Macroscopic Air-Based Molecular Communication using Fluorescent Dyes

Dissertation

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

vorgelegt von
Sunasheer Bhattacharjee

Kiel, 2023

Tag der Einreichung: 10.11.2023
Tag der Disputation: 24.04.2024

Berichterstatter: Prof. Dr.-Ing. Peter Adam Höher,
Christian-Albrechts-Universität zu Kiel

Assoz. Univ.-Prof. DI Dr. Werner Haselmayr,
Johannes Kepler Universität Linz

Declaration

I hereby declare, that I have produced the dissertation *Macroscopic Air-Based Molecular Communication using Fluorescent Dyes* independently with my supervisor's guidance and without improper external assistance and that I have identified all word-for-word quotations of other authors, as well as comments based closely on other authors' ideas, and I have listed the relevant sources. Furthermore, this dissertation has not been, partially or completely, submitted to any other university or institute in the context of an examination procedure. I declare that, the following work has been written in compliance with the rules of good scientific practice of the German Research Foundation. I also confirm that my academic degree has never been withdrawn.

Kiel, 10.11.2023

Sunasheer Bhattacharjee
Gurlittstraße 1
24106 Kiel

Preface

All experiments concerning this dissertation have been carried out at the Chair of Information and Coding Theory, Christian-Albrechts-Universität zu Kiel. First and foremost, I would like to express my heartfelt gratitude to my supervisor Prof. Dr.-Ing. Peter Adam Hoeher, who has been a pillar of strength and support, both professionally and personally. His style of providing absolute freedom to his research assistants in terms of realizing their scientific ideas helped me achieve my goals consistently, especially during the testing times of the pandemic. I would also like to thank my co-referee Assoz. Univ.-Prof. DI Dr. Werner Haselmayr for his valuable time and interest in reviewing my work. Furthermore, I extend my thanks to Prof. Dr. Jan Steinkühler and Prof. Dr. Hermann Kohlstedt for their willingness and time to form the examination committee.

Special shout-out to my colleagues at workplace, Sami Alkubti Almasri, Dr.-Ing. Martin Damrath, Ali Derakhshandeh, Andrej Harlakin, Maurice Hott, Nils Johannsen, Dr.-Ing. Adrian Krohn, Julius Maximilian Placzek, and Max Schurwanz. They always offered me a conducive and a pleasant environment, which not only helped me to carry out my scientific research, but also gave me the opportunity to unwind myself outside university by participating in various fun activities. I am further grateful to my former students, who also immensely contributed towards taking this research forward.

I would like to convey my special gratitude to Dr.-Ing. Martin Damrath, who introduced me to the field of molecular communication, and as a mentor, has been very instrumental in shaping my career as a researcher. His insightful comments and feedback while proofreading this dissertation helped me to improve its quality. My sincere appreciation also goes to Sami Alkubti Almasri for being a guide and a friend when it came to explaining the concepts of Information Theory & Coding, and in general, life in Germany as a foreigner.

Last but not least, I would like to thank my parents Sushobhan Bhattacharjee and Banashree Bhattacharjee, who have always supported me in realizing my dreams. I am, because they are. Finally, I would like to thank my friends Alok, Debayan, Jayasmita, Kallol, Krishna, Saikat, and Vijender for being an inalienable part of my bubble.

Abstract

The work presented in this dissertation explores the possibility of molecular communication to be a viable alternative to radio wave-based wireless communication, where deployment of infrastructure using traditional wired and wireless communication proves to be challenging. Any industry-specific molecular communication system application, prior to its standardization, requires a framework to encompass various aspects of communication theory. This dissertation aims at achieving the goal through the design and development of a macroscopic air-based molecular communication testbed, and analyzing it from the viewpoint of communication engineering. To this end, a new air-based molecular communication testbed is developed based on the principle of optical detection of the information carriers in the form of water-based fluorescent dye solutions, which outperforms many existing macroscopic molecular communication testbed designs in terms of the data transmission distance and the data transmission rate. The end-to-end system model of the testbed is considered in order to define the analytical model expression comprising of the system impulse response and the noise process, incorporating the effects of air drag and gravitational forces. The transmission algorithms include various modulation schemes exploiting concentration (such as on-off keying), time (such as pulse-position modulation), color of the dye (such as molecular shift keying) as well as the spatial domain (such as spatial shift keying) in order to achieve high data transmission rates. Equalization and detection algorithms are employed in order to mitigate the problem of inter-symbol interference and improving the bit error rate (BER) for long range communication in the process. A comparison of the resource requirements with respect to the implemented modulation schemes are also analyzed. Additionally, transmission errors are alleviated by the incorporation of block codes, convolutional codes, line codes, and spatial codes. Emphasis is particularly laid on the implementation of low-complexity codes that require less resources. The BER analyses of all these aspects are performed based on the measurements from the testbed trials as well as by means of Monte Carlo simulations based on the analytical model. Finally, a possibility to improve the data transmission rates further by modifying the transmitter and the receiver side hardware is discussed.

Kurzfassung

Die in dieser Dissertation vorgestellte Arbeit untersucht die Möglichkeit, molekulare Kommunikation als eine brauchbare drahtlose Alternative in Umgebungen, in denen herkömmliche drahtgebundene und drahtlose Kommunikation herausfordernd ist, zu etablieren. Jede industrielle Anwendung eines molekularen Kommunikationssystems erfordert vor ihrer Standardisierung einen Rahmen, der verschiedene Aspekte der Nachrichtentechnik umfasst. Diese Dissertation zielt darauf ab, diesen Rahmen durch den Entwurf und die Entwicklung eines makroskopischen luftbasierten molekularen Kommunikationstestbeds zu schaffen und es aus nachrichtentechnischer Sicht zu analysieren. Zu diesem Zweck wird ein neues luftbasiertes molekulares Kommunikationstestbed entwickelt, das auf dem Prinzip der optischen Detektion der Informationsträger in Form von wasserbasierten Fluoreszenzfarbstofflösungen beruht und viele bestehende makroskopische molekulare Kommunikationstestbeds in Bezug auf die Datenübertragungsdistanz und die Datenübertragungsrate übertrifft. Das End-to-End-Systemmodell des Testbeds wird betrachtet, um das analytische Modell zu definieren, das die Systemimpulsantwort und das Rauschen umfasst und die Auswirkungen des Luftwiderstands und der Gravitationskräfte einbezieht. Die betrachteten Übertragungsalgorithmen umfassen verschiedene Modulationsverfahren, die die Konzentration (z. B. *On-Off Keying*), die Zeit (z. B. *Pulse-Position Modulation*), die Farbe des Farbstoffs (z. B. *Molecular Shift Keying*) und die räumliche Position (z. B. *Spatial Shift Keying*) nutzen, um hohe Datenübertragungsraten zu erzielen. Entzerrungs- und Detektionsalgorithmen werden eingesetzt, um dem Problem der Intersymbolinterferenz entgegen zu wirken und dabei die Bitfehlerrate (BER) für die Übertragung über große Entfernungen zu verbessern. Dabei wird auch ein Vergleich des Ressourcenbedarfs der implementierten Modulationsverfahren analysiert. Darüber hinaus werden Übertragungsfehler durch die Verwendung von Blockcodes, Faltungscodes, Leitungscodes und räumlichen Codes reduziert. Ein besonderes Augenmerk liegt dabei auf der Implementierung von Codes mit geringer Komplexität, die weniger Ressourcen benötigen. Die BER-Analysen all dieser Aspekte werden auf der Grundlage von Testbedmessungen sowie mit Hilfe von Monte-Carlo-Simulationen unter Berücksichtigung des analytischen Modells durchgeführt. Abschließend wird eine Möglichkeit zur

weiteren Verbesserung der Datenübertragungsraten durch Modifikationen der Hardware sowohl auf der Sender- als auch auf der Empfängerseite diskutiert.

Contents

1

Introduction

1.1 A Brief History of Communication

The term communication traces its roots back to the Latin word *communicare*, which means "to share" or "to be in relation with" [Pet12]. The word also shares a direct etymological relation with the concepts of "common", "commune", and "community" that form the basic foundations of the societal fabric existent amongst all forms of living organisms inhabiting this planet [Cob08]. Communication forms an integral part of the lives of the sentient beings, especially that of humans who are always involved in the exchange of information, ideas, thoughts, etc. The individuals involved in the process of communication are most often separated from one another over large spatial distances. However, communication process is also witnessed amongst non-sentient beings. For example, viruses and bacteria exchange genetic information through the process of transduction [SS$^+$90]. Similarly, plants communicate with each other through the exchange of volatile organic compounds (VOCs) in the form of various released gases [CC$^+$18], and surprisingly, through sound waves too [GR13]. In a nutshell, communication is the process of sending messages with useful information from one entity to another, which are separated through physical distance and space. The goal of a successful communication process is to ensure that the entity receiving the messages is successfully able to decipher the exact idea and intention the sender is trying to convey.

Instinctively, humans are inquisitive creatures with an insatiable appetite to learn more, which is followed by their perpetual desire to preserve and share this attained knowledge amongst themselves and with the future generation through information storage and exchange. This is not just limited to the events taking place on the planet, but also the

occurrences taking place outside the planet, so much so that telescopes such as the *James Webb Space Telescope* have been launched into the outer space to gather information about the origin of the Universe, and the existence of other life supporting exoplanets outside the Solar System [Kal18]. The information acquired by these telescopes are then relayed back to the observatories on Earth through wireless transmission of the electro magnetic (EM) signals for further transformation, processing, and interpretation of the useful data. This complex transmission link forms a part of a very sophisticated communication system spanning over very large distances.

The communication capabilities the human civilization currently possesses, has not been accomplished and developed overnight. Over the last few millennia, the humans have witnessed a steady evolution in the area of communication technology. From giving out war cries during the early Stone Ages to establishing deep-space communication links with satellites and probes in modern times, the field has in the true sense come a long way. During the testing times of the *Severe Acute Respiratory Syndrome Coronavirus 2* (SARS-CoV-2) pandemic, when the world was literally brought to a standstill, it was with the help of communication technology that people remained virtually connected through smartphones and the Internet, despite having physical barriers and restrictions. This also allowed most social and economic activities to continue and new business opportunities to flourish, which otherwise would not have been possible in the past.

In the ancient times, smoke signals, flags etc., served as some of the most popular information exchange methods, which were prevalent amongst the sailors, traders and military armies over distances within the line-of-sight range. This slowly evolved into pigeon posts, where birds would be flown to send information over very large distances well beyond the line-of-sight range. In the nineteenth century, a revolution was witnessed in the field of communication with the advent of electricity, which gave rise to more sophisticated forms of long-distance communication using telegraph lines and telephones, making their presence felt for several generations due to their reliability and robustness. This, however, was a prelude to the Information Age. By the turn of the twentieth century, many endeavors in the field of computing and semiconductors had already been made, but a whole new research domain in the field of information and communication technology was still in the offing. This was not until the year 1948 when an American mathematician and electrical engineer named Claude Elwood Shannon published his ground-breaking research work in [Sha48], officially announcing the onset of the "Digital and Information Age", with the information transfer taking place in the form of digital bit streams comprising of binary data symbols "0" and "1". This eventually paved way for the telecommunication fields such

as wireless communication, satellite communication, optical fiber communication, etc. In modern times, these technologies find their ways into telephone, radio, television, computer and mobile phones, with the information transmission media primarily being wired, optical or wireless.

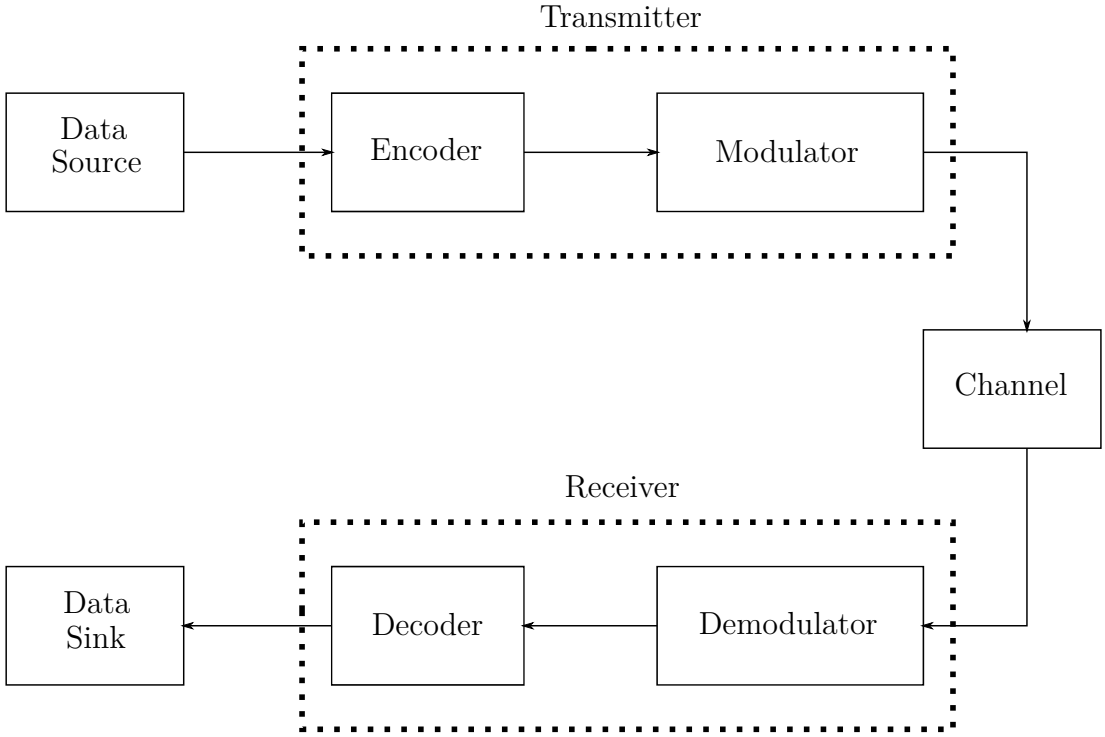


Figure 1.1: A basic model of a digital communication system [Sha48].

The most basic digital communication model can be described by Figure 1.1. The information in the form of audio, video, picture, etc. to be transmitted to the desired recipient is first transformed into a sequence of bits comprising of binary data symbols at the data source. The bit sequence is then divided into smaller groups to form the info words. Each info word is then passed through the encoder adding redundant bits to the info word to form a code word. This strategy is applied in order to minimize the effects of error that may arise due to various unwarranted effects inside the transmission channel, affecting the detection of the originally transmitted bits at the receiver side. The encoder is then followed by a modulator that converts a code word into a time-varying measurable physical quantity, for example, current, voltage, light intensity, magnetic field, etc., making it compatible with the channel and the receiver system over which the data is to be transferred and ultimately detected. The modulated data in the form of a time-varying physical quantity is transmitted over the channel. The channel introduces undesirable effects such as noise and attenuation, distorting the information-carrying physical quantity

that results in erroneous transfer of the information. At the receiver side, the information-carrying physical quantity is demodulated to get the encoded bit sequence comprising of binary data symbols, which is then finally decoded to get back the original bit sequence containing the information. This is then followed by the data sink, which serves as the final destination of the message [Sha48]. This particular communication model serves as the foundational basis to the work that has been presented in this dissertation.

1.2 Modes of Communication

The information transfer between a source and a receiver can be achieved over various modes and means. However, the two most commonly used modes can broadly be classified as follows:

- *Wired Communication* - Wired communication involves a perceivable physical connection between the transmitter and the receiver, often with the help of a wire. The range of communication can span over several meters to kilometers, but is largely restricted to terrestrial and underwater applications. Telegraph, telephone, cable television, Ethernet-based local area network (LAN), optical fiber communication (OFC), etc., are some known communication methods that use a wired connection between the transmitter and the receiver. Telegraphs for example, used electrical signals over transmission lines to transmit messages over large distances. The early telephones also relied on wired transmission systems, but with a higher information-carrying capacity. Later advancements in the field along with the use of twisted pair insulated cables increased the message quality further. The Internet and LAN connections today use high speed unshielded twisted pair cables [TB06] to exchange information between the connected devices. Cable televisions on the other hand use coaxial cables [LF04], which efficiently transmit electrical signals of very high bandwidth. These cables are also capable of providing high immunity to interference from other EM noise sources. OFC [Agr12] developed as a communication methodology some four decades ago has also seen a wide range of utility across many fields. This technology makes use of the EM wavelength in visible spectrum for information and data transfer over a glass fiber called the waveguide. This method of communication ensures an unprecedented amount of speed, low error rates, higher bandwidth and information-carrying capacity, all at the same time.

- *Wireless Communication* - Wireless communication technology, however, achieves

successful information exchange between a transmitter and a receiver without establishing a physical connection through electrical conductors or optical waveguides. Some well-known methods of wireless communication include radio wave communication, satellite communication, acoustic communication, visible light communication (VLC), etc. Radio wave communication finds its application in radios, televisions with antenna receptors, wireless LAN, etc., where EM radiation of very high bandwidth is used for long range communication. The standards for the EM radiation spectra is set by the International Telecommunication Union [Lya16]. Very long distance communication through radio waves is mainly propelled by the ionosphere in the Earth's atmosphere [JB68]. Satellite communication is another widespread method of wireless message exchange involving a satellite acting as both the transmitter and the receiver [Elb08]. A wireless line-of-sight communication link is established with telephones, televisions, probes or with other satellites by sending a very high energy narrow beam microwave, radio wave, or visible light at times. Acoustic communication is also a popular method of wireless communication, which is used in underwater environments where radio waves and visible light experience heavy attenuation [SP09]. Information is relayed to hydrophones by using a sound producing transmitter. Sound navigation and ranging is another acoustic communication technique, which is employed in order to detect objects underwater and also to gauge distances between them. Another variant of optical communication namely VLC evolved with time as a potential wireless communication method. Fluorescent lamps and light emitting diodes are used for both illumination as well as data transmission and localization purposes [Hoe19]. The transmitted light is then received by a photodiode, which converts it into useful information.

Although the advantages of using EM waves for communication purposes are many, they do come with certain disadvantages that are listed as follows:

- High attenuation in closed bounded environments such as inside of a network of pipelines and tunnels [NE$^+$13; GM$^+$15]
- Unreliable and frail connectivity between transmitter and receiver in underwater environments with non-uniform salinity over large distances [SS$^+$22; KP$^+$20]
- Challenging use in micro- and nano-scale applications due to the constraint in the form of antenna size to the wavelength ratio, in addition to supplying necessary power to drive the antenna operation [SI$^+$12; Whe47]

- Limited availability of spectra with high interference [SK+12; WS+21]
- Biocompatibility and suitability for medical applications is an issue [HM+00; ZH05]

A promising alternative to EM wave-based communication has been gaining traction in recent times, which is termed as molecular communication. As the name suggests, information transfer in this method is performed through the release of molecules/particles by the transmitter, which propagate through an air-/fluid-based transmission channel to reach the receiver. At the receiver side, these molecules/particles are then detected in order to successfully decode the transmitted message [NE+13]. Scientists believe that molecular communication is arguably the oldest mode of communication, which was first used by primitive single-celled organisms long before the existence of complex multi-cellular organisms such as plants and animals [VW+16]. Molecular communication can find its application in both micro-scale (in the range of nm to µm) as well as macro-scale (in the range of cm to m) domains, the details of which are discussed in the subsequent sections of this chapter.

1.3 A Synopsis on Molecular Communication

The most ubiquitous form of communication providing information access to the entire world is the EM wave-based communication. Although the benefits of communication using EM waves are many, there are still some application areas where their performance suffers greatly, especially, when it comes to their use in hostile conditions involving pipeline and tunnel networks, salt-water environments, etc. Moreover, their utility is physically not suitable for small-scale applications due to their performance being dependent on attributes such as the antenna size to the wavelength ratio, as also mentioned in Section 1.2 in detail. Therefore, there has been a growing demand in the research community to look for alternative means of communication offering solutions to the stated problems. One such favorable alternative is molecular communication.

1.3.1 Molecular Communication in the Nature

Molecular communication is an information transfer methodology that primarily uses molecules, particles, ions, etc. to send information from a source to a destination that are separated from one another with some physical distance or space [NE+13]. The history of this type of communication dates back several millions to billions of years into the past, where it was the sole technique of information transfer among the living organisms.

For survival strategies and evolutionary processes, the nature generally preferred slow yet energy-efficient and reliable way of communication in comparison to high information transfer rate methods such as the use of EM waves. For instance, free-living cells such as *Archaea*, a type of bacteria populating the planet some 2.5 to 4 billion years ago during the *Archaean Eon*, is believed to have communicated with one another through the process of quorum sensing [RK17]. Quorum sensing refers to the communication process in bacteria, whereby the information-carrying molecules such as N-(3-oxoacyl) homoserine lactone (3-oxo-AHL), N-(3-hydoxyacyl) homoserine lactone (3-hydroxy-AHL), etc., allow a single cell to sense the number of bacteria present in the environment so that the population makes a collective and coordinated response [DC$^+$07]. This indicates the presence of social behavioral interactions among bacteria which include swarming motility, conjugal plasmid transfer, antibiotic resistance, biofilm maturation, and virulence. Other unicellular organisms such as yeast cells of haploid type also use molecular communication to perform certain specific social activities especially in the process of reproduction. The mating process in yeasts is stimulated by the presence of pheromone, which binds to the cell receptors. This leads to the activation of a Heterotrimeric G-protein to indicate its mating readiness to other yeast cells of the opposite sex [CA18; AB$^+$15]. In multicellular systems, the intra and/or intercellular signaling exist in their more complex forms, and are categorized based on their specific tasks and functionalities they undertake [KW$^+$16]:

- *Autocrine Signaling* - Cell signaling process that involves the secretion of chemical messengers that bind to the receptors of the same cell, leading to physical changes in the cell [Pan07].
- *Intracrine Signaling* - Signaling process in cells, where they stimulate themselves to produce chemical messengers that act within the cells, managing intracellular events [WC08].
- *Paracrine Signaling* - Cellular signaling technique, whereby messenger molecules called paracrine factors travel short distances to induce behavioral changes in nearby cells [KB$^+$21].
- *Endocrine Signaling* - Signaling process very similar to paracrine signaling, where messenger molecules travel larger distances through the circulatory system of the body in order to induce a behavioral change in distant target cells [CH07].

Several examples of cellular signaling are observed inside the bodies of living organisms, especially humans. For instance, the transfer of calcium ions with variation in concen-

tration affects the muscle contraction, cell locomotion, hormonal secretion, and neural transmission [CL08]. As observed in neural transmission, where a neurotransmitter named acetylcholine is transported along the axon and released in such a way, so as to cause a response in the other neurons, muscle cells or gland cells with an acetylcholine receptor [TD$^+$13]. Similarly, ciliated airway epithilial cells communicate with each other by the release of inositol 1,4,5-triphosphate (IP_3). The diffusion of IP_3 molecules through gap junctions to adjacent cells results in the release of calcium ions in each adjacent cell facilitating intercellular communication [SM$^+$05].

Molecular communication is also used by more complex organisms such as plants and animals for communicating among themselves over much larger distances. VOCs such as benzyl alcohol, benzyl benzoate, methyl salicylate, o-cymene, and many other compounds help to attract pollinators like bees, facilitating the process of cross-pollination and reproduction in plants. Concurrently, *Anemophily* or the process of pollen transportation through wind movement can also be perceived in the sense of molecular communication, as these pollens carry the genetic information of one plant (source) to another plant (receiver), where they eventually get settled in order to further the process of reproduction, ensuring the continuation of species. Additionally, in some cases, VOCs also play a completely different role from the process of reproduction. Damaged plants also release alarm pheromones and VOCs in order to transmit stress signals to undamaged plants nearby. This alerts the plants to a variety of threats, and allow them to strengthen their defense mechanism [UK$^+$12]. Insects such as ants have signaling techniques, which make use of released pheromones in order to coordinate activities such as nest defense and foraging [CR$^+$19]. For example, conveying the message about perceived threat or quality of food and its availability is achieved by producing a higher concentration trail [RC09]. Bees also possess an elaborate array of pheromone-based molecular communication signaling techniques that trigger the activities involving reproduction, foraging, defense, integration of colony, etc [BC14]. Certain marine animals such as minnows and catfish release alarm pheromones when injured in order to warn the nearby fish of any impending danger [CR06]. More complex living organisms such as mice have their olfactory system designed in a way, which avoids inbreeding among the close relatives [ST$^+$07]. Interestingly, humans also possess an augmented olfactory system, which is designed to sense the pheromones released by the person of the opposite sex, especially by women during their ovulation phase acting as a potential aphrodisiac [GF$^+$05].

1.3.2 Synthetic Molecular Communication

Taking insightful knowledge and inspiration from these naturally occurring phenomena, scientists and researchers from around the globe are currently endeavoring to replicate molecular communication in their laboratories and realize successful information transfer for a variety of potential applications in various disciplines. One important aspect of molecular communication is being looked into that explores its possibility of integration with the future generation of wireless networks such as 6G and beyond [HS+19]. The advantages of molecular communication over classical communication systems are many. Since it is already seen as the foundational basis of communication technique in living organisms, biocompatible systems incorporating molecular communication can be realized without much effort. Furthermore, the transmission is radiation-free, which has always been a topic of debate when it comes to the health implications of radio wave communication on humans and other living organisms [HH03]. Additionally, when realizing communication at a microscopic level, the introduction of a dedicated power supply for such small-scale applications is a problem due to space restrictions. This is where the use of molecular communication seems to be a plausible solution, as it offers a highly energy-efficient information transfer mechanism. For example, if the molecules are diffusing from the transmitter to the receiver, they use the environmental energy to propagate through the transmission channel. Therefore, the energy from the transmitter is only required for the release of molecules into the environment, but no additional energy is required for their movement through the channel. The molecules have the ability to propagate over long distances, especially in those environments that prove to be challenging even for EM wave-based communication. For example, in underwater information transmission applications, in environments with a network of pipelines and shaft systems, or in disaster management operations among mounds of debris, etc. However, molecular communication also comes with a lot of disadvantages. As the molecules are physical particles that need to be transported from the transmitter to the receiver, the speed of information transfer or the throughput gets affected, which particularly gets hindered over very large communication distances. When molecules get carried through the transmission channel by diffusion, the information-carrying signal gets attenuated with distance along with the additional spreading in time domain, leading to inter-symbol interference (ISI). These reasons result in errors in the transmission system. Moreover, it is the stochastic propagation of the information-carrying molecules that makes it an inconsistent and unreliable transmission system, as there is no guarantee that all the released molecules would eventually reach the receiver over an intended period of time. Unless there is a presence of favorable drift from the transmitter towards the receiver,

molecular communication systems also suffer in terms of their maximum transmission range.

Potential Applications of Microscopic Molecular Communication

There are many potential multi-scale applications of molecular communication that are envisaged by the research community, which could lead to its practical realization in the days to come. In the microscopic domain, for example, it is imperative for communication engineers to understand and draw inspiration from molecular communication in cell biology involving cell signaling techniques. The process involves interpretation, replication and manipulation of these signaling methods in a useful manner that suits the application in terms of the physical nature and scale of the environment. This knowledge is especially more helpful in those applications, where miniaturization of macroscopic transceivers to microscopic levels is extremely challenging [BA$^+$21; AB$^+$08]. However, the field of nanotechnology is striving to push the technological boundaries by ushering in the development of devices such as nanomachines and nanorobots in the nanometer scale range [Req03; AB$^+$08]. Subsequently, the fields of biotechnology, biomedical engineering and nanotechnology can be combined together for medical applications, one of the main driving motivations behind the conception of molecular communication [AA$^+$12]. For instance, the interaction between the cells, tissues, organs, and the control of the cell growth and division can be realized with the help of nanomachines [PM$^+$08]. They can also be engineered to be made more biocompatible such as in the form of a genetically modified bacteria or a synthetic cell performing more specialized but simple tasks [AF$^+$12; XY$^+$03; SM$^+$05]. However, due to their small size and limited access to energy, their performance can strongly be hindered. The problem can be mitigated by introducing more nanomachines in the form of a self-organizing swarm [CT04]. The nodes within the swarm can use energy-efficient means such as quorum sensing to communicate with one another [ES$^+$12; MB01]. Additionally, these swarm of nanomachines can be modified to form a nano-scale sensor network [ES$^+$12; AB$^+$08], serving both medical and military applications [NM$^+$12; YW$^+$16]. In the recent times, the scaling of these sensors have taken place to realize their utility in microscopic range [EG$^+$18; Bog08], for instance, in the design and development of biological circuits, which are based on chemical reaction networks and deoxyribonucleic acid transcription processes [ED$^+$19; MP$^+$18]. The use of artificial neural networks (ANNs) for micro-scale applications has been proposed in [BB$^+$11], where these ANNs could be interfaced with a nanomachine to efficiently perform molecular communication. The idea of nanomachines in literature has mostly existed as a theoretical concept, but their capabilities have already

been showcased in [GA+13], where glucose-mediated delivery of insulin is made possible through injection of nanonetwork. Similarly, in another work, systems have been constructed where synthetic biological and nanotechnological components communicate *in vivo* for the purpose of amplifying disease targeting mechanism [VP+11]. This has led to an important application of molecular communication using nanomachines in the area of targeted drug delivery [CM+17]. Targeted drug delivery can be performed by transporting the drug molecules at a specific location such as disease cells or tumors through nano-scale delivery capsules, and releasing them at highly localized sites [NS+14]. Tumor cells are known for releasing messengers as opposed to healthy cells. The nano-scale capsules are capable of identifying the messenger substance, and are able to bind themselves to the specific receptors of the tumor cells. The capsules then release the intended amount of drugs locally on the unhealthy cells, and reduce the potential hazards of targeting the healthy cells [AC04]. Addtionally, nanorobots are also proposed for the detection of other health complications such as brain aneurysm [CS+09]. Nanorobots are alternatives to nano-scale capsules that can be used to transport molecular payloads to cells [DB+12], which help alter the functionalities and behavior of the cells [EW12]. Ongoing research efforts have made it possible to develop these nanorobots in the form of a programmable bacteria replicating computational capabilities of nanodevices by maintaining biocompatibility [WK00]. In another method, flagellated magnetotactic bacteria are being explored, which can harness power from magnetic resonance imaging to power the transportation process [MS09]. In [CH+06; VP+11], it is shown that a swarm of nanomachines communicating with one another are far more effective at targeted drug delivery than non-communicating nanomachines. For example, when a nanomachine detects released molecules from tumor cells, it can share the information with other nanomachines in the vicinity, and coordinate better to cooperatively release drugs at these tumor cells. Molecular communication also sees potential in the field of tissue engineering, which includes the restoration, maintenance, and/or improvement of damaged tissues or entire organs in a living organism [GN02]. In tissue engineering, the stem cells are extracted from a subject, which are then cultured and returned back to replace the lost tissue. Molecular communication can be used to control the growth and division of these cells to form desired structures [NE+13]. Another interesting field where molecular communication can potentially find its use is in the enhancement of the immune system response of the human body [DG+12]. The introduction of artificial immune systems into the human body can be helpful in providing protection against dreadful infectious agents such as parasites, virus, bacteria, and fungi [TK+04], especially at a time when the world is still reeling under the effects of SARS-CoV-2 pandemic. Molecular communication may

enable the communication and coordination of a group of bio-nanomachines in enhanced immune system by tracking the moving objects such as pathogens, and then notifying the external devices of the target location for further necessary action [ON+14]. Apart from the stated potential use cases of molecular communication, additional micro-scale applications include lab-on-a-chip systems [Par10], cell-on-chips devices [ES+06], point-of-care diagnostic chips [YE+06], bio-computers [Nob02], controlling and detecting chemical reactions [Dem06], monitoring and quality control [AB+08], and environmental control and preservation [HN+11].

Potential Applications of Macroscopic Molecular Communication

A plethora of applications in the macro-scale range are also anticipated, where molecular communication can be used effectively. The feasibility of the application of macroscopic molecular communication is more practical due to the availability of the basic communication system units in the market, especially in the form of detectors [McG20]. However, these applications are mostly in those areas, where the performance of the EM wave-based communication is restricted or suffers to a significantly large extent. For example, radio wave sensor networks are not very robust when it comes to infrastructure monitoring applications [SH+10; QG+14], especially in underground facilities and environments [RS+20]. Hence, molecular communication systems are very useful for transmission of information in those surroundings, which involve an extended network of pipelines and ducting systems, with a lot of bends and turns such as in the oil and gas industry [FY+16]. Based on the naturally occurring pheromone-based signaling in ants and moths [Ago92; AC16], a number of research works have come up to study and replicate these biological processes from the standpoint of macroscopic molecular communication in order to solve engineering problems [CG+09a; CG+09b; MD+12]. For example, chemical signaling in codling moth has been studied in the past, where pheromones are released to confuse the male moths for disrupting the mating process, resulting in better pest management and control [NB08]. Another important area of macroscopic molecular communication application is in the field of robotics [dAR+12; PR10]. Molecular communication between robots can be performed through judicious transmission and detection of certain VOC gases, with the help of a whole range of gas detectors that are readily available. More technologically advanced detectors in the form of electric noses can help establish more efficient molecular communication between the robots [FH+19]. An electronic nose is an artificial sensory device emulating the human olfactory system through pattern recognition algorithms, and is used to detect chemical odors and scents [GB99]. These electric noses are capable of

converting the detected odors into measurable electrical signals. Robots equipped with such sensory devices could then be deployed in harsh environments to perform disaster management and rescue missions through swarm navigation [MN$^+$10; AA$^+$17]. Another interesting application of macroscopic molecular communication can be found in the area of transmitting odor over digital platforms [HC$^+$03], especially at a time when the popularity of the virtual reality hardware is on the rise [Ste92]. These devices can successfully simulate audio and visual sensory inputs on digital platforms and media, however, simulating odor is still a challenge. Macroscopic molecular communication can also be used for applications related to the protection of the environment. For example, bacterial biofilms can be used to clean chemical-laden water that is discharged from manufacturing plants or to treat organic wastes [CS$^+$99; KO$^+$98]. Similarly, the principle of molecular communication can be applied to deploy large-scale biosensor networks in order to monitor the environments that may be contaminated with toxins and radioactive materials [NE$^+$13]. In future, it may be possible to integrate bio-nanomachines into silicon-based electrical devices such as mobile phones for the advancement of information technology [Fre99; Bai16]. These devices equipped with a biochip would be capable of analyzing various bio-marker-carrying agents such as human breath, sweat, saliva, blood, etc. through molecular communication [HM$^+$08]. The information regarding the health of the subject could then be directly relayed to the concerned health authorities during an emergency via traditional cellular networks. Since the start of the SARS-CoV-2 pandemic, many works tried to draw a parallel between macroscopic molecular communication and infectiology. The idea has been proposed for the first time in [KA$^+$19], where pathogen-laden human breath released from an infectious human is considered as the transmitter, while healthy human receiving the pathogens is considered to be the receiver. Furthermore, the study is extended in [KA$^+$20a], where virus spread through aerosol transmission is treated as a macroscopic molecular communication problem by incorporating silicon nanowire field effect transistor-based biosensors to model airborne pathogen transmission and detection. Additionally, an end-to-end system model is presented in [GA21b] for the spreading mechanism of infectious diseases by airborne pathogen transmission between two humans. The sprayed pathogen-laden cough/sneeze droplets in the model are also subjected to the effects of gravity and buoyancy to predict the trajectory of these droplets. These methods and analyses from the perspective of macroscopic molecular communication may prove to be useful in tackling future pandemics.

Gauging the potential of the macroscopic molecular communication, various research groups around the world are working together for investigating its possible industry approved applications. One such example is the project “Makroskopische Molekulare Kom-

munikation (MAMOKO)", funded by the German Federal Ministry of Education and Research [MAMOKO]. The project specifically aims to investigate and perform feasibility studies on macroscopic molecular communication from the perspective of those industrial applications that extensively make use of inflammable and/or explosive liquids/gases.

1.4 Motivation and Contributions

As evident from the discussions in the previous sections, molecular communication has many useful applications across various disciplines needing combined scientific efforts from the community of engineers, physicists, chemists, biologists, physicians, etc., for its successful implementation. The field especially sees immense potential in the microscopic domain for the medical applications revolving around targeted drug delivery, lab-on-chip systems, etc. Therefore, over the years, the research endeavors have mostly remained focused on comprehending the underlying principles that govern the functioning of molecular communication in the micro-scale domain. However, molecular communication in the macro-scale domain also has a wide variety of potential applications, but the field largely remains unexplored. Moreover, as opposed to microscopic molecular communication, macroscopic molecular communication is not limited by its computational power, allowing it to implement high data rate-achieving transmission algorithms by simultaneously dealing with the channel issues that may arise in the process. As research on molecular communication is still in its nascent stage, many theoretical works in the micro-scale domain have come up in the recent times attempting to model the propagation of the information-carrying particles from the transmitter to the receiver through the process of diffusion in three-dimensional space. The diffusion process is often modeled with the help of Brownian motion, which to a large extent, explains the movement of information-carrying particles through space in every direction. However, for a successful transmission of information in the macro-scale range, the diffusion process is not enough as it is a very slow process, affecting the propagation speed and maximum achievable throughput of the molecular communication system. Therefore, additional assistance in the form of advection is necessary for a faster transmission of data in a particular direction (from the transmitter towards the receiver). Moreover, the medium over which the transmission of information occurs, plays a crucial role in determining the data transmission rates in the macro-scale domain. Additional factors influencing the throughput of the macroscopic molecular communication system also include the rate at which the information is generated at the transmitter side as well as the sensitivity and responsitivity of the receiver.

Any industry-specific system application of macroscopic molecular communication prior to its standardization would require a framework to encompass various aspects of communication theory. These include the system impulse response analysis and system modeling (taking into account the combined effects of the transmitter, the transmission channel, and the receiver), implementation of high data rate achieving modulation schemes, efficient detection algorithms at the receiver side for achieving data transmissions at very low error rates, channel coding techniques for effective error detection and/or correction, etc. Therefore, the design and development of experimental platforms and testbeds becomes a crucial starting point to perform these tests and analyses in order to understand the system better (for example to attain the highest possible data transmission rate/reliable communication over a certain length of distance), and also to incorporate the aforementioned functionalities more effectively. Furthermore, the testbed lays the foundation for far-reaching innovations and added value. For example, a simulation environment can be set-up which emulates the properties of the testbed. The simulation environment could help bring in more flexibility into the system by offering the choice of seamless addition/alteration of various physical parameters governing the communication process. Towards this goal, this dissertation attempts to investigate macroscopic air-based molecular communication, where the released information-carrying particles from the transmitter side travel through the air medium in order to reach the receiver. The information is decoded by performing optical detection of the information-carrying particles that exhibit fluorescence property. Based on the said principle, a new molecular communication testbed is designed. Additionally, the testbed system model is developed through experimental measurements by considering the effects of the velocity of the released information-carrying particles, the influence of drag and gravitational forces on them, and the noise process originating from various sources inside the testbed, as the distance is increased between the transmitter and the receiver. Various modulation schemes are implemented and studied that exploit not just a single degree of freedom in the form of concentration of the released information-carrying particles, but also additional degrees of freedom in the form of time, fluorescent dye colors and the spatial regions populated by the water-based solution droplets of the information carriers in order to explore the possibility of increasing the data transmission rates. For long-distance communication, the macroscopic air-based molecular communication system also brings in a lot of challenges in the form of signal attenuation and a strong influence of ISI. In addition, the system is prone to amplitude-dependent noise leading to greater unreliability in data transmission. Therefore, these problems are tackled by incorporating and analyzing different detection algorithms and channel equalization techniques for further improving the

reliability of data transmission. Additionally, the investigation of forward error correction (FEC) channel codes, line codes, and spatial codes is carried out, and their potential to enhance transmission reliability is evaluated. Finally, an improvement of optical detection hardware is proposed to explore the possibility of further increasing the data transmission rates.

1.5 State-of-the-Art

Testbed Design In order to make molecular communication a feasible option for applications in various industries, many researchers have contributed to testbed designs at the macroscopic domain considering various types of transmitters, transmission channels, and receivers. This can also be seen as an attempt to complement the already existing theoretical works on molecular communication as a proof of principle in both micro- and macro-scale domains. A detailed survey concerning the practical implementations of molecular communication testbeds can be found in [JA$^+$19; LB$^+$23a; LB$^+$23b]. One of the earlier works concerning the testbed design is presented in [FG$^+$13], where isopropyl alcohol is used as the information carrier between the transmitter consisting of an electronic sprayer and the receiver comprising of a metal-oxide alcohol sensor. The information-carrying droplets released by the sprayer are propelled with the help of a tabletop fan, which propagate through the open air space to ultimately reach the receiver. The channel modeling is based on the advection-diffusion model as presented in [FK$^+$14]. However, nonlinearity, high ISI, and saturation level attained by the receiver proved to be major performance deterrents to the system. A testbed for in-vessel communication simulating portion of the cardiovascular system or a complex network of pipes in industrial complexes and city infrastructures is shown in [FP$^+$17]. The presented work makes use of acids and bases as the information carriers, where peristaltic pumps at the transmitter side release acids and bases, which propagate through a tube with the help of flow-assisted filtered water. A pH probe is used at the receiver side, where the pH value is converted into an electric potential. The system faces challenges in the form of nonlinearity and a long memory due to chemical reactions, apart from lacking a channel model of its own [FG16]. Similarly, a more biocompatible approach is adopted through the use of magnetic nanoparticles in [UK$^+$18]. The transmitter side comprises of a peristaltic pump connected to a syringe that holds the information carrier in the form of superparamagnetic iron oxide nanoparticles (SPIONs). The syringe is connected to a Y-connector with another tube providing a background flow of water. The end of the Y-connector joins the propagation channel tube, which joins the magnetic

susceptometer coil acting as the receiver. The system model is based on the laminar flow of fluid inside a duct. A fluid-based molecular communication testbed, however, offers low data transmission rates due to high viscosity. With respect to air-based molecular communication, many testbeds developed by the researchers utilize VOC and chemical vapor for transmitting the information that are also based on the advection-diffusion model. For example, in [GM$^+$17], the transmitter comprises of an odor emitter that controls the evaporation of liquid chemical analytes and their diffusion in a carrier gas stream through a tube-like propagation channel. The VOC particles are detected with the help of a portable membrane inlet mass spectrometer at the receiver side. However, the response time of the mass spectrometer-based detector is slow, which affects the overall data transmission rates. Therefore, on similar lines as an alternative, a high-speed chemical vapor communication testbed is presented in [OK$^+$18], where a solenoid valve at the transmitter side controls the release of solvent vapors carrying the information into an enclosed flow tube propagation channel. The vapors are carried by a background airflow provided by a fan. The information-carrying particles are then detected by several photoionization detectors at the receiver side with a very high responsitivity. Many testbeds additionally, also exploit the fluorescence properties of the information carriers for optical-based detection at the receiver side. For instance, in [AE$^+$18], the information-carrying fluorescent particles are injected as plumes inside a fluid propagation channel with variable flow rates. At the receiver side, the fluorescent particles are detected through particle image velocimetry with the help of a camera or a submersible optical fluorometer. Various obstacles are also installed inside the propagation channel in order to emulate the real-life scenarios. Based on similar principles, planar laser-induced fluorescence is studied and presented in [AL$^+$19], which describes the advection diffusion-based channel model along with the quantification of the statistical noise distribution and the mutual information. Additionally, works such as [DD$^+$13; TL$^+$18] have also presented their experimental setups based on the camera-based fluorescence detection of microfluidic droplets and fluorescent substances in liquid medium. In spite of optical-based detectors offering high responsitivity, the fluid-based propagation channel restricts the system from achieving high data transmission rates, as previously stated. Apart from the synthetic molecular communication testbeds, several biological testbeds have also been designed over the years. For example, the experimental setup in [KA$^+$13] uses genetically engineered *Escherichia coli* bacteria producing a fluorescent response to a time-varying chemical stimulus. Similarly, in [GK$^+$19], these bacteria react to fluorescence by releasing protons into the environment. The concentration of protons is then measured by a pH electrode to produce information signals.

Modulation Techniques The information in molecular communication is embedded in molecules/particles based on their various properties namely quantity, type, release time, populated spatial regions, and a hybrid of any of these. The most basic and popular modulation scheme is the on-off keying (OOK), which represents a binary data symbol either by releasing or not releasing molecules from the transmitter. The first implementation of OOK for a diffusion-based molecular communication channel is presented in [MM$^+$10], which is later extended to molecular concentration shift keying (MCSK) in order to accommodate higher data symbol cardinality, represented by multiple concentration levels [KY$^+$11]. This scheme is inspired from amplitude shift keying, which is a well-known modulation scheme in digital communication theory. Another modulation scheme which is quite popular in the molecular communication research community is molecular shift keying (MoSK), embedding the information into the type of molecule that is being transmitted [KY$^+$11]. One advantage of MoSK over MCSK is that it is more resilient to the errors that originate due to ISI. In order to make the in-body nano communication feasible, isomer-based MoSK has been proposed in [KC13] that makes use of aldohexose isomers to modulate the information. Similarly, in [AK$^+$18], a modulation scheme is proposed based on the concept of isotopic isomerism, where information is encoded in the molecular isotopes. Practical implementation of type-based modulation schemes in testbeds has also witnessed significant progress in the recent times. For instance, in [FP$^+$17], a time-slotted binary communication system is implemented, where the information is carried by the pH signal by representing binary data symbols “0” and “1” by an acid and a base pulse, respectively. A droplet-based microfluidic system for molecular communication in [BF$^+$20] has been presented with a droplet detector that has the ability to characterize droplets based on their presence/absence, color as well as their size, which can potentially be utilized separately for data encoding purposes. Information is also encoded by using the temporal dimension. The first implementation of such modulation scheme is shown in [MM$^+$10], where the chemical concentration is varied in accordance with a sinusoidal signal of certain frequency, similar to the idea of frequency shift keying in radio wave communication theory. Pulse-position modulation (PPM) is proposed from the perspective of molecular communication in [GL$^+$11], where the information is represented by the position of a pulse within a transmission interval. Rate modulation is another interesting scheme proposed in the same work, where the information is represented by the number of pulses contained within a transmission interval. Furthermore, communication through silence (CTS) (also known as differential pulse-position modulation (DPPM)) is proposed in the same work, where the information is contained in the time interval between two consecutive pulses.

Modulation schemes such as OOK and CTS in the form of time-elapse communication (TEC) have also been implemented practically in a testbed setup with the use of genetically engineered bacteria [KA+13]. In the same work, TEC is further modified as smart TEC by adding error differentiation and differential coding techniques. In another instance, different time-based modulation schemes are implemented for a droplet-based microfluidic testbed, where information is encoded based on the presence/absence, distance and the size of the droplets, and a study is presented based on the information rates and error tolerances of these modulation schemes [HM+19]. Another degree of freedom is attained in the form of the spatial regions populated by the molecules/particles, which can be achieved by making use of multiple transmitters. Molecular space shift keying (MSSK) is presented in [GB+19b], in which the information is modulated into the transmitter antenna index. Finally, a hybrid modulation scheme incorporates one or more aforementioned properties in order to embed the information. For example, OOK and molecular frequency shift keying are combined together in [MM+10] to form multilevel amplitude modulation to encode more number of symbols. Similarly, in [KC13], a modulation scheme named isomer ratio shift keying is introduced that modulates the information through ratios of quantity of a specific isomer. A higher-order modulation scheme named run-length aware hybrid modulation is showcased in [PS+14] that combines both concentration shift keying (CSK) and MoSK to form MCSK. In [GB+19b], MSSK is extended by including different types of molecules, which results in molecular spatial modulation. In [GB+19a], index modulation and PPM are merged together to provide resilience against inter-link interference (ILI). Additionally, index modulation and CSK are united together for designing a modulation scheme, as discussed in [HW+19a]. In terms of the practical implementation of hybrid modulation schemes in a testbed setup, very little work has been done. However, in [BS+21], SPION and fluorescent particles are used simultaneously for transmission of independent streams of data. This helps to increase the channel capacity in comparison to the transmission of only a single type of information particles.

System Design with Multiple-Input Multiple-Output In classical wireless communication, multiple-input multiple-output (MIMO) is a well-known technique used to achieve spatial diversity, beamforming, and to increase the data transmission capacity using multiple transmit and receive antennas [Höh13]. For inclusion of MIMO in diffusion-based molecular communication, the first attempt has been made in [MY+12], with the introduction to the concepts like transmitter diversity, receiver-side diversity combining, and spatial multiplexing. However, the effect of ISI is not considered in this work, which is later

included in [KL+16]. Additionally, modulation scheme such as index modulation has been introduced to simultaneously tackle the problems of ISI and ILI [GB+19b]. MIMO channel modeling based machine learning is presented in [LY+17], which is used in [DY+18] to analyze spatial coding techniques such as the Alamouti-type coding and repetition MIMO coding along with the receiver-side combining strategies such as selection diversity and equal gain combining. One of very few practical implementations of MIMO in a molecular communication testbed is presented in [KL+16], where a single-input single-output (SISO) system is extended to a MIMO system. In order to present the possibility of utilizing molecular communication in an in-body blood vessel network, an in-vessel molecular communication testbed is developed using acid and base, where information is transmitted by changing the pH value [LK+20a]. The low data transmission rates are improved by extending the work in [LK+20b] that incorporates improved modulation schemes such as binary time shift keying.

Channel Coding To increase the robustness of data transmission over a given channel, the method of channel coding is considered, which helps in the detection and/or correction of the erroneous data symbols by introducing redundancy into the data transmission sequence [PP+72]. The first instance where the channel coding technique has been incorporated in a diffusion-based molecular communication channel is [LH12]. In this work, the Hamming code is implemented in order to achieve a coding gain for larger distances compared to the uncoded transmission scheme. However, this is achieved only if the system makes use of a large number of molecules. In the case of fewer molecules, the effect of ISI from the redundant symbols becomes more dominant leading to higher transmission errors. Therefore, it is observed that in the case of molecular communication, channel codes serve a different purpose altogether, which helps counter the ISI. However, some works such as [Dam20] use channel coding strategies in order to make data transmission more robust against the amplitude-dependent noise of the diffusion-based molecular communication channel. In [BL+14; Bai16], minimum energy codes have been introduced that prove to be more energy efficient as compared to the Hamming codes. Additionally, these codes are also less error prone as compared to the Hamming codes. Other popular channel coding schemes from the digital communication theory have also been applied to the diffusion-based molecular communication channel such as convolutional codes [MM+13; LH+15b], Reed-Muller codes [LH+15a], Reed-Solomon codes [DD+17], Bose–Chaudhuri–Hocquenghem codes [KP+18], Luby transform codes [Bai16], etc. Novel coding techniques for molecular communication in the form of ISI-free code has been pro-

posed in [SL+12; SL+13], where the communication reliability is improved by restricting the molecules emitted at a later time instant from reaching the receiver earlier than the molecules emitted at a previous time instant, also referred to as the crossover effect. In the same work, the performance of repetition codes and ISI-free codes is shown to be similar to that of convolutional codes, but at a much lower complexity. The performance of ISI-free codes is further increased by introducing the concept of time gap between successive code words [AK+16; KJ+19]. Additionally, zero-error codes are presented for different types of molecules that avoid the problem of ISI arising from the occurrence of random delays in successive data transmission time slots [AG+18]. In another work, a low complexity error correction method through amplitude-width encoding is presented for potential use in bio-circuits, where the redundancy is embedded in the duty-cycle of the transmitted signal [KS18]. Finally, the implementation and analysis of FEC channel codes, line codes, spreading codes, and spatial codes for a diffusion-based molecular communication system is extensively discussed in [Dam20]. In the same work, novel line codes that are adapted to the transmission channel characteristics are also presented. One of very few practical works discussing the benefits of the channel coding method is presented in [KA+20b], where ISI-mitigating channel code in combination with MIMO application is implemented in an alcohol-based molecular communication testbed. A performance enhancement in terms of character error rate is witnessed as compared to the uncoded transmission scheme as the coding family avoids the continuous transmission of the binary data symbol "1", and uses an adaptive thresholding strategy.

1.6 Author's Contributions

This section presents the list of peer reviewed scientific contributions from the author, which are in agreement with the objectives of this dissertation discussed in Section 1.4.

Testbed Design A proof of concept for the macroscopic air-based molecular communication testbed using fluorescent dye is presented for the first time in [BD+20b], which shows that a simple modulation scheme such as OOK can help achieve a data transmission rate up to 20 bit/s. This is later extended in [BD+20a] to setup a 2 m long physical testbed with a detailed description about the choice of hardware necessary for the individual units comprising of the transmitter, the information carrier, the transmission channel, and the receiver. The work tries to emulate the scenario that could enable the transmission of information within industrial networks incorporating pipes, ducts, etc. Furthermore, mod-

els for the spray nozzle injector at the transmitter side and the camera-based detector at the receiver side are described to simulate the testbed in a particle-based simulator. The simulated models are adapted with respect to the transmitter and the receiver, and are compared with the presented analytical model obtained from the testbed measurements. Additionally, modifications in the testbed design at the transmitter side through improved hardware and PIN photodiode-based detection at the receiver side is presented in [BK$^+$23], showcasing a significant improvement of the data transmission rate up to 50 bit/s with simple OOK.

System Modeling The air-based molecular communication testbed setup is transformed into an end-to-end system model followed by the evaluation of the end-to-end system impulse response in [BD$^+$22b]. The system impulse response from a short spray burst of the water-based dye solution at the transmitter side is measured using the camera-based detector at the receiver side. The true end-to-end system comprises of the influences of a perfect transmitter, a perfect channel and a perfect receiver, as shown in [BD$^+$20a]. However, the response of the system over multiple trials shows randomness, which is attributed to the imperfect sprayer at the transmitter side, turbulence and air drag in the transmission channel, the imperfection of the camera-based detector at the receiver side, etc. Therefore, a discrete-time input into the system, the true end-to-end system impulse response, and the noise model are presented in [BD$^+$20a] to give the complete analytical model of the testbed.

Modulation and Detection For the first time, the transmission of data over the air-based molecular communication testbed is performed by employing various modulation schemes exploiting several degrees of freedom in the form of the information carrier concentration level, time, color, and the spatial regions populated by the sprayed information carrier, presented in [BD$^+$22b]. These include PPM, DPPM, shorten pulse position modulation, multi-pulse pulse position modulation, MoSK, MCSK, spatial shift keying, and spatial concentration shift keying. For the hardware being used in the testbed, the performances of all the modulation schemes are compared to the standard OOK with respect to their maximum achievable data transmission rates, resource requirements and their bit error rate (BER) performance. Modulation schemes such as MoSK and MCSK are especially discussed in great detail in [BD$^+$19; DB$^+$21b]. Various equalization techniques and detection algorithms have also been implemented and compared for OOK modulation scheme. The aforementioned analyses and comparisons have particularly been carried out using the

analytical model as well as the particle-based simulator, introduced in [BD+20a]. Additionally, for the particle-based simulator, a mobility scenario of the transmitter with respect to the receiver is investigated in [SG+21] for studying its effects on the received signal. The results show that mobility impacts the amplitude as well as the width of the received signal completely in tune with the Doppler effect.

Channel Coding Like most communication systems, long-distance molecular communication is also prone to errors originating from the transmission channel. This leads to decreased reliability in data transmission with an increased BER. This is where the utility of channel coding schemes is exploited by adding redundant symbols to a data sequence, resulting in more robust data transmission. While in the literature, several channel coding techniques have been studied for diffusion-based molecular communication systems, very few works have actually showcased their usefulness in macroscopic molecular communication testbeds. In [BD+22a], a new line coding method named weak sequence preventing mapping (WSPM) code is implemented for the first time in a macroscopic molecular communication testbed using fluorescent dyes by taking into account the testbed channel characteristics. This work presents WSPM code as another viable alternative to improving robustness in long-distance data transmission. Additionally, the resource requirement and the BER performance of WSPM code are compared with the other popular block codes from digital communication theory, proving its usefulness over the other channel coding schemes.

Testbed Design with Multiple Transmitters/Receivers In [DB+21a], for the first time, the air-based molecular communication testbed using fluorescent dyes is extended to a 2×2 MIMO system. The transmitters of the 2×2 MIMO setup is realized by two pressure sprayers containing water-based dye solutions, while the emulation of the spatially separated receivers is realized by splitting the field of view of the camera-based detector into two equal halves. The MIMO setup helps to increase the data transmission rate through spatial multiplexing or the robustness of data transmission through array gain. Implementation of spatial multiplexing is carried out by multiplexing the data stream over the pressure sprayers, and by transmitting them in parallel. On the other hand, an array gain is achieved with the help of spatial diversity algorithms such as the Alamouti-type coding and repetition MIMO coding. For a fair comparison with SISO transmission scheme, simple fixed threshold detector is incorporated at the receiver side for detecting the transmitted data. With the help of BER analysis, it is presented that the performances of both the Alamouti-type

coding and repetition MIMO coding exceed that of SISO transmission system.

Miscellaneous Applications/Applicability Inspired by the presented air-based molecular communication testbed using fluorescent dyes and the ongoing SARS-CoV-2 pandemic, the duality between various aspects of macroscopic air-based molecular communication and viral infection spreading has been showcased in [SH+21a]. The airborne aerosol and droplet transmission via various human respiratory processes is modeled as a multiuser molecular communication scenario [SH+21b] with the respiratory event-driven molecular variable concentration shift keying, as discussed in [BD+19]. Experiments are conducted by inducing coughs artificially by an adult test person with the release of saturated aqueous solution of fluorescent dye mixed with saliva. It is also shown that the number of observed particles released is significantly higher in test series without mouth and nose protection than in those with a well-fitted medical mask. Additionally, the concept of mutual information in the sense of infection rate is presented in [SH+21b]. The infection rate is explained in [HD+21] by considering various discrete memoryless channel models, which depends on the distribution of the channel input events as well as on the transition probabilities between channel input and output events. Contrary to the maximization of the mutual information in data communication systems, it becomes imperative to minimize the infection rate between an infected person acting as the transmitter and a non-infected person acting as the receiver, as mentioned in [SH+21b]. Furthermore, the simulation tool for the macroscopic molecular communication is extended through consideration of all the previously mentioned scenarios.

List of Publications by the Author

[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.

[BD+20a] S. Bhattacharjee, M. Damrath, F. Bronner, L. Stratmann, J. P. Drees, F. Dressler, and P. A. Hoeher, "A testbed and simulation framework for air-based molecular communication using fluorescein," in *Proc. ACM Int. Conf. on Nanoscale Computing and Communication (NANOCOM)*, Virtual Conference, Sep. 2020, pp. 1–6.

Abstract: Molecular communication can enable transmission of information within industrial networks comprising of pipes, ducts, etc. This work emulates the system by introducing an air-based macroscopic molecular communication testbed, exploiting the fluorescence property of a water-based solution of an organic compound called fluorescein. An efficient transmitter in the form of an industrial sprayer, coupled with a high-speed camera-based detection, eventually paves way to achieve higher data transmission rates. The transmission distances considered are in the range of several centimeters to meters. Additionally, models for spray nozzle injector and camera receiver are described to simulate the testbed in a particle-based simulator. These simulated models are calibrated to the used transmitter and receiver and are compared with the analytical models obtained from the testbed measurements.

[BD+20b] S. Bhattacharjee, M. Damrath, and P. A. Hoeher, "Design of macroscopic air-based molecular communication concept using fluorescein," in *Proc. ACM Int. Conf. on Nanoscale Computing and Communication (NANOCOM)*, Virtual Conference, Sep. 2020, pp. 1–2.

Abstract: A proof of concept for a macroscopic air-based molecular communication testbed is presented using fluorescein. The compound in solution state when excited by ultra-violet light, acts as information carrier between a sprayer (transmitter) and a camera (receiver). Relatively higher data rates can be achieved compared to traditional air-based testbeds using alcohol. Implemented modulation schemes include on-off keying, pulse position modulation and differential pulse position modulation, which showcase the efficacy of the system in terms of achievable data transmission rates.

[DB+21a] M. Damrath, S. Bhattacharjee, and P. A. Hoeher, "Application of MIMO techniques in macroscopic air-based molecular communication," in *Proc. ACM Int. Conf. on Nanoscale Computing and Communication (NANOCOM)*, Virtual Conference, Sep. 2021, pp. 1–2.

Abstract: In this contribution, an air-based macroscopic 2x2 multiple-input multiple-output (MIMO) molecular communication testbed using fluorescent dyes

is presented and the application of MIMO techniques is investigated. It is demonstrated that the data rate can be doubled using spatial multiplexing. Using repetition MIMO coding and Alamouti-type coding, a spatial diversity/array gain can be achieved, resulting in error-free transmission over the entire transmission distance of the testbed.

[DB+21b] ——, "Investigation of multiple fluorescent dyes in macroscopic air-based molecular communication," *IEEE Trans. Mol. Biol. Multi-Scale Commun.*, vol. 7, no. 2, pp. 78–82, Jun. 2021.

Abstract: An air-based macroscopic molecular communication testbed exploiting fluorescence properties of water-based solutions of Uranine and Rhodamine 6G is presented in this work. The testbed comprises of an industrial sprayer as its transmitter, a 2 m long tube as transmission channel, and a high-speed camera-based detector. The transmission distances considered cover a range over several tens of centimeters to meters. Concerning modulation schemes, molecular shift keying and molecular concentration shift keying are implemented and compared with on-off keying serving as a benchmark. It is shown that the former two can be used to decrease the bit error rate and/or to increase the bit rate. Furthermore, a platform is set for implementing a multiuser scenario, where the same channel is made accessible to multiple users by simultaneous use of different dyes for data transmission with small co-channel interference.

[HD+21] P. A. Hoeher, M. Damrath, S. Bhattacharjee, and M. Schurwanz, "On mutual information analysis of infectious disease transmission via particle propagation," *IEEE Trans. Mol. Biol. Multi-Scale Commun.*, Oct. 2021.

Abstract: Besides mimicking bio-chemical and multi-scale communication mechanisms, molecular communication forms a theoretical framework for virus infection processes. Towards this goal, aerosol and droplet transmission has recently been modeled as a multiuser scenario. In this letter, the "infection performance" is evaluated by means of a mutual information analysis, and by an even simpler probabilistic performance measure which is closely related to absorbed viruses. The so-called infection rate depends on the distribution of the channel input events as well as on the transition probabilities between channel input and output events. The infection rate is investigated analytically for five basic discrete memoryless channel models. Numerical results for the transition probabilities are obtained by Monte Carlo simulations for pathogen-laden particle transmission in four typical indoor environments: two-person office, corridor, classroom, and bus.

Particle transfer contributed significantly to infectious diseases like SARS-CoV-2 and influenza.

[SH+21a] M. Schurwanz, P. A. Hoeher, S. Bhattacharjee, M. Damrath, L. Stratmann, and F. Dressler, "Duality between coronavirus transmission and air-based macroscopic molecular communication," *IEEE Trans. Mol. Biol. Multi-Scale Commun.*, vol. 7, no. 3, pp. 200–208, Apr. 2021.

Abstract: This contribution exploits the duality between a viral infection process and macroscopic air-based molecular communication. Airborne aerosol and droplet transmission through human respiratory processes is modeled as an instance of a multiuser molecular communication scenario employing respiratory-event-driven molecular variable-concentration shift keying. Modeling is aided by experiments that are motivated by a macroscopic air-based molecular communication testbed. In artificially induced coughs, a saturated aqueous solution containing a fluorescent dye mixed with saliva is released by an adult test person. The emitted particles are made visible by means of optical detection exploiting the fluorescent dye. The number of particles recorded is significantly higher in test series without mouth and nose protection than in those with a well-fitting medical mask. A simulation tool for macroscopic molecular communication processes is extended and used for estimating the transmission of infectious aerosols in different environments. Towards this goal, parameters obtained through self experiments are taken. The work is inspired by the recent outbreak of the coronavirus pandemic.

[SH+21b] ——, "Infectious disease transmission via aerosol propagation from a molecular communication perspective: Shannon meets coronavirus," *IEEE Commun. Mag.*, vol. 59, no. 5, pp. 40–46, May 2021.

Abstract: Molecular communication is not only able to mimic biological and chemical communication mechanisms, but also provides a theoretical framework for viral infection processes. In this tutorial, aerosol and droplet transmission is modeled as a multiuser scenario with mobile nodes, related to broadcasting and relaying. In contrast to data communication systems, in the application of pathogen-laden aerosol transmission, mutual information between nodes should be minimized. Toward this goal, several countermeasures are reasoned. The findings are supported by experimental results and by an advanced particle simulation tool. This work is inspired by the recent outbreak of the coronavirus (COVID-19) pandemic, but also applicable to other airborne infectious diseases like influenza.

[SG+21] L. Stratmann, J. T. Gómez, S. Bhattacharjee, M. Damrath, P. A. Hoeher, and F. Dressler, "Impact of mobility on air-based macroscopic molecular communication: A simulation study," in *Proc. ACM Int. Conf. on Nanoscale Computing and Communication (NANOCOM)*, Virtual Conference, Sep. 2021, pp. 1–2.

Abstract: Molecular communication (MC) is considered to become a very relevant communication technology for many industrial and health care applications. Here, we study the impact of mobility on the channel impulse response. Building upon previous work on simulation and experimental insights on air-based MC, we investigate a mobility scenario in which the sender moves with respect to the receiver of the signal. Our findings show the impact of the mobility on the amplitude as well as the width of the received signal. We consider our results as a first step towards general mobility-impact mitigating waveform and protocol designs for MC.

[BD+22a] S. Bhattacharjee, M. Damrath, and P. A. Hoeher, "Channel coding techniques in macroscopic air-based molecular communication," in *Proc. ACM Int. Conf. on Nanoscale Computing and Communication (NANOCOM)*, Barcelona, Spain, Oct. 2022, pp. 1–2.

Abstract: Long distance molecular communication (MC) is prone to channel errors, leading to an increased bit error rate. For solving the issue, a new coding scheme named weak sequence preventive mapping code is designed, conforming to the channel characteristics of the macroscopic air-based MC using fluorescent dye. The scheme not only minimizes the transmission errors, but also conserves the resource in the form of sprayed water-based dye solution.

[BD+22b] S. Bhattacharjee, M. Damrath, L. Stratmann, P. A. Hoeher, and F. Dressler, "Digital communication techniques in macroscopic air-based molecular communication," *IEEE Trans. Mol. Biol. Multi-Scale Commun.*, vol. 8, no. 4, pp. 276–291, Dec. 2022.

Abstract: An air-based macroscopic molecular communication testbed exploiting fluorescence properties of water-based solutions of Uranine and Rhodamine 6G dyes is presented in this work. The testbed comprises of an industrial sprayer as its transmitter, a 2 m long tube as the transmission channel, and a high-speed camera-based detector. The considered transmission distances cover a range over several tens of centimeters to meters. The analytical end-to-end system model is extended to include the noise model in this work. Various modulation schemes

exploiting the concentration levels of the sprayed dyes as the information source have been implemented and their performances are compared to the standard on-off keying. Spatial domain, along with the concentration of the dyes has been introduced as an additional degree of freedom into the testbed to improve the bit rate, and the performance is compared with respect to those modulation schemes using color and concentration levels of the dyes as degrees of freedom. Various equalization techniques and detection algorithms have also been implemented and compared. All comparative analyses are performed with respect to the measurements obtained from the testbed, the analytical models, and the particle-based Pogona simulator.

