MOLECULAR COMMUNICATION: FROM THEORY TO PRACTICE

NARIMAN FARSAD

A DISSERTATION SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILMENT OF THE REQUIREMENTS
FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN ELECTRICAL ENGINEERING AND COMPUTER SCIENCE
YORK UNIVERSITY
TORONTO, ONTARIO
JULY 2015

Abstract

Always-on, always-available digital communication has changed the world allowing us to collaborate and share information in ways unimaginable not long ago. Yet many of the physical principles used in everyday digital communication break down as the size of the devices approach micro- or nano-scale dimensions. As a result, tiny devices, with dimensions of microns or less, need to do something different in order to communicate. Moreover, at meter scales there are areas where use of radio signals is not possible or desirable. An emerging biomimetic technique called molecular communication, which relies on chemical signaling is a promising solution to these problems. Although biologists have studied molecular communication extensively, it is very poorly understood from a telecommunication engineering perspective.

Engineering molecular communication systems is important since micro- and nano-scale systems are the key to unlocking a realm of futuristic possibilities such as: self-repairing machines, micro- and nano-scale robotics, synthetic biological devices, nanomedicine, and artificial immune systems that detect and kill cancer

cells and pathogens. All these transformative applications have one feature in common: they involve not just single devices working independently, but swarms of devices working in concert. Besides solving the communication problem at small scales, use of molecular communication in areas such as robotics, and infrastructure monitoring can unlock new applications in smart cities and disaster search and rescue.

In this dissertation, after providing a comprehensive survey of the field, two areas of study with high potential impact are identified: on-chip molecular communication, and experimental platforms for molecular communication. First, on-chip molecular communication is investigated towards the goal of networking components within lab-on-chip devices and point-of-care diagnostic devices. This has numerous applications in medicine, environmental monitoring systems, and the food industry. Then in the second part of the dissertation, a tabletop demonstrator for molecular communication is designed and built that could be used for research and experimentation. In particular, no macroscale or microscale molecular communication platform capable of reliably transporting sequential data had existed in the past, and this platform is used to send the world's first text message ("O Canada") using chemical signals.

Acknowledgements

This thesis could have not been completed without the encouragement, collaboration, and support of a tremendous number of people.

First and foremost, I would like to extend my sincere thanks to my advisor, Professor Andrew W. Eckford for his encouragement, support, and patience throughout my Ph.D. studies. I would definitely not reach this point without his support and encouragements. It was a privilege for me to work with an extraordinary supervisor like him.

I would like to extend my sincere appreciations to my supervisory committee members, Profs. Gerald Audette and Uyen Trang Nguyen for their great personality, friendly advice and valuable guidance. I wish to thank my thesis defense committee members, Profs. Steve Hranilovic, Ebrahim Ghafar-Zadeh and Sushanta Mitra for accepting to be members of the committee. Their insight and thoughtful comments during the defense have resulted in a manuscript that is clearer, more compelling, and much improved.

A special thanks to my family. Words can not express how grateful I am to my mother, father, and my brother for all the sacrifices that they have made for me throughout the years.

My special thanks goes to my dear wife, Kathy, for her unlimited love and endless support. Kathy has supported me constantly during all these years, and without her support I would not have been as successful during my studies.

I would also like to thank my fellow collaborators, Na-Rae Kim, Song Qiu, Drs. Satoshi Hiyama, Yuki Moritani, H. Birkan Yilmaz, Siyi Wang, Profs. Chan-Byoung Chae, and Weisi Guo for all the help and support they provided in this multidisciplinary research work.

Finally, I would like to thank friends and colleagues, Yousef Akhavan, Arash Mohammadi, Mohammad Sajjadieh, Kayvan Tirdad, Mohammad Reza Faghani, Alireza Moghadam, Hossein Pourmodheji, Mehdi Kargar, Hooman Baradaran, LinChen Wang, and Lida Jabbari for their help and support.

