkably between subcarriers. As a result, the BER is dominated by subcarriers with weak quality. In such scenarios, bit-loading can be applied to improve the BER performance. By means of bit-loading, the number of bits per subcarrier is allocated with respect to the channel quality, so that more bits are transmitted over stronger channels and less bits are transmitted over weaker channels. Typically, the bit-loading is determined by the subcarrier-wise signal-to-noise ratio (SNR) at the receiver side. In molecular communication, the SNR in time domain changes proportional to the received signal strength due to the amplitude-dependent diffusion noise [WH+15]. Diffusion noise can be considered to be similar to shot noise in optics, which is approximately a white noise process [BL90]. Therefore, the noise power in frequency domain is assumed to be constant. Consequently, the channel coefficients (i. e., the average signal strength) can be exploited for bit-loading in frequency domain. The channel strengths of OFDM subchannels is described by the N_{DFT}-point DFT of the channel coefficients $\boldsymbol{H} = \mathcal{F}_{N_{\mathrm{DFT}}}\{\boldsymbol{h}\}$. The number of bits allocated to the fth subcarrier

can be determined by [DK+17b]

$$B_f = \left\lfloor \frac{|H_f|}{\sum_{i \in F} |H_i|} K_{\mathrm{b,DFT}} \right\rfloor, \qquad f \in F, \tag{3.95}$$

where H_f is the fth entry of $\boldsymbol{H}$, $K_{\mathrm{b,DFT}}$ is the total number of bits allocated to the IDFT input block, and F is the set of modulated subcarrier indices that are used for data transmission. Note that in radio-based communication H_f^2 is typically used to describe the signal power. In molecular communication, however, H_f already represents the signal power as in optical communication.

3.3.5 Bit Error Rate Performance

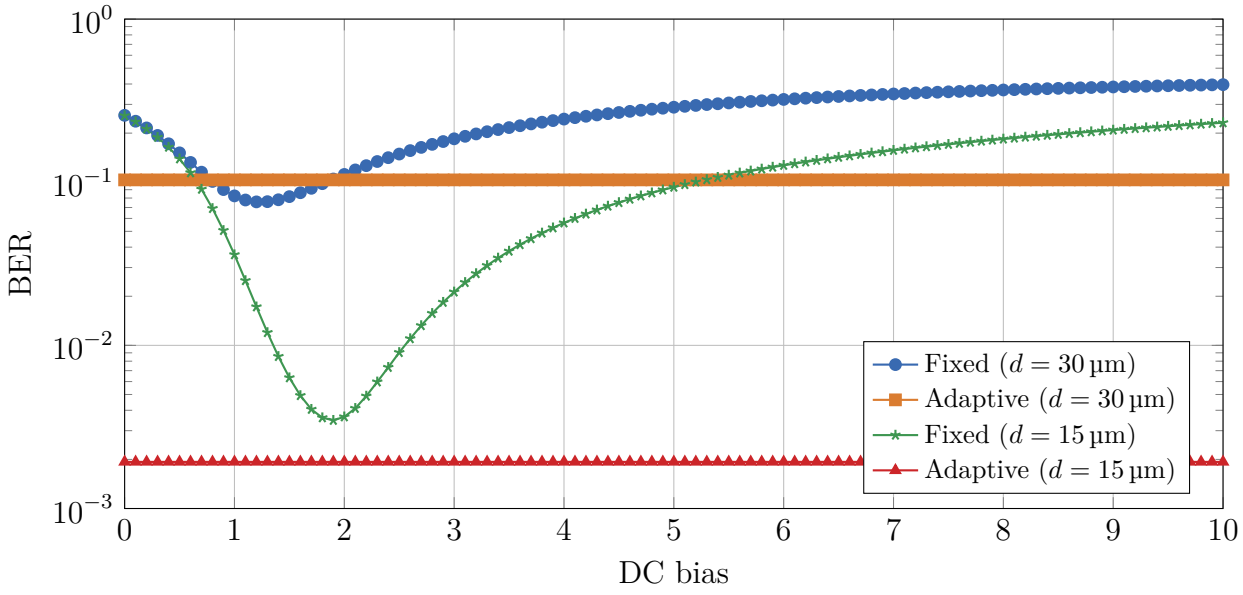


Figure 3.25: Bit error rate comparison between DCO-OFDM scheme with adaptive and fixed DC bias. Parameters according to Appendix C.7.

Effect of DC bias In Fig. 3.25, the effect of DC bias on the BER performance of DCO-OFDM is shown. There exists a trade-off in selecting the DC bias. If the DC bias is chosen too small, a large part of the information is clipped. Consequently, the clipping noise increases for smaller DC biases and the system performance degrades remarkably. On the other hand, the system performance also degrades for a large DC bias. This is due to the amplitude-dependent noise that increases with the DC bias. Note that the DC bias does not carry any information and hence is not used for detection. As a result, the optimal DC bias has to be selected carefully for a DCO-OFDM system and a given scenario. For the scenario shown in Fig. 3.25, the BER performance of an adaptively selected DC bias is shown as a straight line for comparison. At shorter transmission distances, the adaptive DC bias shows a better BER performance than a fixed DC bias. As a result, the adaptive DC bias DCO-OFDM system will be considered in the remaining analysis.

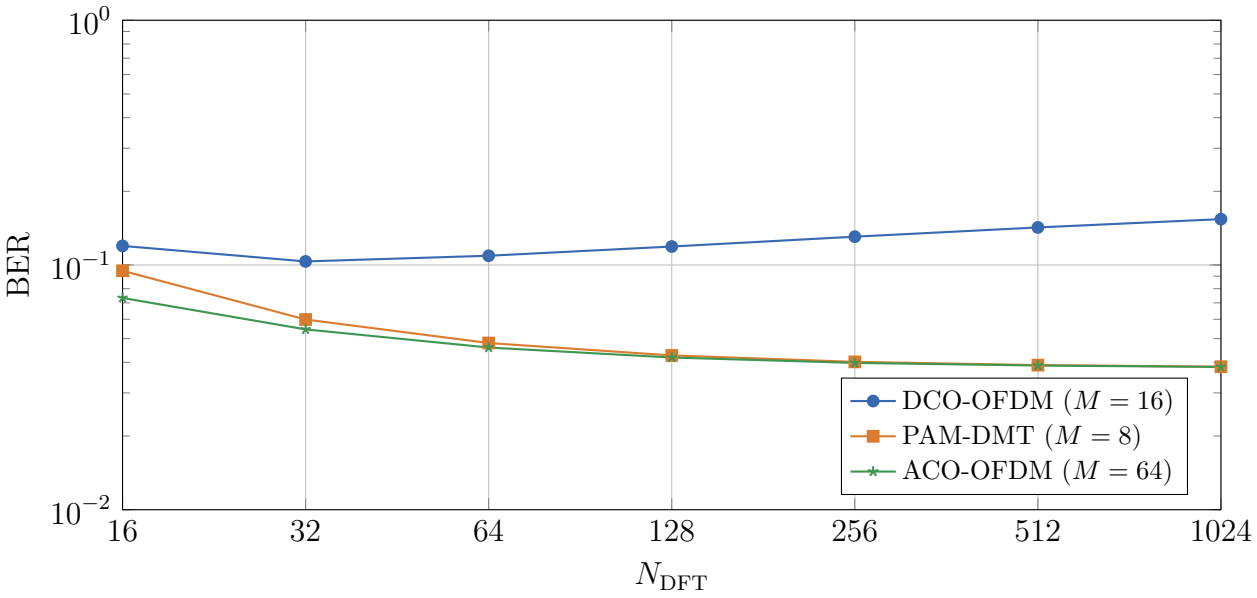


Figure 3.26: Bit error rate comparison between different OFDM schemes for varying DFT size. Parameters according to Appendix C.8.

Effect of OFDM Parameters In Fig. 3.26, the effect of the IDFT length on the OFDM performance is shown. The CP in each IDFT block leads to overhead, which decreases with an increasing IDFT length given a fixed channel memory length L. Considering normalization, the symbol duration and amount of molecules per OFDM symbol will increase consequently with an increasing IDFT length. Therefore, the BER improves for ACO-OFDM and PAM-DMT in Fig. 3.26. In contrast, the BER of DCO-OFDM increases

for $N_{\mathrm{DFT}} > 16$. This behavior is due to the adaptively selected DC bias. For larger IDFT length, the probability of selecting a high DC bias for each IDFT block is larger. Thus, the system performance will be dominated by stronger diffusion noise as discussed previously. In addition it can be stated that PAM-DMT and ACO-OFDM perform similar for large N_{DFT}.

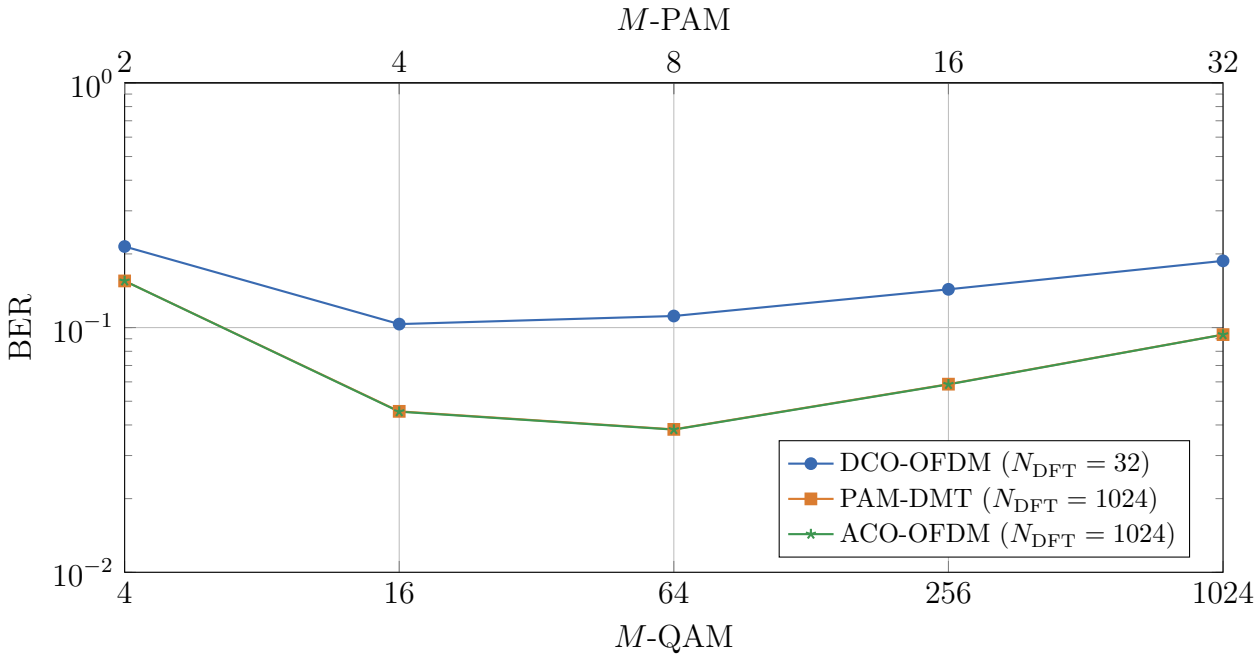


Figure 3.27: Bit error rate comparison between different OFDM schemes for varying modulation order. Parameters according to Appendix C.8.

In Fig. 3.27, the effect of the modulation order M on the OFDM performance is shown. It should be noted that an M-QAM symbol consists of two $\sqrt{M}$-PAM symbols, one in each dimension. There exists a trade-off between two major effects: The first effect describes the degrading BER performance when using higher-order modulation schemes. Due to constant symbol energies for all modulation orders, neighboring symbols are getting closer while increasing the modulation order. Consequently, the modulation scheme is getting less robust against noise. The second effect is due to the normalization. If the modulation order becomes larger, more bits are transmitted per symbol. Consequently, less symbols have to be transmitted in the same time interval, thus the symbol duration and the number of molecules per symbol is getting larger. As a result, the modulation order has to be selected carefully in an OFDM system. In the scenario under investigation, the optimal choice is 16-QAM for DCO-OFDM, 8-PAM for PAM-DMT and 64-QAM for ACO-OFDM.

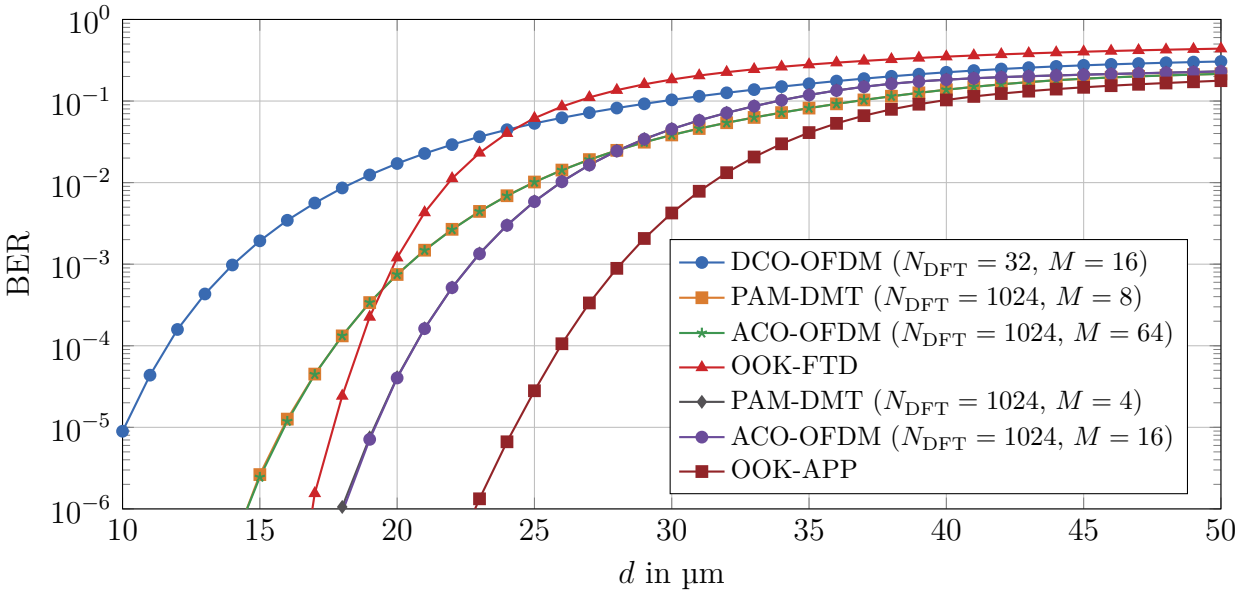


Figure 3.28: Bit error rate comparison between different OFDM schemes and single-carrier OOK transmission for varying transmission distances. Parameters according to Appendix C.8.

Effect of Transmission Distance In Fig. 3.28, the effect of the transmission distance on the BER performances of OFDM is investigated. As a comparison, the performance of an OOK transmission system with the same information rate and number of molecules per info bit is shown with FTD and APP detection. Longer distances decrease the probability of molecules to get absorbed by the receiver. Consequently, the BER performances increase due to a relatively increasing diffusion noise. The performance of ACO-OFDM is approximately equal to the performance of PAM-DMT, while DCO-OFDM performs the worst. The advantage of ACO-OFDM and PAM-DMT over DCO-OFDM is the absence of DC bias and clipping noise in modulated subcarriers. Compared to OOK systems, the performance of OFDM methods falls in-between OOK-FTD and OOK-APP for $d \geq 25\,\mu\mathrm{m}$. This also applies to ACO-OFDM and PAM-DMT at shorter distances, if M and N_{DFT} are selected appropriately. While ACO-OFDM with $N_{\mathrm{DFT}} = 1024$ and $M = 64$ provides the minimum BER for a distance of $d = 30\,\mu\mathrm{m}$ and the OFDM parameters under investigation, ACO-OFDM with $N_{\mathrm{DFT}} = 1024$ and $M = 16$ is selected for a minimum BER at $d = 20\,\mu\mathrm{m}$. The result shows that, depending on the transmission channel, the OFDM system parameters must be carefully adjusted. It should be noted that APP detection suffers from higher computational complexity when M and L increase. In contrast, there is only a single-tap equalizer applied in OFDM.

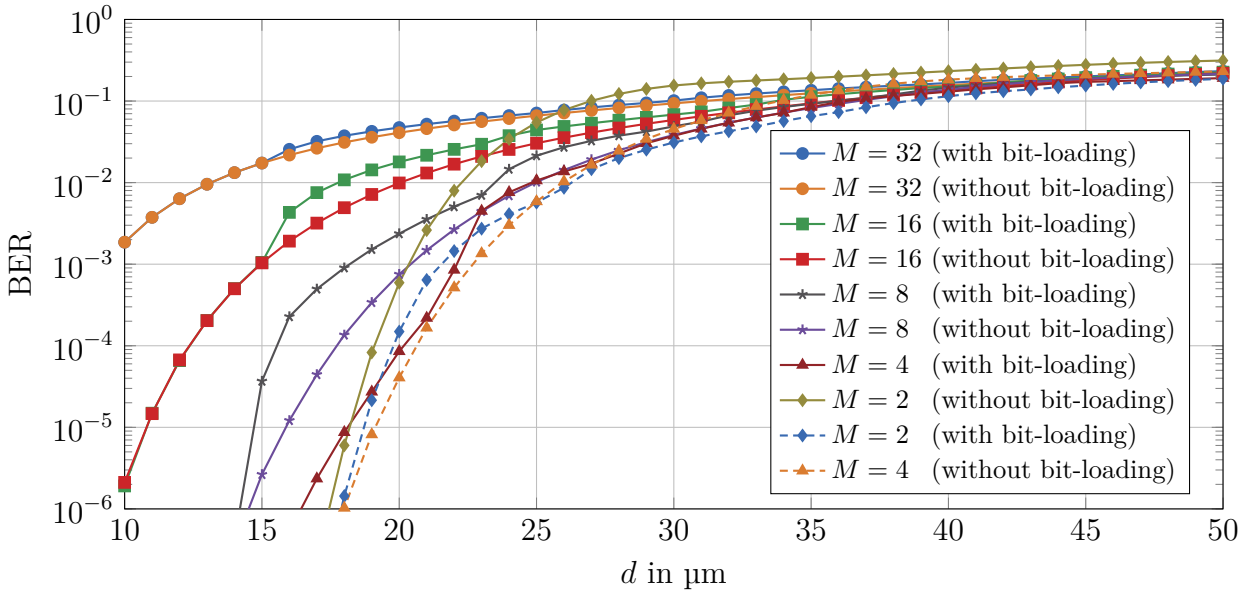


Figure 3.29: Bit error rate comparison between different PAM-DMT schemes with and without bit-loading for varying transmission distances. Parameters according to Appendix C.8.

Effect of Bit-loading In Fig. 3.29, the effect of bit-loading on the BER performance for different distances is shown for PAM-DMT. While the same number of bits per carrier is allocated without bit-loading, with bit-loading it is selected according to (3.95). The number of info bits per IDFT block is kept constant to the case without bit-loading and modulation order M. In the scenario under investigation, only for $M = 2$ bit-loading indicates a significant performance improvement. The performance degradation can be analyzed with Fig. 3.30, which shows the magnitude of the Fourier transformed channel coefficients $\boldsymbol{H}$. Therefore, $|\boldsymbol{H}|$ is symmetrical. For larger distances, the overall subcarrier strength decreases sharply. In addition, the gradient between the subcarrier strengths increases. This leads to strong subcarriers that are loaded with significantly higher amounts of info bits. Here, however, the effect discussed in Fig. 3.27 occurs, in which higher-order constellation points have less distance to neighboring points and are therefore less robust against noise. If this negative effect exceeds the positive effect of bit-loading, the bit-loading system performs worse than the reference system without bit-loading. It should be noted that Fig. 3.30 confirms the typically observed low-pass characteristic of the diffusion-based channel.

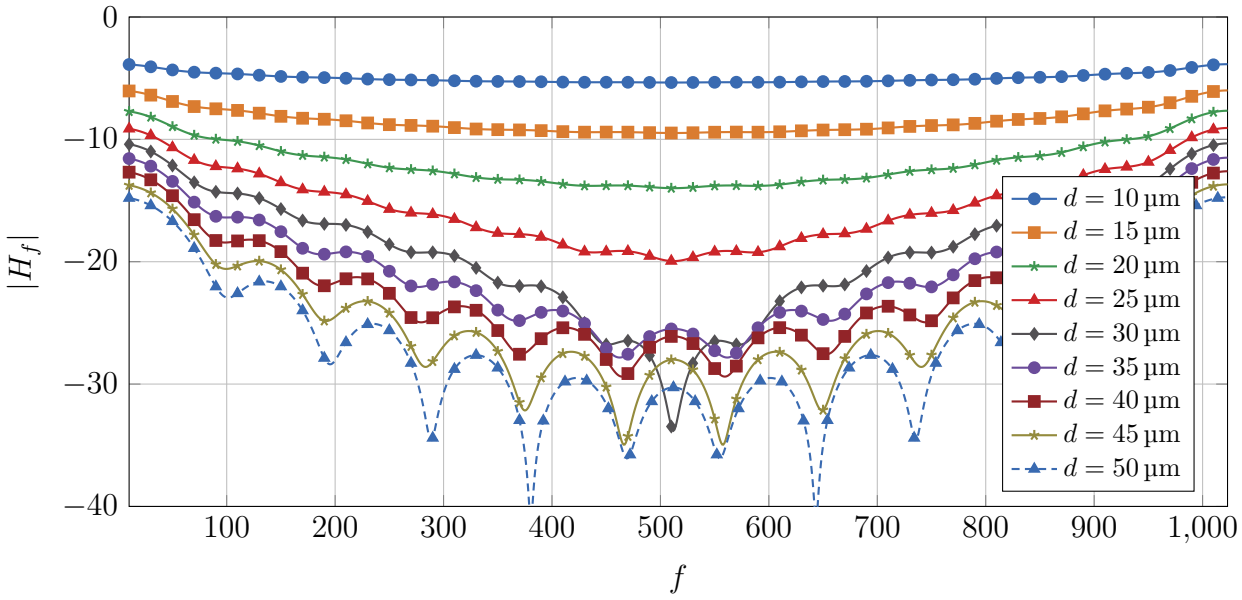


Figure 3.30: Channel strength analysis in Fourier domain for PAM-DMT transmission ($N_{\mathrm{DFT}} = 1024$, $M = 2$). Parameters according to Appendix C.9.

4

Channel Coding Algorithms

Channel codes are typically used to increase the robustness of data transmission against transmission errors. In the following sections, different coding algorithms are analyzed with respect to their performance in the diffusion-based channel. Depending on the type of code, a distinction is made between forward error correction (FEC) classes block codes, convolutional codes, line codes, spreading codes, and spatial codes. The main focus is on binary coding algorithms. In order to reduce the robustness against burst errors and to increase the independence of the error events at the decoder, with the exception of Reed-Solomon and intersymbol interference (ISI)-free codes, the block codes and convolutional codes are implemented with a random interleaver between encoder and modulator and a corresponding deinterleaver between detector and decoder. The ISI-free code, as well as the line codes, spreading codes, and spatial codes are implemented without interleavers, since their code structure is matched to the channel characteristic, which would get lost by interleaving. The Reed-Solomon code is not combined with interleaving, because it is already inherently robust against burst errors. In addition, with the exception of the low-density parity check (LDPC) code, the transmission of short info sequences of length $K_\mathrm{c} = 1000$ is considered, since diffusion-based molecular communication is typically a comparatively slow type of transmission and especially in microscopic applications typically only short messages have to be exchanged.

In order to evaluate and compare the performance of the coding strategies, bit error rate (BER) simulations are performed. Thereby, an on-off keying (OOK) modulated transmit signal and an *a posteriori* probability (APP) detector is considered, which is introduced in Section 3.1.1 and Section 3.2.8, respectively. A distinction is made between three different normalizations [DS+19]:

1. Unnormalized: In the unnormalized case, the symbol duration and the mean number of molecules per symbol remain the same as in the uncoded case.

2. T normalized: The symbol rate is normalized to the information rate of the uncoded transmission. Thus the time required for the transmission of an info bit is the same in the uncoded and coded transmission.

3. T & N normalized: In addition to the information rate, the average number of molecules required per info bit in coded transmission is selected to be equal to the uncoded transmission.

Depending on the application and requirements, the different normalizations can be of interest. In a targeted drug release scenario where the release command is transmitted, a particularly reliable transmission is of interest, while the time component is rather uncritical. In this example, the unnormalized perspective is a useful option. In a macroscopic transmission the goal could be to improve the reliability of the transmission without reducing the transmitted information per time and without using additional extra resources. In this case, the normalized symbol duration and normalized number of released molecules is appropriate.

4.1 Block Codes

Block codes, e. g. comprehensively studied in [LC01], are blockwise encoded and decoded. The info sequence is therefore divided into blocks of length $k_\mathrm{c} \in \mathbb{N}$. Each of these blocks represents an info word $\boldsymbol{u}_\mathrm{c} = [u_0 \ldots u_{k_\mathrm{c}-1}]$ and is independently encoded to a code word $\boldsymbol{x}_\mathrm{c} = [x_0 \ldots x_{n_\mathrm{c}-1}]$ of length $n_\mathrm{c} > k_\mathrm{c}$ with $n_\mathrm{c} \in \mathbb{N}$. With an $(n_\mathrm{c}, k_\mathrm{c})$ block code the 2^{k_c} info words are bijectively assigned to 2^{k_c} code words. The mapping is time-invariant and memoryless and the set of all code words forms the code with the code rate $R = {}^{k_\mathrm{c}}/_{n_\mathrm{c}}$. Since only a subset of the 2^{n_c} possible code words is selected, redundancy is added which can be used for error correction and/or error detection. The maximum performance of a block code can be described by its minimum distance d_min. This specifies the minimum Hamming distance between any two distinct code words, where the Hamming distance describes the number of positions where two code words differ. Thus, a block code with hard-input can detect a maximum of $d_\mathrm{min} - 1$ errors and correct a maximum of $\lfloor (d_\mathrm{min} - 1)/2 \rfloor$ errors. If a block code is systematic, the info word is directly part of the code word and the remaining bits $[p_0 \ldots p_{n_\mathrm{c}-k_\mathrm{c}-1}]$ are called parity bits. A cyclic shift of cyclic code words will result in a valid code word. In a linear block code, a linear combination of code words will lead

to a valid code word. For linear block codes, the encoding can be described by a $k_\mathrm{c} \times n_\mathrm{c}$ generator matrix $\boldsymbol{G}$ and the decoding by a $(n_\mathrm{c} - k_\mathrm{c}) \times n_\mathrm{c}$ parity check matrix $\boldsymbol{P}$

$$\boldsymbol{x}_\mathrm{c} = \boldsymbol{u}_\mathrm{c}\boldsymbol{G}, \tag{4.1}$$

$$\mathbf{0} = \boldsymbol{x}_\mathrm{c}\boldsymbol{P}^\top, \tag{4.2}$$

where $\mathbf{0}$ is the zero vector. Consequently, $\boldsymbol{G}\boldsymbol{P}^\top = \mathbf{0}$ must apply. A generator matrix can be converted into a parity check matrix and vice versa. The following relationship exists between the standard forms of a generator matrix and the corresponding parity check matrix:

$$\boldsymbol{G} = \begin{bmatrix} \boldsymbol{I}_{k_\mathrm{c}} & \boldsymbol{P}_\mathrm{c} \end{bmatrix}, \qquad \boldsymbol{P} = \begin{bmatrix} -\boldsymbol{P}_\mathrm{c}^\top & \boldsymbol{I}_{n_\mathrm{c}-k_\mathrm{c}} \end{bmatrix}, \tag{4.3}$$

where $\boldsymbol{I}_{k_\mathrm{c}}$ is the $k_\mathrm{c} \times k_\mathrm{c}$ identity matrix and $\boldsymbol{P}_\mathrm{c}$ is the $k_\mathrm{c} \times (n_\mathrm{c} - k_\mathrm{c})$ matrix including the parity checks. The Gauss-Jordan elimination [AM87] can be used to convert a matrix into the standard form. Alternatively, some channel codes are described by their generator polynomial $g(z)$. In this case, valid code words are designed in such a way that their polynomial representation is completely divisible by the generator polynomial.

4.1.1 Repetition Code

Probably the most intuitive way to add redundancy for error detection and error correction is to simply repeat the bits. In a repetition code, each info bit is repeated $n_\mathrm{c} - 1$ times:

$$[x_0 \ \ldots \ x_{n_\mathrm{c}-1}] = [u \ \ldots \ u]. \tag{4.4}$$

The repetition code can be represented as a linear systematic $(n_\mathrm{c}, k_\mathrm{c} = 1)$ block code with code rate $R = 1/n_\mathrm{c}$. Its generator matrix and parity check matrix are given as

$$\boldsymbol{G} = \begin{bmatrix} 1 & \cdots & 1 \end{bmatrix}, \qquad \boldsymbol{P} = \begin{bmatrix} 1 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 1 & 0 & \cdots & 0 & 1 \end{bmatrix}. \tag{4.5}$$

The minimum distance of repetition codes is $d_\mathrm{min} = n_\mathrm{c}$. Thus it is able to detect $n_\mathrm{c} - 1$ bit errors and correct $\lfloor (n_\mathrm{c} - 1)/2 \rfloor$ bit errors during a hard-input decoding. The error correction can be performed by a simple majority decision. However, a soft-input decoding [Joh10] is

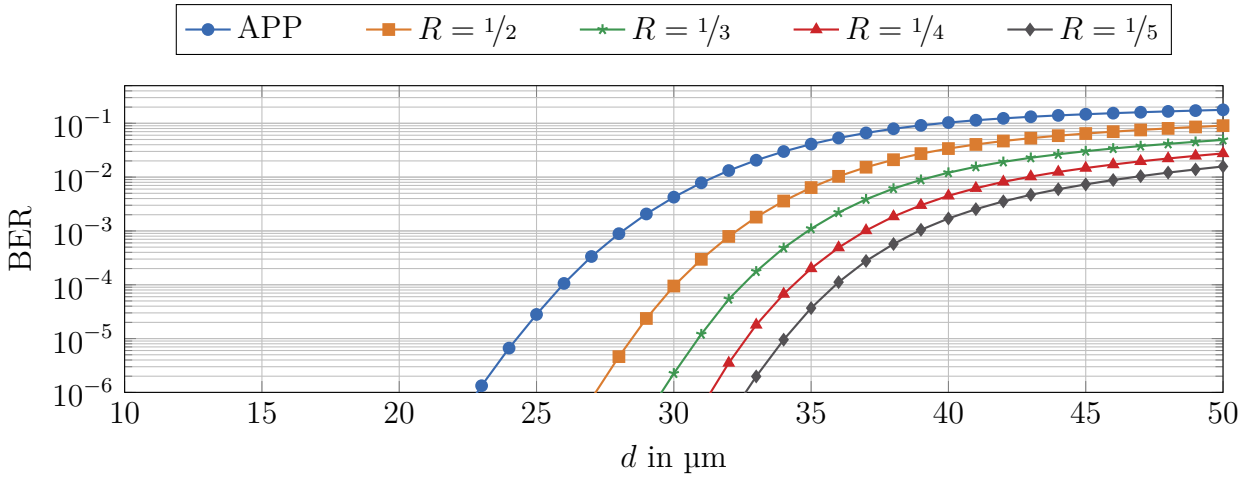
much more powerful. Thereby the log-likelihood ratio (LLR) values of the received code word are summed and up to $n_c - 1$ bit errors can be corrected. Repetition codes with hard-input decoding are studied in [SL+12] in the area of molecular communication. Due to the better performance in general, repetition codes in combination with soft-input decoding are considered in this dissertation.

Numerical Results

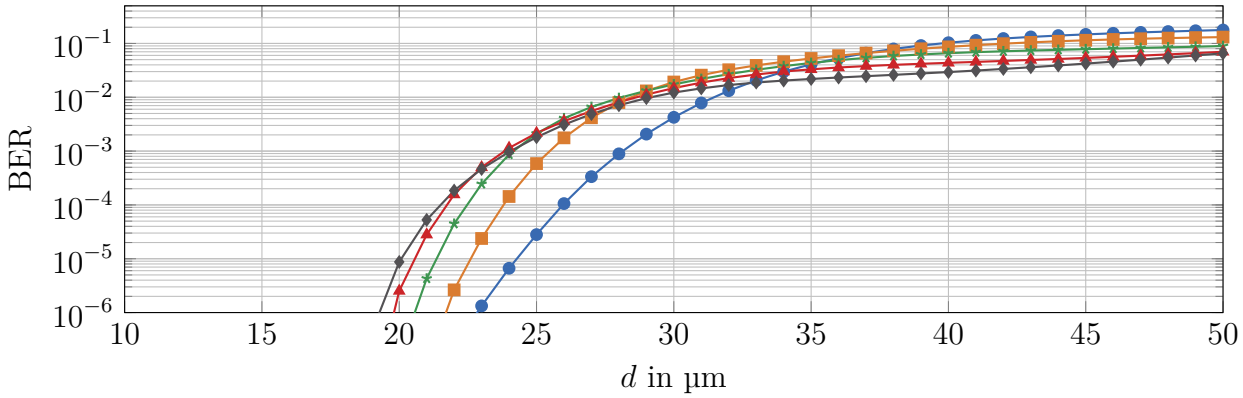
Fig. 4.1 shows numerical results of the BER analysis. To evaluate the gain achieved by a repetition code, the coded transmission is compared to the uncoded transmission. In the unnormalized examination in Fig. 4.1(a), it is evident that a gain is achieved by channel coding and the corresponding error correction. Since the code rate R is inversely proportional to the introduced redundancy, the lower the code rate, the greater the gain achieved in the unnormalized case. Consequently, the $R = 1/5$ repetition code, which has the lowest investigated code rate, shows the highest gain. At a target BER of 10^{-3}, the maximum transmission distance can be increased from 28 µm to 39 µm. The maximum transmission distance can be further increased with lower-rate repetition codes.

If the symbol duration is normalized to a fixed information rate, as in the case shown in Fig. 4.1(b), the general behavior of the coded transmission is inverted. Since the redundancy added by the channel code does not carry any new information, the symbol duration in the coded transmission is shorter than in the uncoded transmission to ensure the same information rate. The lower the code rate, the shorter the symbol duration. A reduction of the symbol duration leads to an increased ISI. In addition, fewer molecules reach the receiver during a transmission interval. This leads to an increase of noise level, since the diffusion-based channel is affected by amplitude-dependent noise. In the scenario under investigation, the reduction of the symbol duration has a stronger influence than the gain from the introduced redundancy. Thus, repetition codes with a higher code rate in Fig. 4.1(b) achieve a better BER performance than repetition codes with a lower code rate. The limitation due to the normalization to the information rate is so strong that none of the repetition codes offers a gain over the uncoded transmission in the low BER regions.

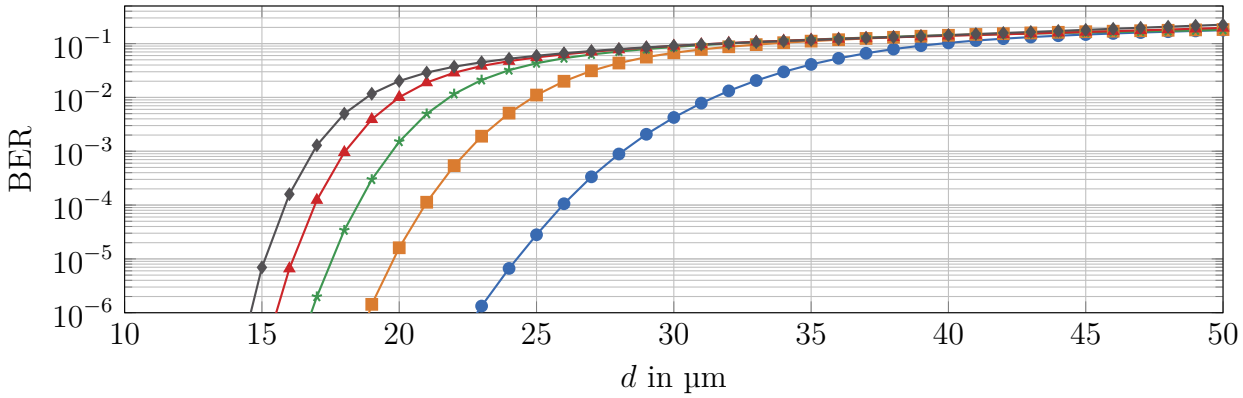
If, in addition to the information rate, the energy in terms of the average released molecules per info bit is normalized as well, the achieved performance of coded transmission becomes even worse. The more redundancy a code has, the fewer molecules per code bit are available on average. Fewer molecules per code bit are equivalent to a lower signal strength. If fewer molecules are released, fewer molecules reach the receiver. Thus the noise power is increased relative to the signal by the characteristic of the amplitude-dependent noise.



(a) Unnormalized.



(b) T normalized.



(c) T & N normalized.

Figure 4.1: Bit error rate performance of repetition encoded transmissions as a function of transmission distance. Parameters according to Appendix C.10.

Consequently, in Fig. 4.1(c), repetition codes with lower code rates show a greater loss compared to Fig. 4.1(b) than repetition codes with higher code rates. At a target BER of 10^{-3}, none of the investigated codes shows an improvement of the maximum transmission range compared to the uncoded case.

From the results in Fig. 4.1, it can be concluded that particularly low-rate repetition codes lead to a significant improvement in transmission with respect to an unnormalized assumption. From a normalized viewpoint, repetition codes should not be used, since, at least for the scenario under investigation, they achieve no gain over the uncoded transmission.

4.1.2 Cyclic Redundancy Check Code

In 1961, W. W. Peterson invented the family of cyclic redundancy check (CRC) codes [PB61], which, as the name suggests, are systematic cyclic codes in which the data is extended by redundant information. The redundant information serves as error check and thus for error detection. CRC codes are particularly well suited for the detection of burst errors [Wic95] and are therefore used in many data transmission standards such as Bluetooth, Ethernet, or universal serial bus. Furthermore, they can also be used for error correction as discussed in this dissertation.

For the construction of an $(n_\mathrm{c}, k_\mathrm{c})$ CRC code word, the info word is extended by an $n_\mathrm{c} - k_\mathrm{c}$ bit long check value. This results in a rate $R = {}^{k_\mathrm{c}}/_{n_\mathrm{c}}$ code. The check value is determined by the remainder of the polynomial division of the info word in polynomial representation and the generator polynomial of order $n_\mathrm{c} - k_\mathrm{c}$. For error detection the received code word is divided again by the generator polynomial. If this polynomial division leads to a remainder, the received code word is erroneous [BG+92].

The choice of the generator polynomial is essential for the error detection performance of CRC codes. Powerful generator polynomials have the form

$$g(z) = (z+1)\alpha_\mathrm{p}(z), \tag{4.6}$$

where $\alpha_\mathrm{p}(z)$ is a primitive polynomial. The factor $(z+1)$ ensures that all odd numbers of bit errors are detectable. The second factor $\alpha_\mathrm{p}(z)$ guarantees that all double errors are detectable if $n_\mathrm{c} < 2^{n_\mathrm{c}-k_\mathrm{c}-1}$ applies [BG+92]. Table 4.1 shows the generator polynomials used in this dissertation [Wic95; DVBS2].

So far, CRC codes are used in molecular communication in [BL+15; BL+16b; BL+16a; Bai16] solely for error detection. Apart from error detection, they provide the possibility of error correction, which will be exploited in this dissertation. Therefore, soft-input decoding

Table 4.1: Generator polynomials of CRC codes.

$n_\mathrm{c} - k_\mathrm{c}$	$g(z)$	Reference
4	$z^4 + z^3 + z^2 + z + 1$	[Wic95]
8	$z^8 + z^7 + z^6 + z^4 + z^2 + 1$	[DVBS2]
32	$z^{32} + z^{30} + z^{22} + z^{15} + z^{12} + z^{11} + z^7 + z^6 + z^5 + z$	[Wic95]

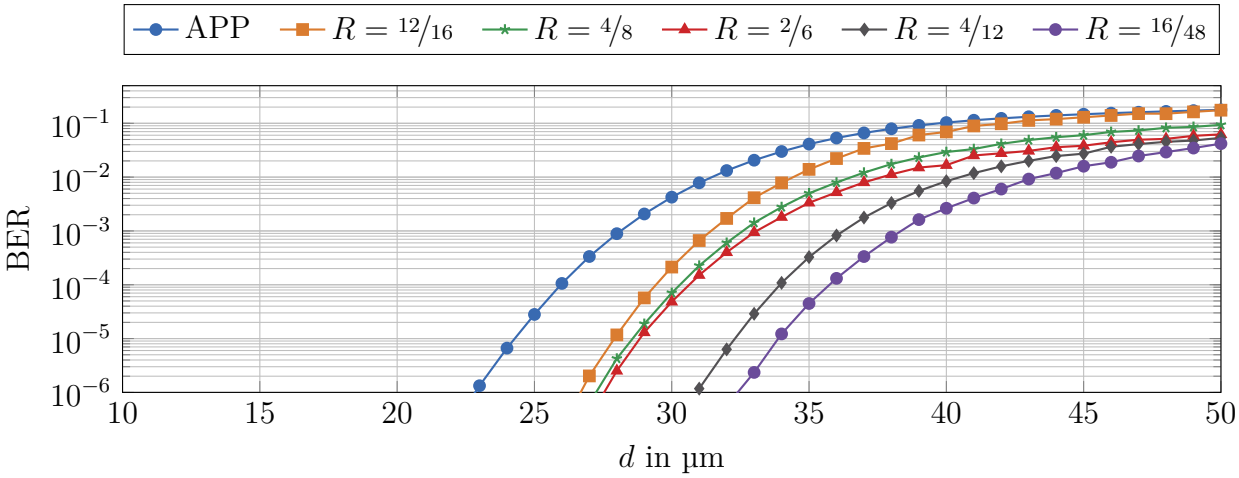
[DW91] is used. When an error is detected, the received word is compared with all possible valid code words and a distance is calculated. As a distance measure between the received word and a code word hypothesis, the absolute LLR values of the mismatching bit positions are accumulated. For the purpose of soft decoding, the most probable code word hypothesis, i. e. the code word hypothesis with the minimum distance to the received word, is selected. The decoded info word can be obtained directly from the systematic part of the corrected code word. The computational complexity for this exhaustive search method depends on the info word length and is $\mathcal{O}\left(2^{k_\mathrm{c}}\right)$. Consequently, soft-input decoding is not practical, especially for long info words. In such cases, it is possible to limit the search to the $n_\mathrm{red} < k_\mathrm{c}$ most unreliable info bits, which reduces the computational complexity to $\mathcal{O}(2^{n_\mathrm{red}})$ [DW91].

Numerical Results

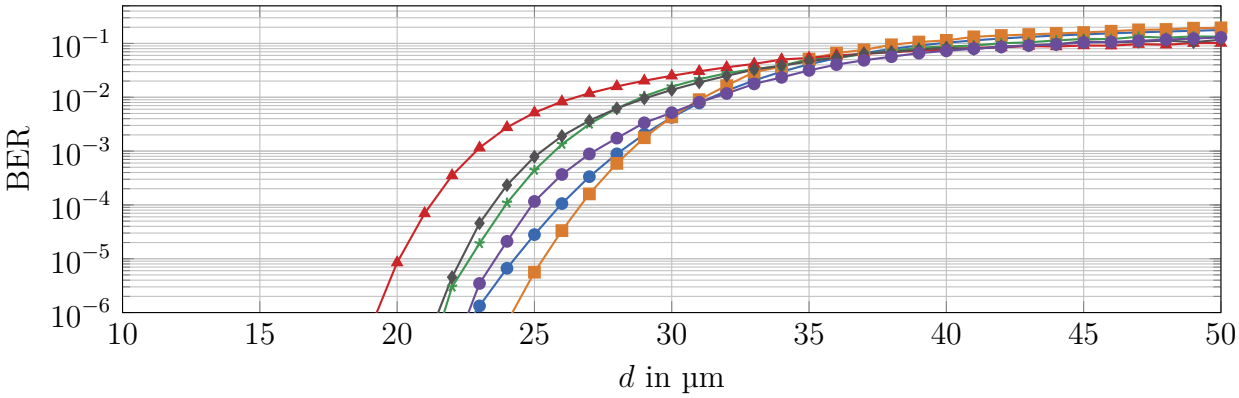
Fig. 4.2 shows the comparison of CRC encoded transmissions to uncoded transmission. The criterion for comparison is the achieved BER at different transmission distances. In order to keep the decoding effort of the CRC codes manageable, especially for those with long code words, $n_\mathrm{red} = 12$ is selected.

In the unnormalized case, shown in Fig. 4.2(a), all CRC coded transmissions lead to an improvement compared to uncoded transmission. In general, lower-rate codes add more redundancy to the information. Consequently, the lower-rate CRC codes in Fig. 4.2(a) show better BER performance than CRC codes with higher rate. Comparing CRC codes with equal rates and different code word lengths, it is observed that longer code word length are advantageous. This is due to the fact that they are capable of correcting more errors within a code word. Among the investigated CRC codes, the $R = {}^{16}\!/_{48}$ CRC code is the most powerful. For a target BER of 10^{-3}, the maximum transmission distance can be increased from 28 µm to 38 µm. It should be noted that the decoder is limited by $n_\mathrm{red} = 12$. Without this limitation the maximum range can be increased further to 41 µm.

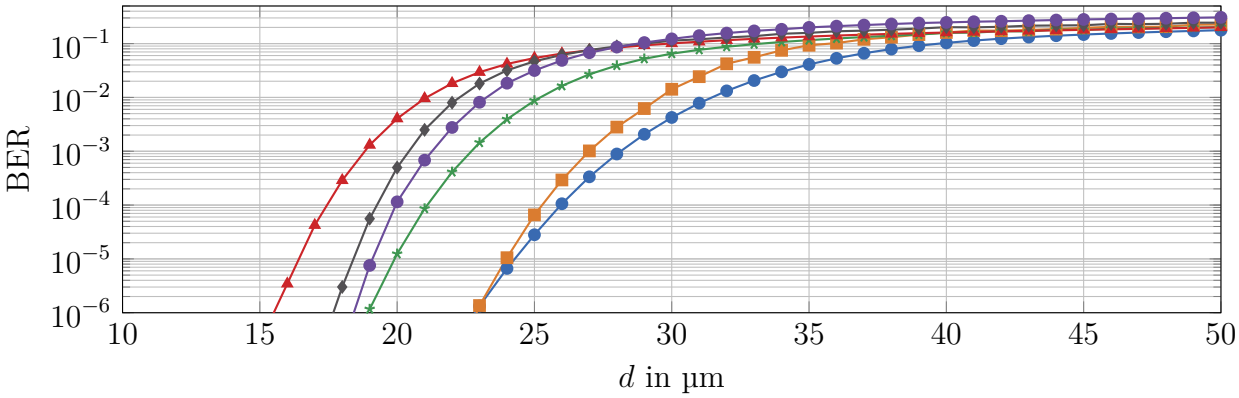
Fig. 4.2(b) shows the scenario for which the symbol duration is normalized to equal information rates for coded and uncoded transmission. Consequently, the symbol duration in a coded transmission is shorter than in an uncoded transmission. This leads to a



(a) Unnormalized.



(b) T normalized.



(c) T & N normalized.

Figure 4.2: Bit error rate performance of CRC encoded transmissions as a function of transmission distance. Parameters according to Appendix C.10.

degradation of the BER performance in coded transmission compared to the unnormalized transmission. Low-rate codes suffer more from this normalization, because their symbol duration is reduced to a greater extent. As a result, the highest rate CRC code with $R = {}^{12}/_{16}$ is found to be the most powerful among the investigated codes. At the target BER of 10^{-3}, the maximum transmission distance can be increased only slightly.

The impact on the BER performance by additionally normalizing the average number of released molecules per info bit is shown in Fig. 4.2(c). A higher degree of redundancy is correspondingly affected by a lower average number of molecules per code bit. Consequently, the low-rate CRC codes show a bigger performance loss compared to Fig. 4.2(b). The limitation due to the normalizations is so large that the high-rate $R = {}^{12}/_{16}$ CRC code provides the best BER performance. Nevertheless, none of the codes shows a gain compared to the uncoded transmission.

The repetition codes discussed in Section 4.1.1 are much less complex than the CRC codes. However, this is also reflected in their performance. If CRC codes are compared with repetition codes with similar rates, especially the CRC codes with longer code word length show a better BER performance.

4.1.3 Single Parity Check Code

One of the simplest block codes considered in diffusion-based molecular communication is the single parity check (SPC) code [MP+18]. It can also be interpreted as an $(n_\mathrm{c} = k_\mathrm{c}+1, k_\mathrm{c})$ CRC code and has a code rate of $R = {}^{k_\mathrm{c}}/_{(k_\mathrm{c}+1)}$. The SPC code is a systematic code in which the code word is formed from the info word and a single parity bit. Parity bits are an alternative to the repetition of info bits to provide redundancy. The parity bit is selected in such a way that the resulting code word has an even or an odd parity. In this dissertation, focus is on SPC with even parity, which results in a linear code. The even parity code word must satisfy the parity equation

$$x_0 \oplus x_1 \oplus \ldots \oplus x_{n_\mathrm{c}-1} = 0, \tag{4.7}$$

where $\oplus$ denotes a modulo two addition. Under the assumption that the parity bit is at the last position of the code word, the resulting generator matrix and parity check matrix are

$$\boldsymbol{G} = \begin{bmatrix} 1 & 0 & \cdots & 0 & 1 \\ 0 & 1 & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & 1 \\ 0 & \cdots & 0 & 1 & 1 \end{bmatrix}, \qquad \boldsymbol{P} = \begin{bmatrix} 1 & \cdots & 1 \end{bmatrix}. \tag{4.8}$$

Tab. 4.2 shows as an example the info word to code word mapping of a $(4, 3)$ SPC code.