[BK+23] S. Bhattacharjee, E. B. Krebs, A. Harlakin, and P. A. Hoeher, "Detection process in macroscopic air-based molecular communication using a PIN photodiode," *IEEE Trans. Mol. Biol. Multi-Scale Commun.*, vol. 9, no. 1, pp. 13–17, Mar. 2023.

Abstract: An air-based macroscopic molecular communication (MC) testbed is presented exploiting the fluorescent properties of a highlighter-ink water-based solution. The testbed consists of a pressure sprayer as its transmitter, a 2 m long tube as the transmission channel, and a newly designed PIN-photodiode-based detector at the receiver side. The considered transmission distances cover a range over several tens of centimeters to meters. The on-off keying (OOK) modulation scheme is implemented in the new setup and a bit error rate (BER) analysis is performed over the channel length. A data transmission rate of 50 bit/s is achieved, which is more than twice that of the erstwhile implemented testbed using a camera-based detector, at a comparable BER. To the best of authors' knowledge, this is perhaps the fastest transmission rate ever achieved across all existing macro-scale MC testbeds employing OOK.

1.7 Dissertation Outline

The content outline and the organization of the dissertation is summarized as follows:

- **Chapter 2** discusses the aspects which concern the design and implementation issues of an air-based macroscopic molecular communication testbed using fluorescent dyes. The various functional units and components that form the building blocks of the testbed are also explained in great detail. Additionally, a performance comparison of

the testbed presented in this dissertation and the traditional macroscopic molecular communication testbed using alcohol is made with a detailed analysis of proven advantages and superiority of the former against the latter.

- **Chapter 3** focuses on the theoretical aspects behind the mass transportation processes of the information carrier through a molecular communication channel. A distinction of these mass transportation processes based on the scale of the application of molecular communication is also made and discussed in this chapter. Moreover, the theoretical models for the transmission of the information carriers in the form of sprayed droplets of uranine and the highlighter ink water-based solutions are defined for the testbed presented in this dissertation. Consequently, the dominant physical effects such as the drag and the gravitational forces influencing the sprayed droplets propagating through the channel are also highlighted. Furthermore, a test of linearity with respect to the choice of the information carrier is showcased for the system under consideration. The end-to-end system model of the testbed is presented that takes into account the system impulse response originating from the combined effects of the transmitter, the propagation channel, the receiver, and the noise process that arises due to the imperfect design of these units. Finally, the true end-to-end system impulse response and the noise models are combined together to define the complete analytical model of the testbed, which is then validated by transmission of a message sequence through testbed measurement scenarios.

- **Chapter 4** presents the receiver side data acquisition principle of the macroscopic molecular communication testbed presented in this dissertation. Simple binary modulation technique like OOK is implemented and its BER over increasing distance is analyzed by considering the testbed measurements and the analytical model using low-complexity threshold detection strategy. Additionally, the theoretical BER is also evaluated by considering the testbed measurements. The performance of different detector types and equalization techniques are compared through BER analysis of the testbed measurement results, and the simulation results obtained from the analytical model. Furthermore, higher-order modulation schemes exploiting additional degrees of freedom such as time, the color of the water-based dye solutions, and the physical space are implemented with an analysis based on their respective BER performances, maximum achievable data transmission rates, and resource requirements.

- **Chapter 5** deals with the implementation of channel coding schemes in the presented testbed, which ensures robust and error resilient data transmission over longer

distances. Various broader types of coding techniques such as block codes, convolutional codes, line codes and spatial codes have been investigated, and their BER performance has been analyzed based on the testbed measurements and simulation results obtained from the analytical model. Furthermore, their respective resource requirements have also been studied in this chapter.

- **Chapter 6** concludes the dissertation by summarizing the content. Additionally, the presented outlook provides some insight into the topics which could be explored in the future.

2

Testbed Setup

This chapter presents a detailed discussion about the design and the implementational aspects of an air-based macroscopic molecular communication testbed using fluorescent dyes. The testbed design principle along with its various units and components mainly comprising of the transmitter, the information carrier, the transmission channel, and the receiver are extensively discussed. Furthermore, a comparison with the traditional molecular communication testbed using alcohol is presented with a detailed analysis of proven advantages and improvements.

2.1 Testbed Units and Components

Efficient transfer of information in any physical communication environment is dependent on its system design aspects and hardware implementation, which in principle, applies to molecular communication too. In order to establish molecular communication as an industry-approved communication methodology, a holistic application framework is necessary which encompasses various domains of communication theory such as the system response analysis and modeling, implementation of high data transmission rate achieving modulation schemes, incorporating efficient detection algorithms to achieve lower error rates, inclusion of channel coding techniques for effective error detection and correction, etc. In this dissertation, efforts have been made for the inclusion of the aforementioned points in a more controlled and confined conditions in order to achieve reliable transmission of information. Towards this goal, a testbed design and its realization helps to conduct preliminary tests and analyses in a seamless fashion, allowing easy extension/alteration of the different communication units, which makes the study of the air-based macroscopic

molecular communication system more efficient. Like any other communication system, this air-based molecular communication testbed comprises of three important units, the transmitter, the transmission channel through which the information carrier propagates, and the receiver.

2.1.1 Transmitter

Sprayer

The *Gloria Prima 3* pressure sprayer is used for spraying out the water-based solution of the information carrier and acts as the transmitter of the testbed, as shown in Figure 2.1. The maximum capacity of the container is 3 L, while the integrated pump generates a maximum pressure of 3 bar. The water-based solution is released from the sprayer through a long spray lance with a conical nozzle tip, mounted on a stand for stable and consistent spray bursts. The nozzle tip of the spray lance is adjusted to produce either "cloudy"- or "jet"-like spray formation. In order to maximize the reach of the sprayed information carrier over larger distances, the nozzle tip is adjusted to emanate "jet"-like spray bursts. This is achieved by turning the nozzle tip 360° from its closed position. At a maximum pressure of 3 bar, the spraying volume rate is around 0.54 L/min. The switching mechanism of the sprayer is electronically controlled by a 12 V *AOMAG* solenoid valve attached at the end of the spray lance, right before its handle. The sprayer switches between two available states, i.e., the "on" state and the "off" state. The valve with a maximum tolerance pressure of 10 bar, is controlled by a waveform generator and a control circuit. The minimum shutting time of the valve is $T_{\text{trans}} = 0.03\,\text{s}$, which is the minimum time required for the valve to transition from the "on" state to the "off" state. This influences the shape of the impulse response of the system in the time domain. The transmitter side can easily be extended to accommodate multiple sprayers with separate water-based solutions of the information carriers, depending on the modulation schemes used for data transmission.

Micro-controller Device

The opening and shutting mechanism of the solenoid valve of the sprayer is operated by a micro-controller device called the *Analog Discovery 2* manufactured by *Digilent*, as shown in Figure 2.2. It is a compact universal serial bus operated device that offers a large variety of useful laboratory functionalities such as waveform generation and digital oscilloscope, apart from offering power supply to the auxiliary electronic devices, and digital inputs and outputs. The solenoid valve is controlled by the waveform generator of the *Analog*

Figure 2.1: Solenoid valve operated *Gloria 3* pressure sprayer as the transmitter with water-based solution of the information carrier.

Figure 2.2: A micro-controller device *Analog Discovery 2* facilitating the opening and shutting mechanisms of the solenoid valve.

Discovery 2 through a metal–oxide semiconductor field-effect transistor from its analog output. The waveform output signal, which is used to produce a transition from the "off" state to the "on" state of the solenoid valve, is a 5 V rectangular-shaped signal. The analog output of the *Analog Discovery 2* has a resolution of 14 bit, and offers up to ±5 V at a sampling rate of 100 megasamples per second with 12 MHz bandwidth.

2.1.2 Information Carrier

The information carrier is chosen based on its property of fluorescence. Fluorescence is a type of luminescence displayed by a substance that has absorbed electromagnetic radiation of certain frequency and wavelength. As a result, the electrons from the atoms are knocked to the higher quantum states by the absorbed photons. The high energy electrons then return back to their ground state by releasing energy in the form of emitted radiation, generally of lower frequency and longer wavelength as compared to the absorbed radiation. The wavelength gap between the initially absorbed radiation and the emitted radiation is termed as the Stokes shift, and is unique to the material displaying fluorescence property. Such fluorescent substances are used as the information carrier, which typically emit radiation in the visible light spectra for the convenience of detection. The transmitter sprayer contains such information carrier in the form of water-based fluorescent dye solutions. Depending on the degree of freedom of the implemented modulation schemes, multiple fluorescent dyes from different transmitter sprayers can be used simultaneously to send information with a higher data symbol cardinality. Therefore, the choice of the fluorescent dye as the information carrier is mainly influenced by the following criteria:

- Solubility of the fluorescent dye in water
- Energy contained in the emitted light from the water-based fluorescent dye solution for a seamless detection at the receiver side, without risking an early saturation of the receiver
- Production of distinct and separable emission spectra in the cases, where multiple dye solutions are used simultaneously for data transmission with minimum interference
- Minimum to no contribution to the production of any offset value during detection process, which may otherwise lead to the background noise
- No adverse effects on health and environment

Keeping the aforementioned points in mind, the water-based solutions of uranine and fluorescent highlighter ink have been chosen as the information carriers. The properties and suitability of these information carriers have been studied very closely in this dissertation for successful data transmission over distances ranging from several tens of centimeters to meters.

Uranine is a disodium salt of fluorescein, which is a red crystalline powdery compound. The salt is slightly soluble in water and alcohol, and finds a plethora of applications particularly in the fields of biochemical and medical research [NU02], earth science [Kas18], plant science [PR+15], etc. When excited by a less harmful and long wavelength ultraviolet (UV)-A light source (also termed as "black light") at approximately 398 nm, a water-based solution of uranine produces a peak emission at approximately 548 nm. The emission spectra of both water-based solution of uranine and the UV-A light source are presented in Figure 2.3. The dye properties do not change while the measurements are being performed for analysis during data transmission, and therefore, are reliably used to represent the information symbols.

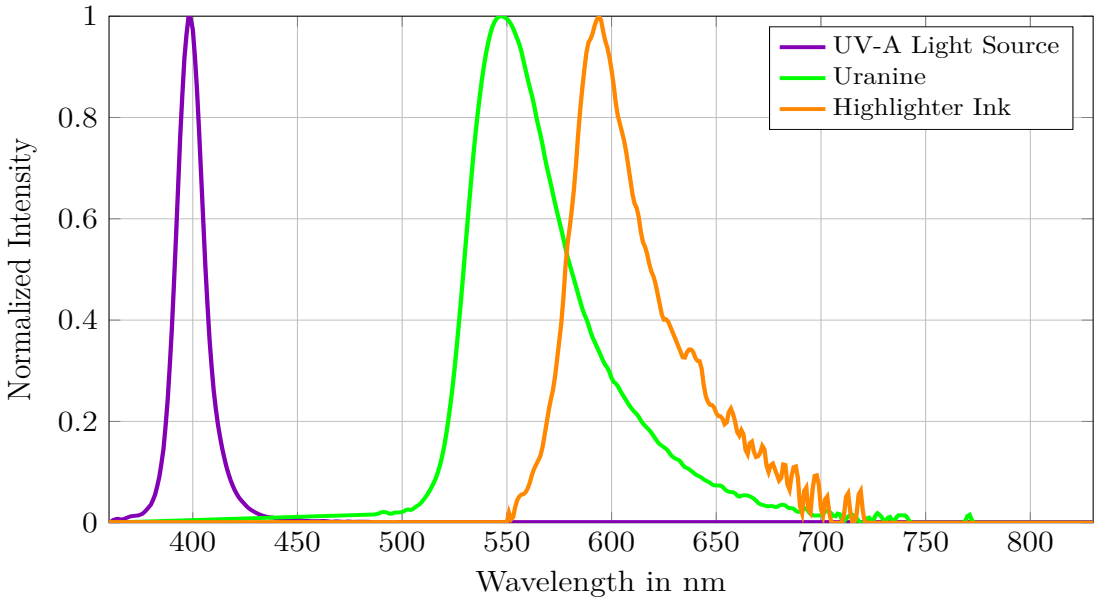


Figure 2.3: Spectra of ultraviolet-A light source and the water-based solutions of uranine and highlighter ink as the information carriers.

Due to their strong fluorescence property, highlighter inks have traditionally been utilized for conducting many scientific experiments and studies, especially in the field of blood vessel imaging [BD21]. Therefore, a water-based solution of *Edding 345* orange highlighter ink is used in this particular work as an additional information carrier for implementing

modulation schemes with higher degrees of freedom. The highlighter ink is available in the form of concentrated aqueous solution, which has a higher solubility in water as compared to uranine. The water-based solution of the highlighter ink produces fluorescence with a peak emission wavelength at approximately 594 nm, when excited by the same UV-A light source at 398 nm. The emission spectra is again depicted in Figure 2.3.

From Figure 2.3, it is clear that the UV-A light source makes uranine and the highlighter ink produce distinguishable emissions over a certain range of wavelength. This property helps the receiver make a decision about the transmitted symbols, when the information carriers are simultaneously released into the channel.

2.1.3 Transmission Channel

Figure 2.4: The transmission channel in the form of a tube-like structure with multiple flaps over a maximum length of 2.38 m, where the detection modules can be mounted.

The transmission channel consists of a tube of length 2.38 m and a diameter of 50 cm, which can be further extended up to several meters by attaching additional tubes. The tube

consists of three rows of six opening flaps, each separated from its neighbor by a distance of 30 cm, as shown in Figure 2.4. The detection modules and chambers can be mounted on these flaps at varying distances from the transmitter, located at one end of the tube. The tube can be made to rest on an aluminum support structure, and tilted up to a maximum angle $\phi_{\text{tilt,max}} = \pm 20^{\circ}$ to facilitate proper drainage of the residual dye solution, which gets collected at the bottom surface of the tube. To further prevent any noise contribution by the residual dye solution, an offset removal technique is employed, described later in Section 2.3 of this chapter. With the flaps closed, the internal environment of the tube is left completely dark in order to avoid interference from the environmental light. Furthermore, turbulence caused due to the movement of air, and its deflection off the interior surface of the tube is also minimized with this design.

2.1.4 Receiver

The receiver module comprises of a high-speed smartphone digital camera *OnePlus 6T*, as shown in Figure 2.5, capturing the images at 480 frames per second (FPS) with a resolution of 1280×720 pixels and a viewing angle of approximately 36.87°. The smartphone digital camera is controlled remotely via instructions passed through Android debug bridge as soon as the spray droplets are released from the solenoid valve controlled by the microcontroller. The region that lies in the field of view of the camera is illuminated by an 18 W UV-A flashlight, as presented in Figure 2.6. The UV-A flashlight excites the liquid droplets containing the uranine and the highlighter ink molecules emitting colors in the wavelength regions, as seen from the spectrometer measurements in Figure 2.3 described in Section 2.1.2.

2.2 Testbed Layout

The front view of the final testbed layout with all the modular components assembled is shown in Figure 2.7. As seen from the figure, the tube rests on an aluminum frame, mounted by the receiver chamber with the UV-A flashlight fitted on its side. The UV-A flashlight is placed along the circumference of the tube at an angle, so that it illuminates major portion of the cross-sectional area of the tube with the light cone passing right through its center. Also, this configuration does not allow the droplets sticking to the inner side walls of the tube from getting illuminated, which could otherwise contribute to the noise. The side view of the testbed layout is depicted in Figure 2.8, which gives a clear picture of the placement

Figure 2.5: *OnePlus 6T* used as the receiver module of the testbed.

Figure 2.6: Ultraviolet-A flashlight.

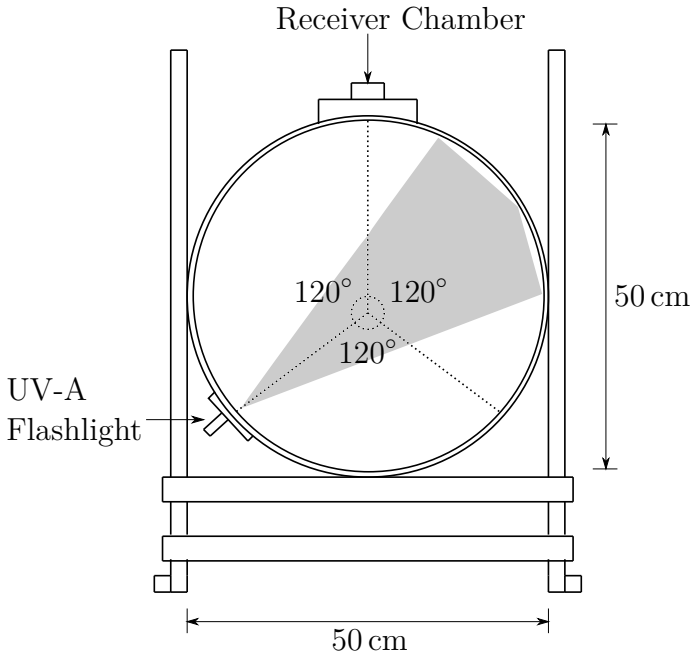


Figure 2.7: Physical testbed front view.

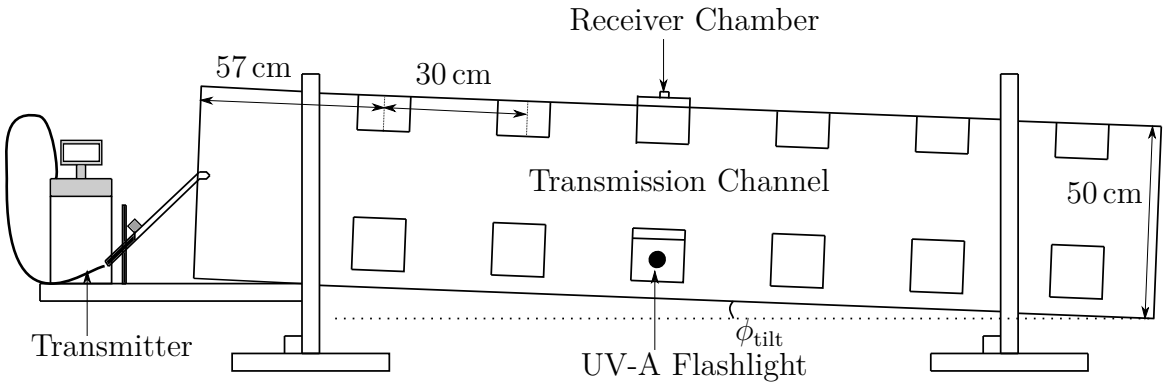


Figure 2.8: Physical testbed side view.

of the various components along the length of the tube. It may be mentioned that the receiver chamber is placed at a certain flap, which also corresponds to the placement of the UV-A flashlight with a separation angle of 120°. This ensures maximum illumination of the information carrier droplets with a clear image captured by the camera at the receiver side. It is also made sure that the camera is kept outside the light cone of the UV-A flashlight in order to minimize the effects of noise and any offset which may arise at the time of capturing images for pixel intensity measurements. The tube is tilted at an angle of ϕ_{tilt} in order to flush out the residual dye solution that collects at the bottom surface of the tube. If the transmission system uses a single fluorescent dye solution as the information carrier, then the residual dye solution could be collected outside the tube and re-used to send new sets of data sequences. This is done in order to conserve the resources used as the information carrier. However, if the system uses multiple fluorescent dye solutions simultaneously to transmit the information, the separation process of the constituent dye solutions becomes more complex, which is not considered in this dissertation. The testbed with all its physical units and components is presented in Figure 2.9.

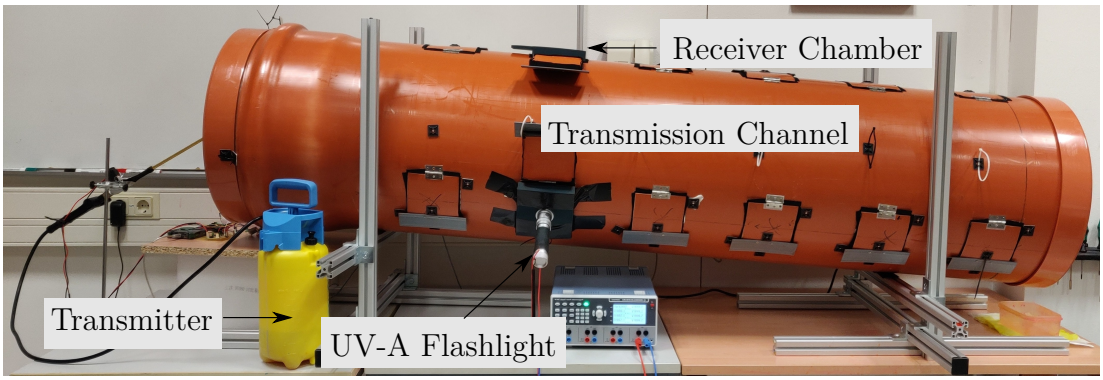


Figure 2.9: Testbed with all its physical units and components for a successful implementation of macroscopic air-based molecular communication.

2.3 Information Carrier Concentration

The choice of the information carrier depends on various aspects such as the wavelength of the light emitted under the influence of UV-A light source, photo-sensitivity of the detector to the light received from the water-based solution droplets, responsitivity of the detector to the fast changing number of droplets observed over a given period of time, etc. However, the aforementioned aspects heavily depend on the concentration level of the information carriers in their respective water-based solutions. At a specific desirable concentration level,

the sprayed water-based solution droplets are detected at the farthest intended transmission distances with significantly high signal-to-noise ratio (SNR). This is necessary in order to perform the measurements leading to the establishment of the communication theory with respect to the maximum reach of the communication channel of the testbed under investigation. Care should also be taken by conserving the information carrier resources by avoiding over-saturation of the water-based solutions.

As already described in Section 2.1.2, uranine and the highlighter ink water-based solutions are used as the information carriers in the system. The light emitted by their sprayed water-based solution droplets, when under the influence of the UV-A light source, can be distinguished in terms of the energy contained in their wavelength components. This phenomenon is more pronounced in the green and the red pixel intensity values of the images captured by the high-speed camera-based detector at the receiver side. The camera captures images at 480 FPS, and the sum of the intensity values of all green and red pixels for the captured frames are added together and plotted with respect to time t. The images are post-processed by taking the first captured image as a reference frame, and subtracting it from all subsequent captured images. This further tackles the problem of the offset that can arise due to any other luminous components or residual dye solution inside the tube, eventually also contributing to the noise process. From the green and the red pixel intensity plots, it is observed that one is more dominant than the other depending on the information carrier being sprayed inside the testbed tube. It is observed (later presented in Section 2.4) that in the case of the sprayed uranine water-based solution droplets, the green pixel intensity recorded by the camera is far more dominant than the red pixel intensity. Conversely, in the case of the sprayed highlighter ink water-based solution droplets, the red pixel intensity is more dominant than the green pixel intensity.

In order to find the optimum concentration levels of each information carrier, the pixel energy is measured at a distance of 88 cm. The concentration levels of the information carriers are gradually increased in separate sprayer containers, each holding 3 L of water. The pixel energy at a given concentration level is then calculated by spraying the water-based solution of the information carrier with a spray duration $T_{\text{on}} = 0.02\,\text{s}$ by integrating the pixel intensity pulse with respect to time over a duration of 0.05 s (the choice of parameters are akin to spray duration T_{on} and symbol duration T, described later in this section and Section 2.4). The idea of correlating energy contained in the pixel intensity pulses to the dye concentration is inspired from the simultaneous measurements of particle image velocimetry and planar laser-induced fluorescence in the work presented in [SG+12]. For both the information carriers, a total of 100 trials are conducted for each concentration level

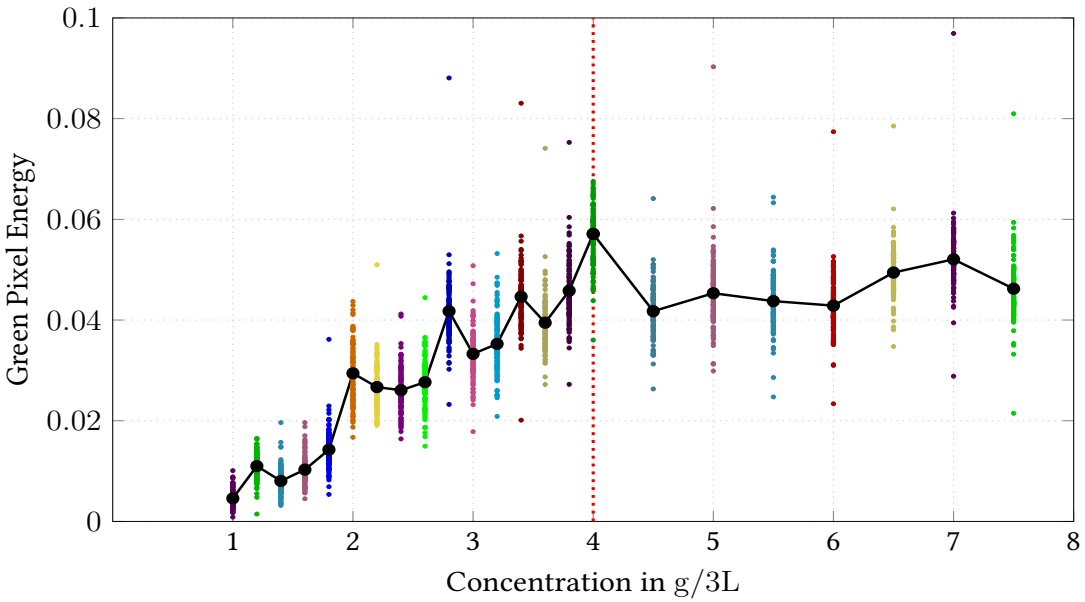


Figure 2.10: Measurement of the green pixel energy values for varying concentration levels of uranine in water-based solution, at a distance of 88 cm.

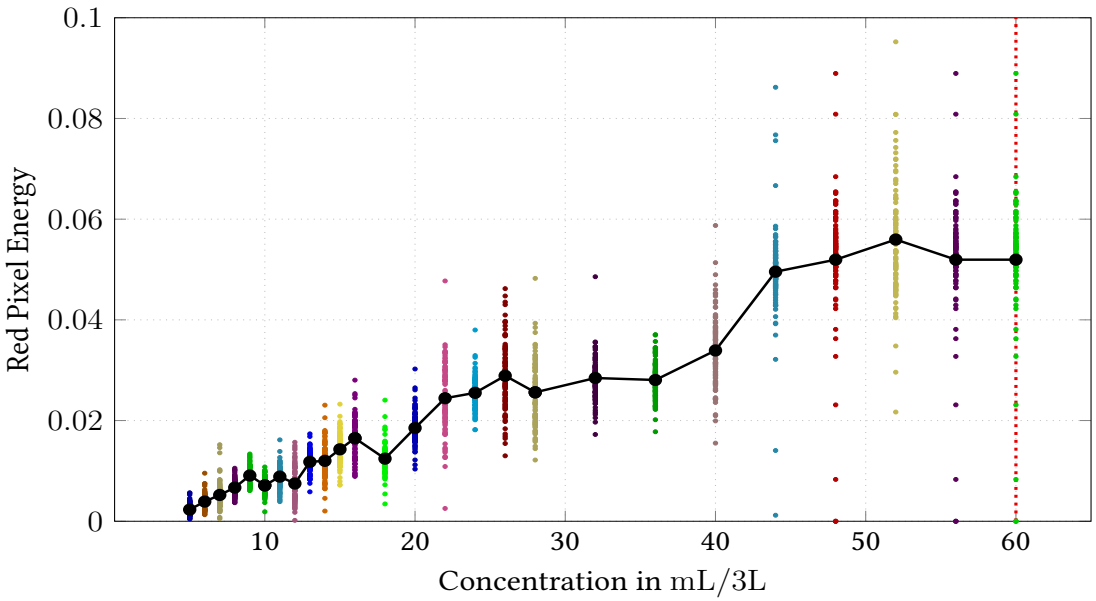
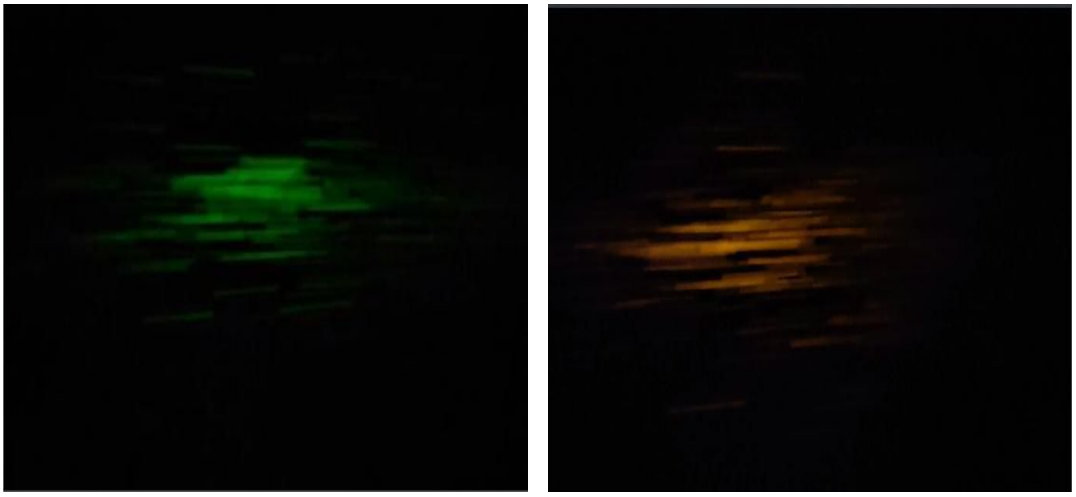


Figure 2.11: Measurement of the red pixel energy values for varying concentration levels of the highlighter ink in water-based solution, at a distance of 88 cm.

and the mean of the energy values is considered. Since uranine is available in its disodium salt form, the concentration is measured in g/3L. It is observed that the saturation level of the green pixel energy is achieved at a concentration level of about 4 g/3L, as shown by the red-dotted line in Figure 2.10. Beyond this point, there is no significant increase in the measured pixel energy values. Henceforth, in this dissertation, the said concentration level is considered for performing any measurements and analyses that involves water-based solution of uranine as the information carrier. Similarly, in case of the highlighter ink, the red pixel energy is measured. Since the highlighter ink is available in its concentrated aqueous solution form, the concentration is measured in mL/3L. It is observed that the red pixel energy continues to increase with increase in concentration level without achieving any apparent saturation level, as shown in Figure 2.11. This is due to the fact that the highlighter ink itself is in the form of a concentrated aqueous solution. Therefore, in its undiluted form, it would ideally produce the maximum pixel energy. However, due to limited availability of resources, in this dissertation, the concentration level of 60 mL/3L is considered for performing any measurements and analyses which involves water-based solution of the highlighter ink as the information carrier. Interestingly, the green pixel energy in water-based solution of uranine at a concentration level 4 g/3L is nearly the same as the red pixel energy of the water-based solution of the highlighter ink with a concentration of 60 mL/3L. The concentration level of the information carrier is also chosen on the basis of the SNR that it produces at the receiver side. This allows the sprayed water-based solution droplets to be conveniently detected by the camera-based detector at the farthest end of the testbed channel, as presented in Section 2.1.3.

With the aforementioned concentration levels of the information carriers, Figure 2.12(a) and Figure 2.12(b) depict uranine and the highlighter ink water-based solution droplets, respectively, when sprayed out of the sprayer at the transmitter side, and viewed through the camera at the receiver side at a distance of 88 cm. The spray duration at the transmitter side is $T_{\text{on}} = 0.02$ s, and is chosen in such a way that the sprayed droplets of the water-based solutions of the information carriers are clearly visible by the camera-based detector at the receiver side with a considerably high SNR at the farthest end of the testbed. The green pixel and the red pixel intensity values in case of uranine is shown in Figure. 2.13. As mentioned earlier, it is clearly seen that the green pixel intensity pulse is far more dominant than the almost negligible red pixel intensity pulse. In case of the highlighter ink, the maximum value of the red pixel intensity is nearly three times that of the maximum value of the green pixel intensity, as observed from Figure 2.14. For the said concentration levels, the maximum value of the green pixel intensity in case of uranine water-based

solution is almost the same as the maximum value of the red pixel intensity in case of the highlighter ink water-based solution.



(a) Uranine **(b)** Highlighter Ink

Figure 2.12: Camera snapshot of water-based solution droplets of uranine and the highlighter ink, at a distance of 88 cm under the influence of the ultraviolet-A flashlight.

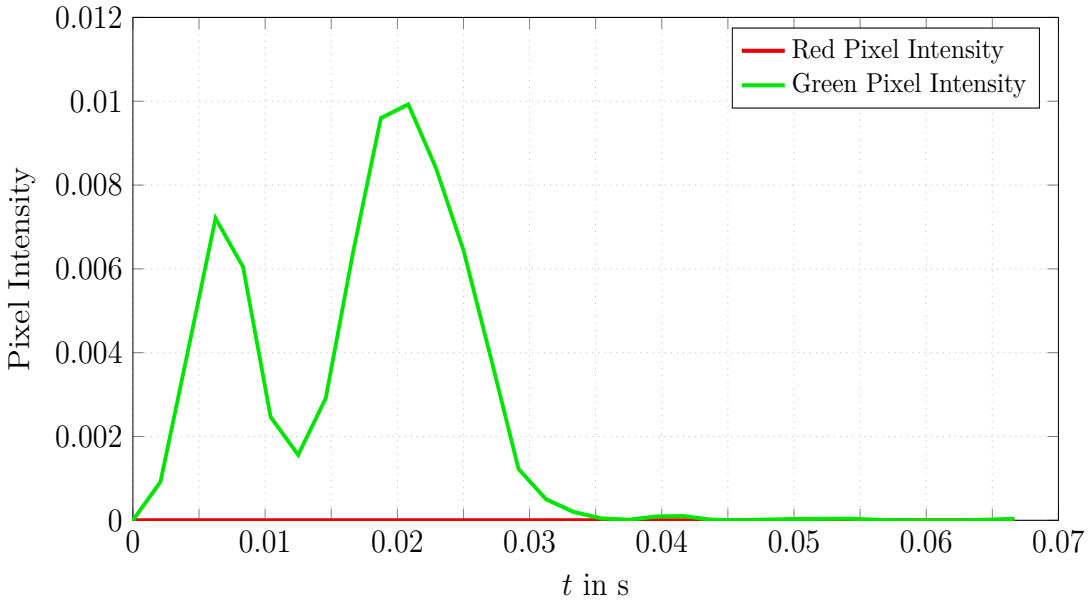


Figure 2.13: Observed green and red pixel intensity values over time for the sprayed water-based solution droplets of uranine, at a distance of 88 cm.

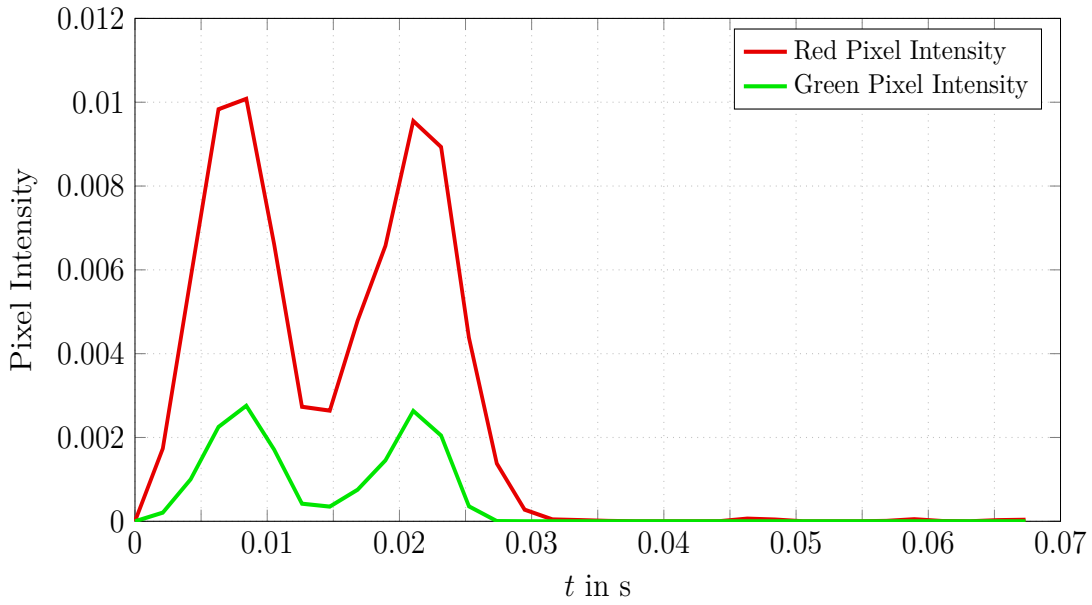


Figure 2.14: Observed green and red pixel intensity values over time for the sprayed water-based solution droplets of the highlighter ink, at a distance of 88 cm.

2.4 Receiver Performance Analysis

The receiver of a communication system plays a huge role when it comes to determining the system performance. It is, therefore, crucial to identify the correct type of receiver for the system under consideration in order to maximize the performance output in terms of the data transmission rates, low inter-symbol interference (ISI), low symbol error rates, etc. This is achieved by performing thorough analyses and investigations of various parameters depending on the type of the receiver.

2.4.1 Receiver Type

The choice of the receiver, to a large extent, depends on the physical quantity that is being measured, which is ultimately used to decipher the useful information sent by the transmitter. This can be done by assessing the interactive behavior of the receiver in relation to the particles carrying the information in a molecular communication system. In this regard, the receivers are broadly classified into two types:

- *Active Receiver* - An active receiver directly interacts with the information-carrying particles in the form of molecules/particles/ions. This allows the receiver to directly quantify these particles into a measurable physical unit. For instance, the molecular communication testbed using alcohol incorporates the *MQ-3* gas sensor as its receiver,

where the resistance of the thin layer of heated-up tin oxide present inside the gas sensor changes as the alcohol molecules come in contact with it, resulting in a measurable voltage [KF$^+$15; Koh89]. In another instance, a molecular communication system comprising of acids and bases for transmitting binary data symbols "0" and "1" encode information in the pH value, which is decoded at the receiver side by an active receiver interacting with the hydrogen and hydroxide ions called a pH probe. The measure of the presence of these ions is converted into a measurable voltage [FP$^+$17]. However, due to this direct interaction between the information particles and the receiver, the risk of it running into an early saturation is very high, resulting in "system nonlinearity" [FG$^+$13; FP$^+$17]. This makes the detection process more complex.

- *Passive Receiver* - A passive receiver, on the other hand, does not interact directly with the information carriers, but rather measures a specific physical property displayed by these particles under certain conditions from a certain distance. For example, a molecular communication system comprising of a culture of genetically engineered *Escherichia coli* bacteria is used that exhibits fluorescence upon receiving a specific signal molecule N-(3-Oxyhexanoyl)-L-homoserine lactone (C6-HSL). The bacteria produces a fluorescent response to the diffusion of C6-HSL molecules, which is then recorded with the help of a passive receiver in the form of fluorescence microscope [KA$^+$13]. In another example, a molecular communication testbed comprising of a droplet-based microfluidic system encodes data on water droplets with dissolved water colors [HM$^+$19]. For imaging and detection of these colored droplets, optical microscope with an integrated high-speed camera is used as a passive receiver. In such systems, due to the absence of direct interaction between the information-carrying particles and the receiver, the risk of the receiver running into saturation is minimum. However, the SNR at the receiver side may get impacted due to the distance between the spatial point where the information-carrying particles are present and the spatial point where the receiver is stationed. Typically, the detection process is less complex in such systems due to minimum or complete absence of nonlinear behavior of the receiver.

2.4.2 Receiver Performance Evaluation

When designing a testbed, the evaluation of the end-to-end system impulse response (henceforth, simply referred to as the system impulse response) performance of any communication

system plays a vital role. The system impulse response refers to the response that a complete end-to-end communication system comprising of the transmitter, the channel, and the receiver has on a Dirac delta function produced at the transmitter side. The block diagram of a molecular communication system model is presented in Figure 2.15 [GA21a; KF$^+$15]. A system with a sprayer as its transmitter, produces a short spray burst to emulate the Dirac delta function. The system impulse response, which is then obtained as a measurable output quantity at the receiver side, contains the combined influences of the transmitter, the channel, and the receiver on the spray burst. Consequently, the system impulse response can be used to compare two different communication systems comprising of the same transmitter and transmission channel, but employing different receivers. This allows the performance evaluation of the receiver through various parameters, often encountered in literature such as the responsitivity of the receiver, the effect of ISI, effects of the receiver reaching its saturation value while measuring a certain quantity, etc. These parameters ultimately decide the error robustness of the communication system, and the achievable data transmission rates with a significant influence on the communication algorithm (modulation/detection/coding) design. For the investigation of these parameters, the system impulse response $h(t)$ of the molecular communication system under investigation is observed over time t. A Dirac impulse in the form of a short spray burst is transmitted in order to trace the response of the system. The response is largely influenced by the responsitivity of the detection process of the communication system under consideration. Thus, taking $h(t)$ into account, two crucial performance metrics are identified in this work, which ensure a fair comparison between the respective responsitivity of the receivers. These are namely the peak's maximum to full-width at half maximum (PMFWHM) and the delay to peak's maximum (DPM). PMFWHM is the ratio between the maximum value of the system impulse response $h_{\max}(t)$ and the full-width at half maximum (FWHM), while DPM is the time τ taken by the system impulse response to reach its maximum value $h_{\max}(t)$ [FG$^+$13]. It is ideal to have PMFWHM to be as large as possible, which implies a much sharper $h(t)$ due to high responsitivity of the receiver. It also suggests that the maximum energy of the impulse is concentrated over a very short period of time. Additionally, it is desirable to have small DPM τ. This means that the time taken by the receiver to sense the very first incoming information carrier to the point of detection of its peak value should be as small as possible. A large value of τ would mean either the information carrier is being transmitted slowly (which affects the data transmission rates) and/or the receiver has a slow response time (a lag which leads to detection errors at the receiver side). Furthermore, interference to total received energy ratio (ITR) is a metric which defines the ratio of the

energy contained within the current symbol duration T of the impulse response $h(t)$ to the energy that is contained beyond T [KY+20]. This metric helps determine the effect of ISI originating from the transmission of the current symbol, which impairs the detection of the next transmitted symbol. ITR is mathematically represented as

$$\mathrm{ITR}(t = T) = \frac{A_{\mathrm{ISI}}}{A_{\mathrm{sym}} + A_{\mathrm{ISI}}} = \frac{\int_{t=T}^{\infty} h(t)\,dt}{\int_{0}^{\infty} h(t)\,dt}, \tag{2.1}$$

where A_{sym} is the total energy contained in $h(t)$, and A_{ISI} is the energy contained in $h(t)$ beyond T. The aforementioned performance metrics to determine the response of the receiver is illustrated in Figure 2.16.

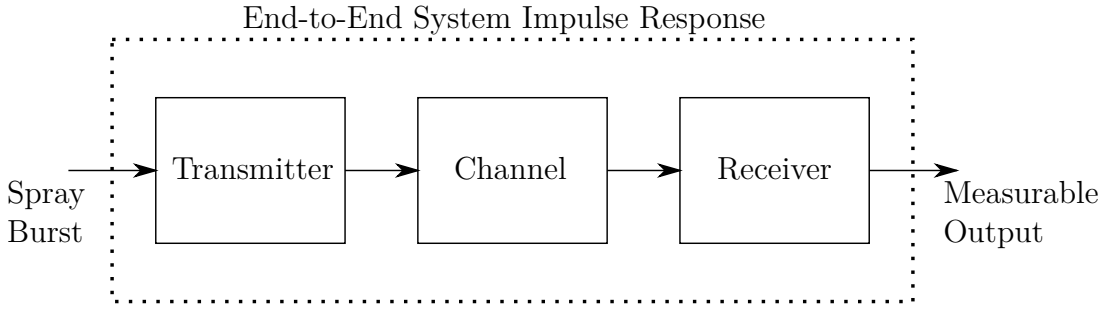


Figure 2.15: Block diagram of the system model of the testbed.

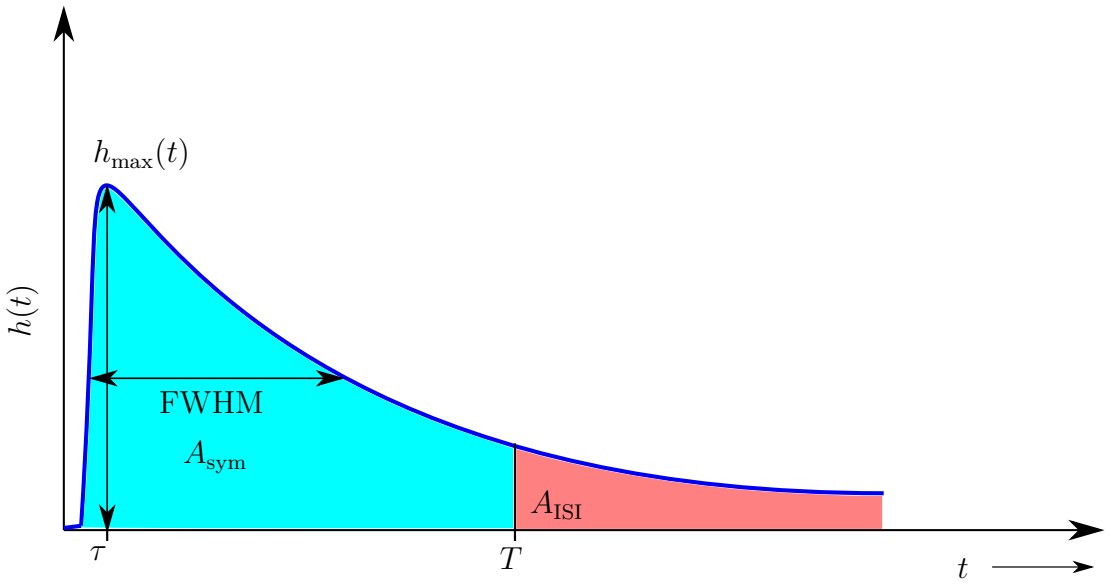


Figure 2.16: Depiction of the system impulse response $h(t)$ along with various performance metrics to determine system response.

In order to compare the system performance of the testbed presented in this work, the

traditionally implemented macroscopic molecular communication testbed using alcohol is considered as a benchmark [FG+13]. The testbed is implemented with modified units and components, which also previously served as a proof of concept for the work presented in [Dam20]. The setup consists of a *Gloria 3* pressure sprayer as the transmitter with an information carrier comprising of a 3.5% concentration of alcohol water-based solution. A fan is used at the transmitter side to transport the sprayed information carrier by the process of advection-diffusion through the channel. The transmission channel is open air space with no boundaries. An *MQ-3* gas sensor is used as the receiver to measure the concentration of the transported alcohol from the transmitter side. The *MQ-3* gas sensor directly interacts with the alcohol particles in the air producing a corresponding desired measurable output, making it an active receiver. The distance between the transmitter and the receiver is 88 cm, while the relative height between the spray nozzle at the transmitter side and the *MQ-3* sensor at the receiver side is 25 cm. The spray duration of $T_{\text{on}} = 0.5\,\text{s}$ is chosen, which produces a clear response at the receiver side with a high SNR, the sampling frequency of *Analog Discovery 2* being 100 Hz. The concentration of the detected alcohol is translated into a measurable voltage, which is considered for the investigation of its system impulse response. Due to a slow response time of the receiver, a total of 15 impulse responses are measured and averaged in order to determine the mean system impulse response, denoted by $h_{\text{alc}}(t)$ as shown in Figure 2.17. In order to make a fair comparison with the system impulse responses of the other systems, $h_{\text{alc}}(t)$ is normalized to its maximum value. It may be mentioned that the symbol duration of $T = 4\,\text{s}$ is chosen for the data transmission in [Dam20]. This period is, however, insufficient for the voltage of the alcohol sensor to fall back to at least half its maximum value after binary data symbol "1" is transmitted. With transmission of long sequences of 1's in quick succession, the sensor quickly runs into saturation, which is to be compensated. This leads to an increased probability of detecting the transmitted symbols erroneously. The $h_{\text{alc}}(t)$ presented in Figure 2.17 takes the compensation of the effect of saturation into account, in accordance with the procedure described in [Dam20]. The problem is also tackled by increasing the T but the data transmission rate suffers, which is seen as a trade-off.

A similar setup is prepared with a macroscopic molecular communication testbed using water-based fluorescent dye solution. The distance is set at 88 cm between the transmitter and the receiver, and a relative height of 25 cm is set between the spray nozzle at the transmitter side, and the camera-based detector at the receiver side. In this case, however, the response produced as pixel intensity value in the camera-based detector is not due to a direct interaction between the sprayed water-based dye solution droplets of the information

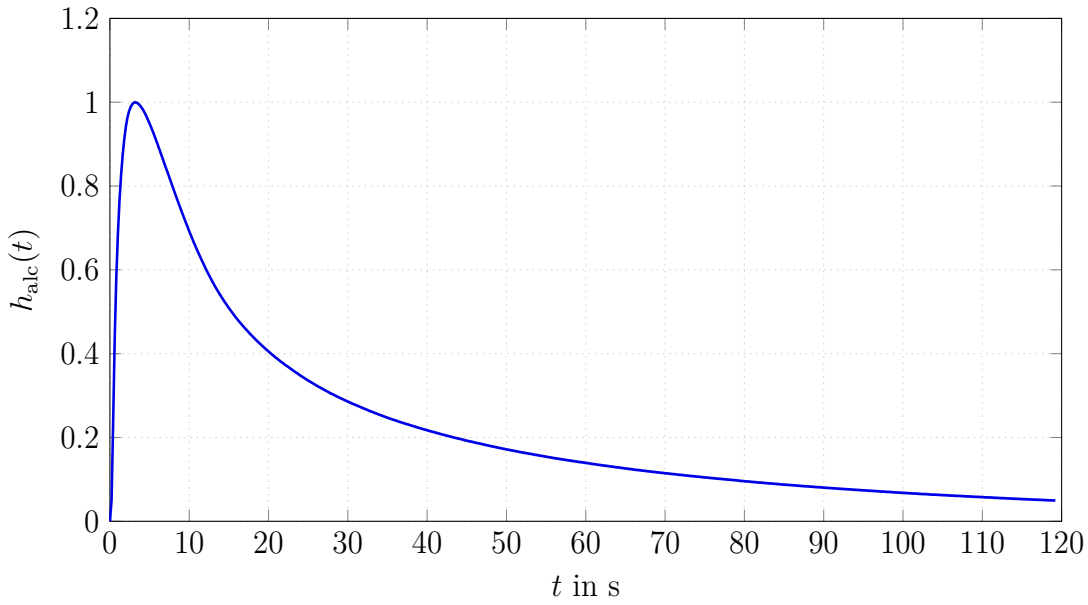


Figure 2.17: Normalized mean system impulse response $h_{\text{alc}}(t)$ over time t for a macroscopic molecular communication testbed using alcohol, at a distance of 88 cm.

carrier and the detector itself. This makes the camera-based detector to be a passive receiver. Furthermore, a spray burst of duration $T_{\text{on}} = 0.02\,\text{s}$ is considered in order to release the information carrier into the transmission channel tube, which is a closed bounded structure. The chosen value of T_{on} still produces a significantly high SNR, which makes it possible to perform detection of the sprayed droplets of the information carrier at the farthest end of the tube. In the case of uranine water-based solution, the green pixel intensity recorded by the camera-based detector at the receiver side is considered for the investigation of its system impulse response $h_{\text{uran}}(t)$. Similarly, in the case of highlighter ink water-based solution, the red pixel intensity recorded by the camera-based detector is considered for the investigation of its system impulse response $h_{\text{ink}}(t)$. In order to make a fair comparison with the measurements of the testbed using alcohol, a total of 15 impulse responses are measured and averaged. Thus, the mean system impulse response is determined and is denoted by $h_{\text{uran}}(t)$ and $h_{\text{ink}}(t)$, as shown in Figure 2.18 and Figure 2.19, respectively. Additionally, $h_{\text{uran}}(t)$ and $h_{\text{ink}}(t)$ are also normalized to their maximum values similar to $h_{\text{alc}}(t)$ for the testbed using alcohol. It is to be noted that the chosen symbol duration for the transmission of data in the testbed using water-based fluorescent dye solution is $T = 0.05\,\text{s}$. As mentioned earlier in the chapter (Section 2.1.1), the minimum shutting time of the valve is $T_{\text{trans}} = 0.03\,\text{s}$, which is the minimum time taken by the solenoid valve to transition from the "on" state to the "off" state. Therefore, $T = T_{\text{on}} + T_{\text{trans}}$.

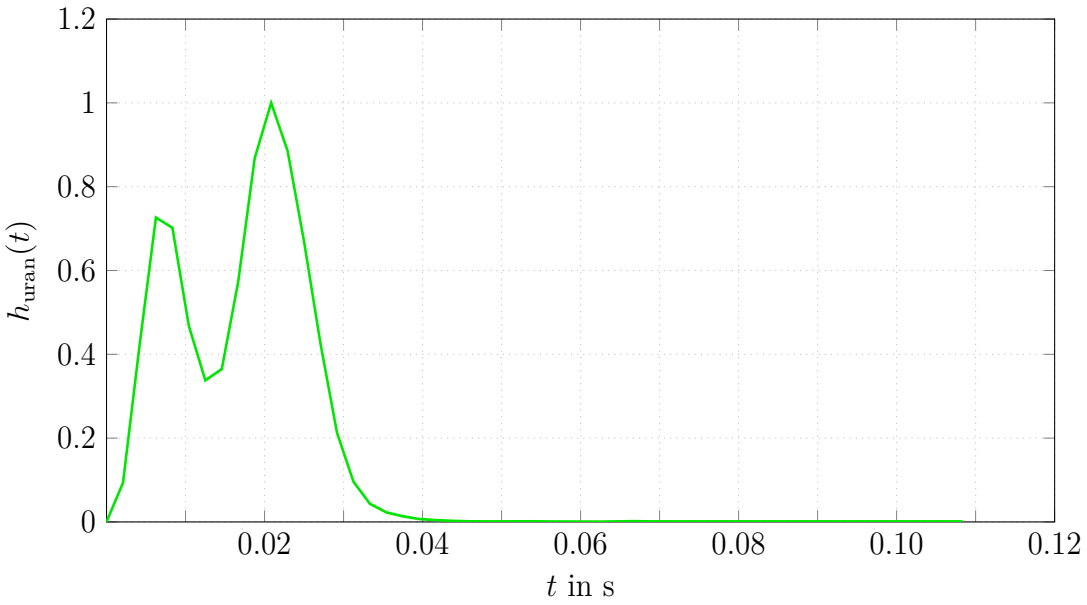


Figure 2.18: Normalized mean system impulse response $h_{\text{uran}}(t)$ over time t for a macroscopic molecular communication testbed using uranine, at a distance of 88 cm.

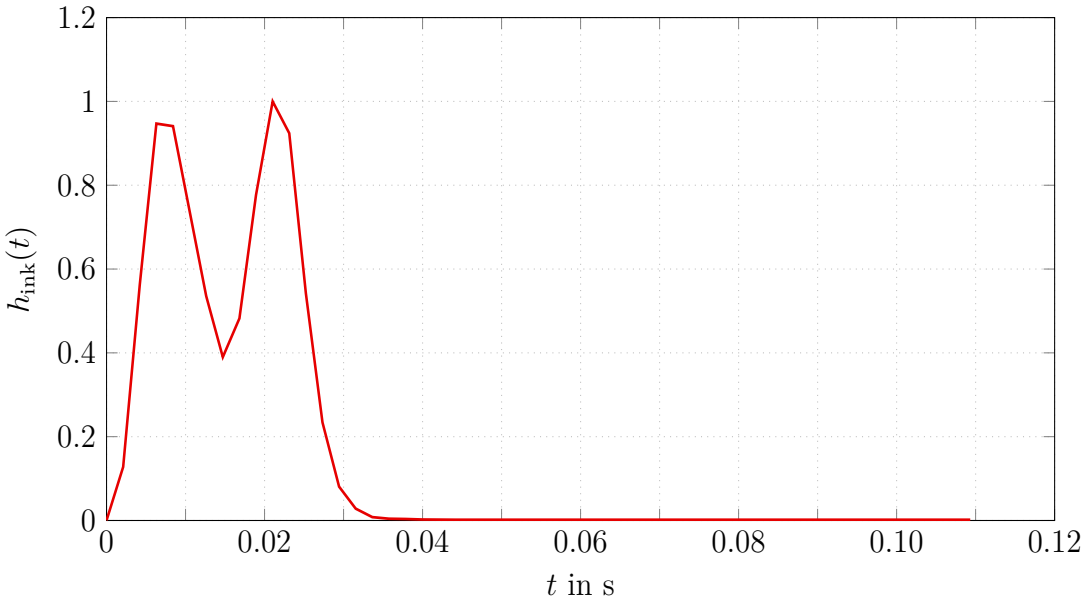


Figure 2.19: Normalized mean system impulse response $h_{\text{ink}}(t)$ over time t for a macroscopic molecular communication testbed using the highlighter ink, at a distance of 88 cm.

Keeping all the discussed parameters and restrictions in mind, Table 2.1 summarizes the performance metrics between the receivers of the testbeds using water-based alcohol and fluorescent dye solutions. It is observed that PMFWHM (measured in unit/s) for the system using fluorescent dye is much higher than the system using alcohol by a factor of almost 1000. Furthermore, DPM is achieved much faster in systems using fluorescent dye over the system using alcohol. Additionally, ITR in the system using alcohol is more than a factor of 100, as compared to ITR in the systems using fluorescent dyes. All these performance metrics make the testbed using fluorescent dye an obvious choice for conducting the macroscopic air-based molecular communication related investigations and analyses over the traditional testbed using alcohol.

Table 2.1: Performance metrics of the system impulse responses of the testbed receivers.

Testbed	Symbol Duration T (s)	PMFWHM (unit/s)	DPM τ (s)	ITR
Alcohol	4	0.0683	3.22	0.8875
Uranine	0.05	45	0.0208	0.0044
Highlighter ink	0.05	47	0.0210	0.0061

2.5 Summary

In this chapter, the testbed design and implementation of the air-based macroscopic molecular communication using fluorescent dyes is presented. Additionally, each testbed unit and component, which primarily comprises of the transmitter, the information carrier, the transmission channel, and the receiver are described in detail.

The chapter also makes an effort to compare the molecular communication testbeds using fluorescent dye that includes water-based solutions of uranine and the highlighter ink with the conventional molecular communication testbed using water-based solution of alcohol. The detailed analyses of proven advantages and improvements of the former against the latter with respect to various performance parameters like PMFWHM, DPM, and ITR are discussed in detail. In conclusion, molecular communication testbed using fluorescent dye has been considered for the first time to establish the complete air-based macroscopic molecular communication theory, presented in the subsequent chapters.

3

System Modeling

In this chapter, the theory behind particle-based molecular communication is presented, which mainly defines the mass transportation process of the information carriers over a given channel based on the scale of its application. Additionally, the theoretical models for the transmission of the information carriers in the form of sprayed droplets of uranine and the highlighter ink water-based solutions are defined for the testbed under investigation in this dissertation. The dominant physical effects, which influence the propagation of the water-based solution droplets of the information carriers through the channel are taken into account. Furthermore, the end-to-end system model is considered to model the system impulse response originating from the combined effects of the transmitter, the propagation channel as well as the receiver, along with the noise process model arising due to the imperfect design of these fundamental units that make up the testbed. Finally, the system impulse response and the noise models are considered to define the complete analytical model of the testbed, which is then validated by transmission of a message sequence through measurement trials inside the testbed.

3.1 Theory of Molecular Communication

The idea of having a communication system is to successfully achieve transfer of information over time and space. This is only possible when a system comprising of a transmitter generates information in the form of some physically quantifiable entity, which is then transmitted over a propagation channel and ultimately measured by the receiver. The physical quantity measured at the receiver side is then interpreted in a meaningful way, which helps to decipher the data that has been transmitted. Conventional communication

systems typically use wires and cables to connect both the transmitter and the receiver tangibly, over which a physically quantifiable entity in the form of an electro magnetic (EM) wave propagates with the useful information. Similarly, in another traditional method of communication using the open-air free space, the EM waves successfully transfer the information through the air medium between the transmitter and the receiver, without any physical connection between the two. However, the speed of transfer of the information is comparable to the speed of light in a given medium. For example, in air, the speed of light is given by $c_{\text{light}} = 299\,702\,547\,\text{m/s}$, and is typically determined by the refractive index of the transmission channel medium.

The rate of information transfer in molecular communication, on the other hand, depends on the speed at which physical transfer of the individual particles carrying the information takes place in a given medium. This typically occurs at a velocity $v << c_{\text{light}}$. However, the propagation of these particles carrying the information depends on the dimensions of the major units that form the molecular communication system. These include the sizes of the transmitter and the receiver, the dimensions and medium of the propagation channel, the physical distance between the transmitter and the receiver, etc. Therefore, based on the aforementioned aspects as well as the scale at which the application takes place, molecular communication can be mainly classified into two types:

- Micro-scale molecular communication applications, which occur in nanometer to micrometer range, for example, as intra and intercellular communication inside human bodies [Alb10]

- Macro-scale molecular communication applications that occur in centimeter to meter range, for example, long distance communication between social insects using pheromones [Ago92]

A comprehensive theoretical framework of molecular communication concerning the testbed presented in this work would be incomplete without mentioning some of the most common phenomena facilitating the transportation of information-carrying particles from the transmitter to the receiver through a transmission channel. These processes include:

- *Diffusion* - Diffusion is a mass transportation process, where the particles move randomly in a fluid-based medium by colliding against each other [Ber93]. The particles mainly utilize the thermal energy of the environment, which can be used in a molecular communication system to power the transmission process of the information-carrying particles from the transmitter to the receiver [AB+08].

- *Advection* - Advection is a faster alternative to the diffusion-based mass transportation process for moving the information carrier particles from the transmitter to the receiver over larger distances [HV07]. This is achieved by introducing flow into the environment in the direction of the receiver [KA09; KA$^+$12]. In the process, additional energy from external sources are necessary for introducing the velocity component into the propagation medium.

- *Turbulent Flow* - Turbulent flows originate due to the random motion of the particles that constitute the fluid propagation medium, over which, the information-carrying particles are transmitted with a certain bulk flow [Csa73]. This random motion of the fluid propagation medium is very similar to the diffusion process of the information-carrying particles. The major difference lies in the fact that in the diffusion process, the information-carrying particles are transported from the transmitter to the receiver due to their own random motion, while in turbulent flow, the random motion of the propagation medium particles helps to transport these information-carrying particles along with the presence of the bulk flow. The turbulence is represented by a unitless quantity called the Reynolds number, which is dependent on the velocity and the viscosity of the fluid propagation medium. Turbulent flow is also noticeable when a stream of high velocity fluid is introduced into a low-velocity fluid. A molecular communication system employing turbulent flow is presented in [AE$^+$18].

- *Convection* - Convection is the mass transportation process of particles of the information carrier in a fluid-based environment caused due to the presence of a temperature gradient in the environment, and can be seen as a combination of diffusion, advection, and turbulent flow. The temperature gradient leads to the thermal expansion and buoyancy in the fluid, which results in the transportation of particles from the region of high pressure to the region of low pressure [MY$^+$95]. The information particles in a molecular communication system could also be transported along with the flow of the fluid medium if the transmitter is situated at a high pressure region, and the receiver is situated at a low pressure region [GA$^+$20].

- *Dispersion* - Dispersion is the process that occurs due to the unequal flow pathways taken by the particles of the fluid propagation medium, when passing through a porous medium such as soil [JD$^+$84], or the different velocities at which fluid particles travel due to the non-uniform laminar flow inside a duct [Tay53]. This process helps to transport the information particles from the transmitter to the receiver that also get dispersed with the particles of the propagation medium, as discussed in the molecular

communication system in [WS+18].

However, in a molecular communication system comprising of an air-based propagation channel medium, the particles of the information carrier in their liquid-based solution form are released out of the transmitter sprayer via pressurized air. These information-carrying droplets are predominantly transported over a certain distance between the transmitter and the receiver through diffusion, advection and/or turbulent flow mass transportation processes. In this chapter, the mass transportation processes with respect to micro-scale molecular communication are described in detail along with their mathematical models. The models are then utilized to formulate the air-based molecular communication in the macro-scale range.

3.2 Mass Transportation Processes in Molecular Communication

3.2.1 Diffusion

Diffusion is the random movement of particles such as molecules, particles and/or ions through random collisions with one another due to the presence of the thermal energy for temperatures above absolute zero [Cus09]. These particles are suspended in a fluid-based medium, and the phenomenon is also referred to as the Brownian motion [NE+13]. The diffusion process is seen as one of the potential methods for transporting information particles from the transmitter to the receiver in a molecular communication system. Its usefulness comes in the form of its energy efficiency. The energy requirement for the propagation of the information particles is solely harnessed from its propagation environment in the form of thermal energy, without the need of any additional external sources of energy. Diffusion-based applications are commonly observed in cellular level biological processes such as in the nerve cells, which release neurotransmitter molecules acetylcholine (ACh) into the neuromuscular junction to transmit messages to the muscle cells [KY+13]. In another example, glucose is released into the bloodstream via diffusion [CY+18]. The epithelial cells of the small intestine absorb glucose molecules through digestion of dietary carbohydrates. These molecules are then released into the bloodstream via diffusion, which are then carried to the cells with the help of glucose transporters. Other examples include deoxyribonucleic acid replication [PS+16], diffusion process in kidneys [FO+00], etc.

The rate of diffusion, also termed as diffusivity, is unique to the physical state of the propagation medium through which the particles are free to move and spread. This phenomenon

is mainly observed in fluid (aqueous and gaseous) media. For instance, the diffusion coefficient in an aqueous medium can be estimated using the Stokes-Einstein equation [KD95] for the size of the diffusing molecules larger than the size of the molecules of the propagation aqueous medium, and is given by

$$D = \frac{k_{\mathrm{B}}\vartheta}{6\pi\eta r_{\mathrm{mol}}}, \tag{3.1}$$

and for the size of the diffusing molecules which have approximately the same size as the molecules of the propagation aqueous medium, is given by

$$D = \frac{k_{\mathrm{B}}\vartheta}{4\pi\eta r_{\mathrm{mol}}}, \tag{3.2}$$

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 aqueous medium, respectively, and r_{mol} is the radius of the diffusing molecule. Similarly, in the gaseous medium, the diffusion coefficient [Kos82] can be expressed as

$$D = \frac{2\vartheta^{3/2}}{3P_{\mathrm{gas}}d_{\mathrm{gas}}^2}\sqrt{\frac{k_{\mathrm{B}}^3}{\pi^3 m_{\mathrm{gas}}}}, \tag{3.3}$$

where P_{gas} is the pressure of the gas, d_{gas} is the diameter of the gas molecule, and m_{gas} is the molar mass of the gas.