Table of Contents

List of Tables

List of Figures

1 Dissertation Overview

1.1 Introduction

The problem of conveying information over a distance has always been an important part of human society. Today, modern telecommunication systems provide an always-on, always-available access to information through use of electrical or electromagnetic (EM) signals. Using these technologies, it is now possible collaborate and share information in ways that were unimaginable just a few decades ago. Yet despite all these achievements, the techniques that are used for communication, such as sending a current down a wire, or using an EM wave in free space, break down in many communication scenarios. For example, there are some macroscale environments that are hostile to EM wave propagations, such as inside networks of tunnels, pipelines, or salt water environments. As another example, at extremely small dimensions, such as among micro- or nano-scaled robots [1, 2], electromagnetic communication is challenging because of constraints such as the ratio of the antenna size to the wavelength of the electromagnetic signal [3].

Engineering micro-scale and nano-scale systems are the key to unlocking a realm of futuristic possibilities. Researchers around the world have been working to unlock new applications that include, but are not limited to: self-repairing machines, micro-scale robotics, synthetic biological devices, nanomedicine, and artificial immune systems that detect and kill cancer cells and pathogens [1, 4]. It should be surprising that, in 2015, these applications are still considered "futuristic", as many of the fundamental building blocks, such as state of the art micro/nano electromechanical systems (MEMS/NEMS), have been available for some time [5]. All these transformative applications of nanotechnology have one feature in common: they involve not just single devices working independently, but swarms of devices working in concert [6]. Therefore, solving the communication problem between nanodevices is an important and vital step towards these futuristic applications.

Inspired by nature, one possible solution to these problems is to use *chemical signals* as carriers of information, a technique known as *molecular communication* [7]. In molecular communication, a transmitter releases small particles such as molecules or lipid vesicles into a fluid or gaseous medium, where the particles propagate until they arrive at a receiver. The receiver then detects and decodes the information encoded in these particles. Figure 1.1 shows a typical molecular communication system.

Besides solving the communication problem at small scales, use of molecular

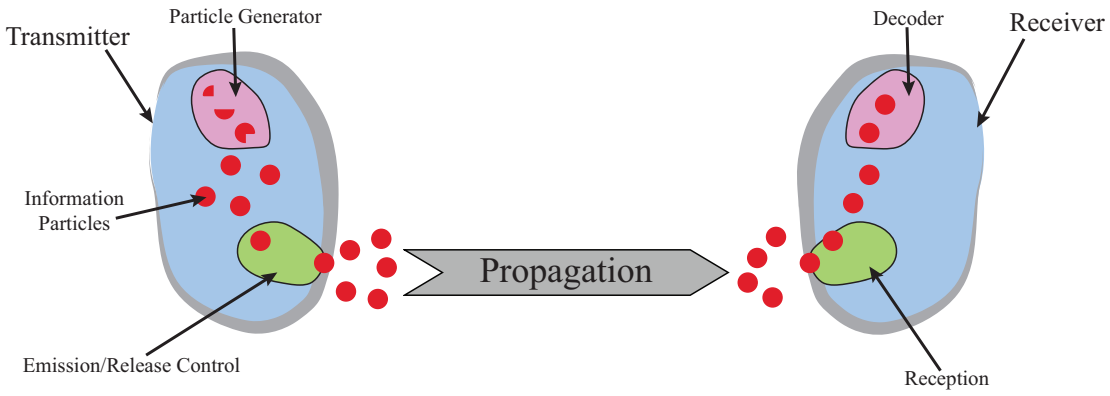


Figure 1.1: A typical molecular communication system.

communication in areas such as robotics, and infrastructure monitoring can unlock new applications in smart cities and disaster search and rescue. For example, inspired by cooperative ants, tiny robots could use molecular communication to collaborative find survivors under collapsed buildings. Typically, such an environment is not friendly to radio propagations and therefore use of radio technology may not be possible. As another example, in smart cities the city infrastructure, which consists of networks of metallic pipes and ducts, needs to be monitored by sensor networks . However, these environments are hostile to radio signals [8] and therefore, chemical signals could be used as an alternative form of communication between these sensors.

In nature, chemical signals are used for inter-cellular and intra-cellular communication at micro- and nano-scales [9], while pheromones are used for long-range

communication between members of the same species (such as social insects) [10]. Therefore, chemical signals can be used for communication at both macroscopic and microscopic scales. Moreover, molecular communication signals are biocompatible, and may require very little energy to generate and propagate [11, 12]. These properties makes chemical signals ideal for many applications, where the use of electromagnetic signals are not possible or not desirable.