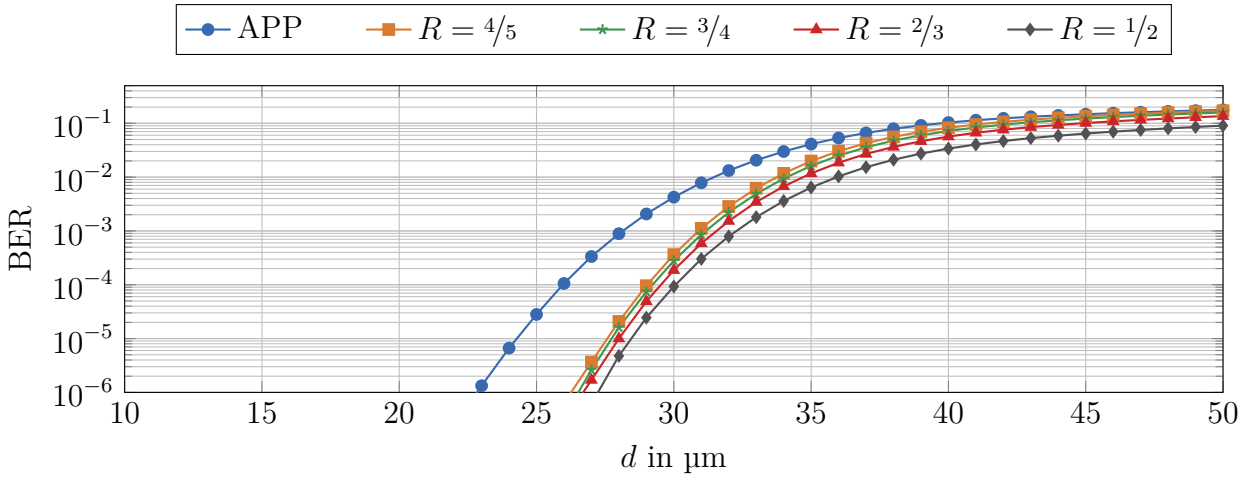
Table 4.2: Mapping table of $(4, 3)$ SPC code.

$\boldsymbol{u}_\mathrm{c} = [u_0\, u_1\, u_2]$	$\boldsymbol{x}_\mathrm{c} = [x_0\, x_1\, x_2\, x_3] = [u_0\, u_1\, u_2\, p_0]$	$\boldsymbol{u}_\mathrm{c} = [u_0\, u_1\, u_2]$	$\boldsymbol{x}_\mathrm{c} = [x_0\, x_1\, x_2\, x_3]$
[0 0 0]	[0 0 0 0]	[1 0 0]	[1 0 0 1]
[0 0 1]	[0 0 1 1]	[1 0 1]	[1 0 1 0]
[0 1 0]	[0 1 0 1]	[1 1 0]	[1 1 0 0]
[0 1 1]	[0 1 1 0]	[1 1 1]	[1 1 1 1]

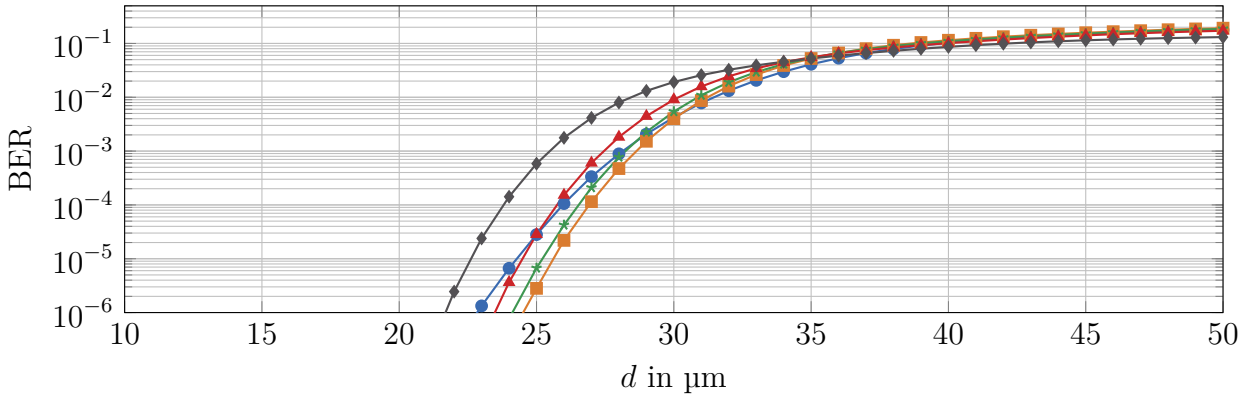
With a minimum distance of $d_\mathrm{min} = 2$, the SPC code can detect a single bit error. However, since the parity equation is violated for each odd number of bit errors, all odd numbers of bit errors can be detected, making the SPC code a powerful tool for error detection. With hard-input decoding, the SPC code cannot correct any bit errors. To provide error correction of a single bit error, soft-input decoding is used in this dissertation. In an erroneous code word, the code bit with the most unreliable LLR (smallest magnitude) is inverted. In fixed threshold detector (FTD), for example, the distance to the threshold value could be used as soft information.

Numerical Results

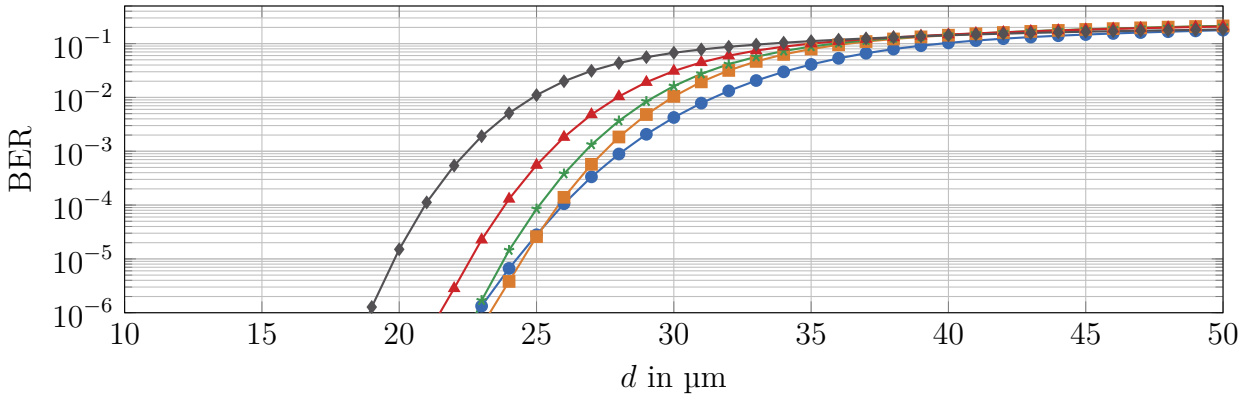
Fig. 4.3 shows the BER performances of SPC codes with different code rates R at varying transmission distances d for different normalizations. Fig. 4.3(a) presents the unnormalized case. As with the repetition and CRC codes, in this case, all coded transmissions achieve a coding gain compared to the uncoded transmission. Furthermore, a code with a lower code rate is also advantageous here. Due to the higher amount of redundancy, more bit errors can be corrected, which is reflected in a better BER performance. For the code rates under investigation in Fig. 4.3(a), the $R = 1/2$ SPC code offers the biggest coding gain. For a target BER of 10^{-3}, the maximum transmission distance can be increased from 28 µm to 32 µm. It should be noted that the $R = 1/2$ SPC code is a special case corresponding to a $R = 1/2$ repetition code, which is discussed in Section 4.1.1.



(a) Unnormalized.



(b) T normalized.



(c) T & N normalized.

Figure 4.3: Bit error rate performance of SPC encoded transmissions as a function of transmission distance. Parameters according to Appendix C.10.

In Fig. 4.3(b), the symbol duration is normalized to the information rate. The same effect as already observed for repetition and CRC codes occurs. The negative impact of shorter symbol durations dominates the coding gain. Consequently, the highest-rate $R = 4/5$ SPC code provides the best BER performance. Compared to uncoded transmission, only gains for code rates $R \geq 2/3$ are achieved. With a target BER of 10^{-3}, the maximum transmission distance can be increased by approximately 1 µm with a $R = 4/5$ SPC code. Even if a further improvement is to be expected for higher-rate SPC codes, the achieved coding gain is negligibly small.

If, in addition, the average number of released molecules per info bit is normalized, the results deteriorate further. The lower the code rate of a SPC code, the fewer molecules are released on average per code bit. Fig. 4.3(c) shows that the gap between the SPC codes and the uncoded transmission has increased further compared to the results in Fig. 4.3(b). At a target BER of 10^{-3}, none of the investigated codes shows an improvement of the maximum transmission range compared to the uncoded case. Higher-rate SPC codes may show a coding gain which would, however, not be remarkably large.

If the results of CRC codes and SPC codes with similar rates are compared, it is observed that CRC codes are more efficient. This is consistent with the observation of CRC codes that codes with longer code word lengths are more efficient.

4.1.4 Hamming Code

The Hamming code is a linear and systematic block code developed by Richard Hamming [Ham50], which has a fixed minimum distance of $d_{\min} = 3$. Since each code word fulfills this minimum distance to exactly one other code word, it is also a perfect code. With hard input decoding, a Hamming code can detect up to two bit errors and correct a single bit error. In contrast to SPC codes, more than one parity bit is used within a code word. In contrast to repetition codes, the code word consists of more than a single info bit. The only exception is the $(3, 1)$ Hamming code, which corresponds to a $(3, 1)$ repetition code. A Hamming code has a fixed relationship between the number of parity bits, info word length, and code word length to construct a perfect block code with $d_{\min} = 3$. For $n_c - k_c$ parity bits, the code word length is $n_c = 2^{n_c - k_c} - 1$. The info word length consequently corresponds to $k_c = n_c - \log_2(n_c + 1)$, where n_c, k_c, and $n_c - k_c$ must be natural numbers. Thus only certain (n_c, k_c) Hamming codes exist, with different code rates. Tab. 4.3 shows some of the existing combinations. If the code word length is increased, the code rate increases and converges towards $R \to 1$. The main difference in the structure of Hamming codes to SPC codes is that a parity bit is not formed over the entire info word but only

Table 4.3: Possible combinations of (n_c, k_c) Hamming codes and corresponding code rates.

$n_c - k_c$	2	3	4	5	6	7	8	9	10
n_c	3	7	15	31	63	127	255	511	1023
k_c	1	4	11	26	57	120	247	502	1013
R	$1/3$	$4/7$	$11/15$	$26/31$	$57/63$	$120/127$	$247/255$	$502/511$	$1013/1023$

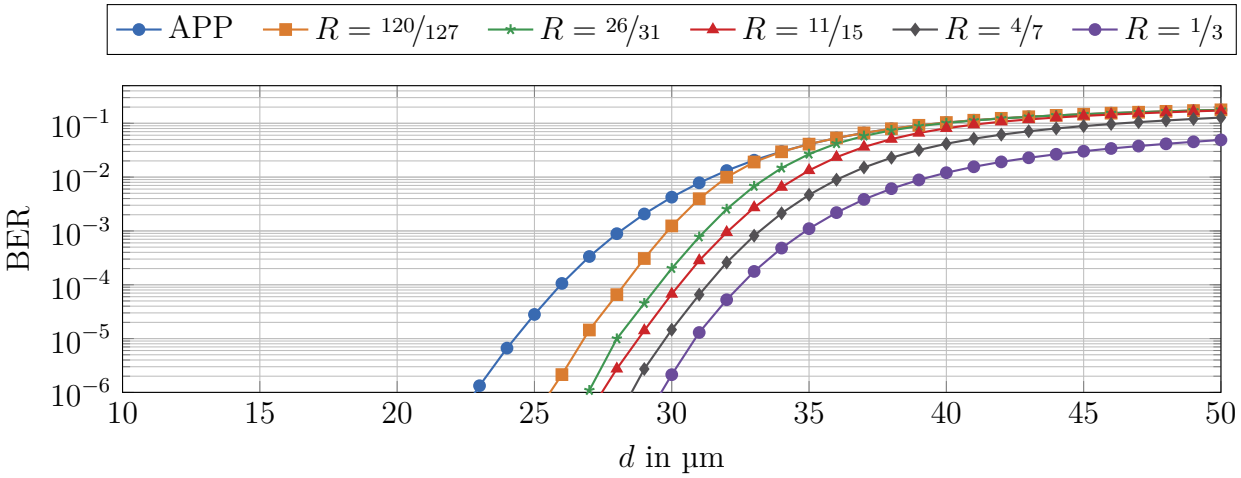
over a part of the info word. Parity bits are chosen to fulfill an even parity equation. The easiest way to describe the Hamming code construction is by the structure of its parity check matrix. The columns consist of all possible bit combinations that are not equal to zero. For a $(7, 4)$ Hamming code, for example, the parity check matrix is

$$\boldsymbol{P} = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}. \tag{4.9}$$

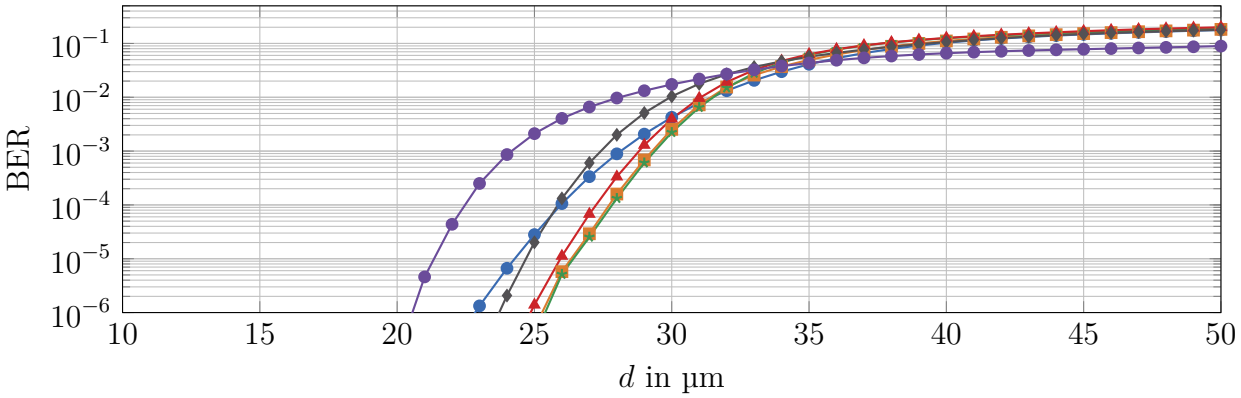
The corresponding generator matrix is

$$\boldsymbol{G} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}. \tag{4.10}$$

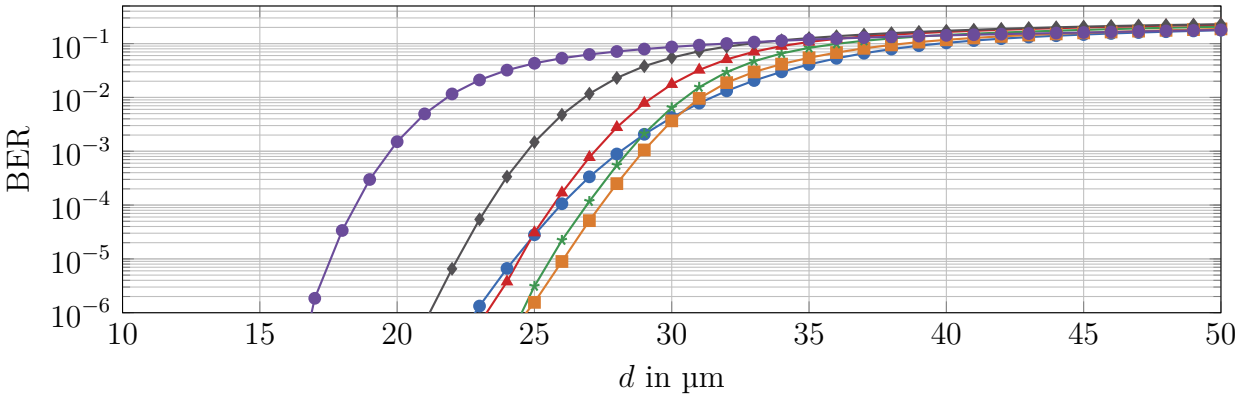
In this example, the three parity bits are placed at positions which are powers of two. This construction leads to an elegant hard-input decoding, because in case of a single bit error the result of the parity check equation (4.2) is the binary representation of the erroneous bit position. By swapping the columns of $\boldsymbol{G}$ and rows of $\boldsymbol{P}$, it is possible to form the Hamming code in such a way that the info word is at the beginning of the code word and the parity bits are at the end of the code word. In this case, the assignment of (4.2) with the erroneous bit positions can be done using a syndrome table [LC01]. Because of their simple structure, Hamming codes using hard-input decoding are widely considered in the diffusion-based molecular communication literature e. g. in [LH12; BL$^+$14; Bai16; LH$^+$15a; LH$^+$15b; DD$^+$17]. In this dissertation, the Hamming code is considered with a soft-input decoding. For this purpose, the suboptimal belief propagation algorithm [Joh10] is used.



(a) Unnormalized.



(b) T normalized.



(c) T & N normalized.

Figure 4.4: Bit error rate performance of Hamming encoded transmissions as a function of transmission distance. Parameters according to Appendix C.10.

Numerical Results

Fig. 4.4 presents the BER performance as a function of the transmission distance for selected Hamming codes using various normalizations. Fig. 4.4(a) shows the unnormalized case. A lower code rate leads to a higher coding gain due to the additional redundancy. The $(3, 1)$ Hamming code with the lowest rate $R = 1/3$ leads to the maximum gain. With a target BER of 10^{-3} the maximum transmission distance can be increased from 28 µm to 35 µm. It should be noted that the $(3, 1)$ Hamming code is a special case, since it is equivalent to a repetition code of rate $R = 1/3$.

Fig. 4.4(b) shows the behavior of the Hamming codes when the symbol duration is normalized to the information rate. Due to the normalization, the general behavior of the codes is reversed as already discussed in the previous sections for repetition, CRC, and SPC codes. High-rate codes are advantageous because they have a longer symbol duration than low-rate codes. This is in line with the characteristic of Hamming codes that they become higher rate with an increasing number of parity bits. Hamming codes with rate $R > 4/7$ still achieve coding gains in Fig. 4.4(b) despite normalization. For the code rates under investigation, the $(127, 120)$ Hamming code improves the maximum transmission distance at the 10^{-3} target BER of approximately 1 µm.

Fig. 4.4(c) shows not only normalization of the symbol duration but also of the number of released molecules per info bit. As described previously, high-rate codes are less affected by this normalization than low-rate codes, because there is less redundancy. This can also be observed in the BER curves in Fig. 4.4(c). While the performance of the $(7, 4)$ Hamming code has further deteriorated compared to the scenario in Fig. 4.4(b), the additional normalization has only a minor effect on the high-rate $(127, 120)$ Hamming code. This still achieves a coding gain of approximately 1 µm compared to uncoded transmission.

In comparison to the results of SPC codes in Fig. 4.3, it is noticeable that Hamming codes are more powerful. Due to the steeper slope in the lower BER region, Hamming codes provide longer transmission distances than SPC codes with comparable code rates. This becomes clear, for example, when the BER curve of the $(15, 11)$ Hamming code is considered. In an unnormalized case the $(15, 11)$ Hamming code outperforms the low rate $(3, 2)$ SPC code. In the normalized cases, it allows a higher maximum range in the lower BER regions than the higher rate $(4, 3)$ SPC code. If the $(15, 11)$ Hamming code is compared to the $(16, 12)$ CRC code at a target BER of 10^{-3}, both show similar performance at a similar rate.

4.1.5 Reed-Solomom Code

In 1960, I. S. Reed and G. Solomon introduced codes which work completely in Galois field algebra [RS60]. This class of linear block codes are known as Reed-Solomon codes and are very efficient against burst and random errors [Ber68; Bla83]. Furthermore they reach the theoretical Singleton bound and are therefore maximum-distance separable [WB99]. Due to their performance, Reed-Solomon codes are used in numerous applications such as deep space communications, magnetic recording systems, or compact discs. The first application in the field of diffusion-based molecular communication is documented in [DD+17].

A special feature of Reed-Solomon codes is that they are non-binary and work on symbol level. This implies that when binary information is transmitted, several bits are combined into a symbol. In this dissertation, systematic Reed-Solomon codes are considered, while in general they can also be constructed as non-systematic codes. An $(n_\mathrm{c}, k_\mathrm{c})$ Reed-Solomon code constructs n_c code symbols from k_c info symbols by adding $n_\mathrm{c}-k_\mathrm{c}$ parity check symbols. This results in a code rate of $R = {}^{k_\mathrm{c}}/_{n_\mathrm{c}}$. Due to its minimum distance of $d_\mathrm{min} = n_\mathrm{c} - k_\mathrm{c} + 1$, it is capable of correcting up to $\lfloor (n_\mathrm{c} - k_\mathrm{c})/2 \rfloor$ symbol errors. The code typically operates in a Galois field of order $n_\mathrm{c}+1$, where $n_\mathrm{c}+1$ is restricted to be a prime power. Only primitive Reed-Solomon codes where $n_\mathrm{c}+1$ is a power of prime number two will be considered in the following. Hence, a symbol consists of $\log_2(n_\mathrm{c}+1)$ bits.

For encoding, the generator polynomial

$$g(z) = \prod_{i=0}^{n_\mathrm{c}-k_\mathrm{c}-1} \left(z - \alpha_\mathrm{p}^{i+1}\right) \tag{4.11}$$

is calculated, where α_p is a primitive element of the Galois field of order $n_\mathrm{c}+1$ [WB99]. From (4.11) it can be seen that the roots of the generator polynomial are defined by $\alpha_\mathrm{p}, \ldots, \alpha_\mathrm{p}^{n_\mathrm{c}-k_\mathrm{c}}$. The code word consists of the k_c info symbols and the $n_\mathrm{c} - k_\mathrm{c}$ parity check symbols. The latter are chosen in such a way that the resulting code word is completely divisible by the generator polynomial.

For decoding, it is first checked whether the received code word is completely divisible by the generator polynomial. If this is the case, a valid code word is received and the info symbols can be extracted directly from the code word. A transmission error is identified when the division results in a remainder. In order to localize the error position, the received code word polynomial is evaluated at the roots of the generator polynomial to obtain syndromes. By applying the Berlekamp-Massey algorithm [WB99], the error locator polynomial can be computed, which generates the sequence of syndromes. The roots of the error locator

polynomial provide information about the error positions in the code word and can be determined by a Chien search [WB99]. In a last step the error magnitudes are found by Forney's algorithm [For65], which is based on Lagrange interpolation. By exploiting the error positions and the error magnitudes, an error correction can be performed.

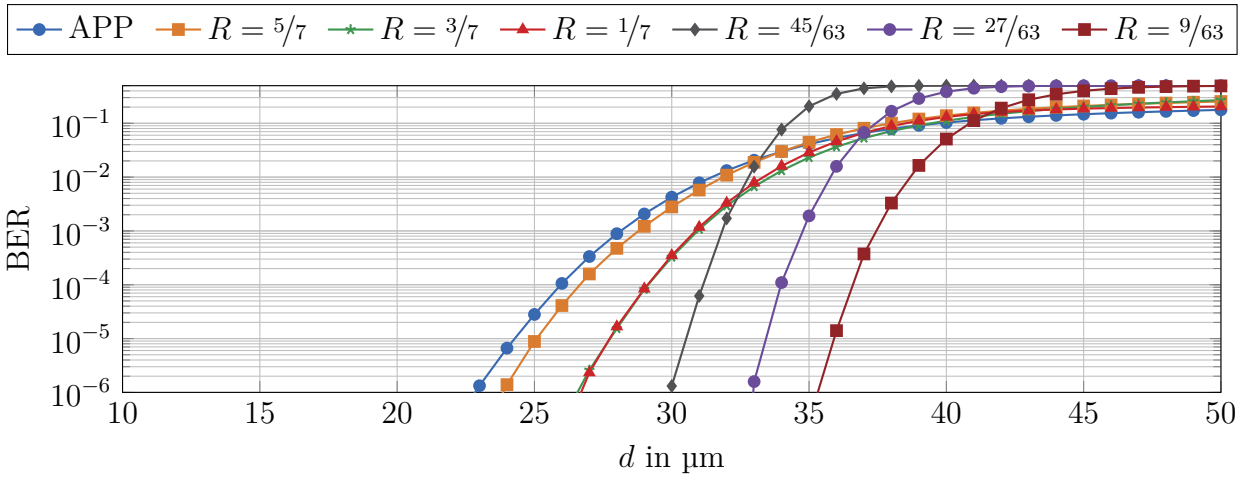
Besides the presented description of the Reed-Solomon code based on polynomials, there are also other approaches in the literature. Among others, the Reed-Solomon code can be described by the finite field Fourier transform or by means of a generator and parity check matrix. For a more detailed description of the presented encoding and decoding steps, as well as other approaches, the reader is referred to [WB99].

Numerical Results

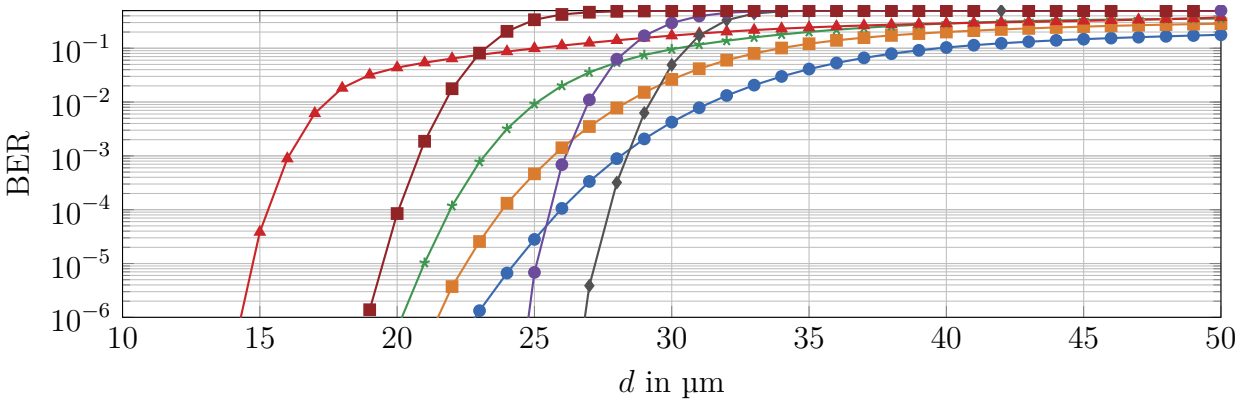
Fig. 4.5 shows the BER performance of a Reed-Solomon encoded transmission system. In the unnormalized case, Fig. 4.5(a), all codes under investigation show a gain over an uncoded transmission. There are two observations that can be made. Firstly, Reed-Solomon codes with longer code word length are more efficient than codes with equal rate and shorter code word length. If the code word length is increased, more bits are combined into one symbol and thus, more bit errors can be corrected within a single symbol error. Secondly, lower-rate codes are more efficient than higher-rate codes with identical code word length, since they contain more redundancy. Among the investigated code rates, the $R = {}^{9}/_{63}$ Reed-Solomon code achieves the largest increase in transmission distance. With a target BER of 10^{-3}, the distance can be improved from 28 µm to 37 µm.

In Fig. 4.5(b), the uncoded and coded transmission is normalized to identical information rates. As a result, lower-rate codes are affected by a larger performance loss, since their symbol duration is reduced more significantly. This effect is so strong that higher-rate codes offer better performance than lower-rate codes. Among the investigated codes, the $R = {}^{45}/_{63}$ Reed-Solomon code achieves the best performance. However, the distance increase compared to uncoded transmission at a target BER of 10^{-3} is negligible. From the obtained results it is to be expected that higher rate Reed-Solomon codes and/or Reed-Solomon codes with longer code word length will lead to a larger increase in transmission distance.

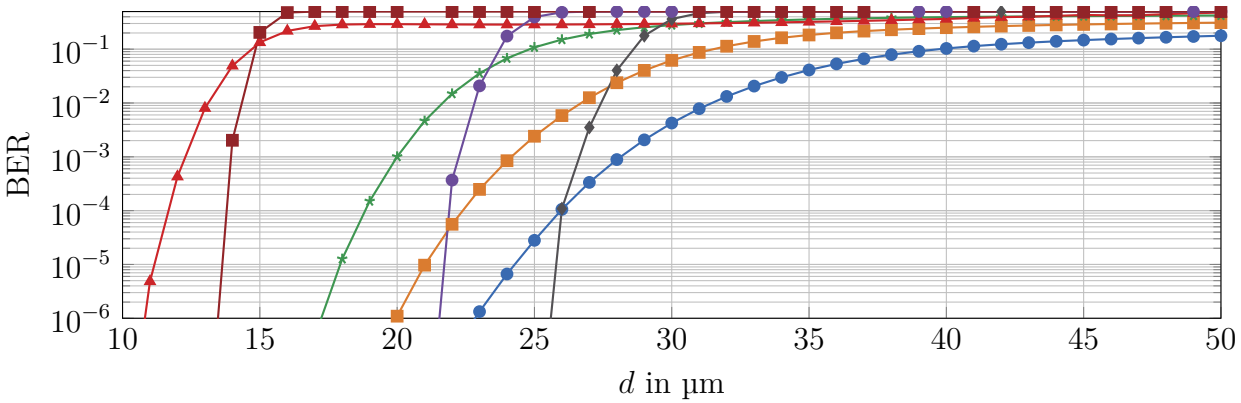
Fig. 4.5(c) shows the case in which, in addition to the information rate, the average number of molecules released per info bit in the uncoded case and coded case is kept constant. Due to the higher redundancy, fewer molecules per code bit are emitted as the code rate decreases. Accordingly, there is a degradation of all coded transmission compared to Fig. 4.5(b). At a target BER of 10^{-3}, even the most powerful Reed-Solomon code with $R = {}^{45}/_{63}$ no longer leads to an improvement of the maximum transmission distance.



(a) Unnormalized.



(b) T normalized.



(c) T & N normalized.

Figure 4.5: Bit error rate performance of Reed-Solomon encoded transmissions as a function of transmission distance. Parameters according to Appendix C.10.

In comparison to CRC and SPC codes considered in this dissertation, Reed-Solomon codes prove to be more powerful, at least for longer code word lengths. In contrast to repetition codes and Hamming codes with similar code rates, however, Reed-Solomon codes perform worse. A possible reason for the difference may be in the decoding process. While repetition codes and Hamming codes are soft-input decoded, Reed-Solomon codes are used with hard-input decoding.

4.1.6 Low-density Parity-check Code

LDPC codes are linear block codes originally invented by Robert Gray Gallager [Gal62], which are very powerful codes and can operate close to the Shannon capacity [RS+01]. The special property of them is the sparse parity check matrix, which describes the code completely. A distinction is generally made between regular and irregular LDPC codes. A regular LDPC code fulfills the condition that the number of ones in each column and the number of ones in each row of the parity check matrix are identical. Typically, irregular LDPC codes have a higher performance than regular LDPC codes [RS+01]. As exemplary

$$\boldsymbol{P} = \begin{bmatrix} 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \end{bmatrix}$$

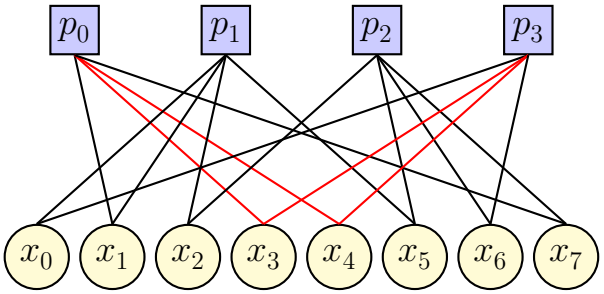


Figure 4.6: Parity check matrix and Tanner graph representation of a regular $(8, 4)$ LDPC code.

shown in Fig. 4.6, the parity check matrix can be visualized by a bipartite Tanner graph consisting of variable nodes x_i and check nodes p_i. A "1" in the parity check matrix symbolizes a connection between a variable node and a check node. The check nodes can be interpreted as individual SPC codes which represent the parity checks. The strong potential of the LDPC codes lies in their iterative soft-input soft-output decoding. According to the turbo principle, extrinsic information can be exchanged between variable nodes and check nodes using message-passing algorithms in order to increase decoding performance, which will be covered in Chapter 5. In this dissertation, the suboptimal graph-based belief-propagation algorithm [MM+98] is used for soft-input decoding. In order to guarantee a

long exchange of independent messages, the length of circles in the parity check matrix should be kept as large as possible. A girth is defined as the circle with the minimum length. It limits the code performance considerably [Joh10]. In the example in Fig. 4.6, there is a girth of length four, which is highlighted in red in the parity check matrix and the Tanner graph.

In diffusion-based molecular communication, LDPC codes are used for the first time in [LH+15a] together with a one step majority logic decoding. As in [DS+19], the focus in this dissertation is on the advanced LDPC codes introduced in the second standard of digital video broadcasting over satellite (DVB-S2) [DVBS2]. As shown in Tab. 4.4, LDPC codes

Table 4.4: LDPC parameters from the DVB-S2 standard [DVBS2].

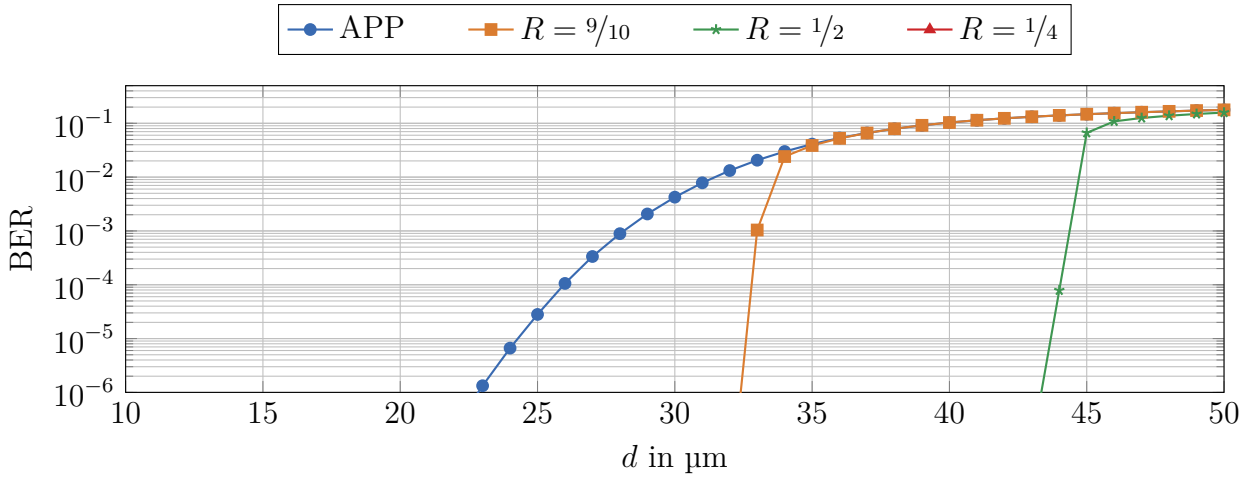
n_c	64800										
R	$1/4$	$1/3$	$2/5$	$1/2$	$3/5$	$2/3$	$3/4$	$4/5$	$5/6$	$8/9$	$9/10$

n_c	16200									
R	$1/5$	$1/3$	$2/5$	$4/9$	$3/5$	$2/3$	$11/15$	$7/9$	$37/45$	$8/9$

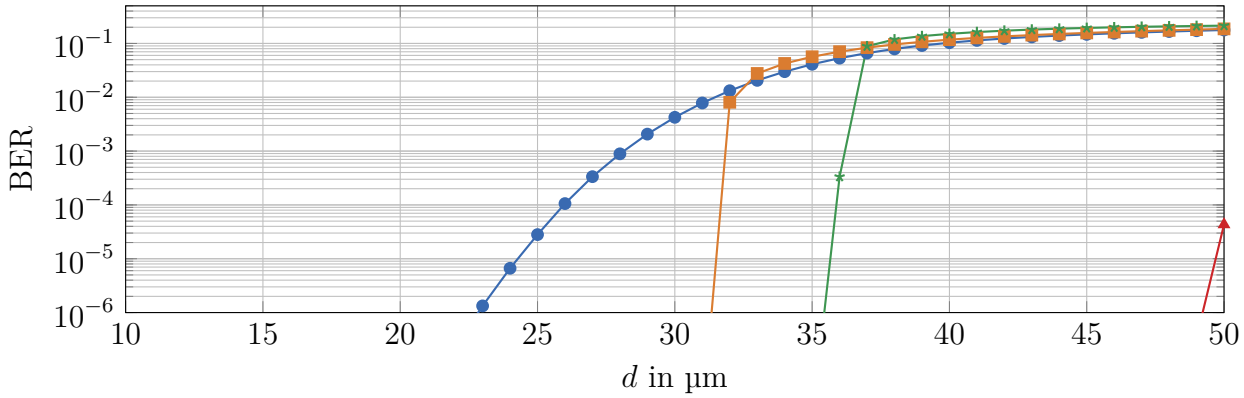
are given with rates from $R = 1/4$ to $R = 9/10$ for moderate block length and from $R = 1/5$ to $R = 8/9$ for short block length. A short block length is defined as $n_c = 16200$ and a moderate block length as $n_c = 64800$. Here, only the LDPC codes with the moderate block length $n_c = 64800$ are considered. For encoding, all parity bits are initialized with zero and determined according to the construction rules and tables given in the DVB-S2 standard. After their determination, the parity bits are accumulated. Since they are systematic LDPC codes, the info bits are concatenated with the parity bits, which results in the final code word.

Numerical Results

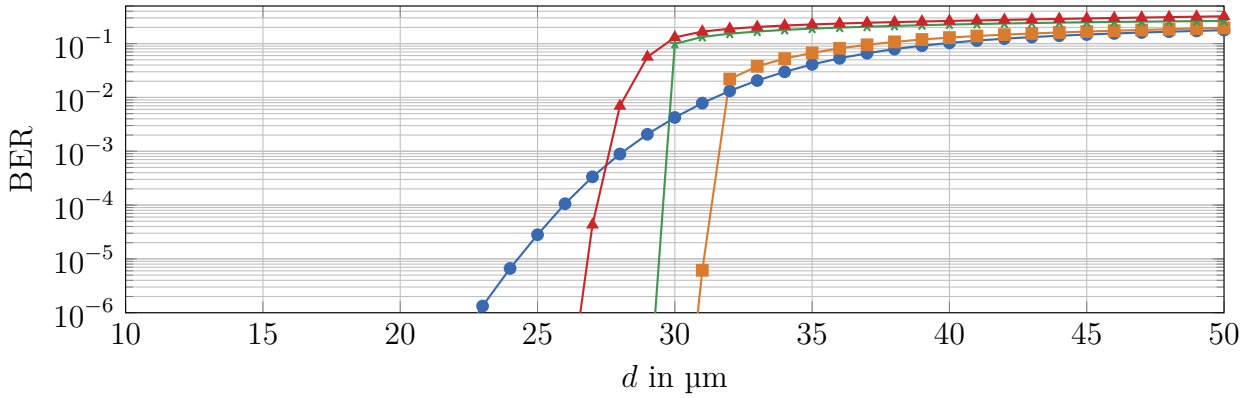
Fig. 4.7 shows the influence of the transmission distance on the BER performance of LDPC coded transmissions under different normalizations. For the analysis, a low-rate $R = 1/4$ LDPC code, a medium-rate $R = 1/2$ LDPC code, and a high-rate $R = 9/10$ LDPC code are considered. Compared to the simpler CRC, SPC, repetition, Hamming, and Reed-Solomon codes from the previous sections, the LDPC codes achieve significantly higher coding gains for comparable rates. In addition, they show a steep slope to the lower BER regions, known as turbo cliff. The reason is the larger code word length of the LDPC codes and their more complex structure, which enables iterative decoding with information exchange between variable nodes and parity check nodes. In the results shown in Fig. 4.7, the maximum



(a) Unnormalized.



(b) T normalized.



(c) T & N normalized.

Figure 4.7: Bit error rate performance of LDPC encoded transmissions as a function of transmission distance. Parameters according to Appendix C.11.

number of internal iterations is limited to 50 iterations.

Fig. 4.7(a) shows the unnormalized case. LDPC codes with lower code rates achieve a larger coding gain due to their increased redundancy. With a target BER of 10^{-3}, the $R = 1/2$ LDPC code already achieves an increase in the maximum transmission distance from 28 µm to 44 µm. The $R = 1/4$ LDPC code further increases the maximum transmission rate to 69 µm which can be observed in Fig. 5.7(a).

In Fig. 4.7(b), the symbol duration is normalized to the information rate. As described previously, this has a negative effect on the BER performance due to shorter symbol durations. The lower the code rate, the larger the negative effect. Thus, the BER performance of the $R = 9/10$ LDPC code changes only slightly compared to the unnormalized analysis. In contrast to CRC, SPC, repetition, Hamming, and Reed-Solomon codes discussed in the previous sections, the general performance behavior of the LDPC codes with respect to the unnormalized case does not reverse. In Fig. 4.7(b), low-rate LDPC codes are still more powerful than high-rate LDPC codes. This shows that in the scenario under consideration the coding gain of LDPC codes by additional redundancy exceeds the limitations of shorter symbol durations. With a target BER of 10^{-3}, the $R = 1/4$ LDPC code achieves a maximum transmission distance of 51 µm, which can be observed in Fig. 5.7(a).

Finally, Fig. 4.7(c) shows in addition to the symbol duration the normalization of the average number of released molecules per info bit. As described in previous sections, this leads to a reduction in signal strength, which has a stronger effect on lower-rate codes. The combination of these two negative effects (reduction of symbol duration and number of molecules) dominates the positive effect (coding gain through additional redundancy). Consequently, higher-rated LDPC codes achieve better BER performances than lower-rated LDPC codes. Despite the normalizations, the $R = 9/10$ LDPC code increases the maximum transmission distance at a target BER of 10^{-3} to 31 µm.

4.1.7 Repeat-Accumulate Code

The repeat-accumulate (RA) codes [DJ+98] have a very simple encoding rule. In spite of their simple structure, they can approach the performance of LDPC codes [RG+04]. In fact, they also have a sparse parity check matrix, so they can also be interpreted as LDPC codes with a special construction rule. As the name suggests, the RA code consists of a repetition code and an accumulator. As shown in Fig. 4.8, it is a serially concatenated turbo-like code structure. The outer component is the $(n_{\mathrm{rc}}, 1)$ repetition code which is serially concatenated to the inner component by an interleaver. The inner component is an accumulator which can be interpreted as a truncated $R = 1$ convolutional code with

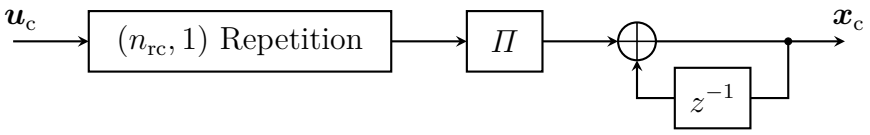


Figure 4.8: Block diagram of a RA encoder.

generator polynomial $g(z) = 1/(1+z)$. Nevertheless, the RA code is typically considered as a linear block code. The total code rate of an $(n_\mathrm{c}, k_\mathrm{c})$ RA code is equal to the rate of the repetition component code $R = {}^{k_\mathrm{c}}/_{(k_\mathrm{c} n_\mathrm{rc})} = {}^{1}/_{n_\mathrm{rc}}$. The RA code in Fig. 4.8 is not systematic. However, by saving a repetition and adding the info word to the resulting parity bits, a systematic RA code can also be generated. In the following, only systematic RA codes are considered.

RA codes can be turbo decoded in the same way as serially concatenated codes [Joh10]. Each component code is decoded separately and the gained extrinsic information is used as *a priori* information of the other component decoder. As already described in Section 4.1.1, the outer repetition decoder sums up the *a priori* LLRs to determine the *a posteriori* LLR of an info bit. The inner accumulator is interpreted as a convolutional code. The extrinsic LLRs can be determined using a Bahl-Cocke-Jelinek-Raviv (BCJR) decoder.

A second decoding approach is to interpret the RA code as LDPC code [Joh10]. The $(n_\mathrm{c} - k_\mathrm{c}) \times n_\mathrm{c}$ parity check matrix $\boldsymbol{P}$ consists of two parts. The first part is an $(n_\mathrm{c} - k_\mathrm{c}) \times k_\mathrm{c}$ matrix that belongs to the repetition code. Its row weight is one and its column weight is equal to the number of repetitions. The second $(n_\mathrm{c} - k_\mathrm{c}) \times (n_\mathrm{c} - k_\mathrm{c})$ submatrix belongs to the accumulator. It is a matrix with a double diagonal. In the case of a $(9, 3)$ RA code, the parity check matrix with and without interleaving, for example, has the following structure:

$$\boldsymbol{P} = \left[\begin{array}{ccc|cccccc} 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \end{array}\right] \overset{\Pi}{\Longrightarrow} \boldsymbol{P} = \left[\begin{array}{ccc|cccccc} 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{array}\right]. \tag{4.12}$$

Interleaving exchanges rows of the first submatrix. The interleaver should be designed in such a way that a girth of length four (highlighted in red in (4.12)) is avoided [Joh10]. Due to the accumulator structure this can be achieved if in the first submatrix two consecutive rows are never assigned to the same info bit. Fig. 4.9 shows the Tanner graph representation

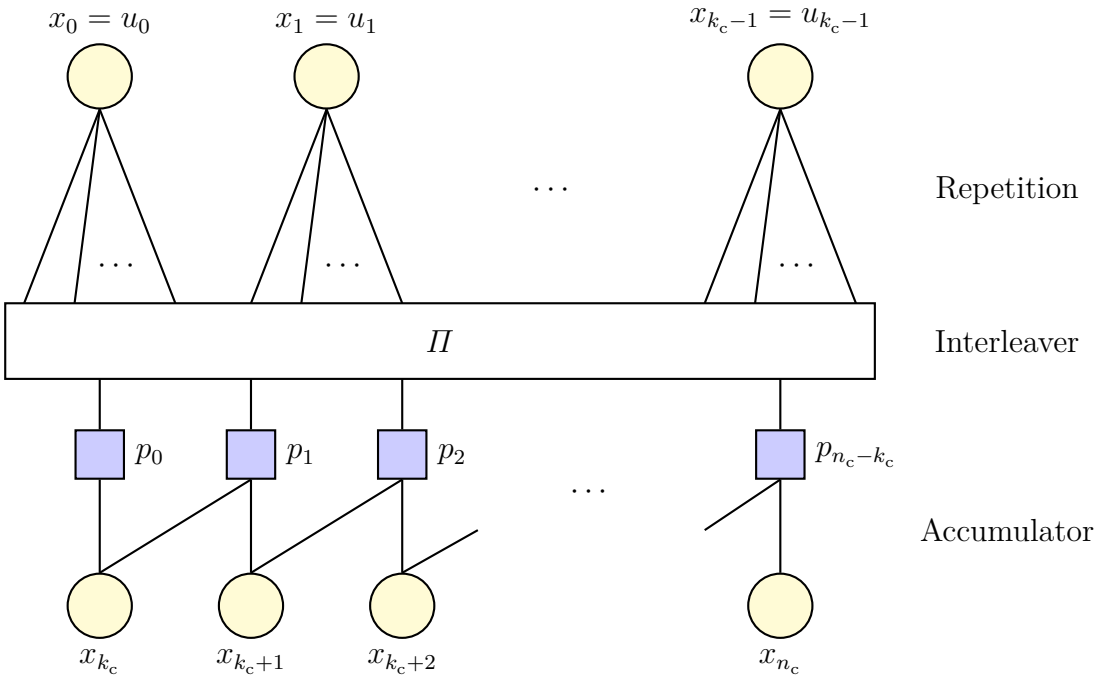


Figure 4.9: Tanner graph representation of an (n_c, k_c) RA code.