To model the transfer of information particles via the diffusion process, a three dimensional scenario is assumed, as shown in Figure 3.1. The unbounded environment is filled with a fluid-based medium of diffusion coefficient D. The described scenario consists of an infinitesimally small point transmitter, which is located in the coordinates origin of the three dimensional space. The interaction of the released information-carrying particles with each other or the homogeneous fluid-based medium is neglected with absence of flow. The diffusion process can then be modeled using Monte Carlo simulations [FY+16; BS10]. The Brownian motion emulating the diffusion process is well described by a Wiener process $W(t)$ at time t, with spatial increments of $W(t+\Delta t)-W(t)$ in every direction, governed by the distribution $\mathcal{N}(0, 2D\Delta t)$, where $\mathcal{N}(\mu, \sigma^2)$ is a Gausssian distribution with mean μ and variance σ^2. For a small step size of $\Delta t \to 0$, a three-dimensional random walk converges

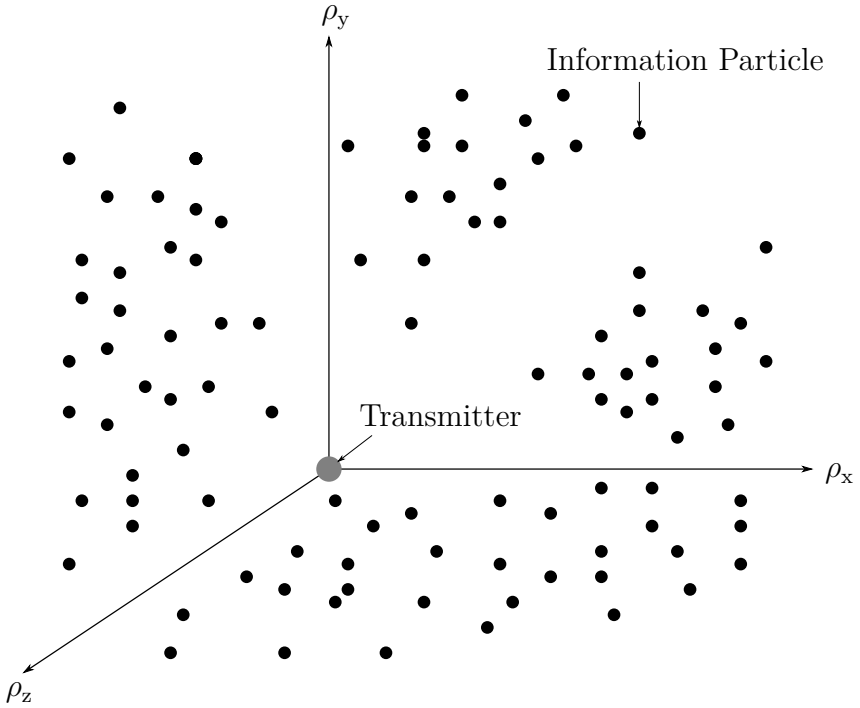


Figure 3.1: Diffusion-based molecular communication scenario.

to the Wiener process [NE+13], and is given by

$$\begin{aligned}\left(\rho_x\left[i_{sim}\right],\rho_y\left[i_{sim}\right],\rho_z\left[i_{sim}\right]\right) &= \left(\rho_x\left[i_{sim}-1\right],\rho_y\left[i_{sim}-1\right],\rho_z\left[i_{sim}-1\right]\right) \\ &\quad+\left(\Delta\rho_x,\Delta\rho_y,\Delta\rho_z\right),\end{aligned} \tag{3.4}$$

$$\Delta\rho_x \sim \mathcal{N}(0,2D\Delta t), \tag{3.5}$$

$$\Delta\rho_y \sim \mathcal{N}(0,2D\Delta t), \tag{3.6}$$

$$\Delta\rho_z \sim \mathcal{N}(0,2D\Delta t), \tag{3.7}$$

where $i_{sim} \in \mathbb{N}_0$ represents the simulation step, $\left(\rho_x\left[i_{sim}\right],\rho_y\left[i_{sim}\right],\rho_z\left[i_{sim}\right]\right)$ is the position of the information particle at time instant $i_{sim}\Delta t$, and $\left(\rho_x\left[0\right],\rho_y\left[0\right],\rho_z\left[0\right]\right)$ is the initial position of the information particle at time instant $t = 0$. Additionally, the diffusion process can also be modeled in a spherical coordinate system as

$$\begin{bmatrix}\rho_x\left[i_{sim}\right]\\ \rho_y\left[i_{sim}\right]\\ \rho_z\left[i_{sim}\right]\end{bmatrix} = \begin{bmatrix}\rho_x\left[i_{sim}-1\right]\\ \rho_y\left[i_{sim}-1\right]\\ \rho_z\left[i_{sim}-1\right]\end{bmatrix} + \Delta\rho\begin{bmatrix}\sin\theta_{i_{sim}}\cos\varphi_{i_{sim}}\\ \sin\theta_{i_{sim}}\sin\varphi_{i_{sim}}\\ \cos\theta_{i_{sim}}\end{bmatrix}, \tag{3.8}$$

where $\Delta\rho$ is the displacement of the information particle over the time interval Δt, and

$\varphi_{i_{\text{sim}}}$ and $\theta_{i_{\text{sim}}}$ describe the three dimensional angular displacements (azimuthal and polar) over time interval Δt.

Another method to model the diffusion process is by using the analytical approach with the help of partial differential equation (PDE). This is represented by Fick's second law, which results in mathematical expressions for one dimension, two dimensions and three dimensions as [SK00; Cra79]

$$\frac{\partial \phi\left(\boldsymbol{\rho} = [\rho_{\text{x}}], t\right)}{\partial t} = D\left(\frac{\partial^2 \phi\left(\rho_{\text{x}}, t\right)}{\partial \rho_{\text{x}}^2}\right), \tag{3.9}$$

$$\frac{\partial \phi\left(\boldsymbol{\rho} = \left[\rho_{\text{x}}, \rho_{\text{y}}\right], t\right)}{\partial t} = D\left(\frac{\partial^2 \phi\left(\rho_{\text{x}}, \rho_{\text{y}}, t\right)}{\partial \rho_{\text{x}}^2} + \frac{\partial^2 \phi\left(\rho_{\text{x}}, \rho_{\text{y}}, t\right)}{\partial \rho_{\text{y}}^2}\right), \tag{3.10}$$

$$\frac{\partial \phi\left(\boldsymbol{\rho} = \left[\rho_{\text{x}}, \rho_{\text{y}}, \rho_{\text{z}}\right], t\right)}{\partial t} = D\left(\frac{\partial^2 \phi\left(\rho_{\text{x}}, \rho_{\text{y}}, \rho_{\text{z}}, t\right)}{\partial \rho_{\text{x}}^2} + \frac{\partial^2 \phi\left(\rho_{\text{x}}, \rho_{\text{y}}, \rho_{\text{z}}, t\right)}{\partial \rho_{\text{y}}^2} + \frac{\partial^2 \phi\left(\rho_{\text{x}}, \rho_{\text{y}}, \rho_{\text{z}}, t\right)}{\partial \rho_{\text{z}}^2}\right), \tag{3.11}$$

where D is the diffusion coefficient, $\phi(\boldsymbol{\rho}, t)$ is the concentration at a given point in space, defined by the dimension dependent location vector $\boldsymbol{\rho}$ at time instant t.

To solve the PDEs given by (3.9), (3.10), and (3.11), a simple but useful initial condition is considered, which corresponds to the release of molecules from a point transmitter [Far15], where it is assumed that M_0 molecules are released by the point transmitter at time $t = 0$. Thus, the initial condition for diffusion process for a dimension dependent space vector can be given as

$$\phi\left(\boldsymbol{\rho}, t\right) = M_0 \delta(\boldsymbol{\rho}), \tag{3.12}$$

where $\delta(\boldsymbol{\rho})$ represents the Dirac delta function for the dimension dependent location vector $\boldsymbol{\rho}$, which is defined as

$$\int_{-\infty}^{\infty} \delta(\boldsymbol{\rho}) d\boldsymbol{\rho} = 1. \tag{3.13}$$

The boundary condition is considered as

$$\lim_{\boldsymbol{\rho} \to \infty} \phi\left(\boldsymbol{\rho}, t\right) = 0, \tag{3.14}$$

where it is assumed that the concentration of the information particles at the far end in space, represented by the dimension dependent space vector $\boldsymbol{\rho}$ is negligible. Solving

the PDEs by considering the stated initial and boundary conditions lead to the following solutions for the diffusion process in one dimension, two dimensions and three dimensions, respectively, and are given as

$$\phi\left(\boldsymbol{\rho} = [\rho_{\mathrm{x}}], t\right) = \frac{M_0}{\sqrt{4\pi D t}} \exp\left(-\frac{\rho_{\mathrm{x}}^2}{4Dt}\right), \tag{3.15}$$

$$\phi\left(\boldsymbol{\rho} = \left[\rho_{\mathrm{x}}, \rho_{\mathrm{y}}\right], t\right) = \frac{M_0}{\sqrt{(4\pi D t)^2}} \exp\left(-\frac{\rho_{\mathrm{x}}^2 + \rho_{\mathrm{y}}^2}{4Dt}\right), \tag{3.16}$$

$$\phi\left(\boldsymbol{\rho} = \left[\rho_{\mathrm{x}}, \rho_{\mathrm{y}}, \rho_{\mathrm{z}}\right], t\right) = \frac{M_0}{\sqrt{(4\pi D t)^3}} \exp\left(-\frac{\rho_{\mathrm{x}}^2 + \rho_{\mathrm{y}}^2 + \rho_{\mathrm{z}}^2}{4Dt}\right). \tag{3.17}$$

The complete derivation of the solution can be found in [McG20]. The diffusion equation can also be solved in many other ways considering different initial and boundary conditions in order to obtain the channel impulse response [Ber93]. This could particularly be solved for the absorbing receivers [Red01]. The capture function of the information particles for a certain receiver in space can also be obtained by integrating the channel impulse response from the PDE over a certain boundary [YH$^+$14]. Also, by differentiating the channel impulse response with respect to time, it is possible to obtain the maximum concentration [YH$^+$14; LA$^+$11].

3.2.2 Advection-Diffusion

To accelerate the process of diffusion, a flow component is introduced in the fluid-based propagation medium, which is termed as the process of advection [AB86]. This is especially helpful in transporting particles over larger distances with a more channelized approach to reach a certain target. Together with diffusion, it is termed as the advection-diffusion process. Advection-based phenomena are often encountered in the nature such as tissue oxygen delivery system by a micro-vascular network inside the human body [FS$^+$08]. In plants, the nutrient uptake by the roots from the groundwater is also seen to be an advection-diffusion process [Avh18]. In a much larger context, advection-diffusion is also responsible for the flow of the wind, which facilitates the process of pollination in plants, enabling fertilization [Bar11]. Furthermore, the distribution of air pollution to monitor the time and space dependent concentration of pollutants across a certain region from the source is also understood better by applying the advection-diffusion models [UA$^+$17]. Based on the diffusion process, the advection-diffusion process can be modeled by introducing flow

velocity components in each spatial dimension, and can be expressed as

$$\begin{bmatrix} \rho_{\mathrm{x}}\left[i_{\mathrm{sim}}\right] \\ \rho_{\mathrm{y}}\left[i_{\mathrm{sim}}\right] \\ \rho_{\mathrm{z}}\left[i_{\mathrm{sim}}\right] \end{bmatrix} = \Delta t \begin{bmatrix} v_{\rho_{\mathrm{x}}}\left[i_{\mathrm{sim}}-1\right] \\ v_{\rho_{\mathrm{y}}}\left[i_{\mathrm{sim}}-1\right] \\ v_{\rho_{\mathrm{z}}}\left[i_{\mathrm{sim}}-1\right] \end{bmatrix} + \begin{bmatrix} \rho_{\mathrm{x}}\left[i_{\mathrm{sim}}-1\right] \\ \rho_{\mathrm{y}}\left[i_{\mathrm{sim}}-1\right] \\ \rho_{\mathrm{z}}\left[i_{\mathrm{sim}}-1\right] \end{bmatrix} + \Delta\rho \begin{bmatrix} \sin\theta_{i_{\mathrm{sim}}} \cos\varphi_{i_{\mathrm{sim}}} \\ \sin\theta_{i_{\mathrm{sim}}} \sin\varphi_{i_{\mathrm{sim}}} \\ \cos\theta_{i_{\mathrm{sim}}} \end{bmatrix}, \tag{3.18}$$

where $v_{\rho_{\mathrm{x}}}$, $v_{\rho_{\mathrm{y}}}$ and $v_{\rho_{\mathrm{z}}}$ represent the flow velocity components in spatial dimensions ρ_{x}, ρ_{y} and ρ_{z}, respectively.

Additionally, the advection-diffusion process can also be modeled using analytical approach. In order to model the propagation of the information particles, the flow due to the advection process is defined as

$$\mathbf{v} \cdot \nabla = v_{\rho_{\mathrm{x}}} \frac{\partial}{\partial x} + v_{\rho_{\mathrm{y}}} \frac{\partial}{\partial y} + v_{\rho_{\mathrm{z}}} \frac{\partial}{\partial z}, \tag{3.19}$$

where $\mathbf{v} = \left[v_{\rho_{\mathrm{x}}}, v_{\rho_{\mathrm{y}}}, v_{\rho_{\mathrm{z}}}\right]$ is the velocity vector in a three dimensional space, and $\nabla = \frac{\partial}{\partial x} + \frac{\partial}{\partial y} + \frac{\partial}{\partial z}$ is the divergence operator.

Therefore, the transport of the information particles through the advection process modeling the flow of the propagation environment, with the help of the velocity vector in three dimensional space is expressed using the advection equation PDE [JA+19] as

$$\frac{\partial \phi(\boldsymbol{\rho}, t)}{\partial t} = -\nabla \cdot (\mathbf{v}\phi(\boldsymbol{\rho}, t)). \tag{3.20}$$

Finally, the Fick's second law for the process of diffusion is extended to include the advection process too, resulting in the mathematical expression for the advection-diffusion process, which is given as

$$\frac{\partial \phi(\boldsymbol{\rho}, t)}{\partial t} = D\nabla^2 \phi(\boldsymbol{\rho}, t) - \nabla \cdot (\mathbf{v}\phi(\boldsymbol{\rho}, t)), \tag{3.21}$$

where $\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ is the Laplacian operator.

With the same initial condition as described in (3.12) and the boundary condition as

described in (3.14), the final solutions for the different dimensions can be given as

$$\phi\left(\boldsymbol{\rho}=[\rho_{\mathrm{x}}],t\right)=\frac{M_0}{\sqrt{4\pi Dt}}\exp\left(-\frac{\left(\rho_{\mathrm{x}}-v_{\rho_{\mathrm{x}}}t\right)^2}{4Dt}\right), \tag{3.22}$$

$$\phi\left(\boldsymbol{\rho}=\left[\rho_{\mathrm{x}},\rho_{\mathrm{y}}\right],t\right)=\frac{M_0}{\sqrt{\left(4\pi Dt\right)^2}}\exp\left(\frac{-\left(\rho_{\mathrm{x}}-v_{\rho_{\mathrm{x}}}t\right)^2-\left(\rho_{\mathrm{y}}-v_{\rho_{\mathrm{y}}}t\right)^2}{4Dt}\right), \tag{3.23}$$

$$\phi\left(\boldsymbol{\rho}=\left[\rho_{\mathrm{x}},\rho_{\mathrm{y}},\rho_{\mathrm{z}}\right],t\right)=\frac{M_0}{\sqrt{\left(4\pi Dt\right)^3}}\exp\left(\frac{-\left(\rho_{\mathrm{x}}-v_{\rho_{\mathrm{x}}}t\right)^2-\left(\rho_{\mathrm{y}}-v_{\rho_{\mathrm{y}}}t\right)^2-\left(\rho_{\mathrm{z}}-v_{\rho_{\mathrm{z}}}t\right)^2}{4Dt}\right). \tag{3.24}$$

3.2.3 Turbulent Flow

Turbulent flow is another mass transportation process in which the bulk flow present in the fluid propagation medium undergoes chaotic mixing, in contrast to laminar flow, where the flow of the fluid propagation medium takes place smoothly over multiple layers [MS00]. The fluid propagation medium eddies around the propagating particles, which makes the particles of the propagation medium at any point to undergo a random change in velocity (both in magnitude as well as in direction), while the majority of the bulk flow takes place along a certain direction [BS+15]. This phenomenon is termed as the turbulent flow, where the random motion of the particles of the fluid propagation medium is very similar to the diffusion process. The difference, however, lies in the fact that in the diffusion process, the propagating particles move randomly due to the thermal energy present in the environment, while in the case of the turbulent flow, the randomness occurs in the particles of the fluid propagation medium itself. A critical number termed as the Reynold's number is used to quantify the presence of turbulence, which is defined as the ratio of the forces due to inertia to the forces due to viscosity of the fluid propagation medium [Rot90], which is given as

$$\mathrm{Re}=\frac{d_{\mathrm{eff}}v_{\mathrm{fluid}}}{\eta}, \tag{3.25}$$

where d_{eff} is the effective linear dimension of the environment, v_{fluid} is the flow velocity in the fluid propagation medium, and η is the kinematic viscosity of the fluid propagation medium. The critical value for Reynold's number commonly encountered in the literature is $\mathrm{Re}=2300$. The flow is laminar for $\mathrm{Re}<2300$ and turbulent for $\mathrm{Re}>=2300$. Turbulent flow is also observed in regular day-to-day life scenarios such as the blood flow in the

arteries of the heart. Any curves or obstructions in the aorta causes the blood flow to become more turbulent [SS76]. In another example, the smoke particles emanating from the ignited tip of an incense stick display a smooth laminar flow for shorter distances, while the flow becomes more chaotic and turbulent for larger distances [IT92]. Similarly, turbulent flow is also observed in flowing rivers due to the presence of obstacles, which turn the laminar flow into turbulent flow [Yok67]. Turbulence in the water increases with an increase in the flow velocity v_{fluid}.

As mentioned in [McG20], the equation to model turbulence in the spherical coordinates can be given as

$$\frac{\partial \phi\left(\boldsymbol{\rho} = \left[\rho_{\text{x}}, \rho_{\text{y}}, \rho_{\text{z}}\right], t\right)}{\partial t} = \frac{D_{\text{eddy}}}{r^2} \frac{\partial}{\partial r}\left(r^2 \frac{\partial \phi(t)}{\partial r}\right), \tag{3.26}$$

where D_{eddy} is the eddy-diffusion coefficient, ϕ is the concentration dependent on time t, and r is the radius of the transmission with the relation

$$r^2 = \rho_{\text{x}}^2 + \rho_{\text{y}}^2 + \rho_{\text{z}}^2. \tag{3.27}$$

The solution for the PDE in three dimensions can be expressed as

$$\phi\left(\boldsymbol{\rho} = \left[\rho_{\text{x}}, \rho_{\text{y}}, \rho_{\text{z}}\right], t\right) = \frac{M_0}{\sqrt{\left(4\pi D_{\text{eddy}} t\right)^3}} \exp\left(-\frac{r^2}{4 D_{\text{eddy}} t}\right). \tag{3.28}$$

Thus, in the presence of obstacles, a flow component that brings about a movement of propagating particles and the propagation medium particles alike, (3.24) can be combined with (3.28) for more accurate models of the mass transportation process.

3.3 Physical Influences Inside the Testbed

As described in Chapter 2, the macroscopic air-based molecular communication testbed under consideration comprises of a pressure sprayer as the transmitter, a hollow tube of length 2.38 m as the transmission channel, and a camera-based detector as its receiver. Before determining the model that best describes the mass transportation process of the sprayed droplets of the water-based solution of the information carrier inside the testbed, the physical influences that govern the trajectory of the sprayed droplets are to be identified. In this dissertation, mainly two major influences are investigated with increasing distance d inside the testbed by performing physical measurements, one being the drag force in the

direction opposite to the sprayed droplets in horizontal ρ_{x} axis, and the other being the influence of gravity in the vertical ρ_{y} axis in a three dimensional space. It is also to be noted that the nozzle tip of the spray lance of the sprayer is adjusted in such a way, so that it produces a more directed and beamformed "jet"-like spray bursts, allowing the sprayed droplets to reach the farthest ends of the transmission channel tube by passing through the ultraviolet (UV)-A flashlight cone. This allows the signal-to-noise ratio (SNR) to reach sufficient levels, which makes it possible for the camera-based detector to conveniently detect the transmitted data symbols. The parameters chosen to conduct the measurements are presented in Table 3.1.

Table 3.1: Parameters for the investigation of the physical influences inside the testbed.

Parameter	Value
Volume of the sprayer container (L)	3
Pressure inside the sprayer container (bar)	3
Spray volume rate (L/min)	0.54
Distance to the receiver d (cm)	$\{58, 88, \ldots, 178\}$
Camera frame rate f_{cam} (FPS)	480

3.3.1 Influence of the Drag Force on Droplet Trajectory

The sprayed water droplets containing the information carrier are emitted from the sprayer at the transmitter side. These droplets travel a distance d in the positive ρ_{x} direction to reach the camera-based detector at the receiver side. Over the course of their travel, these droplets are subjected to drag force from the air medium, which slows down their velocity $v_{\rho_{\mathrm{x}}}$ (henceforth, denoted as v for simplicity), thus increasing the transmission delay. The width of the transmit pulse shape in time domain detected by the camera-based detector is also increased with increasing distance d from the transmitter, as the droplets take much longer time to pass through the entire length of the camera field of view, contributing to inter-symbol interference (ISI).

In this section, the variation of velocity v over the length of the testbed is observed by carrying out the measurements at a particular distance d from the transmitter, and taking the average over 20 trials. A light-emitting diode (LED) is placed in the field of view of the camera-based receiver, which records the incoming droplets. The switching operation of the LED is linked to that of the sprayer solenoid valve, which is controlled by the micro-controller. The LED glows as long as the solenoid valve remains open. The frames are then

counted from the time the LED turns on to the first droplets that reach the center of the camera field of view, the schematic diagram of which is shown in Figure 3.2. Similarly, an offset removal technique is employed at the transmitter side by counting the number of frames from the time instance the LED turns on to the time instance the first droplets appear just outside the sprayer nozzle. Uranine water-based dye solution is considered for conducting the trials for the particular experiment, although, the choice of the fluorescent dye is trivial in this case. Therefore, the measured velocity v_{meas} is given as

$$v_{\text{meas}} = \frac{df_{\text{cam}}}{n_{\text{rec}} - n_{\text{trans}}}, \tag{3.29}$$

where n_{trans} is the number of frames from LED turning on to the first appearance of droplets just outside the sprayer nozzle at the transmitter side, and n_{rec} is the number of frames from LED turning on to the first appearance of the droplets at the center of the field of view of camera at the receiver side.

The mean measured velocities at distances with respect to the position of the flaps in the transmission channel tube is depicted in Figure 3.3. The standard deviation of the velocities is represented by the lines enclosed within a certain range of the values. The velocity v of the droplets increases up to a certain distance to attain a maximum value right outside the spray nozzle [BD$^+$22b]. This is because the droplets are subjected to a constant force from the pressure sprayer, allowing them to accelerate over a short range of distance before attaining a maximum velocity. However, it is observed that as the distance d is further increased, the velocity v decreases linearly, as depicted in Figure 3.3 due to the drag force.

In order to obtain a direct relation between the decrease in velocity v due to drag force with increase in transmission distance d, a linear equation model is fitted (goodness of fit $R^2 = 0.9912$) to obtain v_{fit} over the mean measured velocities after attaining the maximum value, which is shown in Figure 3.3 and is given as

$$v_{\text{fit}} = 15.45\,\text{m/s} - d\;0.010\,67\,/\text{s}. \tag{3.30}$$

It is to be noted that the fitted equation model is only valid for the range of distance that has been considered for the particular analysis, but might be inaccurate for larger distances.

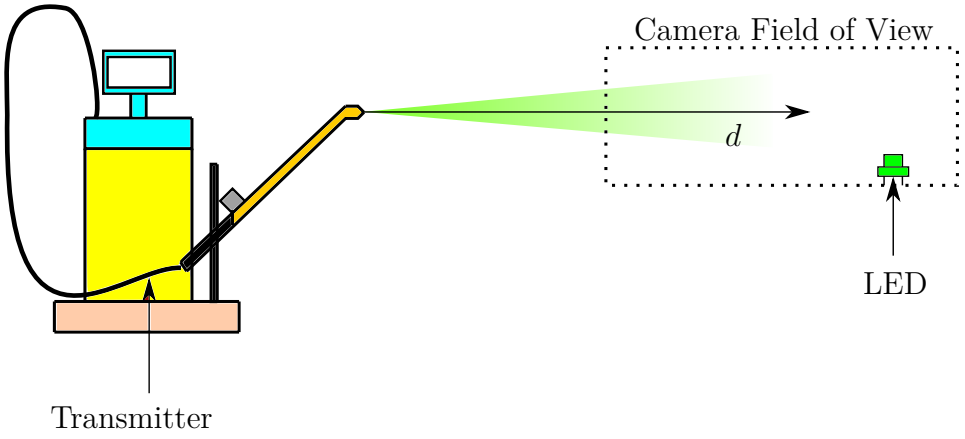


Figure 3.2: Velocity measurement of the sprayed droplets.

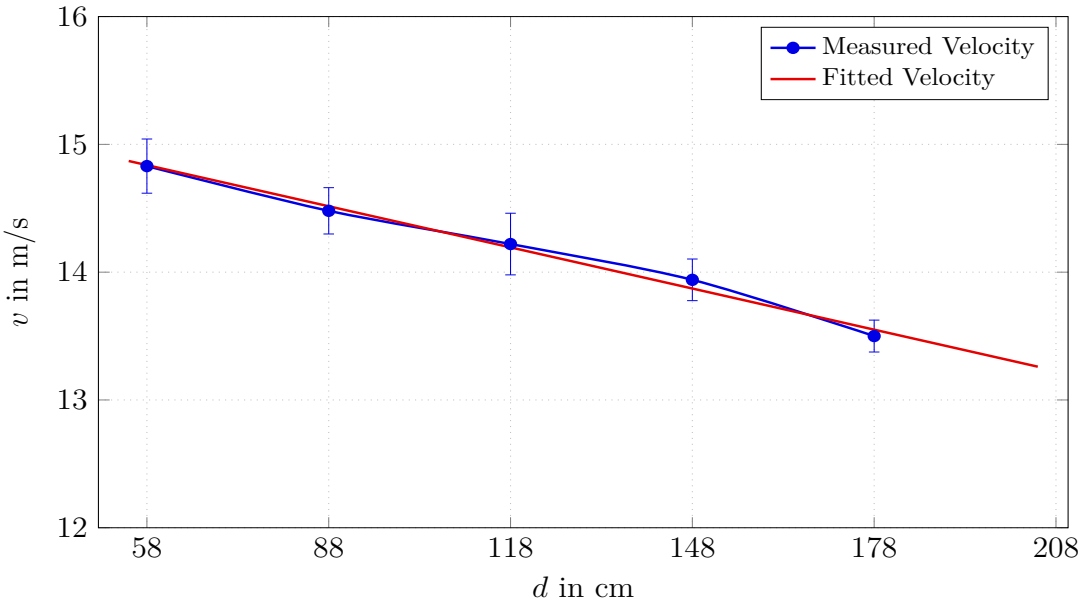


Figure 3.3: Effect of drag force on velocity of droplets v with increase in distance d.

3.3.2 Influence of the Gravitational Force on Droplet Trajectory

The sprayed droplets released as the water-based solution of the information carrier from the transmitter side undergo a projectile motion to reach the camera-based detector at the receiver side. The motion of the droplets in three dimensional space is mainly governed by their velocity components in both vertical and horizontal directions, represented by ρ_x and ρ_y axes, respectively. The increasing vertical component of the velocity v_{ρ_y} in ρ_y axis is due to the gravitational pull of the Earth, while the varying velocity component v in the horizontal ρ_x axis is due to the drag force, as described in Section 3.3.1. The divergence of the stream of sprayed droplets away from the center of the UV-A flashlight cone at the

receiver side majorly influences the pixel intensity values recorded by the camera-based receiver, which in turn impacts the SNR. The droplets pass through the light cone of the UV-A flashlight at a height that varies with distance d between the sprayer nozzle at the transmitter side, and the camera-based detector at the receiver side.

In this section, the experimental evaluation of the variation of height h' of the stream of sprayed droplets from the transmitter side over the length of the testbed tube is observed by carrying out measurements at a distance d. Figure 3.4 depicts the experimental setup, where a sheet of absorbing paper is vertically placed at distance d perpendicular to the direction of the spray nozzle of the sprayer at the transmitter side. The spray nozzle is fixed at a height $h' = 25\,\text{cm}$ from the ground level, as it mimics the exact placement position of the sprayer at the transmission channel tube of the testbed. The measurements are conducted over 20 trials in open environment, and the average height is evaluated at the observation point, depicted in Figure 3.4. A fluorescent impression is left by the sprayed droplets of water-based solution of uranine on the absorbing sheet of paper, as shown in Figure 3.5, the center of which is considered to be the height h' of the sprayed droplets from the ground. It is also noteworthy to mention that the maximum range of the sprayed droplets in the open environment is approximately 3.36 m.

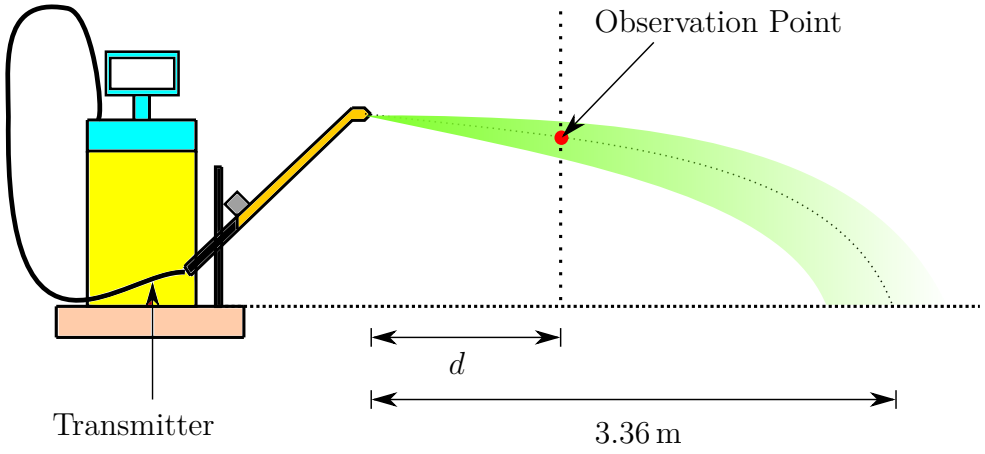


Figure 3.4: Gravitational force influencing the trajectory of the sprayed droplets.

The mean measured heights at distances with respect to the position of the flaps in the transmission channel tube is depicted in Figure 3.6. The standard deviation of the heights is represented by the lines enclosed within a certain range of the values. It is observed that the height h' of the droplets decreases in a parabolic fashion. In order to obtain a direct relation between the change in height h' due to the influence of the gravitational pull of the Earth with increase in transmission distance d, a parabolic equation model is

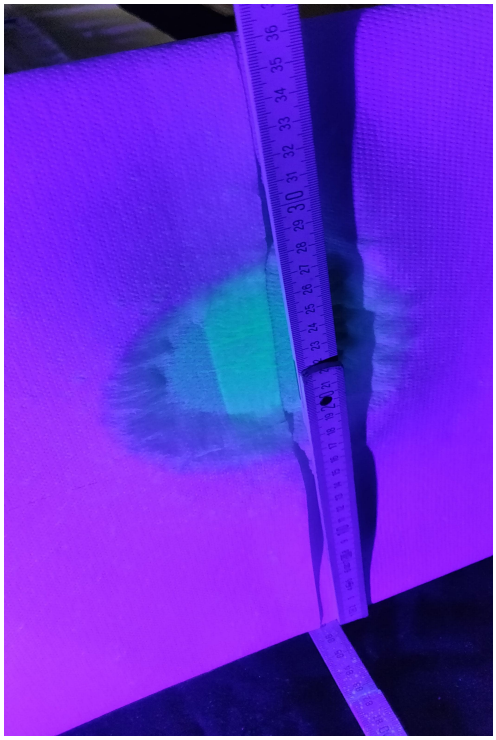


Figure 3.5: Fluorescent impression left by the sprayed droplets of the water-based solution of uranine on the absorbing sheet of paper.

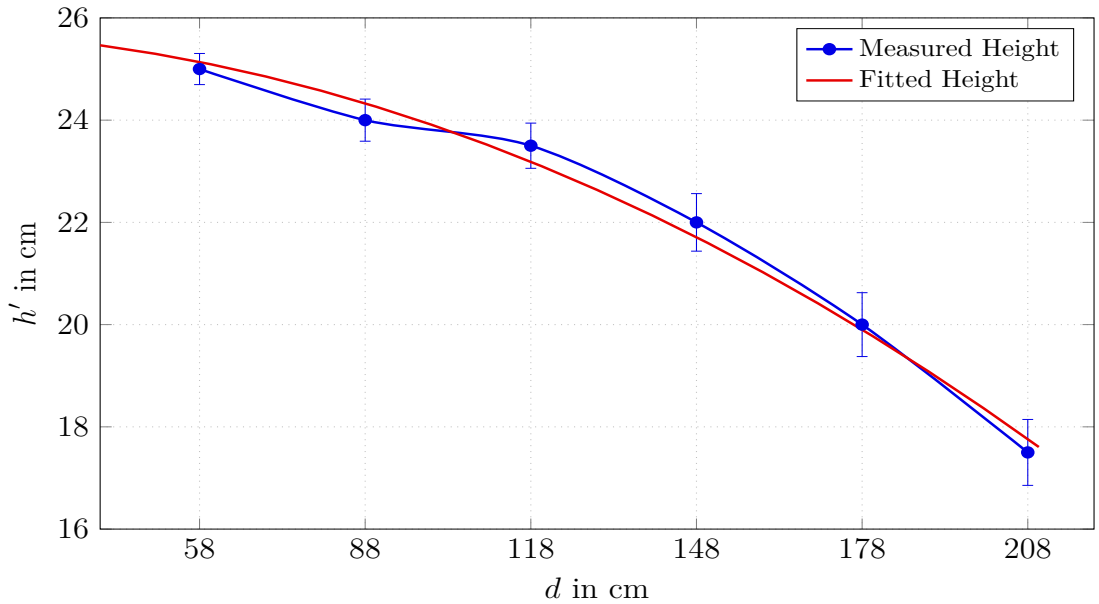


Figure 3.6: Variation of height h' of the sprayed droplets due to the influence of the gravitational force with increase in distance d.

fitted (goodness of fit $R^2 = 0.9902$) to obtain h'_{fit} over the mean measured heights, which is shown in Figure 3.6 and is given as

$$h'_{\text{fit}} = 0.258\,\text{m} - d^2\ 0.0185\,/\text{m}. \tag{3.31}$$

It is to be mentioned that the fitted equation model is only valid for the range of distance that has been considered for the particular analysis, but might be inaccurate for larger distances.

3.4 System Impulse Response Model

The presented testbed in this dissertation consists of a water-based solution of the information carrier at the transmitter side, which is filled inside a pressurized container. During a single spray burst, the pressurized air particles are released from the sprayer at a very high speed, which carry the water-based solution droplets of the information carrier along the length of the transmission channel tube. As mentioned in Section 2.1.1, the nozzle tip of the spray lance is adjusted to produce "jet"-like spray bursts of the droplets, which gives them a more directed and beamformed shape. This eliminates the chances of these droplets hitting the lateral walls of the propagation channel tube of the testbed, before reaching the camera-based receiver. Another advantage of the spray bursts being directed and beamformed is that the sprayed droplets reach the farthest end of the channel tube, making the camera-based detector to detect them with sufficient SNR. Furthermore, the testbed channel tube is also unbounded in the direction of the propagation of the sprayed droplets, which makes it safe to assume that the information-carrying droplets travel unhindered through the free space environment, without hitting any obstacles on the way. The tilt of the testbed channel tube for facilitating proper drainage of the residual solution conforms with the trajectory of the information-carrying droplets, which is in tune with the distance dependent height profile due to the gravitational pull of the Earth, as discussed and depicted in Section 3.3.2. However, the sprayed droplets under the influence of gravity diverge away from the center of the UV-A flashlight cone, thus increasing the relative distance between the camera-based detector and the sprayed droplets. Additionally, there is also a small percentage of these droplets which either get collected at the bottom of the tube or get dispersed laterally beyond the field of view of the camera. This further impacts the light received by the camera-based detector from the total number of sprayed droplets, thus affecting the measured pixel intensity values described in Section 2.3. It is

noteworthy that with a passive receiver in the form of the camera-based detector, it is a challenge to estimate the exact number of molecules of the information carrier present in the water-based solution due to the absence of a direct interaction between the detector and the information-carrying particles. However, there exists a correlation between the number of information-carrying particles with their changing concentration through space and time, and the pixel intensity value recorded by the camera-based detector. This relationship can be established through a calibration process presented in [SG+12].

Finally, in order to find the solution that best describes the mass transportation process of the sprayed water-based droplets carrying the information carrier, it is witnessed that the droplets predominantly move from the sprayer at the transmitter side towards the camera-based detector with a dominant velocity component v along the horizontal ρ_{x} axis in the three dimensional space. However, it is to be noted that in the testbed environment, the velocity v of the sprayed droplets is considered based on the time taken by the sprayed droplets to traverse the length of the section of the transmission channel tube which is visible in the field of view of the camera-based detector. It is also assumed at this point that the transmit sprayer, the channel and the receiver are ideal with no influence of the noise from various sources, as pointed out in the latter part of this chapter in Section 3.6. Moreover, the camera-based detector is assumed to act as a transparent receiver since the sprayed droplets of the information carrier are recorded while they pass through the field of view of the camera. The sprayed information-carrying droplets also do not stay in the environment after detection, as they are drained out of the tube. Any other droplets that collect at the bottom of the tube are not detected by the camera-based detector since they are not illuminated by the UV-A flashlight, lying completely outside its illumination cone. Therefore, taking into account all the aforementioned points, the concentration of the sprayed water-based solution of the information carrier inside the testbed can be modeled using the one dimensional solution of the advection-diffusion mass transportation process based on (3.22). This is popularly known as the inverse Gaussian distribution function in literature [SE+12], and can be expressed as

$$\phi\left(\rho_{\mathrm{x}}, t\right) = \frac{M_0 \rho_{\mathrm{x}}}{\sqrt{4\pi D t^3}} \exp\left(-\frac{\left(\rho_{\mathrm{x}} - vt\right)^2}{4Dt}\right). \tag{3.32}$$

In this dissertation, as mentioned in Section 2.4.2, the pixel intensity value measured by the camera-based detector at the receiver side is considered to be the system impulse response of the testbed with a strong correlation to the changing concentration of the information-carrying particles with respect to space and time. Therefore, at the transmission distance

d, $h(t) \propto \phi\,(\rho_{\mathrm{x}} = d, t)$. It is also to be noted that like any other sensor, the sensor of the camera-based detector at the receiver side also has a certain response time, i.e., how quickly it responds to, and picks up the sudden changes in the illumination brought by the sprayed fluorescent droplets in the environment of the testbed tube. Furthermore, the velocity v of the sprayed droplets decreases with increase in d due to the drag force from the air-based propagation medium, explained in Section 3.3.1. This phenomenon remains unchanged when considering the testbed hardware setup, where v of the sprayed droplets is taken into account only for the tube section that gets captured by the field of view of the camera-based detector mounted on a particular flap. Both the aforementioned phenomena combined together, cause the system impulse response to stretch in time domain, as opposed to the theoretical model described by (3.32), with a detailed explanation in [FK$^+$14]. Therefore, a correction term needs to be introduced in the system impulse response (3.32) in order to compensate for these effects. Moreover, as the pressurized air particles are released from the sprayer container at a high velocity, they carry the water-based solution of the information carrier along with them, which then breaks apart at the tip of the nozzle into multiple droplets to give a spray-like formation. This phenomenon drags along the air from the propagation medium of the testbed tube, generating shear forces responsible for producing small amount of turbulence in the process [VD$^+$18]. However, the Fick's law can still be applied to such cases with the introduction of a correction term in the diffusion coefficient D [Guh08; Far15]. From Figure 2.18 and Figure 2.19, it is also observed that the testbed measurements of the system impulse responses for sprayed uranine as well as the highlighter ink water-based solution droplets display two discernible local maxima. This is due to the fact that as the high velocity air particles are ejected from the pressurized container at the transmitter side, they push out the traces of the water-based solution of the information carriers already present inside the spray lance. This is seen as the initial burst of the sprayed droplets. The second burst, which is much wider in time domain, comes from the droplets being ejected out from the container along with the air particles. This occurrence can be modeled by the summation of two individual inverse Gaussian distribution functions, described by (3.32). Finally, the trajectory of the sprayed droplets gets affected by the gravitational pull of the Earth, increasing the distance between the camera-based detector at the receiver side and the information-carrying droplets, as described in Section 3.3.2. This affects the amplitude of the system impulse response, and is determined by the height of the trajectory given by h'. Therefore, taking into account all the aforementioned factors and (3.32) forming the basis, a more practical and realistic model for the true system

impulse response free from the effects of the noise can be described as

$$h\left(t\right) = h' \frac{\alpha_1}{\sqrt{t^3}} \exp\left(-\beta_1 \frac{\left(d - \gamma_1 t\right)^2}{t}\right) + h' \frac{\alpha_2}{\sqrt{t^3}} \exp\left(-\beta_2 \frac{\left(d - \gamma_2 t\right)^2}{t}\right), \tag{3.33}$$

where α_1 and α_2 are the correction coefficients that take into account the combined effects of the response time of the camera-based detector, the number of sprayed droplets of the information-carrying particles populating the field of view of the camera-based detector undergoing lateral dispersion, and the diffusion coefficient due to the turbulence introduced into the air-based channel by the spraying action and the change in the sprayed droplet size, when traversing through the transmission channel tube. Similarly, β_1 and β_2 are the coefficients that correct the diffusion coefficient due to the effects of turbulence from the spraying action, and the change in the sprayed droplet size as they travel through the transmission channel tube. Finally, γ_1 and γ_2 are the coefficients used to correct the change in the velocity of the sprayed droplets due to the drag force as they travel through the channel.

Estimation of the System Impulse Response Coefficients

The coefficients of the true system impulse response can be estimated for varying conditions and parameters of the testbed such as the pressure, the spray duration, and the distance between the transmitter and the receiver corresponding to a water-based solution of the information carrier. However, in this dissertation, the pressure of the sprayer at the transmitter side is fixed at 3 bar, as this is the maximum pressure at which the sprayer can operate. Also, the spray duration T_{on} is chosen to be 0.02 s for the reasons stated in Chapter 2. Therefore, in this work, only the behavior of the true system impulse response model described by (3.33) is analyzed with increasing distance d, which helps understand the extent up to which a molecular communication link can be established. In this regard, measurements are conducted in the testbed to help estimate the values of the coefficients of the model, and to study their behavior with increasing d between the transmitter and the receiver.

For estimating the coefficients of the true system impulse response for increasing distance d between the transmitter and the receiver, the list of parameters considered to perform the experiments is listed in Table 3.1. For a particular value of d, the system impulse response from the testbed is recorded by performing 100 measurement trials. As examples, both uranine as well as the highlighter ink water-based solutions at $d = 88$ cm are presented

in Figure 3.7 and Figure 3.8, respectively. Finally, the model for the true system impulse response described by (3.33) is fitted through the average of all measured trials from the testbed.

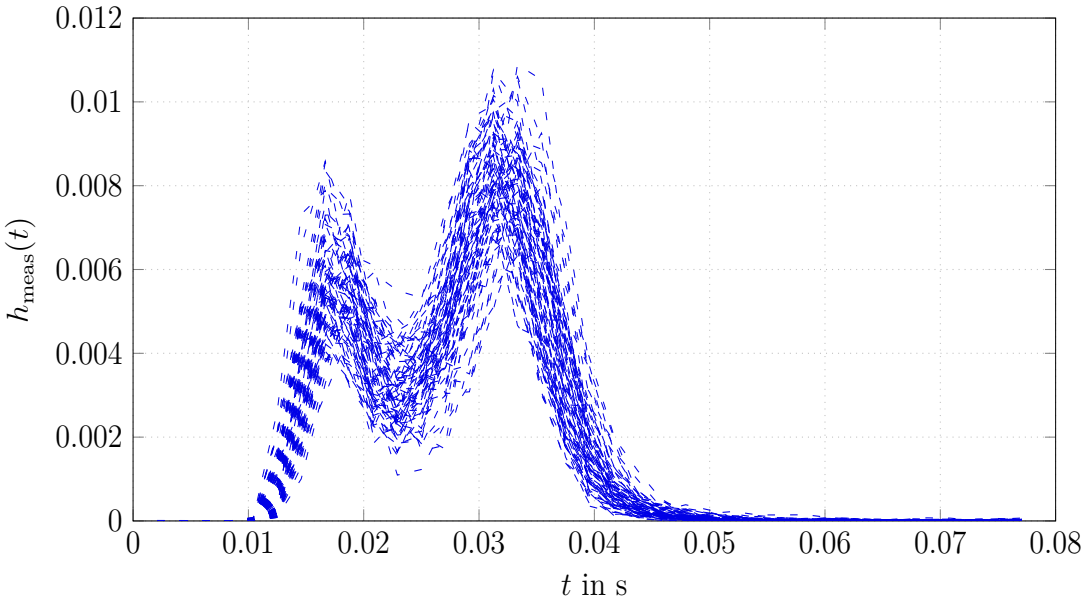


Figure 3.7: System impulse response based on the testbed measurements with 100 trials for uranine water-based solution, at distance $d = 88\,\text{cm}$.

In order to find the coefficients of the true system impulse response model, $h(t)$ is sampled at m multiples of period $T_{\text{cam}} = 1/f_{\text{cam}}$, and is represented as $h(mT_{\text{cam}}, \mathbf{p})$, where $\mathbf{p} = [\alpha_1, \beta_1, \gamma_1, \alpha_2, \beta_2, \gamma_2]^\intercal$, with α_1, β_1, γ_1, α_2, β_2, and γ_2 being the model coefficients. The fitting is performed with the help of the nonlinear least square curve-fitting technique for N samples measured by the camera-based detector, and every sample being represented by a function $f(mT_{\text{cam}})$. The coefficient estimation problem can therefore be formulated as

$$\min_{\mathbf{p}} \sum_{m=1}^{N} \left(f(mT_{\text{cam}}) - h(mT_{\text{cam}}, \mathbf{p}) \right)^2 . \tag{3.34}$$

The problem is finally solved using the Levenberg-Marquardt algorithm [Gav19], offered by the curve-fitting toolbox in MATLAB. For a particular set of coefficients, the goodness of fit between the true system impulse response model, and the average measured system impulse response from the testbed is evaluated in absolute terms by a metric named

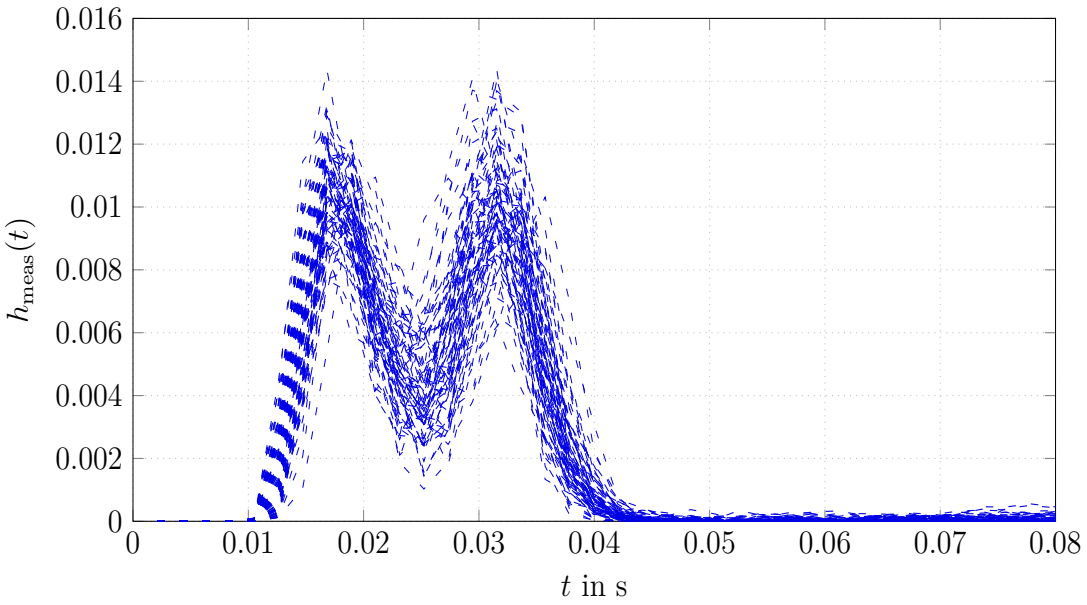


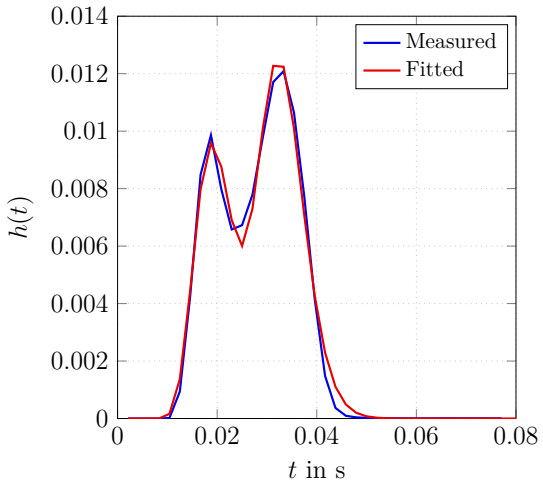
Figure 3.8: System impulse response based on the testbed measurements with 100 trials for the highlighter ink water-based solution, at distance $d = 88\,\text{cm}$.

root-mean-square error (RMSE), which is given as

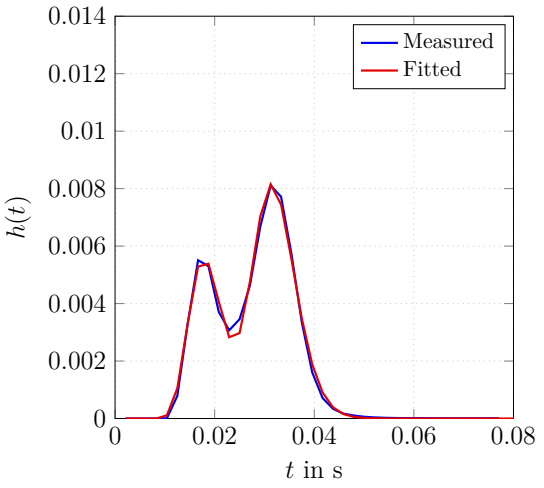
$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{m=1}^{N}\left(f(mT_{\text{cam}}) - h(mT_{\text{cam}}, \mathbf{p})\right)^2}. \tag{3.35}$$

The average measured system impulse response along with the curve-fitted system impulse response for the water-based solutions of both uranine and the highlighter ink with increasing distance d are shown in Figure 3.9 and Figure 3.10, respectively. The model coefficients obtained produce an RMSE in the order of approximately 10^{-4}, as listed in Appendix B, which corroborates the fact that the fitted model very closely resembles the real-life measurements obtained from the testbed. General observation shows that the maximum value of the system impulse response decreases, while its spread increases with increase in d, which is explained in detail later in this section. Moreover, the two discernible local maxima of the system impulse response appear to merge together with increase in d. This is due to the fact that for larger distances, the clear distinction between the consecutive spray bursts from the spray lance dissipate slowly as the individual droplets present in the spray burst travel with different velocities constantly undergoing a lateral dispersion.

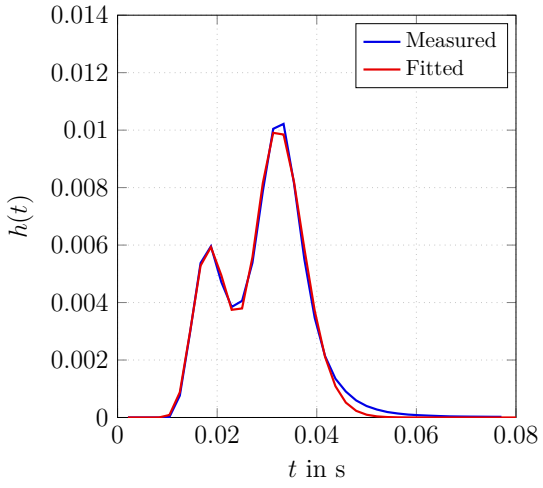
The coefficients of the system impulse response themselves display a variation with increasing distance d between the transmitter and the receiver. The coefficients α_1 and α_2



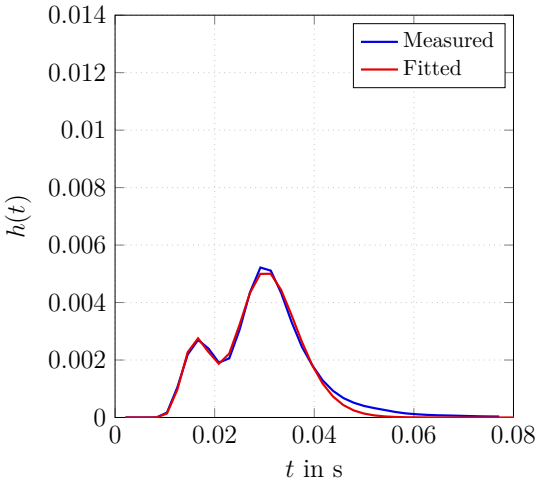
(a) $d = 58\,\text{cm}$



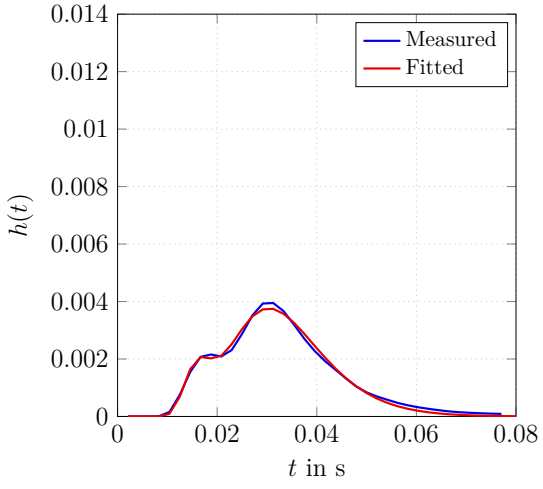
(b) $d = 88\,\text{cm}$



(c) $d = 118\,\text{cm}$

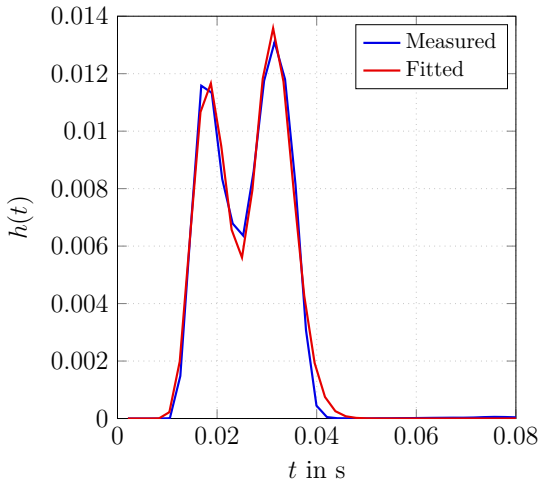


(d) $d = 148\,\text{cm}$

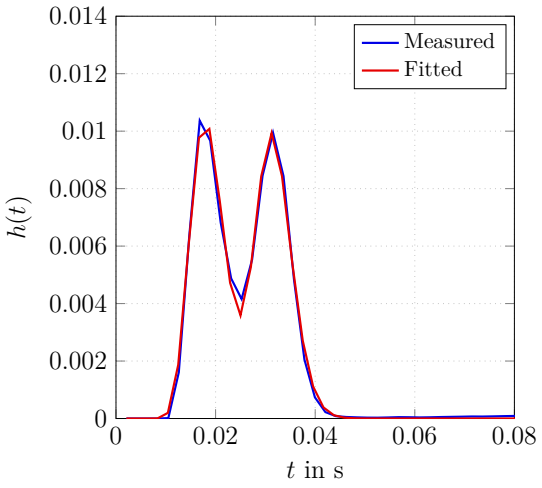


(e) $d = 178\,\text{cm}$

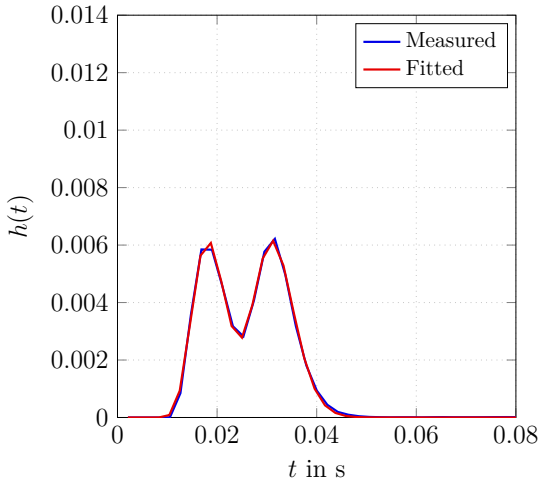
Figure 3.9: Average measured system impulse response from the testbed with the curve-fitted system impulse response for uranine water-based solution for increasing distance d.



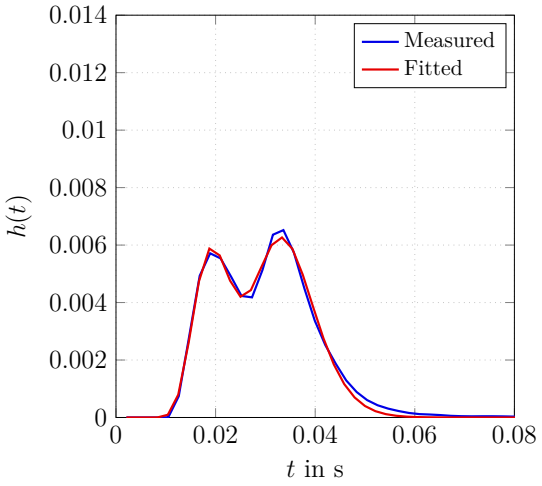
(a) $d = 58\,\text{cm}$



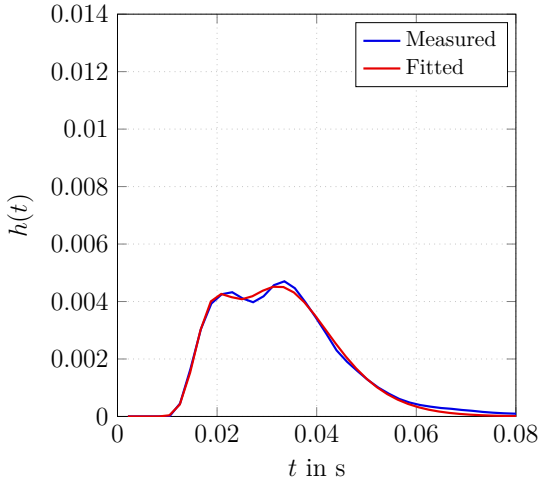
(b) $d = 88\,\text{cm}$



(c) $d = 118\,\text{cm}$

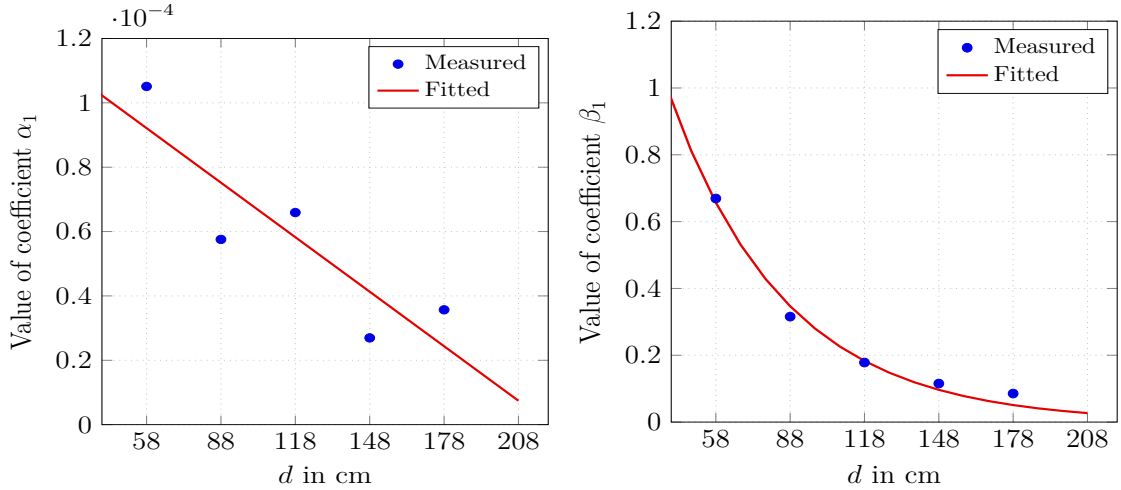


(d) $d = 148\,\text{cm}$



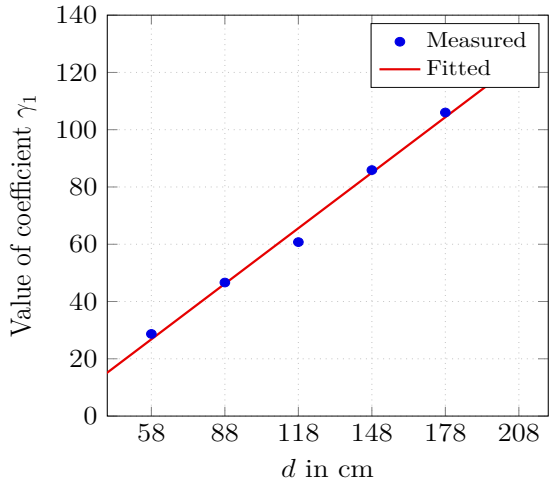
(e) $d = 178\,\text{cm}$

Figure 3.10: Average measured system impulse response from the testbed with the curve-fitted system impulse response for the highlighter ink water-based solution for increasing distance d.

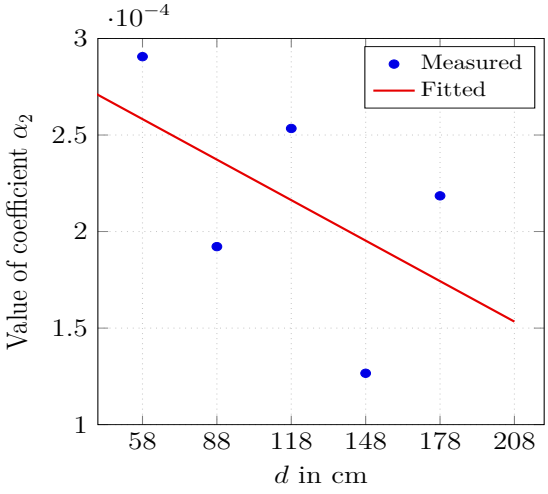


(a) Variation of coefficient α_1.

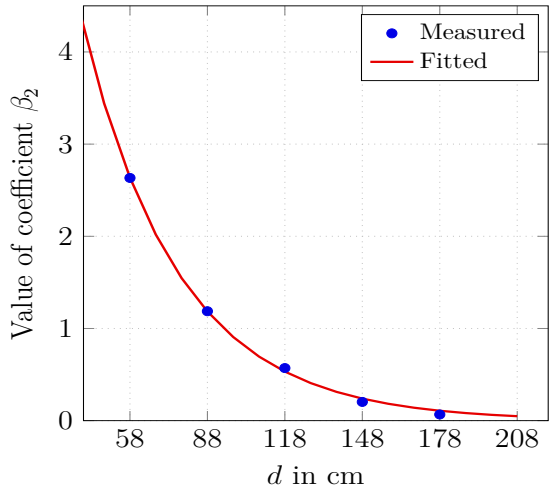
(b) Variation of coefficient β_1.



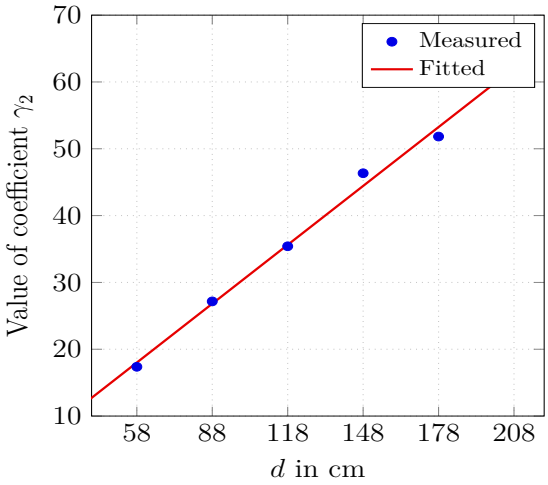
(c) Variation of coefficient γ_1.



(d) Variation of coefficient α_2.

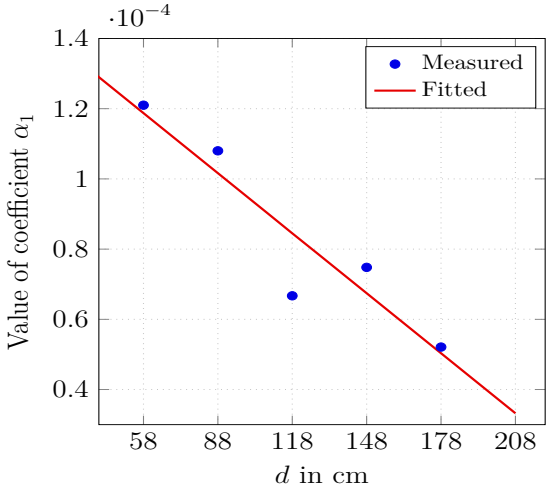


(e) Variation of coefficient β_2.

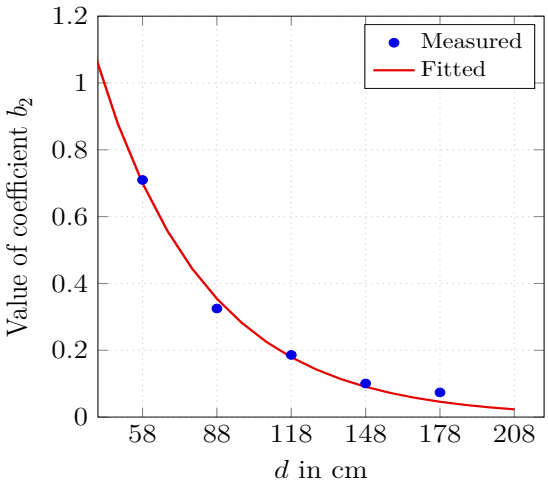


(f) Variation of coefficient γ_2.