Some of the disadvantages of molecular communication are the long delays and low throughputs of this technology. Therefore, molecular communication may not be suitable for applications where a large amount of data needs to be transferred, or applications with low-latency requirements. Therefore, radio-based technology is preferred for applications where high throughput and low latency is required, as long as there are no constraints for using this technology.

Although molecular communication is present in nature and is used by micro-organisms such as microbes to communicate and detect other micro-organisms, it was only recently that engineering a molecular communication system has been proposed as a solution for networking microscale devices [13]. At macroscales, the use of chemical signals for communication has been almost nonexistent, where only very primitive forms of chemical signaling have been developed in robotics [14, 15].

In designing molecular communication systems, different techniques can be employed for encoding messages.

- *Number/concentration of information particles* [16, 17]: Different numbers of information particles indicate different transmission symbols. For example, in a binary transmission, release of 0 information particles can indicate bit 0, while release of $n > 0$ information particles could indicate bit 1. Alternatively, information could be encoded in the concentration of the information particles released.

- *Structure/type of particles* [18,19]: Messages could be encoded in the structure of information particles. For example, different proteins, deoxyribonucleic acid (DNA), and ribonucleic acid (RNA) sequences could encode different transmission messages. Alternatively, release of different types of chemicals could encode the transmission information. For example, different isomers of a molecule could be used to encode different transmission symbols.

- *Release timing* [20–22]: Information can also be encoded in the release timing of information particles. For example, in the time interval between consecutive releases of information particles.

Regardless of the encoding technique used, the particles released by the transmitter must propagate the medium until they reach the receiver.

Many different propagation schemes can be employed in molecular communication [23].

- *Propagation through diffusion* [24]: In this scheme, the information particles propagate via Brownian motion due to collisions with other free molecules present in the fluid environment. Therefore, they diffuse randomly in the environment until they arrive at the destination.

- *Flow-assisted propagation* [25]: Flow can be used along with diffusion to improve the speed of transport by introducing a bias in the random movement of information particles in the direction of the receiver.

- *Propagation through gap junctions* [26]: Gap junctions are intercellular connections between neighbouring cells located at the cell membrane. They are literally gaps that are formed out of two aligned connexon structures which connect the cytoplasm of two adjacent cells. These gap junctions allow the free diffusion of selected information particles between two neighbouring cells. Through interconnecting a series of real or artificial cells connected by gap junctions, a "wire" can be made connecting the transmitter to the receiver.

- *Propagation using molecular motors* [27]: Molecular motors are a special type of molecule (typically proteins) that convert chemical energy into kinetic energy. They could be used to carry information particles from the transmitter to the receiver. For example, inside cells kinesin (a special type of molecular motor) is used to transport various cargoes (typically other proteins or

vesicles) from one location in the cell to another.

- *Propagation using bacteria* [28]: Information particles could be embedded inside bacteria, where chemotaxis is then used to deliver the information chemicals to the destination.

The common theme between all these propagation schemes is the random nature of the propagation. This means that from the receiver's perspective there will be some noise introduced by the random propagation of the information particles.

Until only recently, the underlying mathematical models for molecular communication systems had not been studied from a communication-engineering perspective. Therefore, there are still many open problems to be solved, and there is a need for a mathematical framework from which these communication systems can be designed and developed. Moreover, there are no practical or experimental platforms based on molecular communication, which could be used for research and development, at either the macroscale or the microscale. With this in mind, my research plan is twofold. First, I study on-chip molecular communication systems from a communication theoretic perspective. Then, I introduce the first experimental platform that demonstrates the feasibility of molecular communication. In this dissertation, I also present a comprehensive survey of previous work on molecular communication, highlighting important terminology, concepts, and their underlying

physical processes.

1.2 Dissertation Contributions

In this section, I describe the dissertation's contributions.

1. **A comprehensive survey of molecular communication** [29]: Molecular communication is still in its infancy. This field has been studied as a solution to engineering communication systems only over the past decade. Therefore, most of what has been published was scattered across different journals. In this comprehensive survey, which I plan to submit to a journal for publication, I summarize the recent advancements in the field in a tutorial fashion. Chapters 2 and 3 are dedicated to this comprehensive survey.