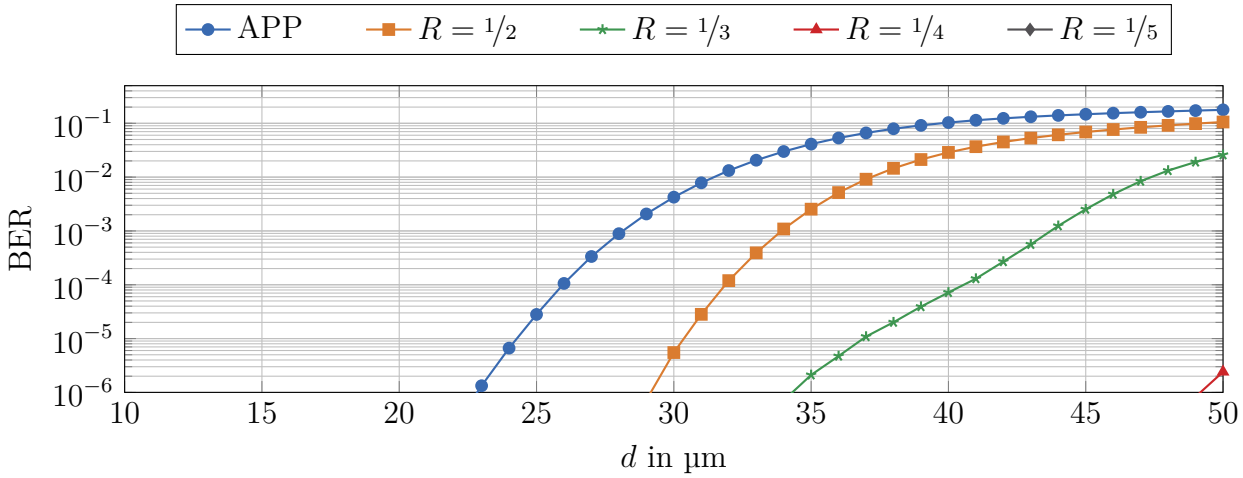
of an (n_c, k_c) RA code parity check matrix. For decoding, message passing algorithms can be used similar to LDPC decoding.

In this dissertation, the RA codes are turbo decoded with the BCJR algorithm. Even if a single iteration has a comparatively higher computational complexity than a message passing decoding, in total less iterations are needed for decoding [Joh10].

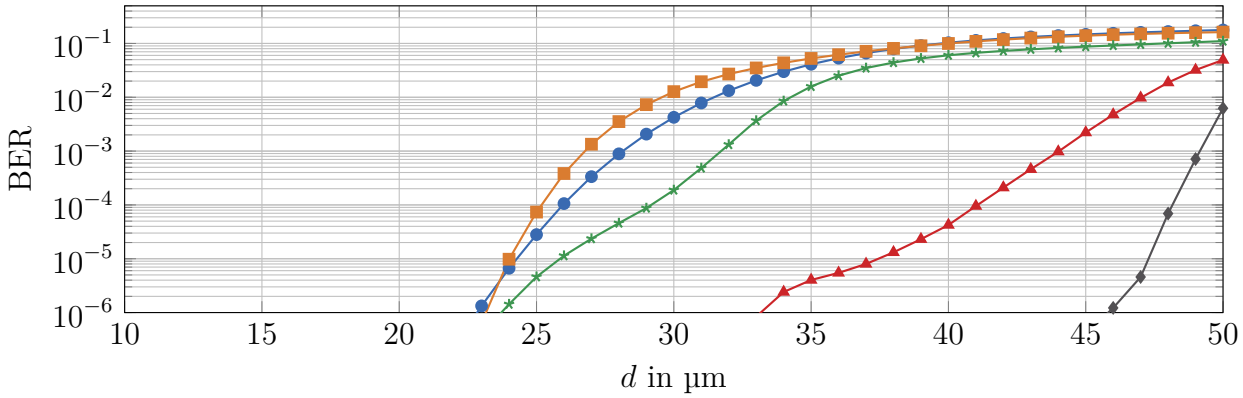
Numerical Results

Fig. 4.10 shows the BER performance of RA coded transmissions over different transmission distances with different normalizations. If compared with repetition and LDPC codes in Fig. 4.1 and Fig. 4.7 for identical code rates, the performance of the RA codes is between the two. Compared to the repetition code, the RA code achieves gains through the accumulator and the internal iterative decoding. The maximum number of internal iterations in the results shown is limited to 100 iterations. In order to reach the performance of the LDPC codes, the code word length should be increased. This increases the interleaver length which has a positive effect on the BER performance. In addition, a larger code word length will lower the error floor observed in Fig. 4.10.

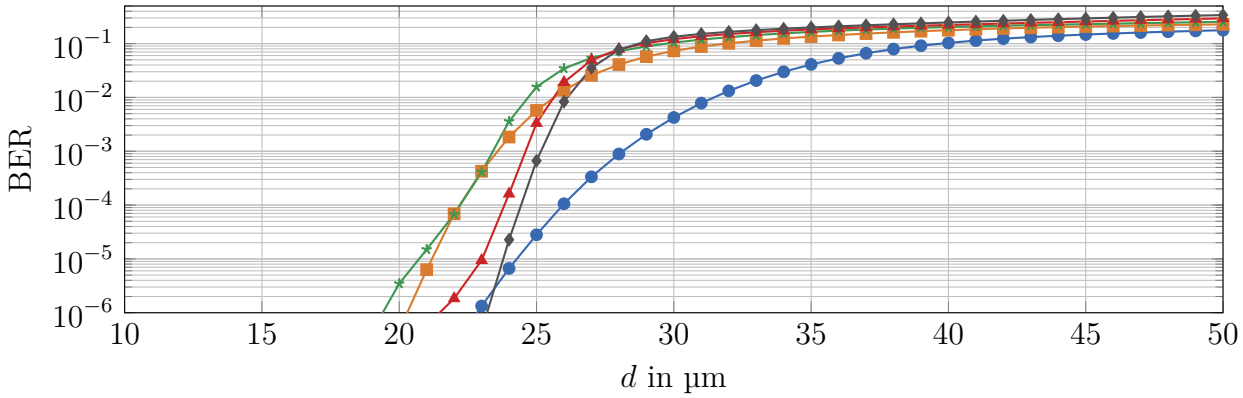
Fig. 4.10(a) deals with the unnormalized case. In the unnormalized case, lower code



(a) Unnormalized.



(b) T normalized.



(c) T & N normalized.

Figure 4.10: Bit error rate performance of RA encoded transmissions as a function of transmission distance. Parameters according to Appendix C.10.

rates result in better BER performance due to more redundancy. For a target BER of 10^{-3}, the maximum transmission distance of 28 µm in the uncoded case can be increased to 44 µm by the $R = 1/3$ RA code. The lower-rate $R = 1/4$ and $R = 1/5$ RA codes increase the maximum transmission distance to 58 µm and 75 µm, respectively, which is outside the depicted distance range.

In Fig. 4.10(b) the symbol rate is normalized to the information rate. As with the LDPC code results in Section 4.1.6, the positive effect of additional redundancy predominates over the negative effect of smaller symbol durations. Thus, RA codes with a low rate have a higher coding gain. For the examined code rates the $R = 1/5$ RA code leads to the largest range increase. With a target BER of 10^{-3}, the maximum transmission distance is increased to 49 µm. Compared to the uncoded transmission, the $R = 1/2$ RA code leads to no improvement in the BER $> 10^{-5}$ region.

In Fig. 4.10(c), in addition to the symbol duration, the average number of released molecules per info bit is normalized as well. As discussed in Section 4.1.3, this leads to a reduction of the signal strength. The lower the code rate, the stronger the reduction. In contrast to the results of the previously considered block codes, low-rate RA codes still show the best performance. Consequently, the performance gain due to additional redundancy is greater than the loss due to normalization of symbol duration and number of molecules. For the target BER of 10^{-3}, none of the investigated RA codes shows a gain over uncoded transmission. Only in the lower BER range, the $R = 1/5$ RA code can achieve improvements.

4.1.8 ISI-free Code

ISI-free codes [SL$^+$12; YC$^+$12; SL$^+$13] are the first channel codes explicitly designed for the ISI mitigation in diffusion-based molecular channels. They were specifically designed for a one-dimensional scenario, where each bit is modulated in the release or absence of a single molecule. In contrast to the three-dimensional scenario, no molecules can get lost, but still may reach the receiver in different order. Such an exchange in the order is called a crossover. In a level τ crossover, a released molecule is received in the τth time slot before or after its desired time slot. An (n_c, k_c, τ) ISI-free code encodes k_c info bits into n_c code bits and is robust against all crossovers with maximum level τ.

In general, crossovers can be divided into two categories. The first category leads to a crossover between two consecutive code words. To counteract these crossovers, each code word ends in τ identical bits and the subsequent code word begins with the same τ identical bits. Consequently, there are exactly two code words for each info word, which

are selected depending on the last code bit of the previous code word. Tab. 4.5 shows, as

Table 4.5: Code word assignment of the $(5, 2, 2)$ ISI-free code [SL+13].

$\boldsymbol{u}_\mathrm{c}$	$\boldsymbol{x}_\mathrm{c}$ beginning with 0	$\boldsymbol{x}_\mathrm{c}$ beginning with 1
[0 0]	[0 0 0 0 0]	[1 1 1 1 1]
[0 1]	[0 0 1 0 0]	[1 1 0 1 1]
[1 0]	[0 0 0 1 1]	[1 1 0 0 0]
[1 1]	[0 0 1 1 1]	[1 1 1 0 0]

an example, the code word assignment for a $(5, 2, 2)$ ISI-free code. The second category contains crossovers within a code word. To be robust against these crossovers, ISI-free codes ensure that a permutation of all possible level τ crossovers still results in a unique code word that is independent of the other code words. For the $(5, 2, 2)$ ISI-free code in Tab. 4.5, for example, any possible permutation does not change the Hamming weight of the code word. This fact can be used for low-complexity decoding using an accumulator and threshold detector:

$$\hat{u}_\mathrm{dec} = \begin{cases} w_\mathrm{H}\left(\boldsymbol{x}_\mathrm{c}\right) & \text{if } w_\mathrm{H}\left(\boldsymbol{x}_\mathrm{c}\right) \leq 2^{k_\mathrm{c}} - 1, \\ n_\mathrm{c} - w_\mathrm{H}\left(\boldsymbol{x}_\mathrm{c}\right) & \text{otherwise,} \end{cases} \tag{4.13}$$

where $\hat{u}_\mathrm{dec}$ is the decimal representation of the estimated info word. Depending on the code construction, a more complex algorithm may be necessary for decoding an $(n_\mathrm{c}, k_\mathrm{c}, \tau)$ ISI-free code [SL+13].

ISI-free codes were further developed into $\left(n_\mathrm{c}, k_\mathrm{c}, \tau, \tau'\right)$ ISI-free codes in [SL+13]. If only a single level τ crossover is guaranteed to be decoded without errors, the number of identical bits at the beginning of the code word can be reduced from τ to τ'. This results in an increase of the code rate. Tab. 4.6 shows, as an example, the code word assignment of a

Table 4.6: Code word assignment of the $(4, 2, 2, 1)$ ISI-free code [SL+13].

$\boldsymbol{u}_\mathrm{c}$	$\boldsymbol{x}_\mathrm{c}$ beginning with 0	$\boldsymbol{x}_\mathrm{c}$ beginning with 1
[0 0]	[0 0 0 0]	[1 1 1 1]
[0 1]	[0 1 0 0]	[1 0 0 0]
[1 0]	[0 0 1 1]	[1 1 0 0]
[1 1]	[0 1 1 1]	[1 0 1 1]

$(4, 2, 2, 1)$ ISI-free code, which can be constructed from the $(5, 2, 2)$ ISI-free code shown in Tab. 4.5. Decoding can be performed analogously to the $(5, 2, 2)$ ISI-free code according to (4.13).

Numerical Results

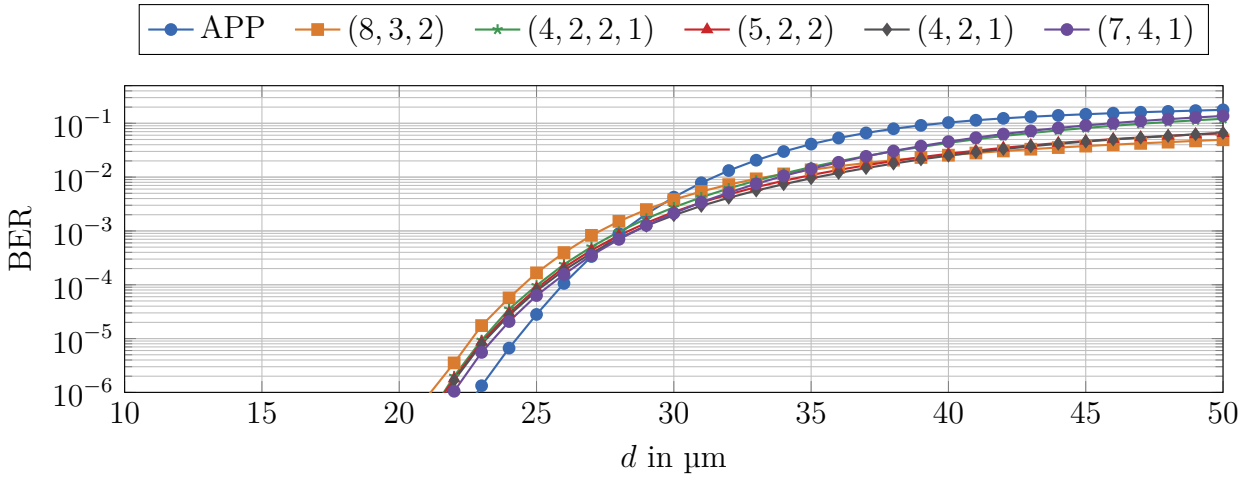
Fig. 4.11 shows BER performances of different ISI-free codes over the transmission distance with different normalizations being considered. In the unnormalized case depicted in Fig. 4.11(a), it can be seen that all considered ISI-free codes show a similar performance. Furthermore, they do not lead to any significant improvement compared to uncoded transmission. In the low BER region, their performance is even worse than the uncoded transmission. The reason for the poor performance of ISI-free codes is that they were originally designed for a one-dimensional scenario. In these scenarios, molecules always reach the receiver and can only be reversed in their order. In the three-dimensional scenario assumed in this dissertation, however, the probability that a molecule will never reach the receiver is larger than zero. In addition, more than a single molecule per bit is released. Besides crossovers, this results especially in the loss of molecules, for which the ISI-free codes are not designed.

The normalizations considered in Fig. 4.11(b) and Fig. 4.11(c) further reduce the BER performance of the ISI-free codes. Low-rate codes are more restricted by normalization than higher-rate codes. As a consequence, the distance between the ISI-free codes with different code rates increases. Also the gap to uncoded transmission increases and the ISI-free code does not lead to any transmission improvement.

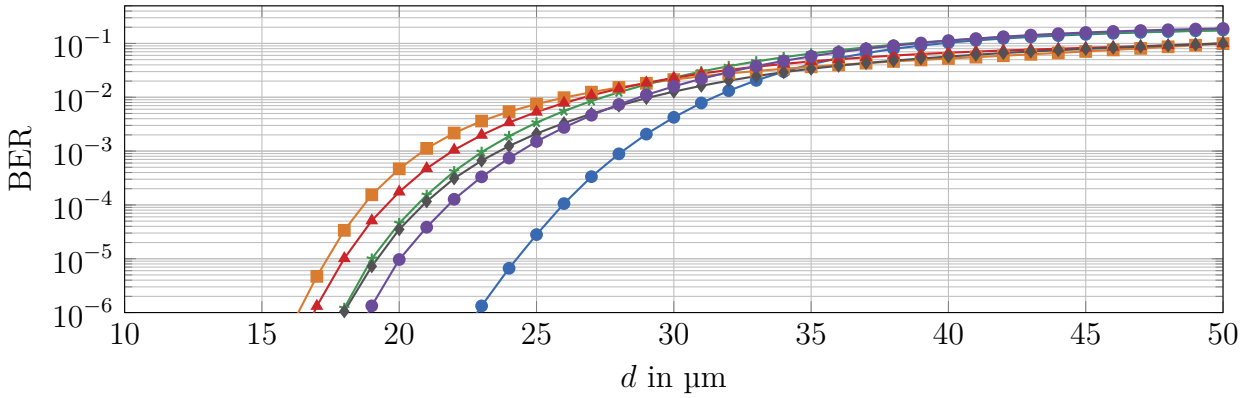
In summary, it can be stated that ISI-free codes are not adapted to the effects of the three-dimensional scenario assumed in this study. Therefore, they do not lead to any coding gain and should not be used in corresponding systems.

4.2 Convolutional Codes

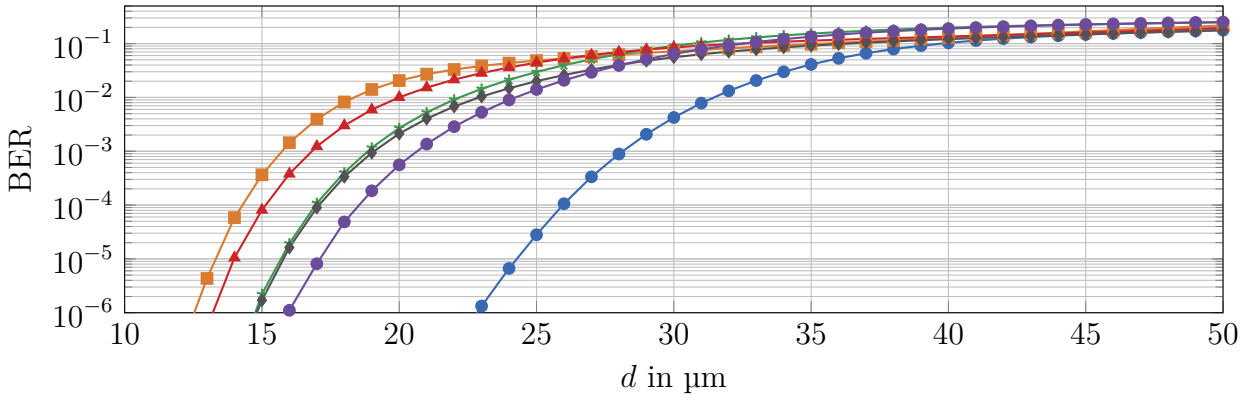
Convolutional codes, e. g. comprehensively studied in [LC01], form the second prominent class of channel codes besides block codes. Just like block codes, convolutional codes add redundancy to be robust against transmission errors. With the help of redundancy, an error correction can be performed. In contrast to block codes, convolutional codes are continuously encoding with theoretically infinite code sequence length. Since in practice, however, information is typically transmitted with a finite length, convolutional codes are rather characterized by a variable code sequence length/block length. Thus, convolutional codes can also be interpreted as special block codes. As shown in Fig 4.12, convolutional codes are typically represented by shift registers and modulo two additions. The shift registers spread the information over several parts of the code sequence and thus create redundancy. This operation is equal to a convolution of the info sequence with the impulse



(a) Unnormalized.



(b) T normalized.



(c) T & N normalized.

Figure 4.11: Bit error rate performance of ISI-free encoded transmissions as a function of transmission distance. Parameters according to Appendix C.10.

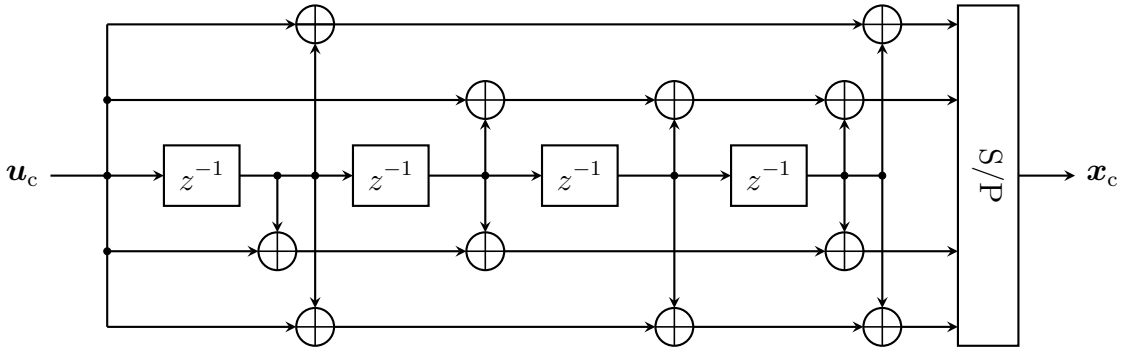


Figure 4.12: Shift-register representation of a $(4, 1, 4)$ convolutional code.

response of the code, hence the name convolutional code. The shift registers introduce a memory to the code words in contrast to block codes. Alternatively, a convolutional code can also be described by its generator polynomials. For the example code in Fig 4.12, the generator polynomials are

$$\begin{aligned} g_0(z) &= 1 + z + z^4, \\ g_1(z) &= 1 + z^2 + z^3 + z^4, \\ g_2(z) &= 1 + z + z^2 + z^4, \\ g_3(z) &= 1 + z + z^3 + z^4. \end{aligned} \tag{4.14}$$

In general, a distinction is made between recursive and non-recursive convolutional codes. The former have an internal feedback path and can thus also be interpreted as infinite impulse response (IIR) filters. Non-recursive convolutional codes, however, as shown in Fig. 4.12, do not have an internal feedback path and can be interpreted as finite impulse response (FIR) filters. In this dissertation, only non-recursive convolutional codes are considered.

An $(n_\mathrm{c}, k_\mathrm{c}, L_\mathrm{c})$ convolutional code encodes in one step k_c info bits to n_c code bits. L_c is the number of shift registers and equal to the memory. Alternatively to the memory length, the constraint length is often specified, which is typically $L_\mathrm{c} + 1$. The resulting code rate is $R = {}^{k_\mathrm{c}}/_{n_\mathrm{c}}$. The example code in Fig. 4.12 is thus a $(4, 1, 4)$ convolutional code with code rate $R = {}^1/_4$. With finite information length, an info sequence of K_c info bits is encoded to a code sequence of N_c code bits. At the beginning of the encoding, the encoder is typically in the zero state, where the state defines the bit sequence in the shift registers. A performance gain can be achieved if the code sequence is terminated. For this purpose, the encoder must end in a well-defined state (usually the zero state). Therefore, the info

sequence is extended by L_c tailbits. Since the tailbits do not contain any information and represent an overhead, the code rate is reduced to $R_\mathrm{tail} = {}^{RK_\mathrm{c}}/_{(K_\mathrm{c}+L_\mathrm{c})}$. With long code sequences, however, this rate loss is negligibly small. As an alternative to tailbits, tailbiting can be used, where the initial state of the encoder is set equal to the final state. This avoids a rate loss, but leads to increased decoding complexity [Höh11].

For decoding, a convolutional code is typically represented by a trellis diagram with its 2^{L_c} states. A state transition is equal to the k_c info bits and the corresponding n_c code bits. Consequently, trellis-based decoders are usually used. In this dissertation, APP decoding is performed using the BCJR algorithm [BC$^+$74], which provides LLRs at its output. The performance of convolutional codes is often estimated by the minimum free distance [LC01], which is equivalent to the minimum distance in block codes. It describes the minimum Hamming distance considering all possible error paths. An error path is a path that deviates from the correct path and remerges again. In general, a larger minimum free distance leads to more powerful convolutional codes.

In contrast to block codes, there are no simple construction rules for designing convolutional codes. Powerful codes are typically determined by an exhaustive search. In addition, convolutional codes have usually a low code rate. In order to generate convolutional codes with high code rates, existing codes can be punctured [Tüc04]. Following a certain pattern, some code bits are suppressed, i. e. not transmitted. This generally leads to a reduction of the free distance. The decoder treats the punctured bits as erasures. Puncturing can also be used to provide unequal error protection. For code rate reduction, generator polynomials can be repeated. By a combination of puncturing and repetitions, code rates can be achieved flexibly. These newly obtained codes usually show a similar performance as optimized convolutional codes with the same code rate [CC$^+$79; YK$^+$84; LP$^+$01]. In addition, the newly designed codes can be decoded using the trellis of the original convolutional code. The original convolutional code is called the mother code.

The first application of convolutional codes in diffusion-based molecular communication is studied in [SL$^+$12; MM$^+$13]. However, a scenario without equalization is considered. As in [DH18], this disseration focuses on the convolutional codes presented in [Tüc04]. From the $(4, 1, 4)$ mother code (equal to the code in Fig 4.12 and (4.14)), 17 subcodes are obtained by puncturing, repeating the generator polynomials and simultaneously maximizing the free distance. The subcodes have code rates from $R = {}^1/_{10}$ to $R = {}^9/_{10}$ in ${}^1/_{20}$ steps.

In Section 5.3 these subcodes are used to design an irregular convolutional code (IRCC). By means of IRCC, the info/code word is partially encoded/decoded by the subcodes of different rates. This allows to achieve arbitrary overall code rates and to shape the extrinsic information transfer (EXIT) function of the decoder, which can be advantageous for iterative detection.

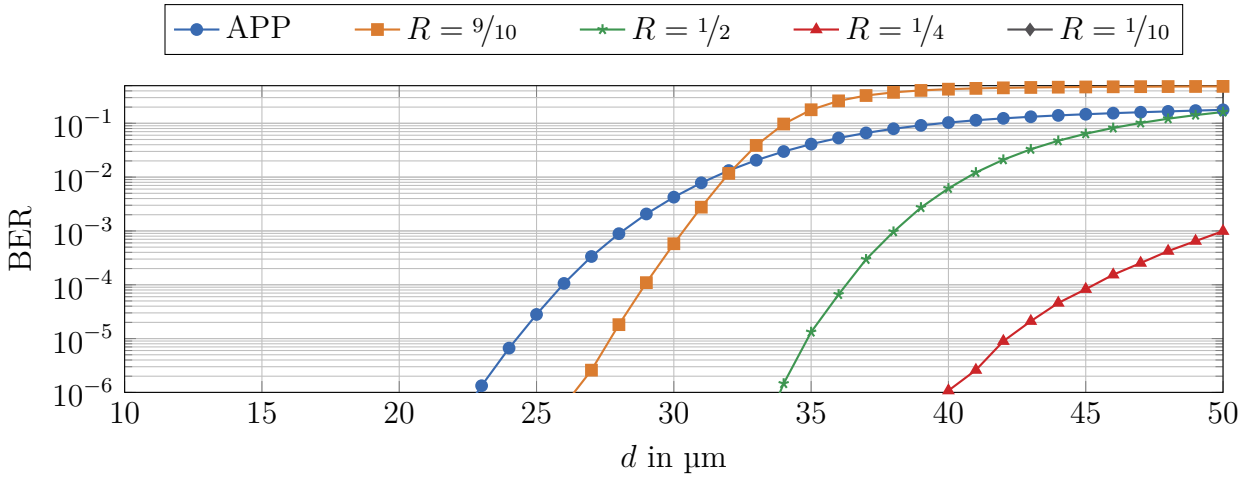
Numerical Results

Fig. 4.13 shows the BER performance of convolutional codes at different code rates over the transmission distance. Different normalizations are considered. Compared to the block codes discussed in Section 4.1, the coding gain achieved by the convolutional codes at comparable code rates is in between. The convolutional code achieves a better performance than the simpler repetition code due to the introduced memory in conjunction with APP decoding. However, the LDPC code outperforms the convolutional code by its internal iterative decoding. The comparison to the RA code, which can also be interpreted as a special kind of convolutional code, is mixed. While the $R = 1/2$ convolutional code performs even better than the $R = 1/2$ RA code, the $R = 1/4$ RA code outperforms the $R = 1/4$ convolutional code. An advantage of the RA code is its internal iterative decoding. A disadvantage is the short memory length.

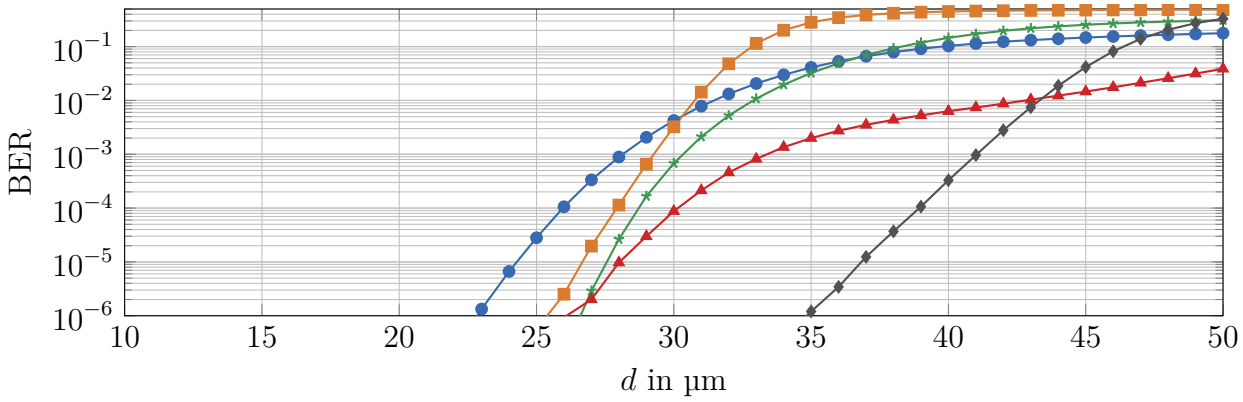
Fig. 4.13(a) shows the unnormalized case. In this case, as with the block codes in Section 4.1, the lower the code rate, the more redundancy the convolutional code offers. More redundancy leads to more powerful codes. Consequently, a convolutional code with lower code rate achieve a larger coding gains. With a target BER of 10^{-3}, the maximum transmission distance can be increased from 28 µm to 50 µm with an $R = 1/4$ convolutional code. For the $R = 1/10$ convolutional code, the maximum transmission distance can be increased to 62 µm, which is depicted in Fig. 5.9.

In Fig. 4.13(b), the symbol duration is normalized to the information rate. Consequently, the symbol duration is reduced more strongly for low-rate codes than for high-rate codes. Nevertheless, the additional redundancy of the coding gain dominates the limitation of the symbol duration. Furthermore, all considered convolutional codes lead to a coding gain compared to uncoded transmission. The maximum coding gain at a target BER of 10^{-3} is achieved by the $R = 1/10$ convolutional code with a maximum transmission distance of 41 µm.

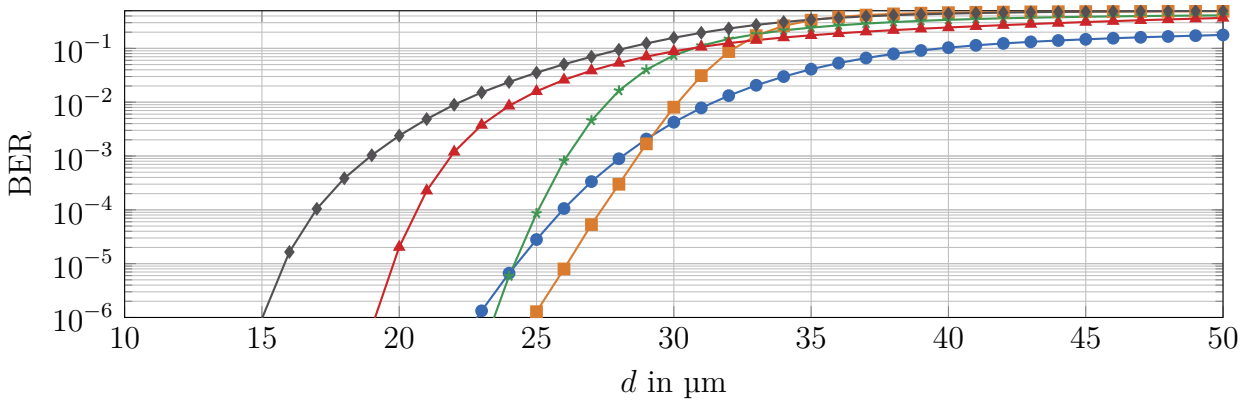
In Fig. 4.13(c), in addition to the symbol duration, the average number of released molecules per info bit is normalized. Here, low-rate codes are more affected by molecule reduction than higher-rate codes as well. If this additional normalization is considered, the



(a) Unnormalized.



(b) T normalized.



(c) T & N normalized.

Figure 4.13: Bit error rate performance of convolutional encoded transmissions as a function of transmission distance. Parameters according to Appendix C.10.

limitations of the normalizations outweigh the positive effect of the gained redundancy. Consequently, high-rate convolutional codes in Fig. 4.13(c) perform better than low-rate convolutional codes. In comparison to uncoded transmission, only high-rate convolutional codes lead to a gain. The $R = 9/10$ code can increase the maximum transmission distance by about 1 µm at a target BER of 10^{-3}. For lower BER, this gain is even larger.

4.3 Line Codes

Line codes [Imm04] are typically used for spectral shaping of a signal before its transmission over a channel. The channel is often a transmission or storage medium. Line codes can pursue different goals. The most prominent goals are the removal of a direct current (DC) bias, the provision of regular level transitions for succesful clock synchronization, and spectral adaptation of the transmission data to the transmission channel. Encoding can be implemented blockwise or continuously.

Although line codes typically do not include error detection, they are interesting for molecular communication. Their ability to adapt the transmitted data to the channel characteristic can be used to reduce transmission errors. The molecular communication channel is affected by amplitude-dependent noise, which, unlike the classical additive white Gaussian noise channel, leads to an unbalanced error probability. Some bit sequences are more prone to errors than other bit sequences. Motivated by this observation and as a replacement for the Hamming distance that is intended for symmetric channels, [KL+12] introduces a new, probability-based distance:

$$d_{\mathrm{p}}\left(\boldsymbol{x}, \hat{\boldsymbol{x}}\right) = -\log\Big(P(\boldsymbol{x} \to \hat{\boldsymbol{x}})\Big), \tag{4.15}$$

where $P(\boldsymbol{x} \to \hat{\boldsymbol{x}})$ is the probability of receiving sequence $\hat{\boldsymbol{x}}$ when $\boldsymbol{x}$ is transmitted. Line codes are used in this dissertation to avoid the most unreliable sequences. It should be mentioned that in a broader sense also modulation methods like pulse-position modulation (PPM) or channel codes like the Hamming code can be interpreted as line codes.

4.3.1 Run-length Limited Code

Run-length limited (RLL) codes [Imm04] are line codes that are mainly used in storage media such as magnetic hard disks or compact discs. They are classically used for reliable clock synchronization and for suppressing high-frequency signals that are often attenuated by the channel. For this purpose, an RLL code limits the minimum and maximum number

of consecutive zeros. An $(n_\mathrm{c}, k_\mathrm{c}, \tau_\mathrm{min}, \tau_\mathrm{max})$ RLL code ensures that at least τ_min zeros and at most τ_max zeros occur between two consecutive ones. It is not necessarily uniquely defined by parameters $(n_\mathrm{c}, k_\mathrm{c}, \tau_\mathrm{min}, \tau_\mathrm{max})$, but by a state diagram or a code book. The block length of the code words does not necessarily have to be constant, but for this dissertation the focus is only on code words with a fixed block length n_c. Tab. 4.7 shows an example of the

Table 4.7: Code table of the $(5, 4, 0, 2)$ RLL code.

$\boldsymbol{u}_\mathrm{c}$	$\boldsymbol{x}_\mathrm{c}$	$\boldsymbol{u}_\mathrm{c}$	$\boldsymbol{x}_\mathrm{c}$	$\boldsymbol{u}_\mathrm{c}$	$\boldsymbol{x}_\mathrm{c}$	$\boldsymbol{u}_\mathrm{c}$	$\boldsymbol{x}_\mathrm{c}$
[0 0 0 0]	[1 1 0 0 1]	[0 1 0 0]	[1 1 1 0 1]	[1 0 0 0]	[1 1 0 1 0]	[1 1 0 0]	[1 1 1 1 0]
[0 0 0 1]	[1 1 0 1 1]	[0 1 0 1]	[1 0 1 0 1]	[1 0 0 1]	[0 1 0 0 1]	[1 1 0 1]	[0 1 1 0 1]
[0 0 1 0]	[1 0 0 1 0]	[0 1 1 0]	[1 0 1 1 0]	[1 0 1 0]	[0 1 0 1 0]	[1 1 1 0]	[0 1 1 1 0]
[0 0 1 1]	[1 0 0 1 1]	[0 1 1 1]	[1 0 1 1 1]	[1 0 1 1]	[0 1 0 1 1]	[1 1 1 1]	[0 1 1 1 1]

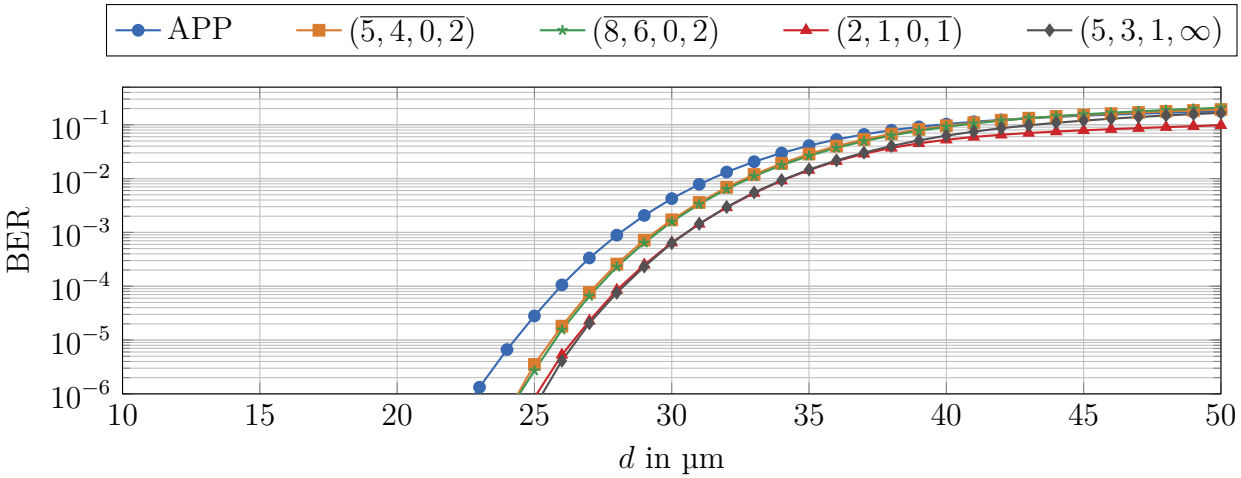
code table of a $(5, 4, 0, 2)$ RLL code.

For molecular communication, RLL codes are interesting from two points of view. Firstly, they can be used to prevent high-frequency signals that are attenuated by the low-pass characteristic of the diffusion-based channel as shown in Fig. 3.30. On the other hand, the transmission of a one via the diffusion-based channel is more error-prone than a zero (see e. g. (3.3)). RLL codes could be used to generate code words in which more zeros than ones are transmitted. If the received sequence does not correspond to a valid code word, soft-input decoding is applied. In this case, decoding is done in favor of the nearest code word, whereby the LLR values serve as a reliability measure of the detected bit.

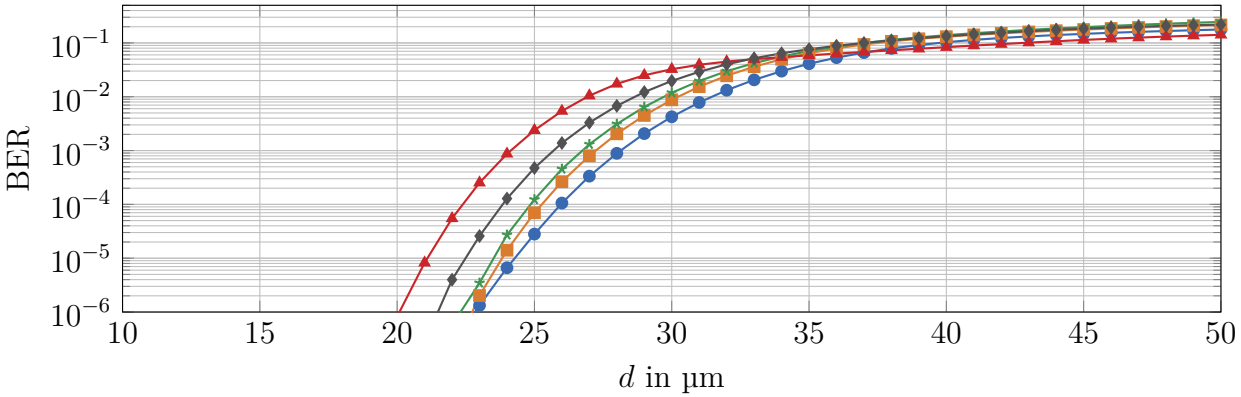
Numerical Results

Fig. 4.14 shows the BER performance of different RLL coded transmissions over varying transmission distances. In addition to the usual defined RLL codes, their bitwise inverted versions, which are denoted by $\left(\overline{n_\mathrm{c}, k_\mathrm{c}, \tau_\mathrm{min}, \tau_\mathrm{max}}\right)$, are also considered. An $\left(\overline{n_\mathrm{c}, k_\mathrm{c}, \tau_\mathrm{min}, \tau_\mathrm{max}}\right)$ RLL limits the minimum and maximum number of ones between two consecutive zeros. Thus the number of ones can be controlled. This can be advantageous in molecular communication. Among the pairs of non-inverted and inverted RLL codes, Fig. 4.14 shows the more powerful ones.

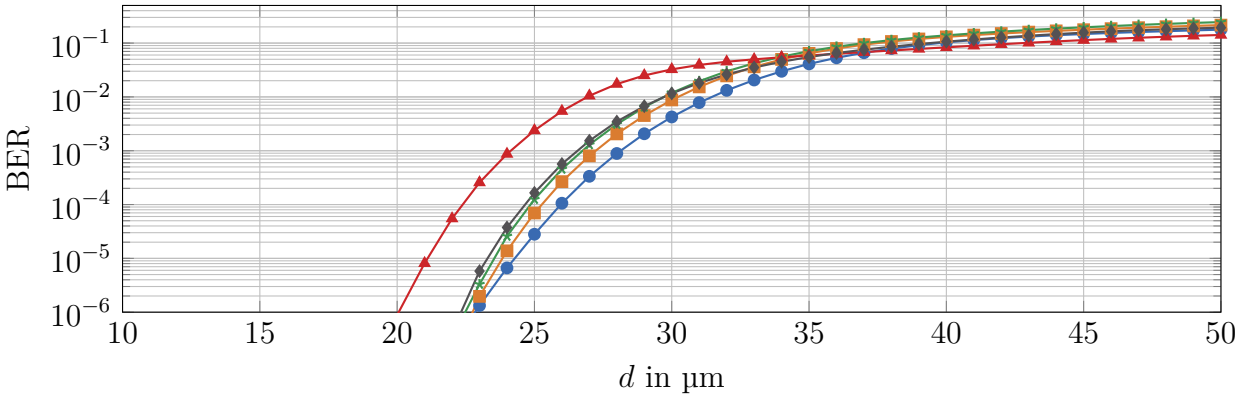
Fig. 4.14(a) shows the BER results for the unnormalized case. It can be observed that all considered RLL codes lead to a coding gain. It can also be seen that a reduced number of ones is an advantage. For RLL codes with no minimum and a maximum run length limitation, the inverted RLL codes are more powerful. If the minimum run length is limited and the maximum unlimited, the classical non-inverted RLL code is advantageous. The two



(a) Unnormalized.



(b) T normalized.



(c) T & N normalized.

Figure 4.14: Bit error rate performance of RLL encoded transmissions as a function of transmission distance. Parameters according to Appendix C.10.

$(\overline{2,1,0,1})$ and $(5,3,1,\infty)$ RLL codes with lower code rate achieve the maximum gain over uncoded transmission. With a target BER of 10^{-3}, they increase the maximum transmission distance from 28 µm to approximately 31 µm.

In Fig. 4.14(b), the symbol duration is normalized to the information rate. As described in Section 4.1, the shorter symbol durations lead to a degradation of the results. The degradation is so strong that no improvement can be achieved compared to uncoded transmission.

Fig. 4.14(c) displays the case in which not only the information rate but also the mean number of molecules per info bit is fixed. While this normalization means an attenuation of the signal for the block and convolutional codes discussed in Section 4.1 and Section 4.2, it can have a positive effect in the case of RLL codes. Compared to Fig. 4.14(b), the $(5,3,1,\infty)$ RLL code benefits the most. Because it requires at least one binary zero between two consecutive binary ones, the code words tend to have fewer binary ones than the uncoded transmission. Thus, more molecules can be released for the fewer binary ones. Nevertheless, the effect of the normalized symbol duration is too strong and no gain can be achieved over the uncoded transmission.

4.3.2 Weak Sequences Preventing Mapping

Weak sequence preventing mapping (WSPM) can be used to create a codebook of an $(n_\mathrm{c}, k_\mathrm{c})$ block code. The idea is to first perform an error analysis of all possible 2^{n_c} code words. Then the 2^{k_c} least error-prone code words are assigned to the 2^{k_c} info words to create a codebook. Fig. 4.15 shows an example of an error analysis for sequences of length $n_\mathrm{c} = 4$. For a $R = 1/4$ WSPM code, the $2^{k_\mathrm{c}} = 2$ code words [0 0 0 0] and [0 0 0 1] would be assigned to the two info words. If the info words are not equally likely, this should be taken into account in the code word assignment, i. e. the most likely info word should be assigned to the least error-prone code word. If a non-valid code word is received, an error correction to the valid code word with the smallest distance can take place. To improve the correction performance, a soft distance is considered taking LLRs into account.

Numerical Results

Before Decoding Fig. 4.16 shows the BER performance of certain WSPM codes over the transmission distance. These are the BER results before decoding, which highlight the impact of avoiding error-prone sequences. The different WSPM codes are compared with respect to different normalizations.

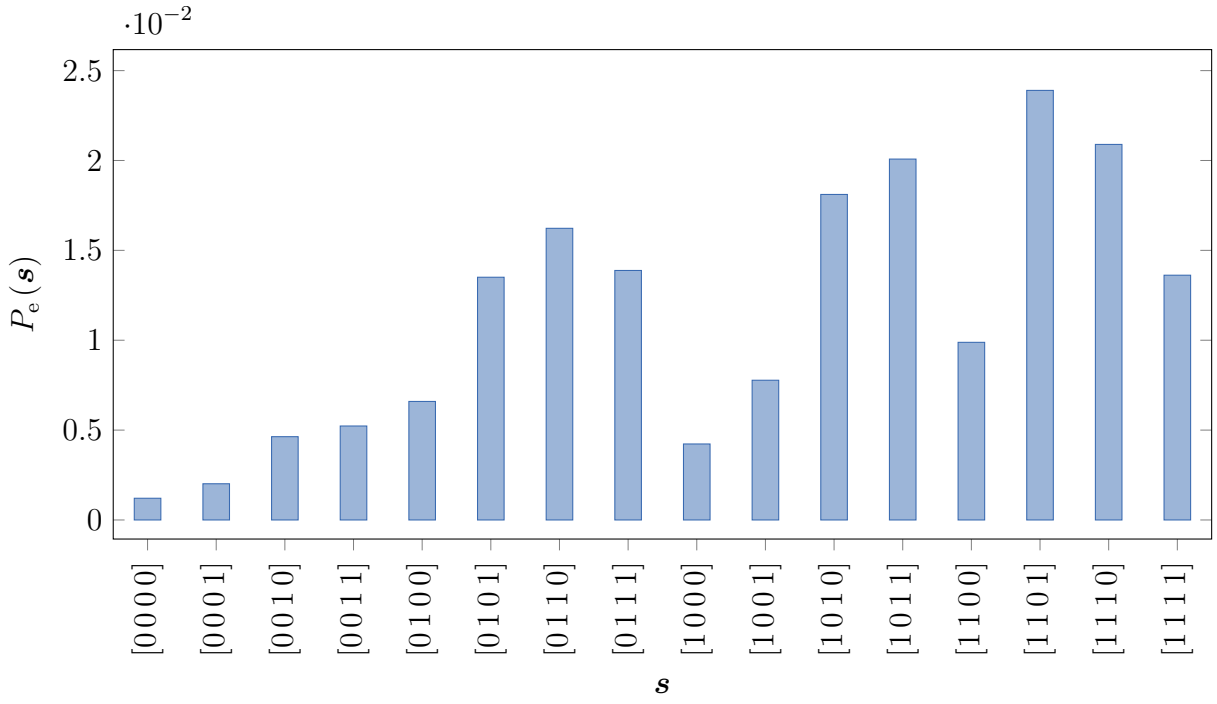
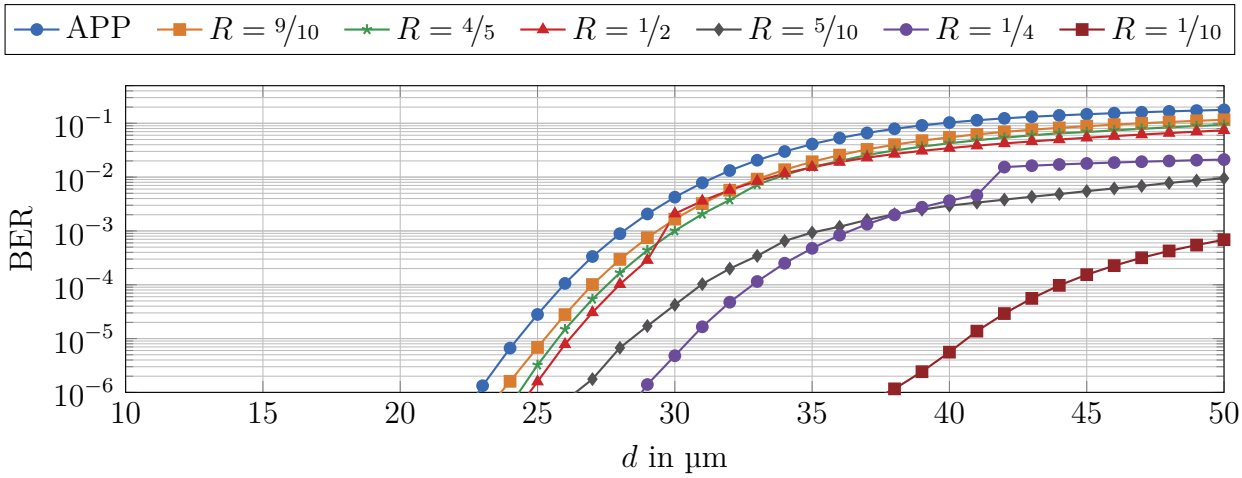


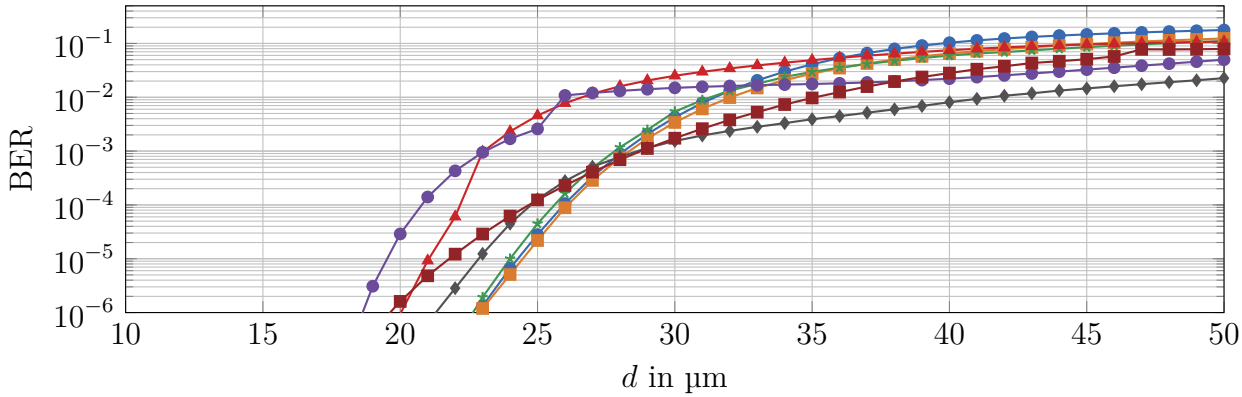
Figure 4.15: Sequence error analysis for $n_c = 4$ and $d = 30\,\mu m$. Parameters according to Appendix C.12.