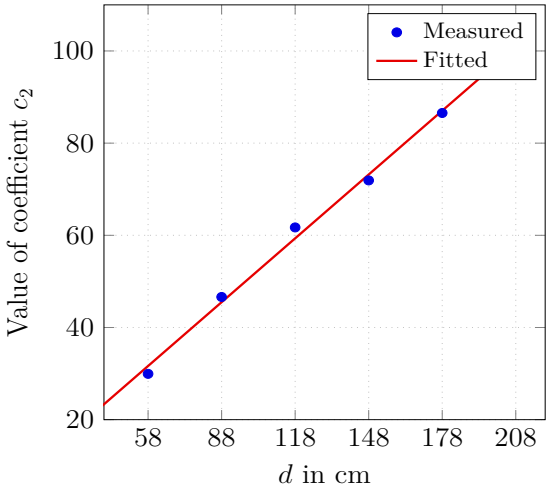
Figure 3.11: Variation of coefficients of true system impulse response model $h(t)$ with distance d for uranine water-based solution.



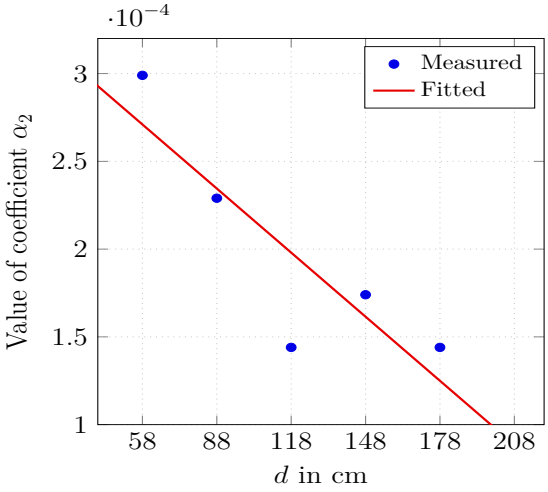
(a) Variation of coefficient α_1.



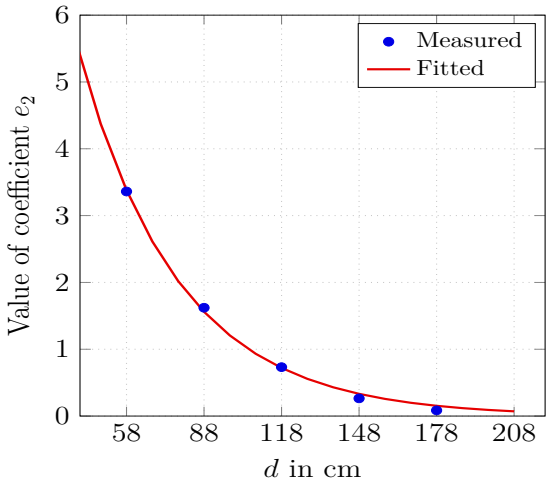
(b) Variation of coefficient β_1.



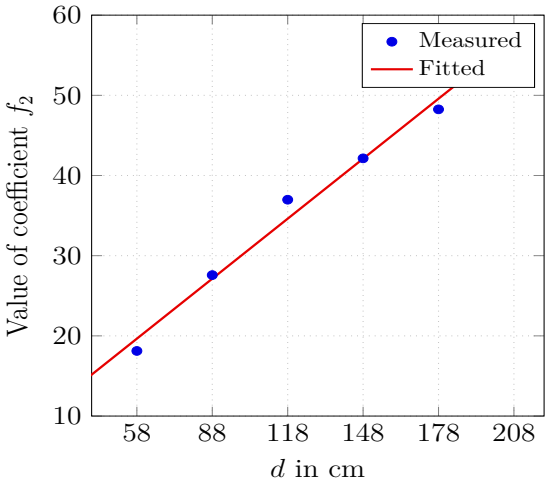
(c) Variation of coefficient γ_1.



(d) Variation of coefficient α_2.



(e) Variation of coefficient β_2.



(f) Variation of coefficient γ_2.

Figure 3.12: Variation of coefficients of true system impulse response model $h(t)$ with distance d for the highlighter ink water-based solution.

for both water-based solutions of uranine and the highlighter ink decrease almost linearly with increase in d, as shown in both Figure 3.11 and Figure 3.12. This is because the coefficients take into account the number of sprayed droplets of the information-carrying particles, which constantly undergo a lateral dispersion. As a result, a fraction of these particles settle down at the bottom of the tube and/or move out of the field of view of the camera with increasing d. Moreover, the diffusion coefficient changes because of the turbulence introduced into the air-based channel due to the spraying action of the information-carrying droplets. Additionally, the coalesced droplets break apart into smaller droplets, as they travel from their release point at the tip of the spraying nozzle through the transmission channel tube. The change in radius of these droplets also affects the diffusion coefficient. Linear equation models for uranine water-based solution are thus fitted in order to get a direct relation between the change in the values of the coefficients α_1 and α_2 with increase in d, and are obtained as

$$\alpha_1 = 1.249 \times 10^{-4}\,\mathrm{s}^{1/2} - d\,5.648 \times 10^{-5}\,\mathrm{s}^{1/2}/\mathrm{m}, \tag{3.36}$$

and

$$\alpha_2 = 2.988 \times 10^{-4}\,\mathrm{s}^{1/2} - d\,6.993 \times 10^{-5}\,\mathrm{s}^{1/2}/\mathrm{m}. \tag{3.37}$$

Similarly, the linear equation models for the highlighter ink water-based solution for coefficients α_1 and α_2 with an increase in d are obtained as

$$\alpha_1 = 1.518 \times 10^{-4}\,\mathrm{s}^{1/2} - d\,5.7 \times 10^{-5}\,\mathrm{s}^{1/2}/\mathrm{m}, \tag{3.38}$$

and

$$\alpha_2 = 3.416 \times 10^{-4}\,\mathrm{s}^{1/2} - d\,1.217 \times 10^{-4}\,\mathrm{s}^{1/2}/\mathrm{m}. \tag{3.39}$$

Coefficients β_1 and β_2 decrease exponentially with increase in distance d for both water-based solutions of uranine and the highlighter ink, as depicted in Figure 3.11 and Figure 3.12, respectively. These coefficients take into account the diffusion coefficient that gets affected by the turbulence caused by the spraying action of the water-based solution of the information-carrying droplets. Moreover, the coalesced droplets break apart into smaller droplets from their point of release at the tip of the spraying nozzle, as they travel through the transmission channel tube. The constantly changing radius of these droplets ultimately introduces a change in the diffusion coefficient with increasing d, which can be represented in the case of uranine water-based solution by fitted exponential equation

models for coefficients β_1 and β_2 as

$$\beta_1 = 2.261 \exp(-d\ 2.131)\,\mathrm{s/m^2}, \tag{3.40}$$

and

$$\beta_2 = 12.38 \exp(-d\ 2.667)\,\mathrm{s/m^2}. \tag{3.41}$$

In similar lines, the exponential models that best describe the change in the values of coefficients β_1 and β_2 with an increase in d for the highlighter ink water-based solution are obtained as

$$\beta_1 = 2.613 \exp(-d\ 2.271)\,\mathrm{s/m^2}, \tag{3.42}$$

and

$$\beta_2 = 15.07 \exp(-d\ 2.577)\,\mathrm{s/m^2}. \tag{3.43}$$

Lastly, coefficients γ_1 and γ_2 increase linearly with increase in distance d for both water-based solutions of uranine and the highlighter ink, as shown in Figure 3.11 and Figure 3.12, respectively. The coefficients take into account the change in the velocity of the sprayed droplets due to the drag force as they travel through the channel. In the testbed environment, however, the velocity v of these droplets is only considered based on the time taken by them to traverse the length of the channel that is visible in the field of view of the camera-based detector, and not from the point of release of the droplets from the spray nozzle, as described in Section 3.3.1. Thus, the increase in the values of the coefficients γ_1 and γ_2 is mostly to offset the effect of the increase in d in the true system impulse response model. Therefore, the change in the coefficient with d for uranine water-based solution is represented by fitted linear equation models for coefficients γ_1 and γ_2, and are obtained as

$$\gamma_1 = -10.68\,\mathrm{m/s} + d\ 64.64\,\mathrm{/s}, \tag{3.44}$$

and

$$\gamma_2 = 0.9693\,\mathrm{m/s} + d\ 29.37\,\mathrm{/s}. \tag{3.45}$$

Similarly, the linear models describing the change in the values of coefficients γ_1 and γ_2 with increasing d for the highlighter ink water-based solution are obtained as

$$\gamma_1 = 4.871\,\mathrm{m/s} + d\ 46.16\,\mathrm{/s}, \tag{3.46}$$

and

$$\gamma_2 = 5.185\,\mathrm{m/s} + d\ 24.94\,/\mathrm{s}. \tag{3.47}$$

3.5 Test of Linearity

In signals and systems theory, linear systems are classified as those, in which the output produced due to a linear combination of the inputs is the same as that of the linear combination of the individual responses to the same inputs [PM14]. Mathematically speaking, for a system $\mathfrak{T}[x]$ defined over time t, the relation between the individual inputs $x_1(t)$ and $x_2(t)$ scaled by two arbitrary constants a_1 and a_2 respectively, and the corresponding outputs $y_1(t)$ and $y_2(t)$ can be defined as

$$y_1(t) = a_1\mathfrak{T}[x_1(t)], \tag{3.48}$$

and

$$y_2(t) = a_2\mathfrak{T}[x_2(t)]. \tag{3.49}$$

The condition of linearity is thus satisfied when

$$\mathfrak{T}[a_1x_1(t) + a_2x_2(t)] = a_1\mathfrak{T}[x_1(t)] + a_2\mathfrak{T}[x_2(t)] = y_1(t) + y_2(t). \tag{3.50}$$

Molecular communication systems relying on fluid flow mechanism for information transfer processes, often risk defying the linearity property that is presented in (3.50), which causes dramatic variations between the inputs and the outputs of a system, requiring a more complex detection process of the transmitted information. Some of the possible reasons leading to this nonlinear behavior are the presence of strong turbulence or wind effects at the transmitter side or in the channel, which could potentially change the direction of flow of the information-carrying particles. Other possible reasons include changes in the properties of the information carrier that affect its concentration, and thus, its ability to be detected at the receiver side. The imperfect receiver design is another cause of nonlinear behavior, making the detection process inconsistent over extended periods of time. However, as long as these deviations are small, well within certain "permissible limit", and can be minimized by averaging out over multiple trial runs of transmitted sequences, the system can still be considered to be close to a linear system.

Concerning a molecular communication system employing a single transmitter, the nonlinear behavior is identified for the first time in the alcohol-based molecular communication

testbed, presented in [FG+13; FK+14; KF+14]. The main reason behind the display of such behavior is due to the imperfection of the gas sensor at the receiver side reaching its saturation level, which leads to an inconsistent detection of long sequences of 1's in the data transmission sequence.

In order to test the system linearity of the introduced molecular communication testbed in this dissertation, the experiment is performed using a single transmitter with the assumption that all the inconsistencies in the end-to-end system leading to the deviation from the true input value can be reduced by averaging over multiple trial runs of a transmission sequence. The experiment to test the linearity of the system is performed separately for the water-based solutions of uranine as well as the highlighter ink, in line with the experiments performed in [OK+18; WH+20]. In order to do so, three transmit sequences are considered, as shown in Table 3.2. Sequences S_1 and S_2 comprise of 20 randomly generated symbols without having an overlapping symbol "1", as the system is incapable of producing symbol "2" due to the availability of only two states of a single sprayer at the transmitter side, i.e., "on" (symbol "1") and "off" (symbol "0").

Table 3.2: Transmit sequences for the test of linearity of the fluorescent dye-based molecular communication system.

Transmit Sequence	Binary Representation
S_1	1 0 0 1 0 1 0 0 1 1 1 0 1 0 1 0 1 0 0 1
S_2	0 1 0 0 1 0 1 0 0 0 0 1 0 1 0 0 0 1 0 0
S_3	1 1 0 1 1 1 1 0 1 1 1 1 1 1 1 0 1 1 0 1

Additionally, sequence S_3 is considered, which is the superposition of the sequences S_1 and S_2. The sequences S_1 and S_2 are transmitted separately, and the pixel intensities generated by the individual transmission of these sequences are recorded by the camera-based detector and superpositioned on one another. Finally, sequence S_3 is transmitted, and the pixel intensity is separately recorded by the camera-based detector. The detection at the receiver side is performed at a distance of $d = 88\,\text{cm}$. To minimize the effect of inconsistencies in the system, the sequences are run over 20 trials, and the mean pixel intensity values are calculated as shown in Figure 3.13 and Figure 3.14 for water-based solutions of uranine and the highlighter ink, respectively. It is observed that the superpositioned received pixel intensity values from the transmission of sequences S_1 and S_2 are almost the same as the received pixel intensity value from the transmission of sequence S_3. Therefore, it is concluded that the molecular communication system making use of only one transmitter is nearly linear and is independent of the choice of the information carrier used to transmit

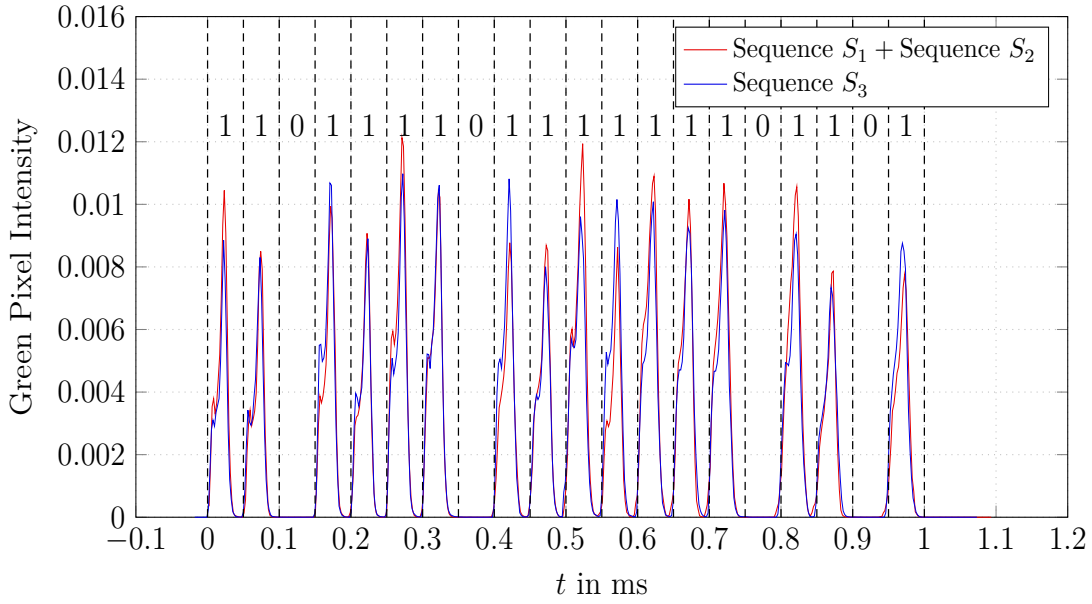


Figure 3.13: Recorded green pixel intensity values of the water-based solution of uranine for the transmit sequences, at distance $d = 88\,\text{cm}$.

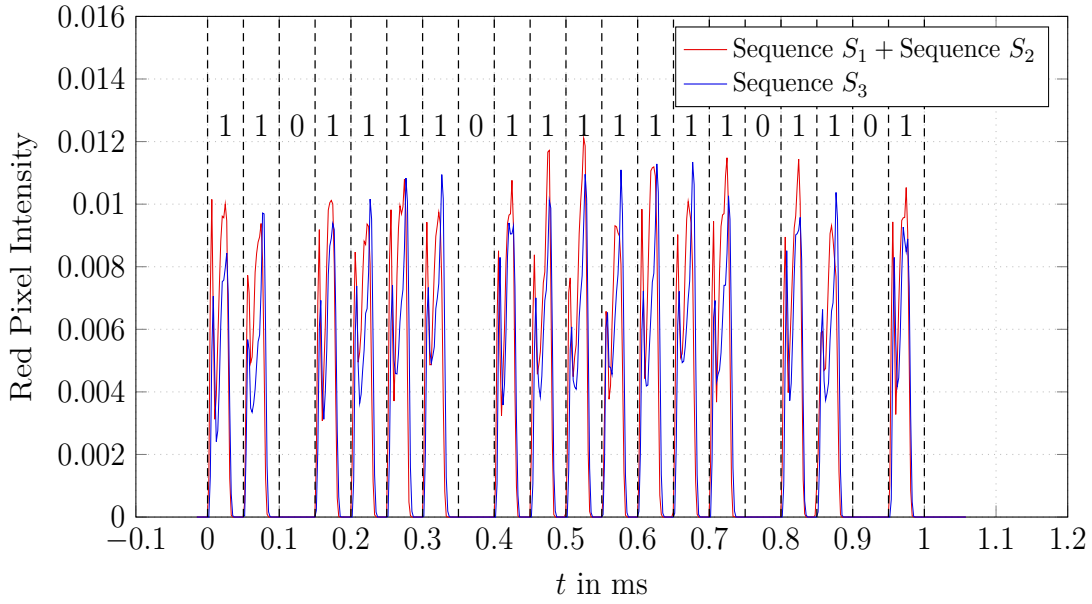


Figure 3.14: Recorded red pixel intensity values of the water-based solution of the highlighter ink for the transmit sequences, at distance $d = 88\,\text{cm}$.

the information.

It is important to mention that those molecular communication systems employing multiple transmitters for modulation purposes by exploiting the color of the information carrier and/or the spatial domain as additional degrees of freedom (described in Chapter 4 in detail), the system can display nonlinear behavior due to the effect of shadowing. Shadowing occurs when the sprayed droplets of one color or belonging to one spatial region may get hidden by the droplets of other color or belonging to the other spatial region, due to an increase in lateral dispersion with distance between the transmitter and the receiver.

3.6 Noise Process

As mentioned in Section 2.4.2 and in Section 3.4, the system impulse response is evaluated by transmitting a short spray burst through the channel, resembling a Dirac delta function, and measuring the response using the camera-based detector at the receiver side. In an ideal scenario, the system impulse response comprises of the influences of a perfect transmitter, a perfect channel, and a perfect receiver on the Dirac delta function, which is described by the model in (3.33). However, the response of the system over multiple trials shows randomness, as seen from multiple measurement trials in Figure 3.7 and Figure 3.8, and the linearity test of the system in Figure 3.13 and Figure 3.14. This phenomenon can be attributed to the following reasons: Imperfect sprayer at the transmitter side producing inconsistent spray bursts, spray nozzle adjustments, small turbulence created by the initial interaction of released high velocity droplets with air, the air drag from the transmission channel medium, the inconsistency in the amount of droplets passing through the UV-A light cone, and the imperfection of the camera-based detector at the receiver side. Some of the sources of randomness are also in line with the noise sources of the molecular communication systems described in [KF$^+$15; GA20], and therefore in this work too, the randomness present in the system is treated as the background noise, which is described by a statistical noise model. The proposed noise model together with the true system impulse response model in (3.33) can be used to describe the complete analytical model for the end-to-end system under investigation. The analytical model could then be used to predict the results much faster and more efficiently by conducting extensive simulations as compared to the time-consuming testbed measurements.

The proposed noise model is assumed to be additive in nature, and in order to establish the noise process, the noise samples in m multiples of period $T_{\text{cam}} = 1/f_{\text{cam}}$ are estimated. Therefore, the estimated noise sample $\hat{n}_m$ from the measured system impulse response

trials $h_{\text{meas}}(t)$, and the true system impulse response $h(t)$ sampled at $t = mT_{\text{cam}}$ can be described as

$$\hat{n}_m = h_{\text{meas}}(mT_{\text{cam}}) - h(mT_{\text{cam}}). \tag{3.51}$$

The estimated continuous-time noise signals of five testbed trials at a transmission distance $d = 88\,\text{cm}$ is shown for uranine and the highlighter ink water-based solutions in Figure 3.15 and Figure 3.16, respectively.

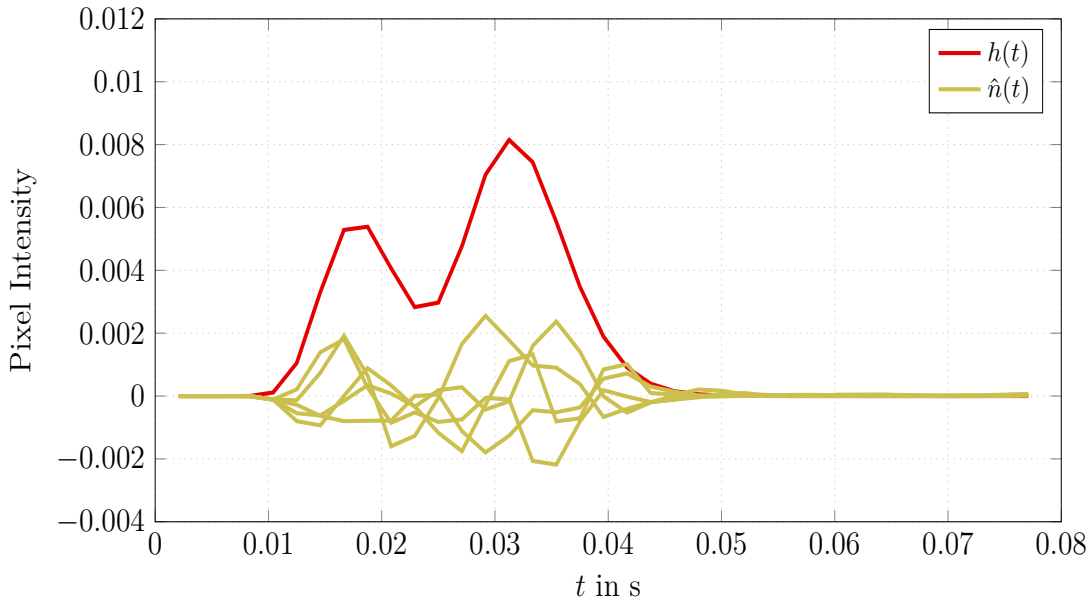


Figure 3.15: True system impulse response model for uranine water-based solution and estimated noise signal of five testbed measurement trials, at a distance $d = 88\,\text{cm}$.

It is also interesting to see how the standard deviation of the noise signals for 100 measurement trials sampled at $t = mT_{\text{cam}}$ vary with time t. If the total number of trials is represented by N_{trial}, the standard deviation of the noise samples at time $t = mT_{\text{cam}}$ is given by

$$\sigma(\hat{n}_{m,N_{\text{trial}}}) = \frac{1}{N_{\text{trial}}} \sum_{i_{\text{trial}}=1}^{N_{\text{trial}}} \hat{n}_{m,i_{\text{trial}}} - \mu_{\hat{n}_{m,N_{\text{trial}}}}, \tag{3.52}$$

where i_{trial} is a trial from N_{trial} total trials, $\hat{n}_{m,i_{\text{trial}}}$ is an estimated noise sample at $t = mT_{\text{cam}}$ for trial i_{trial}, and $\mu_{\hat{n}_{m,N_{\text{trial}}}}$ is the mean value of the estimated noise samples at time $t = mT_{\text{cam}}$ for N_{trial} total number of trials, which is given as

$$\mu_{\hat{n}_{m,N_{\text{trial}}}} = \frac{1}{N_{\text{trial}}} \sum_{i_{\text{trial}}=1}^{N_{\text{trial}}} \hat{n}_{m,i_{\text{trial}}}. \tag{3.53}$$

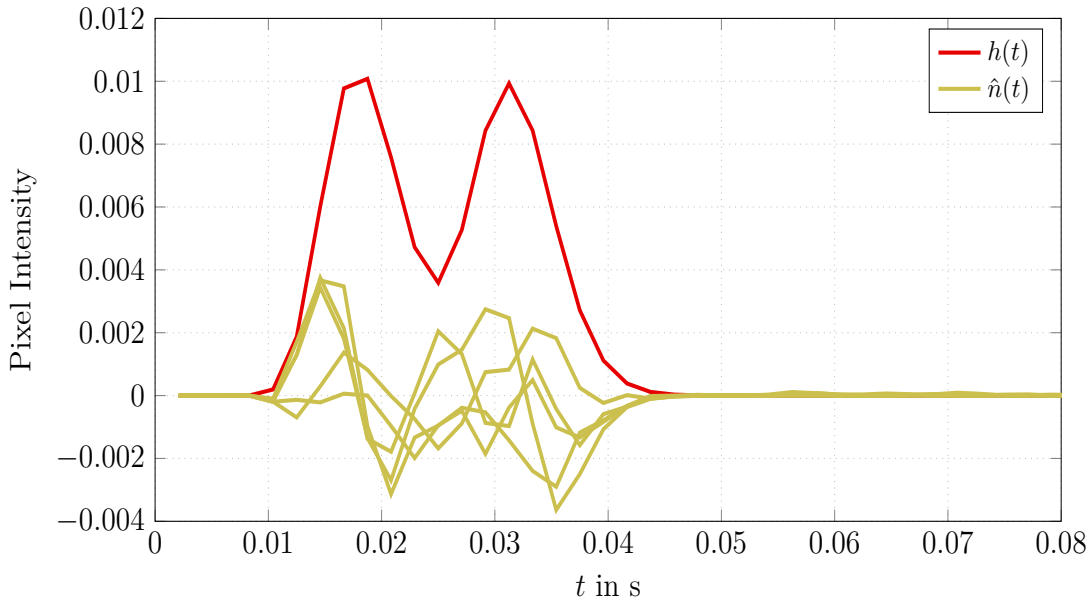


Figure 3.16: True system impulse response model for the highlighter ink water-based solution and estimated noise signal of five testbed measurement trials, at a distance $d = 88\,\text{cm}$.

It is observed that a direct correlation exists between the true system impulse response $h(t)$ sampled at time $t = mT_{\text{cam}}$ scaled by a factor κ, and the standard deviation of the noise samples $\sigma(\hat{n}_{m,N_{\text{trial}}})$, the plots of which for both uranine as well as the highlighter ink water-based solutions at a distance $d = 88\,\text{cm}$ are shown in Figure 3.17 and Figure 3.18, respectively. For the quantification of the correlation that exists between the aforementioned quantities, a Pearson correlation coefficient [Pes15] is calculated, which is defined as

$$\rho_{\text{corr}} = \frac{\text{cov}\left(h(mT_{\text{cam}}), \sigma\left(\hat{n}_{m,N_{\text{trial}}}\right)\right)}{\sigma_{h(mT_{\text{cam}})}\sigma_{\sigma\left(\hat{n}_{m,N_{\text{trial}}}\right)}}, \tag{3.54}$$

where $\text{cov}(\cdot)$ is the covariance between the true system impulse response $h(t)$ sampled at $t = mT_{\text{cam}}$ and the standard deviation of the estimated noise samples $\sigma\left(\hat{n}_{m,N_{\text{trial}}}\right)$ at $t = mT_{\text{cam}}$, $\sigma_{h(mT_{\text{cam}})}$ is the standard deviation of the true system impulse response $h(t)$ sampled at $t = mT_{\text{cam}}$, and $\sigma_{\sigma\left(\hat{n}_{m,N_{\text{trial}}}\right)}$ is the standard deviation of the standard deviation of the estimated noise samples $\sigma\left(\hat{n}_{m,N_{\text{trial}}}\right)$ at $t = mT_{\text{cam}}$. The values of the scaling factor κ and the Pearson correlation coefficient ρ_{corr} with distance d for the water-based solutions of uranine and the highlighter ink are presented in Table 3.3 and Table 3.4, respectively. The value of κ is proportional to the maximum value of the system impulse response $h_{\max}(t)$ achieved, which decreases almost linearly with increase in distance d between the

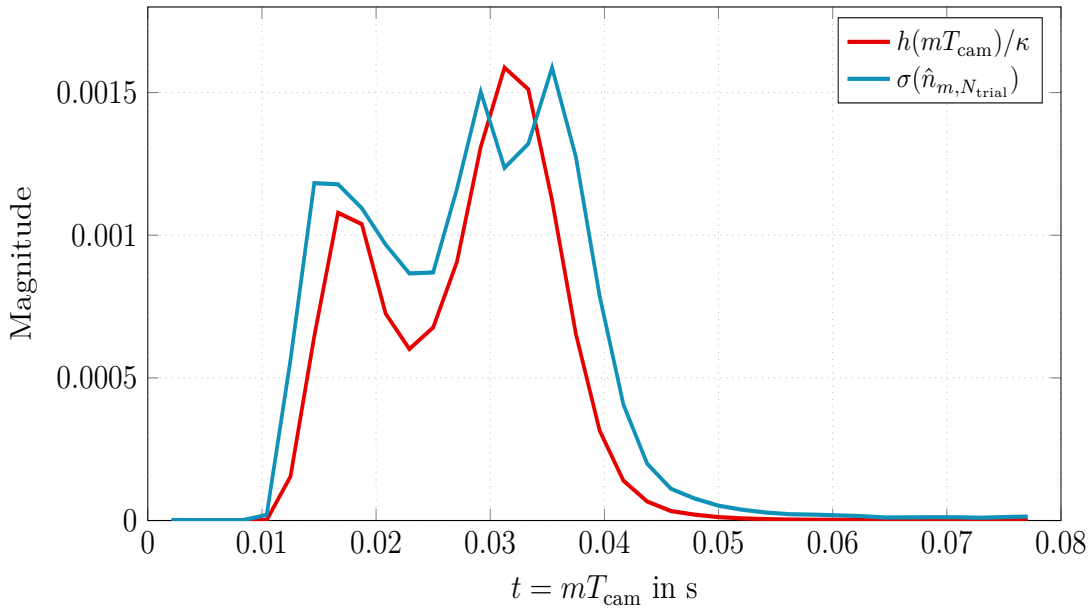


Figure 3.17: Scaled true system impulse response and the standard deviation of the estimated noise samples at $t = mT_{\text{cam}}$ for the uranine water-based solution, at distance $d = 88\,\text{cm}$.

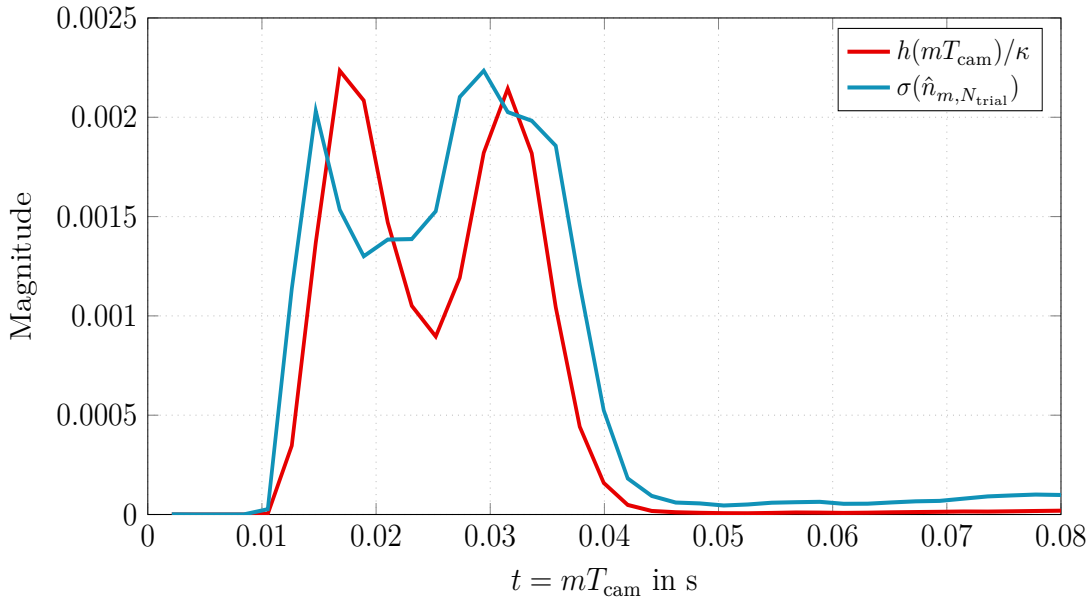


Figure 3.18: Scaled true system impulse response and the standard deviation of the estimated noise samples at $t = mT_{\text{cam}}$ for the highlighter ink water-based solution, at distance $d = 88\,\text{cm}$.

transmitter and the receiver. Owing to a very large value of ρ_{corr} corresponding to d, it is concluded that the noise power is amplitude-dependent, which is also in tune with the diffusion noise in classical diffusion-based molecular communication models, often encountered in the literature [DK+17]. Additionally, the randomness present in the system under investigation is due to the superposition of many uncorrelated processes mentioned earlier in this section. As the central limit theorem states that the superposition of uncorrelated random variables in real-life scenarios roughly leads to a Gaussian distribution [Fis11], the noise process variance can therefore be modeled as a continuous-time amplitude-dependent white Gaussian noise process, and can be expressed as

$$\sigma_n^2(t) = h(t)\,(1 - h(t))\,. \tag{3.55}$$

Table 3.3: The values of the scaling factor κ and the Pearson correlation coefficient ρ_{corr} with distance d for uranine water-based solution.

Distance d (cm)	Scaling Factor κ	Correlation Factor ρ_{corr}
58	7.2601	0.8382
88	5.1068	0.9410
118	4.6945	0.9502
148	3.2124	0.9628
178	1.9398	0.9421

Table 3.4: The values of the scaling factor κ and the Pearson correlation coefficient ρ_{corr} with distance d for the highlighter ink water-based solution.

Distance d (cm)	Scaling Factor κ	Correlation Factor ρ_{corr}
58	5.9209	0.8790
88	4.6385	0.9089
118	3.2324	0.9473
148	3.4119	0.9749
178	2.7211	0.9804

The statistical dependency of the estimated noise sample $\hat{n}_m$ to a previous noise sample $\hat{n}_{m-\tau_{\text{lag}}}$ with a relative time lag of τ_{lag} is described by the auto-correlation function $\text{ACF}\left(\hat{n}_m, \tau_{\text{lag}}\right)$. As examples, this function is illustrated in Figure 3.19 and Figure 3.20 for uranine as well as the highlighter ink water-based solutions for a transmission distance of $d = 88\,\text{cm}$, which show the previous noise sample estimates that have a strong statistical influence on the current noise sample estimate, which makes the noise process itself to be

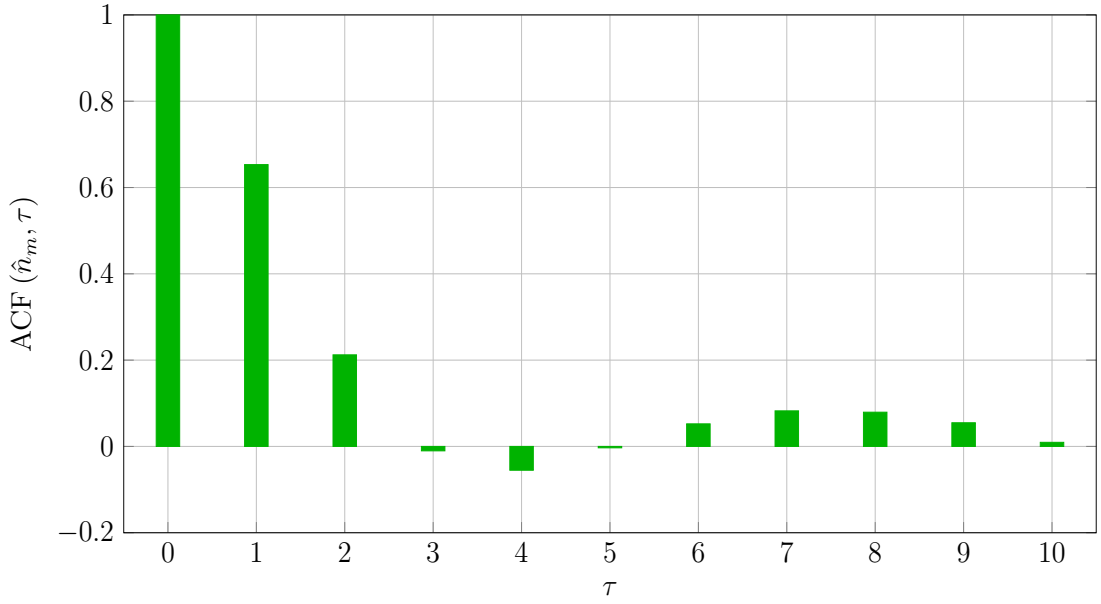


Figure 3.19: Correlation weighting factors averaged over 100 measured trials for uranine water-based solution, at a distance $d = 88\,\text{cm}$.

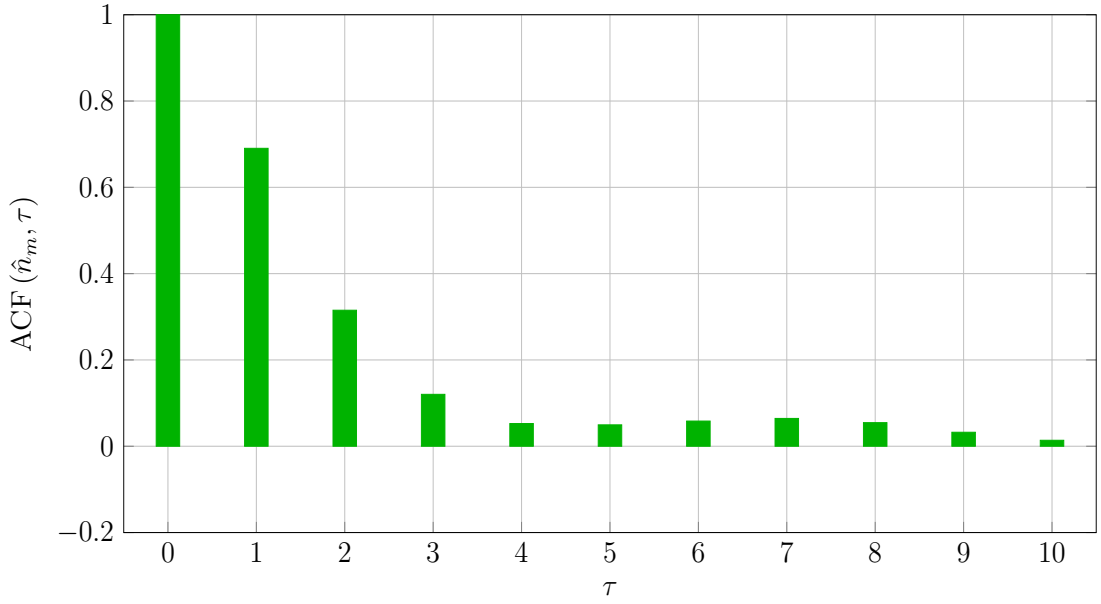


Figure 3.20: Correlation weighting factors averaged over 100 measured trials for the highlighter ink water-based solution, at a distance $d = 88\,\text{cm}$.

correlated in nature. This is in contrast to a white noise process, where the statistical dependency of a noise sample on any of the previous noise samples is absent. It is to be noted that the strong correlation between the current sample and the past noise samples exists due to the appearance of the same droplets multiple times in a sequence of recorded frames of the camera, as they travel along the field of view of the camera. Once the droplets exit the field of view of the camera, there is no longer any contribution by these droplets to the pixel intensity values, recorded in subsequent frames. The observation is also true for the other transmission distances under investigation. As a consequence of these observations specific to the testbed design and the distance between the transmitter and the receiver, the statistical dependence of the noise samples can be emulated using a simple moving average (MA) model. The order w_{MA} of the MA model is determined by taking into account all the weighting factors with a value of at least 0.1 that exist due to the correlation between the current noise sample estimate, and the previous noise sample estimates. As examples, the correlation weighting factors averaged over 100 trials for both uranine as well as the highlighter ink water-based solutions at a distance $d = 88\,\mathrm{cm}$ are presented in Figure 3.19 and Figure 3.20, respectively. Thus, the MA model can be described as

$$n_m = \frac{1}{\sqrt{\left(1 + \sum_{i=1}^{w_{\mathrm{MA}}} \psi_i^2\right)}} \left(\sum_{i=1}^{w_{\mathrm{MA}}} \psi_i \varepsilon_{m-i} + \varepsilon_m \right), \tag{3.56}$$

where ψ_i are the weighting factors that determine the influence of the previous noise samples, and $\varepsilon_{m-i} \sim \mathcal{N}(0,1)$, $\varepsilon_m \sim \mathcal{N}(0,1)$, are zero-mean white Gaussian noise processes with variance of 1. The parameters of the MA model are determined from the measured testbed signals in the sense of a maximum likelihood estimation with the help of the `arima()` function available in MATLAB. Therefore, the final amplitude-dependent and correlated noise process can be given as

$$n(mT_{\mathrm{cam}}) = \sigma_n(mT_{\mathrm{cam}}) n_m. \tag{3.57}$$

Finally, the superposition of the true system impulse response $h(t)$ sampled at mT_{cam}, and the noise model sampled at mT_{cam} for a distance $d = 88\,\mathrm{cm}$ are combined together in order to generate 100 impulses for both uranine as well as the highlighter ink water-based solutions, as depicted in Figure 3.21 and Figure 3.22, respectively. The impulses generated from the models show a clear conformity with the measured impulses from the testbed, as depicted in Figure 3.7 and Figure 3.8.

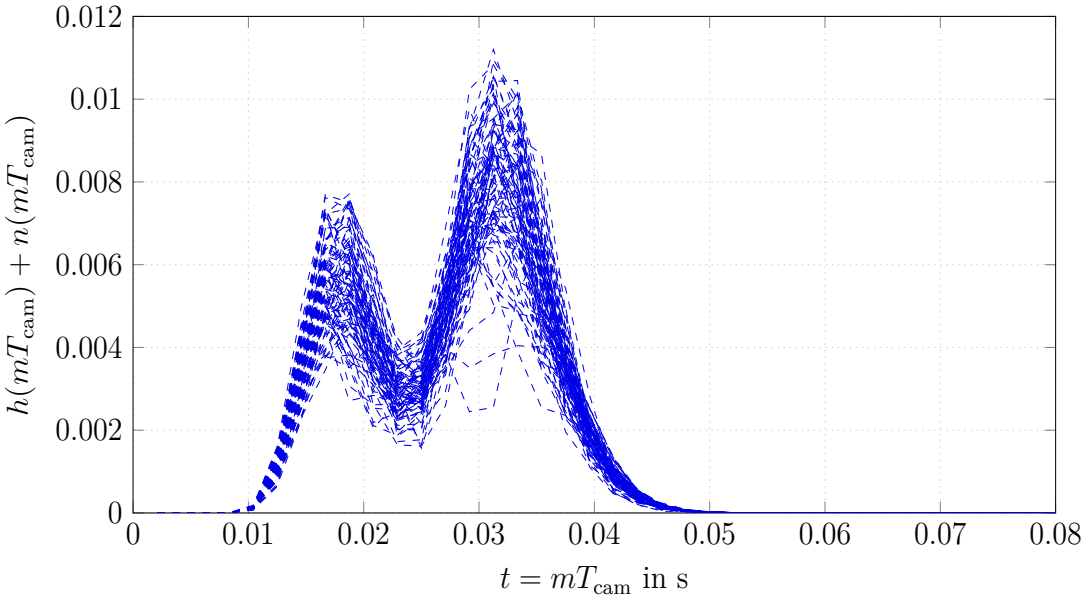


Figure 3.21: System impulse response from the testbed with 100 measured trials for uranine water-based solution, at distance $d = 88\,\text{cm}$.

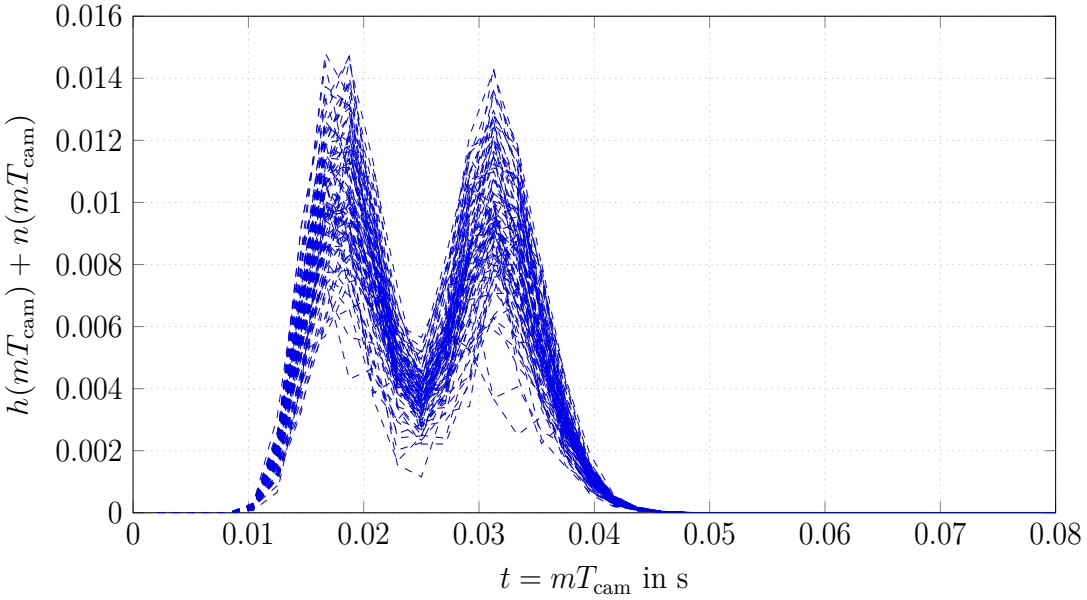


Figure 3.22: System impulse response from the testbed with 100 measured trials for the highlighter ink water-based solution, at distance $d = 88\,\text{cm}$.

3.7 Analytical Model

Finally, the transmission and reception of data through the testbed can be well described by a continuous-time end-to-end system model comprising of the input signal into the system $x(t)$, the system impulse response $h(t)$, the additive noise $n(t)$, which ultimately produces the output signal $y(t)$ coming out of the system, as depicted in Figure 3.23.

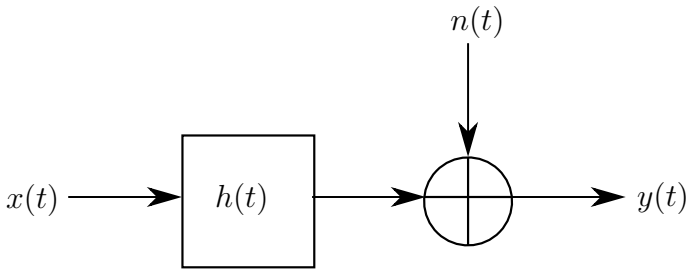


Figure 3.23: Continuous-time end-to-end system model of the testbed.

Mathematically speaking, the system model can be translated into the analytical model, which is a computational representation used to describe and analyze the behavior of a system, process, or a particular phenomenon. The analytical model can also be used as an alternative to performing real-life measurements, which often prove to be very tedious and time-consuming. The analytical model introduces more flexibility by providing an option to perform the various analyses, especially that of bit error rate (BER) by running larger Monte Carlo runs for higher accuracy, which otherwise is a constraint in a physical testbed scenario. However, it must also be noted that analytical models do have their limitations. They often rely on simplified assumptions that may not accurately capture all the real-life complexities. For example, when a system becomes nonlinear, it might be difficult or even impossible to derive closed-form solutions. Keeping all the constraints in mind, the complete end-to-end system of the testbed can be described conforming with Figure 3.23 as

$$y(t) = x(t) * h(t) + n(t), \tag{3.58}$$

where $*$ represents the convolution operation.

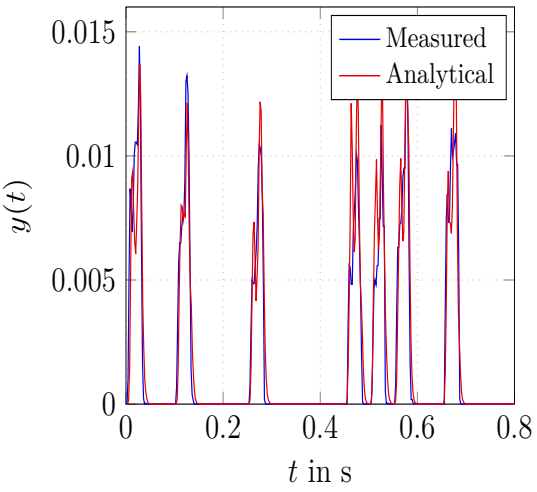
In order to test the validity of the analytical model and a compatibility with the measurements from the testbed, data transmission of the text message "ICT" is performed over the testbed with uranine as well as the highlighter ink water-based solutions, as depicted in Figure 3.24 and Figure 3.25, respectively. This is performed for increasing distance d between the transmitter and the receiver. A source-coded binary data sequence "010010001110100" that corresponds to the American standard code for information interchange is transmitted

over the testbed channel. An initialization symbol "1" precedes the transmission sequence in order to facilitate proper synchronization at the receiver side, indicating the arrival of the message sequence. Symbol "0" is represented by the inactivity of the sprayer over a symbol duration $T = 0.05\,\mathrm{s}$, while symbol "1" is represented by the spraying action at the transmitter side over $T = 0.05\,\mathrm{s}$, with a spray duration of $T_{\mathrm{on}} = 0.02\,\mathrm{s}$. The same text message is also transmitted using the analytical model, as shown in Figure 3.24 and Figure 3.25. It is observed that the model very closely emulates the real-life measurements obtained from the testbed, indicating the fact that the model is suitable for the purpose of BER analysis, when comparing the performances of various modulation schemes, equalization and detection strategies, as well as channel coding schemes by increasing the Monte Carlo trials in a simulation environment. The obtained results closely resemble the testbed measurement results with a higher precision, requiring only a fraction of time and effort as compared to conducting the measurement trials.

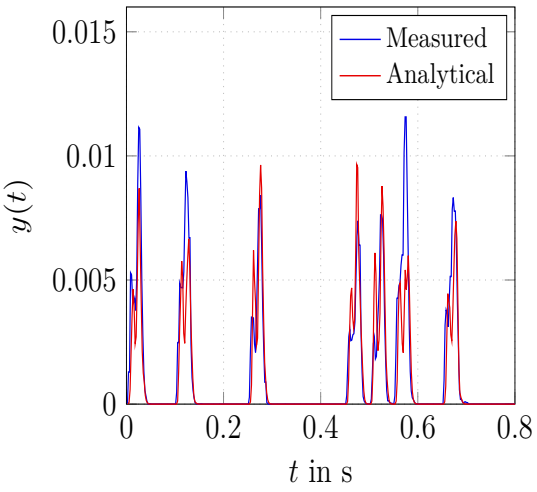
3.8 Summary

This chapter discusses the theory behind the mass transportation process of the information-carrying particles in a transmission channel of a molecular communication system. Micro-scale molecular communication system primarily witnesses the mass transportation of the information-carrying particles in the form of diffusion, advection, turbulent flow, convection, and dispersion. However, when it comes to the macro-scale air-based molecular communication system, diffusion, advection, and turbulence are the predominant ways in which the information-carrying particles get transmitted from the transmitter to the receiver. The testbed presented in this work is modeled to transport the information-carrying water-based solution droplets of uranine and the highlighter ink using the inverse Gaussian distribution function, based on the principle of advection-diffusion. Moreover, by conducting experiments, the dominant physical effects such as the drag force and the gravitational force are seen to influence the propagation of the water-based solution droplets of the information carriers through the channel. These effects are incorporated into the inverse Gaussian distribution function defining the noise-free system impulse response of the molecular communication testbed under investigation. Furthermore, the estimation of the system impulse response coefficients and their variation with increase in distance d between the transmitter and the receiver are carried out in this work based on the testbed measurements and curve-fitting technique using the Levenberg-Marquardt algorithm.

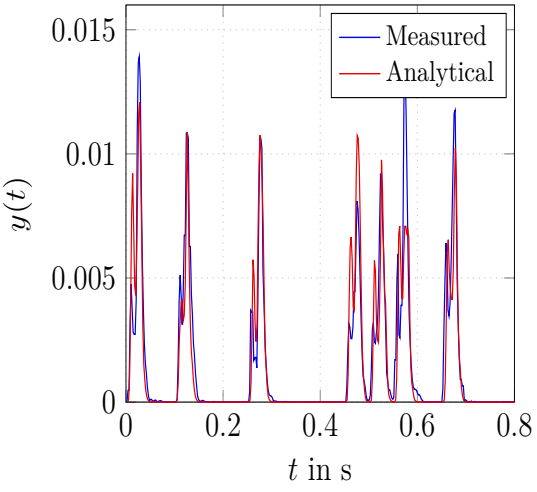
The system linearity of the introduced molecular communication testbed in this disserta-



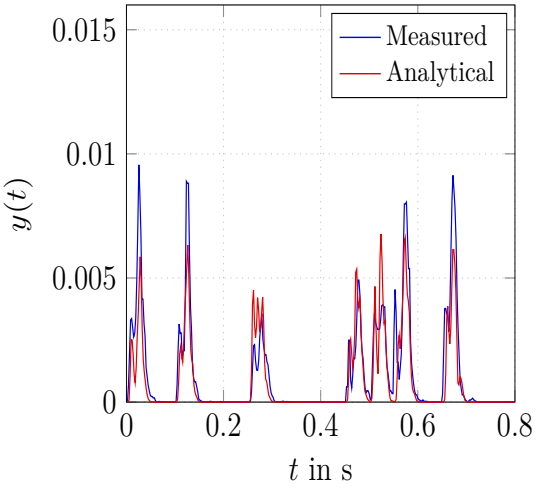
(a) $d = 58\,\text{cm}$



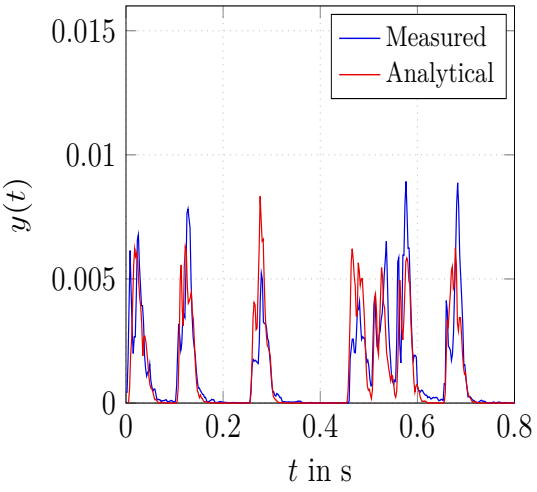
(b) $d = 88\,\text{cm}$



(c) $d = 118\,\text{cm}$

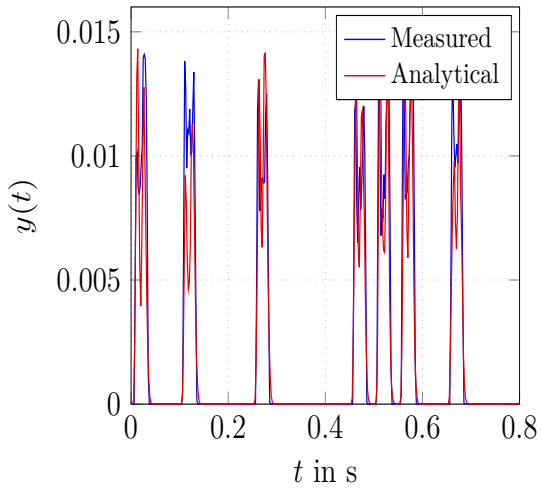


(d) $d = 148\,\text{cm}$

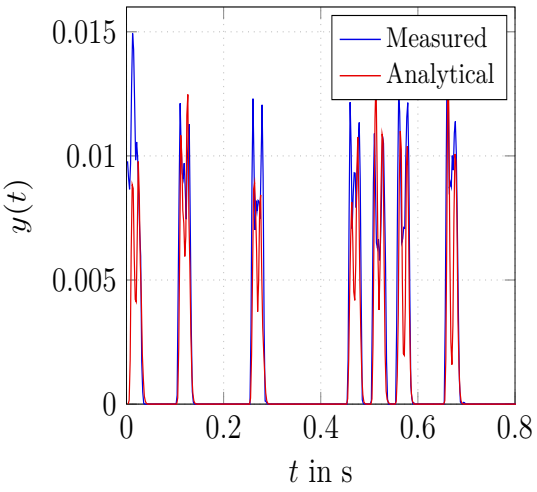


(e) $d = 178\,\text{cm}$

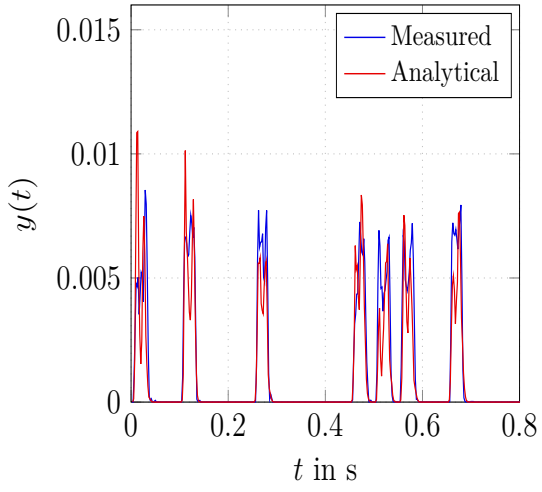
Figure 3.24: Comparison of measured $y(t)$ and analytical $y(t)$ for transmitted message "ICT" for uranine water-based solution with increasing distance d.



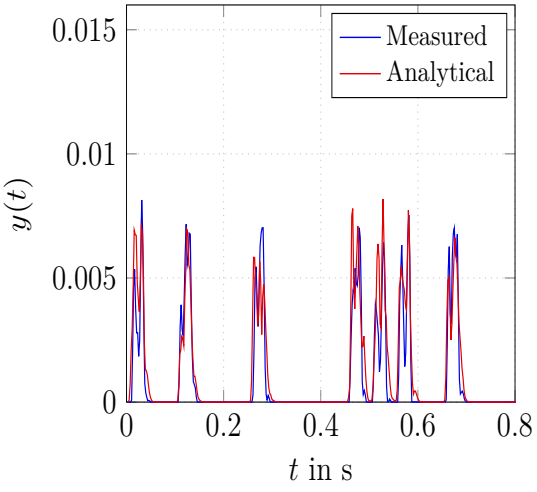
(a) $d = 58\,\text{cm}$



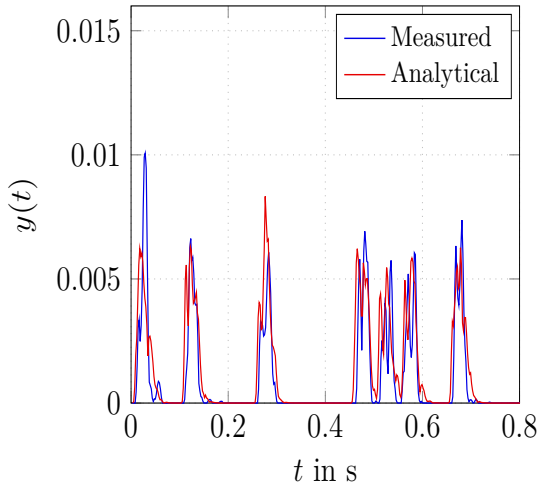
(b) $d = 88\,\text{cm}$



(c) $d = 118\,\text{cm}$



(d) $d = 148\,\text{cm}$



(e) $d = 178\,\text{cm}$

Figure 3.25: Comparison of measured $y(t)$ and analytical $y(t)$ for transmitted message "ICT" for the highlighter ink water-based solution with increasing distance d.

tion is tested by performing experiments with a single transmitter containing a water-based solution of uranine and the highlighter ink. With the assumption that the inconsistencies in the end-to-end system can be reduced through averaging of the transmission sequences over multiple trials, the linearity test is performed by the transmission and superposition of independent data sequences. It is observed that the presented molecular communication system displays almost a linear behavior. The slight deviations and randomness are treated as noise.

The noise in the end-to-end system of the presented macroscopic molecular communication testbed originates due to the imperfect sprayer design at the transmitter side, spray nozzle adjustments, small turbulence due to the interaction of released high velocity droplets with air, the air drag from the transmission channel medium, inconsistency in the amount of sprayed droplets passing through the UV-A light cone, and the imperfection of the camera-based detector at the receiver side. The statistical analysis of the noise is performed to find out its characteristics, which is observed to be amplitude-dependent in nature. This noise process is combined with the system impulse response to define the complete analytical model of the testbed.

The mathematical translation of the end-to-end system model comprising of the input, the true system impulse response, the noise model, and the output is given by the analytical model. The analytical model is helpful in the sense that it can be used as an alternative to performing real-life measurements, which are often time-consuming and require a lot of effort. The validation of the analytical model is done by generating a text message with the model, and comparing it with the same text message transmitted using the water-based solutions of uranine as well as the highlighter ink through the testbed channel. This is done for increasing distance d between the transmitter and the receiver. The observation shows the ability of the analytical model to very closely emulate the real-life measurements performed on the testbed, indicating its suitability in performing BER analyses with respect to the implementation of different modulation schemes, equalization and detection methods as well as the channel coding schemes through Monte Carlo trials in a simulation environment.

4

Modulation and Detection

This chapter deals with the data detection process at the receiver side as well as implementational aspects of various transmission algorithms in the form of modulation schemes incorporating one or more degrees of freedom. This is achieved typically by exploiting the information carrier concentration, time, the color of the information carrier, and the presence of sprayed water-based droplets of the information carrier in the spatial domain to represent the information. The error performance of the information carriers in the form of the water-based solutions of uranine and the highlighter ink are compared for a simple binary modulation scheme incorporating concentration as a degree of freedom, for increasing distance between the transmitter and the receiver. Similarly, the detection algorithms and equalization techniques are also realized, and their performances are compared for a particular information carrier. Finally, the error performance analysis is extended to various higher order modulation schemes for increasing distance between the transmitter and the receiver.

4.1 Receiver-Side Data Acquisition

Before discussing the modulation techniques and detection algorithms, it is imperative to discuss the process by which the transmitted data gets acquired at the receiver side. As explained in Section 2.3, the detection process considers the fact that the sprayed water-based solution droplets of the information carriers under the influence of the ultraviolet (UV)-A light source can be distinguished based on the fluorescence produced by them over a certain wavelength range in the visible light region. This phenomenon is reflected predominantly on the green and red pixel intensity values of the images that are captured

by the high-speed camera-based detector at the receiver side for uranine and the highlighter ink water-based solutions, respectively. However, prior to their interpretation for detection of transmitted data symbols, the images are post-processed by taking the first captured image as a reference frame, and then subtracting it from all the subsequent images that are captured. This step ensures that there is no offset in the recorded pixel intensity values, which may arise due to any other luminous components or residual dye solution inside the tube contributing to the noise process. In the next step, the green and the red pixel intensities for the respective water-based solutions of the information carriers recorded by each pixel in a given frame are summed separately, which are then interpreted as the information signal, ready to undergo the detection process. As example, the transmission of the text message "ICT" in Section 3.7 is considered for illustration purpose. At distance $d = 88\,\text{cm}$, the information signal representation with green pixel intensity value for uranine water-based solution is presented in Figure 4.1. The binary equivalent of the message sequence along with the initialization symbol "1" is depicted, where symbol "0" and symbol "1" are represented by the inactivity and activity of the sprayer over symbol duration $T = 0.05\,\text{s}$ at the transmitter side, respectively.

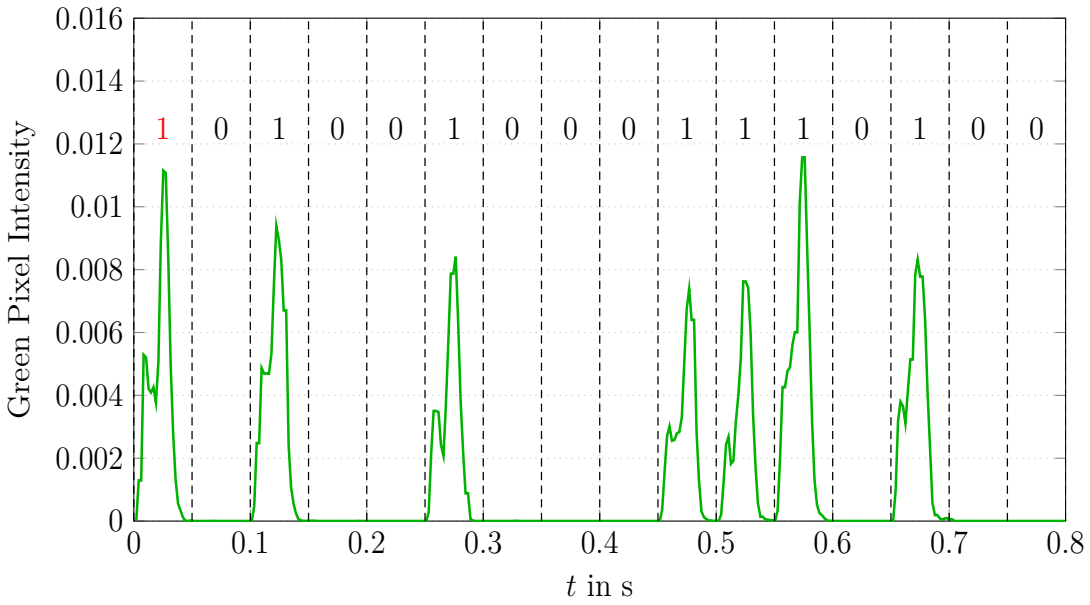


Figure 4.1: The information signal representation of the transmitted message "ICT" with green pixel intensity value for uranine water-based solution, at distance $d = 88\,\text{cm}$.

In the sense of the acquisition of the transmitted data, the energy under the received green and/or red pixel intensity pulses are accumulated over a symbol duration T, and can be mathematically obtained through integration. Hence, the ℓth end-to-end system

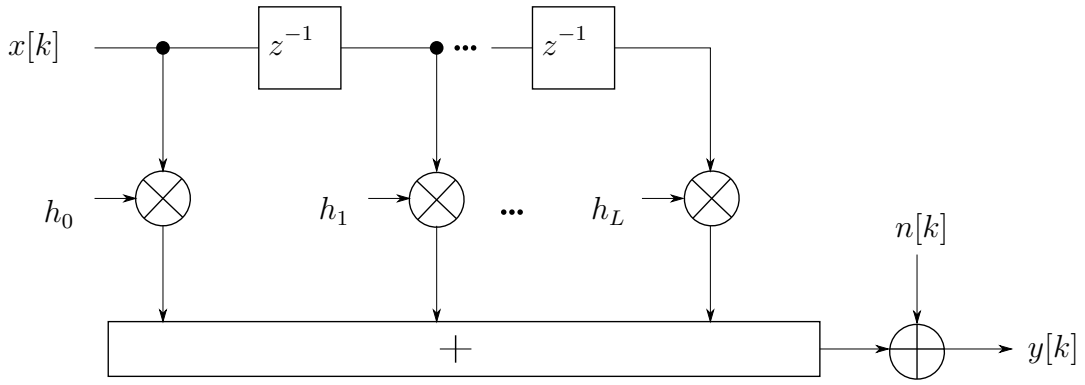


Figure 4.2: Block diagram of the equivalent discrete-time end-to-end system model.