2. **Simulation and modelling on-chip molecular communication** [30–32]: Most of the models developed for molecular communication have assumed channels with infinite boundaries. However, in practice many real-world applications, such as on-chip communication, are inside confined channels. Analytical solutions for confined space diffusion or advection-diffusion are difficult to obtain. In this part, I present a simulator that can simulate molecular communication under different propagation schemes: diffusion, flow-assisted diffusion, and active transport using motor proteins (i.e., special protein molecules

that convert chemical energy into kinetic energy). I then compare the performance of each propagation scheme in terms of achievable information transmission rates for different channel dimensions. I show that active transport can be an effective form of propagation for on-chip applications. This contribution is covered in Chapter 4.

3. **Mathematical models for active transport molecular communication** [33, 34]: Since I show that active transport could be an effective means for engineering on-chip molecular communication systems, I derive analytical models for these systems. Simulating active transport molecular communication systems can be computationally intensive and therefore time consuming. Similarly, wet-lab experimentations can be laborious, expensive, and very time consuming. Therefore, solving complex design and optimization problems are impossible. To overcome these difficulties, I propose mathematical models that could be used to reduce the computational complexity of simulations, and would provide insights that could be used to find optimal design strategies. This contribution is presented in Chapter 5.

4. **Channel design and optimization for active transport molecular communication** [16, 35–38]: Using the derived mathematical models for active transport molecular communication, I find optimal design strategies. In

particular, finding the optimal placement for the transmitter in the channel, the optimal size of information particles, and the optimal channel shape are considered. For all these cases, the optimal design is the one that maximizes the information transmission rate. I show that the optimal transmission zone is along the walls of the channel, and that the optimal channel shape is the circular-shaped channel. I show that it is possible to achieve a considerable gain in information transmission rate when the proposed optimal design strategies are used. Chapter 6 is dedicated to this contribution.

5. **Building the first molecular communication platform** [39–41]: One of the most significant issues in molecular communication is the lack of experimental work. Although there have been demonstrations of engineered cell-to-cell communication, where simple commands are transferred between synthetic or genetically modified cells [42], there are no demonstrations for reliable sequential data communications. Another issue with these works is the expensive and specialized nature of wet-lab experimentation, which makes replicating the results more difficult. Perhaps the most important contribution of this dissertation is presenting the world's first molecular communication platform that is capable of transmitting short text messages reliably over a distance of a few meters. This macroscale system was designed to be inexpensive, and made from off-the-shelf materials, such that other re-

searchers could replicate and experiment with it. For example, with my help my collaborators have shown that molecular communication could be used for communication inside network of metallic ducts, where radio signals failed to propagate. Chapter 7 covers this topic.

6. **Channel models for my tabletop platform** [43–45]: My final contribution in this dissertation is derivation of new mathematical models for realistic molecular communication systems, such as my tabletop platform. Many of the theoretical models used in molecular communication rely on simplifying assumptions. For example, they assume that the information particles are perfectly detected at the receiver. However, in practice, many sensors are imperfect. My contribution in this part of the dissertation is to show that the end-to-end system response of my tabletop platform is nonlinear and different from previous theoretical works. I then use experimental results to derive new theoretical models for my platform, and show that the inherent nonlinearity may be modeled as additive noise which could be filtered. This is a significant contribution since most of the theoretical tools that are traditionally applied to communication systems, assume linearity. Therefore, demonstrating that the nonlinear system could be filtered into a linear one makes use of these tools possible. This contribution is discussed in Chapter 8.

1.3 Dissertation Organization

In this chapter, I provided an overview and a summary of important contributions made in the dissertation.

The rest of the dissertation is organized as follows.

- Chapter 2 gives an overview of molecular communication and its underlying physical and biological processes. In particular, I consider molecular communication at microscale and macroscale, and discuss the physical components of the transmitter, receiver, and the physical processes associated with different propagation schemes.

- Chapter 3 provides a survey of molecular communication from a communication engineering perspective. In particular, I consider different channel models, modulation schemes, error-correcting codes, simulation tools, and some practical implementations.

- In Chapter 4, I consider on-chip molecular communication, and create a simulation environment for diffusion-based, flow-based, and active transport based molecular communication. Using the simulation environment, I calculate the channel capacity under different propagation schemes and show that active transport could be beneficial for on-chip communication.