The unnormalized case in Fig. 4.16(a) illustrates that all investigated WSPM codes achieve a gain compared to uncoded transmission. This gain is in general larger the lower the code rate. This is consistent with the fact that in higher-rate WSPM codes more sequences are seleceted and thus less error-prone sequences can be avoided. Comparing the $R = 1/2$ WSPM with the $R = 5/10$ WSPM code, it can be seen that the latter of the two results in a lower BER due to its larger variety of selectable sequences. From this observation, it can be concluded that n_c should be chosen as large as possible in terms of BER performance. Fig. 4.16(a) shows a significant increases in BER of the $R = 1/2$ WSPM and $R = 1/4$ WSPM code at $d = 30\,\mu m$ and $d = 42\,\mu m$, respectively. These increases are caused by a changed code word assignment. For example, considering the $R = 1/4$ WSPM code, the two least error-prone sequences for $d \leq 41\,\mu m$ are equal to $[0\,0\,0\,0]$ and $[0\,0\,0\,1]$, as shown in Fig. 4.15. For distances $42\,\mu m \leq d \leq 50\,\mu m$, the two least error-prone sequences change to $[0\,0\,0\,0]$ and $[1\,1\,1\,1]$. Among the codes considered, the $R = 1/10$ WSPM code leads to the largest increase in distance for a target BER of 10^{-3}. Compared to uncoded transmission, the maximum transmission distance can be increased from $28\,\mu m$ to $51\,\mu m$.

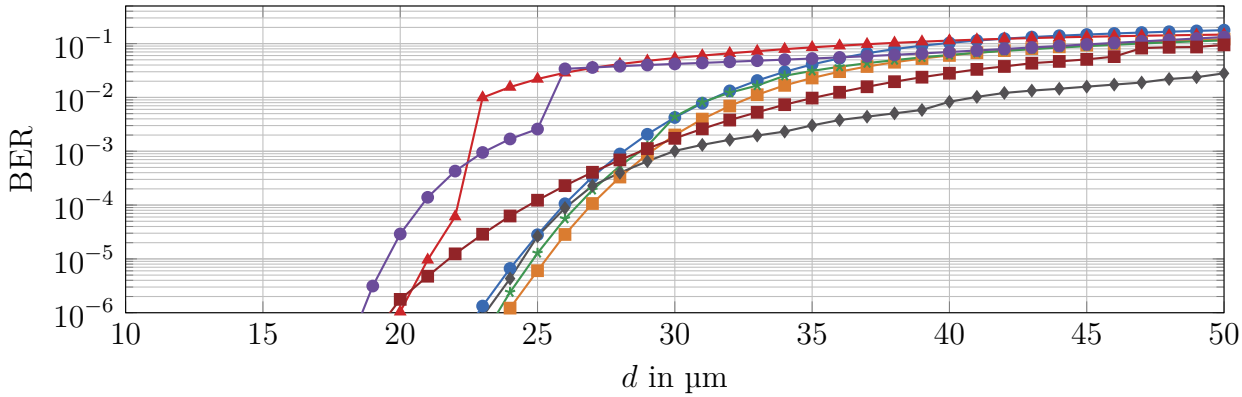
In Fig. 4.16(b), the symbol duration is normalized to a fixed information rate. As discussed in Section 4.1 and Section 4.2, this normalization leads to a degradation of the



(a) Unnormalized.



(b) T normalized.



(c) T & N normalized.

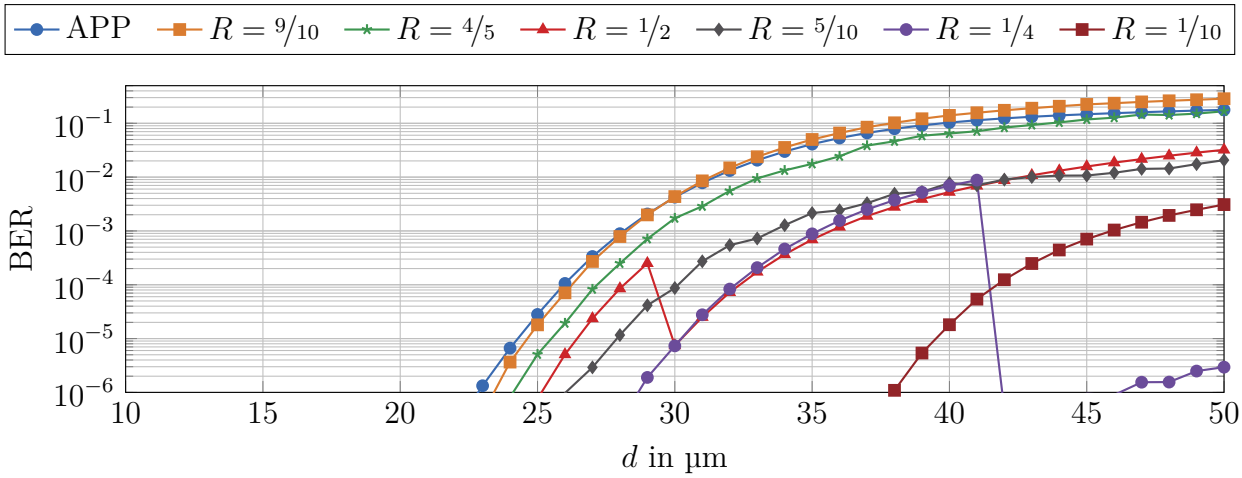
Figure 4.16: Bit error rate performance of WSPM encoded transmissions as a function of transmission distance before decoding. Parameters according to Appendix C.13.

BER performance of the coded transmission, because there is less time per symbol for the molecules to reach the receiver. The lower the code rate, the shorter the symbol duration and thus the larger the degradation. In contrast to the results in Fig. 4.16(a), the WSPM codes in Fig. 4.16(b) show no or only a slight gain over uncoded transmission. At the target BER of 10^{-3}, this distance increase is maximal around 1 µm for the $R = 5/10$ WSPM and $R = 1/10$ WSPM codes.

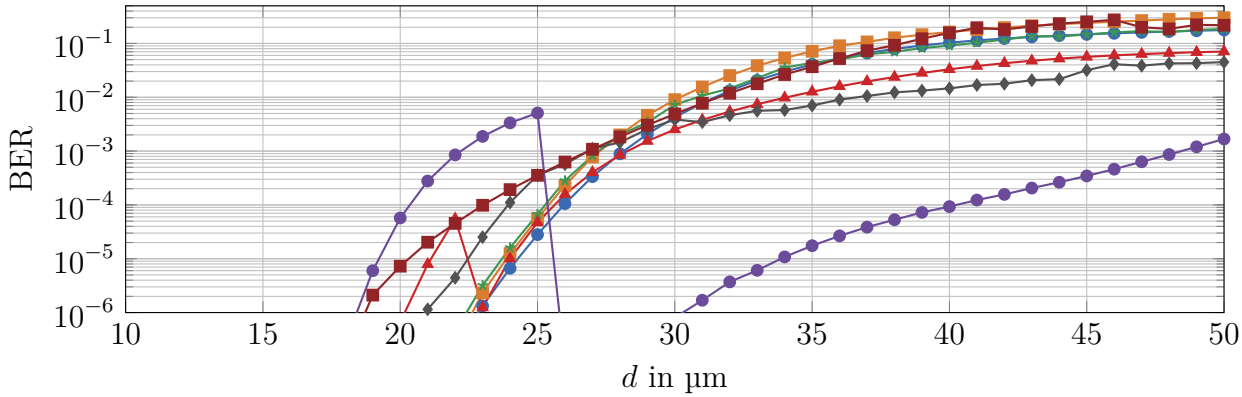
In Fig. 4.16(c), in addition to the symbol duration, the number of molecules released is normalized as well. The normalization ensures that the average number of molecules per info bit is equal to the uncoded case for all codes. Depending on the Hamming weight of the code words, this normalization can have a positive effect (e. g. $R = 5/10$ WSPM code at $d \leq 39\,\mu\text{m}$), a negative effect (e. g. $R = 1/2$ WSPM code at $d \geq 23\,\mu\text{m}$) or no effect (e. g. $R = 1/2$ WSPM code at $d \leq 22\,\mu\text{m}$) on the BER performance of the coded transmission. Among the WSPM codes under consideration and a target BER of 10^{-3}, the $R = 5/10$ WSPM code provides the largest increase. The maximum transmission distance can be increased from 28 µm to 30 µm.

After Decoding Fig. 4.17 present the BER results of WSPM codes achieved at different transmission distances and under consideration of different normalizations. In contrast to Fig. 4.16, the BER performances after decoding are depicted.

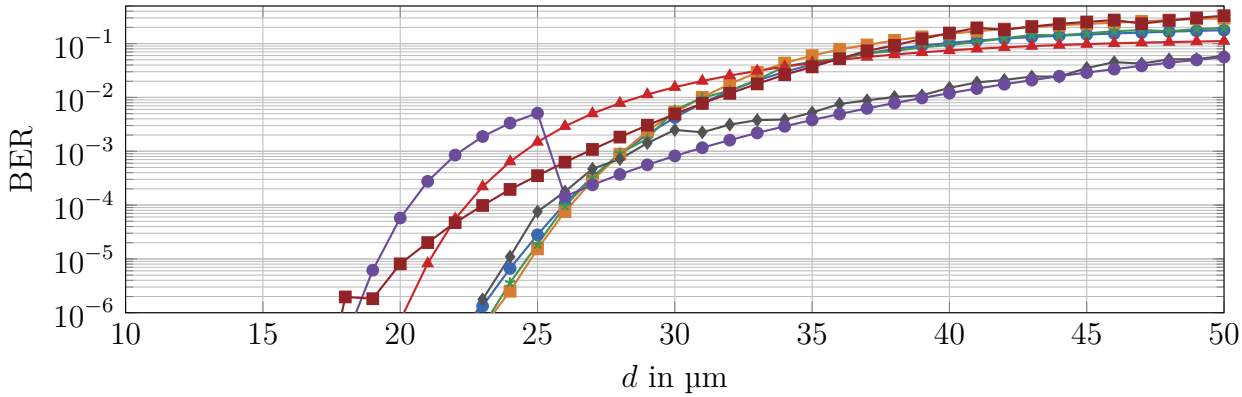
Fig. 4.17(a) shows the unnormalized case. It can be seen that all WSPM codes achieve a gain over uncoded transmission. As expected, a larger gain is achieved by lower code rates. Large discontinuities are particularly noticeable among codes with low code word lengths, i. e. $R = 1/2$ and $R = 1/4$ WSPM code. The reason for these discontinuities is the different potential of error correction depending on the mapping. The codebook is only generated from the error probabilities of the individual possible code words and does not consider the potential of error correction. If the $R = 1/4$ WSPM code is taken exemplarily, the two sequences with the lowest error probability for $d \geq 42\,\mu\text{m}$ are [0 0 0 0] and [1 1 1 1]. For $d \leq 41\,\mu\text{m}$, however, the sequences [0 0 0 0] and [0 0 0 1] are the least error-prone, as shown in Fig. 4.15. The Hamming distance between [0 0 0 0] and [1 1 1 1] is much greater than between [0 0 0 0] and [0 0 0 1], hence the former provides a stronger error correction, which is reflected in a lower BER. If the code word length is longer, there will be more balanced Hamming distances between the code words. Therefore, the $R = 5/10$ WSPM code is much less discontinuous than the $R = 1/2$ WSPM code. Depending on the error correction potential, an improvement or a degradation of the BER through decoding can be observed compared to Fig. 4.16(a). For a target BER of 10^{-3}, the $R = 1/10$ WSPM



(a) Unnormalized.



(b) T normalized.



(c) T & N normalized.

Figure 4.17: Bit error rate performance of WSPM encoded transmissions as a function of transmission distance after decoding. Parameters according to Appendix C.13.

achieves a maximum transmission range increase from 28 µm to 46 µm. In the region of larger distances, the $R = 1/4$ WSPM is equivalent to $R = 1/4$ repetition code without interleaving. It significantly increases the maximum transmission range to 76 µm.

In Fig. 4.17(b) the symbol duration is normalized to the information rate. As already discussed e. g. in Section 4.1, the normalization has a negative effect on the BER performance of the coded transmission. The negative impact is so strong that all coded transmissions at a target BER of 10^{-3} no longer achieve any gain over uncoded transmissions. The only exception is the $R = 1/4$ WSPM code, which allows a maximum transmission distance of 48 µm. For lower distances, however, it is worse than uncoded transmission due to its change to sequences with less Hamming distance between each other. In comparison to Fig. 4.16(b), it can again be observed that decoding can have a positive or negative impact on BER performance depending on the Hamming distance between the code words.

Fig. 4.17(c) displays the BER performance applying both a normalization of the symbol duration and a normalization of the average number of released molecules per info bit. Depending on the Hamming weight of the code words, this can lead to an increase or a reduction of the released molecules per code bit. If the number is increased, this results in an increase of the signal strength and thus an improvement of the BER performance compared to Fig. 4.17(b). Due to the improvement, $R = 9/10$, $R = 4/5$, and $R = 5/10$ WSPM codes achieve a similar maximum transmission distance as uncoded transmission. The $R = 1/4$ WSPM code degrades to 30 µm and is still worse for lower distances than uncoded transmission. Compared to Fig. 4.16(c), the impact of decoding on the BER performance is again visible.

4.3.3 Weak Sequence Preventing Convolutional Code

Similar to WSPM, weak sequence preventing convolutional codes (WSPCCs) are created on the basis of an error analysis. The most error robust sequences are used to design a convolutional code. In principle there are many possibilities to design the convolutional code, here a special one is presented, which is generated by a simple procedure. In total, k_c info bits are mapped to n_c code bits to create a convolutional code with code rate $R = k_\mathrm{c}/n_\mathrm{c}$. The structure of the code consists of 2^{2k_c} states. The connection between the states results from a regular 2^{k_c}-ary state diagram with memory length $L = 2$. Subsequently the 2^{2k_c} states are assigned to the 2^{2k_c} most error robust sequences of length n_c obtained from the error analysis. The state transitions are also assigned to the 2^{2k_c} most error robust sequences in such a way that a convolutional code is generated whose n_c code bits depend on the previous n_c code bits. The APP decoding, which works on the state diagram, is

used for decoding.

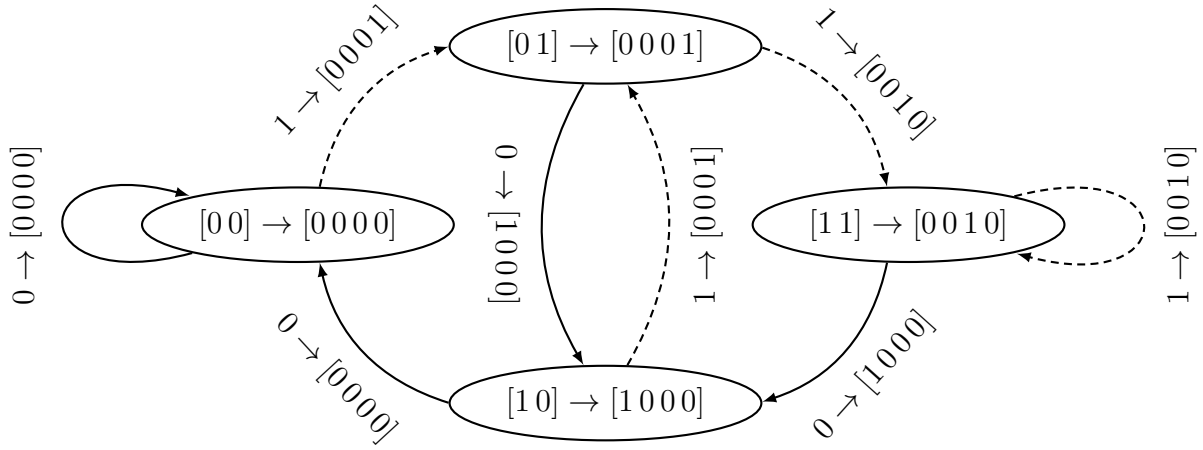


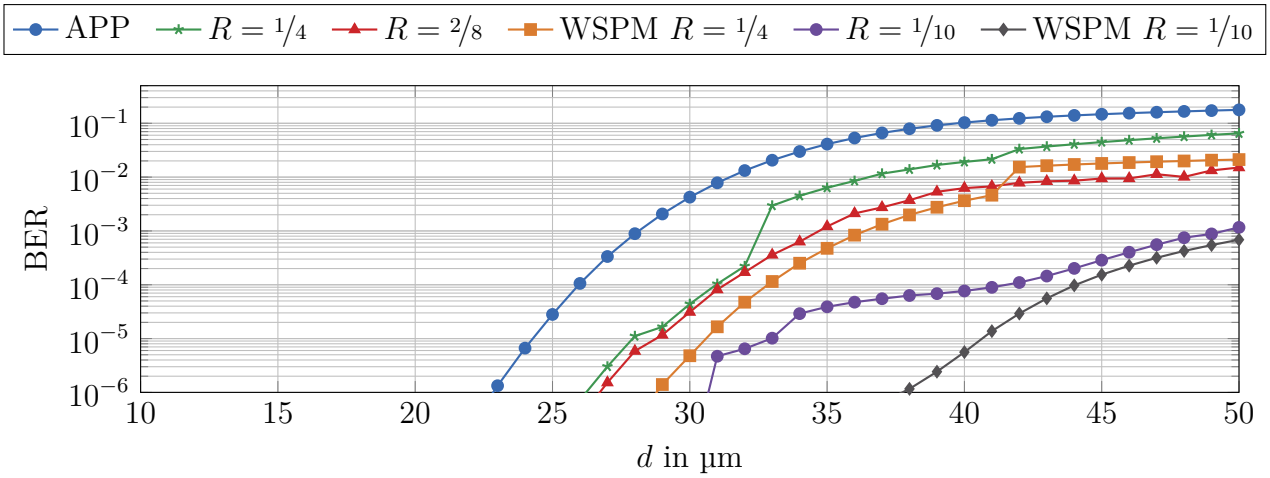
Figure 4.18: State diagram of a WSPCC with $k_c = 1$ and $n_c = 4$ for $d = 30\,\mu m$.

Fig. 4.18 shows an example construction of a $R = 1/4$ WSPCC with $k_c = 1$ and $n_c = 4$. The basic structure results from a regular binary state diagram with memory length $L = 2$. The $2^{2 \cdot 1} = 4$ most error robust sequences from Fig. 4.15, i. e. [0 0 0 0], [0 0 0 1], [1 0 0 0], and [0 0 1 0], are assigned to the four states [0 0], [0 1], [1 0], and [1 1], respectively. The transitions also correspond to the four most error robust sequences, so that the n_c code bits only depend on the previous states, i. e. the past n_c code bits.

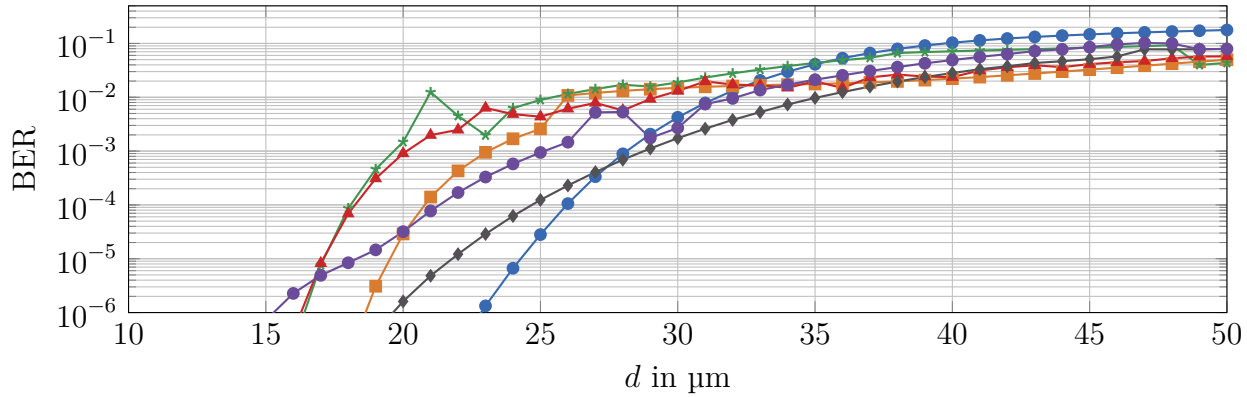
A WSPCC created according to this construction rule has certain limitations in the achievable code rate. Thus only WSPCCs with a code rate $R \leq \frac{1}{2}$ can be designed, whereby with equality, no more gain from the weak sequence preventing is achieved, since all possible sequences with length 2^{n_c} are assigned to the states. In order to generate arbitrary code rates, code puncturing and repetitions can be used, as with the convolutional codes from Section 4.2. However, the focus in this dissertation will not be on puncturing and repetition of WSPCC.

Numerical Results

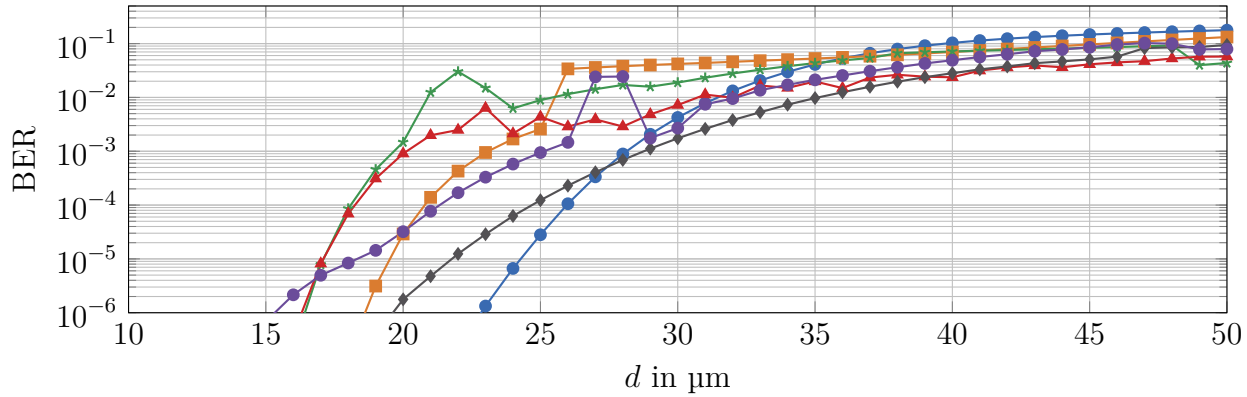
Before Decoding Fig. 4.19 presents the BER results of certain WSPCCs with different rates. The BER performances before decoding are analyzed with respet to different normalizations. In addition, the BER results of the uncoded and WSPM codes of the same rate are included for comparison.



(a) Unnormalized.



(b) T normalized.



(c) T & N normalized.

Figure 4.19: Bit error rate performance of WSPCC encoded transmissions as a function of transmission distance before decoding. Parameters according to Appendix C.13.

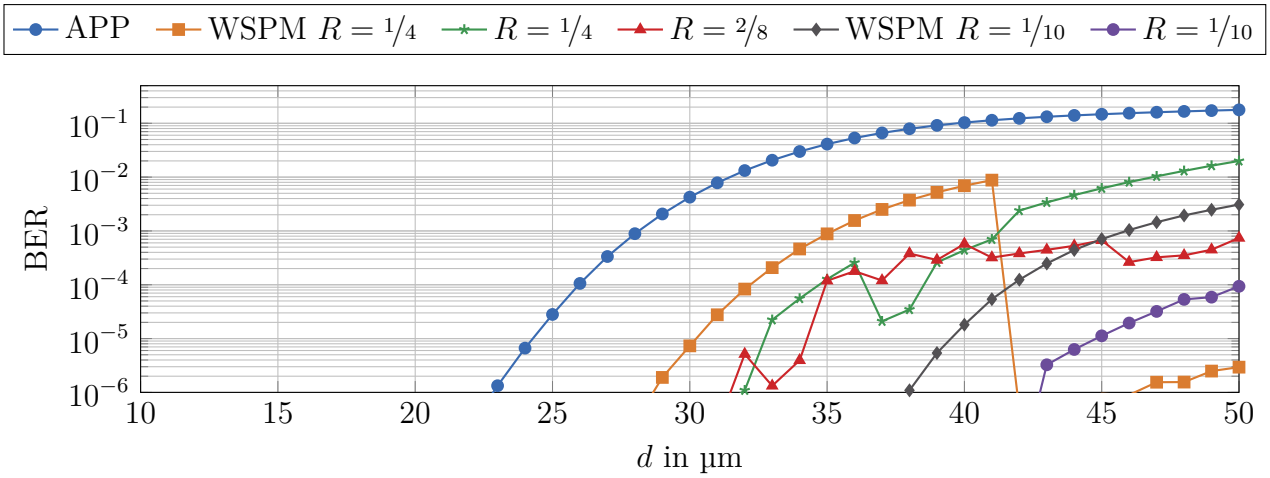
Fig. 4.19(a) shows the results of the unnormalized case. As expected, WSPCCs with lower code rates show better BER performance due to their higher degree of redundancy. It is further demonstrated that by avoiding error-prone sequences, the BER can be reduced compared to uncoded transmission. Using the example of the $R = 1/4$ and $R = 2/8$ WSPCCs, it is demonstrated that WSPCCs with equal code rates and higher n_c are more powerful. As with WSPM codes, this is due to the increasing number of selectable sequences with increasing n_c. As discussed in Section 4.3.2, the reason for the sudden increases in the BER curves is due to a change in the selected code words. For a target BER of 10^{-3}, the $R = 1/10$ WSPCC leads to the greatest increase in transmission distance. Compared to uncoded transmission, the distance can be increased from 28 µm to 49 µm.

In the results of Fig. 4.19(b), the symbol duration is normalized to a fixed information rate. As expected, this normalization leads to a degradation of the BER performance of the coded transmission compared to Fig. 4.19(a). This degradation becomes larger the smaller the code rate becomes, as the symbol duration becomes shorter. In this case, none of the WSPCCs under consideration shows an improvement in the maximum transmission distance at a target BER of 10^{-3} compared to the uncoded case.

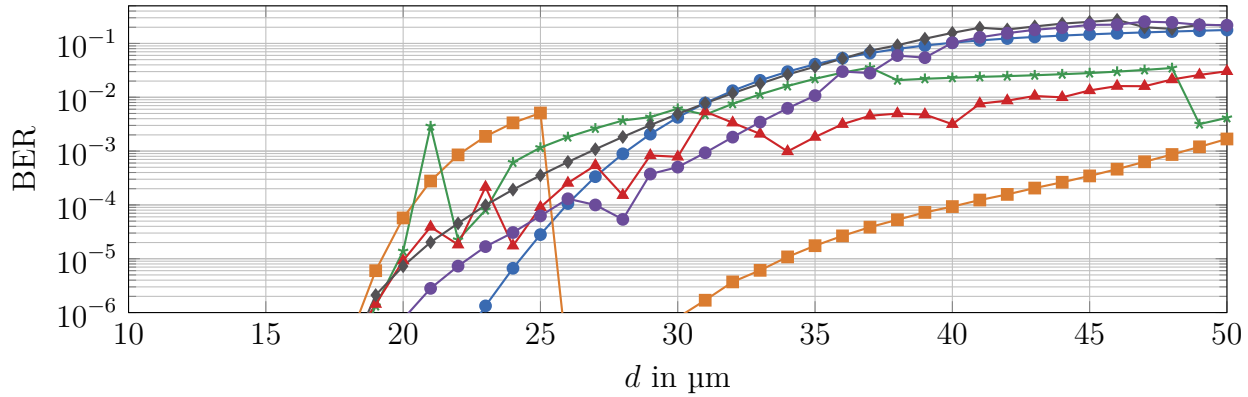
In the case shown in Fig. 4.19(c), in addition to a fixed information rate, the average number of molecules per info bit is normalized as well. Depending on the Hamming weight of the code words, this can lead to an improvement, degradation or no effect. For the WSPCCs under consideration, this additional normalization tends to lead to no significant improvement in bit error rates. Thus, there is no increase in the maximum transmission distance at a target BER of 10^{-3} in this case either.

Compared to WSPM codes of the same rate in Fig. 4.19, it can be observed that WSPCCs have a worse performance especially in the lower BER regions. This result is to be expected because the structure for creating a WSPCC requires more sequences to be selected from the sequence error analysis than in the case of WSPM codes of the same rate. Thus, the overall error probability of WSPCCs is higher than that of comparable WSPM codes. It should be noted that these are the results before decoding. Due to the decoding gain, this behavior can be reversed as examined in the following.

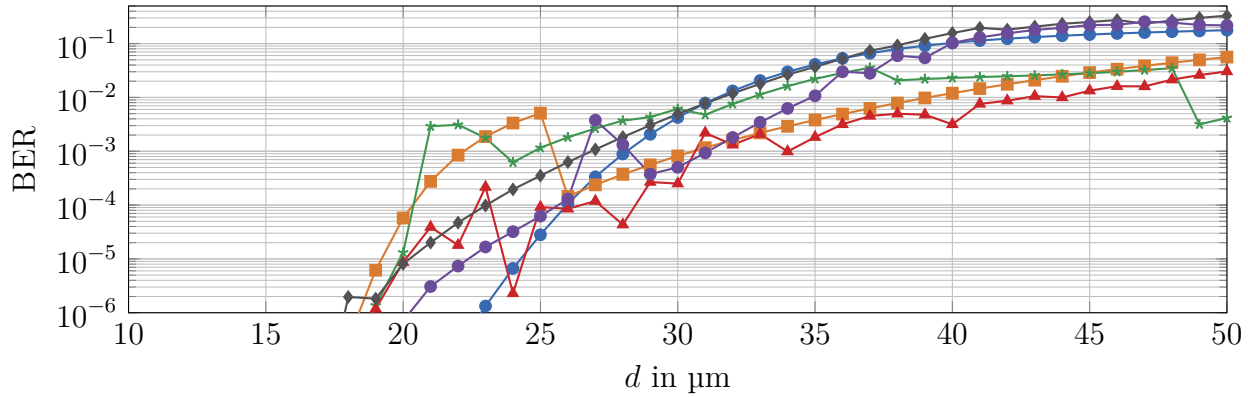
After Decoding Fig. 4.20 presents BER results of WSPCC coded transmissions after decoding. In addition to different transmission distances, different normalizations are considered. As a comparison, the results of the WSPM codes from Section 4.3.2 with comparable code rates are shown as well.



(a) Unnormalized.



(b) T normalized.



(c) T & N normalized.

Figure 4.20: Bit error rate performance of WSPCC encoded transmissions as a function of transmission distance after decoding. Parameters according to Appendix C.13.

Fig. 4.20(a) shows the unnormalized case in which all coded transmissions show a gain over the uncoded transmission. The low rate $R = 1/10$ WSPCC is more powerful than the $R = 1/4$ WSPCC and allows an increase of the maximum transmission distance to 59 µm. In contrast to Fig. 4.19(a), the WSPCC can increase the error correction performance over a wide range compared to the WSPM code, which is reflected in a lower BER. This is due to a higher decoding gain caused by the convolutional code structure. Comparing the $R = 1/4$ with the $R = 2/8$ WSPCC, the $R = 2/8$ code is superior at most transmission distances. From this, it can be concluded that the WSPCC code should be constructed from a sequence error analysis with the largest possible sequence length.

In Fig. 4.20(b), the symbol duration is normalized to the information rate. Despite the normalization, the $R = 2/8$ and $R = 1/10$ WSPCC codes lead to an improvement in transmission distance at a target BER of 10^{-3}. The $R = 1/10$ WSPCC with 31 µm shows the largest gain. In contrast to the results in Fig. 4.19(b), the $R = 1/10$ WSPCC achieves a lower BER than the $R = 1/10$ WSPM code. This is again due to the higher decoding gain originated from the convolutional code structure. The $R = 1/4$ and $R = 2/8$ WSPCC show an improvement only at shorter transmission distances.

Fig. 4.20(c) shows the case where the average number of released molecules per info bit is normalized in addition to the symbol duration. Depending on the Hamming weight of the code sequences, this leads to a signal amplification, and thus, to an improvement of the WSPCC performance compared to Fig. 4.20(b). While the $R = 1/4$ WSPCC deteriorates over long distances, the $R = 2/8$ WSPCC improves. For a target BER of 10^{-3}, $R = 2/8$ and $R = 1/10$ WSPCCs show comparable gains. Both lead to an increase of the maximum transmission distance of approximately 31 µm, which is larger than the gain from $R = 1/4$ convolutional code in Fig. 4.13(c). Thus, both achieve a similar gain as the $R = 1/4$ WSPM, but show a lower BER at certain transmission distances, especially at short ones.

It should be noted that the WSPCC codes are built according to the previously presented rather simple structure. Other structures may be more optimal and lead to a higher gain. The results show that taking the channel characteristic into account can be helpful to design powerful channel codes.

4.4 Spreading Codes

Spreading codes or spreading sequences [PS07] have their name due to their ability to spread information across a wide bandwidth. Their best known field of application is their use in code-division multiple access (CDMA) [PO98] multi-user systems. In contrast to

time-division multiple access and frequency-division multiple access, users transmit their data over a shared frequency spectrum at the same time. At the receiver side, the different users can be separated by a correlation with their corresponding spreading sequence. Consequently, the spreading sequences should have cross-correlations as small as possible to each other. Besides CDMA systems, spreading codes can also be applied in code-division multiplexing (CDM) systems or to counteract ISI. For the latter purpose, a good autocorrelation function is crucial [WO03]. The optimum autocorrelation function would have the shape of a Kronecker delta.

The spreading of the data to a larger bandwidth is achieved by the fact that the chip duration T_c of an element of the spreading sequence is typically smaller than the symbol duration T of the data. The bandwidth is proportional to $1/T_\mathrm{c}$ and $1/T$, respectively, so it is increased accordingly. The ratio $K_\mathrm{s} = T/T_\mathrm{c}$ is referred to as the spreading factor. It should be noted that in molecular communication the spectrum cannot be spread arbitrarily, because the usable bandwidth is limited by the low-pass characteristic of the diffusion-based molecular transmission channel (see Fig. 3.30).

In the following, the use of spreading codes in molecular communication is described. Similar to orthogonal frequency-division multiplexing (OFDM) in Section 3.3, the focus is on concentration-based modulation with only a single type of molecule. Furthermore, a single user transmission is considered, therefore the use of spreading codes is only investigated for ISI suppression. CDM transmission can be interpreted as CDMA, where all transmitters have the same distance to the receiver. A multi-user scenario in conjunction with molecule shift keying (MoSK) is studied in [KD+17].

4.4.1 Walsh-Hadamard Codes

Walsh-Hadamard codes [PS07] are orthogonal codes that can be created by a simple construction rule. They are non-cyclic codes, i. e., a cyclic shift of a code word does not lead to a new code word. Here, Walsh-Hadamard codes are considered whose quadratic Hadamard matrix construction can be described according to the Sylvester method. Binary phase shift keying (BPSK) symbols $\{+1, -1\}$ are considered for the construction. If one follows the Sylvester method, a $2n_\mathrm{c} \times 2n_\mathrm{c}$ Hadamard matrix $\boldsymbol{W}_{2n_\mathrm{c}}$ results from an $n_\mathrm{c} \times n_\mathrm{c}$ Hadamard matrix $\boldsymbol{W}_{n_\mathrm{c}}$ according to

$$\boldsymbol{W}_{2n_\mathrm{c}} = \begin{bmatrix} \boldsymbol{W}_{n_\mathrm{c}} & \boldsymbol{W}_{n_\mathrm{c}} \\ \boldsymbol{W}_{n_\mathrm{c}} & -\boldsymbol{W}_{n_\mathrm{c}} \end{bmatrix}, \tag{4.16}$$

where $\boldsymbol{W}_1 = +1$ applies. The columns or rows of the Hadamard matrix represent Walsh codes and are orthogonal to each other.

4.4.2 Maximum Length Sequences

Maximum length sequences [GW+82] are pseudo-random binary sequences. Each sequence represents a cyclic code, since a cyclic shift leads to a valid code word. The construction of maximum length sequences can be described by linear feedback shift registers, which are fed back in such a way that a sequence with maximum run length is created before it repeats periodically. Accordingly, ν shift registers result in maximum length sequences of length $2^\nu - 1$. In general, a maximum length sequence of length $n_\mathrm{c} = 2^\nu - 1$ has the following properties [Höh11]:

- It consists of $(n_\mathrm{c} - 1)/2$ symbols of type $+1$ and $(n_\mathrm{c} + 1)/2$ symbols of type -1.
- Exactly $1/2^c$ of all runs have the length c, where a run is defined by consecutive symbols with the same sign.
- The code words of a maximum length sequence are not orthogonal due to their odd length, but have a good even periodic autocorrelation function. The even periodic autocorrelation function is n_c for no shift and -1 for shifted values.

Table 4.8: Maximum length sequences with different runlengths.

n_c	Maximum length sequence
3	$[-1\ -1\ +1]$
7	$[-1\ -1\ -1\ +1\ +1\ -1\ +1]$
15	$[-1\ -1\ -1\ -1\ +1\ +1\ +1\ -1\ +1\ +1\ -1\ -1\ +1\ -1\ +1]$

Tab. 4.8 shows exemplary maximum length sequences with different run lengths.

4.4.3 Binary Complementary Sequences

Binary complementary sequences [Gol61] are pairs of BPSK sequences of equal length n_c which are characterized by their combined aperiodic autocorrelation function. While the individual sequences are not ideally autocorrelated, the sum of their autocorrelation functions is ideal and thus

$$R_x[k] + R_y[k] = 2n_\mathrm{c}\delta_k, \tag{4.17}$$

where δ_k is the Kronecker delta and $R_x[k]$ is the aperiodic autocorrelation function

$$R_x[k] = \sum_{j=0}^{n_c-k-1} x_j x_{j+k}. \tag{4.18}$$

Tab. 4.9 shows exemplary pairs of binary complementary sequences with different lengths.

Table 4.9: Binary complementary sequences with different lengths.

n_c	Binary complementary sequences	Aperiodic autocorrelation function
2	[+1 +1]	[+2 +1]
	[+1 −1]	[+2 −1]
4	[+1 +1 +1 −1]	[+4 +1 ±0 −1]
	[+1 +1 −1 +1]	[+4 −1 ±0 +1]
8	[+1 +1 +1 −1 +1 +1 −1 +1]	[+8 −1 ±0 +3 ±0 +1 ±0 +1]
	[+1 +1 +1 −1 −1 −1 +1 −1]	[+8 +1 ±0 −3 ±0 −1 ±0 −1]
10	[+1 +1 −1 +1 −1 +1 −1 −1 +1 +1]	[+10 −3 ±0 −1 ±0 +1 −2 −1 +2 +1]
	[+1 +1 −1 +1 +1 +1 +1 +1 −1 −1]	[+10 +3 ±0 +1 ±0 −1 +2 +1 −2 −1]

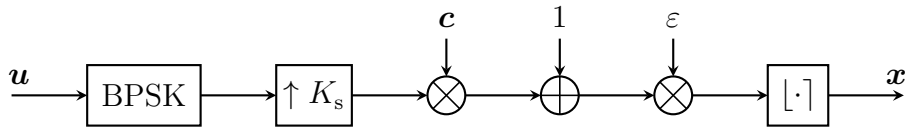
Considering, for example, the sequences of length $n_c = 2$, the aperiodic autocorrelation function accumulates to $[+4 \pm 0]$.

4.4.4 Spreading Codes for Unipolar Channels

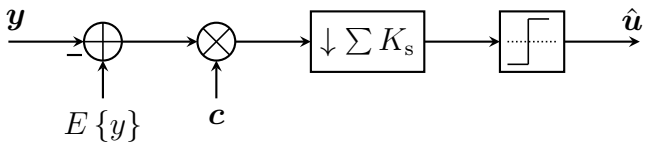
In classical wireless communication, typically bipolar spreading signals are transmitted. In molecular communication, however, only unipolar integer signals can be transmitted in the form of released molecules. A corresponding adaptation of the transmission and reception structure can be adopted from indoor wireless infrared communication [WO03], in which only unipolar signals can be transmitted. If unipolar-bipolar sequencing [OFa89] is applied, the same spreading sequences can be used as in radio-based systems. Fig. 4.21 shows the corresponding transmitter and receiver structure.

In the transmitter (Fig. 4.21(a)), the binary data bits $\boldsymbol{u}$ are first BPSK modulated as in radio-based systems, repeated K_s times and then multiplied element-wise by the spreading sequence $\boldsymbol{c}$. At this point, the signal consists of bipolar symbols $\{-1, +1\}$. To create a unipolar signal, a DC bias of 1 is added. The signal is then multiplied by a factor ε to adjust the mean symbol energy. To ensure the release of (integer-valued) molecules, the signal is rounded to nearest integer before transmission.

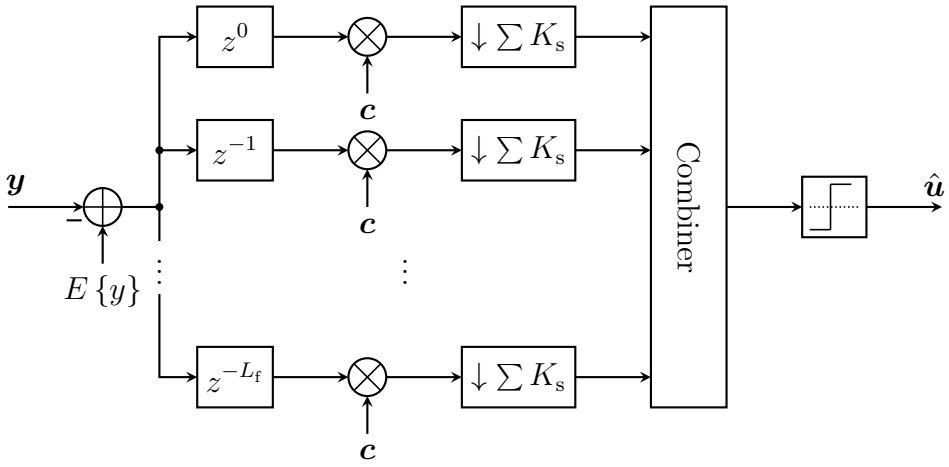
When a unipolar spreading sequence is transmitted and correlated with the corresponding bipolar spreading sequence, the correlation properties of the bipolar spreading sequence



(a) Spreading code transmitter structure.



(b) Spreading code receiver structure.



(c) Rake receiver structure.

Figure 4.21: Block diagram of unipolar spreading code transmitter and receiver structures.

with an additional fixed DC bias are preserved [WO03]. This is taken into account in the receiver structure in Fig. 4.21(b). As a first step, the estimated DC bias $E\{y\}$ is subtracted from the received molecules $\boldsymbol{y}$ in order to regain bipolar signals. The rest of the receiver has the structure of a classical spreading code receiver. The received signal is multiplied element-wise with the spreading sequence $\boldsymbol{c}$ and summed up for the duration of a data symbol (K_s samples). An FTD with a threshold value of $\xi = 0$ can finally be used to estimate the binary data bits $\hat{\boldsymbol{u}}$.

Fig. 4.21(a) and Fig. 4.21(b) show the structure for a single data stream, which can be used to avoid ISI and thereby achieve equalization, if the autocorrelation property of the spreading sequence is appropriate. The great advantage of this equalization-avoiding detection is its low-complexity structure. Above all, it functions independently of channel knowledge if the DC bias is estimated from the received values. If the structure is considered with binary complementary sequences, the data is spread in parallel with both complementary sequences and the result is summed element-wise before transmission. The receiver also correlates the received signal with both complementary sequences and adds the result element-wise before threshold detection. For CDM, each data stream has its own transmitter and receiver structure with its corresponding spreading sequence. All parallel data streams are added before transmission.

4.4.5 Rake Receiver

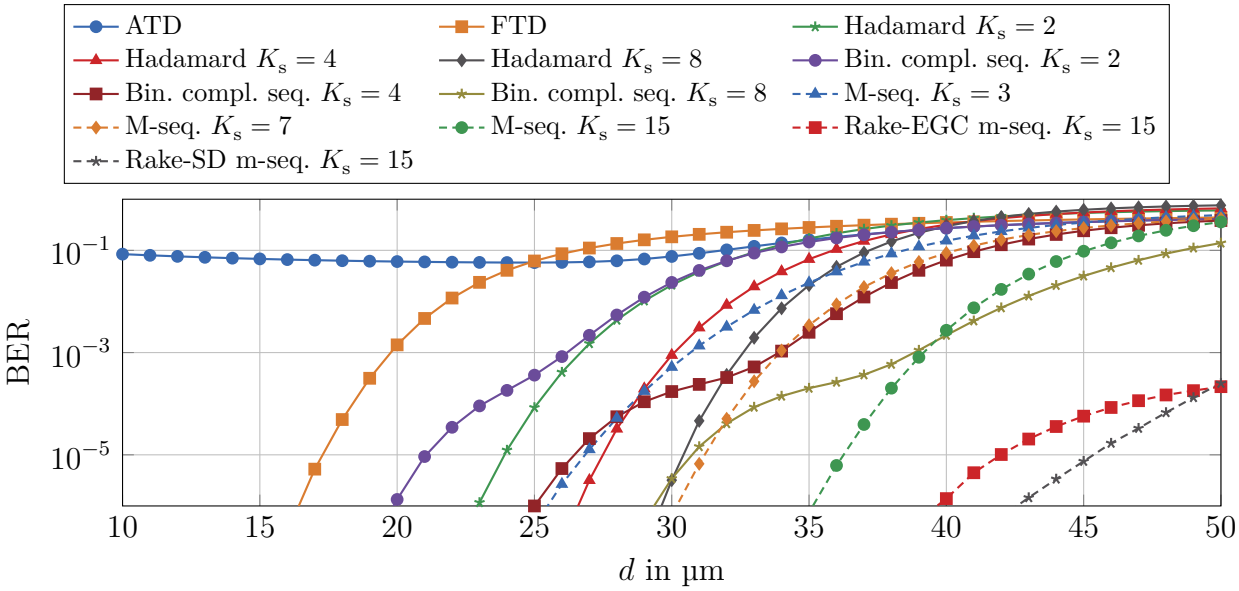
If the transmission channel has multipath propagation with differences in runtimes of more than one chip duration T_c, the received signal is affected by ISI. In case of good autocorrelation properties of the spreading sequence, the multipath interference, i. e. the ISI, can be used to increase the signal strength with the help of a rake receiver [PG58]. In the radio-based channel, a diversity gain can thus be achieved with statistically independent transmission paths. The structure of a rake receiver is shown in Fig. 4.21(c). It consists of at least two so-called fingers. Fig. 4.21(c) shows $L_\mathrm{f} + 1$ fingers according to the equivalent discrete-time channel model (EDTCM) with its $L + 1$ channel coefficients. Each finger correlates a delayed version of the received signal with the spreading sequence. Thus, each finger collects the signal energy of the multipath belonging to the delay, if the spreading sequence has the appropriate autocorrelation property. Finally, the energies of the fingers must be combined before a threshold detection can take place. For the combination, there are different approaches [Höh11]. Optimal is the so-called maximum ratio combining (MRC), in which the fingers are weighted according to their multipath channel strength before combination. Here, however, spreading codes are exploited as low-complexity detectors that

do not require channel knowledge. Therefore equal gain combining (EGC) and selection diversity (SD) are investigated as combiners. With EGC, all fingers are equally weighted before combination. SD, on the other hand, considers only a single finger that provides the strongest signal.