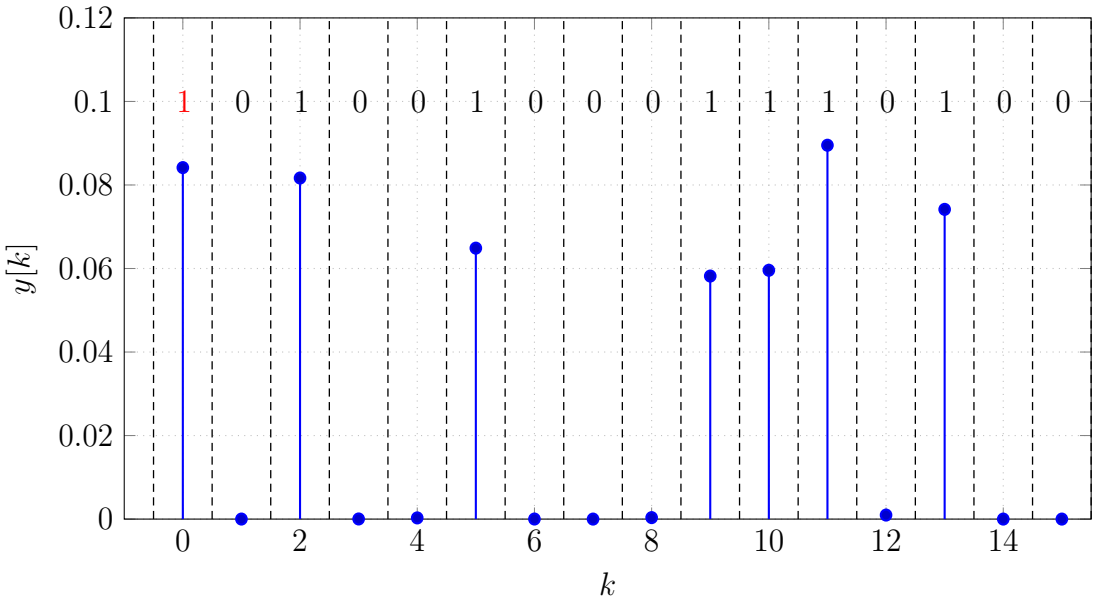


Figure 4.3: The pixel energy samples $y[k]$ for the binary equivalent of the transmitted message sequence “ICT”.

coefficient h_ℓ, $0 \leq \ell \leq L$ is calculated as

$$h_\ell = \int_{\ell T}^{(\ell+1)T} h(t)\, dt, \tag{4.1}$$

where L is the effective end-to-end system memory length. Therefore, the relationship between the transmitted data symbol $x[k]$ and the received pixel energy value $y[k]$ at kth transmission interval can be exactly described in discrete-time domain by a representation of convolution operation with addition of noise term $n[k]$. This is termed as the equivalent discrete-time end-to-end system model, as shown in Figure 4.2. The model is in line with the equivalent discrete-time channel model, introduced for a diffusion-based molecular communication channel in [DK+17]. Therefore, the pixel energy samples for the binary equivalent of the transmitted message sequence "ICT" in Figure 4.1 is depicted in Figure 4.3, which are then subjected to various detection algorithms to detect the transmitted data symbols.

4.2 Simple Binary Transmission and Threshold Detection

Once the data acquisition process is established, data transmission can be achieved by employing low-complexity modulation of the data based on the spraying action at the transmitter side. The modulated data is then detected and deciphered at the receiver side using low-complexity detection algorithms. This section discusses these methods in detail.

4.2.1 On-Off Keying

On-off keying (OOK) is arguably the simplest type of modulation scheme, which is the unipolar version of the more popular binary amplitude shift keying used for several applications in wireless communication. Here, the information is represented by switching the signal on or off. Due to low energy consumption, especially when there is no transmission of information, OOK finds its application in optical transmission and battery-operated remote controls. For a diffusion-based molecular communication channel, the first implementation of OOK is presented in [MM+10], where the binary information is represented by the presence and absence of molecules released from the transmitter side. Therefore, a single type of molecule is sufficient to represent the information. In this dissertation, OOK modulated information is represented using a single transmit sprayer containing either a water-based solution of uranine or the highlighter ink. The chosen symbol duration is $T = 0.05\,\mathrm{s}$ of

which the solenoid valve of the sprayer remains open for a duration of $T_{\text{on}} = 0.02\,\text{s}$ for representing a binary data symbol "1", while it remains closed for a binary data symbol "0". For the parameters like T and T_{on}, their values are chosen in conformity with the reasons stated in Section 2.4.2. For illustration purpose, the depiction of OOK modulated data in terms of the received pixel energy samples $y[k]$ for a data sequence "101101" is presented in Figure 4.4, with the dashed lines representing a symbol duration T.

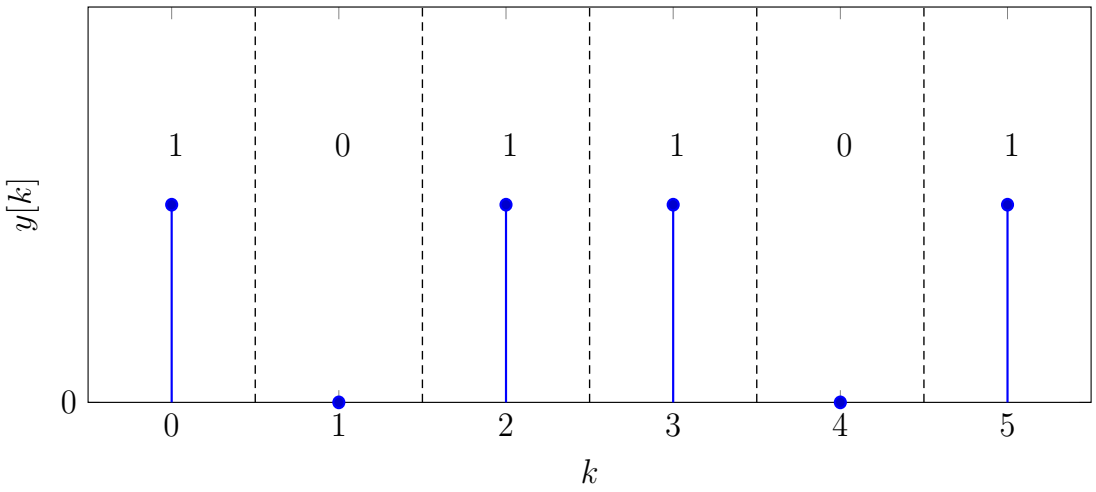


Figure 4.4: Depiction of the received pixel energy samples $y[k]$ for an on-off keying modulated data transmission sequence.

4.2.2 Threshold Detection

Threshold detection refers to a fundamental technique employed for the detection of the transmitted data, which gets affected by noise in the transmission channel. It involves establishing a decision boundary, known as the threshold, to discriminate between different data symbol states or levels. It is a low-complexity detection technique, typically employed in those systems which have limited access to computational power such as the diffusion-based molecular communication system. The threshold detection is also incorporated in the macroscopic molecular communication testbed presented in this dissertation.

Fixed Threshold Detector

With respect to molecular communication, fixed threshold detector (FTD) is the simplest, the least complex, and the most common type of detector found in nature. FTD is operational inside the human bodies, where the transfer of calcium ions above a certain concentration level affects muscle contraction, cell locomotion, hormonal secretion, and

neural transmission [CL08]. In another example, insects such as ants make use of the FTD-based signaling techniques for conveying the message about perceived threat or quality of food by producing a higher concentration trail of pheromones, which is detected by the other ants in the colony [RC09]. In molecular communication, FTD is used to detect the transmitted data symbols. For a diffusion-based molecular communication channel, when the number of absorbed molecules at the receiver side within a transmission interval exceeds a certain threshold, the decision is made accordingly for the corresponding transmitted data symbol [MM$^+$11]. However, in this dissertation, when the received pixel energy value $y[k]$ at the receiver side accumulated over a symbol duration T exceeds a certain threshold ξ, the data symbol estimate $\hat{x}[k]$ corresponding to the transmitted data symbol $x[k]$ is assigned with the decision value. The value of ξ remains unchanged during the detection process of the entire sequence of symbols in the case of FTD, which makes it suitable only for concentration-based modulation schemes. The FTD rule for OOK modulation scheme can be given as

$$\hat{x}[k] = \begin{cases} 0 & \text{if } y[k] \leq \xi, \\ 1 & \text{if } y[k] > \xi. \end{cases} \tag{4.2}$$

Adaptive Threshold Detector

Although FTD is a low-complexity detector, it is not well-suited for countering inter-symbol interference (ISI). Therefore, the problem can be solved by introducing the concept of adaptive threshold detector (ATD), as first introduced for a diffusion-based molecular communication channel in [DH16]. In ATD, the detection threshold changes depending on the received number of molecules. It is a biologically inspired detection algorithm motivated by the synaptic plasticity witnessed in human neural networks. Based on the previous stimuli, the synapses either increase (enhancement) or decrease (depression) their activity for a stimulus. In this dissertation, the threshold value $\xi[k]$ changes with time instance k, depending on the received pixel energy value of the previously transmitted symbol $y[k-1]$. This is then used to determine the symbol estimate $\hat{x}[k]$ corresponding to the transmitted data symbol $x[k]$. The ATD rule for OOK modulation scheme can be given as

$$\hat{x}[k] = \begin{cases} 0 & \text{if } y[k] \leq \xi\,[k] = y[k-1], \\ 1 & \text{if } y[k] > \xi\,[k] = y[k-1]. \end{cases} \tag{4.3}$$

One disadvantage of using ATD is that without ISI in the channel, the detection algorithm performs poorly, especially when there is a sequence with a long chain of binary data symbol "1". On the other hand, ATD is advantageous in the sense that it requires no knowledge of the end-to-end system coefficients h_ℓ, saving computational complexity and channel estimation overhead [Dam20].

4.2.3 Error Performance of the Transmission System

The transmission link using OOK modulation scheme at the transmitter side and a threshold-based detection algorithm at the receiver side is established by separately using water-based solutions of both uranine as well as the highlighter ink. The bit error rate (BER) performance of the information carriers are compared with increasing distance d. The theoretical BER as well as the BER based on the testbed measurements and the analytical model are evaluated and analyzed in this section. The theoretical BER is especially analyzed by taking into account 108 OOK modulated data symbols over 20 trials, detected at the receiver side using FTD. The data transmission sequence length is chosen in a way, so as to make it a possibility to compare up to 16-ary higher-order modulation schemes (8-ary higher-order modulation schemes already implemented in this work, as discussed in Section 4.4). The received energy samples $y[k]$ corresponding to each binary data symbol is separately represented by a probability density function (PDF). Since the received energy samples representing the unipolar data symbols are always non-negative, the received energy samples are only defined for the region $0 \leq y[k] < \infty$. PDF for the received energy samples $y[k]$ (denoted by y for simplicity) representing the binary data symbol "0" is modeled by a half Gaussian distribution $\mathcal{N}_{\text{half}}(0, \sigma_0^2)$, which is formulated as

$$\mathcal{N}_{\text{half}}(0, \sigma_0^2) = \frac{\sqrt{2}}{\sigma_0 \sqrt{\pi}} \exp\left(-\frac{y^2}{2\sigma_0^2}\right), \tag{4.4}$$

where σ_0 is the standard deviation of the distribution. Similarly, PDF for the received energy samples $y[k]$ representing the binary data symbol "1" is modeled by a Gaussian distribution $\mathcal{N}(\mu_1, \sigma_1^2)$, which is formulated as

$$\mathcal{N}(\mu_1, \sigma_1^2) = \frac{1}{\sigma_1 \sqrt{2\pi}} \exp\left(-\frac{(y - \mu_1)^2}{2\sigma_1^2}\right), \tag{4.5}$$

where μ_1 and σ_1 are the mean and the standard deviation of the distribution, respectively. In order to obtain the mean and the standard deviation values of the distributions with

increasing distance d between the transmitter and the receiver, the histogram curve-fitting function `histfit()` is used, which is available in MATLAB. The histogram bin size is chosen to be 50 with each bin having a width of 0.00048 for binary data symbol "0", and 0.0024 for binary data symbol "1". The values of μ_1, σ_0 and σ_1 with increase in d for the molecular communication testbed presented in this work are listed in Appendix C.1. An evident observation shows that $\sigma_1 > \sigma_0$, a clear indication to the presence of the amplitude-dependent noise in the end-to-end system.

PDFs of binary data symbol detection based on the received green pixel energy samples in the case of uranine water-based solution, and the received red pixel energy samples in the case of the highlighter ink water-based solution with increasing distance d are presented in Figure 4.5 and Figure 4.6, respectively. An error is said to have occurred when a transmitted data symbol $x\,[k]$ is incorrectly detected. The general expression of the symbol error probability for an erroneous detection of the transmit symbol $x[k]$ (denoted by x for simplicity) at the receiver side is given as

$$P_{\mathrm{s}} = \sum_{x \in \mathcal{X}} P(x) P(\hat{x} \neq x \,\Big|\, x), \tag{4.6}$$

where $P\,(x)$ is the probability of occurrence of the transmitted data symbol, $P(\hat{x} \neq x \,\Big|\, x)$ is the probability of receiving the erroneous data symbol, and $\mathcal{X}$ is the symbol alphabet. Therefore, for a system incorporating binary data symbols for data transmission purposes, the symbol error probability P_{s} is the same as the bit error probability P_{b}. Thus, P_{b} for OOK modulated binary data symbols is given as

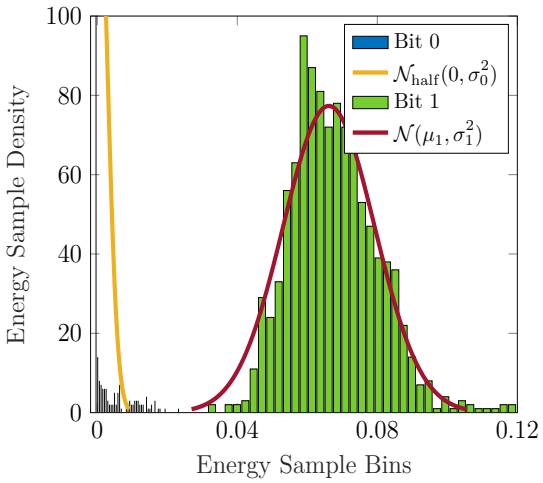
$$P_{\mathrm{b}} = P(0) P(1 \,\Big|\, 0) + P(1) P(0 \,\Big|\, 1). \tag{4.7}$$

The decision threshold ξ is the intersection point of the two distributions $\mathcal{N}_{\mathrm{half}}(0, \sigma_0^2)$ and $\mathcal{N}(\mu_1, \sigma_1^2)$, beyond which, a transmitted symbol is erroneously detected, i.e., if binary data symbol "0" is transmitted, then it is detected as a binary symbol "1" and vice versa. Hence, the theoretical BER is evaluated for equiprobable binary data symbols as

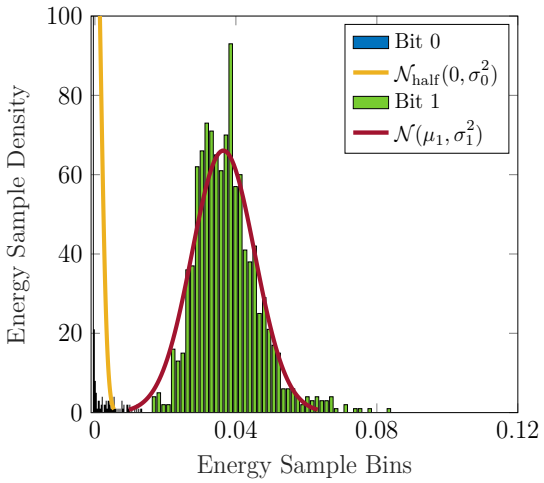
$$P_{\mathrm{b}} = \frac{1}{2} \int_0^{\xi} \frac{1}{\sigma_1 \sqrt{2\pi}} \exp\left(-\frac{(y - \mu_1)^2}{2\sigma_1^2} \right) dy + \frac{1}{2} \int_{\xi}^{\infty} \frac{\sqrt{2}}{\sigma_0 \sqrt{\pi}} \exp\left(-\frac{y^2}{2\sigma_0^2} \right) dy. \tag{4.8}$$

The derivation to evaluate ξ is described in Appendix C.2 in detail.

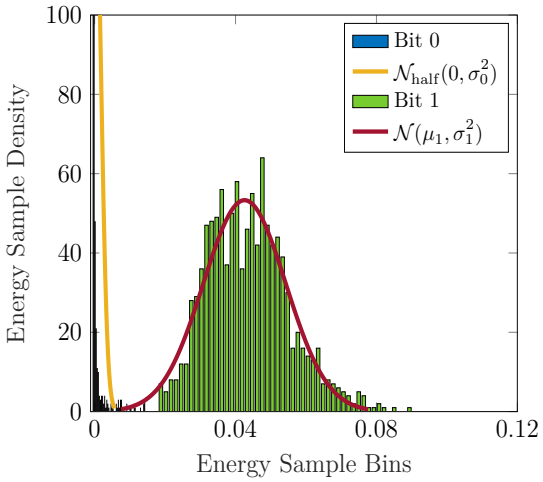
A performance comparison of the theoretical BER for both uranine and the highlighter ink water-based solutions along with BER obtained from the testbed measurements and



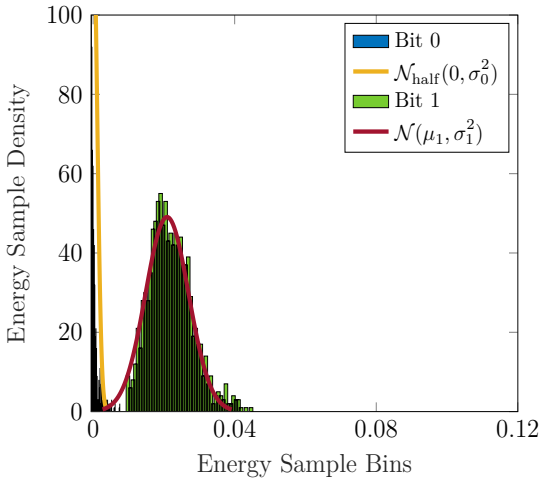
(a) $d = 58\,\text{cm}$



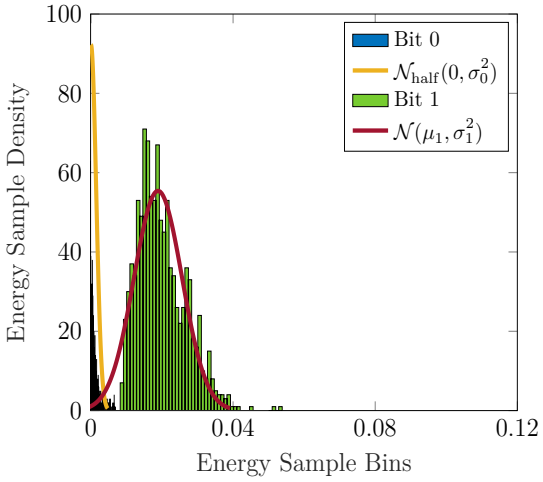
(b) $d = 88\,\text{cm}$



(c) $d = 118\,\text{cm}$

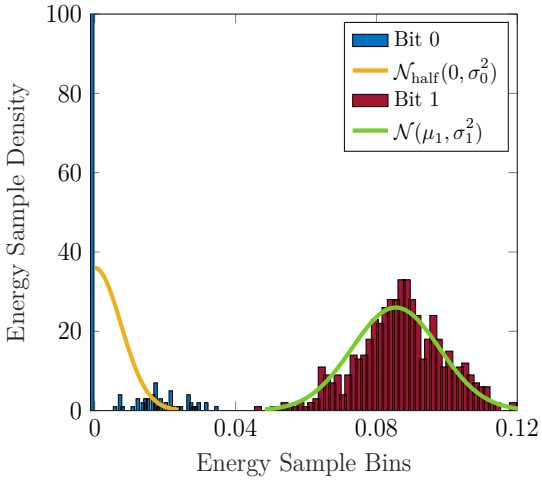


(d) $d = 148\,\text{cm}$

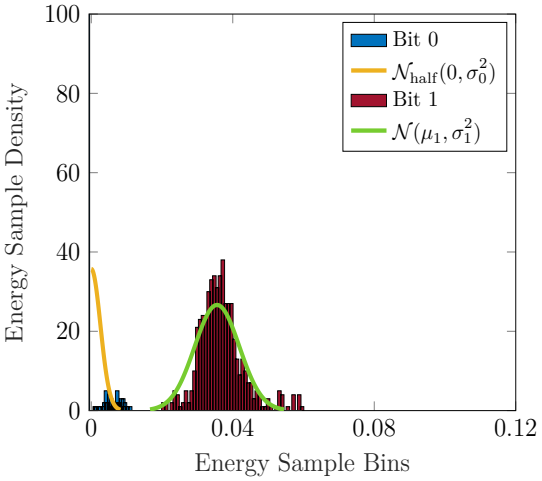


(e) $d = 178\,\text{cm}$

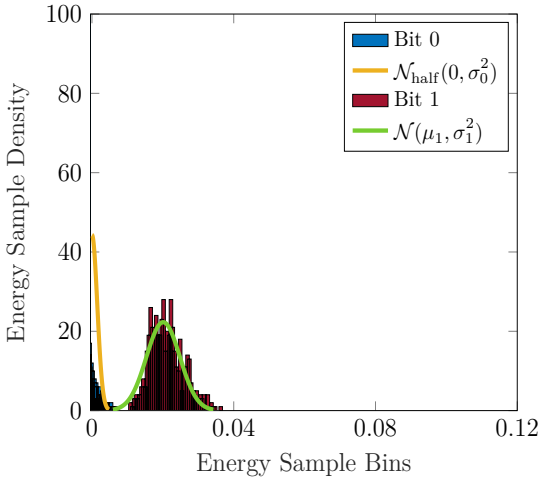
Figure 4.5: Distribution of the received green pixel energy samples of binary data symbols with increasing distance d using water-based solution of uranine.



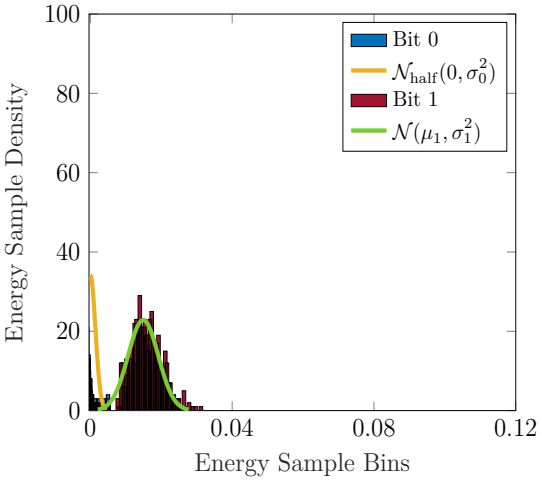
(a) $d = 58\,\text{cm}$



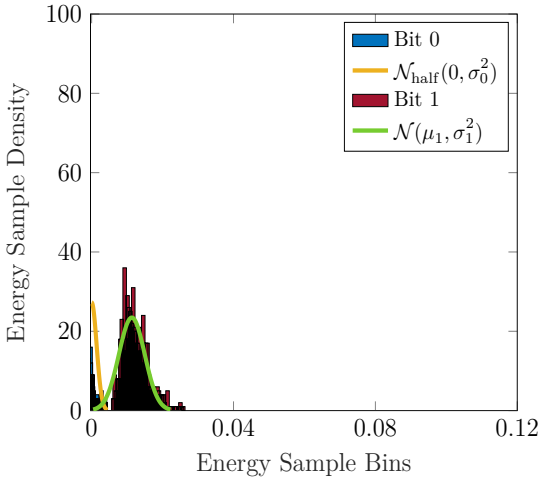
(b) $d = 88\,\text{cm}$



(c) $d = 118\,\text{cm}$



(d) $d = 148\,\text{cm}$



(e) $d = 178\,\text{cm}$

Figure 4.6: Distribution of the received red pixel energy samples of binary data symbols with increasing distance d using water-based solution of the highlighter ink.

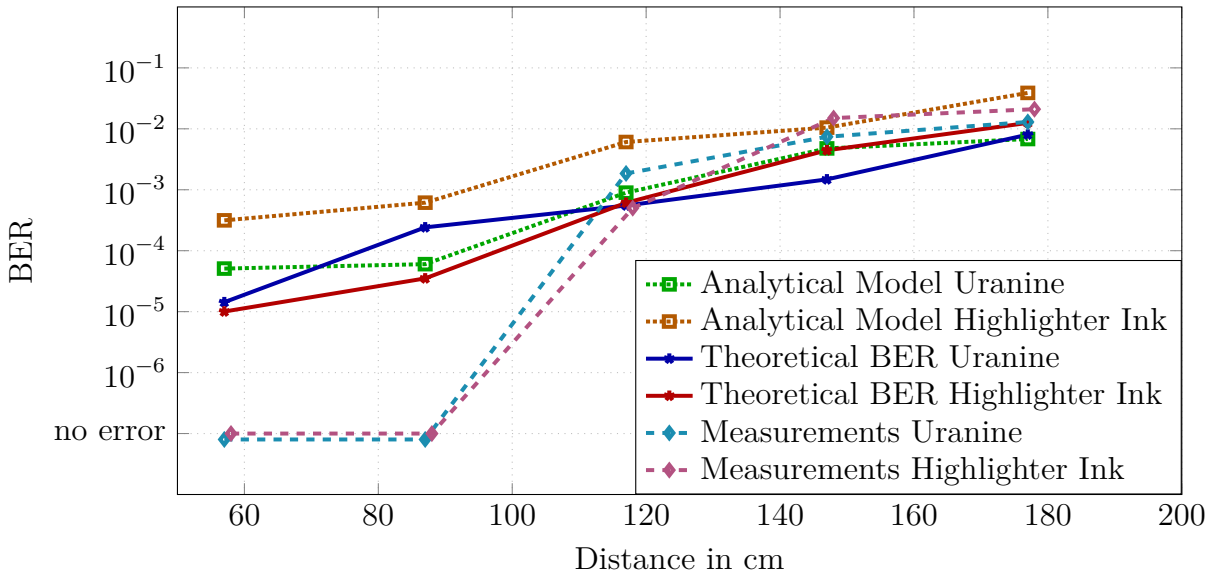


Figure 4.7: Comparison of theoretical bit error rate of the on-off keying modulated binary data symbols for water-based solutions of uranine and the highlighter ink with the bit error rates obtained from the testbed measurements and the analytical model for increasing distance d.

the analytical model is presented in Figure 4.7. In the case of the testbed measurements, a sequence of 108 randomly generated OOK modulated binary data symbols is transmitted over 20 trials for increasing distance d. In the case of the analytical model, where a better observation of the BER performance can be made for shorter values of d, OOK modulated and randomly generated binary data symbols of the same sequence length is transmitted over 100000 trials. As a common observation, it is seen that for both uranine as well as the highlighter ink water-based solutions, the BER performance deteriorates with increase in d. This is because there is a significant reduction in the number of droplets traveling the length of the tube and reaching the field of view of the camera-based detector at the receiver side. The droplets either fall down and collect at the bottom of the tube or simply get dispersed laterally. Also, the trajectory assumed by the spray is influenced by the gravitational pull of the Earth, increasing the distance between the camera-based detector and the droplets, as discussed in Section 3.3.2, thus reducing the intensity of light received by the camera. All these influences decrease the received signal strength while contributing to the noise power. Moreover, the slowing down of the sprayed droplets due to the effect of the drag force further leads to an increase in ISI. For smaller transmission distances, more accurate results are obtained with the help of the theoretical BER and the BER obtained from the analytical model, as opposed to the results obtained from the

testbed measurements. For larger transmission distances, the theoretical BER performance of the uranine water-based solution as the information carrier is better than that of the highlighter ink water-based solution. This observation also holds true when considering the BER results obtained from both the testbed measurements as well as the analytical model. This is due to the presence of the Bayer filter in the optical sensors of the camera-based detector, which is more sensitive to the fluorescence producing wavelength in the case of uranine water-based solution (548 nm) as opposed to the fluorescence producing wavelength in the case of the highlighter ink water-based solution (594 nm) [$HX^{+}10$]. Henceforth, all the testbed measurements are carried out using the water-based solution of uranine, unless mentioned otherwise.

4.3 Channel Equalization

In communication theory, equalization is a technique that is used in order to mitigate the effects of channel distortion on the transmitted signals. For an electro magnetic (EM) wave-based communication, when data is transmitted through a communication channel, the channel can introduce various impairments that include signal attenuation, delay spread, and frequency-selective fading. These impairments can distort the original signal causing ISI, and reducing the quality of the received signal. Equalization, therefore, helps compensate for these distortions and restores the original signal as closely as possible at the receiver side of the transmission system. The macroscopic air-based molecular communication testbed using fluorescent dyes presented in this dissertation displays clear evidence of ISI with increasing distance between the transmitter and the receiver, which proves to be a major deterrent to the BER performance of the various transmission algorithms employing modulation schemes. Hence, popular equalization techniques from the literature are implemented in the molecular communication system to address the issue.

Prior to implementing an equalizer, least squares channel estimation technique can be used as a tool to estimate the end-to-end system coefficients of a molecular communication system. This is a popular method used in classical wireless communication systems to estimate the characteristics of a communication channel. When the signals from the transmitter towards the receiver get distorted due to noise, interference, multipath propagation, etc., channel estimation helps determine the behavior of the channel so that the received signal can be properly equalized or corrected. A popular approach in the estimation of the channel coefficients include the transmission of a known pilot sequence with the combination of the knowledge of the transmitted and the received signals. For the molecular

communication testbed presented in this dissertation and for a system memory length of L, the $L+1$ end-to-end system coefficients can also be estimated using the least squares channel estimation strategy. For a pilot matrix $\mathbf{A}$ of size $(L+1) \times K$, where K is the length of randomly generated known pilot sequence, and the received pixel energy vector $\mathbf{y}$ of size $1 \times K$, the estimated end-to-end system coefficient vector $\hat{\mathbf{h}}$ is given as

$$\hat{\mathbf{h}} = \mathbf{y}\mathbf{A}^{\top}\left(\mathbf{A}\mathbf{A}^{\top}\right)^{-1}. \tag{4.9}$$

It must be mentioned that the estimated end-to-end system coefficients are not normalized in order to preserve the features like the hardware position/configuration related and the distance dependent signal strength.

4.3.1 Zero-Forcing Equalization

Equalization techniques are commonly used in digital communications theory as a measure to counteract the ISI originating in the data transmission channel [PS07]. A linear equalization technique named the zero-forcing (ZF) equalizer has been implemented in this dissertation for the fluorescent dye-based macroscopic molecular communication testbed. Similar to the equivalent discrete-time end-to-end system model, where the memory and the end-to-end system coefficients form the building blocks of a finite impulse response (FIR) filter, the ZF equalizer is also represented in the form of an FIR filter with its coefficients g_j in order to invert the characteristics of the end-to-end system. This helps to cancel out the ISI. The convolution of the end-to-end system coefficients with the ZF filter coefficients in the discrete-time domain results is a unit impulse that can be formulated as [PS07]:

$$\sum_{j=-\infty}^{\infty} g_j h_{\ell-j} \stackrel{!}{=} \begin{cases} 1 & \text{if } \ell = 0, \\ 0 & \text{if } \ell \neq 0. \end{cases} \tag{4.10}$$

In principle, the system behaves like an FIR filter, implying that the ZF equalizer should ideally be implemented as an infinite impulse response (IIR) filter. However, given the complexity involved in the design and realization of an IIR filter with infinite filter coefficients, the ZF equalizer is implemented as an FIR filter with G filter coefficients, as shown in Figure 4.8. This particular constraint affects the performance of the ZF equalizer as the ISI is not removed completely. The system is however modeled on the one dimensional solution of the advection-diffusion mass transportation process which is causal, stable, and minimum-phase, and is similar to the diffusion-based molecular communication channel.

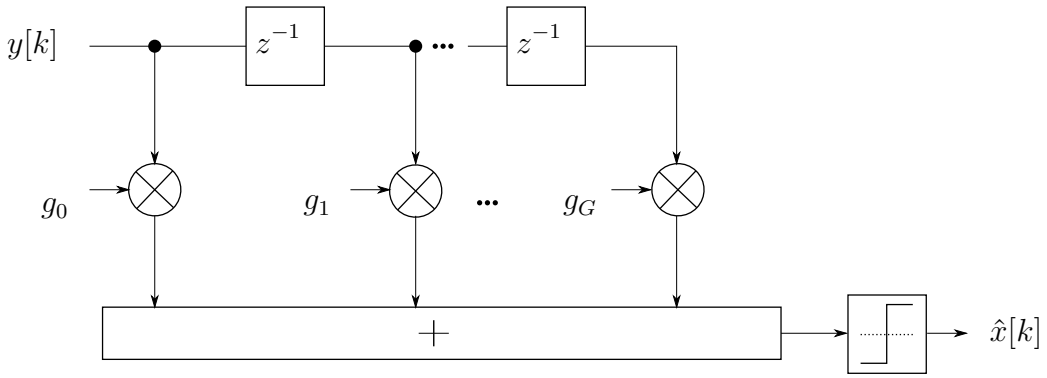


Figure 4.8: Block diagram of the zero-forcing equalizer.

This in turn makes the ZF filter causal, stable, and minimum-phase, with a very small influence of the IIR filter coefficients, which has not been considered in the design process. Finally, the transmit symbol $x[k]$ is estimated as $\hat{x}[k]$ with the help of a threshold detection process, without considering the effects of ISI. Therefore, FTD presented in Section 4.2.2 is used for the final detection of symbols when applying the ZF equalizer, as also seen from Figure 4.8.

4.3.2 Decision-Feedback Equalization

Another popular equalizer in digital communications is the decision-feedback equalizer (DFE) that is used for effectively correcting the current transmit symbol by relying on the decisions made with respect to the previously transmitted symbols. The DFE is a nonlinear type of equalizer whose performance is often better than that of a linear equalizer [PS07], and is implemented for the macroscopic molecular communication testbed presented in this work. It comprises of two parts, a feedforward filter that helps in the truncation of a long system impulse response, and a nonlinear feedback filter that helps in the removal of any residual ISI through the previously detected symbols [PS07]. In this dissertation, however, a DFE is considered without taking into account the feedforward filter, as depicted in Figure 4.9. It is the special case, where the filter coefficients are equal to the coefficients of the system memory. Therefore, the equalized received sample is given as [Dam20]

$$y'[k] = y[k] - \sum_{\ell=1}^{L} \hat{x}[k-\ell]h_\ell, \tag{4.11}$$

where $\hat{x}[k-\ell]h_\ell$ is the estimated ISI available at the input of the DFE through the feedback of the previously decided symbols $\hat{x}\,[k-\ell]$, which is then subtracted from the received pixel

energy sample $y[k]$. The transmitted symbol estimate $\hat{x}[k]$ is finally obtained after $y'[k]$

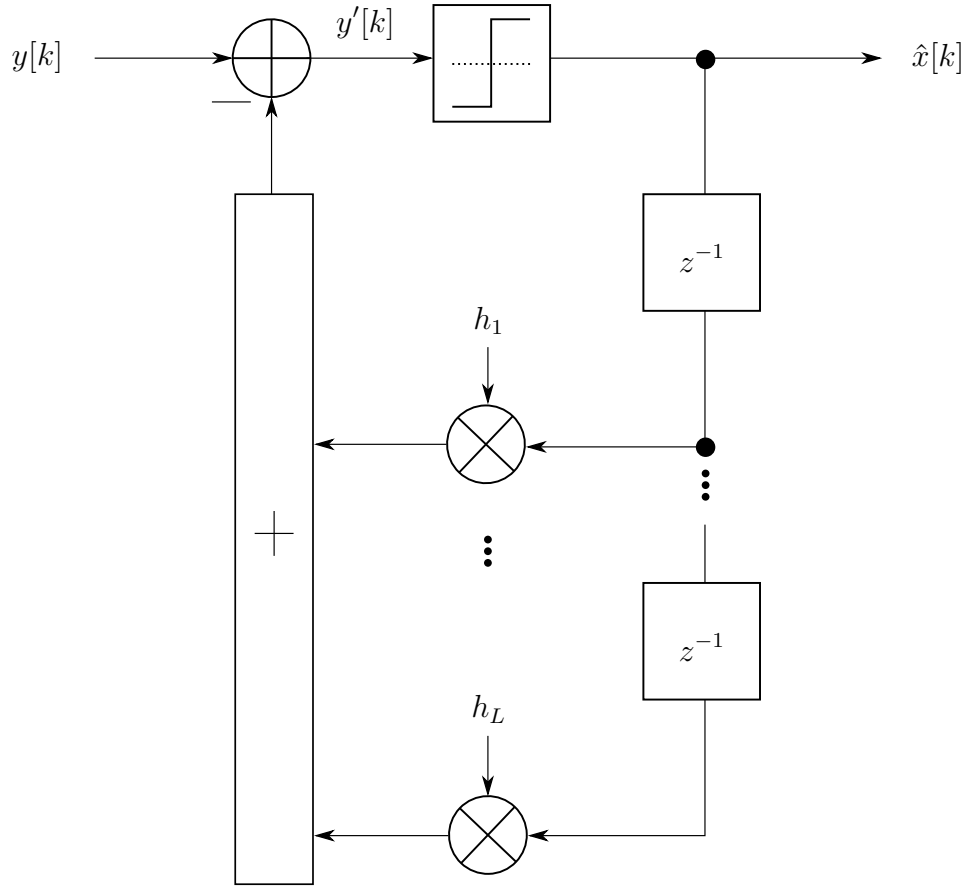


Figure 4.9: Block diagram of the decision-feedback equalizer without feedforward filter.

undergoes a threshold detection process, without taking the effects of ISI into account. For example, in this dissertation, FTD presented in Section 4.2.2 is used for the final detection of symbols when applying DFE. One advantage of using a nonlinear equalizer like DFE over a linear equalizer like ZF is that the former does not amplify the noise as opposed to the latter. However, if an erroneous detection of a symbol occurs, the error then propagates across the entire transmission sequence limiting the performance of the DFE [PS07].

4.3.3 Maximum-Likelihood Sequence Estimation

The maximum-likelihood sequence estimation (MLSE) approach is an optimal detection strategy that helps in the minimization of error present in the transmitted data sequence. With the help of MLSE technique as a detection algorithm for the macroscopic air-based

molecular communication using fluorescent dyes, the most probable transmitted data sequence is determined by considering every possible sequence [PS07], and is mathematically represented as

$$\hat{\boldsymbol{x}} = \arg\max_{\tilde{\boldsymbol{x}}} \left\{ p\left(\boldsymbol{y} \mid \tilde{\boldsymbol{x}}\right) \right\}, \tag{4.12}$$

where $\hat{\boldsymbol{x}}$ is the sequence of the transmitted data symbol estimate, $\tilde{\boldsymbol{x}}$ is the sequence of the data symbol hypothesis, and $\boldsymbol{y}$ is the sequence of the received pixel energy samples at the receiver side. The MLSE detection strategy has many advantages over conventional equalization techniques. For example, the MLSE-based detectors neither amplify the noise as in the case of ZF equalizer, nor do their use lead to any error propagation, as witnessed in DFE. A low-complexity and elegant version of the MLSE-based detection scheme is the Viterbi algorithm [For73], which looks for the most probable path in a trellis diagram leading to the identification of the data sequence that has most-likely been transmitted. A trellis diagram comprises of K trellis segments, where K is the number of transmitted data symbols in a sequence, resulting in one trellis segment for every transmission interval. An example of a trellis segment with binary symbols and a system memory length $L = 2$ is shown in Figure 4.10. A trellis segment consists of M^L states, where M is the cardinality of the transmitted data symbols. These states represent all possible symbol combinations

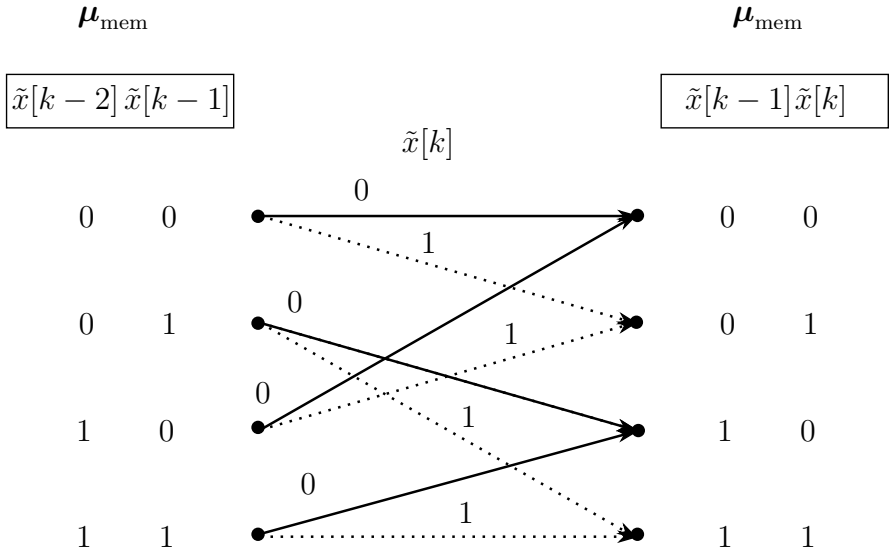


Figure 4.10: Trellis segment of a binary data transmission sequence with system memory length $L = 2$ in maximum-likelihood sequence estimation detection strategy.

in the memory elements of the equivalent discrete-time end-to-end system model. Thus, the trellis segment in Figure 4.10 consists of $2^2 = 4$ states. Two adjacent states in a trellis segment are connected by a branch. Each state can therefore be reached by M branches, and

M branches emanate from a particular state. These branches constitute a transmit data symbol in the current time instance, which is subsequently present in a memory element of the equivalent discrete-time end-to-end system model in the next time instance. In the case of binary data symbols, a solid branch depicts symbol $\tilde{x}[k] = 0$ and the dashed branch depicts symbol $\tilde{x}[k] = 1$. Each branch is assigned with a cost when transitioning from one state to another. The branch metric is denoted by $\lambda_{\boldsymbol{\mu}_{\text{mem}}}^{\tilde{x}[k]}[k]$, where $\boldsymbol{\mu}_{\text{mem}}$ represents the current memory state. The sum of all branch metrics that lead to the next memory state at time k is called the path metric, which is denoted by $\Gamma_{\boldsymbol{\mu}_{\text{mem}}}[k]$. Thus, the implementation of the Viterbi algorithm consists of the following three steps:

1. *Initialization:* The branch metrics of all trellis segments corresponding to the length of the transmitted data sequence are calculated in this step based on the rule $\lambda_{\boldsymbol{\mu}_{\text{mem}}}^{\tilde{x}[k]}[k] = \left| y[k] - \sum_{\ell=0}^{L} h_\ell \tilde{x}[k-\ell] \right|^2$. Furthermore, at the beginning of the transmission, the memory state is assumed to be empty, and the path metric of the first all-zero state is set to zero, while the path metric of every other state is initialized to infinite.

2. *Add, compare, select:* This step is sequentially executed for every new state in the successive trellis segment. Initially, the new path metrics are computed based on the rule $\Gamma_{\boldsymbol{\mu}_{\text{mem}}}[k] = \Gamma_{\boldsymbol{\mu}_{\text{mem}}}[k-1] + \lambda_{\boldsymbol{\mu}_{\text{mem}}}^{\tilde{x}[k]}[k]$ (*add*), following which, the different path metrics leading to the new state are compared with one another (*compare*). Finally, the path with the least cost is preserved as the surviving path, while every other path with higher cost is eliminated (*select*).

3. *Trace-back:* Finally, at the trellis diagram corresponding to the end of the transmitted data sequence, the minimum path metrics amongst all final states are compared. The state with the lowest path metric marks the beginning of the path realizing the most probable sequence. The path from that point onwards is traced back through the surviving trellis paths. Each selected branch is assigned a particular symbol leading to the detection of the transmitted data sequence.

4.3.4 Results

A comparative BER performance analysis of the detection algorithms with increasing distance d is presented in this section. The BER performance based on the testbed measurements using the water-based solution of uranine for a randomly generated binary data sequence length of 108 over 20 trials is presented in Figure 4.11 (testbed measurement parameters listed in Table 3.1). Similarly, for a better observation of the BER performance

in the case of smaller values of d, the simulation results based on the analytical model presented in Figure 4.12 is considered for a randomly generated binary data of same sequence length over 100000 trials. An end-to-end system with a memory length $L = 3$ with OOK modulation scheme is considered for evaluating the BER performance of the detection algorithms. These algorithms can also be applied to higher order modulation schemes, described in Section 4.4. A known pilot sequence of length 108 over 20 trials is randomly generated for separately estimating the end-to-end system coefficients, using the least squares channel estimation technique. For the BER analysis, the length of the filter coefficients for both ZF equalizer and the DFE is considered to be the same as the end-to-end system memory length. As observed from both Figure 4.11 and Figure 4.12, FTD is the worst performer amongst all the detection algorithms, as it does not take ISI correction into account. The ZF equalizer, however, performs better than FTD but performs worse than DFE and MLSE-based equalizer. This is due to the fact that ZF equalizer is represented as a low-complexity FIR filter with finite filter coefficients, inverting the characteristics of the end-to-end system, instead of using an IIR filter in the ideal case. This hinders the performance of the ZF equalizer and the ISI is not completely removed. Furthermore, noise amplification also occurs for smaller end-to-end system coefficients. The DFE, however, only feeds back hard decisions, and does not amplify the noise. Therefore, it has a better BER performance as compared to the ZF equalizer. However, error propagation might still occur in the case of DFE if a decision is made incorrectly. The MLSE-based equalizer performs the best due to its high computational complexity. It is also interesting to see that the BER performance of the detection algorithms converge for higher values of d, indicating the fact that the noise in the system becomes a more dominant error contributor as compared to the ISI.

4.4 Higher-Order Modulation Schemes

Higher-order modulation schemes are advanced techniques used in digital communication systems to transmit and receive higher volume of data over a communication channel. These schemes involve encoding information using various degrees of freedom, allowing for the transmission of multiple binary data symbols per unit time, thus, increasing the data transmission rates and the spectral efficiency. Therefore, in this section, higher-order modulation schemes are discussed that are incorporated in the macroscopic air-based molecular communication testbed presented in this dissertation. Apart from using time, additional degrees of freedom in the form of fluorescent dye colors and the spatial regions

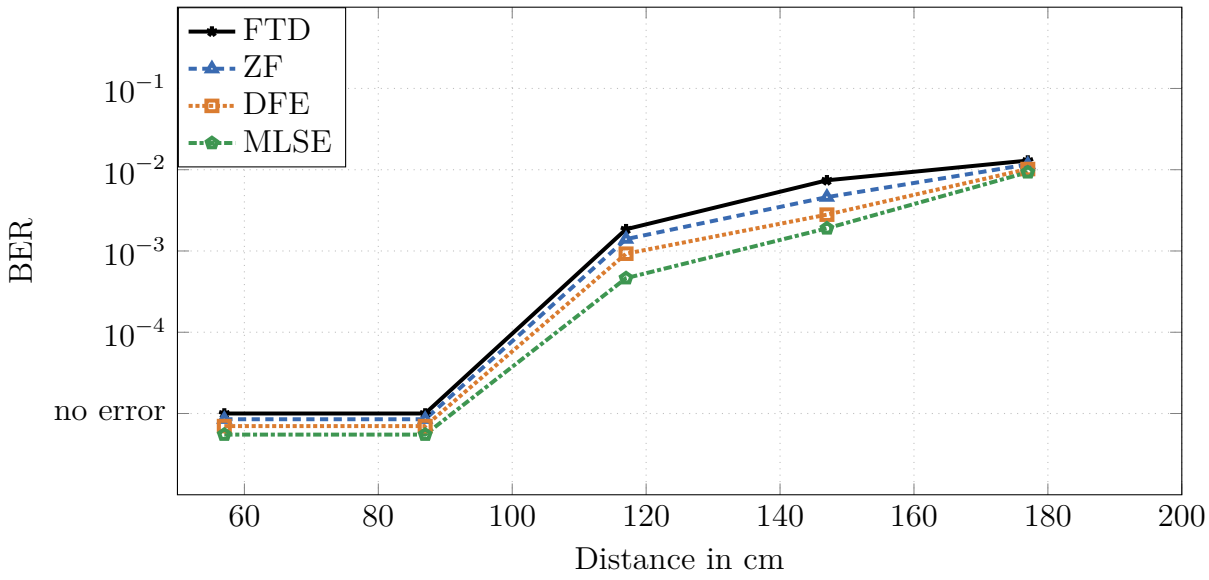


Figure 4.11: Bit error rate performance of the channel equalization techniques on the on-off keying modulated data transmission sequences based on the testbed measurements with increasing distance d.

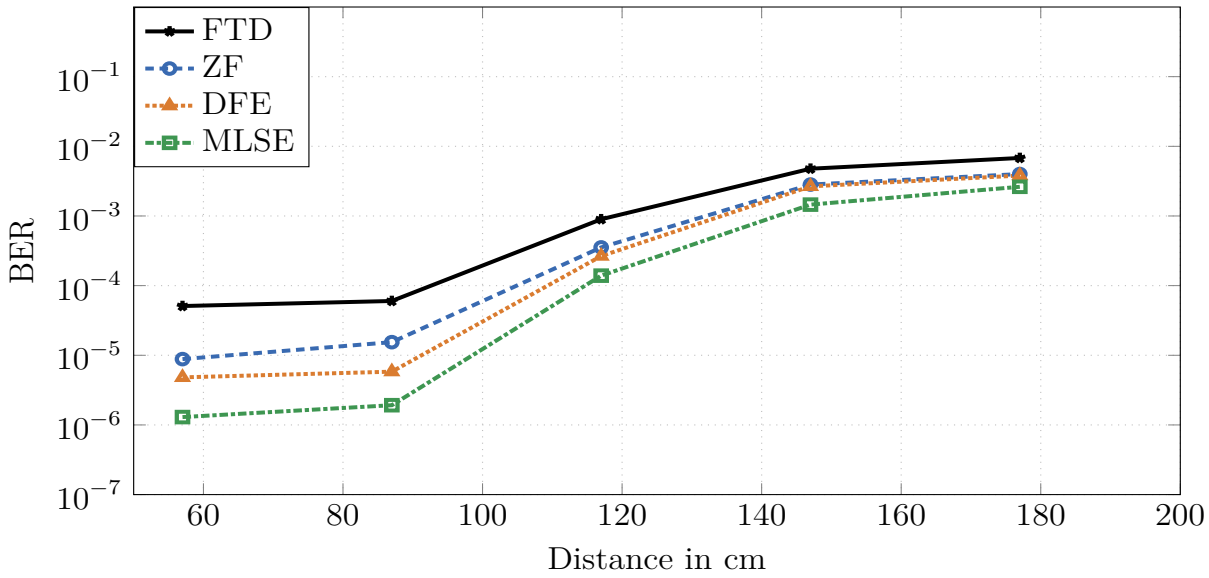


Figure 4.12: Bit error rate performance of the channel equalization techniques on the on-off keying modulated data transmission sequences based on the analytical model with increasing distance d.

populated by the information carriers are exploited to design the modulation schemes for achieving higher data transmission rates.

4.4.1 Pulse-Position Modulation

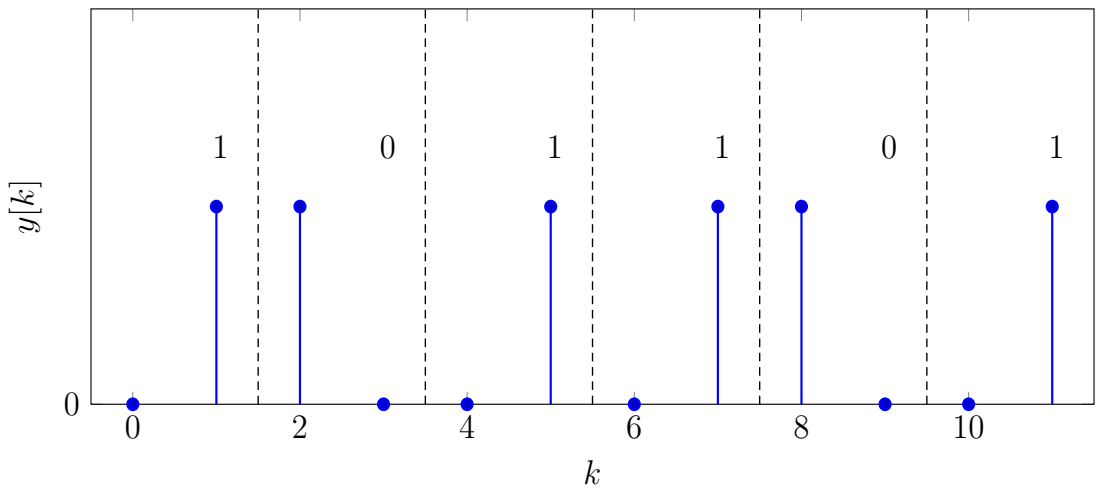


Figure 4.13: Depiction of the received pixel energy samples $y[k]$ for a 2-ary pulse-position modulated data transmission sequence.

Pulse-position modulation (PPM) is a popular modulation technique commonly used in optical communication, where the information is embedded in the position of a signal pulse within a transmission interval. Based on the same principle, PPM has also been introduced for a diffusion-based molecular communication channel in [GL+11], where the time of release of molecules within a transmission interval is utilized to represent the information. In this dissertation, PPM modulated information is represented using a single sprayer containing uranine water-based solution. The M data symbols in an M-ary PPM are encoded by transmitting a spray-pulse occupying one of the M chip slots of duration T_c, within a symbol duration $T = M\,T_\mathrm{c}$. Bearing in mind the hardware restrictions as mentioned in Section 2.4.2, suitable value of T_c is chosen. For this work, $T_\mathrm{c} = 0.025\,\mathrm{s}$ is considered, while the spray duration $T_\mathrm{on} = 0.02\,\mathrm{s}$. Additionally, 2-, 4-, and 8-PPM are implemented in this work. For illustration purpose, the depiction of 2-PPM in terms of the received pixel energy samples $y[k]$ for a binary data sequence "101101" is shown in Figure 4.13, with the dashed lines representing a symbol duration T.

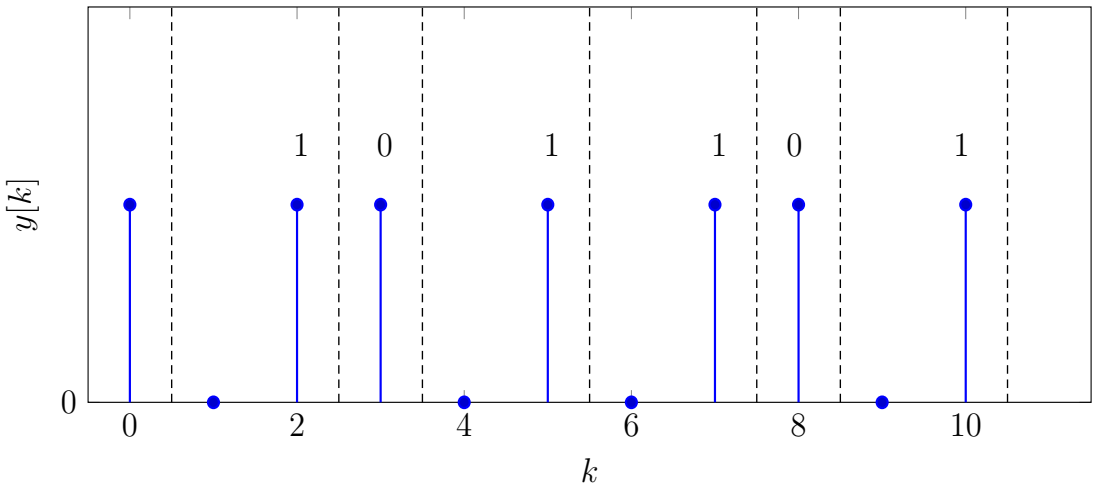


Figure 4.14: Depiction of the received pixel energy samples $y[k]$ for a 2-ary differential pulse-position modulated data transmission sequence.

4.4.2 Differential Pulse-Position Modulation

Differential pulse-position modulation (DPPM) is a commonly used modulation technique concerning applications in both radio-based wireless and optical communication [SK98; SK99], which is implemented by transmitting a reference signal pulse followed by the transmission of the subsequent pulses at various time intervals corresponding to the data symbols. The information is embedded within the time gap between two successive pulses. DPPM has been introduced in the form of communication through silence for a diffusion-based molecular communication channel in [GL+11], and its variation in the form of time-elapse communication for a testbed setup using genetically modified bacteria in [KA+13]. In this dissertation, DPPM is implemented by spraying uranine water-based solution from a single sprayer. A single reference spray-pulse of duration $T_{\text{pulse}} = 0.025\,\text{s}$ is transmitted with each of the subsequent symbols in an M-ary DPPM being transmitted with the same pulse duration after a certain time gap, which is in multiples of $T_{\text{gap}} = T_{\text{pulse}} = 0.025\,\text{s}$ from the previous pulse. The symbols in M-ary DPPM are represented as $s_i = i\,T_{\text{gap}}$, $i \in \{0, \ldots, M-1\}$. In this dissertation, 2-, 4-, and 8-DPPM have been implemented, and the pictorial representation of 2-DPPM in terms of the received pixel energy samples $y[k]$ for a binary data sequence "101101" is showcased in Figure 4.14, with the dashed lines representing a symbol duration $T = i\,T_{\text{gap}}$.

4.4.3 Shorten Pulse-Position Modulation

Shorten pulse-position modulation (SPPM), as first introduced in [SZ09], is a more bandwidth-efficient alternative to PPM modulation scheme that shows promise in applications such as underwater optical communication. In this scheme, the first binary data symbol of an M-ary SPPM is OOK modulated, while the remaining binary data symbols are PPM modulated, making it more bandwidth efficient as compared to the standard PPM. For the first time in [BD+22b], SPPM is implemented with regard to molecular communication. SPPM modulated information is represented using a single sprayer containing uranine water-based solution. The chip slot duration is the same as for PPM at $T_c = 0.025\,\mathrm{s}$, while the symbol duration can be expressed in the form of $T = \left(2^{\log_2 M-1} + \mathrm{mod}\left(2^{\log_2 M-1}, 2\right) + \left(\sqrt{-1}\right)^M\right) T_c$. In this work, 4- and 8-SPPM have been implemented, and the graphical representation of 4-SPPM in terms of the received pixel energy samples $y[k]$ for a binary data sequence "101101" is showcased in Figure 4.15, with the dashed lines representing a symbol duration T.

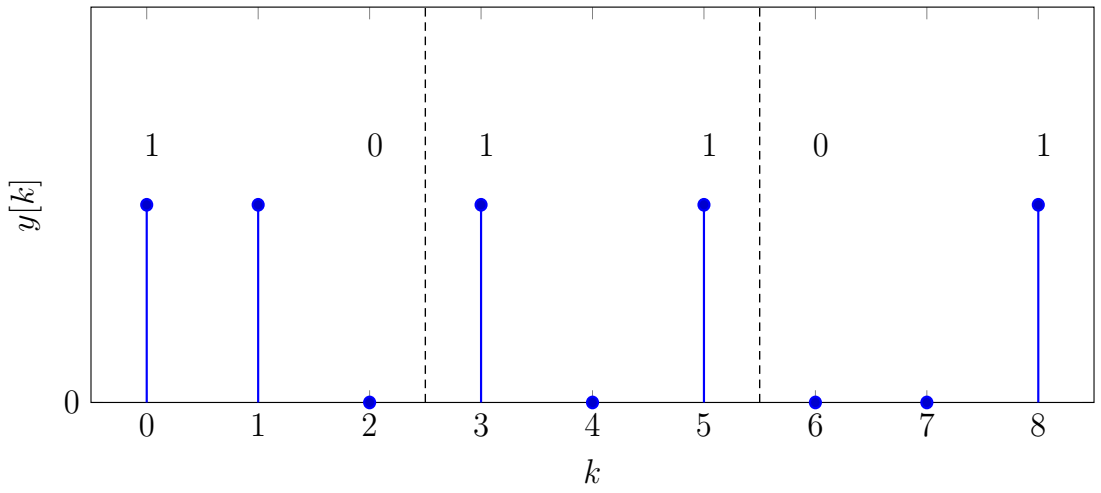


Figure 4.15: Depiction of the received pixel energy samples $y[k]$ for a 4-ary shorten pulse-position modulated data transmission sequence.

4.4.4 Multi-Pulse Pulse-Position Modulation

Another modulation scheme to improve the bandwidth utilization in PPM is proposed in [SN89] as multi-pulse pulse-position modulation (MPPM) for optical communication. Contrary to PPM, where one pulse occupies one of the M chip slots for representation of a data symbol, in an M-ary MPPM scheme, several pulses occupy the chip slots to represent one symbol. If K_c among N_c chip slots are to be occupied by the pulses to

represent a symbol, then up to $\binom{N_c}{K_c}$ symbols can be transmitted. However, the value $\binom{N_c}{K_c}$ is not a power of two, and in order to transmit $\log_2 M$ bits, the additional symbols must be modulated by the conventional PPM technique using a single pulse. In [BD+22b] for the first time, MPPM is implemented for a molecular communication testbed, where a single sprayer containing uranine water-based solution is used to modulate the symbols in an 8-MPPM scheme. Here, $N_c = 4$ and $K_c = 2$ are chosen, while the chip slot duration $T_c = 0.025\,\mathrm{s}$ remains the same as that of PPM. Thus, the symbol duration can be given as $T = M\,T_c/2^{\log_4(M/2)}$. The binary data symbols are mapped as shown exemplarily for the 8-MPPM in Table 4.1, while the pictorial illustration of the scheme in terms of the received pixel energy samples $y[k]$ for a binary data sequence "101101" is showcased in Figure 4.16, with the dashed lines representing a symbol duration T.

Table 4.1: Mapping for the 8-multi-pulse pulse-position modulation scheme.

Data Sequence	$\boldsymbol{x}$
[0 0 0]	[1 1 0 0]
[0 0 1]	[1 0 1 0]
[0 1 0]	[1 0 0 1]
[0 1 1]	[0 1 1 0]
[1 0 0]	[0 1 0 1]
[1 0 1]	[0 0 1 1]
[1 1 0]	[1 0 0 0]
[1 1 1]	[0 0 1 0]

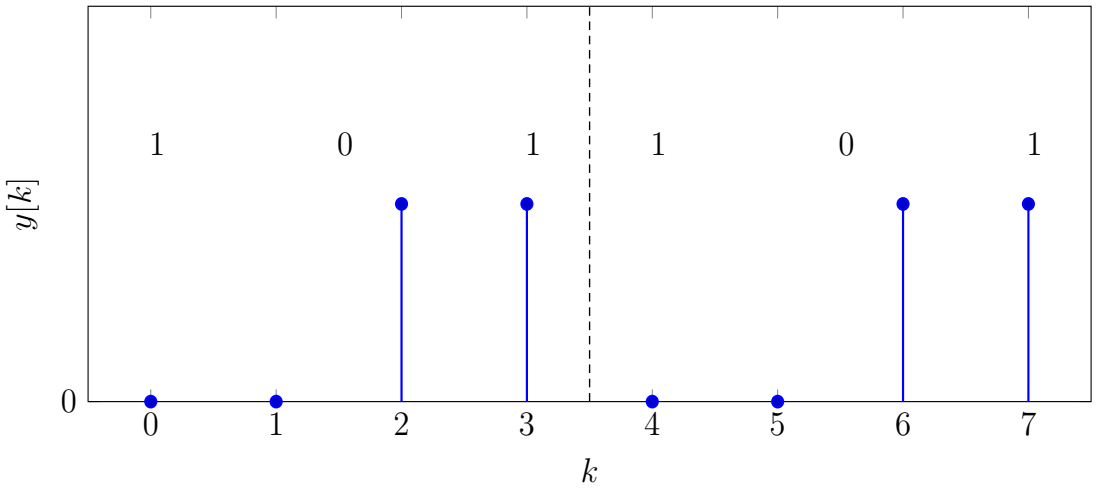


Figure 4.16: Depiction of the received pixel energy samples $y[k]$ for an 8-ary multi-pulse pulse-position modulated data transmission sequence.

4.4.5 Molecular Shift Keying

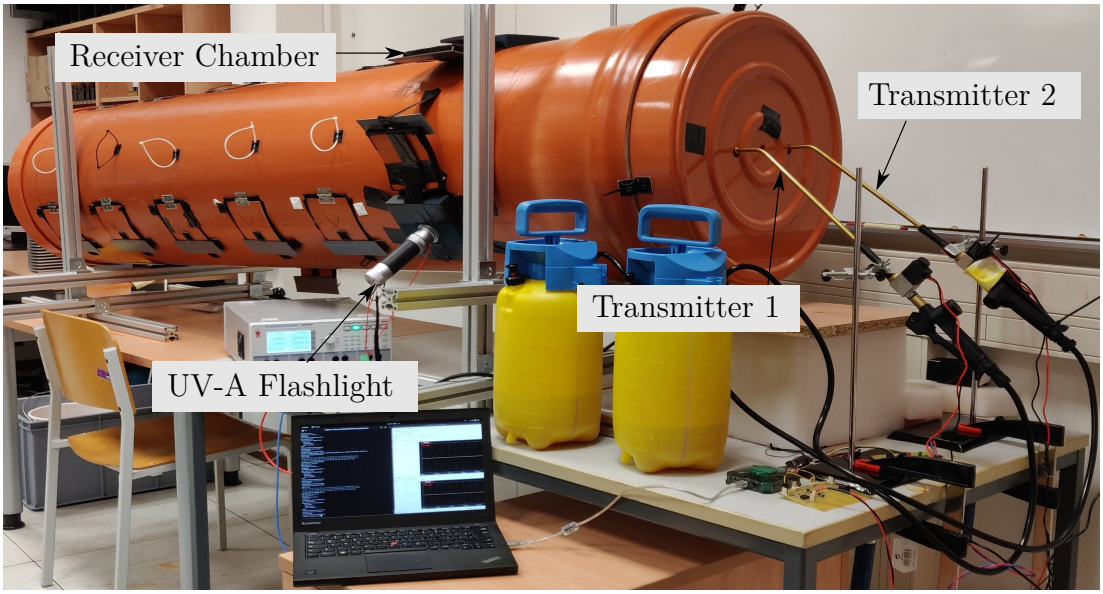
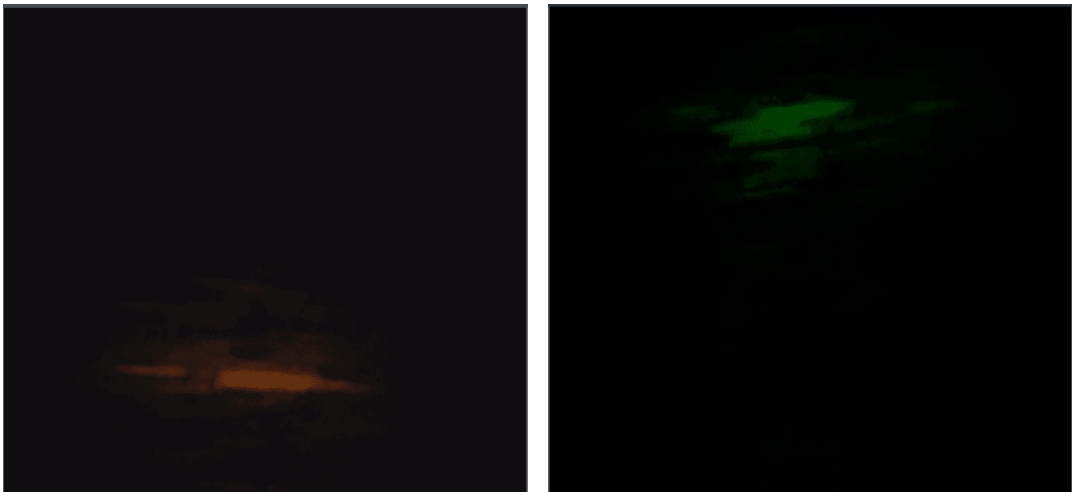


Figure 4.17: Testbed with two different sprayers, one containing uranine water-based solution, while the other containing highlighter ink water-based solution.