- In Chapter 5, I derive different mathematical models for active transport

propagation. These models could be used to lessen the computational complexity of simulating active transport, and provide insights that could be used for design and optimization problems.

- I use the models derived in Chapter 5 to find optimal design strategies for active transport molecular communication in Chapter 6. I show that there is an optimal size for the information particles or their encapsulating vesicles, and I show that the optimal channel shape for this propagation scheme is the circular-shaped channel.

- Chapter 7 presents an experimental platform for molecular communication, that is capable of reliably transmitting short text messages over a distance of a few meters. The system is purposely designed to be inexpensive, and made from off-the-shelf materials such that other researchers could replicate the results.

- Chapter 8 provides new mathematical models for my tabletop platform. I show that the nonlinearity of the system can be modeled as additive noise, which may be filtered.

- Chapter 9 concludes the dissertation and provides some directions for future work.

1.4 Publications

The following are the publications which resulted from this dissertation research and follow up work.

1. N. Farsad, H. B. Yilmaz, A. Eckford, C.-B. Chae, and W. Guo, "A comprehensive survey of recent advancements in molecular communication," *IEEE Commun. Surveys Tuts.*, pp. 1–29, 2015, in preparation for submission

2. W. Guo, N. Farsad, J.-L. Wu, and A. Eckford, "Random walk through thin obstacles with gaps: Approximate hitting distribution," *IEEE Commun. Lett.*, pp. 1–4, 2015, submitted

3. W. Guo, C. Mias, N. Farsad, and J. Wu, "Molecular versus electromagnetic wave propagation loss in macro-scale environments," *IEEE Trans. Mol. Biol. Multi-Scale Commun.*, pp. 1–21, 2015, to appear

4. N. Farsad, A. Eckford, and S. Hiyama, "Design and optimizing of on-chip kinesin substrates for molecular communication," *IEEE Trans. Nanotechnol.*, vol. 14, no. 4, pp. 699–708, July 2015

5. N. Farsad, N.-R. Kim, A. W. Eckford, and C.-B. Chae, "Channel and noise models for nonlinear molecular communication systems," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 12, pp. 2392–2401, Dec 2014

6. N. Farsad, A. W. Eckford, and S. Hiyama, "A Markov chain channel model for active transport molecular communication," *IEEE Trans. Signal Process.*, vol. 62, no. 9, pp. 2424–2436, May 2014

7. S. Qiu, W. Guo, M. Leeson, S. Wang, N. Farsad, and A. W. Eckford, "Nanoparticle communications: from chemical signals in nature to wireless sensor networks," *Nanotechnology Perceptions*, vol. 10, no. 1, pp. 1–13, 2014

8. N. Farsad, W. Guo, and A. W. Eckford, "Tabletop molecular communication: Text messages through chemical signals," *PLoS ONE*, vol. 8, no. 12, p. e82935, Dec. 2013

9. N. Farsad, A. W. Eckford, S. Hiyama, and Y. Moritani, "On-chip molecular communication: Analysis and design," *IEEE Trans. NanoBiosci.*, vol. 11, no. 3, pp. 304–314, Sep. 2012

10. N. Farsad, A. W. Eckford, S. Hiyama, and Y. Moritani, "Quick system design of vesicle-based active transport molecular communication by using a simple transport model," *Nano Commun. Netw.*, vol. 2, no. 4, pp. 175–188, Dec. 2011

11. N. Farsad, W. Guo, C.-B. Chae, and A. W. Eckford, "Stable Distributions as Noise Models for Molecular Communication," in *Proc. IEEE Glob. Commun. Conf. (GLOBECOM)*, 2015, to appear

12. 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 Commun. (ICC)*, 2015, to appear

13. L. Wang, N. Farsad, W. Guo, S. Magierowski, and A. W. Eckford, "Molecular Barcodes: Information Transmission via Persistent Chemical Tags," in *Proc. IEEE Int. Conf. on Commun. (ICC)*, 2015, to appear

14. S. Qiu, N. Farsad, T. Dong, A. W. Eckford, and W. Guo, "Under-Water Molecular Signalling: a Hidden Transmitter and Absent Receivers Problem," in *Proc. IEEE Int. Conf. on Commun. (ICC)*, 2015, to appear