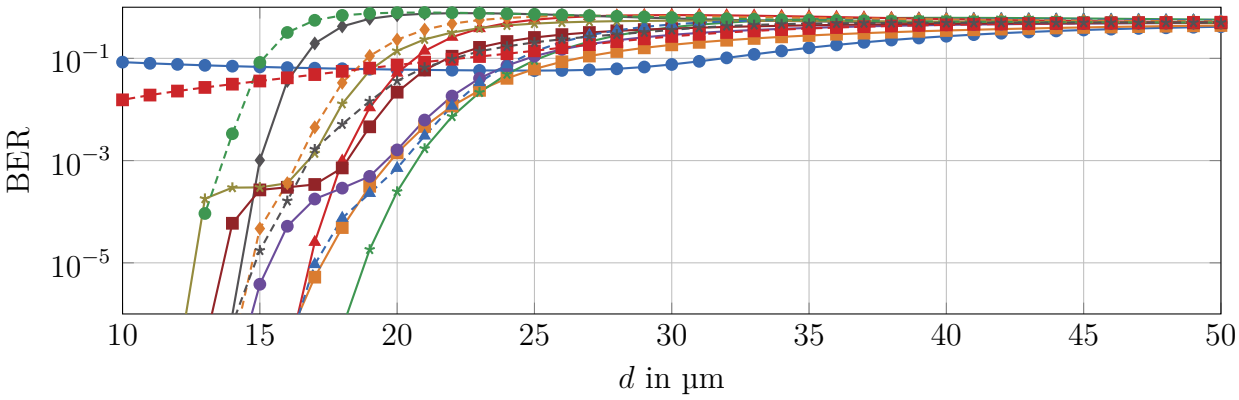
4.4.6 Numerical Results

Fig. 4.22 shows the results of the BER analysis when using spreading codes. In this analysis, spreading codes are used especially to create robustness against ISI by means of their autocorrelation property. The spreading code systems are investigated as an alternative to low-complexity receivers, since they do not require explicit channel knowledge. Therefore, the low-complexity adaptive threshold detector (ATD) and the simple FTD are considered as a comparison. While the ATD also works without explicit channel knowledge, the FTD is based on an optimally selected threshold, which depends on the transmission channel.

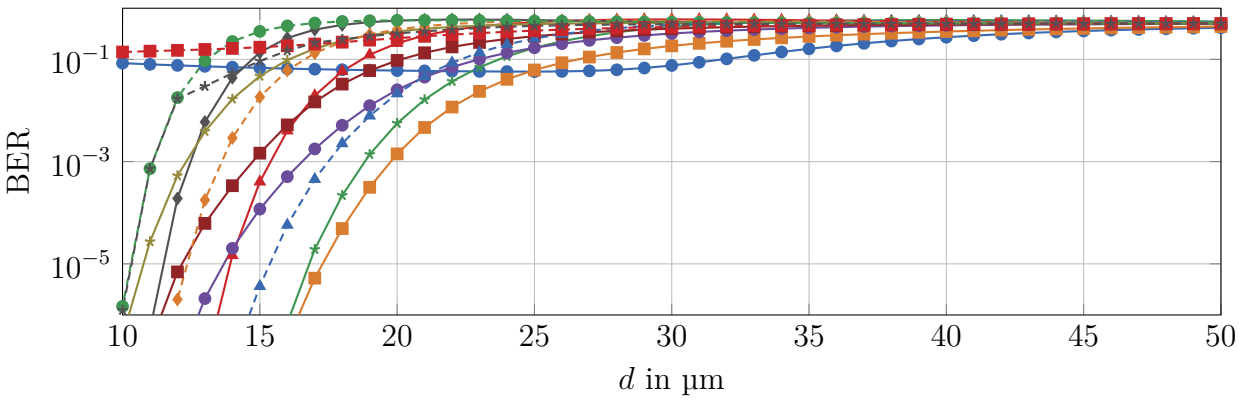
Fig. 4.22(a) shows the unnormalized case where there is no normalization to either a constant information rate nor a constant bit energy. For all spreading codes, it can be seen that their application yields a performance gain compared to FTD and ATD, since the chip duration in this scenario is equal to the symbol duration of the uncoded transmission. Furthermore, it can be seen that the performance of the spreading codes improves with increasing spreading factor. This effect can be explained by the increasing redundancy. In addition, the absolute channel memory duration is the same for all methods. Thus, with a perfect autocorrelation of the spreading codes, ISI is completely avoided if $K_\mathrm{s} \geq L + 1$ applies. Comparing the different considered spreading codes, it becomes clear that Hadamard sequences, except for the special case $K_\mathrm{s} = 2$, perform worst. This result is to be expected, because the autocorrelation function of Hadamard codes is far away from the optimum. A perfect autocorrelation is provided by the binary complementary sequences, which provide the best result at a target BER of 10^{-3} for comparable spreading factors. In addition, the signal strength is twice as high as with the other spreading codes due to the superposition of the two complementary sequences. Maximum-length sequences have an almost perfect autocorrelation function and their performance lies between the other two methods with comparable spreading factors. For the selection of spreading codes considered here and a target BER of 10^{-3}, the maximum-length sequence with $K_\mathrm{s} = 15$ leads to the largest improvement in distance. Compared to FTD, the maximum transmission distance can be increased from 20 µm to about 39 µm without the need for explicit channel knowledge. For this maximum-length sequence, the gain is exemplarily shown when using a rake receiver with EGC or SD, which also works without explicit



(a) Unnormalized.



(b) T normalized.



(c) T & N normalized.

Figure 4.22: Bit error rate performance of spreading encoded transmissions as a function of transmission distance. Parameters according to Appendix C.14.

channel knowledge. For $d < 50\,\mu\text{m}$, the rake receiver with SD leads to the best result by selecting the strongest channel, since weak channels gain too much influence with EGC by the identical weighting. For longer distances, however, the channel impulse response will flatten out, which is beneficial for the rake receiver in combination with EGC. It allows to extend the maximum transmission distance to 72 µm.

In Fig. 4.22(b), the information rate is the same for all methods. As a result, the chip duration becomes shorter as the spreading factor increases. Consequently, in contrast to Fig. 4.22(a), shorter spreading codes work better than longer spreading codes because the diffusion-based channel features a low-pass characteristic (see Fig. 3.30). By fixing the absolute channel memory duration, ISI now occurs in all investigated systems, which also affects the next information symbol. Nevertheless, Hadamard codes with $K_\text{s} = 2$ and maximum-length sequences with $K_\text{s} = 3$ show a small gain over FTD. At a target BER of 10^{-3}, the maximum transmission distance can be increased by about 1 µm by using Hadamard codes with $K_\text{s} = 2$. It should be emphasized that in contrast to FTD, no explicit channel knowledge is required. Considering the exemplary results with rake receivers, the performance can be increased in combination with SD. In contrast, the combination with EGC does not lead to a significant gain, since the number of weak channels that gain too much influence by equal weighting increases with increasing spreading factor. Nevertheless, all spreading codes show a gain for short transmission distances compared to ATD, as ATD inherently builds on existing ISI.

In Fig. 4.22(c), in addition to the information rate, the average number of molecules per info bit is chosen identical in all methods. This additional normalization leads to a reduction of the signal strength in all spreading code systems, since the average number of released molecules per chip decreases. As a result, the performance of all these systems decreases and there is no gain over FTD. Nevertheless, the maximum transmission distance of Hadamard codes with $K_\text{s} = 2$ at a target BER of 10^{-3} is only 1 µm shorter than that of FTD without the need for explicit channel knowledge.

In conclusion, it can be stated that spreading codes enable a low-complexity receiver which, based on the autocorrelation of the code, provides robustness against ISI. Depending on the normalization considered, this leads to a gain compared to FTD, which in contrast to the investigated spreading codes requires channel knowledge for the determination of the optimal threshold. Especially at short distances, the spreading code systems are more powerful than the ATD detector. In addition, a rake receiver can be used to further increase the performance.

4.5 Spatial Codes

If several transmit and receive antennas are used in a wireless transmission system, this is referred to as a multiple-input multiple-output (MIMO) system. In a MIMO system, information can be encoded not only in temporal but also in spatial dimension. A combined coding over the temporal and spatial dimension is called space-time code [PS07]. In classical wireless communication, space-time codes are used to increase the link reliability. By transmitting the information over several transmit antennas, a diversity gain can be achieved. In classical wireless communications, MIMO techniques already appertain to the state-of-the-art. In molecular communication, however, they have rarely been considered (see Section 1.3).

As in [DY$^+$17; DY$^+$18], the focus in this dissertation is on a diffusion-based molecular MIMO channel taking ISI as well as interlink interference (ILI) into account. Assuming a symmetrical 2×2 MIMO scenario, different spatial algorithms at the transmitter and at the receiver sides are studied. At the transmitter side, two different spatial coding techniques, namely Alamouti-type coding and repetition MIMO coding are proposed and analyzed. At the receiver side, focus is on two different receiver combining strategies, SD and EGC. The latter strategy is the same as MRC in symmetrical scenarios. To model the channel with two absorbing receivers, a trained artificial neural network (ANN) is used to acquire the MIMO channel impulse response.

4.5.1 MIMO Scenario

For the analysis of the different spatial codes and combining/selection strategies, the symmetrical 2×2 MIMO scenario from Fig. 4.23 is considered [DY$^+$17; DY$^+$18]. The same assumptions apply as in the single-input single-output (SISO) scenario assumed in Section 2.1, with the difference that the transmitter (Tx) is extended by a second point source and the receiver (Rx) by an identical second absorbing spherical receiver. In analogy to radio transmission, point sources and spherical receivers are referred to as antennas in the following, even if they differ significantly from the functionality of a classical antenna. Since this is a symmetrical scenario, the centers of the respective transmit and receive antennas are aligned with each other. Thus, the distance between Tx_0 and Rx_0 and between Tx_1 and Rx_1 is d. The separation distance between the transmit antennas Tx_0 and Tx_1 and between the receive antennas Rx_0 and Rx_1 is fixed at $d_\mathrm{s} \geq 2r$. In contrast to the SISO scenario in Section 2.1, the receiver also consists of a reflecting cuboid body to which the receive antennas Rx_0 and Rx_1 are attached as depicted in Fig. 4.23. The reflecting body is

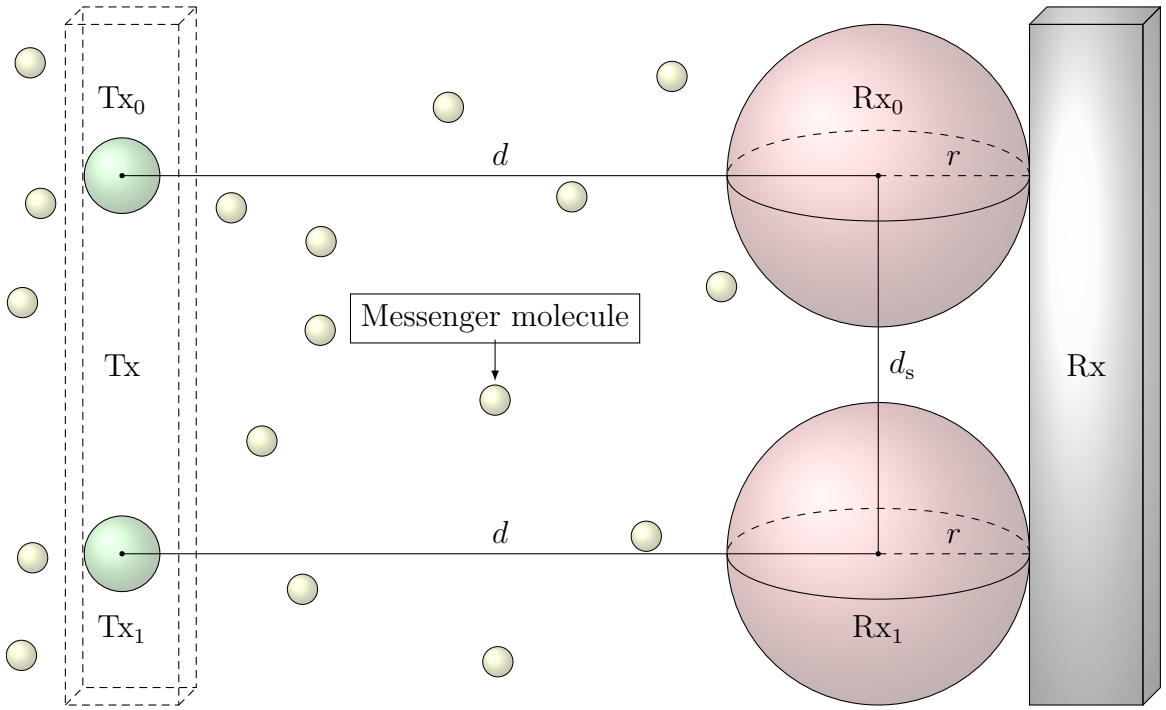


Figure 4.23: Visualization of the diffusion-based molecular 2×2 MIMO scenario under investigation.

assumed to be $2r$ wide, $d_\mathrm{s} + 2r$ high, and $4r$ deep.

The subchannel between transmit antenna Tx_i and receive antenna Rx_j can be described by the channel coefficients h_ℓ^{ji}. The total number of molecules that reach the receive antenna Rx_j during the kth transmission interval is given by the superposition of all associated subchannels. In general, if N_Tx transmit antennas are assumed, the EDTCM at Rx_j is extended to [DY+17; DY+18]

$$y_j[k] = \sum_{i=0}^{N_\mathrm{Tx}-1} \sum_{\ell=0}^{L} h_\ell^{ji} x_i[k-\ell] + \omega_j[k], \tag{4.19}$$

where $x_i[k]$ is the discrete-time representation of the modulated data symbol transmitted by Tx_i in the kth transmission interval and $\omega_j[k]$ represents the amplitude-dependent noise caused by the diffusive propagation of the molecules towards Rx_j. As discussed in Section 2.3.1, $y_j[k]$ follows a Poisson binomial distribution, which can be described by the summation of individual binomial distributions [DY+17; DY+18]

$$y_j[k] \sim \sum_{i=0}^{N_\mathrm{Tx}-1} \sum_{\ell=0}^{L} \mathcal{B}\Big(x_i[k-\ell], h_\ell^{ji}\Big)\,. \tag{4.20}$$

For the derivation and analysis throughout this section, OOK modulation is assumed at the transmit antennas, i. e. $x_i[k] \in \{0, N\}$.

4.5.2 Alamouti-type Code

The Alamouti code [Ala98] is an orthogonal space-time code originally invented for a two-antenna system. The structure of the code can be described by its 2×2 transmission matrix

$$\boldsymbol{T} = \begin{bmatrix} a_k & a_{k+1} \\ -a_{k+1}^* & a_k^* \end{bmatrix}. \tag{4.21}$$

The information is typically represented by a sequence of data symbols $\boldsymbol{a}$, which is generated by mapping the binary data sequence $\boldsymbol{u}$ onto data symbols. The kth data symbol of $\boldsymbol{a}$ is denoted by a_k. The rows of $\boldsymbol{T}$ represent two consecutive transmission intervals $[kT\ (k+1)T]$ and $[(k+1)T\ (k+2)T]$. The columns of $\boldsymbol{T}$ belong to the two transmit antennas Tx_0 and Tx_1. Consequently, in the first time slot $x_0[k] = a_k$ is transmitted via Tx_0 and $x_1[k] = a_{k+1}$ via Tx_1. In the second time slot, $x_0[k+1] = -a_{k+1}^*$ and $x_1[k+1] = a_k^*$ are transmitted accordingly. Thus, the information from both data symbols is transmitted via both transmit antennas, which leads to a diversity gain in radio-based channels. The Alamouti code is an orthogonal space-time code for complex data symbols. Due to the orthogonality $\boldsymbol{T}^{\mathrm{H}}\boldsymbol{T} = c\boldsymbol{I}_2$ is valid, where $\boldsymbol{T}^{\mathrm{H}}$ is the Hermitian of the transmission matrix, c is a constant factor, and $\boldsymbol{I}_2$ is the 2×2 idendity matrix. Therefore, ILI can be completely cancelled at the receiver side. Thus, the maximum-likelihood detector can be simplified, which makes the Alamouti code so famous in radio-based communication systems. If the transmission is affected by ISI, then orthogonality is no longer maintained. In this case, more complex detection algorithms such as maximum-likelihood sequence estimation (MLSE) are necessary [MH04]. Alternatively, orthogonality can still be maintained by modifying the Alamaouti code [LP00]. Nevertheless, the use of an equalization algorithm like MLSE is still necessary. Furthermore, there are additional errors at the edges of the transmission blocks.

The Alamouti code is originally designed for complex-valued bipolar data symbols. In molecular communication, however, the data is typically represented by the amount of emitted molecules. Thus, the data symbols are unipolar (non-negative) and real-valued. The same restrictions apply to optical communication in conjunction with direct detection. In [SV05], it is described how the conventional Alamouti code can be modified to an Alamouti-type code which avoids data symbols with negative signs and complex conjugations. The adaptation to an Alamouti-type code can be done in two steps:

1. Since only real-valued symbols are considered, the operation of complex conjugation can be dropped.

2. Data symbols with negative signs are replaced by their inverse equivalent.

The modified transmission matrix of the Alamouti-type code applied to OOK thus results in [DY+17; DY+18]

$$\boldsymbol{T} = \begin{bmatrix} a_k & a_{k+1} \\ N - a_{k+1} & a_k \end{bmatrix}, \qquad a_k \in \{0, N\}. \tag{4.22}$$

In [SV05], the maximum-likelihood detection metric is derived and it is shown that the Alamouti-type code has the same desirable orthogonal properties as the conventional Alamouti code. Since in molecular communication the signal is typically affected by ISI, orthogonality is no longer maintained. Thus, as in the case of conventional Alamaouti code with ISI, a more complex detection algorithm such as MLSE is required.

4.5.3 Repetiton MIMO Code

In [WB+05], repetition MIMO coding is proposed as a simple and intuitive alternative to the orthogonal Alamaouti code in the area of optical communication. In contrast to the Alamouti code, repetition MIMO coding uses only the spatial dimension but not the temporal dimension. At least in visible light communication, repetition should occur in the spatial domain rather than in time domain [Hoe19]. As the name suggests, exactly the same information is transmitted over each transmit antenna at the same time. The transmission matrix for a 2×2 MIMO system is therefore [DY+17; DY+18]

$$\boldsymbol{T} = \begin{bmatrix} a_k & a_k \end{bmatrix}. \tag{4.23}$$

Since the code offers no orthogonality, the ILI cannot be erased. However, since the same information is transmitted via each transmit antenna, the ILI has a constructive influence on the received signal strength. Another advantage is that even if ISI is present, SISO detection algorithms can be directly used.

4.5.4 Receiver Combining Techniques

If more than a single receive antenna is used in the transmission system, several receive signals are available for detection. From these signals, either suitable signals must be selected or several signals must be combined before detection is performed. The most common combination and selection strategies are SD, EGC and MRC [Höh11], which are considered throughout this dissertation for a symmetrical 2×2 MIMO system.

In SD, the signal of the receive antenna that provides the strongest receive signal is selected. Considering the received signal for OOK modulated signals in the diffusion-based channel affected by ISI, it is difficult to define the strongest received signal. If the symbol $x[k] = 0$ is transmitted, $y[k] = 0$ would represent the strongest received signal. In contrast, if $x[k] = N$ is transmitted, the received signal becomes stronger the more molecules reach the receiver. To avoid this discrepancy, the channel coefficients h_ℓ^{ji} are used as a measure for the received signal strength. The higher the channel coefficient, the stronger the received signal. Since a symmetrical 2×2 MIMO scenario is assumed throughout this dissertation, exactly the same channel conditions are present at the two receive antennas Rx_0 and Rx_1. Therefore, without loss of generality, for SD the signal of Rx_0 is selected in the scenario under consideration [DY+18]:

$$y_{\mathrm{SD}}[k] = y_0[k] = \sum_{\ell=0}^{L} \left(h_\ell^{00} x_0[k-\ell] + h_\ell^{01} x_1[k-\ell] \right) + \omega_0[k]. \tag{4.24}$$

In EGC, not only a single received signal is selected, but all received signals are weighted and combined with each other. The signals are all equally weighted before they are added together, which leads to a very low-complex combining strategy. Without loss of generality, the weights are chosen equal to one [DY+17; DY+18], so that

$$\begin{aligned} y_{\mathrm{EGC}}[k] &= y_0[k] + y_1[k] \\ &= \sum_{\ell=0}^{L} \left[\left(h_\ell^{00} + h_\ell^{10} \right) x_0[k-\ell] + \left(h_\ell^{01} + h_\ell^{11} \right) x_1[k-\ell] \right] + \omega_0[k] + \omega_1[k]. \end{aligned} \tag{4.25}$$

If the weights are not the same but selected with respect to the channel quality, this is referred to as MRC. Received signals from a better channel (larger channel coefficients) are weighted more strongly than signals from a worse channel. MRC and a subsequent detection leads to the same result of a maximum-likelihood receiver. The big disadvantage of MRC compared to EGC is that channel knowledge at the receiver is absolutely necessary. Since a symmetrical 2×2 MIMO scenario is considered here, the channel conditions at the

receiving antennas Rx_0 and Rx_1 are exactly the same. Consequently, both received signals are weighted equally in terms of MRC. Thus, in this symmetrical special case, the solution from EGC is equal to the solution from MRC, which is why only EGC is considered in the remaining. Furthermore, since $h_\ell^{00} = h_\ell^{11}$ and $h_\ell^{01} = h_\ell^{10}$, (4.25) can be simplified as [DY+17; DY+18]

$$y_{\mathrm{EGC}}[k] = \sum_{\ell=0}^{L} h'_\ell \left(x_0[k-\ell] + x_1[k-\ell]\right) + \omega'[k], \tag{4.26}$$

where $h'_\ell \doteq h_\ell^{00} + h_\ell^{10} = h_\ell^{01} + h_\ell^{11}$ and $\omega'[k] \doteq \omega_0[k] + \omega_1[k]$.

4.5.5 Detection

The originally transmitted information can be detected from the combined/selected signals. Since ISI is present in the 2×2 MIMO scenario under consideration, the detection using the MLSE introduced in Section 3.2.6 is considered. For numerical results and implementation of FTD and ATD, the reader is referred to [DY+18]. In contrast to the MLSE presented in Section 3.2.6, the suboptimal Euclidean branch metric [DY+17; DY+18]

$$\lambda_{\boldsymbol{s}_i,\boldsymbol{s}_j}[k] = \left| y[k] - \sum_{\ell=0}^{L} h_\ell \tilde{x}[k-\ell] \right|^2 \tag{4.27}$$

is used for the sake of simplicity. Its performance deviates only slightly from the optimal branch metric from (3.69) and (3.72) [DK+17a]. If the repetition MIMO code is considered, the branch metric for SD is

$$\lambda^{\mathrm{R,SD}}_{\boldsymbol{s}_i,\boldsymbol{s}_j}[k] = \left| y_{\mathrm{SD}}[k] - \sum_{\ell=0}^{L} h'_\ell \tilde{a}_{k-\ell} \right|^2, \tag{4.28}$$

and for EGC the branch metric is

$$\lambda^{\mathrm{R,EGC}}_{\boldsymbol{s}_i,\boldsymbol{s}_j}[k] = \left| y_{\mathrm{EGC}}[k] - \sum_{\ell=0}^{L} 2h'_\ell \tilde{a}_{k-\ell} \right|^2, \tag{4.29}$$

where $[\tilde{a}_k \ \ldots \ \tilde{a}_{k-L}]$ is the vector of symbol hypotheses that corresponds to a transition from state $\boldsymbol{s}_i$ to state $\boldsymbol{s}_j$ [DY+17; DY+18].

The Alamouti-type code spreads the information of two symbols according to (4.22) over two time slots. Accordingly, the branch metric should be evaluated jointly over these two time slots [DY+17; DY+18]. The adaptation of the branch metric for joint evaluation is illustrated in the following by an example with $L = 1$. If ISI is directly taken into account

in the transmission matrix, this results in

$$\boldsymbol{T}_1 = \begin{bmatrix} a_k & a_{k+1} & N - a_{k-1} & a_{k-2} \\ N - a_{k+1} & a_k & a_k & a_{k+1} \end{bmatrix}, \tag{4.30}$$

where $\boldsymbol{T}_L$ is the transmission matrix including L ISI terms. The last two columns in (4.30) belong to the transmission of the previous time slots by Tx_0 and Tx_1 and thus represent the ISI. The number of received molecules at Rx_0 and Rx_1 can be written as

$$\begin{bmatrix} y_0[k] & y_1[k] \\ y_0[k+1] & y_1[k+1] \end{bmatrix} = \boldsymbol{T}_1 \begin{bmatrix} h_0^{00} & h_0^{10} \\ h_0^{01} & h_0^{11} \\ h_1^{00} & h_1^{10} \\ h_1^{01} & h_1^{11} \end{bmatrix} + \begin{bmatrix} \omega_0[k] & \omega_1[k] \\ \omega_0[k+1] & \omega_1[k+1] \end{bmatrix}. \tag{4.31}$$

If EGC is assumed at the receiver side, the branch metric for the Alamouti-type code for the symmetrical 2×2 MIMO scenario and $L = 1$ results in

$$\begin{aligned} \lambda_{\boldsymbol{s}_i,\boldsymbol{s}_j}^{\mathrm{A,EGC}}[k] = &\left| y_{\mathrm{EGC}}[k] - h_0' \left(\tilde{a}_k + \tilde{a}_{k+1}\right) - h_1' \left(\tilde{a}_{k-2} - \tilde{a}_{k-1} + N\right) \right|^2 \\ &+ \left| y_{\mathrm{EGC}}[k+1] - h_0' \left(\tilde{a}_k - \tilde{a}_{k+1} + N\right) - h_1' \left(\tilde{a}_k + \tilde{a}_{k+1}\right) \right|^2. \end{aligned} \tag{4.32}$$

4.5.6 Numerical Results

This section presents numerical results for the assumed 2×2 MIMO scenario [DY+17; DY+18]. To analyze the effect of spatial coding and combining techniques, BER simulations are performed. The transmission channel is modelled as a Poisson binomial channel according to Section 2.3.1. The required channel coefficients are obtained by a trained ANN. In order to be able to evaluate the gain of the individual MIMO methods, comparisons to SISO transmission systems are made. The simulation parameters, in case they are not varied, are summarized in Section C.15. For the SISO comparisons, only a single transmitter and receiver are assumed. In the SISO case, twice as many molecules are released as in the MIMO case at the transmitter. This guarantees a fair comparison between the SISO and 2×2 MIMO scenario in the sense of the released number of molecules per bit, which is interpreted as transmission energy.

Acquiring Channel Coefficients

When a second absorbing spherical receiver is introduced into the system, the probability of a molecule being absorbed no longer depends only on the radial component of the spherical coordinate system. Consequently, the symmetry exploited in the derivation of Section 2.2.2 is no longer valid and there exists no analytical closed-form solution of Fick's second law (2.2). In this case, the channel coefficients h_k^{ji} can be estimated with the help of Monte Carlo simulation of molecule's random walks according to Section 2.2.1, or with the help of a trained ANN according to Section 2.2.3. The latter offers after training the advantage of being able to estimate the channel coefficients for arbitrary system parameter without having to perform new complex random walk based Monte Carlo simulations. Therefore it is used throughout this analysis.

The structure of the ANN [DY+17; DY+18] is inspired by [LY+17]. Due to the symmetry of the 2×2 MIMO scenario under consideration, $h_\ell^{00} = h_\ell^{11}$ and $h_\ell^{10} = h_\ell^{01}$ apply. Thus it is sufficient to construct an ANN for a scenario with only a single active transmit antenna. Without loss of generality, Tx_0 is selected. According to (2.40), two model functions are constructed, each with three fitting parameters. One model function describes the direct link h_ℓ^{00} from Tx_0 to Rx_0 and the other the ILI h_ℓ^{10} from Tx_0 to Rx_1. The schematic structure of the ANN is shown in Fig. 2.2. The input layer consists of four nodes, where each node stands for one of the system parameters d, d_s, r, and D. The output layer consists of six nodes, each node representing one of the 2×3 fitting parameters. The hidden layer contains 40 nodes, which heuristically represents a good compromise between accuracy and complexity of the ANN. Random walk based Monte Carlo simulations are performed and their results are fitted to the two model functions. The resulting fitting parameters serve as a reference for training the ANN as described in Section 2.2.3. A total of 144 test cases

Table 4.10: ANN training parameters.

Parameter		Unit	Value
Diffusion coefficient	D	$\mu\mathrm{m}^2$ / s	$\{100, 200, 300\}$
Receiver radius	r	µm	5
Transmission distance	d	µm	$\{10, 15, 20, 25\}$
Receiver separation distance	d_s	µm	$\{11, 13, 15\}$
Number of molecules	N		3000
Channel realizations	Q		400
Simulation timestep	Δt	µs	100
Simulation time duration	T_end	s	2

are determined for training, resulting from the combination of all system parameters in

Tab. 4.10.

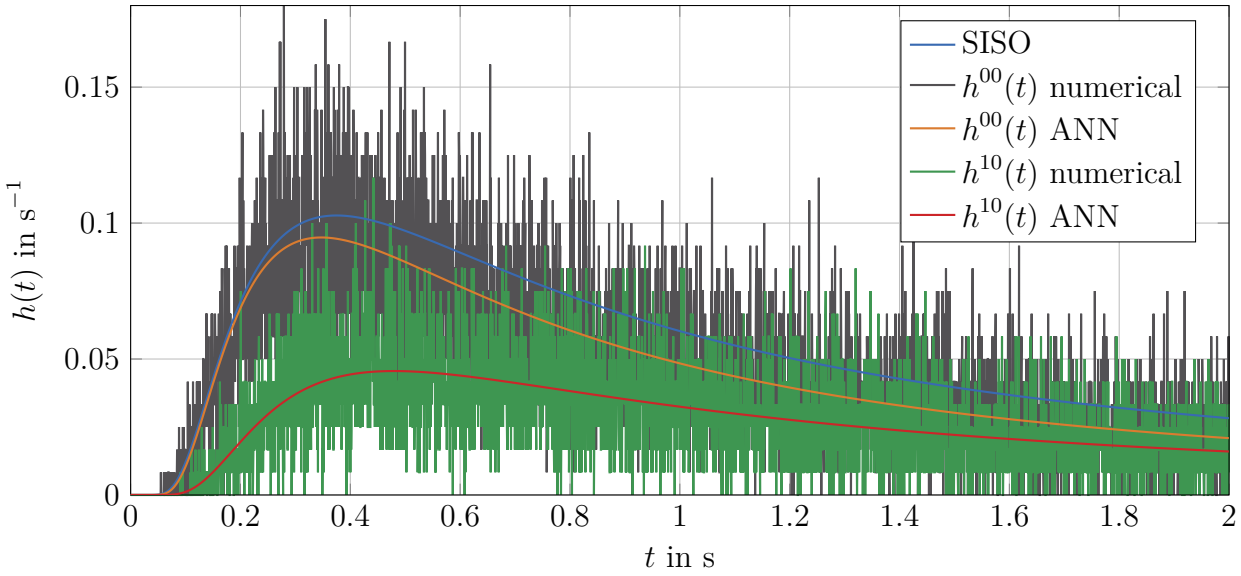


Figure 4.24: Comparison between the channel impulse responses acquired from random walk based Monte Carlo simulations and from ANN for system parameters given in Tab. 4.10 and $D = 100\,\mu\text{m}^2/\text{s}$, $d = 20\,\mu\text{m}$, and $d_\text{s} = 11\,\mu\text{m}$. As a comparison, the channel impulse response of the corresponding SISO scenario is added. Plotting time resolution is set to $500\,\mu\text{s}$.

Fig. 4.24 shows a comparison between the channel impulse responses determined by random walk based Monte Carlo simulations and by the trained ANN. Both the direct link $h^{00}(t)$ and the ILI $h^{10}(t)$ are shown. In addition, the analytical channel impulse response according to (2.37) for a SISO scenario with the parameters of the direct link is depicted. Fig. 4.25 shows the same comparison for the channel coefficients. As expected, the channel impulse response of the direct link is below the channel impulse response of the SISO scenario. With the second absorbing receiver, fewer molecules are absorbed by Rx_0, since a part is already absorbed by Rx_1. The channel impulse response of the ILI is lower than that of the direct link, since the second absorbing receiver Rx_1 with $\sqrt{d^2 + d_\text{s}^2} > d$ has a larger distance to Tx_0. The same observations also apply to the channel coefficients. Both Fig. 4.24 and Fig. 4.25 show that the results determined by the ANN agree well with the random walk based Monte Carlo simulations, which justifies the use of a trained ANN to estimate the channel coefficients.

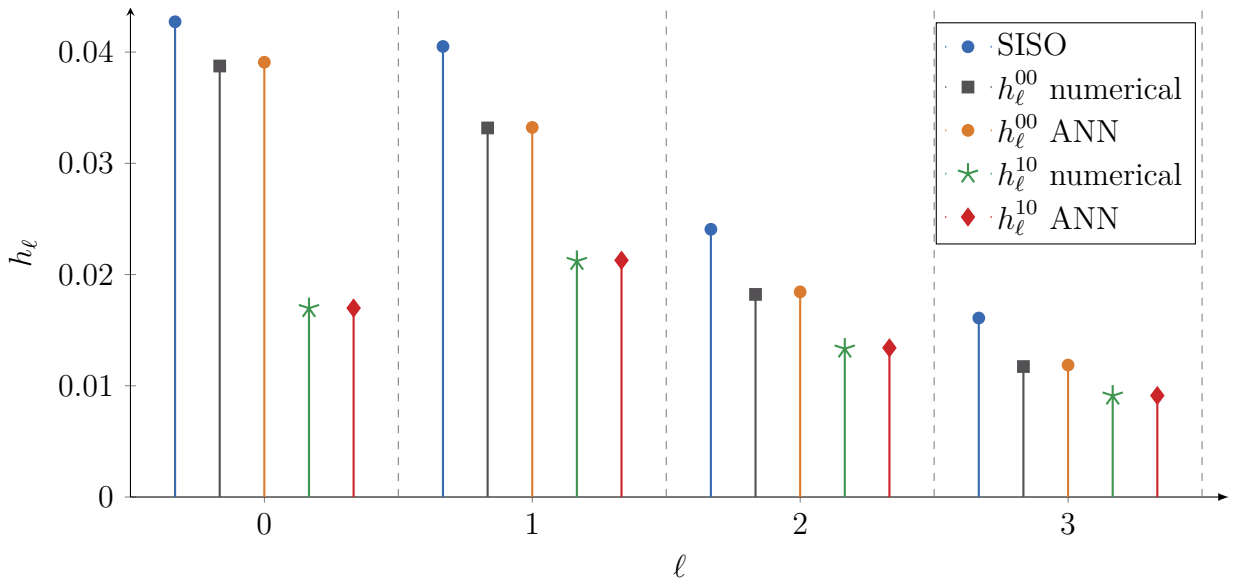


Figure 4.25: Comparison between the channel coefficients acquired from random walk based Monte Carlo simulations and from ANN for system parameters given in Tab. 4.10 and $D = 100\,\mu\text{m}^2/\text{s}$, $d = 20\,\mu\text{m}$, $d_\text{s} = 11\,\mu\text{m}$, $T = 0.6\,\text{s}$, and $L = 3$. As a comparison, the channel coefficients of the corresponding SISO scenario are added.

Spatial Diversity Analysis

In classical wireless communication, spatial diversity is exploited to combat the effect of fading in transmission channels. The resulting diversity order Γ is commonly defined as the asymptotic slope of the BER curve plotted over the signal-to-noise ratio (SNR): $P_\text{b} \propto (1/\text{SNR})^\Gamma$ for $\text{SNR} \to \infty$. Due to the amplitude-dependent diffusion noise in diffusion-based molecular communication, it is not as straightforward to define an SNR as it is in classical wireless communication. The amount of released molecules N, however, is proportional to the SNR [DK+17a]. To obtain logarithmic values which can be interpreted similar to SNR in dB, in Fig. 4.26 the BER is plotted over $10\log_{10}(N)$.

Fig. 4.26 compares the SISO system (with and without power normalization) to the repetition MIMO system with EGC. The focus here is on repetition MIMO with EGC only, since it outperforms the other techniques under investigation as it will be shown in Fig 4.27. When power normalization is not applied, there is the typical gain of about 3 dB (a factor of two in linear scale) between the SISO system and the 2×2 MIMO system with separation distance $d_\text{s} = \infty$. This gain is due to an array gain, which occurs when multiple transmit/receive antennas are used. The transmit power and receive area is doubled, which is equivalent to the use of two transmit and receiving antennas in radio-based wireless

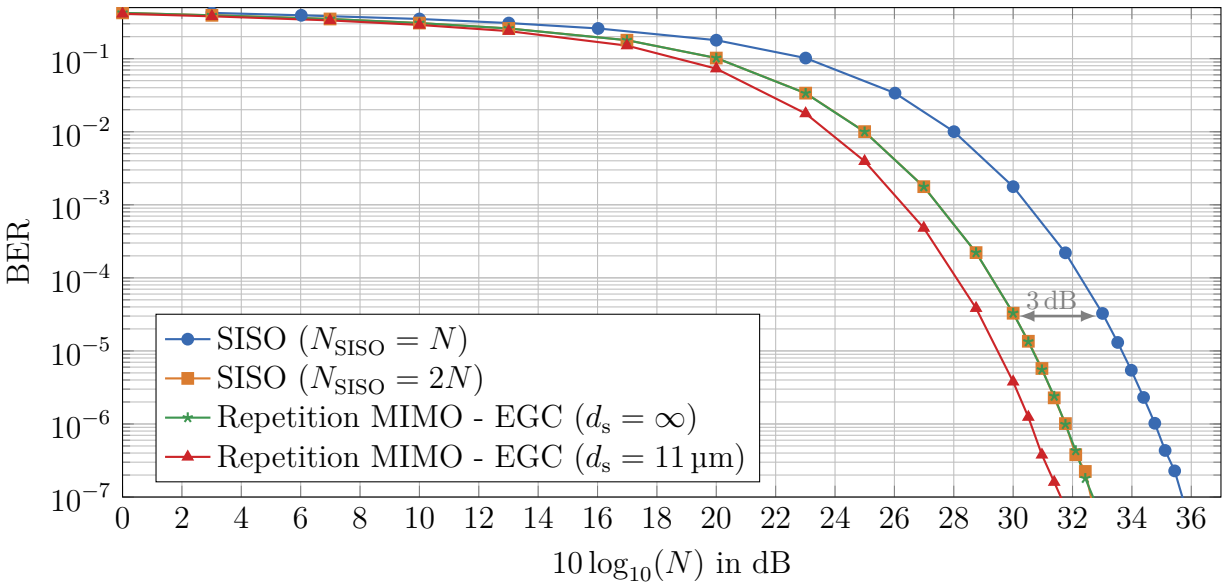


Figure 4.26: Bit error rate performance as a function of the number of molecules. Parameters according to Appendix C.15.

transmissions. It will disappear when power normalization is applied to the SISO scenario as shown in Fig. 4.26. At separation distance $d_{\text{s}} = 11\,\mu\text{m}$, the array gain is even larger than 3 dB. The additional gain comes from the increasing total energy of the channel coefficients. If $d_{\text{s}} \to \infty$, there is no ILI and the channel energy of each direct link is equal to the channel energy of the SISO link. If $d_{\text{s}} = 11\,\mu\text{m}$, however, the channel energy of the direct links is decreasing (due to the second absorbing sphere) while the ILI channels are increasing. The total channel energy of all links will be larger than in the $d_{\text{s}} \to \infty$ case. In order to maximize the gain in the scenario under investigation, the separation distance should be as short as possible.

From Fig. 4.26 it can be observed that the asymptotic slope of the BER curve (i. e., the diversity order) is not changing when it comes to MIMO systems. Consequently, there is no fading or fading-like effect in the standard diffusion-based channel, at least not for the system under investigation. Fading might occur in scenarios, where an obstacle is placed between the transmitter and the receiver. Conclusively, all the performance gains that are achieved by applying spatial diversity algorithms throughout this dissertation are due to an array gain and not due to a spatial diversity gain.

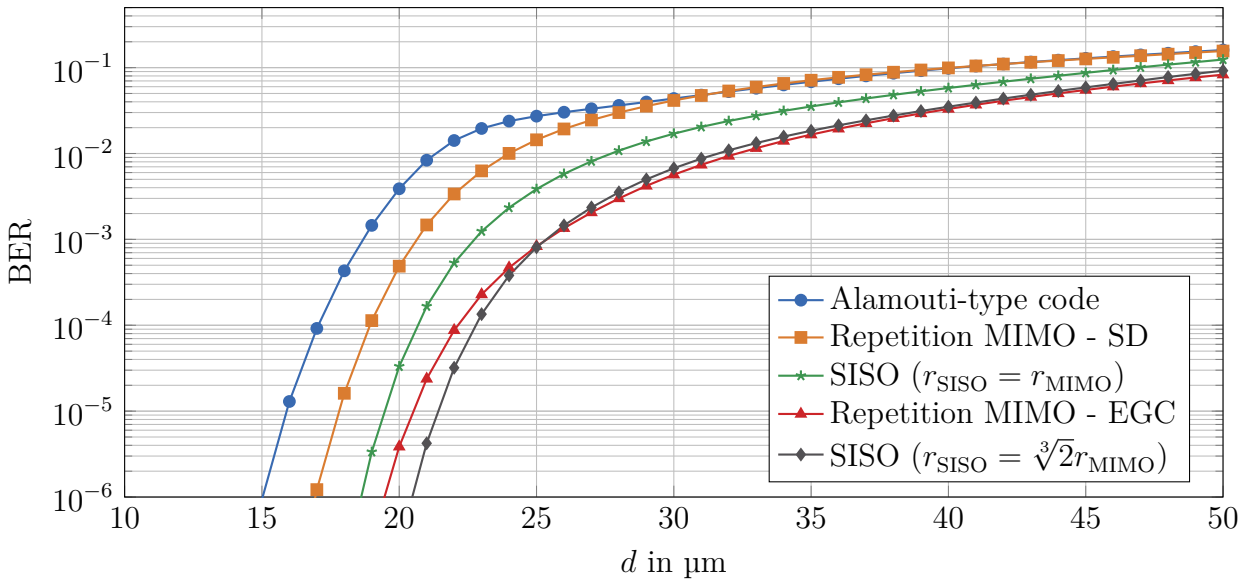


Figure 4.27: Bit error rate performance as a function of transmission distance. Parameters according to Appendix C.15.

Effect of Transmission Distance

Fig. 4.27 shows the effect of the transmission distance on the system performance. In general, a shorter transmission distance results in more molecules being absorbed by the receiver. This increases the signal strength at the receiver. If the symbol duration is fixed, the ISI is also reduced by a shorter transmission distance. As a result, the BER for all methods decreases as the transmission distance decreases. For repetition MIMO with EGC, an array gain can be achieved with respect to the SISO case. The maximum BER improvement is by a factor of approximately 10. Repetition MIMO in conjunction with SD leads to no gain compared to the SISO case, which is mainly due to the assumed power normalization.

Interestingly, the Alamouti-type code does not lead to an array gain for the system under investigation. This can be explained with the ILI present in the system. While ILI in repetition MIMO contributes constructively to the signal strength, ILI in Alamouti-type code is competitive and therefore rather destructive. The Alamouti-type code does not work well for channels with unipolar properties as shown in [SU08] for the free-space optical channel. The main degradation factor in diffusion-based channels is in the ISI and ILI.

To have a fair comparison between the SISO and MIMO systems, the effective receiver volume can be normalized as well. Without this normalization, the radii of the receiving spheres in the SISO and MIMO cases are the same $r = r_{\text{MIMO}} = r_{\text{SISO}}$. Thus the effective

receiver volume is larger than in the SISO case, since two absorbing spheres are present in the environment. In order to normalize the effective receiver volume, the radius of the SISO receiver is normalized to $r_{\mathrm{SISO}} = \sqrt[3]{2}r$. Applying this normalization, all MIMO systems perform worse than the SISO system in the low BER region. Even for repetition MIMO with EGC, no array gain can be observed. Regarding a future implementation in autonomous nanomachines, it might, however, be more practicable to realize several small nanomachines instead of a single large nanomachine.

In conclusion, it can be stated that for the 2×2 MIMO scenario repetition MIMO showed the best performance and that the Alamouti-type code should be avoided in the diffusion-based channel. The focus in this analysis is on a varied transmission distance. The influence of other system parameters, however, leads to comparable results [DY+18].

5

Turbo Principle

In this chapter, joint channel equalization/detection and decoding is considered, which provides a joint analysis of the equalization/detection discussed in Chapter 3 and the channel codes covered in Chapter 4. In the sense of the turbo principle, extrinsic information is exchanged between the two components. By means of iterative processing the overall performance can be improved. With the help of an extrinsic information transfer (EXIT) chart analysis, it is demonstrated how the performance of this iterative processing can be predicted. In addition to low-density parity check (LDPC) codes from Section 4.1.6, a designed irregular convolutional code (IRCC) based on the subcodes in Section 4.2 is considered as channel code.

Furthermore, the EXIT chart analysis is used to estimate the communication channel capacity of molecular communication systems. While a general information theoretical channel capacity expression in diffusion-based molecular communication is unsolved [NE+13], the proposed method offers the possibility to calculate maximum performance limits of investigated transmission systems. Using these performance limits, the quality of transmission algorithms can be evaluated and potential code rates to achieve the communication channel capacity can be determined. Thus, it represents a very powerful tool for the analysis of molecular communication systems. The content of this chapter is mainly based on [DH18; DS+19].

5.1 Bit-Interleaved Coded Modulation

In order to jointly consider equalization and decoding, the bit-interleaved coded modulation (BICM) [CT+98] structure depicted in Fig. 5.1 is applied. The idea is to concatenate

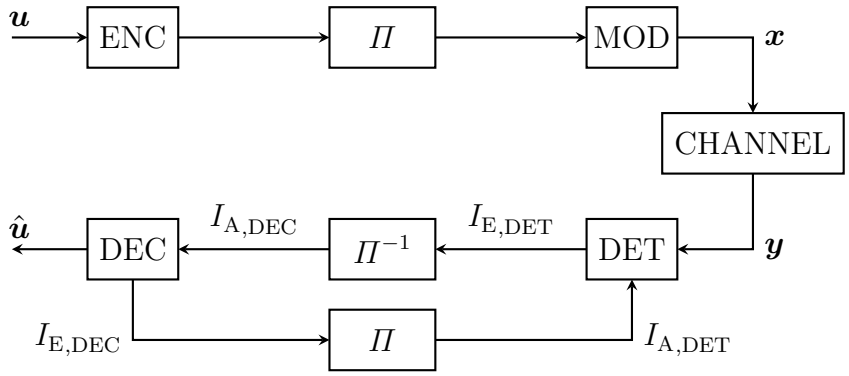


Figure 5.1: Block diagram of a bit-interleaved coded modulation turbo structure with iterative detection.

the channel encoder and the modulator with a bitwise interleaver in between. At the receiver side, the detector compensates the effect of intersymbol interference (ISI). After the deinterleaver, the channel decoder is used to eliminate residual errors. In addition to single detection and decoding, iterative detection and decoding can be performed. This is referred to as BICM with iterative detection (BICM-ID) [tBS+98]. The detector and decoder exchange extrinsic information I_{E} in the sense of the turbo principle in order to iteratively improve their detection/decoding results. For this purpose, a feedback channel from the decoder to the detector is realized with an interleaver in between. The extrinsic information $I_{\mathrm{E,DET}}$ of the detector is used as *a priori* information $I_{\mathrm{A,DEC}}$ in the decoder, while the extrinsic information $I_{\mathrm{E,DEC}}$ of the decoder is used as *a priori* information $I_{\mathrm{A,DET}}$ in the detector. Fig. 5.1 shows a serially concatenated turbo structure consisting of a modulator/detector as the inner component and a channel encoder/decoder as the outer component. If the components are matched to each other, the iterative detection and decoding leads to an information gain in each iteration and consequently to a significant increase of the bit error rate (BER) performance of the system. For iterative detection, soft-input soft-output decoding or soft-input soft-output detection is performed in the components. These provide extrinsic log-likelihood ratio (LLR) values L_{E} which serve as *a priori* LLR values L_{A} in the concatenated component. Within this context, it is important that these are "true" LLR values. If approximated LLR values are used, this leads to an over- or underestimation of the detection or decoding performance and consequently reduces the validity. It should be noted that BICM is typically considered in conjunction with higher-order modulation schemes. In this chapter, however, only on-off keying (OOK) is considered as modulation, but the general structure of BICM-ID is adopted. An analysis

that also considers higher-order modulation schemes can be found in [BD+19]. For detection, the optimal *a posteriori* probability (APP) detector based on the Bahl-Cocke-Jelinek-Raviv (BCJR) algorithm from Section 3.2.8 is applied. The interleaver is assumed as a random interleaver.