Molecular shift keying (MoSK), as the name suggests, embeds the information on the type of molecule that is being transmitted from the transmitter to the receiver, serving as an additional degree of freedom over concentration of molecules. The possibility to exploit an additional degree of freedom for data transmission is also perceived as a significant advantage over traditional EM wave-based wireless communication. The receiver for MoSK has the ability to distinguish between the different data symbols that are being carried by the respective molecules, which leads to the detection of the transmitted information. For a diffusion-based molecular communication channel, the first implementation of MoSK is presented in [KY$^+$11], which draws inspiration from hydrofluorocarbons messenger molecule structure for designing different molecule types to represent the information. Therefore, for an M-ary MoSK, M different types of molecules can be used to represent M different data symbols, each symbol comprising of $\log_2 M$ bits. In this dissertation, based on the same principle, both uranine as well as the highlighter ink water-based solutions are released from the transmitter side sprayers in order to represent different data symbols due to their property of emitting light of different wavelengths, when their sprayed water-based solution droplets come under the influence of the UV-A light source. At the receiver side, the green pixel intensity captured by the high-speed camera-based detector in the case of sprayed uranine water-based solution droplets is far more dominant than the red pixel intensity,



(a) Binary data symbol "0" **(b)** Binary data symbol "1"

Figure 4.18: Camera snapshot of water-based solution droplets of the highlighter ink and uranine for the transmission of binary data symbols in 2-ary molecular shift keying, at a distance $d = 88\,\text{cm}$.

while in the case of sprayed highlighter ink water-based solution droplets, the red pixel intensity is more dominant than the green pixel intensity (presented in Section 2.4.2 in detail). This helps the detector distinguish between the transmitted data symbols. The work presents the implementation of 2-MoSK realized by two sprayers, one filled with a uranine water-based solution, while the other filled with the highlighter ink water-based solution, as shown in Figure 4.17. The separation distance between the two sprayers at the transmitter side is set at 16 cm, which ensures the visibility of the sprayed droplets released by both the sprayers in the field of view of the camera-based detector at the receiver side. For a binary data symbol "0", the solenoid valve of the highlighter ink water-based solution is opened for a duration of $T_{\text{on}} = 0.02\,\text{s}$ right at the start of the transmission interval. Similarly, for a binary data symbol "1", the solenoid valve of the uranine water-based solution is opened for the same duration. The visibility of these water-based dye solutions in the field of view of the camera-based detector at a distance $d = 88\,\text{cm}$ for the transmission of respective data symbols is presented in Figure 4.18. The symbol duration in 2-MoSK is again chosen to be $T = 0.05\,\text{s}$. The graphical representation of 2-MoSK in terms of the received pixel energy samples $y[k]$ for a binary data sequence "101101" is presented in Figure 4.19, with the dashed lines representing a symbol duration T.

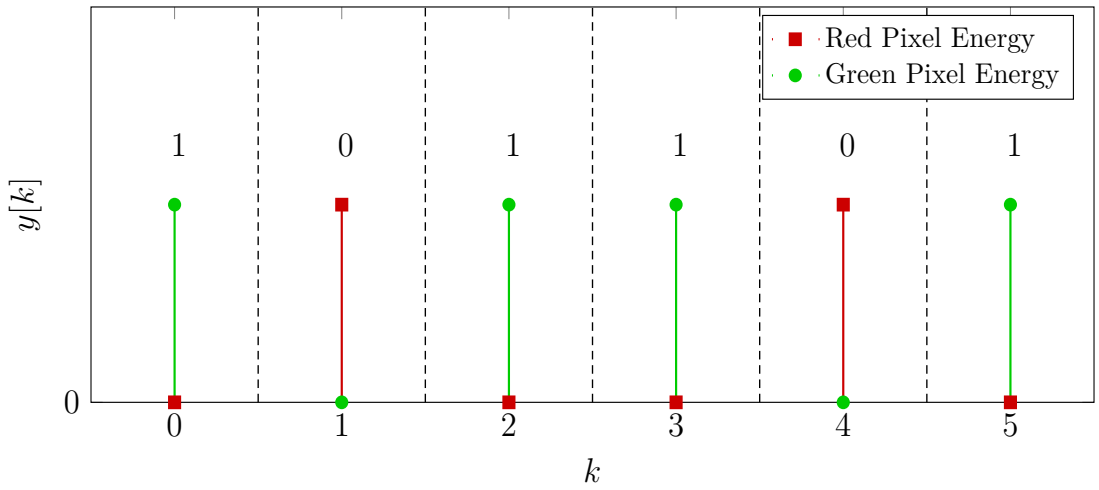


Figure 4.19: Depiction of the received pixel energy samples $y[k]$ for a 2-ary molecular shift keying modulated data transmission sequence.

4.4.6 Molecular Concentration Shift Keying

In molecular concentration shift keying (MCSK), the degrees of freedom such as the concentration level as well as the type of released molecules are combined together to represent the transmitted data symbols with a higher bandwidth efficiency. The receiver in MCSK has the ability to distinguish between the different data symbols based on the type of molecules as well as their respective concentration levels. The first instance of the implementation of MCSK for a diffusion-based molecular communication channel is presented in [KC13], where molecule isomers at respective concentration levels are used. In $M = U^{Z_{\text{mol}}}$-ary MCSK, M data symbols are formed with the help of U possible concentration levels and Z_{mol} molecular types. Therefore, each of $M = U^{Z_{\text{mol}}}$ symbols can be represented by $\log_2(M) = \log_2(U^{Z_{\text{mol}}})$ bits. For a simplified version of MCSK, which is considered in this work due to the availability of only two sprayer states, i.e., "on" and "off", each of $M = 2^{Z_{\text{mol}}}$ symbols can be represented by $\log_2(M) = \log_2(2^{Z_{\text{mol}}})$ bits, where $Z_{\text{mol}} \geq 2$. In this dissertation, the implementation of 4-MCSK is realized by two sprayers, one filled with a uranine water-based solution while the other filled with a highlighter ink water-based solution. The testbed setup for the implementation of 4-MCSK is similar to that of 2-MoSK, as shown in Figure 4.17. Thus, up to four different data transmission symbols can be formed by modulating each sprayer with OOK modulation scheme. Therefore, the symbol representing a bit combination of "00" is transmitted when both the sprayers are switched off, "01" is transmitted when only the sprayer containing the water-based solution of the highlighter ink is turned on, "10" is transmitted when only the sprayer containing

the water-based solution of uranine is turned on, and "11" is transmitted when both the sprayers are turned on. The spray duration of each sprayer is $T_{\text{on}} = 0.02\,\text{s}$, while the symbol duration is $T = 0.05\,\text{s}$. The illustration of the information signals with respect to the green and the red pixel intensity values recorded through spraying action of both the sprayers containing their respective water-based dye solutions is shown in Figure 4.20.

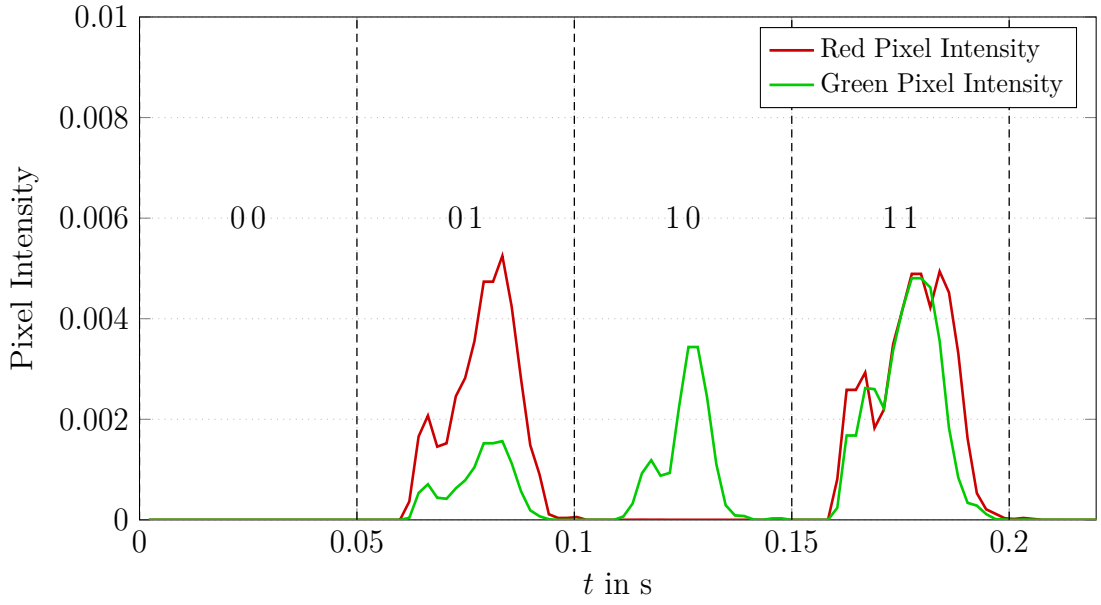


Figure 4.20: Received information signal representation of the transmitted data symbols in 4-ary molecular concentration shift keying, with the green as well as the red pixel intensity values while spraying the water-based solutions of both uranine as well as the highlighter ink, at distance $d = 88\,\text{cm}$.

As also mentioned in Section 4.4.5, the green pixel intensity value captured by the high-speed camera-based detector in the case of sprayed uranine water-based solution droplets is far more dominant than the red pixel intensity value, which is almost negligible. This is in contrast with the sprayed highlighter ink water-based solution droplets, where the maximum value of the red pixel intensity is almost three times more dominant than the maximum value of the green pixel intensity. Therefore, when the water-based solutions of either the highlighter ink or both the dyes are sprayed simultaneously, the presence of inter-color interference (ICI) is observed in the recorded green pixel intensity value, which affects the detection performance. To reduce the effect of ICI, the scaled red pixel intensity value is subtracted from the green pixel intensity value before detection, which is presented in Figure 4.21. Finally, the diagrammatic illustration of 4-MCSK in terms of the received pixel energy samples $y[k]$ for a binary data sequence "101101" after the removal of ICI is presented in Figure 4.22, with the dashed lines representing a symbol duration T.

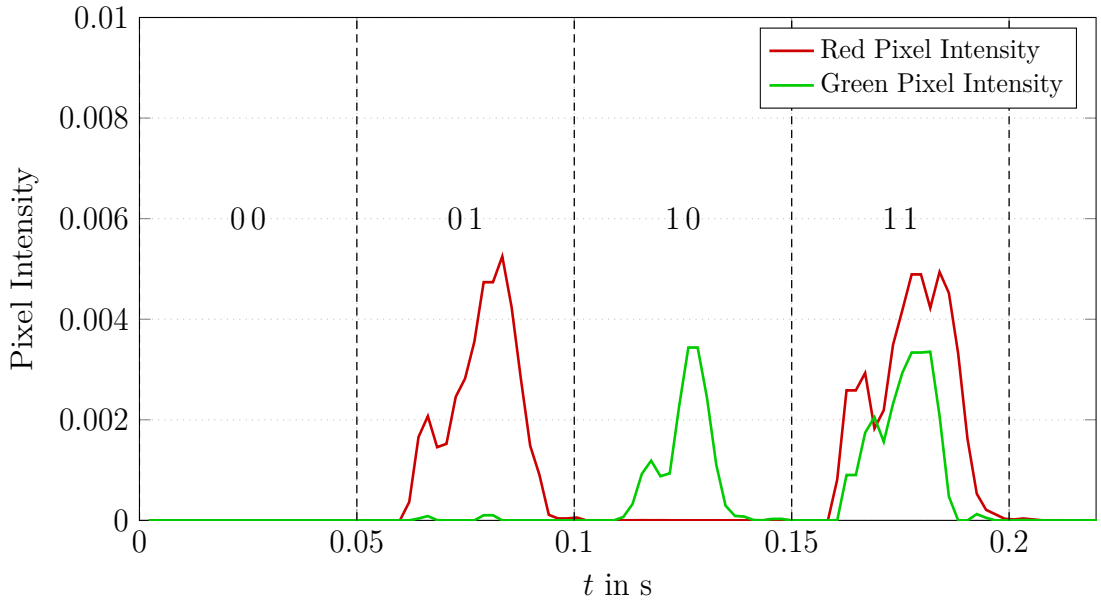


Figure 4.21: Received information signal representation of the transmitted data symbols in 4-ary molecular concentration shift keying, at distance $d = 88\,\text{cm}$ after removing inter-color interference.

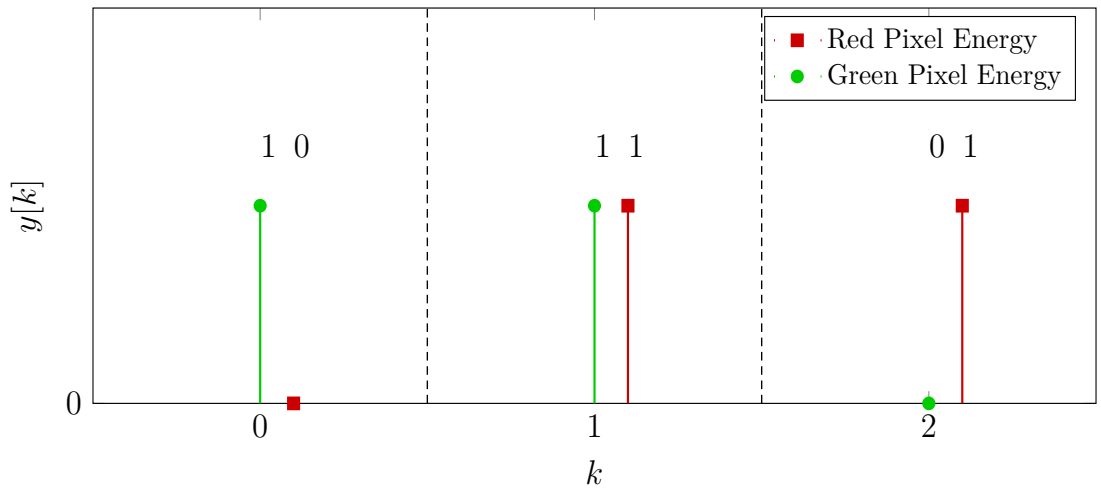
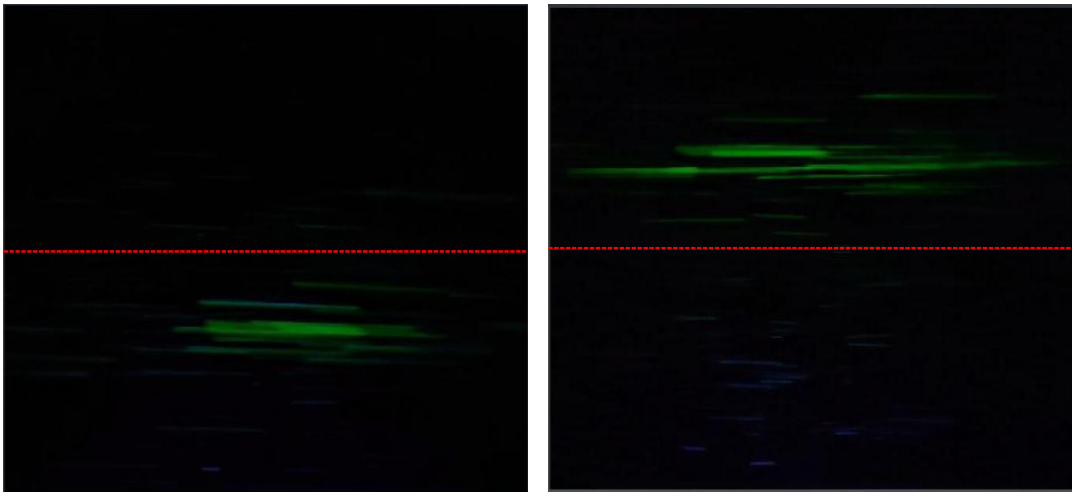


Figure 4.22: Depiction of the received pixel energy samples $y[k]$ for a 4-ary molecular concentration shift keying modulated data transmission sequence.

4.4.7 Spatial Shift Keying



(a) Binary data symbol “0” **(b)** Binary data symbol “1”

Figure 4.23: Camera snapshot of water-based solution droplets of uranine for the transmission of binary data symbols in 2-ary spatial shift keying, at a distance $d = 88\,\text{cm}$.

Another interesting modulation scheme which is similar to MoSK is spatial shift keying (SSK) that modulates the data depending on the presence of the transmitted information-carrying particles in a particular region of space. The exploitation of the spatial element serves as an alternative degree of freedom to concentration and type of released information-carrying particles. The idea is inspired from SSK modulation scheme for multiple-input multiple-output (MIMO) wireless communication links. SSK has been introduced for a diffusion-based molecular communication channel in [GB$^+$19b], where the transmitter antenna indices releasing the information-carrying molecules are used to encode the information, while an underwater prototype using a dilute hydrochloric acid solution is presented for the practical implementation of this particular scheme in [HW$^+$19b]. Therefore, for an M-ary SSK, M different transmit antennas can be used to release the information-carrying molecules for M different data symbols representing $\log_2 M$ bits. In this dissertation, based on the same principle, uranine water-based solution is released from two different sprayers at the transmitter side in order to represent different data symbols. Therefore, the 2-SSK is implemented by representing the binary data symbols based on the transmitter sprayer that is used to release the water-based dye solutions. The testbed setup is similar to the one illustrated in Figure 4.17. The transmitted information is detected at the receiver side by making a decision about the section of the field of view of the camera-based detector,

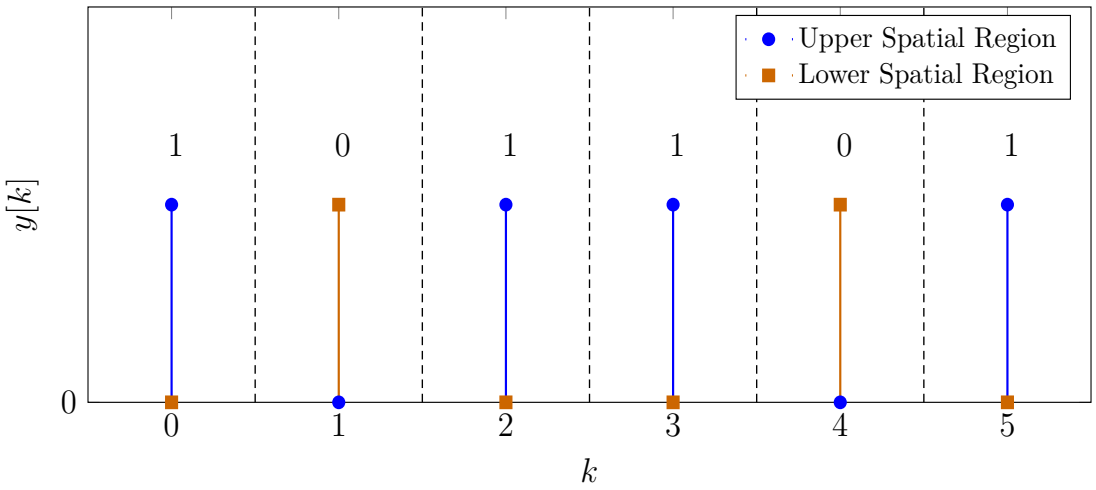


Figure 4.24: Depiction of the received pixel energy samples $y[k]$ for a 2-ary spatial shift keying modulated data transmission sequence.

where most of the received sprayed droplets are observed over a given transmission duration. For a binary data symbol "0", the solenoid valve of the sprayer is opened for $T_{\text{on}} = 0.02\,\text{s}$, for which the lower half region of the field of view of the camera-based detector receives most of the sprayed fluorescent droplets. Similarly, for a binary data symbol "1", the upper half region of the field of view of the camera-based detector receives most of the sprayed fluorescent droplets, which can be seen in Figure 4.23 for a distance of $d = 88\,\text{cm}$. The symbol duration is $T = 0.05\,\text{s}$. The pictorial depiction of 2-SSK in terms of the received pixel energy samples $y[k]$ for a binary data sequence "101101" is presented in Figure 4.24, with the dashed lines representing a symbol duration T.

4.4.8 Spatial Concentration Shift Keying

Similar to MCSK, spatial concentration shift keying (SCSK) combines together the degrees of freedom namely the concentration levels as well as the visibility of the molecules in a particular region of space in order to represent the transmitted data symbols with a higher bandwidth efficiency. The first implementation of SCSK modulation scheme for a diffusion-based molecular communication channel is presented as a spatial modulation-based molecular communication in [HW+19a], where the information is embedded into the index of a transmitter nanomachine and the concentration of the released molecules. This concept is very similar to the idea of spatial multiplexing in MIMO systems for molecular communication [KL+16]. In $M = U^{Z_{\text{space}}}$-ary SCSK, M data symbols are formed by exploiting U possible concentration levels and Z_{space} possible spatial regions. Therefore,

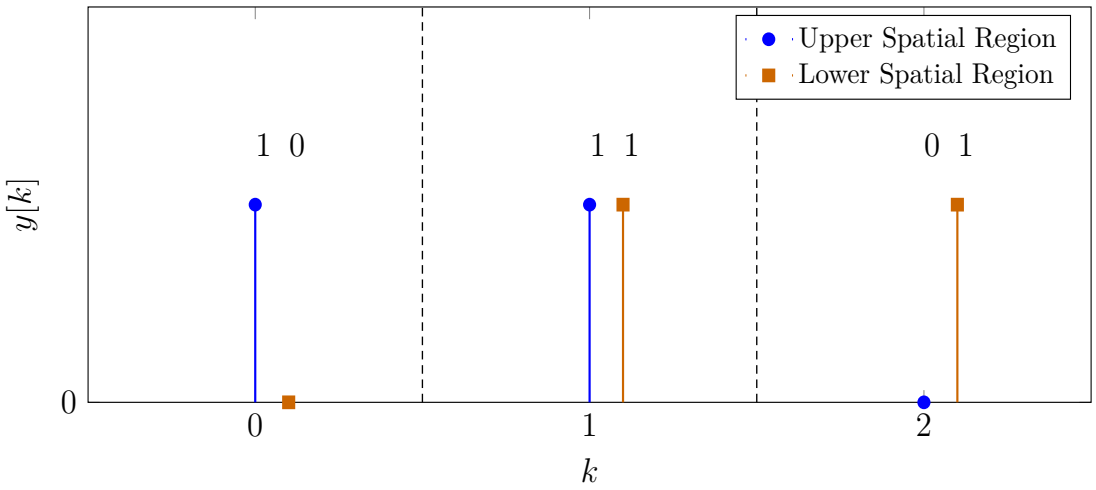


Figure 4.25: Depiction of the received pixel energy samples $y[k]$ for a 4-ary spatial concentration shift keying modulated data transmission sequence.

each of $M = U^{Z_{\text{space}}}$ symbols can be represented by $\log_2(M) = \log_2(U^{Z_{\text{space}}})$ bits. For a simplified version of SCSK, which is considered for this work due to the availability of only two sprayer states, i.e., "on" and "off", each of $M = 2^{Z_{\text{space}}}$ symbols can be represented by $\log_2(M) = \log_2(2^{Z_{\text{space}}})$ bits, where $Z \geq 2$. In this dissertation, the implementation of 4-SCSK is realized by two sprayers, both filled with water-based solutions of uranine. The testbed setup for the implementation of 4-SCSK is similar to that of 2-MoSK, as shown in Figure 4.17. Thus, up to four different data transmission symbols can be formed by modulating the sprayers according to OOK. Therefore, the data symbols are represented by the cases when either both the sprayers are switched off, or one of the two sprayers is switched on, or both the sprayers are switched on for a period of $T_{\text{on}} = 0.02\,\text{s}$ with a symbol duration $T = 0.05\,\text{s}$. This results in the sprayed droplets being not visible at all, or being visible in the lower half spatial region, or being visible in the upper half spatial region, or being visible in both the spatial regions of the field of view of the camera-based detector. The concept is very similar to the implementation of spatial multiplexing in a 2×2 MIMO system, where the two sprayers act as two transmit antennas while the upper and lower regions in the field of view of the camera-based detector act as two receive antennas. The pictorial representation of 4-SCSK in terms of the received pixel energy samples $y[k]$ for a binary data sequence "101101" is presented in Figure 4.25, with the dashed lines representing a symbol duration T.

4.4.9 Hardware Constrained Performance Comparison

Considering the hardware restrictions at the transmitter side of the testbed, as explained in Section 2.1, the minimum shutting time of the sprayer solenoid valve is $T_{\text{trans}} = 0.03\,\text{s}$, the spray duration of the sprayer is $T_{\text{on}} = 0.02\,\text{s}$, and the volume of the water-based solution of the information carrier released during T_{on} duration is $V_{\text{on}} = 0.18\,\text{mL}$. These restrictions reflect on the choice of parameters for the implemented modulation schemes such as the symbol duration T, the chip duration T_{c}, etc. Based on these hardware restrictions and the choice of parameters, the general expressions for the maximum achievable average data transmission rate R_b, and the average volume sprayed per bit V for the implemented modulation schemes can be derived, which are listed in Table 4.2. For instance, if V is not crucial, the design goal may be focused on maximizing R_b. On the other hand, V may be considered as a design criterion if resources are limited and R_b is not a pivotal parameter. Therefore, a performance comparison of the implemented modulation schemes based on the said parameters is essentially gauged by increasing the data symbol cardinality M. Figure 4.26 shows the variation of R_b with $\log_2 M$, and it is observed that for the modulation schemes using a single dye as the information carrier, a general trend is observed where R_b increases or remains the same up to a certain point, before gradually decreasing with $\log_2 M$. This is because the symbol duration T increases with M, and its effect is more dominant as compared to the number of data bits transmitted per symbol. This observation, however, is not seen in the modulation schemes like MoSK, MCSK, SSK, and SCSK, as the symbols are represented by additional degrees of freedom in terms of the color of fluorescent dyes and the spatial position.

Similarly, Figure 4.27 depicts the variation of the average volume sprayed per bit V with the data symbol cardinality M in terms of $\log_2 M$. Apart from SPPM, where V increases initially with increase in the value of $\log_2 M$ before decreasing quasi-exponentially, every other modulation scheme display a common characteristic, where V decreases quasi-exponentially with increase in $\log_2 M$. However, this observation is not seen in the case of MCSK and SCSK, as can also be inferred from their expressions of V which remains constant, as listed in Table 4.2.

Therefore, a more meaningful performance metric is the ratio between the maximum achievable average data transmission rate R_b and the average volume sprayed per bit V given by R_b/V, which takes both these design criteria into account. It is observed from Figure 4.28 that in the case of modulation schemes using a single dye as the information carrier, R_b/V increases up to a certain point before gradually decreasing with increasing symbol cardinality M, represented in terms of $\log_2 M$. This is because with increase in

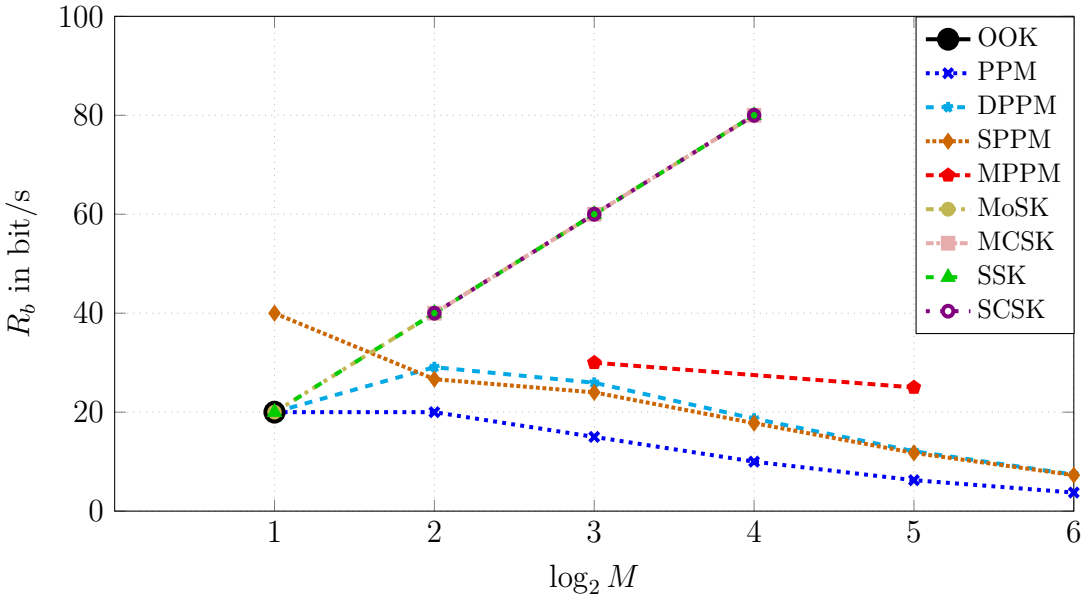


Figure 4.26: Variation of the maximum achievable average data transmission rate R_b with data symbol cardinality M, represented in terms of $\log_2 M$.

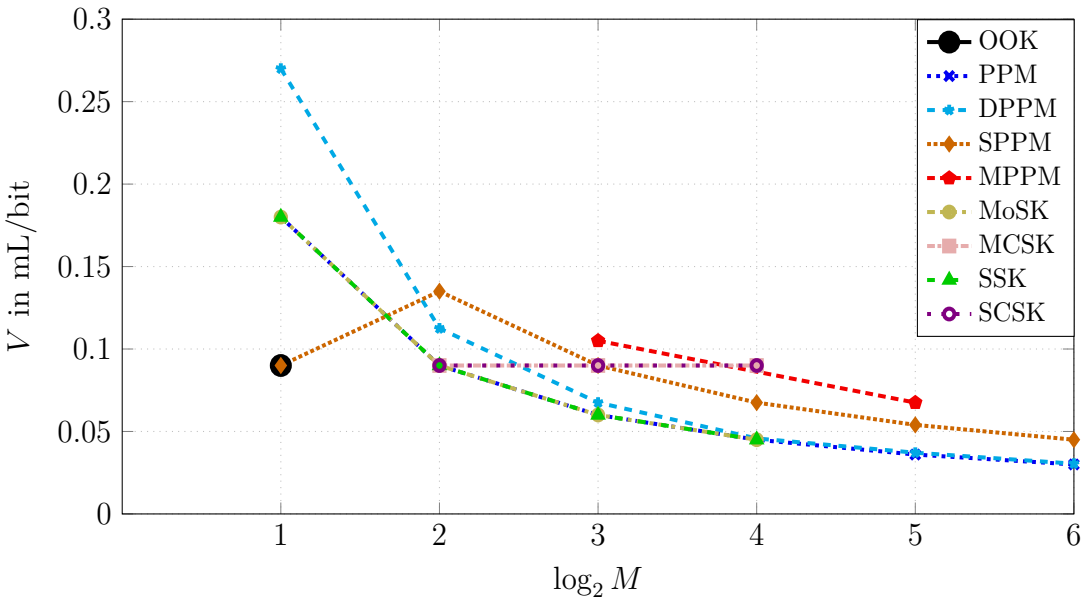


Figure 4.27: Variation of the average volume sprayed per bit V with data symbol cardinality M, represented in terms of $\log_2 M$.

Table 4.2: The hardware constrained maximum achievable average data transmission rate R_b and the average volume sprayed per bit V with respect to the data symbol cardinality M. The general expressions are presented for the implemented modulation schemes in this dissertation.

Modulation	R_b	V
OOK	$\frac{1}{T}$	$\frac{V_{\text{on}}}{2}$
PPM	$\frac{\log_2 M}{MT_{\text{c}}}$	$\frac{V_{\text{on}}}{\log_2 M}$
DPPM	$\frac{2M \log_2 M}{M(M-1)T_{\text{gap}} + 2(M+1)T_{\text{pulse}}}$	$\frac{V_{\text{on}}}{\log_2 M} + \frac{V_{\text{on}}}{M \log_2 M}$
SPPM	$\frac{\log_2 M}{\left(2^{\log_2 M-1} + \text{mod}\left(2^{\log_2 M-1}, 2\right) + \left(\sqrt{-1}\right)^M\right) T_{\text{c}}}$	$\frac{\left(\frac{M}{2} + 2\left(\left(2^{\log_2 M-1} + \text{mod}\left(2^{\log_2 M-1}, 2\right) + \left(\sqrt{-1}\right)^M\right) - 1\right)\right) V_{\text{on}}}{M \log_2 M}$
MPPM	$\frac{\log_2 \left(\binom{M/p}{2} + p\right)}{MT_{\text{c}}/p}$, where $p = 2^{\log_4(M/2)}$	$\frac{\left(2\binom{M/p}{2} + p\right) V_{\text{on}}}{M \log_2 M}$, where $p = 2^{\log_4(M/2)}$
MoSK	$\frac{\log_2 M}{T}$	$\frac{V_{\text{on}}}{\log_2 M}$
MCSK	$\frac{\log_2 M}{T}$	$\frac{V_{\text{on}}}{2}$
SSK	$\frac{\log_2 M}{T}$	$\frac{V_{\text{on}}}{\log_2 M}$
SCSK	$\frac{\log_2 M}{T}$	$\frac{V_{\text{on}}}{2}$

M, the influence of the parameter V dominates over the influence of the parameter R_b. However, this observation is not seen in the case of modulation schemes using additional degrees of freedom to represent the information like MoSK, MCSK, SSK and SCSK, as the influence of parameter R_b is more dominant than the influence of parameter V.

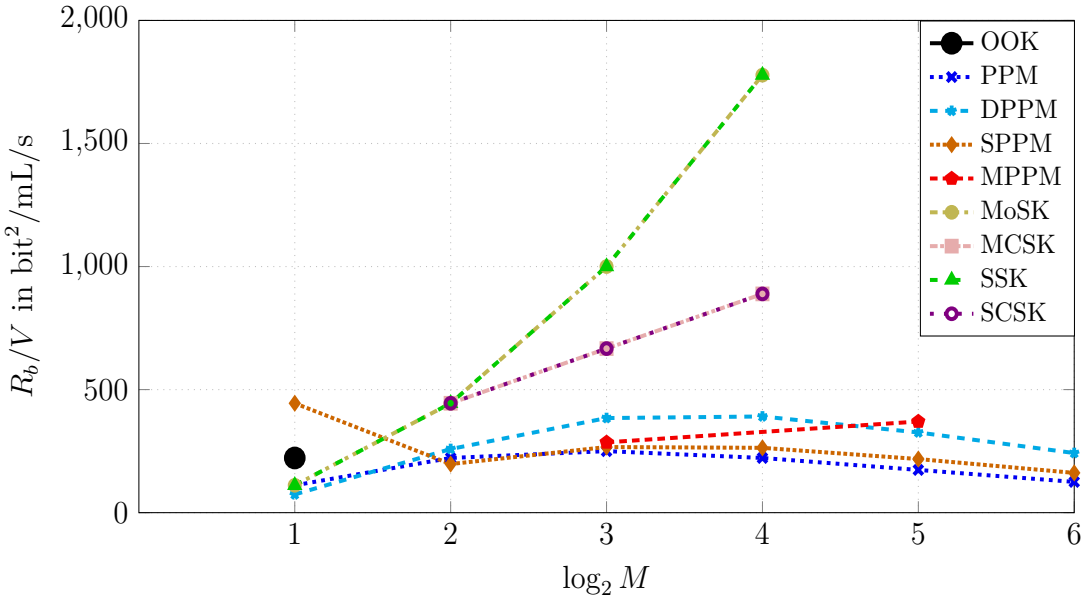


Figure 4.28: Variation of ratio of the maximum achievable average data transmission rate to the average volume sprayed per bit R_b/V with data symbol cardinality M, represented in terms of $\log_2 M$.

4.5 Results

The results of the BER analysis with increasing transmission distance d for the implemented modulation schemes are discussed in this section. The BER performance based on the testbed measurements using water-based solution of uranine for a randomly generated binary data sequence of length 108 over 20 trials is presented in Figure 4.29 (testbed measurement parameters listed in Table 3.1). For a better estimate of the BER performance for smaller values of d, the simulation results based on the analytical model presented in Figure 4.30 is considered for a randomly generated binary data sequence of the same length over 100000 trials. As the detection technique, threshold detection is used for the implemented modulation schemes, as the emphasis has particularly been laid on implementing low-complexity detection methods. These detection algorithms can also be used for applications concerning micro-scale molecular communication systems, which have limited

access to computational power. Therefore, in the case of OOK, 4-MCSK, 4-SCSK and 8-MPPM, FTD detection technique is employed. On the other hand, ATD proves to be more efficient as a detection strategy in the case of 2-PPM, 2-DPPM, 2-MoSK, 2-SSK, 4-PPM, 4-DPPM, 8-PPM, and 8-DPPM. Modulation schemes such as 4-SPPM and 8-SPPM use a combination of FTD and ATD in their detection processes. As a common observation, it is seen that the BER performance deteriorates with increase in d. This is because there is a significant reduction in the number of droplets traveling the length of the tube and reaching the field of view of the camera-based detector at the receiver side. The droplets either fall down and collect at the bottom of the tube or simply get dispersed laterally. Also, the trajectory assumed by the spray is influenced by the gravitational force, increasing the distance between the camera-based detector and the droplets, as discussed in Section 3.3.2, thus reducing the intensity of light received by the camera. All these influences decrease the received signal strength while contributing to the noise power. Moreover, the slowing down of the sprayed droplets due to the effect of air drag further leads to an increase in ISI.

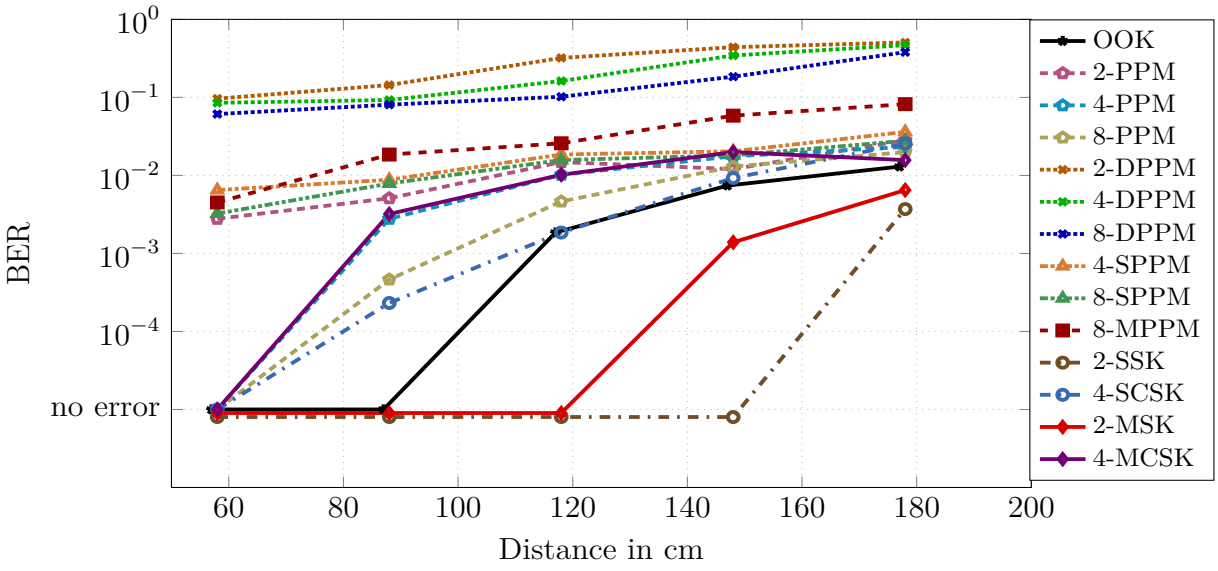


Figure 4.29: Bit error rate performance of the modulation schemes under investigation based on the testbed measurements with increasing distance d.

The BER performances of the implemented higher-order modulation schemes using various degrees of freedom are compared with respect to OOK modulated received energy samples, recorded by the camera-based detector. As seen from Figure 4.29, DPPM displays the worst BER performance amongst all the implemented modulation schemes. This is because in the case of DPPM, the time gap between two successive pulses determines the

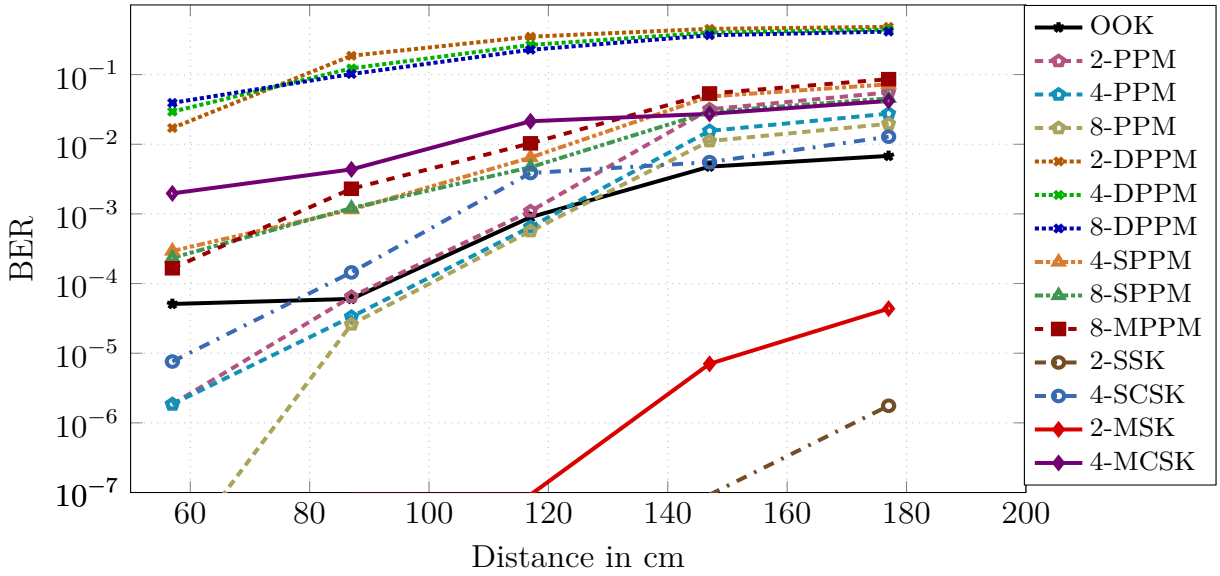


Figure 4.30: Bit error rate performance of the modulation schemes under investigation based on the analytical model with increasing distance d.

transmitted data symbol. However, the detection of a successive pulse at an undesirable time slot leads to an error, and from there onwards, every subsequent pulse detected most-likely falls in an undesirable time slot leading to a severe error propagation. This observation is also corroborated by the simulation results based on the analytical model depicted in Figure 4.30. However, with increase in symbol cardinality M, the BER performance does improve as the time gap between the successive pulses corresponding to the transmitted data symbol increases, as described in Section 4.4.2. This is because with increase in the time gap, the effect of ISI is reduced, thus reducing the chances of detection of the pulses in an incorrect time slot. In the case of PPM and as observed from the testbed measurement results in Figure 4.29, BER performance is worse than that of OOK. This is due to the fact that the fixed chip duration T_c is small, which leads to an increased ISI in a chip slot within a symbol duration T for PPM with lower values of M. As M increases, $T = M\,T_c$ also increases, thus reducing the effects of ISI. This in turn improves the BER performance. However, in the case of the simulation results based on the analytical model shown in Figure 4.30, a more accurate result shows that PPM performs better than OOK for smaller values of d. The reason behind such observation is the fact that PPM employs ATD detection strategy, which performs better than FTD used in the detection of OOK modulated data symbols by efficiently mitigating the ISI. From Figure 4.29 as well as Figure 4.30, it is evident that the modulation schemes SPPM and MPPM perform better than DPPM, but worse than PPM since these schemes employ a multi-pulse system. In

these modulation schemes, more than one pulse is transmitted over multiple chip slots within T, unlike PPM where only a single pulse per T is transmitted. This makes it more difficult for the detector to distinguish the transmitted symbols based on the position of the pulses in the chip slots within T, thus increasing the BER. Again, with increase in the value of M, the BER performance improves as the overall T increases with T_c remaining fixed. This is similar to the observation in the case of PPM. In the case of modulation schemes using more than one degrees of freedom like 4-MCSK displays a better BER performance than most other modulation schemes but performs worse than OOK. This is observed from the plots based on both the testbed measurement results as well as the simulation results based on the analytical model in Figure 4.29 and Figure 4.30, respectively. This is due to the effect of shadowing that takes place amongst the water-based solution droplets of the sprayed information carriers. The droplets of one color gets hidden by the droplets of the other color due to an increase in lateral dispersion of the sprayed droplets with increase in distance d between the transmitter and the receiver. Moreover, in 4-MCSK, the detection of the red pixel energy samples at the receiver side using FTD contributes to the error more than the detection of the green pixel energy samples due to the reasons stated in Section 4.2.3. However, the BER performance of 4-SCSK is better than that of 4-MCSK. This is due to the fact that although the effect of shadowing is still present in the case of 4-SCSK, its ill effects are not very pronounced as only a single color water-based fluorescent dye solution in the form of uranine is used. It is also to be mentioned that the effect of shadowing in the analytical model has been treated as a noise process. The BER performances of 2-MoSK and 2-SSK are the best in comparison to all the implemented modulation schemes as obtained from the results based on the testbed measurements as well as the simulation results based on the analytical model, showcased in Figure 4.29 and Figure 4.30, respectively. This is because the information decoded by the camera-based detector is assumed to be obtained from two reliable independent sources in the form of two separate colors of the sprayed water-based dye solutions or two separate spatial regions, where the sprayed water-based dye solution droplets are present. The dye colors or the spatial regions represent the same transmitted binary data symbol in a complementary fashion. For example, in the case of 2-MoSK, binary data symbol “0” is represented by spraying the highlighter ink water-based solution, while at the same time, not spraying the uranine water-based solution, and vice versa when representing binary data symbol “1”. Similarly, in the case of 2-SSK, binary data symbol “0” is represented by the sprayed water-based solution of uranine, whose droplets populate the lower half region of the field of view of the camera-based detector, while at the same time, not populating the upper half

region of the field of view of the camera-based detector, and vice versa when representing binary data symbol "1". This improves the confidence of the detection algorithm as the information is obtained from two independent sources representing the same event, resulting in improved BER performance. The BER performance of 2-MoSK is worse than that of 2-SSK. This is due to the presence of the Bayer filter in the optical sensors of the camera-based detector, which is less sensitive to the fluorescence producing wavelength in the case of the highlighter ink water-based solution as opposed to the fluorescence producing wavelength in the case of the uranine water-based solution (as also mentioned in Section 4.2.3). A further improvement in the BER performance of all the modulation schemes can also be achieved by combining the channel equalization methods presented in Section 4.3.

4.6 Summary

This chapter discusses the data acquisition principle that is employed at the receiver side of the presented macroscopic air-based molecular communication testbed using fluorescent dyes. The green pixel intensity value is preferred as the information-carrying signal, when water-based solution of uranine is used. Similarly, the red pixel intensity value is preferred as the information-carrying signal when water-based solution of the highlighter ink is used. The energy under the received green and/or red pixel intensity pulses are then accumulated over a symbol duration T, before undergoing the detection process. This process is represented with the help of an equivalent discrete-time end-to-end system model.

Simple binary transmission algorithms like the OOK modulation scheme using only a single degree of freedom in the form of concentration is presented in this chapter. Similarly, low-complexity threshold detection strategies like FTD and ATD, which can also be utilized in those applications that have limited access to computational power are additionally discussed. The error performance of the transmission link incorporating OOK at the transmitter side and FTD at the receiver side is investigated by separately using water-based solutions of both uranine as well as the highlighter ink. The theoretical BERs of both the information carriers are evaluated, and are compared to their respective BER performances based on the testbed measurements and the simulation results using the analytical model. The investigations are carried out for an increasing distance d between the transmitter and the receiver. It is observed that for larger values of d, the BER performance of the uranine water-based solution as the information carrier outperforms that of the highlighter ink water-based solution due to the optical sensors of the camera-based detector being more sensitive to the fluorescent wavelength produced by the former against the latter.

Additionally, the problem of ISI in the presented air-based molecular communication testbed with increasing d is tackled by the use of popular equalization techniques like ZF, DFE and MLSE. It is observed that the MLSE-based equalizer outperforms all other equalization methods due to its high computational complexity.

Finally, higher-order modulation schemes are implemented in this chapter in order to transmit and receive higher volumes of data over the transmission channel of the presented molecular communication testbed. For this purpose, the information is encoded utilizing several degrees of freedom in the form of concentration, time, color, and the spatial region populated by the sprayed droplets of the water-based solution of the information carriers. Modulation schemes like PPM, DPPM, SPPM, MPPM, MoSK, MCSK, SSK, and SCSK have been implemented and investigated in this chapter. A hardware constrained performance comparison for the implemented modulation schemes have also been analyzed in order to give a perspective to the user when choosing a particular modulation scheme, and the parameters like the maximum achievable average data transmission rates and the average resource requirements need to be optimized. Additionally, BER analysis of the implemented modulation schemes have also been included in this chapter with respect to the results obtained from the testbed measurements as well as the simulation results obtained from the analytical model. It is observed that MoSK and SSK perform the best amongst all the implemented modulation schemes, when a binary data symbol is represented by two independent sources complementing one another. Modulation schemes like 4-MCSK and 4-SCSK display poor BER performance as compared to OOK, however their data transmission rates are double that of OOK (40 bit/s for 4-MCSK and 4-SCSK as compared to 20 bit/s for OOK). The worst performing modulation schemes are 2-DPPM, 4-DPPM, and 8-DPPM as they suffer from the problems of error propagation. Finally, 2-PPM, 4-PPM, 8-PPM, 4-SPPM, 8-SPPM, and 8-MPPM modulation schemes perform worse than OOK in terms of the BER performance due to the dominant effect of ISI.

5

Channel Coding Schemes

This chapter deals with the implementation of the channel coding schemes in the macroscopic air-based molecular communication testbed using fluorescent dyes, with the emphasis being on the binary data symbol coding algorithms. The incorporation of the channel codes ensures robustness against the data transmission errors over large distances, which may arise due to the effects of inter-symbol interference (ISI) and noise, specified in Chapter 3. In this dissertation, analysis of various forward error correction (FEC) codes have been carried out that do not require a feedback from the channel decoder to the channel encoder. FEC coding techniques are broadly classified under two larger coding families such as the block codes and the convolutional codes. As alternatives to FEC codes, a type of line code and spatial codes have also been investigated. The inclusion of the channel coding schemes in the molecular communication testbed results in an improved bit error rate (BER) performance due to their error detection, correction, and avoidance (in the case of line codes) capabilities compared to the uncoded transmission scheme. Simultaneously, resource requirements in order to achieve an improved BER performance has also been studied in this chapter with respect to the implemented FEC and line codes.

The evaluation of the channel coding schemes is performed through BER analysis by conducting the testbed measurements as well as the simulations based on the analytical model, described in Section 3.7. Therefore, an uncoded on-off keying (OOK) modulated transmission signal is considered as the performance benchmark, while fixed threshold detector (FTD) is used for the detection of the energy samples of the received signal, as described in Section 4.2.2. Similarly, after the application of channel coding, the energy contained within the pulses of the coded data is evaluated through integration over time for determining the threshold before performing FTD-based detection. It is also to be

mentioned that in this dissertation, the focus is on the application of the unnormalized case of channel codes, i.e., the symbol duration T and the spray duration T_{on} remain the same as in the case of the uncoded scheme. This is because this work aims at achieving long distance transmission with minimum transmission errors, which is ensured in the case of unnormalized channel coding scenario. The normalization of the transmission rates of the coded data to that of the uncoded data ensures that both the coded and the uncoded data are transmitted and received at the same time. Similarly, when normalization of the resource requirements of the coded data to that of the uncoded data is performed, both the scenarios utilize exactly the same amount of resources for the transmission process. However, both these advantages for the coded scenario come at a cost of reduced BER performance as compared to the uncoded scheme. This observation is corroborated by the detailed studies presented in [Dam20].

5.1 Block Codes

Block codes, as the name suggests, use the strategy of blockwise encoding and decoding [LC01]. The information sequence is divided into blocks called the info word $\boldsymbol{u}_{\text{c}} = [u_0 \dots u_{k_{\text{c}}-1}]$ of length $k_{\text{c}} \in \mathbb{N}$. Each info word is then individually encoded into a code word $\boldsymbol{x}_{\text{c}} = [x_0 \dots x_{n_{\text{c}}-1}]$ of length $n_{\text{c}} \in \mathbb{N}$, where $n_{\text{c}} > k_{\text{c}}$. Thus, a bijective mapping exists between $2^{k_{\text{c}}}$ info words and $2^{k_{\text{c}}}$ code words out of a possible $2^{n_{\text{c}}}$ code words. This redundancy offered by $n_{\text{c}} - k_{\text{c}}$ bits helps in the detection and/or correction of the transmission errors. The block code is said to have a code rate $R = k_{\text{c}}/n_{\text{c}}$, that determines the gain incurred over the uncoded scheme in terms of the BER performance for a target distance between the transmitter and the receiver or vice versa. Additionally, the error detection and correction capabilities of a block code can be determined from a metric called the minimum Hamming distance d_{min}, which describes the minimum number of positions at which two code words differ from one another. Hence, with hard-input decoding, a block code guarantees the detection of up to $d_{\text{min}} - 1$ erroneous bits, and a correction of up to $\lfloor (d_{\text{min}} - 1)/2 \rfloor$ erroneous bits. Furthermore, a block code is considered to be systematic if the info word is found embedded within its code word with the remaining bits $\left[p_0 \dots p_{n_{\text{c}}-k_{\text{c}}-1}\right]$ as its parity bits. Additionally, a linear block code is one, where a linear combination of two valid code words leads to another valid code word. A linear block code encoding is made possible with a $k_{\text{c}} \times n_{\text{c}}$ generator matrix $\boldsymbol{G}$, and the decoding is performed by an $(n_{\text{c}} - k_{\text{c}}) \times n_{\text{c}}$ parity check matrix $\boldsymbol{P}$ provided the condition $\boldsymbol{G}\boldsymbol{P}^{\mathsf{T}} = \boldsymbol{0}$ is satisfied, where $\boldsymbol{0}$ is the zero vector. A generator matrix and a parity check matrix are inter-convertible,

and the following relation between them for systematic codes applies:

$$\boldsymbol{G} = \begin{bmatrix} \boldsymbol{I}_{k_\mathrm{c}} & \boldsymbol{P}_\mathrm{c} \end{bmatrix}, \quad \boldsymbol{P} = \begin{bmatrix} -\boldsymbol{P}_\mathrm{c}^\mathsf{T} & \boldsymbol{I}_{n_\mathrm{c}-k_\mathrm{c}} \end{bmatrix}, \tag{5.1}$$

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 with parity check equations. Furthermore, some block codes can also be represented using the generator polynomial $g(z)$. The valid code words are designed in a way that the polynomial representing the code is completely divisible by the generator polynomial. In this section, the following gives an overview of the block codes considered and implemented in this dissertation.

5.1.1 Repetition Code

Repetition coding is arguably the simplest channel coding technique, where a code word is formed by repeating the bits present in the info word n_c times. Repetition code is an $(n_\mathrm{c}, k_\mathrm{c} = 1)$ linear systematic block code, with a code rate of $R = 1/n_\mathrm{c}$. Its generator matrix $\boldsymbol{G}$ and parity check matrix $\boldsymbol{P}$ can be expressed as

$$\boldsymbol{G} = \begin{bmatrix} 1 & \dots & 1 \end{bmatrix}, \quad \boldsymbol{P} = \begin{bmatrix} 1 & 1 & 0 & \dots & 0 \\ 1 & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 1 & 0 & \dots & 0 & 1 \end{bmatrix}. \tag{5.2}$$

The repetition code has the minimum Hamming distance $d_\mathrm{min} = n_\mathrm{c}$. Owing to its low-complexity implementation, the first instance where repetition code is employed in a diffusion-based molecular communication is presented in [SL+12].

In this particular work, a $(2, 1)$ repetition code and a $(3, 1)$ repetition code are investigated, and their info word to code word mapping tables are given in Table 5.1 and Table 5.2, respectively. A hard-input decoding guarantees the detection of a single bit error in the case of the $(2, 1)$ repetition code, while in the case of the $(3, 1)$ repetition code, up to two bit errors can be detected and a single bit error can be corrected for sure. However, the performance of soft-input decoding is better than hard-input decoding, and therefore, has been considered for this work. The decoding of repetition code is performed on the energy samples of the received signal using FTD. Therefore, the energy samples of every code bit corresponding to an info bit is added up, and is used for calculating the FTD threshold.

Table 5.1: Mapping table of the $(2, 1)$ repetition code.

$\boldsymbol{u}_\mathrm{c} = [u_0]$	$\boldsymbol{x}_\mathrm{c} = [x_0\, x_1]$
0	[0 0]
1	[1 1]

Table 5.2: Mapping table of the $(3, 1)$ repetition code.

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

5.1.2 Single Parity Check Code

Single parity check (SPC) code is a systematic block code, where the code word is produced by appending an additional parity bit to the info word. SPC codes can be represented in the form of $(n_\mathrm{c} = k_\mathrm{c} + 1, k_\mathrm{c})$ block code, with $R = k_\mathrm{c}/(k_\mathrm{c} + 1)$. The redundancy is added with the help of a parity bit, which is selected in a way that the code bits in a code word have an even or an odd parity. For even parity SPC code, which also happens to be a linear block code, the parity equation must be satisfied, which is of the form

$$x_0 \oplus x_1 \oplus \cdots \oplus x_{n_\mathrm{c}-1} = 0, \tag{5.3}$$

where $\oplus$ is a modulo two addition. In this dissertation, only SPC code with an even parity is considered. The resulting generator matrix $\boldsymbol{G}$, and parity check matrix $\boldsymbol{P}$ for an SPC code are

$$\boldsymbol{G} = \begin{bmatrix} 1 & 0 & \dots & 0 & 1 \\ 0 & 1 & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & 1 \\ 0 & \dots & 0 & 1 & 1 \end{bmatrix}, \quad \boldsymbol{P} = \begin{bmatrix} 1 & \dots & 1 \end{bmatrix}. \tag{5.4}$$

From the standpoint of diffusion-based molecular communication channel, SPC code is explored in [MP+18] through design of genetic circuits.

In this work, an even parity $(3, 2)$ SPC code is implemented having a minimum Hamming distance $d_\mathrm{min} = 2$, with the info word to code word mapping shown in Table 5.3, making it possible to detect a single bit error with hard-input decoding. However, soft-input decoding makes it possible to even correct the error, and thus, has been considered in this dissertation. The decoding of $(3, 2)$ SPC code is performed on the energy samples of the received signal using FTD. In the case of an erroneous code word, where the parity equation is not satisfied,

the code bit with the smallest distance between its energy sample and the threshold is inverted.

Table 5.3: Mapping table of $(3, 2)$ single parity check code.

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

5.1.3 Cyclic Redundancy Check Code

Cyclic redundancy check (CRC) code is a systematic cyclic block code best suited for detecting and correcting burst errors in various data transmission standards such as Ethernet, Bluetooth, etc. [Wic95]. An $(n_\mathrm{c}, k_\mathrm{c})$ CRC code word with $R = k_\mathrm{c}/n_\mathrm{c}$ is formed by extending the info word by appending $n_\mathrm{c} - k_\mathrm{c}$ bits of the check value. The check value is obtained through the remainder of the division of the info word in its polynomial form by a generator polynomial of the order $n_\mathrm{c} - k_\mathrm{c}$. The error is detected if the division of the received code word with the generator polynomial again produces a remainder [BG$^+$92]. From the context of molecular communication, CRC codes have been investigated and used in works related to error detection strategies in bacterial quorum sensing in [Bai16], and for error detection and correction applications in diffusion-based micro-scale channel models in [Dam20].

In this work, $(8, 4)$ and $(16, 8)$ CRC codes are implemented using polynomials $g(z) = z^4 + z^3 + z^2 + z + 1$ [Wic95], and $g(z) = z^8 + z^7 + z^6 + z^4 + z^2 + 1$ [DVBS2], respectively. $(8, 4)$ CRC code has a minimum Hamming distance $d_\mathrm{min} = 2$, while $(16, 8)$ CRC code has $d_\mathrm{min} = 4$. As an example, the mapping between the info words and the code words for $(8, 4)$ CRC code is presented in Table 5.4. For detection purposes, soft-input decoding is performed on the energy samples of the received signal using FTD. For each code bit, the distance between the energy sample and the threshold is calculated, which is interpreted as a soft value. The distances are then correlated with the code word hypotheses, and the one with the highest correlation is selected, where the info word is obtained from the systematic part of the code word.

Table 5.4: Mapping table of $(8, 4)$ cyclic redundancy check code.

$\boldsymbol{u}_\mathrm{c} = [u_0\, u_1\, u_2\, u_3]$	$\boldsymbol{x}_\mathrm{c} = [x_0\, x_1\, x_2\, x_3\, x_4\, x_5\, x_6\, x_7]$
[0 0 0 0]	[0 0 0 0 0 0 0 0]
[0 0 0 1]	[0 0 0 1 1 1 1 1]
[0 0 1 0]	[0 0 1 0 0 0 0 1]
[0 0 1 1]	[0 0 1 1 1 1 1 0]
[0 1 0 0]	[0 1 0 0 0 0 1 0]
[0 1 0 1]	[0 1 0 1 1 1 0 1]
[0 1 1 0]	[0 1 1 0 0 0 1 1]
[0 1 1 1]	[0 1 1 1 1 1 0 0]
[1 0 0 0]	[1 0 0 0 0 1 0 0]
[1 0 0 1]	[1 0 0 1 1 0 1 1]
[1 0 1 0]	[1 0 1 0 0 1 0 1]
[1 0 1 1]	[1 0 1 1 1 0 1 0]
[1 1 0 0]	[1 1 0 0 0 1 1 0]
[1 1 0 1]	[1 1 0 1 1 0 0 1]
[1 1 1 0]	[1 1 1 0 0 1 1 1]
[1 1 1 1]	[1 1 1 1 1 0 0 0]

5.1.4 Hamming Code

The Hamming code is a linear systematic block code with a minimum Hamming distance value fixed at $d_\mathrm{min} = 3$ [Ham50]. This guarantees the detection of up to two erroneous bits and a correction of a single bit error using hard-input decoding. In Hamming codes, the relationship between the number of parity bits, the info word length, and the code word length is fixed as d_min remains constant for all the code rates. An $(n_\mathrm{c}, k_\mathrm{c})$ Hamming code consists of $n_\mathrm{c} - k_\mathrm{c}$ parity bits with code word length $n_\mathrm{c} = 2^{n_\mathrm{c}-k_\mathrm{c}} - 1$, and info word length $k_\mathrm{c} = n_\mathrm{c} - \log_2(n_\mathrm{c} + 1)$, where $\{n_\mathrm{c}, k_\mathrm{c}, n_\mathrm{c} - k_\mathrm{c}\} \in \mathbb{N}$ must be satisfied. In this work, the parity bits of a Hamming code are placed at positions which are in powers of two. As an example, the generator matrix $\boldsymbol{G}$ and parity check matrix $\boldsymbol{P}$ for $(7, 4)$ Hamming code are

$$\boldsymbol{G} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}, \quad \boldsymbol{P} = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}. \tag{5.5}$$

Hamming coding has also been extensively studied from the perspective of molecular communication in the past for diffusion-based channels, as presented in works [LH+15a; Bai16; Dam20].

In this dissertation, $(7, 4)$ Hamming code and $(15, 11)$ Hamming code are implemented. As example, the info word to code word mapping table for $(7, 4)$ Hamming code is shown in Table 5.5. Although, with hard-input decoding, the Hamming code can detect up to two bit errors and correct up to a single bit error for sure, soft-input decoding gives better error rate performance, and therefore is used in this work. Soft-input decoding is performed on the energy samples of the received signal using FTD. For every code bit, the distance between the energy sample and the threshold is used as a soft value. For decoding, the distances are correlated with the code word hypotheses. The decoded info word is then obtained from the systematic part of the code word hypothesis that gives the highest correlation, a decoding strategy that is very similar to the one employed for CRC codes, described in Section 5.1.3.

Table 5.5: Mapping table of $(7, 4)$ Hamming code.

$\boldsymbol{u}_\mathrm{c} = [u_0\, u_1\, u_2\, u_3]$	$\boldsymbol{x}_\mathrm{c} = [x_0\, x_1\, x_2\, x_3\, x_4\, x_5\, x_6]$
[0 0 0 0]	[0 0 0 0 0 0 0]
[0 0 0 1]	[1 0 1 0 0 0 1]
[0 0 1 0]	[1 1 1 0 0 1 0]
[0 0 1 1]	[0 1 0 0 0 1 1]
[0 1 0 0]	[0 1 1 0 1 0 0]
[0 1 0 1]	[1 1 0 0 1 0 1]
[0 1 1 0]	[1 0 0 0 1 1 0]
[0 1 1 1]	[0 0 1 0 1 1 1]
[1 0 0 0]	[1 1 0 1 0 0 0]
[1 0 0 1]	[0 1 1 1 0 0 1]
[1 0 1 0]	[0 0 1 1 0 1 0]
[1 0 1 1]	[1 0 0 1 0 1 1]
[1 1 0 0]	[1 0 1 1 1 0 0]
[1 1 0 1]	[0 0 0 1 1 0 1]
[1 1 1 0]	[0 1 0 1 1 1 0]
[1 1 1 1]	[1 1 1 1 1 1 1]

5.2 Convolutional Codes

Convolutional codes are another broad classification of channel codes, which tackle the errors originating in the transmission channel by making use of the error detection and correction strategies through addition of redundant bits. Contrary to block codes, where encoding is performed on discrete block of bits, convolutional codes are constantly encoding

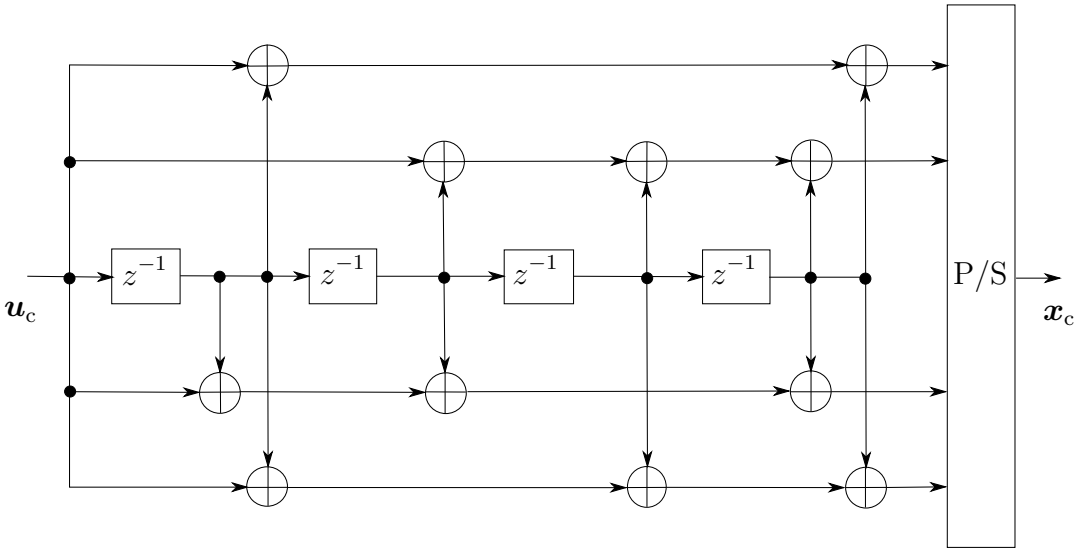


Figure 5.1: $(4, 1, 4)$ convolutional code represented by shift registers and modulo two additions.

the continuous stream of info bit sequence, which are not partitioned into discrete info word blocks [LC01]. A convolutional encoder representation is shown in Figure 5.1, which generally constitutes of shift registers and modulo two additions. Unlike the block codes in their systematic form, where the parity bits are appended to the info bits to form the code word, a convolutional code uses the same info bit with the introduction of memory, and then spreads it over several parts of the code sequence in order to add redundancy. A convolutional code can also be described by its generator polynomials. $(4, 1, 4)$ convolutional code presented in Figure 5.1 has the following generator polynomials [Tüc04; LC01]:

$$\begin{aligned} g_0(z) &= 1 + z + z^4, \\ g_1(z) &= 1 + z^2 + z^3 + z^4, \\ g_2(z) &= 1 + z + z^2 + z^4, \\ g_3(z) &= 1 + z + z^3 + z^4. \end{aligned} \tag{5.6}$$

The presented convolutional code is a non-recursive type, as it does not have any internal feedback path. A typical $(n_\mathrm{c}, k_\mathrm{c}, L_\mathrm{c})$ convolutional code encodes k_c info bits to produce n_c code bits with a memory length of L_c (represented by the shift registers), and a code rate of $R = k_\mathrm{c}/n_\mathrm{c}$. The $(4, 1, 4)$ convolutional code in Figure 5.1 has $R = 1/4$.