15. C. Lee, B. Koo, N.-R. Kim, H. B. Yilmaz, N. Farsad, A. Eckford, and C.-B. Chae, "Molecular MIMO communication link," in *Proc. IEEE Int. Conf. on Comput. Commun. (INFOCOM)*, 2015

16. N. Farsad, W. Guo, and A. W. Eckford, "Molecular communication link," in *Proc. IEEE Int. Conf. on Comput. Commun. Workshops (INFOCOM WKSHPS)*, 2014, pp. 107–108

17. 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 Commun. (ICC)*, 2014

18. 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 Commun. (ICC)*, 2014

19. N. Farsad, A. W. Eckford, and S. Hiyama, "Modelling and design of polygon-shaped kinesin substrates for molecular communication," in *Proc. IEEE Int. Conf. on Nanotechnol. (IEEE-NANO)*, Aug. 2012, pp. 1–5

20. ——, "A mathematical channel optimization formula for active transport molecular communication," in *Proc. IEEE Int. Conf. on Commun. Workshops (ICC WKSHPS)*, Ottawa, Canada, 2012

21. ——, "Channel design and optimization of active transport molecular communication," in *Proc. Int. Conf. on Bio-Inspired Models of Netw., Inf. and Comput. Syst. (BIONETICS)*, York, England, 2011

22. N. Farsad, A. W. Eckford, S. Hiyama, and Y. Moritani, "A simple mathematical model for information rate of active transport molecular communication," in *Proc. IEEE Int. Conf. on Comput. Commun. Workshops (INFOCOM WKSHPS)*, Shanghai, P. R. China, 2011, pp. 473–478

23. A. W. Eckford, N. Farsad, S. Hiyama, and Y. Moritani, "Microchannel molecular communication with nanoscale carriers: Brownian motion versus active

transport," in *Proc. IEEE Int. Conf. on Nanotechnol. (IEEE-NANO)*, Seoul, South Korea, 2010, pp. 854 – 858

24. N. Farsad, A. W. Eckford, S. Hiyama, and Y. Moritani, "Information rates of active propagation in microchannel molecular communication," in *Proc. Int. Conf. on Bio-Inspired Models of Netw., Inf. and Comput. Syst. (BIONETICS)*, Boston, MA, 2010, p. 7

2 Overview of Molecular Communication

The goal of a communication system is to transfer information across space and time. To achieve this, a signal needs to be generated by a transmitter in accordance to the information that is intended to be transferred to a receiver. This signal then propagates to a receiver, where the intended information is decoded from the received signal. Therefore, any communication system can be broken down into three major components: the transmitter, the receiver, and the channel.

Figure 2.1 shows the block diagram representation of these three modules and their submodules. The transmitter has some information that it wants to transmit to a receiver. Any discrete message can be represented with a string of binary numbers, so the transmitter uses a source encoder to encode the information message as a binary sequence. The transmitter can also add error-correcting codes using the channel encoder block, which introduces redundancy by adding extra bits. The receiver can use the added redundancy to mitigate the errors that may be introduced by the channel. Finally, the transmitter must modulate the channel symbols

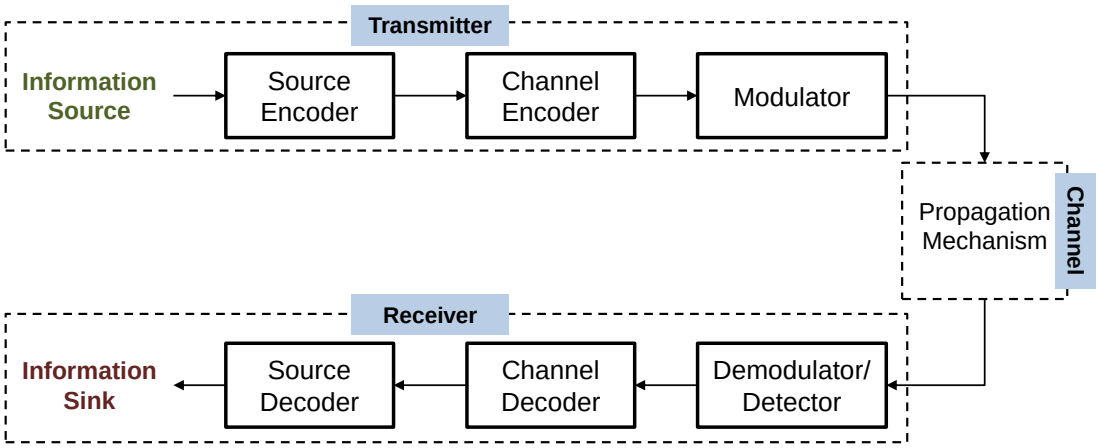


Figure 2.1: Block diagram of a typical communication system.