5.2 Extrinsic Information Transfer Chart

By utilizing the EXIT chart analysis [tBri99], it is possible to predict the extrinsic information exchange between concatenated component decoders/detectors. Consequently, the EXIT chart provides information about the convergence of a turbo structure. For this purpose, the extrinsic information I_{E} is generated separately for each component given a certain amount of *a priori* information I_{A}. *A priori* LLRs are delivered as inputs to

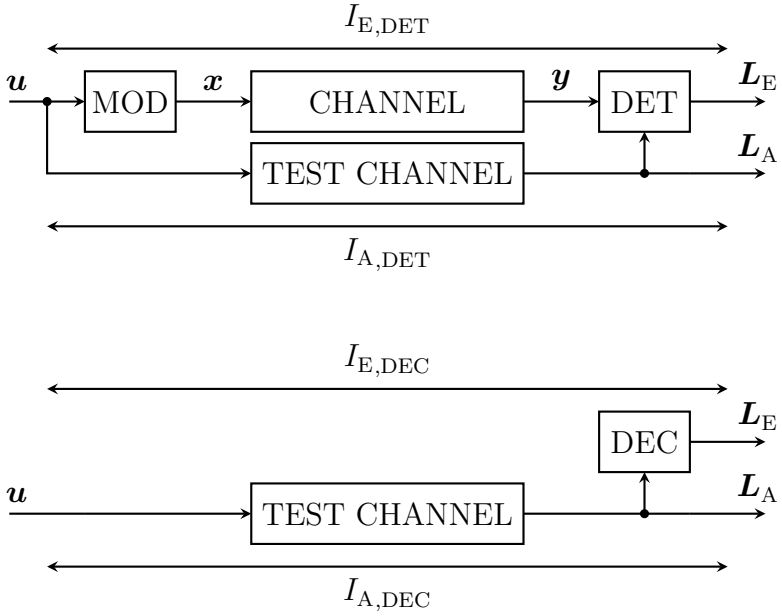


Figure 5.2: Setup for the EXIT chart generation of a serially concatenated bit-interleaved coded modulation system.

the component decoder/detector, see Fig. 5.2. Usually, the generation is emulated by an additive white Gaussian noise test channel [tBri01b], a binary erasure test channel, or a binary symmetric test channel [Hag04]. The extrinsic LLRs provided from the output of the components are used to determine the extrinsic information I_{E}. For instance, this can be realized by estimating the conditional probability density function and numerically integrating [tBri01b], by exploiting the ergodicity theorem [TH02], or by averaging over a

function of the absolute LLRs [LH+04]. In the EXIT chart, the extrinsic information I_E is depicted over the *a priori* information I_A, separately for each component detector/decoder, see Fig. 5.5 and Fig. 5.6. In order to analyze the information exchange, the *a priori* information axis of the first component is equal to the extrinsic information axis of the second component and vice versa.

In Fig. 5.5, the first component is the detector and the second component the channel decoder (see Fig. 5.1). If there is an open tunnel in between the two EXIT functions (i. e., the EXIT function of the inner component is above the EXIT function of the outer component without intersecting), it is possible to predict the iterative information exchange between the components. In conclusion, EXIT charts can be used to estimate the behavior/performance of turbo systems and to optimize component codes.

5.2.1 Bit Error Rate Prediction

With help of the EXIT chart, it is possible to approximate the residual BER (i. e., the BER after iterative decoding), when there is no open tunnel or when the number of iterations between the components is not sufficient to iterate through a tunnel. Assuming Gaussian distributed statistical independent LLRs, the BER can be approximated as

$$P_b \approx \frac{1}{2}\mathrm{erfc}\sqrt{\frac{\sigma_A^2 + \sigma_E^2}{8}}, \tag{5.1}$$

where σ_A^2 and σ_E^2 are the variances of the *a priori* and extrinsic LLRs, respectively [tBri01a]. The relation between I_A/I_E and σ_A^2/σ_E^2 is given by the $I = J(\sigma)$ function. This function can be approximated as [tBK+04]

$$J(\sigma) \approx \begin{cases} -0.0421061\sigma^3 + 0.209252\sigma^2 - 0.00640081\sigma & \text{if } 0 \le \sigma \le 1.6363 \\ 1 - \exp(0.00181491\sigma^3 - 0.142675\sigma^2 - 0.0822054\sigma + 0.0549608) & \text{if } 1.6363 < \sigma < 10 \\ 1 & \text{if } \sigma \ge 10 \end{cases}.$$

$$J^{-1}(I) \approx \begin{cases} 1.09542I^2 + 0.214217I + 2.33727\sqrt{I} & \text{if } I \le 0.3646 \\ -0.706692\log(0.386013(1-I)) + 1.75017I & \text{if } 0.3646 < I < 1 \end{cases} \tag{5.2}$$

5.2.2 Communication Channel Capacity

With EXIT charts, it is possible to estimate the communication channel capacity C for the system under investigation. The communication channel capacity includes the channel, modulation scheme, and input distributions of the information bits [AK+04]. The area

under the inner decoder in the EXIT chart is bounded by

$$A_{\mathrm{i}} \leq \frac{C}{mR_{\mathrm{i}}}, \tag{5.3}$$

where $m = \log_2(M)$ with M being the modulation order of the investigated modulation scheme and R_{i} denoting the code rate of the inner component. Equality is fulfilled for $R_{\mathrm{i}} = 1$ and (5.3) has been proven for the memoryless binary erasure channel (BEC) [AK+04]. For other channels, (5.3) can at least be used as a good approximation [Höh11]. Consequently, analyzing the area under the EXIT chart of an optimal detector (with rate $R_{\mathrm{i}} = 1$) will approximately lead to the communication channel capacity including the diffusion-based molecular communication channel. Following Shannon's capacity theorem [Sha48], the rate of the outer decoder should be

$$R \leq \frac{C}{m}. \tag{5.4}$$

In the EXIT chart, the area under the outer decoder curve is equal to the code rate R. As a result, the tunnel area (i. e., the area between the inner detector and outer decoder EXIT functions) represents a rate loss and should be as small as possible without any intersections [AK+04].

5.3 EXIT Chart Aided Code Design

In order to minimize rate loss and to maximize iterative detection (ID) gain, the EXIT functions of the decoder and the detector must match well. While an adjustment can be done by variations of modulation and doping [PS06] for a given channel code, the focus in this dissertation is on the matching of the decoder EXIT function to a given detector EXIT function [DH18]. Torwards this goal, an IRCC matching to the EXIT function of the detector is constructed given a family of $V = 21$ convolutional subcodes varying from code rate $R_5 = 1/10$ to $R_{21} = 9/10$ with step size $1/20$ as suggested in [Tüc04]. These codes correspond to the convolutional codes considered in Section 4.2. In order to achieve even lower code rates, and hence larger distances, new subcodes are generated by repetition of the $R_5 = 1/10$ subcode. The new subcodes have the rate $R_1 = 1/100$, $R_2 = 1/50$, $R_3 = 1/40$, and $R_4 = 1/20$, respectively.

By means of IRCC matching, the info/code word is partially encoded/decoded by the subcodes. The combination of all code words forms the final IRCC code word. The set of weighting factors $\psi_v \in [0, 1]$ determines the amount of code bits taken from the vth

subcode in the IRCC code word, subject to the constraint

$$\sum_{v=1}^{V} \psi_v = 1. \tag{5.5}$$

Consequently, the average code rate R of the IRCC is

$$R = \sum_{v=1}^{V} \psi_v R_v, \tag{5.6}$$

where R_v is the code rate of the vth subcode. By combining subcodes, almost arbitrary average code rates can be achieved. In addition, the shape of the EXIT function of the decoder can be designed, since it is a weighted combination of the EXIT functions of the subcodes:

$$I_{\mathrm{E,DEC}}(I_{\mathrm{A,DEC}}) = \sum_{v=1}^{V} \psi_v I_{\mathrm{E,DEC},v}(I_{\mathrm{A,DEC}}). \tag{5.7}$$

The weighting factors $\boldsymbol{\psi} = [\psi_1 \ldots \psi_V]$ can be chosen with respect to different cost functions such as maximizing the mutual information. In this dissertation, the optimization criterion is to maximize the average code rate

$$\boldsymbol{\psi}_{\mathrm{opt}} = \arg\max_{\forall \boldsymbol{\psi} \in \mathcal{A}} \sum_{v=1}^{V} \psi_v R_v, \tag{5.8}$$

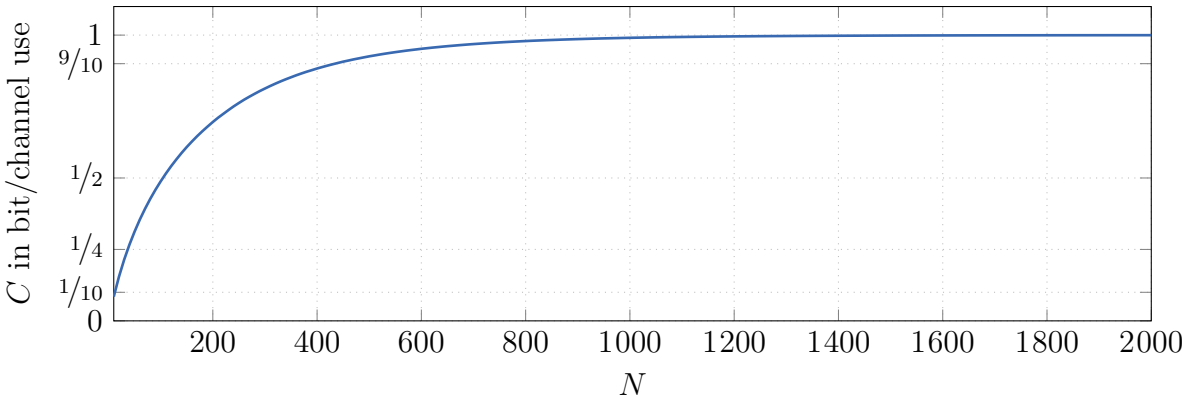
where $\mathcal{A}$ is the set of all $\boldsymbol{\psi}$ that fulfill constraint (5.5) and lead to an open tunnel, i. e., the resulting EXIT function of the decoder lies below the EXIT function of the detector.

5.4 Numerical Results

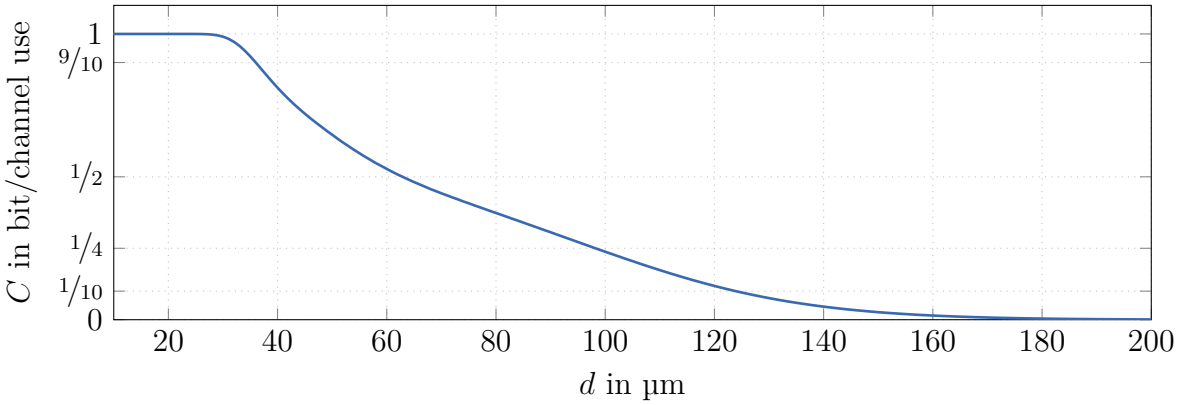
A BEC test channel is used to generate the EXIT charts during the numerical analyses. The extrinsic information is determined by averaging over a function of the absolute LLRs [LH+04].

5.4.1 Communication Channel Capacity Analysis

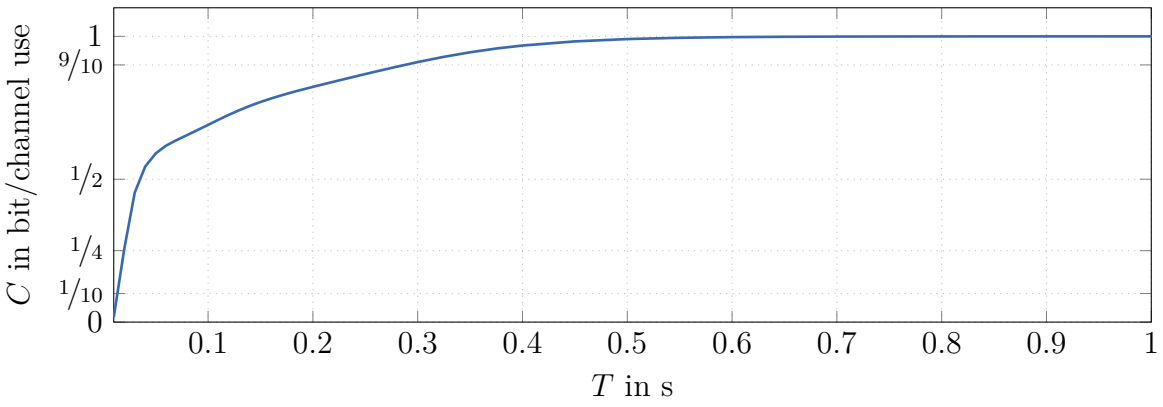
As described in Section 5.2.2, the EXIT chart of the APP detector can be used to estimate the communication channel capacity of the diffusion-based molecular communication channel. Fig. 5.3 shows the resulting communication channel capacity when the number of released molecules, the transmission distance, and the symbol duration is changed. Due



(a) Number of molecules.



(b) Transmission distance.



(c) Symbol duration.

Figure 5.3: Communication channel capacity of the diffusion-based molecular communication channel derived from an EXIT chart analysis. Parameters according to Appendix C.16.

to the focus on OOK with modulation order $M = 2$, the communication channel capacity cannot exceed $C = 1\,\text{bit/channel use}$. As shown in Fig. 5.3(a), when the number of released molecules is increased, the communication channel capacity approaches its maximum. A higher amount of molecules decreases the stochastic behavior in the diffusion-based channel, which is represented by diffusion noise. Assuming equiprobable symbols, the received number of molecules at the receiver has the following statistical moments [DK+17a]:

$$\mu = 0.5N \sum_{\ell=0}^{L} h_\ell, \tag{5.9}$$

$$\sigma^2 = 0.5N \sum_{\ell=0}^{L} (h_\ell - h_\ell^2). \tag{5.10}$$

Since μ can be interpreted as the signal strength and σ as the diffusion noise strength, diffusion noise grows with a factor $1/\sqrt{N}$ compared to the signal strength. This factor explains the shape of the capacity curve in Fig. 5.3(a).

Fig. 5.3(b) shows the communication channel capacity results when the transmission distance is varied. Intuitively, the communication channel capacity reaches its maximum for small distances only. When the distance is increased, the channel coefficients h_ℓ become smaller. According to (5.9) and (5.10), the diffusion noise strength is growing compared to the signal strength. As a result, the communication channel capacity is decreasing for larger distances. In contrast to Fig. 5.3(a), the shape of the capacity curve in Fig. 5.3(b) is non-trivial to describe. While N is only a factor in (5.9) and (5.10), d changes the channel coefficients according to (2.38).

In Fig. 5.3(c), the influence of the symbol duration on the communication channel capacity is shown. An increasing symbol duration is beneficial, because the number of received molecules is accumulated over a longer time period. Additionally, summing over a certain time period leads to an averaging effect. As a result, the channel coefficients get larger and the received samples are less noisy. Consequently, the communication channel capacity reaches its maximum for longer symbol durations and decreases for shorter symbol durations. Similar to changes in the transmission distance, changing the symbol duration will affect the channel coefficients in (5.9) and (5.10) according to (2.38). Thus it is again non-trivial to describe the shape of the capacity curve and it will need further investigations.

Fig. 5.4 shows the communication channel capacity analysis for different code rates under T- and N-normalized conditions introduced in Chapter 4. The results show the theoretically expected performance of the BICM-ID systems. For $R = 9/10$ error-free transmission is possible for $d \leq 34\,\mu\text{m}$. Code rate $R = 1/2$ offers error-free transmission up to $d \leq 39.5\,\mu\text{m}$.

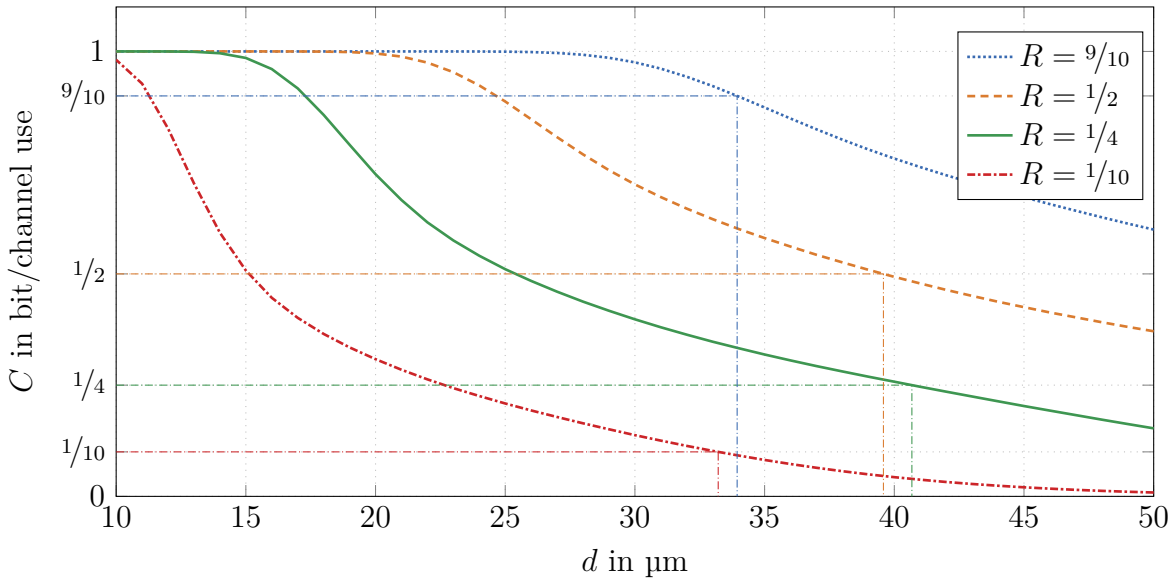


Figure 5.4: Communication channel capacity derived from an EXIT chart analysis over different transmission distances for different code rates, when normalization of information rate and amount of released molecules is applied. Parameters according to Appendix C.17.

The best performance is obtained for $R = 1/4$ with error-free transmission up to $d \leq 40.5\,\mu\text{m}$. From these observations it could be concluded that a lower code rate theoretically improves the BER performance of a BICM-ID system more. However, in the scenario under consideration this seems to be true only up to a certain point. If the communication channel capacity for the code rate $R = 1/10$ is considered, the maximum transmission distance is $d = 33\,\mu\text{m}$ and is therefore lower than for all other considered code rates. The EXIT chart and the communication channel capacity derived from it can be used to select suitable code rates for a system design.

5.4.2 EXIT Chart Analysis

Fig. 5.5 shows the EXIT function of the APP detector for three different transmission distances and the EXIT function of the LDPC codes from the second standard of digital video broadcasting over satellite (DVB-S2) considered in this dissertation in Section 4.1.6. For the LDPC codes, the minimum code rate $R = 1/4$, the maximum code rate $R = 9/10$ and an intermediate code rate $R = 1/2$ are considered. In general, it can be observed that the EXIT function of the APP detector has an increasing characteristic due to the memory affected diffusion-based transmission channel. Thus, the detector in the scenario

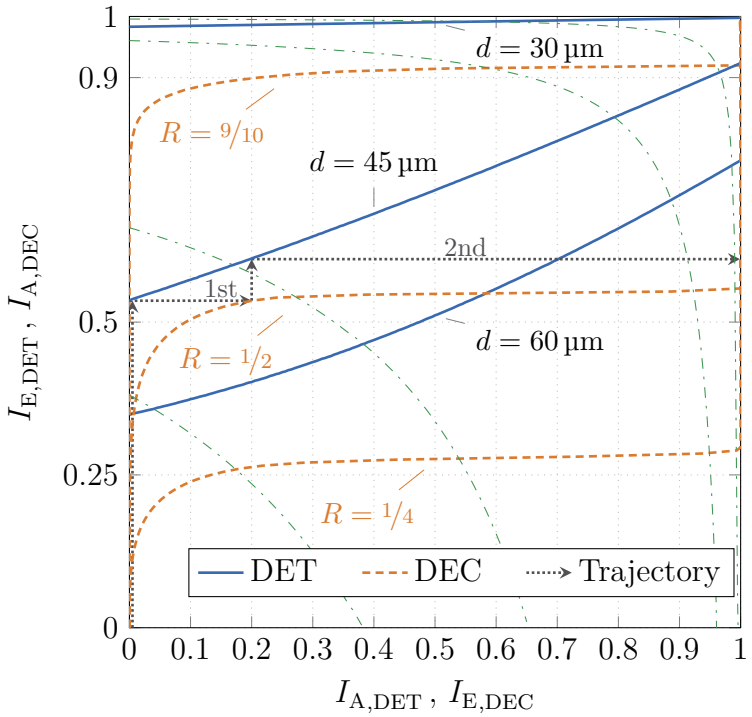


Figure 5.5: EXIT chart for different transmission distances and LDPC codes with different code rates. Parameters according to Appendix C.18.

under investigation is suitable for an application by virtue of the turbo principle, as it shows a potential for performance improvement through ID. Furthermore, the area under the detector EXIT function becomes larger with decreasing transmission distance. This observation is consistent with the communication channel capacity from (5.3) and Fig. 5.3(b). For short transmission distances, the communication channel capacity converges to $C = 1\,\text{bit/channel use}$, which is equivalent to an area $A_\text{i} = 1$ in the EXIT chart. Furthermore, in Fig. 5.5, the approximated BER regions according to (5.1) are represented by dash-dotted lines. From bottom left to top right they represent $P_\text{b} = 0.2$, $P_\text{b} = 10^{-1}$, $P_\text{b} = 10^{-2}$, and $P_\text{b} = 10^{-3}$. With their help the BER of the APP detector without ID, which is equivalent to $I_\text{A,DET} = 0$, can be estimated. For $d = 60\,\mu\text{m}$ $I_\text{E,DET}$ is close to $P_\text{b} = 0.2$ and for $d = 30\,\mu\text{m}$ it is close to $P_\text{b} = 10^{-3}$. Both observations are verified by numerical simulations in Fig. 5.7.

As mentioned in Section 5.2.2, the areas under the decoder EXIT functions in Fig. 5.5 are equal to the code rate R. The EXIT functions of the LDPC codes show a comparatively flat shape over a wide range. If the *a priori* information $I_\text{A,DEC}$ provided by the detector is high enough, the point $I_\text{E,DEC} = 1$ and thus an error-free decoding is achieved even without ID. This is given, for example, for a distance of $d = 30\,\mu\text{m}$ in combination with all investigated LDPC codes. Furthermore, the information gain by ID and the required iterations can be estimated. For illustration purpose, the trajectory of the information exchange between the APP detector and an $R = 1/2$ LDPC code at a distance of $d = 45\,\mu\text{m}$ is shown in Fig. 5.5. In this case, there is an open tunnel and through ID the decoder can converge towards an error-free decoding. Only two iterations are necessary to achieve this. Without ID, which is equivalent to a single iteration, the decoding process ends near the $P_\text{b} = 10^{-1}$ region. Both observations are confirmed by numerical simulations in Fig. 5.7(b). Similarly, the BER can be estimated for closed tunnel constellations where there is an intersection between the detector and decoder EXIT function for $I_\text{A,DET} < 1$.

5.4.3 EXIT Chart Aided Code Design

To demonstrate the construction of an IRCC code, an $R \approx 1/10$ IRCC is matched to the detector EXIT function. Without normalization, the theoretically achievable maximum transmission distance which allows an error-free transmission is $d \approx 123\,\mu\text{m}$, determined by the communication channel capacity (see Fig. 5.3(b)). In the following, the design point for the EXIT chart matching is set to $d = 120\,\mu\text{m}$. Fig. 5.6 shows the EXIT function of the APP detector for the design point $d = 120\,\mu\text{m}$. For the scenario under investigation, the EXIT function has a horizontal flat shape. For IRCC matching described in Section 5.3,

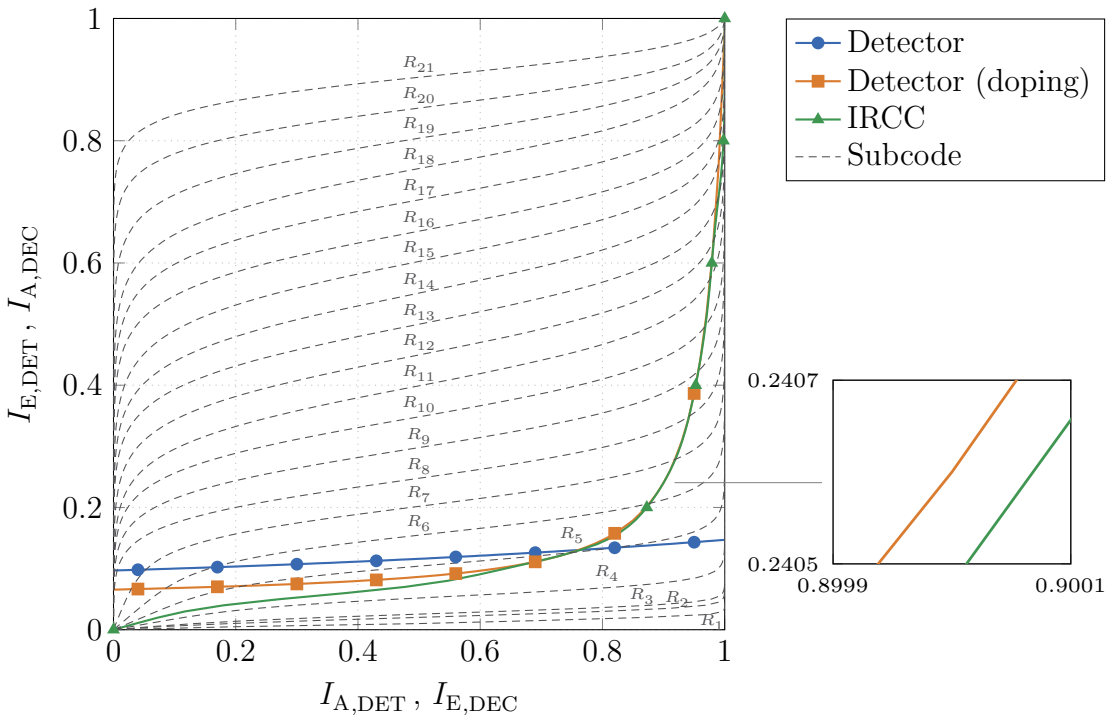


Figure 5.6: EXIT chart with matched IRCC codes for $d = 120\,\mu\mathrm{m}$. Parameters according to Appendix C.19.

however, a monotonically increasing EXIT function is required to reduce the tunnel area and the resulting rate loss. Therefore, code doping [PS06] with a doping periode of two is applied. By means of code doping, each second code bit is replaced by doping bits that are the outputs of an inner recursive encoder with rate $R_\mathrm{i} = 1$. The inner code in this dissertation is an ($L_\mathrm{c} = 1$, $G_\mathrm{fb} = 3$, $G_\mathrm{ff} = 2$) recursive convolutional code, where G_fb is the feedback polynomial and G_ff is the feedforward polynomial in octal notation. The inner recursive encoder will create dependencies between adjacent code bits, which will lead to a rising shape of the EXIT function of the detector as shown in Fig. 5.6. Please note that doping bits could be transmitted by a different type of molecule. Consequently, the effective symbol duration of the doped bits will increase which lowers the ISI degradation. As a result, doped bits are less prone to errors, which will further improve the system performance.

Fig. 5.6 shows the EXIT functions of the $V = 21$ convolutional subcodes which serve as a basis for the IRCC construction as described in Section 5.3. In order to design an IRCC that matches the detector EXIT function with doping, the set of $\boldsymbol{\psi}$ are obtained by solving (5.8) with an exhaustive search. The resulting $\boldsymbol{\psi}$ are summarized in Tab. 5.1. The average

Table 5.1: $R = 0.1149$ IRCC subcode weighting factors (rounded to the fourth decimal place).

v	ψ_v	R_v	v	ψ_v	R_v	v	ψ_v	R_v
1	0	$1/100$	8	0.0342	$1/4$	15	0.0005	$3/5$
2	0	$1/50$	9	0.0158	$3/10$	16	0	$13/20$
3	0.0014	$1/40$	10	0.0029	$7/20$	17	0.0204	$7/10$
4	0.4938	$1/20$	11	0.0187	$2/5$	18	0	$3/4$
5	0.3135	$1/10$	12	0.0053	$9/20$	19	0	$4/5$
6	0.0488	$3/20$	13	0.0079	$1/2$	20	0	$17/20$
7	0.0326	$1/5$	14	0.0041	$11/20$	21	0	$9/10$

code rate is as desired $R \approx 1/10$. As shown in Fig. 5.6, the EXIT function of the IRCC is rapidly approaching the EXIT function of the detector without intersecting.

5.4.4 BER Analysis

In the following, the BER simulation results of uncoded and BICM transmission systems without and with iterative turbo detection over the diffusion-based channel are presented. It is distinguished between the LDPC codes discussed in Section 5.4.2 and the designed IRCC in Section 5.4.3. The iterative processing is terminated whenever the maximum number of $I_\mathrm{max} = 100$ iterations is reached or whenever there is no significant information

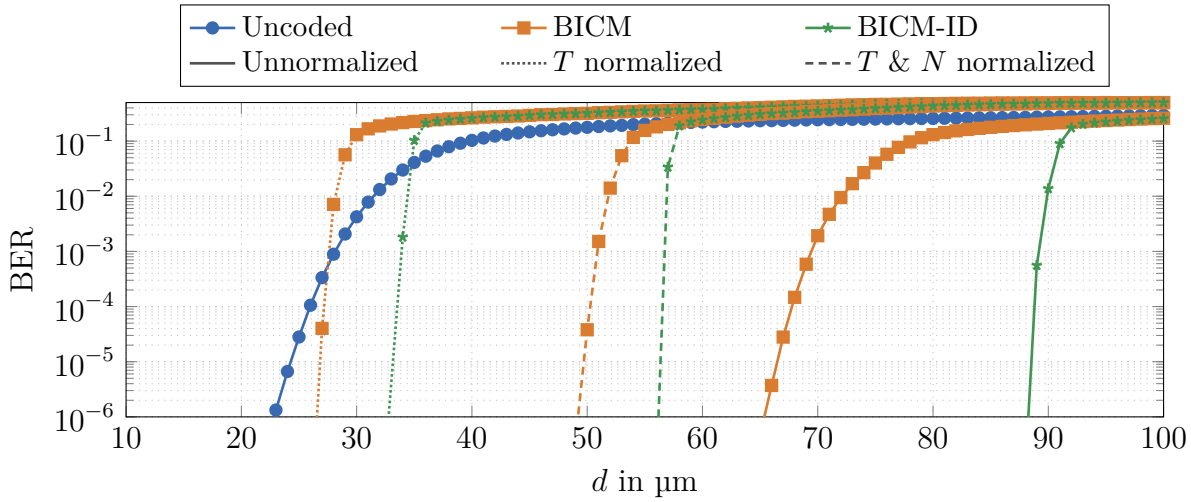
gain between two consecutive iterations, i. e. $|\sigma_{\mathrm{E}}^2[i+1] - \sigma_{\mathrm{E}}^2[i]| < 10^{-4}$. The focus is on the system performance when the transmission distance is changed.

LDPC Code In addition to the unnormalized results, T-normalized and $T\&N$-normalized results are investigated as introduced in Chapter 4. With $R = 1/4$, $R = 1/2$, and $R = 9/10$, the turbo structure is analyzed with the lowest, a medium, and the largest possible coding rate according to the DVB-S2 standard, respectively.

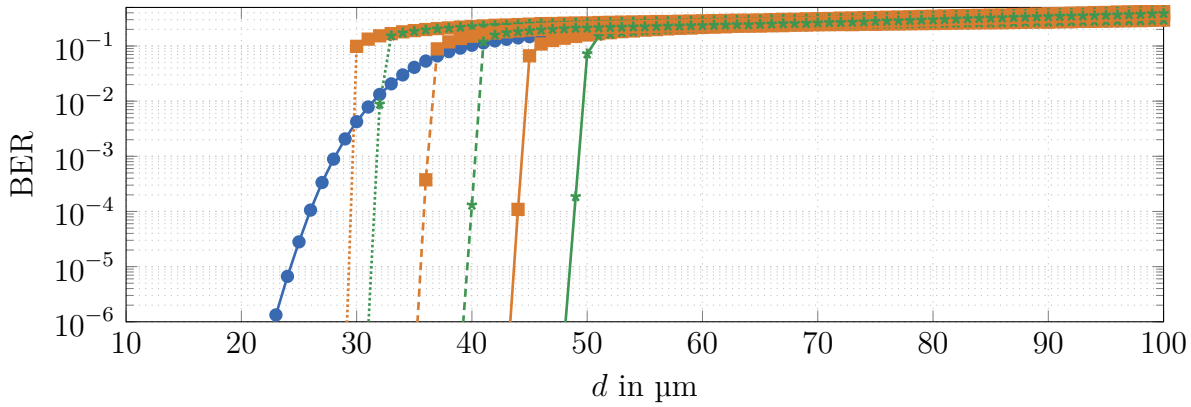
The uncoded transmission in Fig. 5.7 shows the expected behavior as discussed in Section 5.4.1. When the transmission distance is increased, the BER increases as well. A BER of 10^{-3} is achieved for $d \approx 28\,\mu\mathrm{m}$. This value will serve as comparison point to the BICM system performance. As in Fig. 5.3(b), the transmission for $d \leq 25\,\mu\mathrm{m}$ is almost error-free.

In Fig. 5.7(a), the BER results from the BICM system with $R = 1/4$ are shown. A channel code with $R = 1/4$ is a strong channel code that consists of one-quarter info bits and three-quarters redundancy. As a result, the unnormalized BICM system improves the maximum transmission distance remarkably. It reaches almost error-free transmission without ID at $d \leq 65\,\mu\mathrm{m}$ and with ID at $d \leq 88\,\mu\mathrm{m}$. This iterative gain of $23\,\mu\mathrm{m}$ can be derived from Fig. 5.5. For higher distances, the detector EXIT functions have a steeper slope. Consequently, iterative processing is particularly efficient. In addition, even for a distance $d = 60\,\mu\mathrm{m}$, there is an open tunnel between the detector and the LDPC code. Furthermore, with $R = 1/4$ it is possible to get an error-free transmission up to $C = 1/4$. In Fig. 5.3(b), this point is reached for $d \approx 98.5\,\mu\mathrm{m}$, which is almost achieved by the LDPC code under investigation, even without code matching. When T is normalized to a fixed information rate, the BICM system performance degrades, but still outperforms the uncoded transmission. Without ID an error-free transmission is achieved for $d \leq 49\,\mu\mathrm{m}$, with ID it is achieved for $d \leq 56\,\mu\mathrm{m}$. If not only the information rate but also the total number of molecules per info bit is fixed ($T\&N$-normalization), the BICM performance further degrades. Error-free transmission is reached for $d \leq 26\,\mu\mathrm{m}$ without ID and $d \leq 33\,\mu\mathrm{m}$ with ID. The iterative gain is reduced to $7\,\mu\mathrm{m}$. Compared to the uncoded transmission with BER operating point of 10^{-3}, the BICM-ID system with all normalizations still achieves a gain of $5\,\mu\mathrm{m}$, while without ID there is a loss of $2\,\mu\mathrm{m}$.

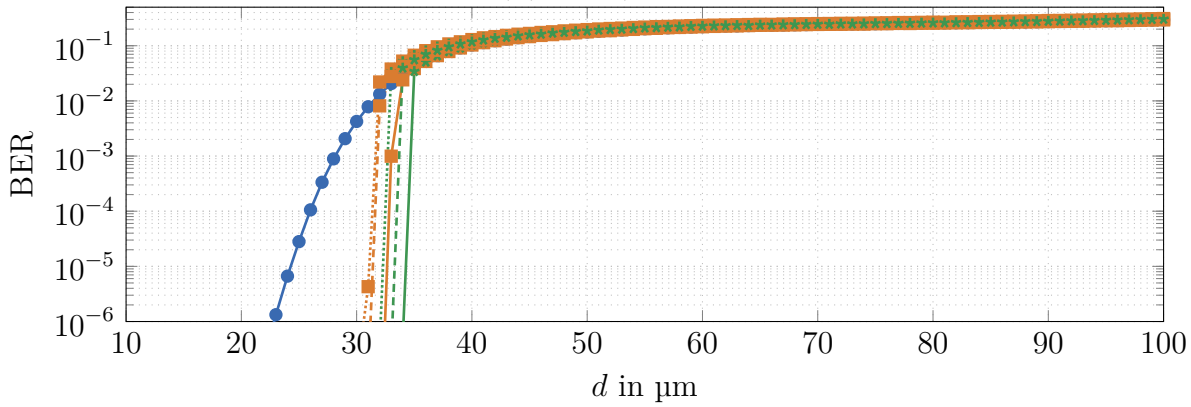
Fig. 5.7(b) shows the BER simulation results of the BICM system with $R = 1/2$. For $R = 1/2$, half of the code word consists of info bits and half of the code word contains redundancy. Consequently, without normalization the channel code is less robust than the $R = 1/4$ LDPC code. As a result, error-free transmission is only possible for $d \leq 48\,\mu\mathrm{m}$ with



(a) $R = 1/4$.



(b) $R = 1/2$.



(c) $R = 9/10$.

Figure 5.7: Bit error rate performance as a function of the transmission distance for uncoded transmission and BICM systems with LDPC codes. BICM results are given with and without applying normalization to information rate and amount of released molecules. Parameters according to Appendix C.20.

ID and $d \leq 43\,\mu\text{m}$ without ID. This result is also expected by the EXIT chart behavior in Fig. 5.5 for $d = 45\,\mu\text{m}$. While the first iteration (equivalent to the case without ID) does not lead to $I_\text{E} = 1$, the second iteration already does. This is represented in Fig. 5.7(b) by an error-free transmission for BICM-ID, but an erroneous transmission for BICM at $d = 45\,\mu\text{m}$. In Fig. 5.3(b), error-free transmission is theoretically reached for $d \approx 62.5\,\mu\text{m}$. The gap of about $14.5\,\mu\text{m}$ in Fig. 5.7(b) is mainly due to the large tunnel areas in the EXIT chart. If the channel code is matched to the EXIT function of the detector, i. e. if the tunnel area is minimized, this gap will further reduce. When the information rate is kept fixed, error-free transmission is reached for $d \leq 39\,\mu\text{m}$ with ID and for $d \leq 35\,\mu\text{m}$ without ID. When T and N are normalized, error-free transmission is obtained for $d \leq 31\,\mu\text{m}$ with ID and $d \leq 29\,\mu\text{m}$ without ID. The distance gain compared to uncoded transmission is $1\,\mu\text{m}$ without ID and $3\,\mu\text{m}$ with ID. For the BICM-ID system, $R = 1/4$ outperforms the $R = 1/2$ system, as it is expected from Fig. 5.3(b) and Fig. 5.4.

In Fig. 5.7(c), the BER results from the BICM system with $R = 9/10$ are shown. $R = 9/10$ is a very high code rate that does not offer a lot of redundancy. Consequently, it cannot correct a lot of bit errors, but when it comes to normalization the system parameters are close to the unnormalized system parameters. As a result, the curves in Fig. 5.7(c) are relatively close to each other. In the unnormalized case, almost error-free transmission is achieved at $d \leq 32\,\mu\text{m}$ without ID and $d \leq 34\,\mu\text{m}$ with ID. From Fig. 5.3(b), the theoretical communication channel capacity will be achieved at $d \approx 35.5\,\mu\text{m}$, so there is only a gap of about $1.5\,\mu\text{m}$. Consequently, the tunnel area between the detector and decoder EXIT functions is smaller than the tunnel area when the other two coding rates are applied. This is confirmed in Fig. 5.5. At distances around $30\,\mu\text{m}$, the detector EXIT function is more flat than for higher distances. Since the decoder EXIT functions are mostly flat as well, the resulting tunnel area is smaller which leads to a smaller rate loss. When the symbol rate is fixed, error-free transmission is achieved for $d \leq 31\,\mu\text{m}$ without ID and $d \leq 33\,\mu\text{m}$ with ID. As expected, these results are relatively close to the unnormalized results and the performance is worse than the performance of lower code rates. Applying normalization to T and N leads to an error-free transmission for $d \leq 30\,\mu\text{m}$ without ID and $d \leq 32\,\mu\text{m}$ with ID. In this case, the BICM system with $R = 9/10$ performs best, when no iterative detection is performed. With iterative detection instead, BICM-ID with $R = 9/10$ outperforms the BICM-ID with $R = 1/2$, but lies below the performance of $R = 1/4$. The reason for this unexpected good performance (in reference to Fig. 5.4) can be seen in Fig. 5.5. The EXIT functions of the LDPC codes are mostly flat. Normalization with $R = 9/10$ does not change the shape of the EXIT function of the detector significantly, which is for $d = 30\,\mu\text{m}$ almost

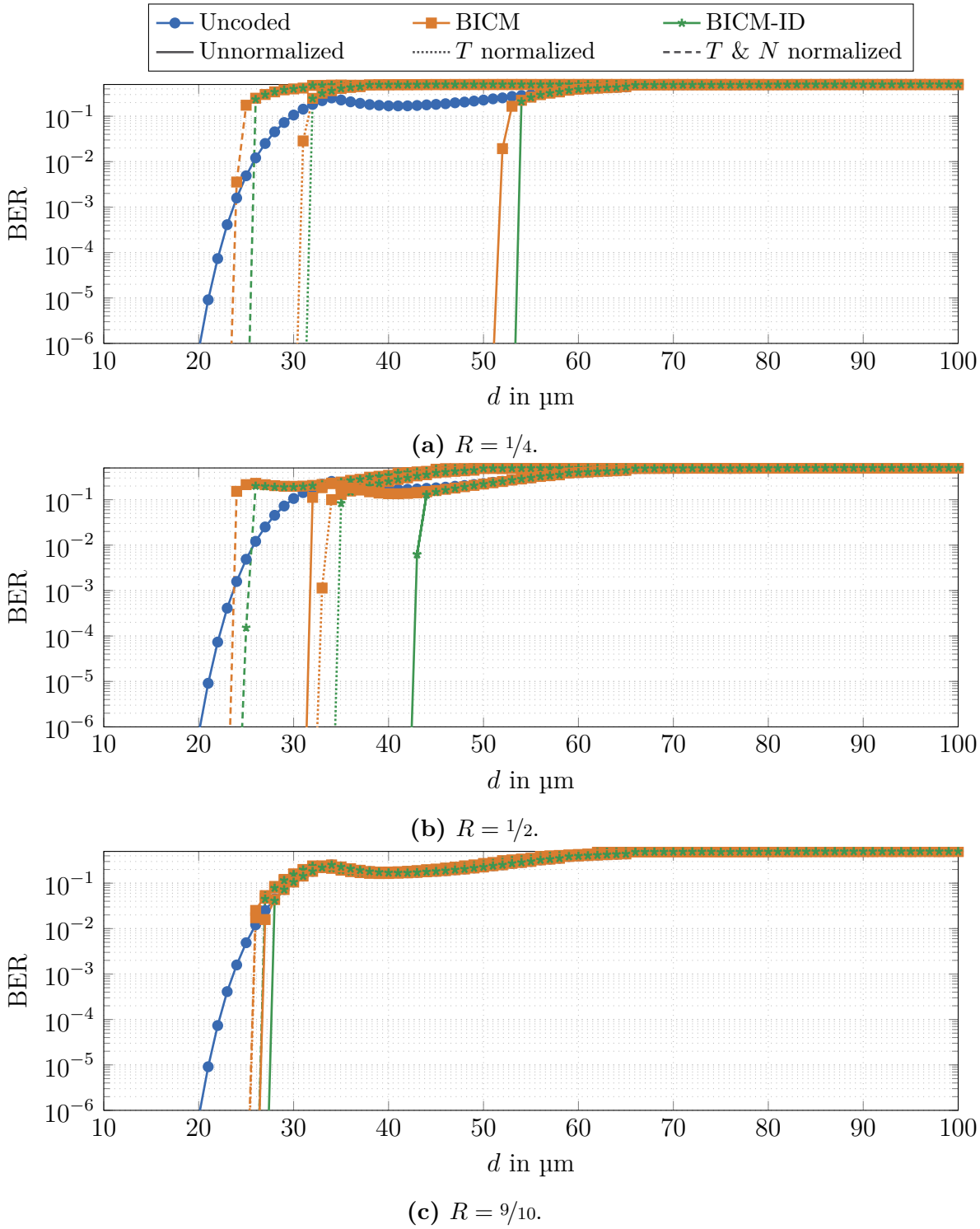
flat. As a result, the shapes of the EXIT functions fit well together and create a tunnel with a small area. Normalizing with $R = 1/2$ and $R = 1/4$ leads to a steeper slope of the EXIT function of the detector. Consequently, the EXIT function of the decoder will intersect earlier with the EXIT function of the detector. For distances that still offer an open tunnel, the area will be larger than the tunnel area given $R = 9/10$.

LDPC Code with PLD Fig. 5.8 shows the BER performance of the LDPC codes in conjunction with the piecewise linear detector (PLD) introduced in Section 3.2.3. As in the derivation of the PLD algorithm, a channel memory length of $L = 1$ is assumed. The uncoded transmission is similar to the discussed results in Section 3.2.9. The increase in maximum transmission distance is considered at a target BER of 10^{-3}. This is achieved with an uncoded transmission at $d = 23\,\mu\text{m}$. As a comparison, the APP detector leads to a maximum transmission distance of $32\,\mu\text{m}$ (see Fig. 3.22).

As in Fig. 5.7, Fig. 5.8 analyzes the LDPC codes from the DVB-S2 standard with code rates $R = 1/4$, $R = 1/2$, and $R = 9/10$, respectively. The main findings, however, do not change. All investigated codes lead in combination with BICM-ID to a gain over uncoded transmission. The amount of the obtained gain depends on the shape of the respective EXIT function of the PLD, which is different from the EXIT function of the APP detector. In the unnormalized case, the maximum transmission distance can be increased from $23\,\mu\text{m}$ to $53\,\mu\text{m}$ by applying the $R = 1/4$ LDPC code. If normalized to a constant information rate, the $R = 1/2$ LDPC code with $34\,\mu\text{m}$ provides the maximum transmission distance. With additional normalization of the average number of released molecules per infobit, the $R = 9/10$ LDPC code with a maximum transmission distance of $26\,\mu\text{m}$ is most powerful.

The PLD detector is a low-complexity detection algorithm. The reduction in computational complexity causes a performance loss compared to an APP detector. Nevertheless, PLD provides soft information at its output in the form of LLRs. Consequently, it offers a computational low-complex detection algorithm that can be applied in BICM-ID systems.

IRCC Code Fig. 5.9 shows the BER performances for various scenarios. For a fair comparison of different schemes, the focus is on a target BER of 10^{-3} as a reference point. In the scenario under investigation, an uncoded transmission scheme will reach this BER at $d \approx 28\,\mu\text{m}$. When using the convolutional subcode with rate $R_5 = 1/10$, the maximum transmission distance is extended to $d \approx 62\,\mu\text{m}$. In conjunction with BICM-ID, the transmission distance can be further increased to $d \approx 95\,\mu\text{m}$. At this distance the EXIT function of the detector will lie above the EXIT function of the decoder, which results in the characteristic



(a) $R = 1/4$.

(b) $R = 1/2$.

(c) $R = 9/10$.

Figure 5.8: Bit error rate performance as a function of the transmission distance for uncoded transmission and BICM systems with LDPC codes and PLD. BICM results are given with and without applying normalization to information rate and amount of released molecules. Parameters according to Appendix C.21.