The construction of convolutional codes is more complex as compared to block codes, and therefore, strong convolutional codes need to be obtained by performing an exhaustive

search. Convolutional codes with high code rate R can be designed with a process called code puncturing [Tüc04]. In this technique, some code bits are omitted in a periodic fashion and are not transmitted at all, which the decoder interprets as erasures. For decreasing the value of R, the generator polynomials can be re-used. Thus, convolutional codes of varied R could be generated by a combination of puncturing and re-use of polynomials. The original convolutional code from which various values of R are obtained is often termed as the mother code.

Decoding a convolutional code is generally performed with the help of a trellis-based decoder having 2^{L_c} states. The most popular trellis-based decoders encountered in the literature are the Viterbi algorithm [For73] and the Bahl-Cocke-Jelinek-Raviv (BCJR) algorithm [BC$^+$74]. Similar to the block codes, where the error detection and correction capabilities of a code is determined by its minimum Hamming distance $d_{\min}$, convolutional codes use a metric called the minimum free distance (also denoted as $d_{\min}$ in this dissertation) [LC01]. The value of the minimum free distance represents the smallest difference in path metrics between the all-zeroes output and the path with the smallest non-zero path metric going from the initial all-zeroes state to some future all-zeroes state. The larger the value of minimum free distance, the more error detection and correction capabilities a convolutional code possesses. The implementation and investigation of a convolutional code with respect to the diffusion-based molecular communication is presented in [SL$^+$12; YC$^+$12; MM$^+$13]. Furthermore, an irregular convolutional code has been implemented in [DH18; Dam20] in order to shape the extrinsic transfer function used for iterative decoding.

In this dissertation, based on the mother code in Figure 5.1 and (5.6), convolutional codes with code rates $R = 3/4$ and $R = 1/2$ are implemented in the air-based molecular communication testbed using fluorescent dyes. For decoding purposes of the convolutional codes, the Viterbi algorithm is preferred over the BCJR algorithm, as the former offers similar decoding performance as the latter with a much lower complexity [CC$^+$19]. With the help of the Viterbi algorithm, the info bit sequence which had the maximum likelihood of getting transmitted is determined, the details of which is explained in Section 4.3.3.

5.3 Line Codes

Line codes were first introduced for the purpose of transmitting digital information using the telephone lines [Gib18]. By definition, the process of conversion of digital data into real-time temporal waveforms in baseband is the process of line coding. Various transmission channel characteristics along with different applications and performance requirements

have led to the development of different line codes, which are commonly found in the literature such as the non-return to zero, the return to zero, alternate mark inversion, Manchester code, etc. [Lat95]. However, the utility and choice of the line codes depends on the following considerations as mentioned in [Gib18]:

- *Timing* - The line coded waveform contains the timing information for the purpose of synchronization of the receiver with the transmitter, without getting affected by the data transmission sequence
- *Direct current content* - The line coded waveform does not contain any direct current (DC), which causes DC wander leading to the variation of the baseline of the received signal with time
- *Power spectrum* - The power spectrum and bandwidth of the waveforms are adapted to the frequency response of the transmission channel in order to avoid signal distortion
- *Performance monitoring* - The error detection and error avoidance capabilities during transmission channel use are realized for the purpose of performance monitoring, as line codes are able to adapt the transmitted data to the channel characteristics
- *Probability of error* - In reliable line codes, the average error probability is maintained at a minimum level for the given transmitter power
- *Transparency* - Line codes allow all combinations of the data transmission sequences, with a possibility to use alternatives to the undesirable sequence patterns

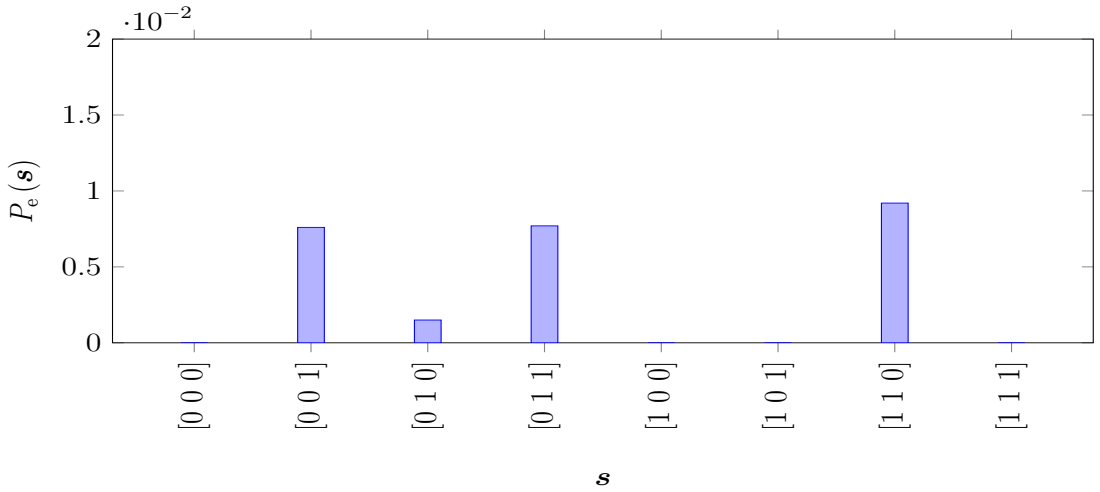
Line codes can particularly be used in the applications of molecular communication due to their ability to adapt data to the transmission channel characteristics, which helps minimize transmission errors. The amplitude-dependent noise, which typically affects the molecular communication channel, gives rise to a transmission sequence-dependent error probability. This means that some binary symbol sequences are more error-prone than the others. Therefore, inspired by this fact, a new probability-based distance metric has been introduced in [KL$^+$12] as a replacement to the Hamming distance:

$$d_{\mathrm{p}}\left(\boldsymbol{x}, \hat{\boldsymbol{x}}\right) = -\log\left(P\left(\boldsymbol{x} \to \hat{\boldsymbol{x}}\right)\right), \tag{5.7}$$

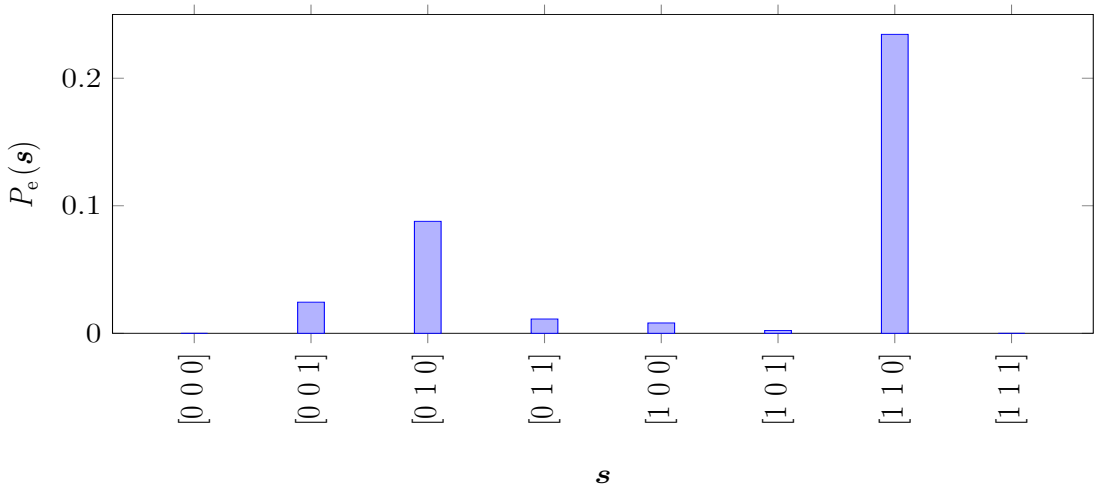
where $P\left(\boldsymbol{x} \to \hat{\boldsymbol{x}}\right)$ is the probability of receiving sequence $\hat{\boldsymbol{x}}$, when $\boldsymbol{x}$ is the transmitted sequence. A detailed investigation of line codes for diffusion-based microscopic molecular

communication channel is presented in [Dam20]. In the mentioned work, run-length limited code is implemented to prevent high-frequency signals, which experience attenuation due to the low-pass characteristic of the diffusion-based channel. Additionally, it also helps generate such code words which contain more zeros than ones. This is because the transmission of binary symbol "1" is more error-prone than the transmission of binary symbol "0".

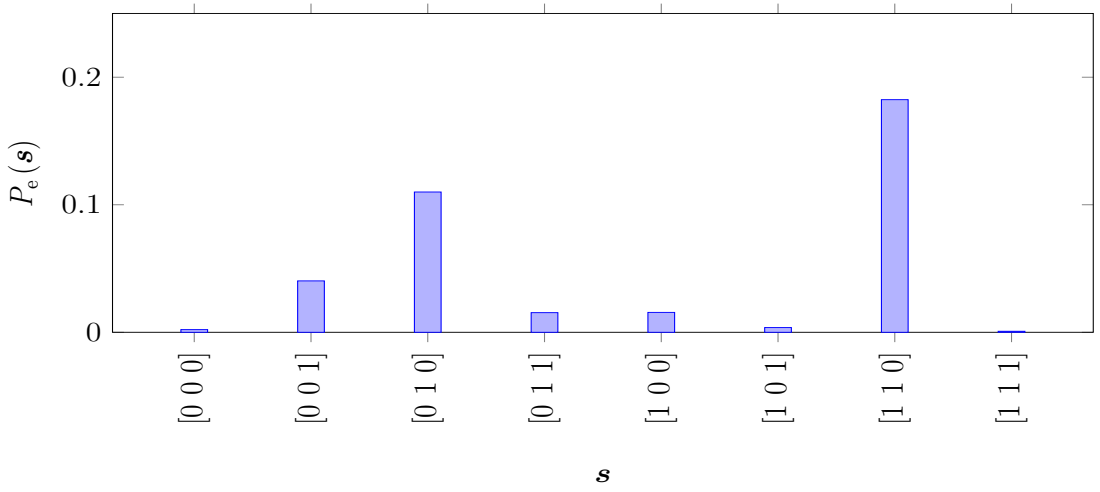
The work in [Dam20] also presents and investigates another novel line coding technique dubbed weak sequence preventing mapping (WSPM), which is introduced to mitigate the transmission errors by omitting the most error-prone code words. In this line coding technique, a codebook for an (n_c, k_c) block code is created by performing an error analysis of all 2^{n_c} possible received words. Then, from these received words, 2^{k_c} least error-prone code words are selected and assigned to 2^{k_c} info words. By applying the same principle, $(3, 2)$ WSPM code is designed and implemented in this dissertation for the macroscopic air-based molecular communication testbed under investigation. This leads to the use of only those code words, which are least prone to transmission errors due to the ill effects of ISI and noise. A water-based dye solution of uranine is used as the information carrier to perform the experiments in this section. An error analysis over 50 testbed measurement trials for the sequence length $n_c = 3$ is performed on a continuous sequence of randomly generated OOK modulated binary data symbols of length 108 (testbed measurement parameters listed in Table 3.1). The analysis is performed for the transmission distances $d = 118\,\text{cm}$, $d = 148\,\text{cm}$ and $d = 178\,\text{cm}$, as shown in Figure 5.2. Any other transmission distance less than $d = 118\,\text{cm}$ is not considered for the measurement analysis, as for these distances, OOK transmission scheme with FTD does not produce any error. For a distance $d = 118\,\text{cm}$, Figure 5.2(a) illustrates that the sequence [1 1 0] is the most error-prone sequence. However, the result is inconclusive as for $d = 118\,\text{cm}$, substantial amount of errors are not produced for the OOK modulated transmitted data. Concrete evidences with regard to the most error-prone sequence is obtained for distances $d = 148\,\text{cm}$ and $d = 178\,\text{cm}$, as illustrated in Figure 5.2(b) and Figure 5.2(c), respectively. This corroborates the fact that [1 1 0] is indeed the most error-prone sequence in the case of testbed measurements. Out of the $2^3 = 8$ possible code words, $2^2 = 4$ code words are selected and assigned to the four info words, as shown in Table 5.6. The selected code words reduce the probability of occurrence of [1 1 0] in a continuous stream of code word sequence. Compared to $(3, 2)$ SPC code, $(3, 2)$ WSPM code is beneficial on many fronts. Firstly, the chances of transmitting the most error-prone sequence [1 1 0] is notably reduced. This is observed by considering an example, where two possible code words are transmitted in succession. In the case of $(3, 2)$ SPC code, the likelihood of occurrence of the sequence [1 1 0] is 75%, while in the case of



(a) $d = 118\,\mathrm{cm}$



(b) $d = 148\,\mathrm{cm}$



(c) $d = 178\,\mathrm{cm}$

Figure 5.2: Sequence error analysis for length $n_c = 3$ for the weak sequence preventing mapping code based on the testbed measurements.

$(3, 2)$ WSPM code, the likelihood of the occurrence of the same sequence is merely 6.25%. Secondly, in the case of $(3, 2)$ WSPM code, the water-based dye solution is sprayed on average 0.375 times per info bit, which offers the best resource efficiency as compared to the other implemented FEC codes, which are summarized in Table 5.7. Interestingly, the performance is even better than the uncoded OOK modulation scheme, which requires to spray 0.5 times per info bit on average.

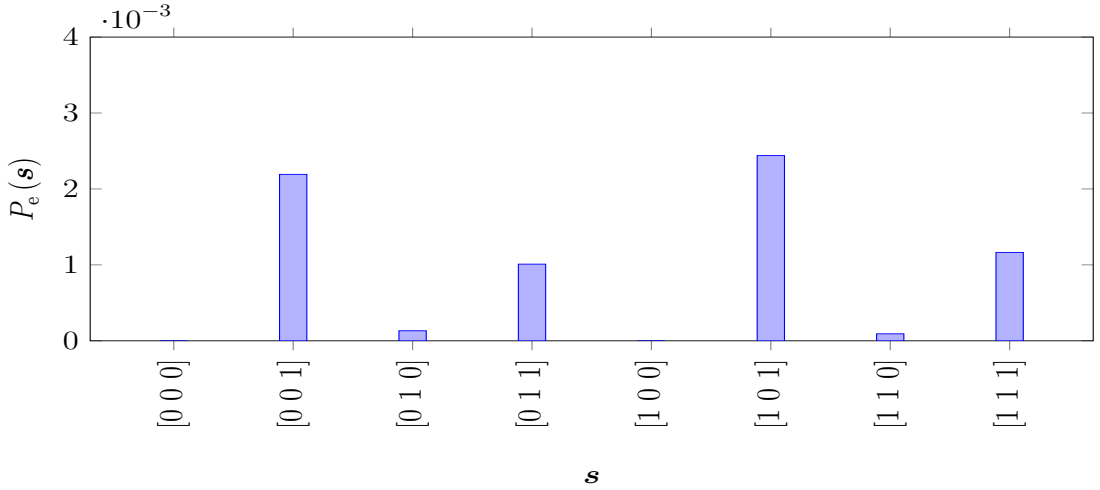
Table 5.6: Code table of $(3, 2)$ weak sequence preventing mapping code inferred from the analysis based on the testbed measurements.

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

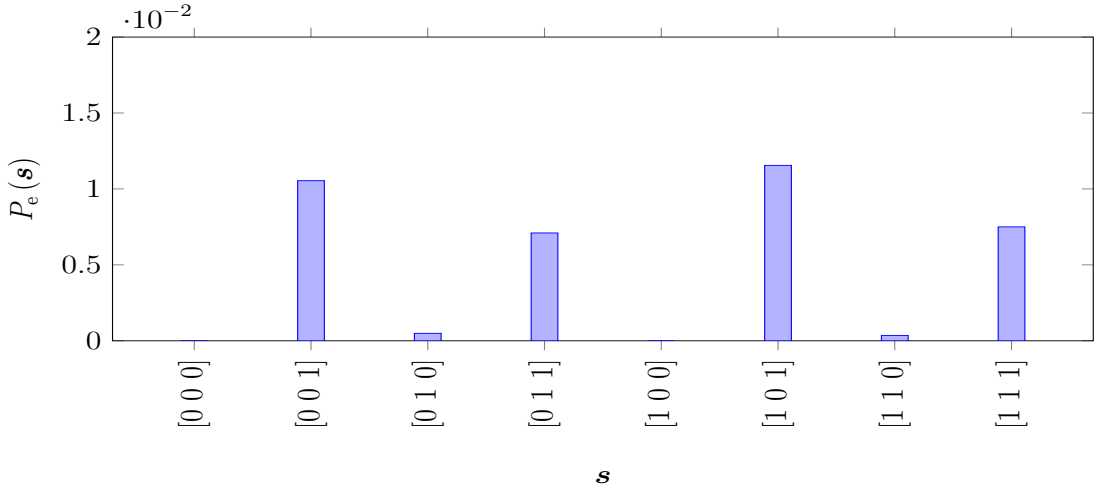
Table 5.7: Comparison of the channel coding schemes and the on-off keying based on the performance parameters.

Channel Code	d_min	R	Sprays per info bit
(2,1) Repetition	2	1/2	1
(3,1) Repetition	3	1/3	1.5
(3,2) SPC	2	2/3	0.75
(8,4) CRC	2	1/2	1
(16,8) CRC	4	1/2	1
(15,11) Hamming	3	11/15	0.682
(7,4) Hamming	3	4/7	0.875
(4,3) Convolutional	4	3/4	0.674
(2,1) Convolutional	7	1/2	1
(3,2) WSPM	1	2/3	0.375
OOK (uncoded)	-	1	0.5

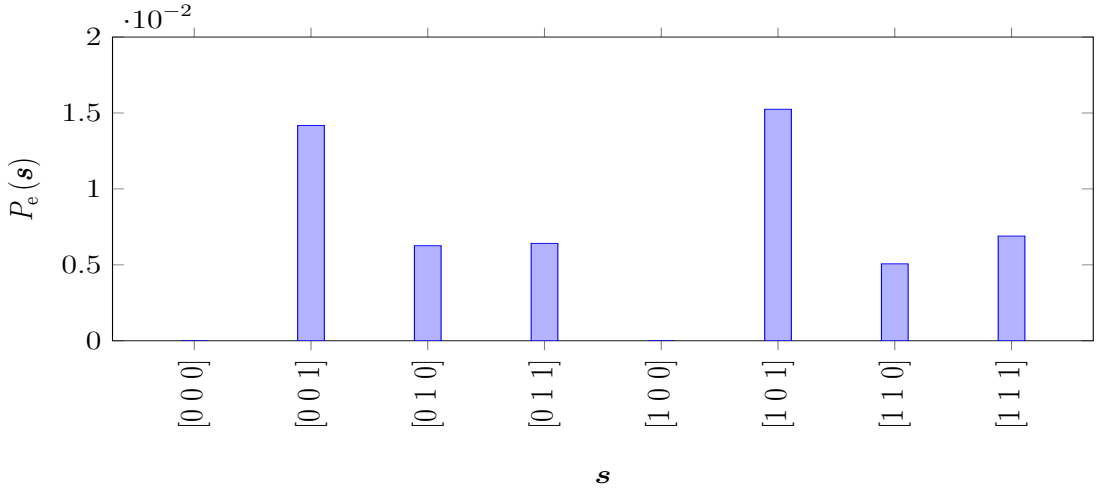
For the analytical model presented in Section 3.7, the error analysis is performed in order to find out the least error prone code words for $(3, 2)$ WSPM code. Similar to the testbed measurements, the error analysis using the analytical model is also performed for the sequence length $n_\mathrm{c} = 3$. The distances $d = 118\,\mathrm{cm}$, $d = 148\,\mathrm{cm}$ and $d = 178\,\mathrm{cm}$ are considered for the analysis by performing transmission of a continuous sequence of randomly generated OOK modulated binary data symbols of length 108 over 100000 trials, as shown in Figure 5.3. It is observed that [1 0 1] is the most error-prone sequence, when the analytical model is considered. The deviation from the error analysis performed on



(a) $d = 118\,\mathrm{cm}$



(b) $d = 148\,\mathrm{cm}$



(c) $d = 178\,\mathrm{cm}$

Figure 5.3: Sequence error analysis for length $n_c = 3$ for the weak sequence preventing mapping code based on the analytical model.

the testbed measurements is due to the fact that the analytical model only considers the effect of ISI on the current data transmission symbol from the previously transmitted data symbol. However, the error analysis on the testbed measurements is performed considering ISI effect of more than one previously transmitted data symbols on the current data symbol. Out of the $2^3 = 8$ possible code words, $2^2 = 4$ code words are selected and assigned to the four info words, as shown in Table 5.8. The selected code words reduce the probability of occurrence of [1 0 1] in a continuous stream of code word sequence.

Table 5.8: Code table of the (3,2) weak sequence preventing mapping code inferred from the error analysis based on the analytical model.

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

When considering the example, where two possible code words are transmitted in succession, the chances of transmitting the most error-prone sequence [1 0 1] is zero in the case of $(3, 2)$ WSPM code. However, in the case of the $(3, 2)$ SPC code, the likelihood of occurrence of the sequence [1 0 1] is 75%. For the analytical model, $(3, 2)$ WSPM code sprays the water-based dye solution 0.625 times per info bit on average, which is higher than the testbed measurements, but still offers the best resource efficiency as compared to the other implemented FEC codes. The decoding of the $(3, 2)$ WSPM code is performed by FTD-based detection of the received energy samples. An error correction of the invalid code word to the true code word is performed by considering the soft value in the form of the distance between the energy sample and the threshold, and finding the highest correlation among all the available code words.

5.4 Results

The results of the BER analysis with the increasing transmission distance d for the implemented channel coding schemes are discussed in this section. The BER performance based on the testbed measurements using the water-based solution of uranine for a randomly generated info word of length 108 over 20 trials is presented in Figure 5.4 (testbed measurement parameters listed in Table 3.1). For a better observation of the BER performance for smaller values of d, the simulation results based on the analytical model presented in

Figure 5.5 is considered for a randomly generated info word of the same length over 100000 trials. As a common observation, it is seen that the BER performance deteriorates with increase in d. This is because there is a significant reduction in the number of droplets traveling the length of the tube and reaching the field of view of the camera-based detector at the receiver side. The droplets either fall down and collect at the bottom of the tube or simply get dispersed laterally. Also, the trajectory assumed by the spray is influenced by the gravitational force, increasing the distance between the camera-based detector and the droplets, as discussed in Section 3.3.2, thus, reducing the intensity of light received by the camera. All these influences decrease the received signal strength while contributing to the noise power. Moreover, the slowing down of the sprayed droplets due to the effect of air drag further leads to an increase in ISI.

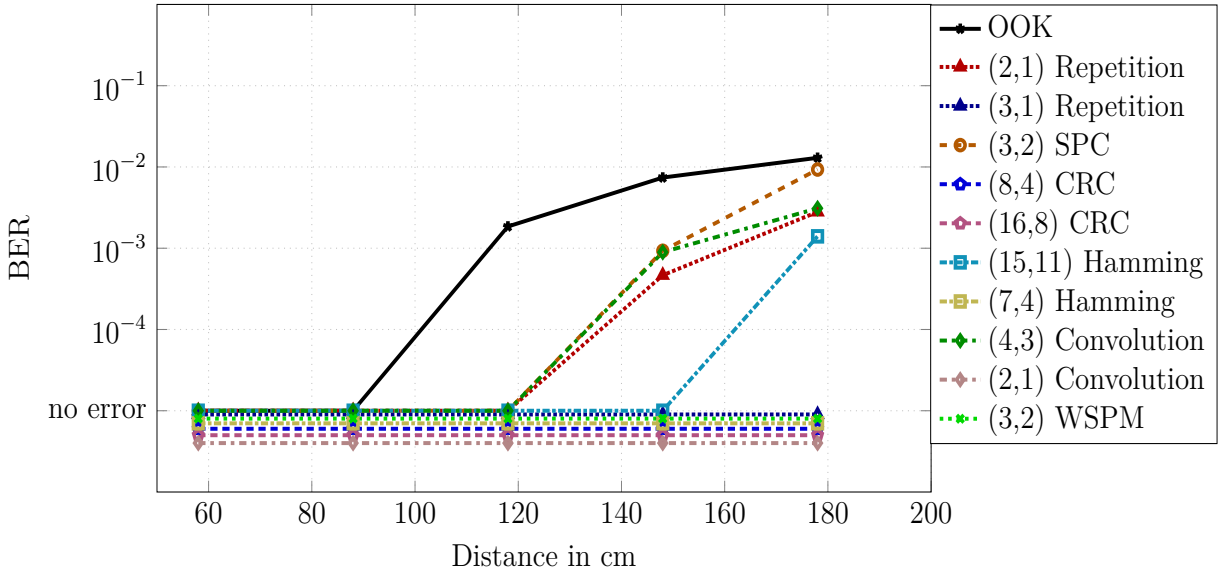


Figure 5.4: Bit error rate performance of the encoded transmission data sequences based on the testbed measurements with increasing distance d.

The BER performances of the channel coding schemes with FTD are compared with respect to the OOK modulated received energy samples recorded by the camera-based detector. It is evident from Figure 5.4 that $(3, 1)$ repetition code offers a better BER performance as compared to $(2, 1)$ repetition code, and produces no error up to $d = 178\,\text{cm}$, which is also corroborated by the simulation results using the analytical model depicted in Figure 5.5. This is due to the fact that $(3, 1)$ repetition code has a lower value of code rate R as compared to $(2, 1)$ repetition code, with more redundant bits capable of detecting and correcting more erroneous bits. In addition, a lower value of R increases the energy per info bit in the transmission system under investigation. From the measurement

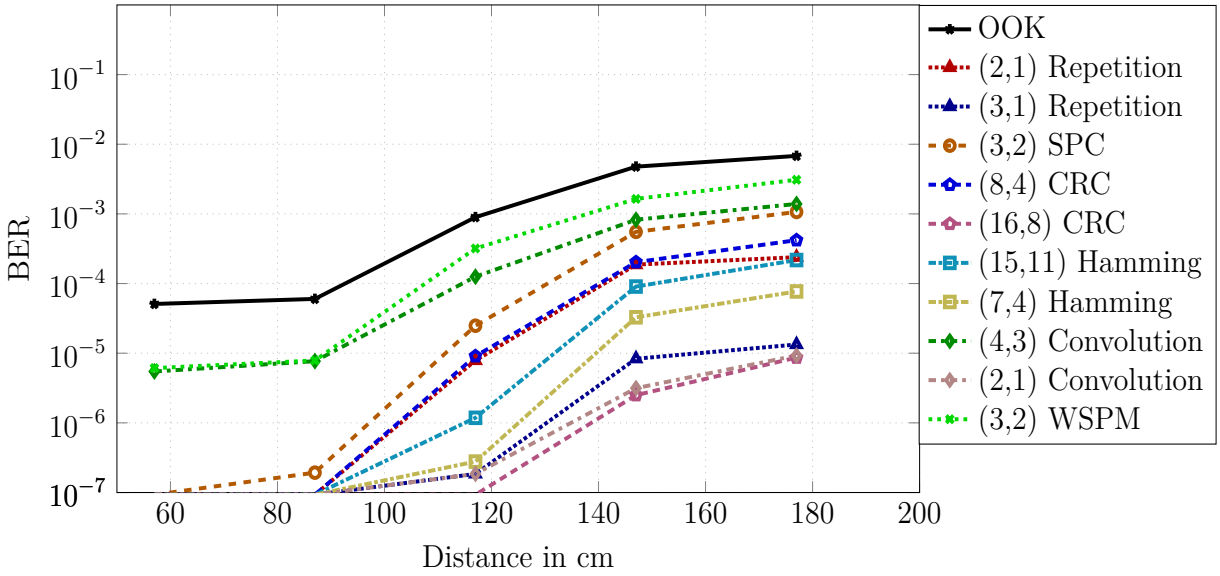


Figure 5.5: Bit error rate performance of the encoded transmission data sequences based on the analytical model with increasing distance d.

results obtained from the testbed, the performances of $(8, 4)$ and $(16, 8)$ CRC codes are very similar and also give zero error for the transmission distances under consideration, as shown in Figure 5.4. A more accurate observation can be made from the simulation results of the analytical model shown in Figure 5.5, which depicts that $(16, 8)$ CRC code outperforms $(8, 4)$ CRC code in terms of the BER performance. This is due to the fact that $(16, 8)$ CRC code has a minimum Hamming distance $d_{\text{min}} = 4$, which enhances its coding gain. Overall, the soft-input decoding of CRC code is quite powerful, however, more computationally complex, which is another reason why no error is obtained in the case of $(8, 4)$ CRC code for the measurement results acquired from the testbed. The performance of $(3, 2)$ SPC code is the worst among all the implemented coding schemes for the testbed measurements, as presented in Figure 5.4. This is due to the fact that the likelihood of transmission of the most error-prone sequence [1 1 0] is quite high. In order to tackle this issue, $(3, 2)$ WSPM code is designed keeping in mind the channel characteristics. The likelihood of transmission of the most error-prone sequence [1 1 0] is significantly reduced in this case, due to which, no error is observed for the transmission distances taken into account. However, this is not the case for the results obtained from the simulation results using the analytical model, presented in Figure 5.5. $(3, 2)$ WSPM code performs worse than $(3, 2)$ SPC code due to the fact that although the transmission of the most error-prone sequence [1 0 1] is avoided, the probability of occurrences of the other more error-prone sequences such as [0 0 1] are still pretty high, leading to the deterioration of the BER

performance. $(7, 4)$ Hamming code performs better than $(15, 11)$ Hamming code in the case of both testbed measurement results as well as the simulation results based on the analytical model presented in Figure 5.4 and Figure 5.5, respectively. $(7, 4)$ Hamming code in fact, produces no error for d up to 178 cm with respect to the testbed measurements. This is again because $(7, 4)$ Hamming code has higher redundancy and higher energy per info bit as compared to $(15, 11)$ Hamming code. It is also interesting to see that the value of R for $(7, 4)$ Hamming code, although higher than $(2, 1)$ repetition code and $(8, 4)$ CRC code, has a better BER performance than them, as evident from the simulation results based on the analytical model in Figure 5.5. This is because $d_{\min} = 3$ remains fixed for all Hamming codes, which helps boost their coding gain. Also, soft-input decoding employed is quite powerful, although being more computationally complex. This is also evident from Figure 5.5 that $(15, 11)$ Hamming code has a much higher value of R as compared to $(2, 1)$ repetition code, but their BER performances at $d = 178$ cm are quite similar. Derived from $(4, 1, 4)$ convolutional mother code, $(2, 1)$ convolutional code is error-free in the case of the results obtained from the testbed measurements, depicted in Figure 5.4, compared to $(3, 4)$ convolutional code for the transmission distances under consideration. This is because $(2, 1)$ convolutional code has a lower value of R and higher energy per info bit. $(2, 1)$ convolutional code also performs better than $(2, 1)$ repetition code, as the former offers a minimum free distance of $d_{\min} = 7$ that helps improve its coding gain as compared to the latter, which offers a minimum Hamming distance of $d_{\min} = 2$. Additionally, $(2, 1)$ convolutional code is decoded using the Viterbi algorithm, which is based on maximum-likelihood sequence estimation (MLSE), offering a higher computational power with a better BER performance.

5.5 Spatial Codes

Multiple-input multiple-output (MIMO) technique is a very popular method incorporated in wireless communication systems, where several transmit and receive antennas are used for data transmission purposes. Currently, MIMO has become an inalienable part of many modern wireless communication standards such as the 5G, Wi-Fi, long term evolution, and worldwide interoperability for microwave access [FK08]. The information in MIMO is encoded by making use of an additional degree of freedom in the form of spatial domain. Multiple antennas in a MIMO system with respect to wireless communication can be utilized in the following forms to achieve improvements in system performance, as mentioned in [MS+09]:

- *Spatial Multiplexing* - The process in which independent data sequences are transmit-

ted over multiple antennas called layers. This method ensures higher data transmission rates as compared to the systems employing single antenna or the single-input single-output (SISO) configuration through multiplexing gain.

- *Spatial Diversity* - This technique makes use of the multiple antennas by transmitting redundant signals representing the same data transmission sequence in order to improve the BER. With the help of space-time coding (two-dimensional coding in time and spatial domains), a diversity gain is achieved.

- *Smart Antennas* - This method makes use of steerable antennas in order to improve the signal-to-noise ratio (SNR), and to subdue the inter-link interference (ILI) for a multiuser scenario at the receiver side. Beamforming technique is used to steer the beam patterns towards a desired receive antenna in order to attain an increased SNR through array gain.

In the context of molecular communication, MIMO for a diffusion-based molecular communication has been introduced for the first time in [MY+12], where various aspects such as transmitter diversity, receiver-side diversity combining techniques, and spatial multiplexing are discussed. Since then, many theoretical works have come up in this field of research such as machine learning-based channel modeling [LY+17], introduction of modulation schemes [GB+19b], receiver design optimization strategies as well as channel estimation, equalization and detection processes [RS19] (details in Section 1.5). However, very few practical implementations of MIMO concepts from the perspective of molecular communication have been realized in the past. The very first attempt to extend a SISO alcohol-based molecular communication testbed setup to a MIMO setup has been showcased in [KL+16; LK+20a; LK+20b]. In similar lines, this dissertation presents a 2×2 MIMO scenario for an air-based molecular communication testbed using fluorescent dyes. The goal in this work is to achieve an improvement in the BER performance of the data transmission process. Therefore, different spatial algorithms at the transmitter and the receiver sides are employed and investigated while taking both the effects of ISI as well as ILI into account. As presented in [DY+17; DY+18], two types of spatial coding techniques namely the Alamouti-type coding and repetition MIMO coding are implemented at the transmitter side. At the receiver side, two different receiver combining techniques namely equal gain combining (EGC) and selection diversity (SD) are used.

5.5.1 MIMO Testbed Setup

In order to tap into the advantages of spatial diversity offered by a MIMO configuration, different spatial codes and combining/selection strategies presented in [DY+17; DY+18] for a diffusion-based molecular communication channel are implemented in a 2×2 MIMO setup for the macroscopic air-based molecular communication testbed presented in this dissertation. The MIMO configuration of the testbed as depicted in Figure 5.6 is very similar to the SISO scenario, discussed in Section 2.2. The difference, however, lies in the fact that the sprayer acting as the transmitter (Tx) is extended by an additional sprayer with both the containers filled with water-based solutions of uranine, while the field of view of the receiver (Rx) side camera is split into two equal halves, each acting as an independent receiver similar to its previous implementation in Section 4.4.7. Analogous to the radio-based wireless transmission systems, the two sprayers at the transmitter side and the split field of view of the camera-based detector at the receiver side are referred to as the antennas from this point on, even though they significantly differ in terms of their operations compared to a classical antenna. In a symmetrical arrangement, the distance between the spray nozzles of the transmit antennas Tx_0 and Tx_1 is fixed at 16 cm, which ensures the reception of the sprayed droplets by both the receive antennas Rx_0 and Rx_1 of the camera-based detector. The distance between Tx_0 and Rx_0, and between Tx_1 and Rx_1 is d.

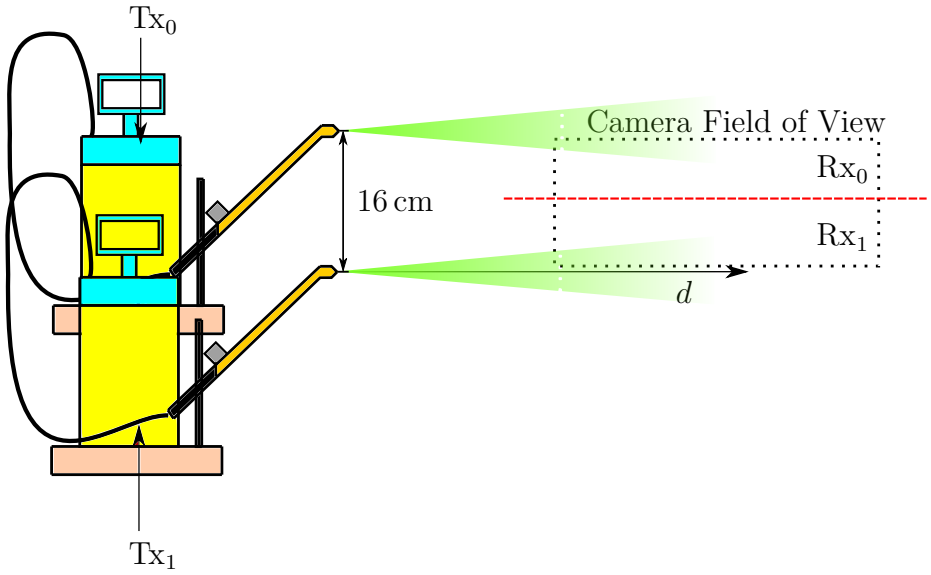


Figure 5.6: A 2×2 multiple-input multiple-output setup of the testbed with two transmit antennas Tx_0 and Tx_1, and two receive antennas Rx_0 and Rx_1, with a separation distance d between the transmit antennas and the receive antennas.

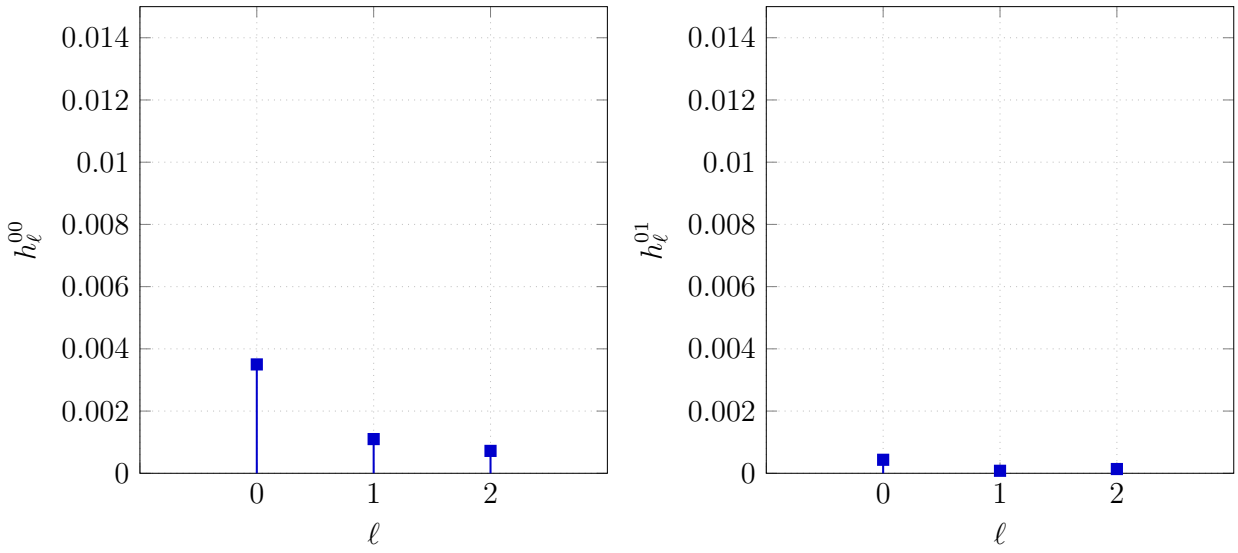
The end-to-end system coefficient h_ℓ^{ji} describes the end-to-end subsystems between the transmit antenna Tx_i and the receive antenna Rx_j. The pixel energy sample from the sprayed information carrier droplets recorded at the receive antenna Rx_j during the kth transmission interval can be expressed by the superposition of all available end-to-end sub systems. Therefore, the equivalent discrete-time end-to-end system model in Section 4.1 at Rx_j for N_{Tx} transmit antennas can be generalized as according to [DY+17; DY+18] by the expression

$$y_j[k] = \sum_{i=0}^{N_{\mathrm{Tx}}-1} \sum_{\ell=0}^{L} h_\ell^{ji} x_i[k-\ell] + \omega_j[k], \tag{5.8}$$

where $x_i[k]$ is the data symbol transmitted by Tx_i in kth transmission interval and $\omega_j[k]$ represents the noise term at Rx_j arising due to the reasons such as imperfection of transmit antennas, turbulence, air drag, etc., as mentioned in Section 3.6 in detail. It must be mentioned that the end-to-end system coefficients are not normalized in order to preserve the features like the hardware position/configuration related and the distance dependent signal strength. All the investigations are carried out with OOK modulated data sequence at the transmit antennas, where $x_i[k] \in \{0, 1\}$. As an example, the end-to-end system coefficients of the 2×2 MIMO configuration at distance $d = 178\,\mathrm{cm}$ is depicted in Figure 5.7. The end-to-end system coefficients represented by h_ℓ^{ji} for the MIMO subsystems for system memory length $L = 2$ are evaluated through the testbed measurements averaged over 20 trials. The effects of ILI is more pronounced at the considered distance, as the sprayed droplets from the antennas experience lateral dispersion, which gets larger with increase in d between the transmit antennas and the receive antennas. It is observed that the system coefficients h_ℓ^{10} and h_ℓ^{11} are stronger as compared to h_ℓ^{00} and h_ℓ^{01}. Since there is only one ultraviolet (UV)-A light source illuminating the sprayed water-based solution droplets at the receive antennas, the received droplets which lie closer to the UV-A light source in the spatial domain glow much brighter than those droplets that lie much farther away from the light source, resulting in the higher magnitude of the system coefficients.

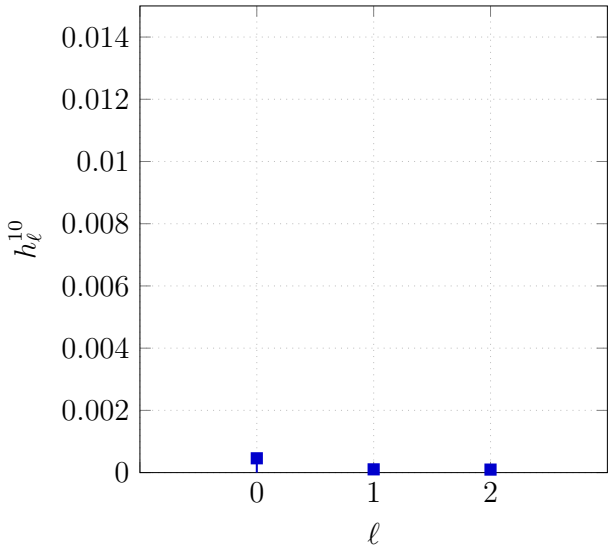
5.5.2 Alamouti-type Code

Spatial diversity achieving algorithms in MIMO systems make use of the multiple transmit antennas to send redundant information to the receive antennas in order to achieve better BER performance. Some MIMO coding techniques introduce the redundancy in both spatial and time domains, and are therefore, referred to as space-time coding. In this regard, the Alamouti space-time block code, originally introduced for a two-antenna wireless communication systems in [Ala98], is an orthogonal space-time coding strategy.

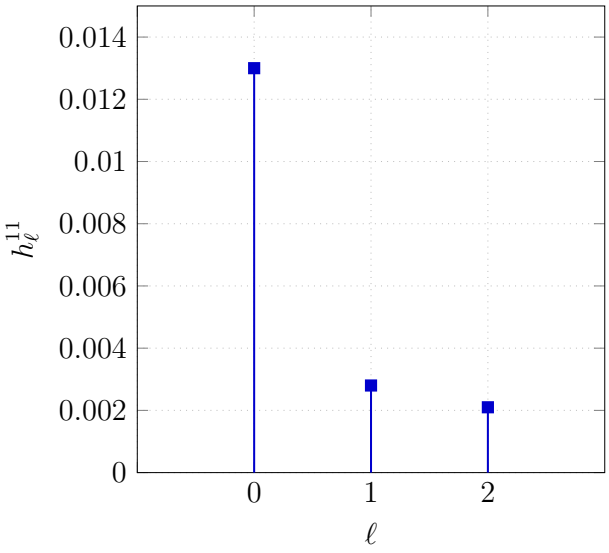


(a) System coefficient h_ℓ^{00}

(b) System coefficient h_ℓ^{01}



(c) System coefficient h_ℓ^{10}



(d) System coefficient h_ℓ^{11}

Figure 5.7: The end-to-end system coefficients of the 2×2 multiple-input multiple-output configuration for the presented testbed, at distance $d = 178\,\text{cm}$.

The information can typically be represented by a sequence of data with symbol cardinality M. However, in this dissertation, only binary data symbols are considered, where the kth info bit u_k from an info word $\boldsymbol{u}$ represents the information. Therefore, the transmission matrix for a 2×2 Alamouti space-time block code can be described by

$$\boldsymbol{T} = \begin{bmatrix} u_k & u_{k+1} \\ -u_{k+1}^* & u_k^* \end{bmatrix}. \tag{5.9}$$

The rows of $\boldsymbol{T}$ represent two consecutive transmission intervals kT and $(k+1)T$, while the columns constitute the two transmit antennas Tx_0 and Tx_1. Therefore, it is understood that in the first time slot, $x_0[k] = u_k$ is transmitted over Tx_0 and $x_1[k] = u_{k+1}$ is transmitted over Tx_1. Similarly, in the next time slot, $x_0[k+1] = -u_{k+1}^*$ and $x_1[k+1] = u_k^*$ are transmitted over the antennas Tx_0 and Tx_1, respectively. The information from both info bits are transmitted over the two transmit antennas in two successive transmission intervals, leading to a diversity gain. Due to the orthogonality introduced by the Alamouti space-time block code at the receive antennas, the condition $\boldsymbol{T}^{\mathsf{H}}\boldsymbol{T} = c\boldsymbol{I}_2$ holds true, where $\boldsymbol{T}^{\mathsf{H}}$ is the Hermitian of the transmission matrix, c is a constant, and $\boldsymbol{I}_2$ is a 2×2 identity matrix. Thus, a complete cancellation of ILI is obtained at the receive antennas, further simplifying the maximum-likelihood detection process. In practical scenarios, the orthogonality often gets lost when the transmission channel gets affected by the ISI, necessitating the use of more complex detection algorithms such as the MLSE [MH04]. The Alamouti space-time block code has been originally conceived for a complex-valued bipolar data symbols, while the data in molecular communication, as in the case of optical communication, is represented by unipolar real-valued data symbols. In order to accommodate these conditions, the Alamouti space-time block code for wireless communication systems can be modified to an Alamouti-type code, as presented in [SV05] for free-space optical communication channels, and extended for diffusion-based molecular communication channels in [DY$^+$17; DY$^+$18]. The Alamouti-type code is formed by taking into account only the real-valued data symbols without the complex conjugate operation as well as replacing the data symbols having negative signs with their inverse equivalent symbols. Therefore, the modified transmission matrix of the Alamouti-type code for OOK modulation in air-based molecular communication testbed can be expressed as

$$\boldsymbol{T} = \begin{bmatrix} u_k & u_{k+1} \\ 1 - u_{k+1} & u_k \end{bmatrix}, u_k \in \{0, 1\}. \tag{5.10}$$

It is to be noted that for the testbed using water-based solution of uranine as the information carrier, the transmission data is affected by ISI leading to a loss in orthogonality that Alamouti-type code typically offers. This entails the use of more complex detection algorithms such as the MLSE at the receive antennas similar to the conventional Alamouti space-time block code with ISI. Since the Alamouti-type code spreads the information of two data symbols over two time slots, the branch metric in MLSE introduced in Section 4.3.3 should be jointly evaluated for both time slots, as also mentioned in [DY+17; DY+18]. For instance, if ISI for the system memory length $L = 1$ is taken into account, the transmission matrix is given as

$$\boldsymbol{T}_1 = \begin{bmatrix} u_k & u_{k+1} & 1 - u_{k-1} & u_{k-2} \\ 1 - u_{k+1} & u_k & u_k & u_{k+1} \end{bmatrix}, \tag{5.11}$$

where $\boldsymbol{T}_L$ is the transmission matrix taking into account L ISI terms. The ISI terms belonging to the previous transmission slots of transmit antennas Tx_0 and Tx_1 are given by the last two columns in (5.11). Therefore, the pixel energy samples due to the sprayed information carrier droplets at the receive antennas Rx_0 and Rx_1 can be expressed as

$$\begin{bmatrix} y_0[k] & y_1[k] \\ y_0[k+1] & y_1[k+1] \end{bmatrix} = \boldsymbol{T}_1 \begin{bmatrix} h_0^{00} & h_0^{10} \\ h_0^{01} & h_0^{11} \\ h_1^{00} & h_1^{10} \\ h_1^{01} & h_1^{11} \end{bmatrix} + \begin{bmatrix} \omega_0[k] & \omega_1[k] \\ \omega_0[k+1] & \omega_1[k+1] \end{bmatrix}. \tag{5.12}$$

5.5.3 Repetiton MIMO Code

Another technique to achieve diversity gain is to simply code the information over the spatial dimension instead of using the time dimension. Based on the exact same principle, a spatial coding strategy namely repetition MIMO coding is introduced for a laser-based optical communication in [WB+05]. The coding design of the repetition MIMO is much simpler at the transmit and the receive antennas as compared to the orthogonal Alamouti space-time block code, making the former a more intuitive choice against the latter. Repetition MIMO is also popularly used in applications related to visible light communication [Hoe19]. The information in repetition MIMO is replicated over all the transmit antennas and transmitted simultaneously to the corresponding receive antennas using OOK modulation scheme. The information can typically be represented by a sequence of data symbols of cardinality M. However, in this dissertation, only binary data symbols are considered, where the kth info bit u_k from an info word $\boldsymbol{u}$ represents the information. Therefore, the transmission matrix

for a 2×2 MIMO system can be written in the matrix form

$$\boldsymbol{T} = \begin{bmatrix} u_k & u_k \end{bmatrix}. \tag{5.13}$$

Repetition MIMO coding does not provide the advantages of orthogonality. Thus, ILI cannot be removed at the receive antennas. However, it is to be noted that since the same information is transmitted simultaneously over every transmit antenna, the ILI acts rather constructively at the receive antennas from the viewpoint of the received signal strength. On the other hand, ISI at the receive antennas can be removed quite conveniently by implementing the detection algorithms for SISO systems.

5.5.4 Receiver Combining Strategies

The main advantage of using a MIMO setup is the availability of multiple receive antennas that can acquire the information-carrying signal in the spatial domain, which can then be utilized for detection purposes. The idea is then to either consider one of these acquired signals or to combine several of these acquired signals together, before employing the detection algorithm. To this end, the most popular selection and combining strategies discussed in the literature are SD and EGC [Höh13], which have also been taken into consideration for the 2×2 MIMO system presented in this dissertation.

In the case of SD, the receive antenna that acquires the information-carrying signal with the maximum strength is selected to perform the detection algorithm. The strength of the received signal is determined by the end-to-end system coefficients h_ℓ^{ji}. Therefore, the higher the value of system coefficient, the stronger the signal acquired by the receive antenna. However, for the 2×2 MIMO scenario introduced in this work, it is observed that the system coefficients h_ℓ^{10} and h_ℓ^{11} are stronger as compared to h_ℓ^{00} and h_ℓ^{01} for the reasons stated in Section 5.5.1. According to the MIMO testbed setup, the receive antenna Rx_1 receives the droplets which end up lying closer to the UV-A light source as compared to the droplets received by antenna Rx_0, leading to employment of the detection algorithm on the acquired signal with a higher SNR. This results in the signal Rx_1 to be selected for the scenario under consideration as

$$y_{\mathrm{SD}}[k] = y_1[k] = \sum_{\ell=0}^{L} \left(h_\ell^{10} x_0[k-\ell] + h_\ell^{11} x_1[k-\ell] \right) + \omega_1[k]. \tag{5.14}$$

For detection purposes, the FTD rule for SD is used which is given as

$$\hat{x}_0[k] = \hat{x}_1[k] = \begin{cases} 0 & \text{if } y_{\text{SD}}[k] \leq \xi, \\ 1 & \text{if } y_{\text{SD}}[k] > \xi. \end{cases} \tag{5.15}$$

In the case of EGC, all the acquired information signals by the receive antennas are equally weighted and summed together, resulting in a very low-complexity combining strategy. Without loss of generality as well as choosing the weights to be equal to one gives

$$\begin{aligned} y_{\text{EGC}}[k] &= y_0[k] + y_1[k] \\ &= \sum_{\ell=0}^{L} \left[\left(h_\ell^{00} + h_\ell^{10} \right) x_0[k-\ell] + \left(h_\ell^{01} + h_\ell^{11} \right) x_1[k-\ell] \right] + \omega_0[k] + \omega_1[k]. \end{aligned} \tag{5.16}$$

For detection purposes, the FTD rule for EGC is used that is given as

$$\hat{x}_0[k] = \hat{x}_1[k] = \begin{cases} 0 & \text{if } y_{\text{EGC}}[k] \leq \xi, \\ 1 & \text{if } y_{\text{EGC}}[k] > \xi. \end{cases} \tag{5.17}$$

The optimum combining strategy is based upon assigning different weights to the acquired information signal by the receive antennas with respect to the channel quality. This is known as the maximum ratio combining (MRC) strategy [Höh13], popularly used in MIMO scenarios for radio wave-based wireless communication systems. However, in this work, the emphasis has particularly been laid on showcasing and implementing low-complexity combining strategies such as SD and EGC, which do not require the knowledge of the channel, and can also be of great importance to applications concerning the micro-scale molecular communication systems.

5.5.5 Results

The results of the BER analysis with increasing transmission distance d for the implemented MIMO spatial codes are discussed in this section. The analysis is performed based on the testbed measurements using the water-based solution of uranine as the information carrier as well as the simulation results based on the analytical model, presented in Section 3.7. An end-to-end system with a memory length $L = 2$ is considered for the analysis. It is important to mention that the resource normalization at the transmit antennas is not considered for the spatial diversity algorithms for this work, as it is shown in [DY+18] that when power normalization is applied in a diffusion-based molecular communication

channel, no diversity gain is achieved in the case of implemented MIMO algorithms over the SISO scenario. Therefore, without power normalization, an array gain is achieved over diversity gain, which is also the case for the spatial diversity algorithms used in the MIMO system considered in this dissertation.

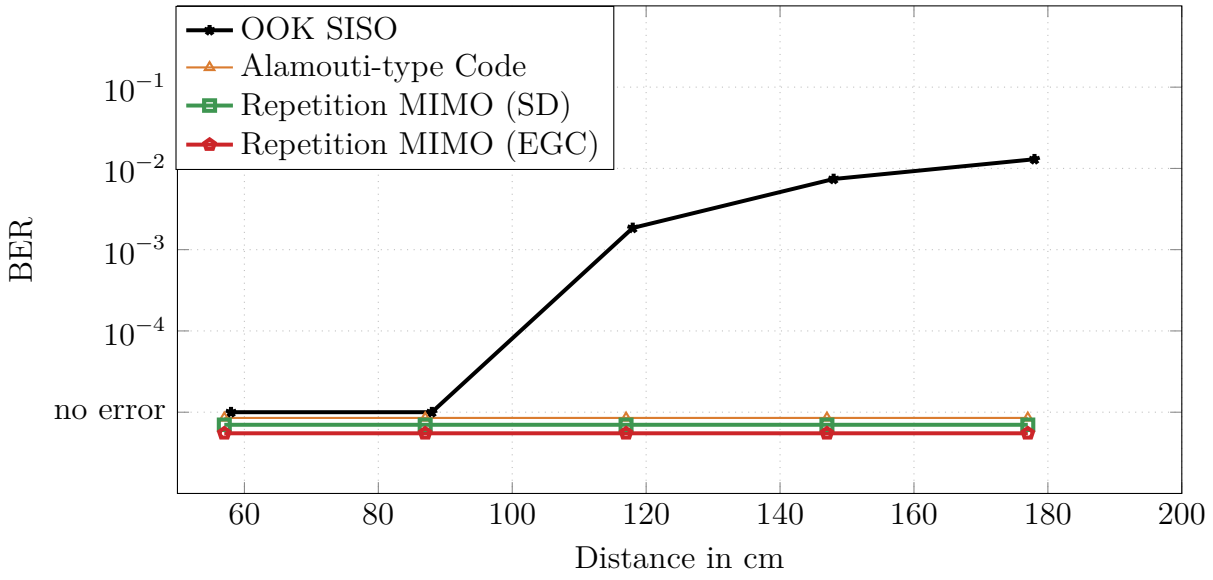


Figure 5.8: Bit error rate performance of the spatial encoded transmission data sequences based on the testbed measurements with increasing distance d.

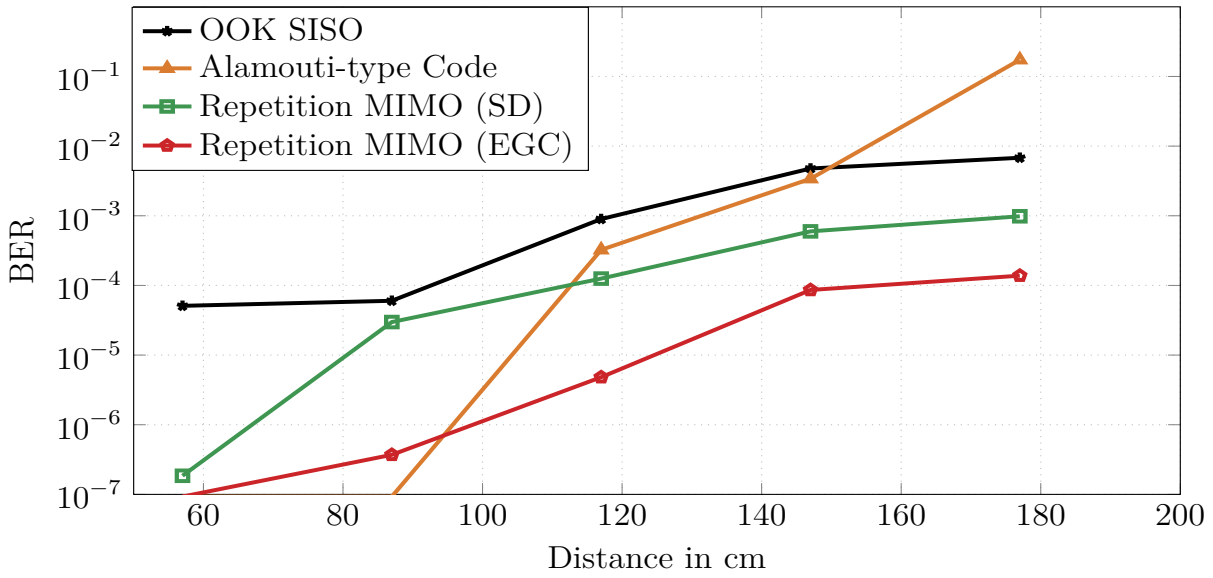


Figure 5.9: Bit error rate performance of the spatial encoded transmission data sequences based on the analytical model with increasing distance d.

The BER performance from the testbed measurements for a transmission sequence length

of 108 randomly generated binary data symbols over 20 trials is presented in Figure 5.8 (testbed measurement parameters listed in Table 3.1). It is observed that compared to the OOK modulated SISO transmission system, all spatial diversity techniques such as the Alamouti-type coding as well as the repetition MIMO coding result in no error for increase in distance d. However, when the simulation results from the analytical model presented in Figure 5.9 is considered for a transmission sequence length of 108 randomly generated binary data symbols and 100000 trials, it is observed that the repetition MIMO coding with EGC receiver combining strategy performs best in terms of the BER performance with increasing d. This is because of the array gain achieved by the use of two different transmitters requiring twice the amount of resources as compared to the SISO transmission system. Similarly, the BER performance of repetition MIMO code employing SD receiver combining strategy performs better than the SISO transmission system with increasing d. This is due to the fact that in a 2×2 MIMO configuration of the testbed setup, one of the transmitters is closer to the UV-A light source illuminating the sprayed droplets of the water-based solution of the information carrier as compared to the transmitter used in the SISO configuration of the testbed setup. This results in a proportionately higher SNR at the receive antenna in the case of 2×2 MIMO setup, thus improving the BER performance when using SD combining strategy at the receiver side. Additionally, it is also observed that the Alamouti-type code achieves an improved BER performance as compared to repetition MIMO code with both EGC and SD combining strategies for smaller values of d. This is because for smaller d, ILI is negligible, while a more complex detection algorithm such as MLSE outperforms FTD-based detection algorithm at the receiver side. However, as d increases, the BER performance worsens despite employing MLSE detection algorithm. This is due to the fact that in repetition MIMO with EGC, ILI contributes constructively to the received signal, while in the Alamouti-type code, ILI has a rather destructive effect. It is also to be mentioned that ILI for the analytical model has been treated as a noise process. As a general observation, it is also seen that for all the implemented spatial coding schemes, the BER performance worsens with increase in d. This is due to the ill effects of the decreasing SNR and increasing ISI with d. A further improvement in BER performance for repetition MIMO code can be achieved by also combining the channel equalization methods presented in Section 4.3.

5.6 Summary

In this chapter, error minimization techniques in the form of channel coding have been employed for the presented macroscopic air-based molecular communication testbed using a water-based solution of uranine as the information carrier. The performance analysis of different FEC channel codes belonging to the families of block codes and convolutional codes, apart from line codes and spatial codes have been investigated from the standpoint of the presented macroscopic molecular communication testbed, based on the measurements as well as the analytical model. $(3, 2)$ WSPM line code especially, has been implemented for the very first time in a molecular communication testbed, which minimizes the transmission errors by the use of code words that are adapted to the channel characteristics. It is also shown that this low-complexity line coding technique not only helps minimize the transmission errors but also uses the least amount of resources on average, as compared to the other implemented FEC codes. A BER performance analysis of the FEC codes and the line code have been presented with respect to the increasing distance between the transmitter and the receiver. The analysis is performed based on the testbed measurements as well as the simulation results obtained from the analytical model, presented in Section 3.7.

Similarly, the implementation of spatial codes has been achieved by extending the SISO setup of the testbed to a 2×2 MIMO system. The Alamouti-type code as well as repetition MIMO code have been presented as two spatial coding techniques that help improve the BER performance of the transmission system with increasing distance between the transmitter and the receiver by exploiting the spatial domain. The array gain involved in the spatial coding techniques leads to a better BER performance compared to the SISO transmission system, which is presented with the help of a BER analysis, making use of both the testbed measurements as well as the simulation results acquired from the analytical model. Repetition MIMO incorporating EGC as a low-complexity combining strategy particularly displays the best BER performance amongst all the implemented spatial coding techniques due to the constructive effect of ILI. Moreover, it also offers advantage over the optimal MRC combining strategy, as it completely eliminates the requirement of the channel knowledge at the receiver side, as the performance of the MRC gets hindered with inaccurate estimation of the channel.

6

Conclusion and Outlook

6.1 Conclusion

The work presented in this dissertation presents the new practical implementation of macroscopic molecular communication testbed using water-based solution of fluorescent dye as information carrier, sprayed over an air-based transmission channel medium. Due to the limited availability of resources in the literature with respect to the practical implementation of molecular communication at a macro-scale level, this work attempts at discussing its various aspects from the perspective of communication theory that would eventually be helpful in its standardization for industry-specific applications. Infrastructure monitoring applications such as the underground sewage facilities and oil/gas industries, where pipelines and ducting systems are extensively deployed, can especially reap benefits of molecular communication, where otherwise conventional radio-based wireless communication performs poorly. Therefore, the focus of this work is to investigate a complete end-to-end data transmission system that includes the transmitter, the information carrier, the transmission channel, and the receiver encompassing the end-to-end system model, the transmission and reception algorithms, as well as the error mitigation strategies. Additionally, for the considered transmission distance, one of the main objectives of this work is also to attain high data transmission rates at low bit error rate (BER) in order to make molecular communication at macro-scale level a more feasible choice of data communication.

The testbed design and implementation of the air-based macroscopic molecular communication using fluorescent dye is discussed in this dissertation. At the transmitter side, a pressure sprayer containing the water-based solution of uranine or the highlighter ink as the information carrier is operated with the help of a micro-controller. The sprayed

droplets of the water-based information-carrying dye solution, which pose no threat to health or environment, are transmitted through a transmission channel tube, which is then detected using a camera at the receiver side. The information signal is extracted in the form of green or red pixel intensity values of the camera-based detector. The performance of the overall transmission system especially depends on the responsitivity of the receiver, which is gauged with the help of certain performance metrics namely peak's maximum to full-width at half maximum, delay to peak's maximum and interference to total received energy ratio. In order to compare the system performance of the molecular communication testbed using fluorescent dyes, the traditionally implemented molecular communication testbed using alcohol is considered as a benchmark. It is demonstrated that the aforementioned performance metrics show defining results which are in favor of the molecular communication testbed using fluorescent dyes, as opposed to the molecular communication testbed using alcohol.