(i.e., the output of the channel encoder) onto a carrier signal and release the signal for propagation in the channel.

The channel is the environment in which the transmitted signal propagates from the sender to the receiver. In traditional communication system, a channel is typically a wire where electrical signals propagate, or free space where electromagnetic signals propagate. The transmitted signals are the electrical currents or electromagnetic waves. In molecular communication, small particles called *information particles* form the carrier signal conveying the information. Information particles are a few nanometers to a few micrometers in size. Information particles could be biological compounds, such as proteins, or synthetic compounds, such as gold nano particles. The channel in molecular communication is an aqueous or a gaseous environment where the tiny information particles can freely propagate.

In communication systems, the channel may introduce noise into the system, where the noise is any distortion that results in degradation of the signal at the receiver. For example, the noise can result from the signal fading as it propagates, or interference from other signals. Noise can also be introduced by the transmitter and the receiver themselves (e.g., thermal noise in the electronic components). In radio-based communication systems, the source of the noise is typically the fading of electromagnetic signals, and the interference of different electromagnetic waves, for example due to multipath. In molecular communication, the source of the noise is the random nature of the underlying physical processes such as random propagation due to Brownian motion.

When the transmitted signal arrives at the receiver, the receiver must first demodulate and detect the channel symbols. The estimated channel symbols are then decoded using a channel decoder, where some of the errors introduced by the transmitter, the channel and the receiver may be corrected. The output of the channel decoder goes through a source decoder, where the receiver estimates what information the transmitter has sent. If this estimation is correct, then the communication session has been successful.

Some of the physical components that may be required to implement molecular communication systems are shown in Figure 2.2. At the transmitter a physical process is required for generation or storage of information particles. There may

also be a need for a mechanism that controls the release of information particles. This mechanism acts as the modulator. Finally, there should be a processing unit that controls the different processes within the transmitter. The processing unit can operate chemically (e.g., chemical pathways), electrically (e.g., microcontroller), or through other means. To function, the transmitter may need a source of energy. For example, the transmitter can be a synthetic cell that gets its power from the environment, or from an electrical source.

After the information-carrying particles are released in the channel, a propagation mechanism is necessary to transport them to the receiver. This mechanism can be diffusion based, flow based, or an engineered transport system using molecular motors. At the receiver, there must be a sensor, receptor, or a detector that can measure some arrival property of the received information particles. This property can be simply the presence or absence of the information particles, their concentration, time of arrival, or any other measurable parameter. If coding techniques are used there may also be a need for a central processing unit for decoding and deciphering the received signal. Finally, an energy source may be required for the receiver to function.

In nature, molecular communication is employed over short-range (nm scale), mid-range (μm to cm scale), or long-range (cm to m scale) communication [53, 54]. For example, neurotransmitters use passive transport (free diffusion) to commu-