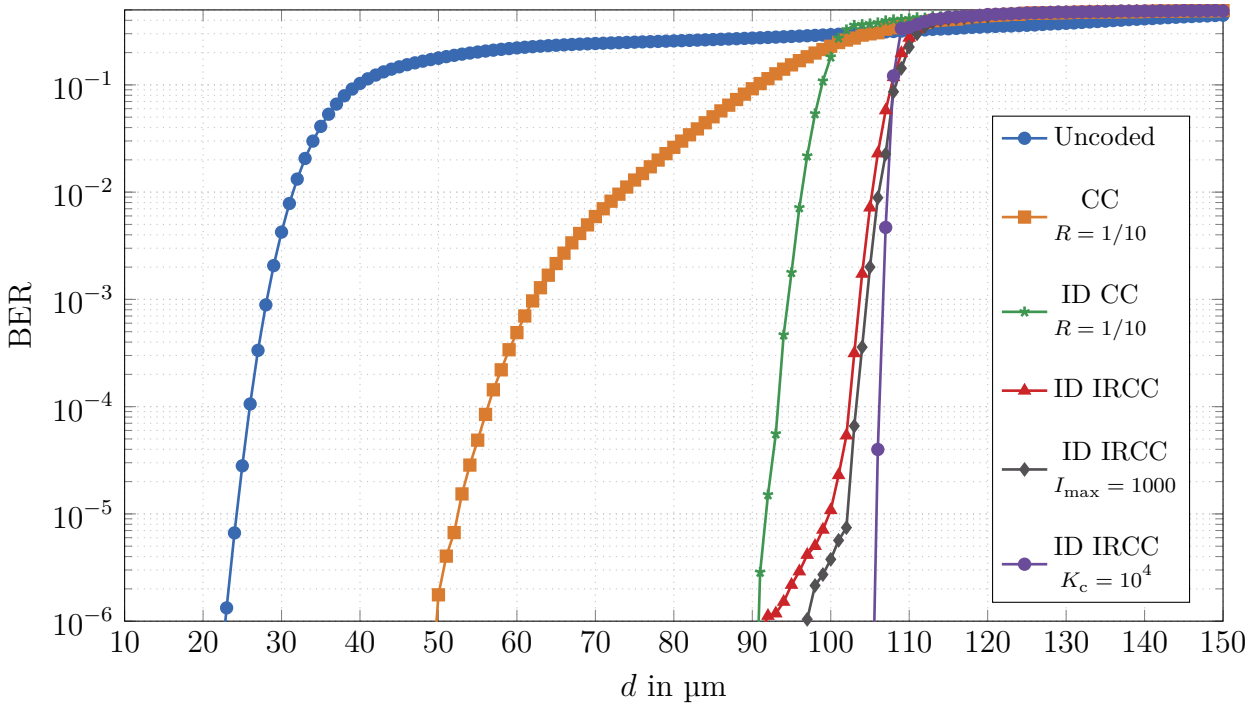


Figure 5.9: Bit error ratio performance as a function of the transmission distance for uncoded transmission, coded transmission with conventional convolutional code (CC), and BICM-ID systems with CC and matched IRCC. Parameters according to Appendix C.22.

waterfall shape. When applying the matched IRCC code in conjunction with BICM-ID, the maximum transmission distance is further increased to $d \approx 104\,\mu\text{m}$. The remaining gap to the design point of $d = 120\,\mu\text{m}$ is due to several reasons. First of all, the number of iterations in this dissertation is limited to $I_{\max} = 100$, which is not always enough to iterate through the open tunnel in Fig. 5.6. Furthermore, the EXIT chart technique presumes infinitely long sequences, which is contradictory to the focus on $K_{\text{c}} = 1000$ short sequences in this dissertation. In addition, the finite length of the pseudo-random interleaver causes correlation between the exchanged messages. When increasing the info word length to $K_{\text{c}} = 10^4$, the latter two effects are reduced. The maximum transmission distance is further increased to $d \approx 107\,\mu\text{m}$, which is, however, still $13\,\mu\text{m}$ away from the design point. A maximum number of $I_{\max} = 1000$ iterations leads to almost similar results as $I_{\max} = 100$. The maximum transmission distance is increased by approximately $1\,\mu\text{m}$. It should be noted that not all $I_{\max}$ iterations are required in every run. Finally, the data rate in molecular communication is typically quite low, so that comparatively much time is available for iterative processing. The results demonstrate that the performance of a transmission system can be significantly improved if the channel code is adapted to the EXIT function of the detector.

6

Conclusion and Outlook

6.1 Conclusion

This dissertation deals with molecular communication from a communication engineering perspective. The focus is on a diffusion-based molecular communication scenario in which the information-carrying molecules propagate passively through free diffusion in three-dimensional unbounded space. The dissertation investigates a complete transmission system consisting of a transmitter, receiver and transmission channel. In addition to channel modeling, transmitter and receiver algorithms are evaluated for their use in diffusion-based molecular communication. The main focus is on a reliable transmission of information in the overall system.

An equivalent discrete-time channel model (EDTCM) based on Poisson binomial distribution is established as a basis for the analyses performed in this dissertation. The EDTCM offers the molecular communication community a compact accessible representation of the sophisticated diffusion-based transmission channel. Because of its similarity to the EDTCM commonly used in wireless communication, it also provides easy access for communications engineers to the field of molecular communication. With known probability distributions, it can be used to significantly reduce computational time and the accuracy of molecular propagation simulations compared to random walk based propagation simulations. In order to obtain the most accurate result by random walk based propagation, the simulation time step must be chosen as small as possible. While in random walk based simulations the computational time is proportional to the number of molecules and simulation time steps, the EDTCM scales only proportional to the number of channel coefficients. Approximations of the channel model based on Gaussian distribution, Poisson distribution and binomial dis-

tribution are presented. These give the possibility of a simplified description of the channel model without causing large deviations from the Poisson binomial EDTCM. Four different methods are considered for the characterization of the channel. Random walk and artificial neural network (ANN)-based characterization is always possible, while the analytical and testbed-based determination is only solvable for some specific scenarios. It is demonstrated that the presented testbed is suitable as proof of concept of the diffusion-based transmission channel, while for more in-depth analysis further effects of the environment and hardware have to be considered.

Based on the channel model, transmission algorithms are evaluated with respect to their robustness against the diffusion-based channel effects, which are mainly strong intersymbol interference (ISI) and amplitude-dependent noise. The focus is on modulation schemes, channel equalization and detection, and forward error correction (FEC) channel coding. In addition to the existing modulation schemes on-off keying (OOK), concentration shift keying (CSK), pulse-position modulation (PPM), molecule shift keying (MoSK), and molecule concentration shift keying (MCSK), a novel modulation scheme is proposed dubbed variable concentration shift keying (VCSK). It is shown that VCSK is superior to CSK in terms of bit error rate (BER) performance by adapting to the amplitude-dependent noise of the diffusion-based channel. This result demonstrates that the consideration of the channel properties can lead to a significant improvement. A comparison of all modulation schemes with a constant information rate shows that as many different molecule types and as few concentration levels as possible should be exploited for modulation.

The analysis of the receiver algorithms shows that channel equalization counteracts the typical ISI of the diffusion-based transmission channel and significantly improves the error-free detection performance. Particularly noteworthy is the derivation of the *a posteriori* probability (APP) detector which is the optimal detector of the assumed channel model and therefore can be used as a performance benchmark in diffusion-based molecular communication systems. It also forms the basis for the calculation of log-likelihood ratio (LLR) values and thus for the determination of the communication channel capacity. Furthermore, adaptive threshold detector (ATD) is proposed as a low-complexity biologically inspired detection algorithm, which outperforms simple fixed threshold detector (FTD) in the presence of significant ISI in terms of BER performance without requiring explicit channel knowledge. These properties make it a promising candidate for implementation in a performance-limited nanomachine. A low-complex detector that is able to take *a priori* information into account and provides LLRs at its output is proposed with the piecewise linear detector (PLD). With orthogonal frequency-division multiplexing (OFDM) a

multi-carrier modulation scheme is introduced to the field of molecular communication, which has its advantages in simple channel equalization and detection algorithms. The BER performance, as well as the computational complexity, is between the optimal APP detector and the simple FTD. Thus, a significant improvement over the FTD detector can be achieved with significantly less effort than for the APP detector. Therefore, OFDM represents a good compromise. It is demonstrated that the BER performance of an OFDM system strongly depends on its parameters and that asymmetrically clipped optical OFDM (ACO-OFDM) is most reliable considering a constant information rate.

Based on a system with OOK modulation and APP detection, numerous channel codes are investigated for their use in molecular communication. The methods considered range from block codes and convolutional codes to line codes and the introduction of spreading codes and spatial codes in diffusion-based molecular communication. The latter are used in multiple-input multiple-output (MIMO) systems. Besides state-of-the-art codes, novel line codes adapted to the error statistics of the diffusion-based transmission channel are proposed, which exceeds the performance of the line codes known from the literature. These codes demonstrate that the channel with its properties should be considered in the code word selection. The performance of the channel codes is investigated under different normalizations. While for the unnormalized case all channel codes lead to a gain, this gain is removed for many channel codes if the information rate, the average number of emitted molecules per info bit, and the receiver volume are constant. For the scenario under investigation, the low-density parity check (LDPC) code is proven to be the most powerful. In the case of MIMO transmission with spatial coding it is demonstrated that repetition MIMO should be used instead of Alamouti-type coding and that no diversity gain can be achieved in the diffusion-based channel under consideration. The comprehensive analysis provides the molecular communication community with a broad spectrum of coding algorithms from which a system designer can select the appropriate channel code according to the requirements.

Furthermore, an extended system structure based on the turbo principle is investigated. Therefore modulator/detector and channel code are concatenated by an interleaver. It is shown that iterative exchange of extrinsic information between detector and decoder can significantly increase transmission reliability. An extrinsic information transfer (EXIT) chart analysis is used to estimate the communication channel capacity and thus, the theoretical performance limit. Using this method, performance limits for molecular communication systems can be determined and necessary code rates for transmissions can be obtained. In addition, the EXIT chart analysis is used to predict the gain of iterative processing and to

design an irregular convolutional code (IRCC) adapted to the channel. It is demonstrated that a large performance gain close to the communication channel capacity can be achieved by turbo processing using a suitable code.

The comprehensive analyses and results of this dissertation give the molecular communication community an extensive view of an overall system design including channel modeling, modulation, channel equalization, detection, and channel coding. Using these results, algorithms can be selected according to the application scenario and requirements in order to ensure the most reliable information transmission.

6.2 Outlook

Since the research area of molecular communication is still a relatively young research area, there are open research topics in many different directions. This outlook is restricted to possible further research related to the topics covered in this dissertation.

The system considered in this dissertation and the corresponding assumptions are focused on the pure effect of diffusion. Therefore, many side effects occurring in reality are neglected, which, however, will have an influence on a transmission system. To create a more realistic representation of a practical system, the channel model can be extended by considering additional effects. For example, bounded environments with arbitrary shapes, obstacles inside the environment, as well as drift and background molecules in the transmission medium could be modeled. On the transmitter side, the transmitter can be investigated with a spatial extension and different release characteristics. On the receiver side, modelling the binding process of molecules to receptors or different spatial receiver shapes are of interest. Furthermore, movable transmitters and receivers can be studied.

The proof of concept implementation of the testbed has shown that in macroscopic communication further effects occur which influence the transmission system. These additional effects should also be considered in the channel model, in order to ensure the most accurate description of the processes occurring in reality. These effects include turbulences, location-dependent drift, air circulation in rooms, gravity, as well as the modeling of transmitter and receiver hardware imperfections. While the testbed presented in this dissertation serves only as proof of concept, new and more professional measurement setups should be developed in the future. These are especially helpful for the extension of channel models and for the validation of transmission algorithms. Practical measurement setups are currently missing especially in the field of microscopic molecular communication.

The results presented throughout the dissertation could be re-evaluated by new channel

models and the algorithms could be adapted to newly discovered effects. Furthermore, the diffusion-based transmission channel offers further potential for the search of powerful codes adapted to amplitude-dependent noise. This dissertation provides, using the newly introduced line codes, a first approach in this direction. These line codes show that there is great potential for improvement in the appropriate selection and coupling of code words. An optimal selection and coupling is still an open topic for further work. The search for particularly low-complexity and biologically feasible transmission algorithms is especially crucial for microsocopic molecular communication in which strong energy limitations apply. The ATD and PLD presented in this dissertation are good examples of such an algorithms. The search for other low-complexity soft detectors and soft decoders are still open tasks.

A

Notation

Acronyms

ACO-OFDM	Asymmetrically clipped optical orthogonal frequency-division multiplexing
ANN	Artificial neural network
APP	*A posteriori* probability
ASK	Amplitude shift keying
ATD	Adaptive threshold detector
BCJR	Bahl-Cocke-Jelinek-Raviv
BEC	Binary erasure channel
BER	Bit error rate
BICM	Bit-interleaved coded modulation
BICM-ID	Bit-interleaved coded modulation with iterative detection
BPSK	Binary phase shift keying
CC	Convolutional code
CDM	Code-division multiplexing
CDMA	Code-division multiple access
CP	Cyclic prefix
CRC	Cyclic redundancy check
CSK	Concentration shift keying
DC	Direct current
DCO-OFDM	Direct current biased optical orthogonal frequency-division multiplexing

DFE	Decision-feedback equalizer
DFT	Discrete Fourier transform
DNA	Deoxyribonucleic acid
DVB-S2	Second standard of digital video broadcasting over satellite
EDTCM	Equivalent discrete-time channel model
EGC	Equal gain combining
EXIT	Extrinsic information transfer
FEC	Forward error correction
FIR	Finite impulse response
FTD	Fixed threshold detector
ID	Iterative detection
IDFT	Inverse discrete Fourier transform
IIR	Infinite impulse response
ILI	Interlink interference
IRCC	Irregular convolutional code
ISI	Intersymbol interference
LDPC	Low-density parity check
LLR	Log-likelihood ratio
MCSK	Molecule concentration shift keying
MIMO	Multiple-input multiple-output
MKLD	Mean Kullback-Leibler divergence
MLSE	Maximum-likelihood sequence estimation
MRC	Maximum ratio combining
MoSK	Molecule shift keying
OFDM	Orthogonal frequency-division multiplexing
OOK	On-off keying
PAM	Pulse amplitude modulation
PAM-DMT	Pulse amplitude modulated discrete multitone transmission
PLD	Piecewise linear detector
PPM	Pulse-position modulation
QAM	Quadrature amplitude modulation
RA	Repeat-accumulate
RLL	Run-length limited
RMSE	Root-mean-square error
RNG	Random number generator

RSSE	Reduced-state sequence estimation
SD	Selection diversity
SER	Symbol error rate
SISO	Single-input single-output
SNR	Signal-to-noise ratio
SPC	Single parity check
VCSK	Variable concentration shift keying
WSPCC	Weak sequence preventing convolutional code
WSPM	Weak sequence preventing mapping
ZF	Zero-forcing

Mathematical Notation

$\mathbf{0}$	Zero vector
a	Data symbol
$\tilde{a}$	Data symbol hypothesis
$\boldsymbol{a}$	Sequence of data symbols
A_{i}	Area below EXIT function of inner component
A_{ξ}	Constant value
A_{v}	Constant value
$b_{\cdot}$	Fitting coefficient
$b_{\mathrm{ref},\cdot}$	ANN training coefficient
B_i	Number of bits allocated to ith carrier
$\mathcal{B}(N,h)$	Binomial distribution with N trials and success probability h
$\mathcal{B}(y \leq \xi \,\vert N,h\,)$	Binomial probability mass function with N trials and success probability h
c	Constant value
$\boldsymbol{c}$	Spreading sequence
C	Communication channel capacity
d	Transmission distance
d_{s}	Receiver separation distance
$d_{\min}$	Minimum distance of block code
$d_{\mathrm{p}}\,(\cdot,\cdot)$	Probability-based distance
D	Diffusion coefficient
$D(\cdot\vert\vert\cdot)$	Kullback-Leibler divergence
$\bar{D}(\cdot\vert\vert\cdot)$	Mean Kullback-Leibler divergence

E_b	Energy per info bit / number of molecules per info bit
E_s	Energy per symbol / number of molecules per symbol
$E\{\cdot\}$	Expectation value
f	Frequency index in OFDM system
$f(\cdot)$	Function
$f_\mathrm{m}(\cdot)$	Model function
F	Set of modulated subcarrier indices
$F(\cdot)$	Transform of function $f(\cdot)$
$\mathcal{F}\{\cdot\}$	Fourier transform
$\mathcal{F}^{-1}\{\cdot\}$	Inverse Fourier transform
$\mathcal{F}_{N_\mathrm{DFT}}\{\cdot\}$	N_DFT-point Fourier transform
g	ZF filter coefficient
$g(z)$	Generator polynomial
G	Number of ZF filter coefficients
G_ff	Feedforward polynomial
G_fb	Feedback polynomial
$G(z)$	Z-transform of ZF impulse response
$\boldsymbol{G}$	Generator matrix
h	Channel coefficient
h_ℓ^{ij}	ℓth channel coefficient of channel between jth transmitter and ith receiver
$h(t)$	Channel impulse response
$\boldsymbol{h}$	Vector of channel coefficients
H	Channel coefficient in frequency domain
$\boldsymbol{H}$	Vector of channel coefficients in frequency domain
$H(z)$	Z-transform of channel impulse response
I	Information
I_A	*A priori* information
I_E	Extrinsic information
I_max	Maximum turbo iterations
$\boldsymbol{I}$	Identity matrix
j	Imaginary number
J	Reduced channel memory length
$J(\sigma)$	Function that describes relationship between information and standard deviation of LLRs
$J^{-1}(I)$	Inverse $J(\sigma)$ function

k	Discrete time
k_B	Boltzmann constant
k_c	Info word length
K	Number of symbols
K_c	Number of info bits
K_s	Spreading factor
$K_\mathrm{b,DFT}$	Number of bits allocated to the IDFT input block in OFDM
L	Channel memory length
L_A	*A priori* LLR
L_c	Memory length / number of shift registers of a CC
L_CP	Cyclic prefix length
L_f	Number of fingers in rake receiver
L_E	Extrinsic LLR
$L\left(u[k]\right)$	*A priori* LLR
$L\left(u[k]\|\boldsymbol{y}\right)$	*A posteriori* LLR
$L\left(\boldsymbol{y}\|u[k]\right)$	Extrinsic LLR
$\mathcal{L}\left\{\cdot\right\}$	Laplace transform
$\mathcal{L}^{-1}\left\{\cdot\right\}$	Inverse Laplace transform
m	Bits per symbol
M	Modulation order
n	Time index in OFDM system
n_c	Code word length
n_rc	Code word length of repetition code
n_red	Number of considered code bits in CRC soft decoding
N	Number of molecules
$N_\mathrm{A}, N_\mathrm{B}$	Number of molecules of type A and type B
N_c	Number of code bits
N_DFT	DFT size
N_Tx	Number of transmitters
$N_{\Delta t}$	Number of observations in random walk simulation
$\mathbb{N}$	Natural numbers excluding 0
$\mathbb{N}_0$	Natural numbers including 0
$\mathcal{N}\left(\mu, \sigma^2\right)$	Gaussian distribution with mean μ and variance σ^2
$\mathcal{N}\left(y \leq \xi \,\middle\|\, \mu, \sigma^2\right)$	Gaussian cumulative distribution with mean μ and variance σ^2
$\mathcal{O}(\cdot)$	Proportional computational complexity

p_i	ith parity bit
$p_{\boldsymbol{s}_i,\boldsymbol{s}_j}$	Probability for state transition $\boldsymbol{s}_i$ to $\boldsymbol{s}_j$
P_b	Bit error probability
P_s	Symbol error probability
$\boldsymbol{P}$	Parity check matrix
$\boldsymbol{P}_\mathrm{c}$	Submatrix that includes parity checks
$P(\cdot)$	Probability
$P(\cdot\|\cdot)$	Conditional probability
$P_\mathrm{e}(\boldsymbol{s})$	Error probability of sequence $\boldsymbol{s}$
$\mathcal{P}(\lambda)$	Poisson distribution with event rate λ
$\mathcal{P}(y \leq \xi\,\|\lambda)$	Poisson probability mass function with event rate λ
$P_\mathcal{D}(n)$	$\mathcal{D}$ distributed probability mass function
Q	Channel realizations
r	Receiver radius
r_mol	Molecule radius
R	Code rate
R_r	Sensor resistance at reference point
R_s	Sensor resistance
R_tail	Code rate including tailing
$R_x[k]$	Autocorrelation function
$\mathcal{R}$	Poisson binomial distribution
s	Laplace transform variable
s_OFDM	Transmit OFDM symbol
$\boldsymbol{s}$	State vector
$\boldsymbol{s}_\mathrm{OFDM}$	Transmit OFDM symbol sequence
$\mathcal{S}_0$	Set of states with hypothesis zero
$\mathcal{S}_1$	Set of states with hypothesis one
t	Time
T_b	Bit duration
T_c	Chip duration
T	Symbol duration
T_end	Simulation time duration
$\boldsymbol{T}$	Transmission matrix
u	Info bit
$\tilde{u}$	Info bit hypothesis

$\hat{u}$	Estimated info bit
$\hat{u}_{\mathrm{dec}}$	Decimal representation of estimated info word
$\boldsymbol{u}$	Info bit sequence
$\tilde{\boldsymbol{u}}$	Info bit sequence hypotheses
$\hat{\boldsymbol{u}}$	Estimated info bit sequence
$\boldsymbol{u}_{\mathrm{c}}$	Info word
U	Number of concentration levels
U_{L}	Voltage over load resistance
v	Subcode index
V	Number of subcodes
$w_{\cdot,\cdot}^{(\cdot)}$	ANN weighting factor
$w_{\mathrm{H}}(\cdot)$	Hamming weight
$\boldsymbol{W}_i$	Hadamard matrix
$W(\cdot)$	Wiener process
x	Transmit symbol
$x^{(i)}$	ith transmit symbol from symbol alphabet $\mathcal{X}$
$\tilde{x}$	Transmit symbol hypothesis
$\hat{x}$	Estimated transmit symbol
$\boldsymbol{x}$	Transmit symbol sequence
$\tilde{\boldsymbol{x}}$	Transmit symbol hypotheses
$\hat{\boldsymbol{x}}$	Estimated transmit symbol sequence
$\boldsymbol{x}_{\mathrm{c}}$	Code word
X	Transmit symbol in frequency domain
$\boldsymbol{X}$	Transmit symbol sequence in frequency domain
$\mathcal{X}$	Transmit symbol alphabet
y	Receive sample / number of absorbed molecules at the receiver
y'	Receive sample after post-processing
y_{EGC}	EGC receive sample
$y_{\mathrm{IS}}[k]$	Intersection point
y_{SD}	SD receive sample
$\boldsymbol{y}$	Vector with received samples
Y	Receive sample in frequency domain
$\boldsymbol{Y}$	Vector with received samples in frequency domain
z	Z-transform variable
z^i	Delay element for i discrete timesteps

z_i	Value in ith branch of EDTCM
$z[k]$	kth mean value of EDTCM
Z	Number of molecule types
α	Distance between two neighboring symbols
α_i	Distance factor between adjacent symbols in VCSK
$\alpha_{\boldsymbol{s}_i}[k]$	Forward recursion probability
α_{p}	Primitive element of Galois field
$\alpha_{\mathrm{p}}(z)$	Primitive polynomial of Galois field
$\beta_{\boldsymbol{s}_i}[k]$	Backward recursion probability
$\gamma_{\boldsymbol{s}_i,\boldsymbol{s}_j}[k]$	Transition probability from state $\boldsymbol{s}_i$ to $\boldsymbol{s}_j$
Γ	Diversity order
$\Gamma_{\boldsymbol{s}_i}[k]$	Path metric at state $\boldsymbol{s}_i$
$\delta(\cdot)$	Dirac delta function
$\delta_{\cdot}$	Kronecker delta function
Δt	Discrete timestep
$\Delta\rho_{\mathrm{x}}, \Delta\rho_{\mathrm{y}}, \Delta\rho_{\mathrm{z}}$	Discrete step in x, y, and z dimension
ε	Power normalization factor
ζ	Fourier transform variable
ϑ	Temperature
λ	Event rate
$\lambda_{\boldsymbol{s}_i,\boldsymbol{s}_j}$	Event rate for state transition $\boldsymbol{s}_i$ to $\boldsymbol{s}_j$
$\lambda_{\boldsymbol{s}_i,\boldsymbol{s}_j}[k]$	Branch metric from state $\boldsymbol{s}_i$ to $\boldsymbol{s}_j$
$\lambda^{\mathrm{A,EGC}}_{\boldsymbol{s}_i,\boldsymbol{s}_j}[k]$	Branch metric from state $\boldsymbol{s}_i$ to $\boldsymbol{s}_j$ for Alamouti-type coding with EGC
$\lambda^{\mathrm{R,EGC}}_{\boldsymbol{s}_i,\boldsymbol{s}_j}[k]$	Branch metric from state $\boldsymbol{s}_i$ to $\boldsymbol{s}_j$ for repetition MIMO coding with EGC
$\lambda^{\mathrm{R,SD}}_{\boldsymbol{s}_i,\boldsymbol{s}_j}[k]$	Branch metric from state $\boldsymbol{s}_i$ to $\boldsymbol{s}_j$ for repetition MIMO coding with SD
μ	Mean value
$\mu_{\boldsymbol{s}_i,\boldsymbol{s}_j}$	Mean value for for state transition $\boldsymbol{s}_i$ to $\boldsymbol{s}_j$
η	Viscosity
ν	Number of shift registers in maximum length sequence generator
$\nu_{\boldsymbol{s}_i,\boldsymbol{s}_j}$	Number of released molecules for state transition $\boldsymbol{s}_i$ to $\boldsymbol{s}_j$
ξ	Threshold value
$\xi(\cdot,\cdot)$	Function
$\Xi(\cdot,\cdot)$	Transform of function $f(\cdot)\cdot$
$\Xi_{\boldsymbol{s}_i,\boldsymbol{s}_j}[n]$	Subfunction in Poisson binomial probability
Π	Interleaver

Π^{-1}	Deinterleaver
$\rho_{\mathrm{x}}, \rho_{\mathrm{y}}, \rho_{\mathrm{z}}$	Location in x, y, and z dimension
ρ, φ, θ	Location in radial, azimuthal, and polar dimension
$\boldsymbol{\rho}$	Location vector
σ	Standard deviation
σ^2	Variance
σ_{A}^2	Variances of *a priori* LLRs
σ_{E}^2	Variances of extrinsic LLRs
$\sigma_{\boldsymbol{s}_i,\boldsymbol{s}_j}^2$	Variance for state transition $\boldsymbol{s}_i$ to $\boldsymbol{s}_j$
τ	Level of crossover
τ'	Reduced level of crossover
$\tau_{\max}$	Maximum number of zeros between two consecutive ones
$\tau_{\min}$	Minimum number of zeros between two consecutive ones
$\upsilon(\cdot,\cdot)$	Function
$\Upsilon(\cdot,\cdot)$	Transform of function $\upsilon(\cdot,\cdot)$
$\phi_{\cdot}^{(\cdot)}$	ANN bias value
$\phi(t)$	Concentration at sensor
$\phi(\rho,t)$	Concentration at location ρ at time t
$\Phi(\rho,s)$	Laplace transform of concentration at location ρ
ψ_i	Weighting factor of ith subcode
$\boldsymbol{\psi}$	Vector of subcode weighting factors
ω	Noise sample
∇^2	Laplace operator

B

Proofs

B.1 Solution of Fick's Second Law

In the following, the separate derivation proves that (2.4) is a valid solution of (2.2). The derivation with respect to time results in

$$\frac{\partial}{\partial t}\left[\frac{A}{\sqrt{t^3}}\exp\left(-\frac{\rho_\mathrm{x}^2+\rho_\mathrm{y}^2+\rho_\mathrm{z}^2}{4Dt}\right)\right]=\frac{A}{\sqrt{t^5}}\exp\left(-\frac{\rho_\mathrm{x}^2+\rho_\mathrm{y}^2+\rho_\mathrm{z}^2}{4Dt}\right)\left[\frac{\rho_\mathrm{x}^2+\rho_\mathrm{y}^2+\rho_\mathrm{z}^2}{4Dt}-\frac{3}{2}\right]. \tag{B.1}$$

The second partial derivative to spatial direction ρ_x leads to

$$\frac{\partial^2}{\partial\rho_\mathrm{x}^2}\left[\frac{A}{\sqrt{t^3}}\exp\left(-\frac{\rho_\mathrm{x}^2+\rho_\mathrm{y}^2+\rho_\mathrm{z}^2}{4Dt}\right)\right]=\frac{\partial}{\partial\rho_\mathrm{x}}\left[-\frac{A\rho_\mathrm{x}}{2D\sqrt{t^5}}\exp\left(-\frac{\rho_\mathrm{x}^2+\rho_\mathrm{y}^2+\rho_\mathrm{z}^2}{4Dt}\right)\right] \tag{B.2}$$

$$=\frac{A}{D\sqrt{t^5}}\exp\left(-\frac{\rho_\mathrm{x}^2+\rho_\mathrm{y}^2+\rho_\mathrm{z}^2}{4Dt}\right)\left[\frac{\rho_\mathrm{x}^2}{4Dt}-\frac{1}{2}\right]. \tag{B.3}$$

As a result, the partial derivative in all spatial directions results in

$$\nabla^2\left[\frac{A}{\sqrt{t^3}}\exp\left(-\frac{\rho_\mathrm{x}^2+\rho_\mathrm{y}^2+\rho_\mathrm{z}^2}{4Dt}\right)\right]=\frac{A}{D\sqrt{t^5}}\exp\left(-\frac{\rho_\mathrm{x}^2+\rho_\mathrm{y}^2+\rho_\mathrm{z}^2}{4Dt}\right)\left[\frac{\rho_\mathrm{x}^2+\rho_\mathrm{y}^2+\rho_\mathrm{z}^2}{4Dt}-\frac{3}{2}\right]. \tag{B.4}$$

Thus it is shown that (2.4) is a valid solution of (2.2)

$$\frac{\partial}{\partial t}\left[\frac{A}{\sqrt{t^3}}\exp\left(-\frac{\rho_\mathrm{x}^2+\rho_\mathrm{y}^2+\rho_\mathrm{z}^2}{4Dt}\right)\right]=D\nabla^2\left[\frac{A}{\sqrt{t^3}}\exp\left(-\frac{\rho_\mathrm{x}^2+\rho_\mathrm{y}^2+\rho_\mathrm{z}^2}{4Dt}\right)\right]. \tag{B.5}$$

B.2 Identity of Second Derivative

In the following, the validity of the identity applied in (2.11) is shown by rewriting the left term of the identity:

$$\frac{1}{\rho^2}\frac{\partial}{\partial\rho}\left[\rho^2\frac{\partial f\left(\rho\right)}{\partial\rho}\right] = \frac{1}{\rho^2}\left[2\rho\frac{\partial f\left(\rho\right)}{\partial\rho} + \rho^2\frac{\partial^2 f\left(\rho\right)}{\partial\rho^2}\right] \tag{B.6}$$

$$= \frac{1}{\rho}\left[\frac{\partial f\left(\rho\right)}{\partial\rho} + \frac{\partial f\left(\rho\right)}{\partial\rho} + \rho\frac{\partial^2 f\left(\rho\right)}{\partial\rho^2}\right] \tag{B.7}$$

$$= \frac{1}{\rho}\frac{\partial}{\partial\rho}\left[f\left(\rho\right) + \rho\frac{\partial f\left(\rho\right)}{\partial\rho}\right] \tag{B.8}$$

$$= \frac{1}{\rho}\frac{\partial^2}{\partial\rho^2}\Big[\rho\, f\left(\rho\right)\Big]. \tag{B.9}$$

B.3 Intersection Point of Two Gaussian Distributions

In the following, the intersection point of two Gaussian distributions is derived. The first distribution is given as $\mathcal{N}\left(\mu_\mathrm{L}, \sigma_\mathrm{L}^2\right)$ and the second as $\mathcal{N}\left(\mu_\mathrm{R}, \sigma_\mathrm{R}^2\right)$, where $\mu_\mathrm{L} < \mu_\mathrm{R}$. In the case of amplitude-dependent noise in the diffusion-based transmission channel $\sigma_\mathrm{L}^2 < \sigma_\mathrm{R}^2$ applies. For detection, the larger intersection point is usually of interest, while the smaller one is often negative or adds only a negligible small amount to the error probability. Equating both distributions results in

$$\frac{1}{\sqrt{2\pi\sigma_\mathrm{L}^2}}\exp\left(-\frac{\left(y-\mu_\mathrm{L}\right)^2}{2\sigma_\mathrm{L}^2}\right) = \frac{1}{\sqrt{2\pi\sigma_\mathrm{R}^2}}\exp\left(-\frac{\left(y-\mu_\mathrm{R}\right)^2}{2\sigma_\mathrm{R}^2}\right), \tag{B.10}$$

$$-\frac{\left(y-\mu_\mathrm{L}\right)^2}{2\sigma_\mathrm{L}^2} - \frac{1}{2}\log\left(\sigma_\mathrm{L}^2\right) = -\frac{\left(y-\mu_\mathrm{R}\right)^2}{2\sigma_\mathrm{R}^2} - \frac{1}{2}\log\left(\sigma_\mathrm{R}^2\right), \tag{B.11}$$

$$\frac{y^2+2\mu_\mathrm{L}y+\mu_\mathrm{L}^2}{2\sigma_\mathrm{L}^2} + \frac{1}{2}\log\left(\sigma_\mathrm{L}^2\right) = \frac{y^2+2\mu_\mathrm{R}y+\mu_\mathrm{R}^2}{2\sigma_\mathrm{R}^2} + \frac{1}{2}\log\left(\sigma_\mathrm{R}^2\right). \tag{B.12}$$

Solving for y yields

$$0 = y^2 \left(\frac{1}{\sigma_{\mathrm{L}}^2} - \frac{1}{\sigma_{\mathrm{R}}^2} \right) - 2y \left(\frac{\mu_{\mathrm{L}}}{\sigma_{\mathrm{L}}^2} - \frac{\mu_{\mathrm{R}}}{\sigma_{\mathrm{R}}^2} \right) + \frac{\mu_{\mathrm{L}}^2}{\sigma_{\mathrm{L}}^2} - \frac{\mu_{\mathrm{R}}^2}{\sigma_{\mathrm{R}}^2} + \log \left(\frac{\sigma_{\mathrm{L}}^2}{\sigma_{\mathrm{R}}^2} \right), \tag{B.13}$$

$$= y^2 - 2 \frac{\sigma_{\mathrm{R}}^2 \mu_{\mathrm{L}} - \sigma_{\mathrm{L}}^2 \mu_{\mathrm{R}}}{\sigma_{\mathrm{R}}^2 - \sigma_{\mathrm{L}}^2} y + \frac{\sigma_{\mathrm{R}}^2 \mu_{\mathrm{L}}^2 - \sigma_{\mathrm{L}}^2 \mu_{\mathrm{R}}^2 + \sigma_{\mathrm{L}}^2 \sigma_{\mathrm{R}}^2 \log \left(\frac{\sigma_{\mathrm{L}}^2}{\sigma_{\mathrm{R}}^2} \right)}{\sigma_{\mathrm{R}}^2 - \sigma_{\mathrm{L}}^2}, \tag{B.14}$$

$$y = \frac{\sigma_{\mathrm{R}}^2 \mu_{\mathrm{L}} - \sigma_{\mathrm{L}}^2 \mu_{\mathrm{R}}}{\sigma_{\mathrm{R}}^2 - \sigma_{\mathrm{L}}^2} \pm \sqrt{\left(\frac{\sigma_{\mathrm{R}}^2 \mu_{\mathrm{L}} - \sigma_{\mathrm{L}}^2 \mu_{\mathrm{R}}}{\sigma_{\mathrm{R}}^2 - \sigma_{\mathrm{L}}^2} \right)^2 - \frac{\sigma_{\mathrm{R}}^2 \mu_{\mathrm{L}}^2 - \sigma_{\mathrm{L}}^2 \mu_{\mathrm{R}}^2 + \sigma_{\mathrm{L}}^2 \sigma_{\mathrm{R}}^2 \log \left(\frac{\sigma_{\mathrm{L}}^2}{\sigma_{\mathrm{R}}^2} \right)}{\sigma_{\mathrm{R}}^2 - \sigma_{\mathrm{L}}^2}}. \tag{B.15}$$

C

Simulation Parameters

In the following the simulation parameters of all numerical results shown in this dissertation are summarized. The specified values are to be considered as default values if they are not varied throughout the simulation.

C.1 Figure 2.8

Parameter		Value
Molecule radius	r_{mol}	$0.5\,\mathrm{nm}$
Viscosity	η	$10^{-3}\,\mathrm{kg/m/s}$
Temperature	ϑ	$(25 + 273.15)\,\mathrm{K}$
Diffusion coefficient	D	$4.367\,64 \times 10^{-10}\,\mathrm{m^2/s}$
Receiver radius	r	$4.5\,\mu\mathrm{m}$
Symbol duration	T	$0.5\,\mathrm{s}$
Channel memory length	L	5

C.2 Figure 3.3, 3.5, 3.8, 3.9, 3.10

Parameter		Value
Channel memory length	L	0
Number of released molecules per info bit	E_{b}	500
Number of symbols	K	1000
Channel realizations	Q	10^5

C.3 Figure 3.7

Parameter		Value
Channel memory length	L	0
Number of molecules	N	1000

C.4 Figure 3.11

Parameter		Value
Molecule radius	r_{mol}	$0.5\,\mathrm{nm}$
Viscosity	η	$10^{-3}\,\mathrm{kg/m/s}$
Temperature	ϑ	$(25 + 273.15)\,\mathrm{K}$
Diffusion coefficient	D	$4.367\,64 \times 10^{-10}\,\mathrm{m^2/s}$
Receiver radius	r	$4.5\,\mu\mathrm{m}$
Bit duration	T_{b}	$0.5\,\mathrm{s}$
Absolute channel memory duration	$(L+1)T$	$3\,\mathrm{s}$
Number of released molecules per info bit	E_{b}	500
Number of symbols	K	1000
Channel realizations	Q	10^5

C.5 Figure 3.13

Parameter		Value
Molecule radius	r_{mol}	$0.5\,\mathrm{nm}$
Viscosity	η	$10^{-3}\,\mathrm{kg/m/s}$
Temperature	ϑ	$(25 + 273.15)\,\mathrm{K}$
Diffusion coefficient	D	$4.367\,64 \times 10^{-10}\,\mathrm{m^2/s}$
Receiver radius	r	$4.5\,\mu\mathrm{m}$
Bit duration	T_{b}	$0.5\,\mathrm{s}$
Channel memory length	L	1
Number of molecules	N	1000
Number of symbols	K	1000
Channel realizations	Q	10^5

C.6 Figure 3.15, 3.19, 3.20, 3.21, 3.22

Parameter		Value
Molecule radius	r_{mol}	$0.5\,\mathrm{nm}$
Viscosity	η	$10^{-3}\,\mathrm{kg/m/s}$
Temperature	ϑ	$(25 + 273.15)\,\mathrm{K}$
Diffusion coefficient	D	$4.367\,64 \times 10^{-10}\,\mathrm{m^2/s}$
Receiver radius	r	$4.5\,\mu\mathrm{m}$
Transmission distance	d	$30\,\mu\mathrm{m}$
Symbol duration	T	$0.5\,\mathrm{s}$
Channel memory length	L	5
Number of molecules	N	1000
Number of symbols	K	1000
Channel realizations	Q	10^5

C.7 Figure 3.25

Parameter		Value
Molecule radius	r_{mol}	$0.5\,\mathrm{nm}$
Viscosity	η	$10^{-3}\,\mathrm{kg/m/s}$
Temperature	ϑ	$(25 + 273.15)\,\mathrm{K}$
Diffusion coefficient	D	$4.367\,64 \times 10^{-10}\,\mathrm{m^2/s}$
Receiver radius	r	$4.5\,\mu\mathrm{m}$
Bit duration	T_{b}	$0.5\,\mathrm{s}$
Absolute channel memory duration	$(L+1)T$	$3\,\mathrm{s}$
Number of released molecules per info bit	E_{b}	500
Number of information bits	K_{c}	1000
Channel realizations	Q	10^5
DFT size	N_{DFT}	32
Modulation order	M	16

C.8 Figure 3.26, 3.27, 3.28, 3.29

Parameter		Value
Molecule radius	r_{mol}	$0.5\,\mathrm{nm}$
Viscosity	η	$10^{-3}\,\mathrm{kg/m/s}$
Temperature	ϑ	$(25 + 273.15)\,\mathrm{K}$
Diffusion coefficient	D	$4.367\,64 \times 10^{-10}\,\mathrm{m^2/s}$
Receiver radius	r	$4.5\,\mu\mathrm{m}$
Bit duration	T_{b}	$0.5\,\mathrm{s}$
Absolute channel memory duration	$(L+1)T$	$3\,\mathrm{s}$
Number of released molecules per info bit	E_{b}	500
Number of information bits	K_{c}	1000
Channel realizations	Q	10^5
DFT size	N_{DFT}	1024

C.9 Figure 3.30

Parameter		Value
Molecule radius	r_{mol}	$0.5\,\mathrm{nm}$
Viscosity	η	$10^{-3}\,\mathrm{kg/m/s}$
Temperature	ϑ	$(25 + 273.15)\,\mathrm{K}$
Diffusion coefficient	D	$4.367\,64 \times 10^{-10}\,\mathrm{m^2/s}$
Receiver radius	r	$4.5\,\mu\mathrm{m}$
Bit duration	T_{b}	$0.5\,\mathrm{s}$
Absolute channel memory duration	$(L+1)T$	$3\,\mathrm{s}$

C.10 Figure 4.1, 4.2, 4.3, 4.4, 4.5, 4.10, 4.11, 4.13, 4.14

Parameter		Value		
		(a)	(b)	(c)
Molecule radius	r_{mol}	$0.5\,\mathrm{nm}$		
Viscosity	η	$10^{-3}\,\mathrm{kg/m/s}$		
Temperature	ϑ	$(25 + 273.15)\,\mathrm{K}$		
Diffusion coefficient	D	$4.367\,64 \times 10^{-10}\,\mathrm{m^2/s}$		
Receiver radius	r	$4.5\,\mu\mathrm{m}$		
Symbol duration	T	$0.5\,\mathrm{s}$	$0.5R\,\mathrm{s}$	$0.5R\,\mathrm{s}$
Channel memory length	L	5		
Number of molecules	N	1000	1000	$\lfloor 1000R \rceil$
Number of information bits	K_{c}	1000		
Channel realizations	Q	10^5		

C.11 Figure 4.7

Parameter		Value		
		(a)	(b)	(c)
Molecule radius	r_{mol}	$0.5\,\mathrm{nm}$		
Viscosity	η	$10^{-3}\,\mathrm{kg/m/s}$		
Temperature	ϑ	$(25 + 273.15)\,\mathrm{K}$		
Diffusion coefficient	D	$4.367\,64 \times 10^{-10}\,\mathrm{m^2/s}$		
Receiver radius	r	$4.5\,\mu\mathrm{m}$		
Symbol duration	T	$0.5\,\mathrm{s}$	$0.5R\,\mathrm{s}$	$0.5R\,\mathrm{s}$
Channel memory length	L	5		
Number of molecules	N	1000	1000	$\lfloor 1000R \rceil$
Number of symbols	K	64800		
Channel realizations	Q	7000		

C.12 Figure 4.15

Parameter		Value
Molecule radius	r_{mol}	$0.5\,\mathrm{nm}$
Viscosity	η	$10^{-3}\,\mathrm{kg/m/s}$
Temperature	ϑ	$(25 + 273.15)\,\mathrm{K}$
Diffusion coefficient	D	$4.367\,64 \times 10^{-10}\,\mathrm{m^2/s}$
Receiver radius	r	$4.5\,\mu\mathrm{m}$
Symbol duration	T	$0.5\,\mathrm{s}$
Channel memory length	L	5
Number of molecules	N	1000
Number of symbols	K	1000
Channel realizations	Q	10^5

C.13 Figure 4.16, 4.17, 4.19, 4.20

Parameter		Value		
		(a)	(b)	(c)
Molecule radius	r_{mol}	$0.5\,\mathrm{nm}$		
Viscosity	η	$10^{-3}\,\mathrm{kg/m/s}$		
Temperature	ϑ	$(25 + 273.15)\,\mathrm{K}$		
Diffusion coefficient	D	$4.367\,64 \times 10^{-10}\,\mathrm{m^2/s}$		
Receiver radius	r	$4.5\,\mu\mathrm{m}$		
Symbol duration	T	$0.5\,\mathrm{s}$	$0.5R\,\mathrm{s}$	$0.5R\,\mathrm{s}$
Channel memory length	L	5		
Number of released molecules per info bit	E_{b}	$N = 1000$	$N = 1000$	500
Number of information bits	K_{c}	1000		
Channel realizations	Q	10^5		

C.14 Figure 4.22

Parameter		Value		
		(a)	(b)	(c)
Molecule radius	r_{mol}	$0.5\,\mathrm{nm}$		
Viscosity	η	$10^{-3}\,\mathrm{kg/m/s}$		
Temperature	ϑ	$(25 + 273.15)\,\mathrm{K}$		
Diffusion coefficient	D	$4.367\,64 \times 10^{-10}\,\mathrm{m^2/s}$		
Receiver radius	r	$4.5\,\mu\mathrm{m}$		
Symbol duration	T	$0.5K_{\mathrm{s}}\,\mathrm{s}$	$0.5\,\mathrm{s}$	$0.5\,\mathrm{s}$
Channel memory length	L	5		
Number of molecules	N	1000	1000	$\lfloor 1000/K_{\mathrm{s}} \rceil$
Number of information bits	K_{c}	1000		
Channel realizations	Q	10^5		

C.15 Figure 4.26, 4.27

Parameter		Value
Diffusion coefficient	D	$1 \times 10^{-10}\,\mathrm{m^2/s}$
Receiver radius	r	$5\,\mu\mathrm{m}$
Transmission distance	d	$20\,\mu\mathrm{m}$
Receiver separation distance	d_{s}	$11\,\mu\mathrm{m}$
Symbol duration	T	$0.6\,\mathrm{s}$
Channel memory length	L	3
Number of molecules	N	1000
Number of information bits	K_{c}	10^6
Channel realizations	Q	10^3

C.16 Figure 5.3

Parameter		Value
Molecule radius	r_{mol}	$0.5\,\mathrm{nm}$
Viscosity	η	$10^{-3}\,\mathrm{kg/m/s}$
Temperature	ϑ	$(25 + 273.15)\,\mathrm{K}$
Diffusion coefficient	D	$4.367\,64 \times 10^{-10}\,\mathrm{m^2/s}$
Receiver radius	r	$4.5\,\mu\mathrm{m}$
Transmission distance	d	$30\,\mu\mathrm{m}$
Symbol duration	T	$0.5\,\mathrm{s}$
Channel memory length	L	5
Number of molecules	N	1000
Number of symbols	K	1000
Channel realizations	Q	10^5

C.17 Figure 5.4

Parameter		Value
Molecule radius	r_{mol}	$0.5\,\mathrm{nm}$
Viscosity	η	$10^{-3}\,\mathrm{kg/m/s}$
Temperature	ϑ	$(25 + 273.15)\,\mathrm{K}$
Diffusion coefficient	D	$4.367\,64 \times 10^{-10}\,\mathrm{m^2/s}$
Receiver radius	r	$4.5\,\mu\mathrm{m}$
Symbol duration	T	$0.5R\,\mathrm{s}$
Channel memory length	L	5
Number of molecules	N	$\lfloor 1000R \rceil$
Number of symbols	K	1000
Channel realizations	Q	10^5

C.18 Figure 5.5

Parameter		Value
Molecule radius	r_{mol}	$0.5\,\mathrm{nm}$
Viscosity	η	$10^{-3}\,\mathrm{kg/m/s}$
Temperature	ϑ	$(25 + 273.15)\,\mathrm{K}$
Diffusion coefficient	D	$4.367\,64 \times 10^{-10}\,\mathrm{m}^2/\mathrm{s}$
Receiver radius	r	$4.5\,\mathrm{\mu m}$
Symbol duration	T	$0.5\,\mathrm{s}$
Channel memory length	L	5
Number of molecules	N	1000
Number of symbols	K	648000
Channel realizations	Q	30

C.19 Figure 5.6

Parameter		Value
Molecule radius	r_{mol}	$0.5\,\mathrm{nm}$
Viscosity	η	$10^{-3}\,\mathrm{kg/m/s}$
Temperature	ϑ	$(25 + 273.15)\,\mathrm{K}$
Diffusion coefficient	D	$4.367\,64 \times 10^{-10}\,\mathrm{m}^2/\mathrm{s}$
Receiver radius	r	$4.5\,\mathrm{\mu m}$
Transmission distance	d	$120\,\mathrm{\mu m}$
Symbol duration	T	$0.5\,\mathrm{s}$
Channel memory length	L	5
Number of molecules	N	1000
Number of symbols	K	1000
Channel realizations	Q	10^5

C.20 Figure 5.7

Parameter		Value		
		-	T	$T\&N$
Molecule radius	r_{mol}	$0.5\,\mathrm{nm}$		
Viscosity	η	$10^{-3}\,\mathrm{kg/m/s}$		
Temperature	ϑ	$(25 + 273.15)\,\mathrm{K}$		
Diffusion coefficient	D	$4.367\,64 \times 10^{-10}\,\mathrm{m^2/s}$		
Receiver radius	r	$4.5\,\mu\mathrm{m}$		
Symbol duration	T	$0.5\,\mathrm{s}$	$0.5R\,\mathrm{s}$	$0.5R\,\mathrm{s}$
Channel memory length	L	5		
Number of molecules	N	1000	1000	$\lfloor 1000R \rceil$
Number of symbols	K	64800		
Channel realizations	Q	7000		

C.21 Figure 5.8

Parameter		Value		
		-	T	$T\&N$
Molecule radius	r_{mol}	$0.5\,\mathrm{nm}$		
Viscosity	η	$10^{-3}\,\mathrm{kg/m/s}$		
Temperature	ϑ	$(25 + 273.15)\,\mathrm{K}$		
Diffusion coefficient	D	$4.367\,64 \times 10^{-10}\,\mathrm{m^2/s}$		
Receiver radius	r	$4.5\,\mu\mathrm{m}$		
Symbol duration	T	$0.5\,\mathrm{s}$	$0.5R\,\mathrm{s}$	$0.5R\,\mathrm{s}$
Channel memory length	L	1		
Number of molecules	N	1000	1000	$\lfloor 1000R \rceil$
Number of symbols	K	64800		
Channel realizations	Q	7000		

C.22 Figure 5.9

Parameter		Value
Molecule radius	r_{mol}	$0.5\,\mathrm{nm}$
Viscosity	η	$10^{-3}\,\mathrm{kg/m/s}$
Temperature	ϑ	$(25 + 273.15)\,\mathrm{K}$
Diffusion coefficient	D	$4.367\,64 \times 10^{-10}\,\mathrm{m}^2/\mathrm{s}$
Receiver radius	r	$4.5\,\mu\mathrm{m}$
Symbol duration	T	$0.5\,\mathrm{s}$
Channel memory length	L	5
Number of molecules	N	1000
Number of information bits	K_{c}	1000
Channel realizations	Q	10^5
Maximum turbo iterations	$I_{\max}$	100

Bibliography

[AD82] J. H. Ahrens and U. Dieter, "Computer generation of Poisson deviates from modified normal distributions," *ACM Transactions on Mathematical Software*, vol. 8, no. 2, pp. 163–179, Jun. 1982.