The theory behind the mass transportation process of the water-based solution droplets of the dye can be perceived from the standpoint of particle-based molecular communication. In this regard, some of the most common phenomena facilitating the transportation of information-carrying particles from the transmitter to the receiver that are often mentioned in the literature include diffusion, advection, turbulent flow, convection, and dispersion. However, when molecular communication takes place in an air-based medium with information carrier in liquid-based solution droplets, the released droplets from the transmitter side sprayer mostly get carried towards the receiver through diffusion, advection and/or turbulent flow. In this work, the transportation of the water-based dye solution droplets are modeled using the advection-diffusion process by taking into account physical influences such as the drag and gravitational forces. The end-to-end system model of the testbed is considered to model the system impulse response, which is a direct consequence of the combined effects of the transmitter, the propagation channel as well as the receiver. Moreover, the molecular communication system that is under investigation is also shown to be nearly linear, and is independent of the choice of the information carrier in the form of the water-based solutions of uranine or the highlighter ink. Any uncertainty and randomness in the system impulse response due to inconsistent spray bursts at the transmitter side, turbulence causing initial interaction of released high velocity droplets with air, imperfection of the camera-based detector at the receiver side, etc., is modeled with the help of an amplitude-dependent noise process. Finally, the system impulse response and the noise model are combined together to formulate the complete analytical model of the testbed, which is validated by the transmission of a message sequence through measurement trials

from the testbed. The analytical model helps to run computer simulations with very large number of trials in order to obtain more accurate results, which would otherwise be a tedious process when obtaining the results based on the testbed measurements.

The data acquisition principle pertaining to the presented air-based macroscopic molecular communication testbed using fluorescent dyes is also focused in this dissertation. Obtaining the green pixel intensity value for the water-based solution of uranine and obtaining the red pixel intensity value for the highlighter ink water-based solution are elucidated, followed by the accumulation of energy under these pixel intensity pulses over a symbol duration, a process delineated through an equivalent discrete-time end-to-end system model. Simple binary transmitter-side modulation schemes and low-complexity receiver-side threshold-based detection algorithms are presented with a possibility to include equalization techniques for further improvement of BER performance. Additionally, higher-order modulation schemes incorporating various degrees of freedom are implemented. A performance analysis in terms of their hardware constrained data transmission rates as well as their resource requirements are showcased, which gives a unique perspective to the user when choosing a particular modulation scheme for data transmission purposes. Moreover, investigative analysis of the implemented modulation schemes in terms of their BER performance is also presented in this dissertation based on the results obtained by conducting the testbed measurements as well as the simulation results obtained through the analytical model.

From the wireless communication and information theory literature, it is a well-known fact that channel coding schemes are used to successfully mitigate the transmission errors that originate in the channel. The same goes for the macroscopic air-based molecular communication system presented in this work, where the transmission channel errors occur due to the reasons such as the effects of inter-symbol interference (ISI), noise processes, etc. The analysis of various forward error correction (FEC) codes that are broadly classified under two larger coding families of block codes (like repetition code, single parity check code, cyclic redundancy check code, and the Hamming code) and convolutional codes have been carried out. A specific type of line code namely weak sequence preventing mapping code has been implemented in a macroscopic molecular communication testbed, which minimizes the transmission errors by the use of code words that are adapted to the channel characteristics. This low-complexity line coding technique additionally helps to minimize the average resource requirement as compared to the other implemented FEC codes. Furthermore, spatial codes have also been implemented by extending the single-input single-output (SISO) setup of the macroscopic air-based molecular communication

testbed to a 2×2 multiple-input multiple-output (MIMO) system. The performances of the Alamouti-type code and the repetition MIMO code have specifically been investigated with respect to their roles in improving BER performances by exploiting the spatial domain. The array gain in the spatial codes helps to improve the BER performance with respect to the SISO transmission system as observed from both the testbed measurements as well as the simulation results acquired from the analytical model. Low-complexity combining strategies such as equal gain combining also leads to an improvement in BER performance due to the constructive effect of the inter-link interference (ILI) at the receiver side.

6.2 Outlook

The field of molecular communication is still at a very nascent stage and developing continuously. This offers several research opportunities to the community in multiple areas. This outlook discusses possibilities of further research on the topics presented in this dissertation.

The end-to-end system of the macroscopic air-based molecular communication testbed presented in this work is based on the principle of advection-diffusion process due to the transmission channel itself being a bounded tube, reducing the effects of turbulence. In order to expand the investigation, scenarios that cause turbulence such as the presence of obstacles as well as drift in random direction can help make the system model more holistic. Additionally, reflecting surfaces inside the transmission channel as well as various spray characteristics and shapes can be investigated. Moreover, the presence of the water-based information-carrying droplets in the background can be another scenario for which the noise model can be formulated. Mobility of transmitters and receivers can also be studied for the testbed in order to perceive its impact on the system impulse response. The same setup can be extended to include a network of tube-like structures emulating a pipeline network, commonly found in oil/gas industries and sewage systems employing relaying and routing strategies.

The dissertation discusses about the use of spatial domain and color of dye as two additional degrees of freedom along with the concentration of the sprayed water-based dye solution for the representation of transmit data symbols. Future research work can combine all the three degrees of freedom together in order to represent data symbols of a higher symbol cardinality, thus increasing the transmission rates further. The analytical models, when simultaneously using two different colored dyes or spatial regions, can be improved by modeling the effects of shadowing and ILI. Innovative modulation techniques and channel coding schemes can also be designed that not only serve their respective purposes efficiently,

but also minimize the average volume of the sprayed water-based solution of the information carrier per bit.

Finally, the transmitter and the receiver hardware can themselves be modified and improved in order to achieve higher data transmission rates. For example, the work presented in [BK+23] makes use of a faster switching solenoid valve at the transmitter side, controlling the opening and shutting mechanism of the spray nozzle, and releasing the water-based solution of the information carrier. This is used in conjunction with an improved detector with a PIN-photodiode at the receiver side, which enhances the detection capability of the sprayed droplets. The end-to-end system model for this system can be proposed to formulate the analytical model including the noise process for simulation purposes. The modulation schemes with higher degrees of freedom can additionally be used to increase the data transmission rates by manifold. Furthermore, the range of data transmission between the transmitter and the receiver can be increased for a target BER when equalization techniques and channel coding schemes are introduced into the system.

A

Notation

Acronyms

ANN	Artificial neural network
ATD	Adaptive threshold detector
BCJR	Bahl-Cocke-Jelinek-Raviv
BER	Bit error rate
C6-HSL	N-(3-Oxyhexanoyl)-L-homoserine lactone
CRC	Cyclic redundancy check
CSK	Concentration shift keying
CTS	Communication through silence
DC	Direct current
DFE	Decision-feedback equalizer
DPM	Delay to peak's maximum
DPPM	Differential pulse-position modulation
EGC	Equal gain combining
EM	Electro magnetic
FEC	Forward error correction
FIR	Finite impulse response
FPS	Frames per second
FTD	Fixed threshold detector
FWHM	Full-width at half maximum
ICI	Inter-color interference
IIR	Infinite impulse response

ILI	Inter-link interference
ISI	Inter-symbol interference
ITR	Interference to total received energy ratio
LAN	Local area network
LED	Light-emitting diode
MA	Moving average
MCSK	Molecular concentration shift keying
MIMO	Multiple-input multiple-output
MLSE	Maximum-likelihood sequence estimation
MoSK	Molecular shift keying
MPPM	Multi-pulse pulse-position modulation
MRC	Maximum ratio combining
MSSK	Molecular space shift keying
3-hydroxy-AHL	N-(3-hydoxyacyl) homoserine lactone
3-oxo-AHL	N-(3-oxoacyl) homoserine lactone
OFC	Optical fiber communication
OOK	On-off keying
PDE	Partial differential equation
PDF	Probability density function
PMFWHM	Peak's maximum to full-width at half maximum
PPM	Pulse-position modulation
RMSE	Root-mean-square error
Rx	Receiver
SARS-CoV-2	*Severe Acute Respiratory Syndrome Coronavirus 2*
SCSK	Spatial concentration shift keying
SD	Selection diversity
SISO	Single-input single-output
SNR	Signal-to-noise ratio
SPC	Single parity check
SPION	Superparamagnetic iron oxide nanoparticle
SPPM	Shorten pulse-position modulation
SSK	Spatial shift keying
TEC	Time-elapse communication
Tx	Transmitter
UV	Ultraviolet

VLC	Visible light communication
VOC	Volatile organic compound
WSPM	Weak sequence preventing mapping
ZF	Zero-forcing

Mathematical Notation

$\mathbf{0}$	Zero vector
a_1, a_2	Scaling constants
$\mathbf{A}$	Pilot matrix
A_{ISI}	Total energy in system impulse response curve
A_{sym}	Energy in system impulse response curve beyond symbol duration
$\mathrm{ACF}(\cdot)$	Auto-correlation function
c_{light}	Speed of light in air
$\mathrm{cov}(\cdot)$	Covariance
d	Transmission distance
d_{eff}	Effective linear dimension of the environment
d_{gas}	Diameter of a gas molecule
d_{min}	Minimum Hamming/free distance of block/convolutional code
$d_{\mathrm{p}}(\cdot,\cdot)$	Probability-based distance
D	Diffusion coefficient
D_{eddy}	Eddy-diffusion coefficient
$f(\cdot)$	Function
f_{cam}	Camera frame rate
g	ZF filter coefficient
$g(z)$	Generator polynomial
G	Number of ZF filter coefficients
$\boldsymbol{G}$	Generator matrix
h	End-to-end system coefficient
$\hat{\mathbf{h}}$	Estimated end-to-end system coefficient vector
h_ℓ^{ij}	ℓth system coefficient of end-to-end system between jth transmit antenna and ith receive antenna
$h(t)$	System impulse response
$h_{\mathrm{alc}}(t)$	System impulse response for alcohol
$h_{\mathrm{ink}}(t)$	System impulse response for the highlighter ink

$h_{\max}(t)$	Maximum system impulse response value
$h_{\text{meas}}(t)$	Measured system impulse response
$h_{\text{uran}}(t)$	System impulse response for uranine
$h(mT_{\text{cam}})$	$h(t)$ sampled at time $t = mT_{\text{cam}}$
h'	Height of the stream of sprayed droplets
h'_{fit}	Fitted height
i_{sim}	Simulation step
i_{trial}	Trial index
$\boldsymbol{I}$	Identity matrix
k	Discrete time
k_{B}	Boltzmann constant
k_{c}	Info word length
K	Sequence length of the transmitted pilot/data symbols
K_{c}	Total chip slots occupied by MPPM pulses
L	System memory length
L_{c}	Memory length / number of shift registers of a convolutional code
m	Multiples of sampling period
m_{gas}	Molar mass of a gas
M	Symbol cardinality
M_0	Number of molecules released by the point transmitter at time $t = 0$
n_{c}	Code word length
n_m	Noise sample generated using the MA model
$\hat{n}_m$	Estimated noise sample
n_{rec}	Number of frames from LED turning on to the first appearance of droplets at the center of the field of view of camera
n_{trans}	Number of frames from LED turning on to the first appearance of droplets just outside the sprayer nozzle
$n(t)$	Continuous-time additive noise
N	Number of samples
N_{c}	Total available chip slots
N_{trial}	Total number of trials
N_{Tx}	Number of transmitters
$\mathbb{N}$	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

$\mathcal{N}_{\text{half}}(\cdot,\cdot)$	Half Gaussian distribution
p_i	ith parity bit
$p(\cdot)$	Probability of a transmitted data sequence in MLSE
$\mathbf{p}$	Vector with the coefficients of the system impulse response model
$\boldsymbol{P}$	Parity check matrix
$\boldsymbol{P}_{\text{c}}$	Submatrix with parity checks
$P(\cdot)$	Probability
P_{b}	Bit error probability
$P_{\text{e}}(\boldsymbol{s})$	Error probability of sequence $\boldsymbol{s}$
P_{s}	Symbol error probability
P_{gas}	Pressure of a gas
r	Transmission radius
r_{mol}	Radius of the diffusing molecule
R	Code rate
R^2	Goodness of fit
R_b	Maximum achievable average data transmission rate
s_i	DPPM symbol
S_1, S_2, S_3	Data transmission sequences for the test of linearity
t	Time
T	Symbol duration
T_{c}	Chip slot duration
T_{cam}	Sampling period of the camera-based detector
T_{on}	Spray duration
T_{gap}	Time gap in DPPM
T_{pulse}	Reference spray-pulse in DPPM
T_{trans}	Minimum shutting time of the valve
$\boldsymbol{T}$	Transmission matrix
$\mathfrak{T}[\cdot]$	System
u	Info bit
$\boldsymbol{u}_{\text{c}}$	Info word
U	Possible concentration levels in MCSK and SCSK
v	Velocity of the information carrying particles
v_{fit}	Fitted velocity
v_{fluid}	Flow velocity in the fluid propagation medium
v_{meas}	Measured velocity from the experiments

$v_{\rho_\mathrm{x}}, v_{\rho_\mathrm{y}}, v_{\rho_\mathrm{z}}$	Flow velocity components in x, y, and z dimensions
$\mathbf{v}$	Velocity vector
V_on	Volume of the water-based solution of the information carrier released over T_on duration
V	Average volume sprayed per bit
w_MA	MA model order
$W(\cdot)$	Wiener process
x	Transmit symbol
$\tilde{x}$	Transmit symbol hypothesis
$\hat{x}$	Estimated transmit symbol
$\boldsymbol{x}$	Transmit symbol sequence
$\tilde{\boldsymbol{x}}$	Transmit symbol sequence hypothesis
$\boldsymbol{x}_\mathrm{c}$	Code word
$\mathcal{X}$	Transmit symbol alphabet
y	Received pixel energy sample
y'	Received pixel energy sample after post-processing
$y(t)$	Continuous-time output signal of the end-to-end system
$\mathbf{y}$	Received pixel energy sample vector
y_EGC	EGC receive sample
y_SD	SD receive sample
z	Z-transform variable
Z_mol	Different molecular types in MCSK
Z_space	Different spatial regions in SCSK
$\alpha_1, \beta_1, \gamma_1$	Coefficients of the system impulse response model for the first local maximum
$\alpha_2, \beta_2, \gamma_2$	Coefficients of the system impulse response model for the second local maximum
$\Gamma_{\boldsymbol{\mu}_\mathrm{mem}}[k]$	Path metric
$\delta(\cdot)$	Dirac delta function
Δt	Discrete time step
$\Delta\rho$	Discrete displacement step in spherical coordinate system
$\Delta\rho_\mathrm{x}, \Delta\rho_\mathrm{y}, \Delta\rho_\mathrm{z}$	Discrete step in x, y, and z dimensions
ε_m	Zero-mean white Gaussian noise process sample with a variance of 1
η	Viscosity of the aqueous medium
ϑ	Temperature of the medium

κ	Scaling factor
$\lambda_{\boldsymbol{\mu}_{\text{mem}}}^{\tilde{x}[k]}[k]$	Branch metric
μ	Mean
μ_1	Mean of the distribution corresponding to data symbol "1"
$\mu_{\hat{n}_{m,N_{\text{trial}}}}$	Mean value of the estimated noise samples at time $t = mT_{\text{cam}}$ for N_{trial} total trials
$\boldsymbol{\mu}_{\text{mem}}$	Memory state
ξ	FTD threshold value
ρ_{corr}	Pearson correlation coefficient
ρ, φ, θ	Location in radial, azimuthal, and polar dimension
$\boldsymbol{\rho}$	Location vector
$\rho_{\text{x}}, \rho_{\text{y}}, \rho_{\text{z}}$	Location in x, y, and z dimensions
σ	Standard deviation
σ_0	Standard deviation of the distribution corresponding to data symbol "0"
σ_1	Standard deviation of the distribution corresponding to data symbol "1"
$\sigma(\hat{n}_{m,N_{\text{trial}}})$	Standard deviation of the estimated noise samples at time $t = mT_{\text{cam}}$ for N_{trial} total trials
σ_n	Standard deviation of the continuous-time amplitude-dependent white Gausssian noise process
σ^2	Variance
$\sigma^2\left(\hat{n}_{m,N_{\text{trial}}}\right)$	Variance of the estimated noise samples at time $t = mT_{\text{cam}}$ for N_{trial} total trials
$\sigma_n^2(t)$	Variance of the continuous-time amplitude-dependent white Gausssian noise process
τ	Time taken by the system impulse response to reach its maximum value
τ_{lag}	Time lag
$\phi(\boldsymbol{\rho}, t)$	Concentration at a point $\boldsymbol{\rho}$ in space at time instant t
ϕ_{tilt}	Angular tilt of the transmission channel tube
$\phi_{\text{tilt,max}}$	Maximum angular tilt of the transmission channel tube
ψ_i	ith weighting factor for the MA model used in the generation of the noise sample
ξ	Detection threshold
ω	Noise sample
∇	Divergence operator
∇^2	Laplacian operator

B

System Impulse Response Coefficients

Table B.1: Obtained model coefficients with increasing distance d by performing curve-fitting through measured system impulse response averaged over 100 trials for uranine water-based solution.

Distance d (cm)	α_1	β_1	γ_1	α_2	β_2	γ_2	RMSE
58	1.05×10^{-4}	0.67	28.68	2.91×10^{-4}	2.63	17.37	3.87×10^{-4}
88	5.76×10^{-5}	0.32	46.63	1.92×10^{-4}	1.19	27.16	1.75×10^{-4}
118	6.59×10^{-5}	0.18	60.57	2.53×10^{-4}	0.57	35.41	1.99×10^{-4}
148	2.70×10^{-5}	0.12	85.9	1.27×10^{-4}	0.20	46.34	1.50×10^{-4}
178	3.57×10^{-5}	0.09	106	2.19×10^{-4}	0.07	51.83	2.26×10^{-4}

Table B.2: Obtained model coefficients with increasing distance d by performing curve-fitting through measured system impulse response averaged over 100 trials for highlighter ink water-based solution.

Distance d (cm)	α_1	β_1	γ_1	α_2	β_2	γ_2	RMSE
58	1.21×10^{-4}	0.71	29.94	2.99×10^{-4}	3.36	18.12	3.70×10^{-4}
88	1.08×10^{-4}	0.33	46.62	2.29×10^{-4}	1.62	27.58	2.28×10^{-4}
118	6.67×10^{-5}	0.19	61.71	1.44×10^{-4}	0.73	36.98	6.60×10^{-5}
148	7.48×10^{-5}	0.10	71.91	1.74×10^{-4}	0.26	42.14	1.46×10^{-4}
178	5.21×10^{-5}	0.07	86.54	1.44×10^{-4}	0.08	48.25	1.03×10^{-4}

C

Evaluation of Theoretical BER

C.1 Variation of Distribution Parameters with Increasing Distance

In the following, the standard deviation of the half Gaussian distribution for binary data symbol "0" is denoted by $\mathcal{N}_{\text{half}}(0, \sigma_0^2)$. Similarly, the mean and the standard deviation of Gaussian distribution for binary data symbol "1" is denoted by $\mathcal{N}(\mu_1, \sigma_1^2)$. The variation of these distribution parameters with respect to increasing distance d for both uranine as well as the highlighter ink water-based solutions as the information carriers are summarized in Table C.1 and Table C.2, respectively. The distribution parameters are attained based on the testbed measurements and the histogram curve-fitting function `histfit()` available on MATLAB.

Table C.1: The variation of standard deviation σ_0 of $\mathcal{N}_{\text{half}}(0, \sigma_0^2)$ and the variation of mean μ_1 and standard deviation σ_1 of $\mathcal{N}(\mu_1, \sigma_1^2)$ with increasing distance d for uranine water-based solution.

Distance d (cm)	Standard Deviation σ_0	Mean μ_1	Standard Deviation σ_1
58	0.0028	0.0662	0.0131
88	0.0017	0.0364	0.0089
118	0.0018	0.0426	0.0117
148	0.0012	0.0209	0.0061
178	0.0015	0.0189	0.0067

Table C.2: The variation of standard deviation σ_0 of $\mathcal{N}_{\text{half}}(0, \sigma_0^2)$ and the variation of mean μ_1 and standard deviation σ_1 of $\mathcal{N}(\mu_1, \sigma_1^2)$ with increasing distance d for the highlighter ink water-based solution.

Distance d (cm)	Standard Deviation σ_0	Mean μ_1	Standard Deviation σ_1
58	0.0074	0.0855	0.0124
88	0.0026	0.0356	0.0063
118	0.0015	0.0201	0.0047
148	0.0014	0.0149	0.0043
178	0.0014	0.0113	0.0036

C.2 Derivation of Intersection Point of Two Distributions

In the following, the intersection point between a half Gaussian distribution and a Gaussian distribution for the evaluation of the theoretical bit error rate (BER) is derived. The half Gaussian distribution for binary data symbol "0" is represented by $\mathcal{N}_{\text{half}}(0, \sigma_0^2)$ and the Gaussian distribution for binary data symbol "1" is represented by $\mathcal{N}(\mu_1, \sigma_1^2)$. Since the received energy sample $y[k]$ is always non-negative, the distributions are only defined for the region $0 < y[k] < \infty$. For detection, the non-zero intersection point is of interest. For simplicity reasons, $y[k]$ is represented by y, and equating both the distributions results in

$$\frac{\sqrt{2}}{\sigma_0\sqrt{\pi}} \exp\left(-\frac{y^2}{2\sigma_0^2}\right) = \frac{1}{\sigma_1\sqrt{2\pi}} \exp\left(-\frac{(y-\mu_1)^2}{2\sigma_1^2}\right), \tag{C.1}$$

$$\ln\left(\frac{\sqrt{2}}{\sigma_0\sqrt{\pi}}\right) - \ln\left(\frac{1}{\sigma_1\sqrt{2\pi}}\right) = \frac{y^2}{2\sigma_0^2} - \frac{(y-\mu_1)^2}{2\sigma_1^2}, \tag{C.2}$$

$$\ln\left(2\frac{\sigma_1}{\sigma_0}\right) = \frac{\sigma_1^2 y^2 - \sigma_0^2 y^2 + 2\mu_1\sigma_0^2 y - \mu_1^2\sigma_0^2}{2\sigma_0^2\sigma_1^2}. \tag{C.3}$$

Solving for y yields

$$0 = \left(\frac{\sigma_1^2 - \sigma_0^2}{2\sigma_0^2\sigma_1^2}\right) y^2 + \left(\frac{\mu_1}{\sigma_1^2}\right) y - \frac{\mu_1^2}{2\sigma_1^2} - \ln\left(2\frac{\sigma_1}{\sigma_0}\right), \tag{C.4}$$

$$y = \frac{-\frac{\mu_1}{\sigma_1^2} + \sqrt{\frac{\mu_1^2}{\sigma_1^4} + 4\left(\frac{\sigma_1^2 - \sigma_0^2}{2\sigma_0^2\sigma_1^2}\right)\left(\frac{\mu_1^2}{2\sigma_1^2} + \ln\left(2\frac{\sigma_1}{\sigma_0}\right)\right)}}{\left(\frac{\sigma_1^2 - \sigma_0^2}{\sigma_0^2\sigma_1^2}\right)}. \tag{C.5}$$

Bibliography

[AG+18] N. Abadi, A. A. Gohari, M. Mirmohseni, and M. Nasiri-Kenari, "Zero-error codes for multi-type molecular communication in random delay channel," in *Proc. Iran Workshop on Commun. Inf. Theory (IWCIT)*, Tehran, Iran, Apr. 2018, pp. 1–6.

[AL+19] M. Abbaszadeh, W. Li, L. Lin, I. White, P. Denissenko, P. J. Thomas, and W. Guo, "Mutual information and noise distributions of molecular signals using laser induced fluorescence," in *Proc. IEEE Global Communications Conf. (GLOBECOM)*, Waikoloa, HI, USA, Dec. 2019, pp. 1–6.

[Ago92] W. C. Agosta, *Chemical Communication: The Language of Pheromones*. Henry Holt and Company, 1992.

[Agr12] G. P. Agrawal, *Fiber-Optic Communication Systems*. John Wiley & Sons, 2012.

[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.

[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.

[Alb10] B. Alberts, "Cell biology: The endless frontier," *Molecular Biology of the Cell*, vol. 21, no. 22, pp. 3785–3785, Nov. 2010.

[AB+15] B. Alberts, D. Bray, K. Hopkin, A. D. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter, *Essential Cell Biology*. Garland Science, 2015.

[ES+06] J. El-Ali, P. K. Sorger, and K. F. Jensen, "Cells on chips," *Nature*, vol. 442, no. 7101, pp. 403–411, Jul. 2006.

[AC04] T. M. Allen and P. R. Cullis, "Drug delivery systems: Entering the mainstream," *Science*, vol. 303, no. 5665, pp. 1818–1822, Mar. 2004.

[AC16] J. D. Allison and R. T. Cardé, *Pheromone Communication in Moths: Evolution, Behavior, and Application*. University of California Press Oakland, 2016.

[AK+18] G. Ardelt, C. Külls, and H. Hellbrück, "Towards intrinsic molecular communication using isotopic isomerism," *Open Journal of Internet Of Things (OJIOT)*, vol. 4, no. 1, pp. 135–143, 2018.

[AB86] H. Aref and S. Balachandar, "Chaotic advection in a Stokes flow," *The Physics of Fluids*, vol. 29, no. 11, pp. 3515–3521, Nov. 1986.

[AA+17] A. A. Arnab, S. S. Afrin, F. Fahad, and H. U. Zaman, "A cost effective way to build a web controlled search and CO detector rover," in *Proc. IEEE Annual Computing and Communication Workshop and Conference (CCWC)*, Las Vegas, NV, USA, Jan. 2017, pp. 1–6.

[AA+12] B. Atakan, O. Akan, and S. Balasubramaniam, "Body area nanonetworks with molecular communications in nanomedicine," *IEEE Commun. Mag.*, vol. 50, no. 1, pp. 28–34, Jan. 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.

[Avh18] P. Avhale, "Advection diffusion equation for nutrient uptake by aquatic plant root with nonlinear boundary condition," *Int. J. Math. Trends Technol*, vol. 53, pp. 90–103, Jan. 2018.

[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.

[BB+11] S. Balasubramaniam, N. T. Boyle, A. Della-Chiesa, F. Walsh, A. Mardinoglu, D. Botvich, and A. Prina-Mello, "Development of artificial neuronal networks for molecular communication," *Nano Communication Networks*, vol. 2, no. 2-3, pp. 150–160, Jun. 2011.

[BS+15] D. Barkley, B. Song, V. Mukund, G. Lemoult, M. Avila, and B. Hof, "The rise of fully turbulent flow," *Nature*, vol. 526, no. 7574, pp. 550–553, Oct. 2015.

[Bar11] E. M. Barrows, *Animal Behavior Desk Reference: A Dictionary of Animal Behavior, Ecology, and Evolution*, 3rd ed. CRC Press, 2011.

[BF+20] M. Bartunik, M. Fleischer, W. Haselmayr, and J. Kirchner, "Colour-specific microfluidic droplet detection for molecular communication," in *Proc. ACM Int. Conf. on Nanoscale Computing and Communication (NANOCOM)*, Virtual Conference, Sep. 2020, pp. 1–6.

[BS+21] M. Bartunik, M. Streb, H. Unterweger, J. Haller, and J. Kirchner, "Increasing the channel capacity: Parallel data transmission in a testbed for molecular communication," in *Proc. ACM Int. Conf. on Nanoscale Computing and Communication (NANOCOM)*, Virtual Conference, Sep. 2021, pp. 1–6.

[Ber93] H. C. Berg, *Random Walks in Biology*. Princeton University Press, 1993.

[BS10] J. Berthier and P. Silberzan, *Microfluidics for Biotechnology*. Artech House, 2010.

[BG+92] D. P. Bertsekas, R. G. Gallager, and P. Humblet, *Data Networks*. Prentice-Hall International New Jersey, 1992, vol. 2.

[BD+20a] S. Bhattacharjee, M. Damrath, F. Bronner, L. Stratmann, J. P. Drees, F. Dressler, and P. A. Hoeher, "A testbed and simulation framework for air-based molecular communication using fluorescein," in *Proc. ACM Int. Conf. on Nanoscale Computing and Communication (NANOCOM)*, Virtual Conference, Sep. 2020, pp. 1–6.

[BD+20b] S. Bhattacharjee, M. Damrath, and P. A. Hoeher, "Design of macroscopic air-based molecular communication concept using fluorescein," in *Proc. ACM Int. Conf. on Nanoscale Computing and Communication (NANOCOM)*, Virtual Conference, Sep. 2020, pp. 1–2.

[BD+19] ——, "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.

[BD+22a] S. Bhattacharjee, M. Damrath, and P. A. Hoeher, "Channel coding techniques in macroscopic air-based molecular communication," in *Proc. ACM Int. Conf. on Nanoscale Computing and Communication (NANOCOM)*, Barcelona, Spain, Oct. 2022, pp. 1–2.

[BD+22b] S. Bhattacharjee, M. Damrath, L. Stratmann, P. A. Hoeher, and F. Dressler, "Digital communication techniques in macroscopic air-based molecular communication," *IEEE Trans. Mol. Biol. Multi-Scale Commun.*, vol. 8, no. 4, pp. 276–291, Dec. 2022.

[BK+23] S. Bhattacharjee, E. B. Krebs, A. Harlakin, and P. A. Hoeher, "Detection process in macroscopic air-based molecular communication using a PIN photodiode," *IEEE Trans. Mol. Biol. Multi-Scale Commun.*, vol. 9, no. 1, pp. 13–17, Mar. 2023.

[BA+21] D. Bi, A. Almpanis, A. Noel, Y. Deng, and R. Schober, "A survey of molecular communication in cell biology: Establishing a new hierarchy for interdisciplinary applications," *IEEE Communications Surveys and Tutorials*, vol. 23, no. 3, pp. 1494–1545, Mar. 2021.

[Bog08] R. Bogue, "Nanosensors: A review of recent progress," *Sensor Review*, vol. 28, no. 1, pp. 12–17, Jan. 2008.

[BD21] L. Bollinger and R. Dickie, "Low-cost fluorescent microvascular visualization in ambystoma mexicanum," *Microscopy and Microanalysis*, vol. 27, no. S1, pp. 3160–3162, Jul. 2021.

[BC14] L. Bortolotti and C. Costa, "Chemical communication in the honey bee society," *Neurobiology of Chemical Commun.*, 2014.

[CR06] N. A. Campbell and J. B. Reece, *Biología*. Ed. Médica Panamericana, 2006.

[CC+18] S. Catola, M. Centritto, P. Cascone, A. Ranieri, F. Loreto, L. Calamai, R. Balestrini, and E. Guerrieri, "Effects of single or combined water deficit and aphid attack on tomato volatile organic compound (VOC) emission and plant-plant communication," *Environmental and Experimental Botany*, vol. 153, pp. 54–62, Sep. 2018.

[CH+06] A. Cavalcanti, T. Hogg, B. Shirinzadeh, and H. C. Liaw, "Nanorobot communication techniques: A comprehensive tutorial," in *Proc. Int. Conf. on Control, Automation, Robotics and Vision*, Dec. 2006, pp. 1–6.

[CS+09] A. Cavalcanti, B. Shirinzadeh, T. Fukuda, and S. Ikeda, "Nanorobot for brain aneurysm," *The International Journal of Robotics Research*, vol. 28, no. 4, pp. 558–570, Apr. 2009.

[CA18] D. Chakravorty and S. M. Assmann, "G protein subunit phosphorylation as a regulatory mechanism in heterotrimeric G protein signaling in mammals, yeast, and plants," *Biochemical Journal*, vol. 475, no. 21, pp. 3331–3357, Nov. 2018.

[CR+19] J. M. Chalissery, A. Renyard, R. Gries, D. Hoefele, S. K. Alamsetti, and G. Gries, "Ants sense, and follow, trail pheromones of ant community members," *Insects*, vol. 10, no. 11, p. 383, Nov. 2019.

[CC+19] K. Charles-Darby, R. Carrasco-Alvarez, and R. Parra-Michel, "Complexity reduction of mlse and map equalizers using modified prolate basis expansion," *Electronics*, vol. 8, no. 11, p. 1333, 2019.

[CY+18] C. Chen, Y. Yin, Q. Tu, and H. Yang, "Glucose and amino acid in enterocyte: Absorption, metabolism and maturation," *Front. Biosci.*, vol. 23, pp. 1721–1739, Mar. 2018.

[CL08] H. Cheng and W. Lederer, "Calcium sparks," *Physiological Reviews*, vol. 88, no. 4, pp. 1491–1545, Oct. 2008.

[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, May 2017.

[Cob08] P. Cobley, "Communication: Definitions and concepts," *The Int. Enc. of Comm.*, Jun. 2008.

[CG+09a] M. Cole, J. W. Gardner, S. Pathak, T. C. Pearce, and Z. Rácz, "Towards a biosynthetic infochemical communication system," *Procedia Chemistry*, vol. 1, no. 1, pp. 305–308, Sep. 2009.

[CG+09b] M. Cole, J. Gardner, Z. Ráczu, S. Pathak, T. Pearce, J. Challiss, D. Markovic, A. Guerrero, L. Muñoz, G. Carot, *et al.*, "Biomimetic insect infochemical communication system," in *Proc. IEEE Sensors*, Christchurch, New Zealand, Oct. 2009, pp. 1358–1361.

[CT04] T. C. Collier and C. Taylor, "Self-organization in sensor networks," *Journal of Parallel and Distributed Computing*, vol. 64, no. 7, pp. 866–873, Jul. 2004.

[CH07] G. M. Cooper and R. E. Hausman, *The Cell: A Molecular Approach*. ASM Press Washington, DC, 2007, vol. 4.

[CS+99] J. W. Costerton, P. S. Stewart, and E. P. Greenberg, "Bacterial biofilms: A common cause of persistent infections," *Science*, vol. 284, no. 5418, pp. 1318–1322, May 1999.

[Cra79] J. Crank, *The Mathematics of Diffusion*. Oxford University Press, 1979.

[Csa73] G. T. Csanady, *Turbulent Diffusion in the Environment*. Springer Science & Business Media, 1973.

[Cus09] E. L. Cussler, *Diffusion: Mass Transfer in Fluid Systems*. Cambridge University Press, 2009.

[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.

[DK+17] 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.

[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.

[Dam20] M. Damrath, "Channel Coding in Molecular Communication," Ph.D. dissertation, Kiel University, 2020.

[DB+21a] M. Damrath, S. Bhattacharjee, and P. A. Hoeher, "Application of MIMO techniques in macroscopic air-based molecular communication," in *Proc. ACM Int. Conf. on Nanoscale Computing and Communication (NANOCOM)*, Virtual Conference, Sep. 2021, pp. 1–2.

[DB+21b] ——, "Investigation of multiple fluorescent dyes in macroscopic air-based molecular communication," *IEEE Trans. Mol. Biol. Multi-Scale Commun.*, vol. 7, no. 2, pp. 78–82, Jun. 2021.

[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)*, IEEE, Dec. 2018, pp. 1–5.

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

[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.

[DG+12] E. De Leo, L. Galluccio, A. Lombardo, and G. Morabito, "Networked labs-on-a-chip (NLoC): Introducing networking technologies in microfluidic systems," *Nano Communication Networks*, vol. 3, no. 4, pp. 217–228, Dec. 2012.

[dAR+12] A. D. C. de Albornoz, A. B. Rodríguez, A. L. Lopez, and A. R. G. Ramirez, "A microcontroller-based mobile robotic platform for odor detection," in *Proc. ISSNIP Biosignals and Biorobotics Conference: Biosignals and Robotics for Better and Safer Living (BRC)*, Jan. 2012, pp. 1–6.

[Dem06] A. J. Demello, "Control and detection of chemical reactions in microfluidic systems," *Nature*, vol. 442, no. 7101, pp. 394–402, Jul. 2006.

[DC+07] S. P. Diggle, S. A. Crusz, and M. Cámara, "Quorum sensing," *Current Biology*, vol. 17, no. 21, R907–R910, Nov. 2007.

[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.

[DB+12] S. M. Douglas, I. Bachelet, and G. M. Church, "A logic-gated nanorobot for targeted transport of molecular payloads," *Science*, vol. 335, no. 6070, pp. 831–834, Feb. 2012.

[ED+19] M. Egan, T. Q. Duong, and M. Di Renzo, "Biological circuits for detection in MoSK-based molecular communication," *IEEE Access*, vol. 7, pp. 21 094–21 102, Feb. 2019.

[ES+12] A. Einolghozati, M. Sardari, and F. Fekri, "Collective sensing-capacity of bacteria populations," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Jul. 2012, pp. 2959–2963.

[EW12] J. Elbaz and I. Willner, "DNA origami: Nanorobots grab cellular control," *Nature Materials*, vol. 11, no. 4, pp. 276–277, Apr. 2012.

[Elb08] B. R. Elbert, *Introduction to Satellite Communication*. Artech House, 2008.

[EG+18] A. Erofeev, P. Gorelkin, A. Garanina, A. Alova, M. Efremova, N. Vorobyeva, C. Edwards, Y. Korchev, and A. Majouga, "Novel method for rapid toxicity screening of magnetic nanoparticles," *Scientific Reports*, vol. 8, no. 1, pp. 1–11, May 2018.

[FH+19] H. Fan, V. Hernandez Bennetts, E. Schaffernicht, and A. J. Lilienthal, "Towards gas discrimination and mapping in emergency response scenarios using a mobile robot with an electronic nose," *Sensors*, vol. 19, no. 3, p. 685, Feb. 2019.

[FS+08] Q. Fang, S. Sakadžić, L. Ruvinskaya, A. Devor, A. M. Dale, and D. A. Boas, "Oxygen advection and diffusion in a three-dimensional vascular anatomical network," *Optics Express*, vol. 16, no. 22, pp. 17 530–17 541, Oct. 2008.

[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, Feb. 2016.

[Far15] N. Farsad, "Molecular Communication: From Theory to Practice," Ph.D. dissertation, York University, 2015.

[FG16] N. Farsad and A. Goldsmith, "A molecular communication system using acids, bases and hydrogen ions," in *Proc. IEEE Signal Processing Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Edinburgh, UK, Jul. 2016, pp. 1–6.

[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.

[FK08] K. Fazel and S. Kaiser, *Multi-Carrier and Spread Spectrum Systems: From OFDM and MC-CDMA to LTE and WiMAX*. John Wiley & Sons, 2008.

[Fis11] H. Fischer, *A History of the Central Limit Theorem: From Classical to Modern Probability Theory*. Springer, 2011.

[For73] G. D. Forney, "The Viterbi algorithm," *Proc. IEEE*, vol. 61, no. 3, pp. 268–278, Mar. 1973.

[Fre99] R. A. Freitas, *Nanomedicine, Volume I: Basic Capabilities*. Landes Bioscience Georgetown, TX, 1999, vol. 1.

[FO+00] Y. Fukuda, I. Ohashi, K. Hanafusa, T. Nakagawa, S.-i. Ohtani, Y. An-naka, T. Hayashi, and H. Shibuya, "Anisotropic diffusion in kidney: Apparent diffusion coefficient measurements for clinical use," *J Magn Reson Imaging*, vol. 11, no. 2, pp. 156–160, Feb. 2000.

[GR13] M. Gagliano and M. Renton, "Love thy neighbour: Facilitation through an alternative signalling modality in plants," *BMC Ecology*, vol. 13, no. 1, pp. 1–6, Dec. 2013.

[GB99] J. W. Gardner and P. N. Bartlett, *Electronic Noses: Principles and Applications*. Oxford University Press, 1999.

[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.

[Gav19] H. P. Gavin, "The Levenberg-Marquardt algorithm for nonlinear least squares curve-fitting problems," Tech. Rep., Aug. 2019.

[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.

[Gib18] J. D. Gibson, *The communications handbook*. CRC Press, 2018.

[GF+05] K. Grammer, B. Fink, and N. Neave, "Human pheromones and sexual attraction," *European Journal of Obstetrics & Gynecology and Reproductive Biology*, vol. 118, no. 2, pp. 135–142, Feb. 2005.

[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.

[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.

[Guh08] A. Guha, "Transport and deposition of particles in turbulent and laminar flow," *Annu. Rev. Fluid Mech.*, vol. 40, pp. 311–341, Jan. 2008.

[GA21a] F. Gulec and B. Atakan, "A droplet-based signal reconstruction approach to channel modeling in molecular communication," *IEEE Trans. Mol. Biol. Multi-Scale Commun.*, vol. 7, no. 1, pp. 64–68, Mar. 2021.

[GA21b] ——, "A molecular communication perspective on airborne pathogen transmission and reception via droplets generated by coughing and sneezing," *IEEE Trans. Mol. Biol. Multi-Scale Commun.*, vol. 7, no. 3, pp. 175–184, May 2021.

[GA20] ——, "Distance estimation methods for a practical macroscale molecular communication system," *Nano Communication Networks*, vol. 24, p. 100 300, May 2020.

[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.

[GA+20] W. Guo, I. Atthanayake, and P. Thomas, "Vertical underwater molecular communications via buoyancy: Gaussian velocity distribution of signal," in *Proc. IEEE Int. Conf. on Communications (ICC)*, Dublin, Ireland, Jun. 2020, pp. 1–6.

[GB+19a] M. C. Gursoy, E. Basar, A. E. Pusane, and T. Tugcu, "Pulse position-based spatial modulation for molecular communications," *IEEE Commun. Lett.*, vol. 23, no. 4, pp. 596–599, Apr. 2019.

[GB+19b] 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.

[HM+19] M. Hamidović, U. Marta, G. Fink, R. Wille, A. Springer, and W. Haselmayr, "Information encoding in droplet-based microfluidic systems: First practical study," in *Proc. ACM Int. Conf. on Nanoscale Computing and Communication (NANOCOM)*, Dublin, Ireland, Sep. 2019, pp. 1–6.

[Ham50] R. W. Hamming, "Error detecting and error correcting codes," *Bell Syst. Technical J.*, vol. 29, no. 2, pp. 147–160, Apr. 1950.

[HC+03] D. Harel, L. Carmel, and D. Lancet, "Towards an odor communication system," *Computational Biology and Chemistry*, vol. 27, no. 2, pp. 121–133, May 2003.

[HS+19] W. Haselmayr, A. Springer, G. Fischer, C. Alexiou, H. Boche, P. A. Hoeher, F. Dressler, and R. Schober, "Integration of molecular communications into future generation wireless networks," in *Proc. 6G Wireless Summit*, Lapland, Finland, Mar. 2019.

[HN+11] M. Hirabayashi, A. Nishikawa, F. Tanaka, M. Hagiya, H. Kojima, and K. Oiwa, "Design of molecular-based network robots-toward the environmental control," in *Proc. IEEE Int. Conf. on Nano.*, Portland, OR, USA, Aug. 2011, pp. 313–318.

[HM+00] A. Hirata, S.-I. Matsuyama, and T. Shiozawa, "Temperature rises in the human eye exposed to EM waves in the frequency range 0.6-6 GHz," *IEEE Trans. on Electro. Comp.*, vol. 42, no. 4, pp. 386–393, Nov. 2000.

[HM+08] S. Hiyama, Y. Moritani, and T. Suda, "Molecular transport system in molecular communication," *NTT DOCOMO Technical Journal*, vol. 10, no. 3, pp. 49–53, Dec. 2008.

[Hoe19] P. A. Hoeher, *Visible Light Communications: Theoretical and Practical Foundations*. Carl Hanser Verlag GmbH Co KG, 2019.

[HD+21] P. A. Hoeher, M. Damrath, S. Bhattacharjee, and M. Schurwanz, "On mutual information analysis of infectious disease transmission via particle propagation," *IEEE Trans. Mol. Biol. Multi-Scale Commun.*, Oct. 2021.

[Höh13] P. A. Höher, *Grundlagen der Digitalen Informationsübertragung*, 2nd Edition. Springer, 2013.

[HH03] K.-A. Hossmann and D. Hermann, "Effects of electromagnetic radiation of mobile phones on the central nervous system," *Bioelectromagnetics*, vol. 24, no. 1, pp. 49–62, Jan. 2003.

[HX+10] L. Hua, L. Xie, and H. Chen, "A color interpolation algorithm for Bayer pattern digital cameras based on green components and color difference space," in *Proc. IEEE Int. Conf. Prog. Inf. Comput.*, Shanghai, China, Dec. 2010, pp. 791–795.

[HW+19a] 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.

[HW+19b] Y. Huang, M. Wen, L.-L. Yang, C.-B. Chae, X. Chen, and Y. Tang, "Space shift keying for molecular communication: Theory and experiment," in *Proc. IEEE Global Communications Conf. (GLOBECOM)*, Waikoloa, HI, USA, Dec. 2019, pp. 1–6.

[HV07] W. Hundsdorfer and J. G. Verwer, *Numerical Solution of Time-Dependent Advection-Diffusion-Reaction Equations*. Springer Science & Business Media, 2007, vol. 33.

[IT92] T. Igarashi and H. Takasaki, "Fluid flow around three rectangular blocks in a flat-plate laminar boundary layer," *Experimental Heat Transfer*, vol. 5, no. 1, pp. 17–31, Jan. 1992.

[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.

[JD+84] I. Javandel, L. Doughty, and C. Tsang, *Groundwater Transport: Handbook of Mathematical Models*. American Geophysical Union, 1984, vol. 10.

[JB68] E. C. Jordan and K. G. Balmain, *Electromagnetic Waves and Radiating Systems*. Prentice-Hall, 1968.

[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.

[KA09] S. Kadloor and R. Adve, "A framework to study the molecular communication system," in *Proc. Int. Conf. Comp. Commun. Networks*, Aug. 2009, pp. 1–6.

[Kal18] J. Kalirai, "Scientific discovery with the James Webb space telescope," *Contemporary Physics*, vol. 59, no. 3, pp. 251–290, Jul. 2018.

[Kas18] W. Kass, *Tracing Technique in Geohydrology*. Routledge, 2018.

[KO+98] C. Kerr, K. Osborn, G. Robson, and P. Handley, "The relationship between pipe material and biofilm formation in a laboratory model system," *Journal of Applied Microbiology*, vol. 85, no. S1, 29S–38S, Dec. 1998.

[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.

[KA+19] M. Khalid, O. Amin, S. Ahmed, B. Shihada, and M.-S. Alouini, "Communication through breath: Aerosol transmission," *IEEE Commun. Mag.*, vol. 57, no. 2, pp. 33–39, Feb. 2019.

[KA+20a] ——, "Modeling of viral aerosol transmission and detection," *IEEE Trans. Commun.*, vol. 68, no. 8, pp. 4859–4873, May 2020.

[KD95] A. L. Kholodenko and J. F. Douglas, "Generalized Stokes-Einstein equation for spherical particle suspensions," *Physical Review E*, vol. 51, no. 2, p. 1081, Feb. 1995.

[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.

[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.

[KF+14] N.-R. Kim, N. Farsad, C.-B. Chae, and A. W. Eckford, "A realistic channel model for molecular communication with imperfect receivers," in *Proc. IEEE Int. Conf. on Communications (ICC)*, Sydney, NSW, Australia, Jun. 2014, pp. 3987–3992.

[KA+20b] A. O. Kislal, B. C. Akdeniz, C. Lee, A. E. Pusane, T. Tugcu, and C.-B. Chae, "ISI-mitigating channel codes for molecular communication via diffusion," *IEEE Access*, vol. 8, pp. 24 588–24 599, Jan. 2020.

[KP+18] A. O. Kışlal, A. E. Pusane, and T. Tuğcu, "A comparative analysis of channel coding for molecular communication," in *Proc. Signal Processing and Communications Applications Conference (SIU)*, Izmir, Turkey, May 2018, pp. 1–4.

[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.

[Koh89] D. Kohl, "Surface processes in the detection of reducing gases with SnO_2-based devices," *Sensors and Actuators*, vol. 18, no. 1, pp. 71–113, Jun. 1989.

[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.

[Kos82] N. Kosov, "Elementary kinetic theory of diffusion in gases," *Journal of Engineering Physics*, vol. 42, no. 2, pp. 181–192, 1982.

[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.

[KS18] B. Krishnaswamy and R. Sivakumar, "Amplitude-width encoding for error correction in bacterial communication networks," in *Proc. of ACM Int. Conf. on Nanoscale Computing and Communication (NANOCOM)*, Reykjavik, Iceland, Sep. 2018, pp. 1–7.

[KB+21] M. N. Kumar, S. Biradar, and R. Babu, "Cell signaling and apoptosis in animals," in *Advances in Animal Genomics*, Elsevier, 2021, pp. 199–218.

[KP+20] S. Kumar, S. Prince, J. Venkata Aravind, and S. Kumar G, "Analysis on the effect of salinity in underwater wireless optical communication," *Marine Georesources & Geotechnology*, vol. 38, no. 3, pp. 291–301, Mar. 2020.

[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+20] M. Ş. Kuran, H. B. Yilmaz, I. Demirkol, N. Farsad, and A. Goldsmith, "A survey on modulation techniques in molecular communication via diffusion," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 1, pp. 7–28, Dec. 2020.

[KW+16] I. Kyrou, M. O. Weickert, S. Gharanei, H. S. Randeva, and B. K. Tan, "Fibroblast growth factors: New insights, new targets in the management of diabetes," *Minerva Endocrinologica*, vol. 42, no. 3, pp. 248–270, Jul. 2016.

[LF04] D. Large and J. Farmer, *Modern Cable Television Technology*. Elsevier, 2004.

[Lat95] B. P. Lathi, *Modern digital and analog communication systems*. Oxford University Press, Inc., 1995.

[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.

[LK+20a] C. Lee, B.-H. Koo, and C.-B. Chae, "In-vessel molecular MIMO communications," in *Proc. IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*, Seoul, South Korea, Apr. 2020, pp. 1–2.

[LK+20b] ——, "Molecular MIMO communications platform with BTSK for in-vessel network systems," in *International Conference on Bio-inspired Information and Communication Technologies*, Shanghai, China, Jul. 2020, pp. 289–293.

[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.

[LC01] S. Lin and D. J. Costello, *Error control coding*. Prentice Hall New York, 2001, vol. 2.

[LA+11] I. Llatser, E. Alarcón, and M. Pierobon, "Diffusion-based channel characterization in molecular nanonetworks," in *Proc. IEEE Conf. Comp. Commun. Workshops (INFOCOM WKSHPS)*, Shanghai, China, Apr. 2011, pp. 467–472.

[LB+23a] S. Lotter, L. Brand, V. Jamali, M. Schäfer, H. M. Loos, H. Unterweger, S. Greiner, J. Kirchner, C. Alexiou, D. Drummer, *et al.*, "Experimental research in synthetic molecular communications–part i: Overview and short-range systems," *IEEE Nanotech. Mag.*, vol. 17, no. 3, pp. 42–53, Apr. 2023.

[LB+23b] ——, "Experimental research in synthetic molecular communications–part ii: Long-range communication," *IEEE Nanotech. Mag.*, vol. 17, no. 3, pp. 54–65, Apr. 2023.

[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.

[Lya16] F. Lyall, *International Communications: The International Telecommunication Union and the Universal Postal Union*. Routledge, 2016.

[MM+10] M. U. Mahfuz, D. Makrakis, and H. T. Mouftah, "On the characterization of binary concentration-encoded molecular communication in nanonetworks," *Nano Communication Networks*, vol. 1, no. 4, pp. 289–300, Dec. 2010.

[MM+11] ——, "On the detection of binary concentration-encoded unicast molecular communication in nanonetworks," in *Proc. Int. Conf. on Bio-Inpsired Syst. Signal Process. (BIOSIGNALS)*, 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.

[MS09] T. E. Mallouk and A. Sen, "Powering nanorobots," *Scientific American*, vol. 300, no. 5, pp. 72–77, May 2009.

[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).

[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.

[MN+10] A. Marjovi, J. Nunes, P. Sousa, R. Faria, and L. Marques, "An olfactory-based robot swarm navigation method," in *Proc. IEEE Int. Conf. on Robotics and Automation*, Anchorage, AK, USA, May 2010, pp. 4958–4963.

[MS00] J. Mathieu and J. Scott, *An Introduction to Turbulent Flow*. Cambridge University Press, 2000.

[McG20] D. T. McGuiness, "Macro-Scale Molecular Communications," Ph.D. dissertation, The University of Liverpool (United Kingdom), 2020.

[MY+12] 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.

[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.

[MS+09] J. Mietzner, R. Schober, L. Lampe, W. H. Gerstacker, and P. A. Hoeher, "Multiple-antenna techniques for wireless communications-a comprehensive literature survey," *IEEE Commun. Surveys Tutorials*, vol. 11, no. 2, pp. 87–105, Jun. 2009.

[MB01] M. B. Miller and B. L. Bassler, "Quorum sensing in bacteria," *Annual Review of Microbiology*, vol. 55, no. 1, pp. 165–199, Oct. 2001.

[MD+12] L. Muñoz, N. Dimov, G. Carot-Sans, W. P. Bula, A. Guerrero, and H. J. G. E. Gardeniers, "Mimicking insect communication: Release and detection of pheromone, biosynthesized by an alcohol acetyl transferase immobilized in a microreactor," *PloS One*, vol. 7, no. 11, e47751, Nov. 2012.

[MY+95] B. R. Munson, D. F. Young, and T. H. Okiishi, "Fundamentals of fluid mechanics," *Oceanographic Literature Review*, vol. 10, no. 42, p. 831, 1995.

[NM+12] T. Nakano, M. Moore, F. Wei, A. Vasilakos, and J. Shuai, "Molecular communication and networking: Opportunities and challenges," *IEEE Trans. Nanobiosci.*, vol. 11, no. 2, pp. 135–148, Jun. 2012.

[NS+14] T. Nakano, T. Suda, Y. Okaie, M. Moore, and A. Vasilakos, "Molecular communication among biological nanomachines: A layered architecture and research issues," *IEEE Trans. Nanobiosci.*, vol. 13, no. 3, pp. 169–197, Sep. 2014.

[NE+13] T. Nakano, A. W. Eckford, and T. Haraguchi, *Molecular Communication*. Cambridge University Press, 2013.

[NB08] R. M. Negri and D. L. Bernik, "Tracking the sex pheromone of codling moth against a background of host volatiles with an electronic nose," *Crop Protection*, vol. 27, no. 10, pp. 1295–1302, Oct. 2008.

[Nob02] D. Noble, "Opinion: The rise of computational biology," *Nature Reviews Molecular Cell Biology*, vol. 3, no. 6, pp. 459–463, Jun. 2002.

[NU02] E. J. Noga and P. Udomkusonsri, "Fluorescein: A rapid, sensitive, nonlethal method for detecting skin ulceration in fish," *Veterinary Pathology*, vol. 39, no. 6, pp. 726–731, Nov. 2002.

[ON+14] Y. Okaie, T. Nakano, T. Hara, T. Obuchi, K. Hosoda, Y. Hiraoka, and S. Nishio, "Cooperative target tracking by a mobile bionanosensor network," *IEEE Trans. Nanobiosci.*, vol. 13, no. 3, pp. 267–277, Jul. 2014.

[OK+18] M. Ozmen, E. Kennedy, J. Rose, P. Shakya, J. K. Rosenstein, and C. Rose, "High speed chemical vapor communication using photoionization detectors," in *Proc. IEEE Global Communications Conf. (GLOBECOM)*, Abu Dhabi, United Arab Emirates, Dec. 2018, pp. 1–6.

[Pan07] N. K. Pandit, *Introduction to the Pharmaceutical Sciences*. Lippincott Williams & Wilkins, 2007.

[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.

[PM+08] M. Patil, D. S. Mehta, and S. Guvva, "Future impact of nanotechnology on medicine and dentistry," *Journal of Indian Society of Periodontology*, vol. 12, no. 2, p. 34, May 2008.

[Pes15] M. H. Pesaran, *Time Series and Panel Data Econometrics*. Oxford University Press, 2015.

[Pet12] J. D. Peters, *Speaking into the Air: A History of the Idea of Communication*. University of Chicago Press, 2012.

[PP+72] W. W. Peterson, W. Peterson, E. J. Weldon, and E. J. Weldon, *Error-Correcting Codes*. MIT Press, 1972.

[PR+15] S. Pfautsch, J. Renard, M. G. Tjoelker, and A. Salih, "Phloem as capacitor: Radial transfer of water into xylem of tree stems occurs via symplastic transport in ray parenchyma," *Plant Physiology*, vol. 167, no. 3, pp. 963–971, Mar. 2015.

[PS+16] T. Pichugina, T. Sugawara, A. Kaykov, W. Schierding, K. Masuda, J. Uewaki, R. Grand, J. Allison, R. Martienssen, P. Nurse, *et al.*, "A diffusion model for the coordination of DNA replication in Schizosaccharomyces pombe," *Scientific Reports*, vol. 6, no. 1, pp. 1–11, Jan. 2016.

[PM14] J. G. Proakis and D. G. Manolakis, *Digital Signal Processing*, 5th ed. Pearson Education, 2014.

[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, "Run-length aware hybrid modulation scheme for diffusion-based molecular communication," in *Proc. Int. Symp. on Comm. and Inf. Tech. (ISCIT)*, Incheon, South Korea, Sep. 2014, pp. 439–442.

[PR10] A. H. Purnamadjaja and R. A. Russell, "Bi-directional pheromone communication between robots," *Robotica*, vol. 28, no. 1, pp. 69–79, Jan. 2010.

[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.

[RK17] A. Rajput and M. Kumar, "Computational exploration of putative LuxR solos in archaea and their functional implications in quorum sensing," *Frontiers in Microbiology*, vol. 8, p. 798, May 2017.

[RS+20] A. Ranjan, H. Sahu, and P. Misra, "Modeling and measurements for wireless communication networks in underground mine environments," *Measurement*, vol. 149, p. 106 980, Jan. 2020.

[Red01] S. Redner, *A Guide to First-Passage Processes*. Cambridge University Press, 2001.

[Req03] A. A. Requicha, "Nanorobots, NEMS, and nanoassembly," *Proc. of the IEEE*, vol. 91, no. 11, pp. 1922–1933, Nov. 2003.

[RC09] V. H. Resh and R. T. Cardé, *Encyclopedia of Insects*. Academic Press, 2009.

[Rot90] N. Rott, "Note on the history of the Reynolds number," *Annual Review of Fluid Mechanics*, vol. 22, no. 1, pp. 1–12, Jan. 1990.

[RS19] S. R. Rouzegar and U. Spagnolini, "Diffusive MIMO molecular communications: Channel estimation, equalization, and detection," *IEEE Trans. Commun.*, vol. 67, no. 7, pp. 4872–4884, Jul. 2019.

[SG+12] P. Sarathi, R. Gurka, G. A. Kopp, and P. J. Sullivan, "A calibration scheme for quantitative concentration measurements using simultaneous PIV and PLIF," *Experiments in Fluids*, vol. 52, no. 1, pp. 247–259, Jan. 2012.

[SK00] K. Schulten and I. Kosztin, *Lectures in theoretical biophysics*, Univ. Illinois at Urbana–Champaign, Urbana, IL, USA, Apr. 2000.

[SH+21a] M. Schurwanz, P. A. Hoeher, S. Bhattacharjee, M. Damrath, L. Stratmann, and F. Dressler, "Duality between coronavirus transmission and air-based macroscopic molecular communication," *IEEE Trans. Mol. Biol. Multi-Scale Commun.*, vol. 7, no. 3, pp. 200–208, Apr. 2021.

[SH+21b] ——, "Infectious disease transmission via aerosol propagation from a molecular communication perspective: Shannon meets coronavirus," *IEEE Commun. Mag.*, vol. 59, no. 5, pp. 40–46, May 2021.

[SI+12] S. Sengupta, M. E. Ibele, and A. Sen, "Fantastic voyage: Designing self-powered nanorobots," *Angewandte Chemie International Edition*, vol. 51, no. 34, pp. 8434–8445, Aug. 2012.

[Sha48] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Technical J.*, vol. 27, no. 3, pp. 379–423, Jul. 1948.

[ST+07] A. L. Sherborne, M. D. Thom, S. Paterson, F. Jury, W. E. Ollier, P. Stockley, R. J. Beynon, and J. L. Hurst, "The genetic basis of inbreeding avoidance in house mice," *Current Biology*, vol. 17, no. 23, pp. 2061–2066, Dec. 2007.

[SL+12] P. J. Shih, C. H. Lee, and P. C. Yeh, "Channel codes for mitigating inter-symbol 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.

[SK98] D.-s. Shiu and J. M. Kahn, "Differential pulse position modulation for power-efficient wireless infrared communication," in *Proc. IEEE Global Communications Conf. (GLOBECOM)*, Sydney, NSW, Australia, Nov. 1998, pp. 219–224.

[SK99] ——, "Differential pulse-position modulation for power-efficient optical communication," *IEEE Trans. Commun.*, vol. 47, no. 8, pp. 1201–1210, Aug. 1999.

[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.

[SS+22] M. Singh, M. L. Singh, R. Singh, H. Kaur, P. Chohan, and S. Kaur, "Comprehensive experimental and statistical analysis of the effects of salinity gradient, temperature gradient, air bubbles on the performance of underwater wireless optical communication link," *Journal of Modern Optics*, vol. 69, no. 17, pp. 978–994, Aug. 2022.

[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.

[SH+10] F. Stajano, N. Hoult, I. Wassell, P. Bennett, C. Middleton, and K. Soga, "Smart bridges, smart tunnels: Transforming wireless sensor networks from research prototypes into robust engineering infrastructure," *Ad Hoc Networks*, vol. 8, no. 8, pp. 872–888, Nov. 2010.

[SS76] P. D. Stein and H. N. Sabbah, "Turbulent blood flow in the ascending aorta of humans with normal and diseased aortic valves," *Circulation Research*, vol. 39, no. 1, pp. 58–65, Jul. 1976.

[Ste92] J. Steuer, "Defining virtual reality: Dimensions determining telepresence," *Journal of Comm.*, vol. 42, no. 4, pp. 73–93, Dec. 1992.

[SS+90] J. B. Stock, A. M. Stock, and J. M. Mottonen, "Signal transduction in bacteria," *Nature*, vol. 344, no. 6265, pp. 395–400, Mar. 1990.

[SP09] M. Stojanovic and J. Preisig, "Underwater acoustic communication channels: Propagation models and statistical characterization," *IEEE Commun. Mag.*, vol. 47, no. 1, pp. 84–89, Jan. 2009.

[SG+21] L. Stratmann, J. T. Gómez, S. Bhattacharjee, M. Damrath, P. A. Hoeher, and F. Dressler, "Impact of mobility on air-based macroscopic molecular communication: A simulation study," in *Proc. ACM Int. Conf. on Nanoscale Computing and Communication (NANOCOM)*, Virtual Conference, Sep. 2021, pp. 1–2.

[SK+12] M. Suchański, P. Kaniewski, R. Matyszkiel, and P. Gajewski, "Dynamic spectrum management in legacy military communication systems," in *Proc. IEEE Mil. Comm. and Inf. Sys. Conf. (MCC)*, Gdansk, Poland, Oct. 2012, pp. 1–5.

[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. Genetic and Evol. Comp. Conf. (GECCO)*, Washington, DC, USA, Jun. 2005, pp. 1–5.

[SN89] H. Sugiyama and K. Nosu, "MPPM: A method for improving the band-utilization efficiency in optical PPM," *Journal of Lightwave Technology*, vol. 7, no. 3, pp. 465–472, Mar. 1989.

[SZ09] M. Sui and Z. Zhou, "The modified PPM modulation for underwater wireless optical communication," in *Int. Conf. Commun. Softw. and Netw.*, Chengdu, China, Feb. 2009, pp. 173–177.

[TB06] A. Tang and B. S. Bordoni, *Internet*. StatPearls Publishing, 2006.

[Tay53] G. I. Taylor, "Dispersion of soluble matter in solvent flowing slowly through a tube," *Proc. R. Soc. Lond. A*, vol. 219, no. 1137, pp. 186–203, Aug. 1953.

[TK+04] J. Timmis, T. Knight, L. N. de Castro, and E. Hart, "An overview of artificial immune systems," *Computation in Cells and Tissues*, pp. 51–91, 2004.

[TD+13] P. Tiwari, S. Dwivedi, M. P. Singh, R. Mishra, and A. Chandy, "Basic and modern concepts on cholinergic receptor: A review," *Asian Pacific Journal of Tropical Disease*, vol. 3, no. 5, pp. 413–420, Oct. 2013.

[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.

[UK+12] H. Ueda, Y. Kikuta, and K. Matsuda, "Plant communication: Mediated by individual or blended VOCs?" *Plant Signaling and Behavior*, vol. 7, no. 2, pp. 222–226, Feb. 2012.

[UA+17] S. Ulfah, S. Awalludin, *et al.*, "Advection-diffusion model for the simulation of air pollution distribution from a point source emission," in *J. Phys.: Conf. Ser*, vol. 948, Jakarta, Indonesia, Oct. 2017, p. 012 067.

[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.

[VD+18] D. D. Van der Voort, N. J. Dam, H. J. Clercx, and W. v. d. Water, "Characterization of spray-induced turbulence using fluorescence PIV," *Experiments in Fluids*, vol. 59, no. 7, pp. 1–7, Jul. 2018.

[VW+16] M. Van Wolferen, A. Wagner, C. Van Der Does, and S.-V. Albers, "The archaeal Ced system imports DNA," *Proc. Nat. Acad. Sci.*, vol. 113, no. 9, pp. 2496–2501, Mar. 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.

[WH+20] J. Wang, D. Hu, C. Shetty, and H. Hassanieh, "Understanding and embracing the complexities of the molecular communication channel in liquids," in *Proc. ACM Int. Conf. Mobile Comput. Netw.*, Apr. 2020, pp. 1–15.

[WS+21] Y.-Y. Wang, W.-J. Sun, D.-X. Yan, K. Dai, and Z.-M. Li, "Ultralight carbon nanotube/graphene/polyimide foam with heterogeneous interfaces for efficient electromagnetic interference shielding and electromagnetic wave absorption," *Carbon*, vol. 176, pp. 118–125, May 2021.

[WK00] R. Weiss and T. F. Knight, "Engineered communications for microbial robotics," in *Int. Works. on DNA-Based Comp.*, Springer, Jun. 2000, pp. 1–16.

[Whe47] H. A. Wheeler, "Fundamental limitations of small antennas," *Proc. of the IRE*, vol. 35, no. 12, pp. 1479–1484, Dec. 1947.

[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.

[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.

[WC08] R. J. Wordinger and A. F. Clark, "Growth factors and neurotrophic factors as targets," in *Ocular Therapeutics*, Elsevier, 2008, pp. 87–116.

[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.

[YE+06] P. Yager, T. Edwards, E. Fu, K. Helton, K. Nelson, M. R. Tam, and B. H. Weigl, "Microfluidic diagnostic technologies for global public health," *Nature*, vol. 442, no. 7101, pp. 412–418, Jul. 2006.

[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.

[YW+16] H.-H. Yen, X. Wang, and D. Wang, "QoS aware molecular activation and communication scheme in molecular nanoscale sensor networks," in *Proc. IEEE Int. Conf. on e-Health Net., App. and Serv. (HEALTHCOM)*, Munich, Germany, Sep. 2016, pp. 1–6.

[YH+14] H. B. Yilmaz, A. C. 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.

[Yok67] S. Yokosi, "The structure of river turbulence," *Bulletin of the Disaster Prevention Research Institute*, vol. 17, no. 2, pp. 1–29, Oct. 1967.

[ZH05] A. Zamanian and C. Hardiman, "Electromagnetic radiation and human health: A review of sources and effects," *High Freq. Electron.*, vol. 4, no. 3, pp. 16–26, Jul. 2005.