[AA15] A. Aijaz and A.-H. Aghvami, "Error performance of diffusion-based molecular communication using pulse-based modulation," *IEEE Trans. Nanobiosci.*, vol. 14, no. 1, pp. 146–151, Jan. 2015.

[AP+18] B. C. Akdeniz, A. E. Pusane, and T. Tugcu, "Optimal reception delay in diffusion-based molecular communication," *IEEE Commun. Lett.*, vol. 22, no. 1, pp. 57–60, Jan. 2018.

[AK+16] P. Akhkandi, A. Keshavarz-Haddad, and A. Jamshidi, "A new channel code for decreasing inter-symbol-interference in diffusion based molecular communications," in *Proc. Int. Symp. on Telecommunications (IST)*, Tehran, Iran, Sep. 2016, pp. 277–281.

[AY+15] A. Akkaya, H. B. Yilmaz, C. Chae, and T. Tugcu, "Effect of receptor density and size on signal reception in molecular communication via diffusion with an absorbing receiver," *IEEE Commun. Lett.*, vol. 19, no. 2, pp. 155–158, Feb. 2015.

[AF+12] I. F. Akyildiz, F. Fekri, R. Sivakumar, C. R. Forest, and B. K. Hammer, "MoNaCo: Fundamentals of molecular nano-communication networks," *IEEE Trans. Wireless Commun.*, vol. 19, no. 5, pp. 12–18, Oct. 2012.

[AB+08] I. F. Akyildiz, F. Brunetti, and C. Blázquez, "Nanonetworks: A new communication paradigm," *Computer Networks*, vol. 52, no. 12, pp. 2260–2279, Aug. 2008.

[Ala98] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1451–1458, Oct. 1998.

[AA+17] G. H. Alshammri, M. S. Alzaidi, W. K. M. Ahmed, and V. B. Lawrence, "Low-complexity memory-assisted adaptive-threshold detection scheme for on-off-keying diffusion-based molecular communications," in *Proc. IEEE Sarnoff Symposium*, Sep. 2017, pp. 1–6.

[AM87] S. C. Althoen and R. Mclaughlin, "Gauss-Jordan reduction: A brief history," *The American Mathematical Monthly*, vol. 94, no. 2, pp. 130–142, Feb. 1987.

[AG+13] H. Arjmandi, A. Gohari, M. N. Kenari, and F. Bateni, "Diffusion-based nanonetworking: A new modulation technique and performance analysis," *IEEE Commun. Lett.*, vol. 17, no. 4, pp. 645–648, Apr. 2013.

[AL06] J. Armstrong and A. J. Lowery, "Power efficient optical OFDM," *Electron. Lett.*, vol. 42, no. 6, pp. 370–372, Mar. 2006.

[AK+04] A. Ashikhmin, G. Kramer, and S. ten Brink, "Extrinsic information transfer functions: Model and erasure channel properties," *IEEE Trans. Inf. Theory*, vol. 50, no. 11, pp. 2657–2673, Nov. 2004.

[AG+12] B. Atakan, S. Galmes, and O. B. Akan, "Nanoscale communication with molecular arrays in nanonetworks," *IEEE Trans. Nanobiosci.*, vol. 11, no. 2, pp. 149–160, Jun. 2012.

[AE+18] I. Atthanayake, S. Esfahani, P. Denissenko, I. Guymer, P. J. Thomas, and W. Guo, "Experimental molecular communications in obstacle rich fluids," in *Proc. ACM Int. Conf. on Nanoscale Computing and Communication (NANOCOM)*, Reykjavik, Iceland, Sep. 2018, pp. 1–2.

[BC+74] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inf. Theory*, vol. 20, no. 2, pp. 284–287, Mar. 1974.

[BL+14] C. Bai, M. S. Leeson, and M. D. Higgins, "Minimum energy channel codes for molecular communications," *Electron. Lett.*, vol. 50, no. 23, pp. 1669–1671, Nov. 2014.

[Bai16] C. Bai, "Error Control in Bacterial Quorum Communications," Ph.D. dissertation, University of Warwick, 2016.

[BL+16a] C. Bai, M. Leeson, M. D. Higgins, and Y. Lu, "Throughput and energy efficiency-based packet size optimisation of ARQ protocols in bacterial quorum communications," *Transactions on Emerging Telecommunications Technologies*, vol. 27, no. 8, pp. 1128–1143, Aug. 2016.

[BL+16b] C. Bai, M. S. Leeson, and M. D. Higgins, "Analysis of ARQ protocols for bacterial quorum communications," *Nano Communication Networks*, vol. 7, pp. 65–79, Mar. 2016.

[BL+15] ——, "Performance of SW-ARQ in bacterial quorum communications," *Nano Communication Networks*, vol. 6, no. 1, pp. 3–14, Mar. 2015.

[BK+02] I. T. Baldwin, A. Kessler, and R. Halitschke, "Volatile signaling in plant–plant–herbivore interactions: What is real?" *Current Opinion in Plant Biology*, vol. 5, no. 4, pp. 351–354, Aug. 2002.

[BL90] J. R. Barry and E. A. Lee, "Performance of coherent optical receivers," *Proc. IEEE*, vol. 78, no. 8, pp. 1369–1394, Aug. 1990.

[Bas99] B. L. Bassler, "How bacteria talk to each other: Regulation of gene expression by quorum sensing," *Current Opinion in Microbiology*, vol. 2, no. 6, pp. 582–587, Dec. 1999.

[Ber93] H. C. Berg, *Random Walks in Biology*. Princeton University Press, 1993.

[Ber68] E. Berlekamp, *Algebraic Coding Theory*. World Scientific, 1968.

[BG+92] D. P. Bertsekas, R. G. Gallager, and P. Humblet, *Data Networks*. Prentice-Hall International New Jersey, 1992, vol. 2.

[BD+19] S. Bhattacharjee, M. Damrath, and P. A. Hoeher, "EXIT chart analysis of higher order modulation schemes in molecular communications," in *Proc. ACM Int. Conf. on Nanoscale Computing and Communication (NANOCOM)*, Dublin, Ireland, Sep. 2019, pp. 1–6.

[Bla83] R. E. Blahut, *Theory and Practice of Error Control Codes*. Addison-Wesley, 1983.

[BW63] W. H. Bossert and E. O. Wilson, "The analysis of olfactory communication among animals," *Journal of Theoretical Biology*, vol. 5, no. 3, pp. 443–469, Nov. 1963.

[CC+79] J. Cain, G. Clark, and J. Geist, "Punctured convolutional codes of rate (n-1)/n and simplified maximum likelihood decoding," *IEEE Trans. Inf. Theory*, vol. 25, no. 1, pp. 97–100, Jan. 1979.

[CT+98] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Trans. Inf. Theory*, vol. 44, no. 3, pp. 927–946, May 1998.

[CT+15] T. N. Cao, D. P. Trinh, Y. Jeong, and H. Shin, "Anomalous diffusion in molecular communication," *IEEE Commun. Lett.*, vol. 19, no. 10, pp. 1674–1677, Oct. 2015.

[Car02] E. Carafoli, "Calcium signaling: A tale for all seasons," *Proc. National Academy of Sci.*, vol. 99, no. 3, pp. 1115–1122, Feb. 2002.

[CK96] J. B. Carruthers and J. M. Kahn, "Multiple-subcarrier modulation for nondirected wireless infrared communication," *IEEE J. Sel. Areas Commun.*, vol. 14, no. 3, pp. 538–546, Apr. 1996.

[CM+17] U. A. K. Chude-Okonkwo, R. Malekian, B. T. Maharaj, and A. V. Vasilakos, "Molecular communication and nanonetwork for targeted drug delivery: A survey," *IEEE Commun. Surveys Tutorials*, vol. 19, no. 4, pp. 3046–3096, fourthquarter 2017.

[CA10] L. C. Cobo and I. F. Akyildiz, "Bacteria-based communication in nanonetworks," *Nano Communication Networks*, vol. 1, no. 4, pp. 244–256, Dec. 2010.

[Cus09] E. L. Cussler, *Diffusion: Mass Transfer in Fluid Systems*, 3rd ed. Cambridge university press, 2009.

[DH18] M. Damrath and P. A. Hoeher, "EXIT-chart-aided code matching in molecular communications," in *Proc. Int. Symp. on Turbo Codes & Iterative Information Processing (ISTC)*, Hong Kong, Dec. 2018, pp. 1–5.

[DH16] ——, "Low-complexity adaptive threshold detection for molecular communication," *IEEE Trans. Nanobiosci.*, vol. 15, no. 3, pp. 200–208, Apr. 2016.

[DK+17a] M. Damrath, S. Korte, and P. A. Hoeher, "Equivalent discrete-time channel modeling for molecular communication with emphasize on an absorbing receiver," *IEEE Trans. Nanobiosci.*, vol. 16, no. 1, pp. 60–68, Jan. 2017.

[DK+17b] M. Damrath, J. J. Koshy, and P. A. Hoeher, "Application of OFDM in diffusion-based molecular communication," *IEEE Trans. Mol. Biol. Multi-Scale Commun.*, vol. 3, no. 4, pp. 254–258, Dec. 2017.

[DS+19] M. Damrath, M. Schurwanz, and P. A. Hoeher, "The turbo principle in molecular communications," in *Proc. Int. ITG-Conf. on Systems, Communications and Coding (SCC)*, Rostock, Germany, Feb. 2019, pp. 1–6.

[DY+18] M. Damrath, H. B. Yilmaz, C. Chae, and P. A. Hoeher, "Array gain analysis in molecular MIMO communications," *IEEE Access*, vol. 6, pp. 61 091–61 102, Oct. 2018.

[DY+17] ——, "Spatial coding techniques for molecular MIMO," in *Proc. IEEE Information Theory Workshop (ITW)*, Kaohsiung, Taiwan, Nov. 2017, pp. 324–328.

[DH+18] M. Damrath, P. A. Hoeher, and G. J. Forkel, "Piecewise linear detection for direct superposition modulation," *Digital Communications and Networks*, vol. 4, no. 2, pp. 98–105, Apr. 2018.

[DH+17] ——, "Symbol detection based on Voronoi surfaces with emphasis on superposition modulation," *Digital Communications and Networks*, vol. 3, no. 3, pp. 141–149, Aug. 2017.

[DH97] F. Dan Foresee and M. T. Hagan, "Gauss-Newton approximation to Bayesian learning," in *Proc. IEEE Int. Conf. on Neural Networks (ICNN)*, Houston, TX, USA, Jun. 1997, pp. 1930–1935.

[DD+13] E. De Leo, L. Donvito, L. Galluccio, A. Lombardo, G. Morabito, and L. M. Zanoli, "Communications and switching in microfluidic systems: Pure hydrodynamic control for networking labs-on-a-chip," *IEEE Trans. Commun.*, vol. 61, no. 11, pp. 4663–4677, Nov. 2013.

[DVBS2] *Digital video broadcasting (DVB); Second generation framing structure, channel coding and modulation systems for broadcasting, interactive services, news gathering and other broadband satellite applications*, ETSI EN 302 307, Nov. 2014.

[DD+17] M. B. Dissanayake, Y. Deng, A. Nallanathan, E. M. N. Ekanayake, and M. Elkashlan, "Reed Solomon codes for molecular communication with a full absorption receiver," *IEEE Commun. Lett.*, vol. 21, no. 6, pp. 1245–1248, Jun. 2017.

[DJ+98] D. Divsalar, H. Jin, and R. J. McEliece, "Coding theorems for "turbo-like" codes," in *Proc. Allerton Conf. on Communication Control and Computing*, Monticello, IL, USA, Sep. 1998, pp. 201–210.

[Dre62] S. Dreyfus, "The numerical solution of variational problems," *Journal of Mathematical Analysis and Applications*, vol. 5, no. 1, pp. 30–45, Aug. 1962.

[EM+11] A. Enomoto, M. J. Moore, T. Suda, and K. Oiwa, "Design of self-organizing microtubule networks for molecular communication," *Nano Communication Networks*, vol. 2, no. 1, pp. 16–24, Mar. 2011.

[EQ88] M. Eyuboglu and S. Qureshi, "Reduced-state sequence estimation with set partitioning and decision feedback," *IEEE Trans. Commun.*, vol. 36, no. 1, pp. 13–20, Jan. 1988.

[FK+14] N. Farsad, N. Kim, A. W. Eckford, and C. Chae, "Channel and noise models for nonlinear molecular communication systems," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 12, pp. 2392–2401, Dec. 2014.

[FP+17] N. Farsad, D. Pan, and A. Goldsmith, "A novel experimental platform for in-vessel multi-chemical molecular communications," in *Proc. IEEE Global Communications Conf. (GLOBECOM)*, Singapore, Dec. 2017, pp. 1–6.

[FY+16] N. Farsad, H. B. Yilmaz, A. Eckford, C. Chae, and W. Guo, "A comprehensive survey of recent advancements in molecular communication," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 3, pp. 1887–1919, thirdquarter 2016.

[FG+13] N. Farsad, W. Guo, and A. W. Eckford, "Tabletop molecular communication: Text messages through chemical signals," *PloS One*, vol. 8, no. 12, Dec. 2013.

[FF+14] L. Felicetti, M. Femminella, G. Reali, P. Gresele, M. Malvestiti, and J. N. Daigle, "Modeling CD40-based molecular communications in blood vessels," *IEEE Trans. Nanobiosci.*, vol. 13, no. 3, pp. 230–243, Sep. 2014.

[For65] G. Forney, "On decoding BCH codes," *IEEE Trans. Inf. Theory*, vol. 11, no. 4, pp. 549–557, Oct. 1965.

[For73] G. D. Forney, "The Viterbi algorithm," *Proc. IEEE*, vol. 61, no. 3, pp. 268–278, Mar. 1973.

[FS+18] T. Furubayashi, Y. Sakatani, T. Nakano, A. Eckford, and N. Ichihashi, "Design and wet-laboratory implementation of reliable end-to-end molecular communication," *Wireless Networks*, vol. 24, no. 5, pp. 1809–1819, Jul. 2018.

[Gal62] R. Gallager, "Low-density parity-check codes," *IRE Trans. Inform. Theory*, vol. 8, no. 1, pp. 21–28, Jan. 1962.

[GL+11] N. Garralda, I. Llatser, A. Cabellos-Aparicio, E. Alarcón, and M. Pierobon, "Diffusion-based physical channel identification in molecular nanonetworks," *Nano Communication Networks*, vol. 2, no. 4, pp. 196–204, Dec. 2011.

[GA04] Y. George and O. Amrani, "Bit loading algorithms for OFDM," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Chicago, IL, USA, Jun. 2004, pp. 391–391.

[GM+17] S. Giannoukos, A. Marshall, S. Taylor, and J. Smith, "Molecular communication over gas stream channels using portable mass spectrometry," *Journal of The American Society for Mass Spectrometry*, vol. 28, no. 11, pp. 2371–2383, Jul. 2017.

[Gol61] M. Golay, "Complementary series," *IRE Trans. Inf. Theory*, vol. 7, no. 2, pp. 82–87, Apr. 1961.

[GW+82] S. Golomb, L. Welch, R. Goldstein, and A. Hales, *Shift Register Sequences*. Aegean Park Press, 1982.

[GR+14] I. S. Gradshteyn, I. M. Ryzhik, D. Zwillinger, and V. Moll, *Table of Integrals, Series, and Products*, 8th ed. Elsevier Science, 2014.

[GK+19] L. Grebenstein, J. Kirchner, R. S. Peixoto, W. Zimmermann, F. Irnstorfer, W. Wicke, A. Ahmadzadeh, V. Jamali, G. Fischer, R. Weigel, A. Burkovski, and R. Schober, "Biological optical-to-chemical signal conversion interface: A small-scale modulator for molecular communications," *IEEE Trans. Nanobiosci.*, vol. 18, no. 1, pp. 31–42, Jan. 2019.

[GA10] M. Gregori and I. Akyildiz, "A new nanonetwork architecture using flagellated bacteria and catalytic nanomotors," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 4, pp. 612–619, May 2010.

[GN02] L. G. Griffith and G. Naughton, "Tissue engineering–current challenges and expanding opportunities," *Science*, vol. 295, no. 5557, pp. 1009–1014, Feb. 2002.

[GA+13] Z. Gu, A. A. Aimetti, Q. Wang, T. T. Dang, Y. Zhang, O. Veiseh, H. Cheng, R. S. Langer, and D. G. Anderson, "Injectable nano-network for glucose-mediated insulin delivery," *ACS Nano*, vol. 7, no. 5, pp. 4194–4201, May 2013.

[GD+18] W. Guo, Y. Deng, H. B. Yilmaz, N. Farsad, M. Elkashlan, A. Eckford, A. Nallanathan, and C. Chae, "SMIET: Simultaneous molecular information and energy transfer," *IEEE Wireless Commun.*, vol. 25, no. 1, pp. 106–113, Feb. 2018.

[GM+15] W. Guo, C. Mias, N. Farsad, and J. L. Wu, "Molecular versus electromagnetic wave propagation loss in macro-scale environments," *IEEE Trans. Mol. Biol. Multi-Scale Commun.*, vol. 1, no. 1, pp. 18–25, Mar. 2015.

[GB+19a] M. C. Gursoy, E. Basar, A. E. Pusane, and T. Tugcu, "Index modulation for molecular communication via diffusion systems," *IEEE Trans. Commun.*, vol. 67, no. 5, pp. 3337–3350, May 2019.

[GB+19b] ——, "Pulse position-based spatial modulation for molecular communications," *IEEE Commun. Lett.*, vol. 23, no. 4, pp. 596–599, Apr. 2019.

[HM94] M. T. Hagan and M. B. Menhaj, "Training feedforward networks with the Marquardt algorithm," *IEEE Trans. Neural Netw.*, vol. 5, no. 6, pp. 989–993, Nov. 1994.

[Hag04] J. Hagenauer, "The EXIT chart - Introduction to extrinsic information transfer," in *Proc. European Signal Processing Conf. (EUSIPCO)*, Vienna, Austria, Sep. 2004, pp. 1541–1548.

[Ham50] R. W. Hamming, "Error detecting and error correcting codes," *Bell Syst. Technical J.*, vol. 29, no. 2, pp. 147–160, Apr. 1950.

[HP+13] P. C. Hansen, V. Pereyra, and G. Scherer, *Least Squares Data Fitting with Applications*. JHU Press, 2013.

[HM+16] P. He, Y. Mao, Q. Liu, and K. Yang, "Improving reliability performance of diffusion-based molecular communication with adaptive threshold variation algorithm," *International Journal of Communication Systems*, vol. 29, no. 18, pp. 2669–2680, Dec. 2016.

[HT+08] S. Hiyama, S. Takeuchi, R. Gojo, T. Shima, and K. Sutoh, "Biomolecular motor-based cargo transporters with loading/unloading mechanisms on a micro-patterned DNA array," in *Proc. IEEE Int. Conf. on Micro Electro Mechanical Systems*, Wuhan, China, Jan. 2008, pp. 144–147.

[HM+10] S. Hiyama, Y. Moritani, R. Gojo, S. Takeuchi, and K. Sutoh, "Biomolecular-motor-based autonomous delivery of lipid vesicles as nano-or microscale reactors on a chip," *Lab on a Chip*, vol. 10, no. 20, pp. 2741–2748, Aug. 2010.

[Hoe19] P. A. Hoeher, *Visible Light Communications: Theoretical and Practical Foundations*. Carl Hanser Verlag, 2019.

[Höh11] P. A. Höher, *Grundlagen der Digitalen Informationsübertragung*. Springer, 2011.

[HI+15] M. K. Hojo, K. Ishii, M. Sakura, K. Yamaguchi, S. Shigenobu, and M. Ozaki, "Antennal RNA-sequencing analysis reveals evolutionary aspects of chemosensory proteins in the carpenter ant, Camponotus japonicus," *Scientific Reports*, vol. 5, p. 13 541, Aug. 2015.

[HW+19] Y. Huang, M. Wen, L. Yang, C. Chae, and F. Ji, "Spatial modulation for molecular communication," *IEEE Trans. Nanobiosci.*, vol. 18, no. 3, pp. 381–395, Jul. 2019.

[Imm04] K. A. S. Immink, *Codes for Mass Data Storage Systems*. Shannon Foundation Publisher, 2004.

[JA+17] V. Jamali, A. Ahmadzadeh, and R. Schober, "On the design of matched filters for molecule counting receivers," *IEEE Commun. Lett.*, vol. 21, no. 8, pp. 1711–1714, Aug. 2017.

[JA+19] V. Jamali, A. Ahmadzadeh, W. Wicke, A. Noel, and R. Schober, "Channel modeling for diffusive molecular communication — A tutorial review," *Proc. IEEE*, vol. 107, no. 7, pp. 1256–1301, Jul. 2019.

[JF+18] V. Jamali, N. Farsad, R. Schober, and A. Goldsmith, "Non-coherent detection for diffusive molecular communication systems," *IEEE Trans. Commun.*, vol. 66, no. 6, pp. 2515–2531, Jun. 2018.

[Joh10] S. J. Johnson, *Iterative Error Correction: Turbo, Low-Density Parity-Check and Repeat-Accumulate Codes*. Cambridge University Press, 2010.

[KS88] V. Kachitvichyanukul and B. W. Schmeiser, "Binomial random variate generation," *Communications of the ACM*, vol. 31, no. 2, pp. 216–222, Feb. 1988.

[KA+12] S. Kadloor, R. S. Adve, and A. W. Eckford, "Molecular communication using Brownian motion with drift," *IEEE Trans. Nanobiosci.*, vol. 11, no. 2, pp. 89–99, Jun. 2012.

[KS+18] E. Kennedy, P. Shakya, M. Ozmen, C. Rose, and J. K. Rosenstein, "Spatiotemporal information preservation in turbulent vapor plumes," *Applied Physics Letters*, vol. 112, no. 26, p. 264 103, Jun. 2018. [Online]. Available: `https://doi.org/10.1063/1.5037710`.

[KJ+19] A. Keshavarz-Haddad, A. Jamshidi, and P. Akhkandi, "Inter-symbol interference reduction channel codes based on time gap in diffusion-based molecular communications," *Nano Communication Networks*, vol. 19, pp. 148–156, Mar. 2019.

[KA13] D. Kilinc and O. Akan, "Receiver design for molecular communication," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 12, pp. 705–714, Dec. 2013.

[KC13] N. Kim and C. Chae, "Novel modulation techniques using isomers as messenger molecules for nano communication networks via diffusion," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 12, pp. 847–856, Dec. 2013.

[KE+14] N. Kim, A. W. Eckford, and C. Chae, "Symbol interval optimization for molecular communication with drift," *IEEE Trans. Nanobiosci.*, vol. 13, no. 3, pp. 223–229, Sep. 2014.

[KF+15] N. R. Kim, N. Farsad, C. B. Chae, and A. W. Eckford, "A universal channel model for molecular communication systems with metal-oxide detectors," in *Proc. IEEE Int. Conf. on Communications (ICC)*, London, UK, Jun. 2015, pp. 1054–1059.

[KL+12] P.-Y. Ko, Y.-C. Lee, P. Yeh, C.-h. Lee, and K. Chen, "A new paradigm for channel coding in diffusion-based molecular communications: Molecular coding distance function," in *Proc. IEEE Global Communications Conf. (GLOBECOM)*, Anaheim, CA, USA, Dec. 2012, pp. 3748–3753.

[KL+16] B. H. Koo, C. Lee, H. B. Yilmaz, N. Farsad, A. Eckford, and C. B. Chae, "Molecular MIMO: From theory to prototype," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 3, pp. 600–614, Mar. 2016.

[KD+17] S. Korte, M. Damrath, and P. A. Hoeher, "Multiple channel access techniques for diffusion-based molecular communications," in *Proc. Int. ITG-Conf. on Systems, Communications and Coding (SCC)*, Hamburg, Germany, Feb. 2017, pp. 1–6.

[KA+13] B. Krishnaswamy, C. M. Austin, J. P. Bardill, D. Russakow, G. L. Holst, B. K. Hammer, C. R. Forest, and R. Sivakumar, "Time-elapse communication: Bacterial communication on a microfluidic chip," *IEEE Trans. Commun.*, vol. 61, no. 12, pp. 5139–5151, Dec. 2013.

[KL51] S. Kullback and R. A. Leibler, "On information and sufficiency," *The Annals of Mathematical Statistics*, vol. 22, no. 1, pp. 79–86, Mar. 1951.

[KY+13] M. Ş. Kuran, H. B. Yilmaz, and T. Tugcu, "A tunnel-based approach for signal shaping in molecular communication," in *Proc. IEEE Int. Conf. on Communications (ICC)*, Budapest, Hungary, Jun. 2013, pp. 776–781.

[KY+11] M. Kuran, H. Yilmaz, T. Tugcu, and I. Akyildiz, "Modulation techniques for communication via diffusion in nanonetworks," in *Proc. IEEE Int. Conf. on Communications (ICC)*, Kyoto, Japan, Jun. 2011, pp. 1–5.

[KY+10] M. Ş. Kuran, H. B. Yilmaz, T. Tugcu, and B. Özerman, "Energy model for communication via diffusion in nanonetworks," *Nano Communication Networks*, vol. 1, no. 2, pp. 86–95, Jun. 2010.

[LH+04] I. Land, P. A. Hoeher, and S. Gligorevic, "Computation of symbol-wise mutual information in transmission systems with LogAPP decoders and application to EXIT charts," in *Proc. Int. ITG-Conf. on Systems, Communications and Coding (SCC)*, Erlangen, Germany, Jan. 2004, pp. 195–202.

[LP+01] J. Lassing, A. Persson, T. Ottosson, and E. Ström, "On low-rate convolutional codes for code-spread code division multiple access," 2001. [Online]. Available: `https://www.researchgate.net/profile/Erik_Stroem/publication/2573542_On_Low-Rate_Convolutional_Codes_for_Code-Spread_Code_Division_Multiple_Access/links/0046352025a93cb9d7000000.pdf` (visited on 03/22/2020).

[LY+17] C. Lee, H. B. Yilmaz, C. Chae, N. Farsad, and A. Goldsmith, "Machine learning based channel modeling for molecular MIMO communications," in *Proc. IEEE Signal Processing Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Sapporo, Japan, Jul. 2017, pp. 1–5.

[LR+09] S. C. J. Lee, S. Randel, F. Breyer, and A. M. J. Koonen, "PAM-DMT for intensity-modulated and direct-detection optical communication systems," *IEEE Photon. Technol. Lett.*, vol. 21, no. 23, pp. 1749–1751, Dec. 2009.

[LH12] M. S. Leeson and M. D. Higgins, "Forward error correction for molecular communications," *Nano Communication Networks*, vol. 3, no. 3, pp. 161–167, Sep. 2012.

[LS+16] B. Li, M. Sun, S. Wang, W. Guo, and C. Zhao, "Local convexity inspired low-complexity noncoherent signal detector for nanoscale molecular com-

munications," *IEEE Trans. Commun.*, vol. 64, no. 5, pp. 2079–2091, May 2016.

[LC01] S. Lin and D. J. Costello, *Error Control Coding*. Prentice Hall, 2001, vol. 2.

[LL+12] W.-A. Lin, Y.-C. Lee, P.-C. Yeh, and C.-H. Lee, "Signal detection and ISI cancellation for quantity-based amplitude modulation in diffusion-based molecular communications," in *Proc. IEEE Global Communications Conf. (GLOBECOM)*, Anaheim, CA, USA, Dec. 2012, pp. 4362–4367.

[LP00] E. Lindskog and A. Paulraj, "A transmit diversity scheme for channels with intersymbol interference," in *Proc. IEEE Int. Conf. on Communications (ICC)*, New Orleans, LA, USA, Jun. 2000, pp. 307–311.

[LC+13] I. Llatser, A. Cabellos-Aparicio, M. Pierobon, and E. Alarcon, "Detection techniques for diffusion-based molecular communication," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 12, pp. 726–734, Dec. 2013.

[LB+00] H. Lodish, A. Berk, S. L. Zipursky, P. Matsudaira, D. Baltimore, and J. Darnell, *Molecular Cell Biology*, 4th edition. W. H. Freeman, 2000.

[LH+15a] Y. Lu, M. D. Higgins, and M. S. Leeson, "Comparison of channel coding schemes for molecular communications systems," *IEEE Trans. Commun.*, vol. 63, no. 11, pp. 3991–4001, Nov. 2015.

[LH+15b] ——, "Self-orthogonal convolutional codes (SOCCs) for diffusion-based molecular communication systems," in *Proc. IEEE Int. Conf. on Communications (ICC)*, London, UK, Jun. 2015, pp. 1049–1053.

[LH+16] Y. Lu, M. D. Higgins, A. Noel, M. S. Leeson, and Y. Chen, "The effect of two receivers on broadcast molecular communication systems," *IEEE Trans. Nanobiosci.*, vol. 15, no. 8, pp. 891–900, Dec. 2016.

[Mac92] D. J. C. MacKay, "Bayesian interpolation," *Neural Computation*, vol. 4, no. 3, pp. 415–447, May 1992.

[MM+10] M. Mahfuz, D. Makrakis, and H. Mouftah, "Spatiotemporal distribution and modulation schemes for concentration-encoded medium-to-long range molecular communication," in *Proc. of Biennial Symposium on Communications (QBSC)*, Kingston, ON, Canada, May 2010, pp. 100–105.

[MM+12] M. Mahfuz, D. Makrakis, and H. Mouftah, "Strength based receiver architecture and communication range and rate dependent signal detection characteristics of concentration encoded molecular communication," in *Proc. Int. Conf. Broadband, Wireless Computing, Communication and Applications (BWCCA)*, Victoria, BC, Canada, Nov. 2012, pp. 28–35.

[MM+11] M. U. Mahfuz, D. Makrakis, and H. T. Mouftah, "On the detection of binary concentration-encoded unicast molecular communication in nanonetworks," in *Proc. 4th Int. Conf. on Bio-Inpsired Syst. Signal Process. (BIOSIGNALS-2011)*, Rome, Italy, Jan. 2011, pp. 26–29.

[MM+13] ——, "Performance analysis of convolutional coding techniques in diffusion-based concentration-encoded PAM molecular communication systems," *Bio-NanoScience*, vol. 3, no. 3, pp. 270–284, Sep. 2013.

[MAMOKO] Bundesministerium für Bildung und Forschung. (2019). "Makroskopische Molekulare Kommunikation (MAMOKO)," [Online]. Available: `https://www.forschung-it-sicherheit-kommunikationssysteme.de/projekte/mamoko` (visited on 09/06/2019).

[ML+18] M. A. Mangoud, M. Lestas, and T. Saeed, "Molecular motors MIMO communications for nanonetworks applications," in *Proc. IEEE Wireless Communications and Networking Conf. (WCNC)*, Barcelona, Spain, Apr. 2018, pp. 1–5.

[MP+18] A. Marcone, M. Pierobon, and M. Magarini, "Parity-check coding based on genetic circuits for engineered molecular communication between biological cells," *IEEE Trans. Commun.*, vol. 66, no. 12, pp. 6221–6236, Dec. 2018.

[Mar63] D. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," *Journal of the Society for Industrial and Applied Mathematics*, vol. 11, no. 2, pp. 431–441, Jun. 1963.

[MM+98] R. J. McEliece, D. J. C. MacKay, and Jung-Fu Cheng, "Turbo decoding as an instance of Pearl's "belief propagation" algorithm," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 2, pp. 140–152, Feb. 1998.

[MY+12a] L. S. Meng, P. C. Yeh, K. C. Chen, and I. F. Akyildiz, "MIMO communications based on molecular diffusion," in *Proc. IEEE Global Communications Conf. (GLOBECOM)*, Anaheim, CA, USA, Dec. 2012, pp. 5380–5385.

[MY+14] L.-S. Meng, P.-C. Yeh, K.-C. Chen, and I. Akyildiz, "On receiver design for diffusion-based molecular communication," *IEEE Trans. Signal Process.*, vol. 62, no. 22, pp. 6032–6044, Nov. 2014.

[MY+12b] ——, "Optimal detection for diffusion-based communications in the presence of ISI," in *Proc. IEEE Global Communications Conf. (GLOBECOM)*, Anaheim, CA, USA, Dec. 2012, pp. 3819–3824.

[MH04] J. Mietzner and P. A. Hoeher, "Boosting the performance of wireless communication systems: Theory and practice of multiple-antenna techniques," *IEEE Commun. Mag.*, vol. 42, no. 10, pp. 40–47, Oct. 2004.

[ME+06] M. Moore, A. Enomoto, T. Nakano, R. Egashira, T. Suda, A. Kayasuga, H. Kojima, H. Sakakibara, and K. Oiwa, "A design of a molecular communication system for nanomachines using molecular motors," in *Proc. IEEE Int. Conf. on Pervasive Computing and Communications Workshops (PERCOMW)*, Pisa, Italy, Mar. 2006, pp. 1–6.

[MN13] M. Moore and T. Nakano, "Oscillation and synchronization of molecular machines by the diffusion of inhibitory molecules," *IEEE Trans. Nanotechnol.*, vol. 12, no. 4, pp. 601–608, Jul. 2013.

[MS+16] M. Movahednasab, M. Soleimanifar, A. Gohari, M. Nasiri-Kenari, and U. Mitra, "Adaptive transmission rate with a fixed threshold decoder for diffusion-based molecular communication," *IEEE Trans. Commun.*, vol. 64, no. 1, pp. 236–248, Jan. 2016.

[MQ3] *Technical data MQ-3 gas sensor*, Hanwei Electronics Co., Ltd. [Online]. Available: `https://www.sparkfun.com/datasheets/Sensors/MQ-3.pdf`.

[MC+17] Y. Murin, M. Chowdhury, N. Farsad, and A. Goldsmith, "Diversity gain of one-shot communication over molecular timing channels," in *Proc. IEEE Global Communications Conf. (GLOBECOM)*, Singapore, Dec. 2017, pp. 1–6.

[NS+05] T. Nakano, T. Suda, M. Moore, R. Egashira, A. Enomoto, and K. Arima, "Molecular communication for nanomachines using intercellular calcium signaling," in *Proc. IEEE Conf. Nanotechnology*, Nagoya, Japan, Jul. 2005, pp. 478–481.

[NE+13] T. Nakano, A. W. Eckford, and T. Haraguchi, *Molecular Communication*. Cambridge University Press, 2013.

[NC+14a] A. Noel, K. Cheung, and R. Schober, "Improving receiver performance of diffusive molecular communication with enzymes," *IEEE Trans. Nanobiosci.*, vol. 13, no. 1, pp. 31–43, Mar. 2014.

[NC+14b] ——, "Optimal receiver design for diffusive molecular communication with flow and additive noise," *IEEE Trans. Nanobiosci.*, vol. 13, no. 3, pp. 350–362, Sep. 2014.

[NE17] A. Noel and A. W. Eckford, "Asynchronous peak detection for demodulation in molecular communication," in *Proc. IEEE Int. Conf. on Communications (ICC)*, Paris, France, May 2017, pp. 1–6.

[NC+14c] A. Noel, K. C. Cheung, and R. Schober, "Overcoming noise and multiuser interference in diffusive molecular communication," in *Proc. of ACM Int. Conf. on Nanoscale Computing and Communication (NANOCOM)*, Atlanta, GA, USA, May 2014, pp. 1–9.

[OFa89] T. O'Farrell, "Code-division Multiple-access (CDMA) Techniques in Optical Fibre Local Area Networks," Ph.D. dissertation, University of Manchester, 1989.

[OE12] M. E. Ortiz and D. Endy, "Engineered cell-cell communication via DNA messaging," *Journal of biological engineering*, vol. 6, no. 1, p. 16, Dec. 2012.

[Par10] J. Park, "Lab-on-a-chip technology for integrative bioengineering," in *Proc. IEEE Int. Conf. on Nanotechnology*, Seoul, South Korea, Aug. 2010, pp. 156–159.

[PB61] W. W. Peterson and D. T. Brown, "Cyclic codes for error detection," *Proc. of the IRE*, vol. 49, no. 1, pp. 228–235, Jan. 1961.

[PS06] S. Pfletschinger and F. Sanzi, "Error floor removal for bit-interleaved coded modulation with iterative detection," *IEEE Trans. Wireless Commun.*, vol. 5, no. 11, pp. 3174–3181, Nov. 2006.

[PA11] M. Pierobon and I. Akyildiz, "Diffusion-based noise analysis for molecular communication in nanonetworks," *IEEE Trans. Signal Process.*, vol. 59, no. 6, pp. 2532–2547, Jun. 2011.

[DW91] Di-Ping Chou and S. G. Wilson, "Soft-decision decoding of CRC codes," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Budapest, Hungary, Jun. 1991, pp. 91–91.

[PO98] R. Prasad and T. Ojanpera, "A survey on CDMA: Evolution towards wideband CDMA," in *Proc. IEEE Int. Symp. on Spread Spectrum Techniques and Applications (SSTA)*, Sun City, South Africa, Sep. 1998, pp. 323–331.

[PG58] R. Price and P. E. Green, "A communication technique for multipath channels," *Proc. IRE*, vol. 46, no. 3, pp. 555–570, Mar. 1958.

[PS07] J. G. Proakis and M. Salehi, *Digital Communications*. McGraw-hill New York, 2007, vol. 5.

[PS+14] S. Pudasaini, S. Shin, and K. S. Kwak, "Robust Modulation Technique for Diffusion-based Molecular Communication in Nanonetworks," *arXiv e-prints*, arXiv:1401.3938, Jan. 2014.

[QG+14] S. Qiu, W. Guo, S. Wang, N. Farsad, and A. Eckford, "A molecular communication link for monitoring in confined environments," in *Proc. IEEE Int. Conf. on Communications (ICC)*, Sydney, NSW, Australia, Jun. 2014, pp. 718–723.

[RS60] I. S. Reed and G. Solomon, "Polynomial codes over certain finite fields," *Journal of the Society for Industrial and Applied Mathematics*, vol. 8, no. 2, pp. 300–304, Jun. 1960.

[RS+01] T. J. Richardson, M. A. Shokrollahi, and R. L. Urbanke, "Design of capacity-approaching irregular low-density parity-check codes," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 619–637, Feb. 2001.

[RG+04] A. Roumy, S. Guemghar, G. Caire, and S. Verdu, "Design methods for irregular repeat-accumulate codes," *IEEE Trans. Inf. Theory*, vol. 50, no. 8, pp. 1711–1727, Aug. 2004.

[RS17] S. M. Rouzegar and U. Spagnolini, "Channel estimation for diffusive MIMO molecular communications," in *Proc. European Conf. on Networks and Communications (EuCNC)*, Oulu, Finland, Jun. 2017, pp. 1–5.

[SU08] M. Safari and M. Uysal, "Do we really need OSTBCs for free-space optical communication with direct detection?" *IEEE Trans. Wireless Commun.*, vol. 7, no. 11, pp. 4445–4448, Nov. 2008.

[SW+18] M. Schäfer, W. Wicke, R. Rabenstein, and R. Schober, "An nD model for a cylindrical diffusion-advection problem with an orthogonal force component," in *Proc. IEEE Int. Conf. on Digital Signal Processing (DSP)*, Shanghai, China, Nov. 2018, pp. 1–5.

[SW+19] ——, "Analytical models for particle diffusion and flow in a horizontal cylinder with a vertical force," in *Proc. IEEE Int. Conf. on Communications (ICC)*, Shanghai, China, May 2019, pp. 1–7.

[SK00] K. Schulten and I. Kosztin, "Lectures in theoretical biophysics," *University of Illinois*, vol. 117, 2000.

[SJ18] A. Shahbazi and A. Jamshidi, "Adaptive weighted signal detection for nanoscale molecular communications," in *Proc. Int. Zurich Seminar on Information and Communication (IZS)*, Zurich, Switzerland, Feb. 2018, pp. 149–152.

[SM+12] H. ShahMohammadian, G. G. Messier, and S. Magierowski, "Optimum receiver for molecule shift keying modulation in diffusion-based molecular communication channels," *Nano Communication Networks*, vol. 3, no. 3, pp. 183–195, Sep. 2012.

[Sha48] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Technical J.*, vol. 27, no. 3, pp. 379–423, Jul. 1948.

[SL+12] P. J. Shih, C. H. Lee, and P. C. Yeh, "Channel codes for mitigating intersymbol interference in diffusion-based molecular communications," in *Proc. IEEE Global Communications Conf. (GLOBECOM)*, Anaheim, CA, USA, Dec. 2012, pp. 4228–4232.

[SL+13] P. J. Shih, C. H. Lee, P. C. Yeh, and K. C. Chen, "Channel codes for reliability enhancement in molecular communication," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 12, pp. 857–867, Dec. 2013.

[SV05] M. K. Simon and V. A. Vilnrotter, "Alamouti-type space-time coding for free-space optical communication with direct detection," *IEEE Trans. Wireless Commun.*, vol. 4, no. 1, pp. 35–39, Jan. 2005.

[SM+14] A. Singhal, R. Mallik, and B. Lall, "Effect of molecular noise in diffusion-based molecular communication," *IEEE Wireless Commun. Lett.*, vol. 3, no. 5, pp. 489–492, Oct. 2014.

[SE+12] K. Srinivas, A. Eckford, and R. Adve, "Molecular communication in fluid media: The additive inverse Gaussian noise channel," *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4678–4692, Jul. 2012.

[SM+05] T. Suda, M. Moore, T. Nakano, R. Egashira, A. Enomoto, S. Hiyama, and Y. Moritani, "Exploratory research on molecular communication between nanomachines," in *Proc. Late Breaking Papers Genetic and Evol. Comp. Conf. (GECCO)*, Washington, DC, USA, Jun. 2005, pp. 1–5.

[tBri01a] S. ten Brink, "Code characteristic matching for iteratively decoding of serially concatenated codes," *Annales Des Télécommunications*, vol. 56, no. 7, pp. 394–408, Jul. 2001.

[tBri01b] ——, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE Trans. Commun.*, vol. 49, no. 10, pp. 1727–1737, Oct. 2001.

[tBri99] ——, "Convergence of iterative decoding," *Electron. Lett.*, vol. 35, no. 10, pp. 806–808, May 1999.

[tBK+04] S. ten Brink, G. Kramer, and A. Ashikhmin, "Design of low-density parity-check codes for modulation and detection," *IEEE Trans. Commun.*, vol. 52, no. 4, pp. 670–678, Apr. 2004.

[tBS+98] S. ten Brink, J. Speidel, and R.-H. Yan, "Iterative demapping and decoding for multilevel modulation," in *Proc. IEEE Global Communications Conf. (GLOBECOM)*, Sydney, NSW, Australia, Nov. 1998, pp. 579–584.

[TP+15] B. Tepekule, A. E. Pusane, M. S. Kuran, and T. Tugcu, "A novel pre-equalization method for molecular communication via diffusion in nanonetworks," *IEEE Commun. Lett.*, vol. 19, no. 8, pp. 1311–1314, Aug. 2015.

[TP+14] B. Tepekule, A. E. Pusane, H. B. Yilmaz, and T. Tugcu, "Energy efficient ISI mitigation for communication via diffusion," in *Proc. IEEE Int. Black Sea Conf. on Communications and Networking (BlackSeaCom)*, Odessa, Ukraine, May 2014, pp. 33–37.

[TL+18] N. Tuccitto, G. Li-Destri, G. M. L. Messina, and G. Marletta, "Reactive messengers for digital molecular communication with variable transmitter–receiver distance," *Physical Chemistry Chemical Physics*, vol. 20, no. 48, pp. 30 312–30 320, Nov. 2018.

[Tüc04] M. Tüchler, "Design of serially concatenated systems depending on the block length," *IEEE Trans. Commun.*, vol. 52, no. 2, pp. 209–218, Feb. 2004.

[TH02] M. Tüchler and J. Hagenauer, "EXIT charts of irregular codes," in *Proc. Conf. on Inform. Sciences and Systems (CISS)*, Princeton University, NJ, USA, Mar. 2002, pp. 748–753.

[UK+18] H. Unterweger, J. Kirchner, W. Wicke, A. Ahmadzadeh, D. Ahmed, V. Jamali, C. Alexiou, G. Fischer, and R. Schober, "Experimental molecular communication testbed based on magnetic nanoparticles in duct flow," in *Proc. IEEE Signal Processing Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Kalamata, Greece, Jun. 2018, pp. 1–5.

[vWW+16] M. van Wolferen, A. Wagner, C. van der Does, and S.-V. Albers, "The archaeal Ced system imports DNA," *Proc. of the National Academy of Sciences*, vol. 113, no. 9, pp. 2496–2501, Feb. 2016.

[VP+11] G. Von Maltzahn, J.-H. Park, K. Y. Lin, N. Singh, C. Schwöppe, R. Mesters, W. E. Berdel, E. Ruoslahti, M. J. Sailor, and S. N. Bhatia, "Nanoparticles that communicate in vivo to amplify tumour targeting," *Nature materials*, vol. 10, no. 7, pp. 545–552, Jun. 2011.

[WC99] S. Walczak and N. Cerpa, "Heuristic principles for the design of artificial neural networks," *Information and Software Technology*, vol. 41, no. 2, pp. 107–117, Jan. 1999.

[WH+15] X. Wang, M. D. Higgins, and M. S. Leeson, "Distance estimation schemes for diffusion based molecular communication systems," *IEEE Commun. Lett.*, vol. 19, no. 3, pp. 399–402, Mar. 2015.

[WA+19] W. Wicke, A. Ahmadzadeh, V. Jamali, H. Unterweger, C. Alexiou, and R. Schober, "Magnetic nanoparticle-based molecular communication in microfluidic environments," *IEEE Trans. Nanobiosci.*, vol. 18, no. 2, pp. 156–169, Apr. 2019.

[WS+18] W. Wicke, T. Schwering, A. Ahmadzadeh, V. Jamali, A. Noel, and R. Schober, "Modeling duct flow for molecular communication," in *Proc. IEEE Global Communications Conf. (GLOBECOM)*, Abu Dhabi, United Arab Emirates, Dec. 2018, pp. 206–212.

[Wic95] S. B. Wicker, *Error Control Systems for Digital Communication and Storage*. Prentice Hall Englewood Cliffs, 1995, vol. 1.

[WB99] S. B. Wicker and V. K. Bhargava, *Reed-Solomon Codes and their Applications*. John Wiley & Sons, 1999.

[WB+05] S. G. Wilson, M. Brandt-Pearce, Q. Cao, and M. Baedke, "Optical repetition MIMO transmission with multipulse PPM," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 9, pp. 1901–1910, Sep. 2005.

[WO03] K. K. Wong and T. O'Farrell, "Spread spectrum techniques for indoor wireless IR communications," *IEEE Trans. Wireless Commun.*, vol. 10, no. 2, pp. 54–63, Apr. 2003.

[XY+03] Y. Xia, P. Yang, Y. Sun, Y. Wu, B. Mayers, B. Gates, Y. Yin, F. Kim, and H. Yan, "One-dimensional nanostructures: Synthesis, characterization, and applications," *Advanced Materials*, vol. 15, no. 5, pp. 353–389, Mar. 2003.

[YC+18] H. Yan, G. Chang, Z. Ma, and L. Lin, "Derivative-based signal detection for high data rate molecular communication system," *IEEE Commun. Lett.*, vol. 22, no. 9, pp. 1782–1785, Sep. 2018.

[YK+84] Y. Yasuda, K. Kashiki, and Y. Hirata, "High-rate punctured convolutional codes for soft decision Viterbi decoding," *IEEE Trans. Commun.*, vol. 32, no. 3, pp. 315–319, Mar. 1984.

[YC+12] P. C. Yeh, K. C. Chen, Y. C. Lee, L. S. Meng, P. J. Shih, P. Y. Ko, W. A. Lin, and C. H. Lee, "A new frontier of wireless communication theory: Diffusion-based molecular communications," *IEEE Trans. Wireless Commun.*, vol. 19, no. 5, pp. 28–35, Oct. 2012.

[YH+14] H. Yilmaz, A. Heren, T. Tugcu, and C.-B. Chae, "Three-dimensional channel characteristics for molecular communications with an absorbing receiver," *IEEE Commun. Lett.*, vol. 18, no. 6, pp. 929–932, Jun. 2014.

[ZL+18] H. Zhai, Q. Liu, A. V. Vasilakos, and K. Yang, "Anti-ISI demodulation scheme and its experiment-based evaluation for diffusion-based molecular communication," *IEEE Trans. Nanobiosci.*, vol. 17, no. 2, pp. 126–133, Apr. 2018.

[ZZ14] M. Zhang and Z. Zhang, "An optimum DC-biasing for DCO-OFDM system," *IEEE Commun. Lett.*, vol. 18, no. 8, pp. 1351–1354, Aug. 2014.

[ZR02] R. S. Zucker and W. G. Regehr, "Short-term synaptic plasticity," *Annual Review of Physiology*, vol. 64, no. 1, pp. 355–405, Mar. 2002.