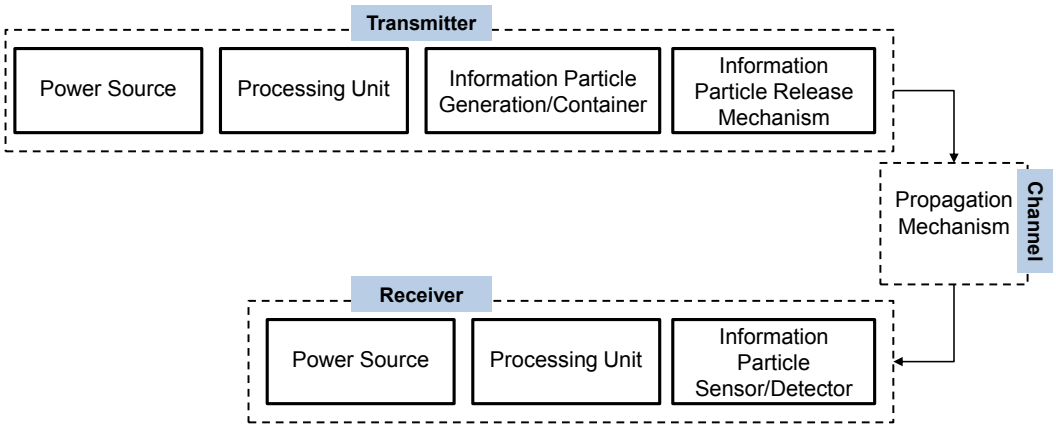


Figure 2.2: Physical components required for molecular communication.

nicate over short-range; inside cells motor proteins are used to actively transport cargoes over the mid-range; and hormones are transported over the long-range using flow (e.g., blood flow from the heart). In this dissertation, the short- and mid-range are referred to as microscale molecular communication, and the long-range is called macroscale molecular communication. The physical properties of matter change from macroscales to microscales, hence microscale and macroscale molecular communication are considered separately in the next two subsections, and different mechanisms that can be implemented at each scale are discussed.

2.1 Microscale Molecular Communication

On December 29, 1959, physicist Richard Feynman gave a lecture at American Physical Society meeting in Caltech titled "There's Plenty of Room at the Bottom" [55]. The goal of the talk was to inspire scientists find a solution to "the problem of manipulating and controlling things on a small scale". To motivate his vision for futuristic possibilities of miniaturization he asked: "Why cannot we write the entire 24 volumes of the Encyclopaedia Brittanica on the head of a pin?"

Over the past couple of decades there have been considerable advancements in the fields of nanotechnology, biotechnology, and microrobotics, where design and engineering of *micro- or nano-scale devices*, performing simple and specific tasks, is beginning to take shape. In order to achieve complex goals cooperation of these devices is essential, therefore a micro- or nano-scale networks [4,6] (typically called *nanonetworks*) of multiple micro- or nano-devices must be formed. This leads to a need for miniaturization of current communication systems, and poses a question similar to the one Feynman asked half a century ago: how much can we shrink a communication system?

The current technology, which is used everyday by billions of people worldwide in devices such as cellphones, computers, televisions, radios, etc., relies entirely on electronic or electromagnetic signals (including the optical range). However,

shrinking the current technology to very small dimensions is not a simple task (e.g., shrinking the size of a transceiver antenna to nano-scales). Currently, two prominent modes of communication are considered at small scales: electronic or electromagnetic communication with the help of novel materials such as carbon nanotubes [56], or molecular communication [13].

One advantage of miniaturizing traditional electronic or electromagnetic communication is the availability of theoretical foundations that communication engineers have developed over the past century. However, there remain numerous obstacles to overcome before a fully functional communication system can be realized using novel materials such as carbon nanotubes or graphene. For molecular commutation, a comprehensive theoretical framework is not known to communication engineers. However, fully functional molecular communication systems have already been selected by evolutionary processes, and used by living organisms over billions of years.

Molecular communication has two main advantages over the electromagnetic communication for nanonetworks. First, it is biocompatible, which means that it would be more suitable for medical applications. Second, it is energy efficient and has a very low heat dispersion. Therefore, molecular communication becomes a more suitable solution for nanonetworks.

Over the next subsections the different physical, chemical and biological pro-

cesses that underlie molecular communication systems are discussed. Since one of the main differences between molecular communication and radio based communication is the carrier signal, information particles and their propagation are discussed first.

2.1.1 Information Particles

The structure and size of information particles affect how they propagate in a given environment. For example, increasing the size of the information particle changes the diffusion coefficient and hence the diffusion process. To make molecular communication channel more reliable, information particles need to be chemically stable and robust against environmental noise and interference from other molecules and particles. Moreover, information particles may degrade or denature over time in the environment. For example, molecular deformation and cleavage caused by enzymatic attacks or changes in pH in the environment may degrade the information particles [23].

Due to degradation of particles, the transmitter can release a large number of information particles, or alternatively encapsulate the information-carrying particles inside *liposomes* that separate the environment and the information particles. Liposomes are lipid bilayer structures that can encapsulate different proteins and particles [57]. Figure 2.3 shows the structure of the liposome. It contains two layers

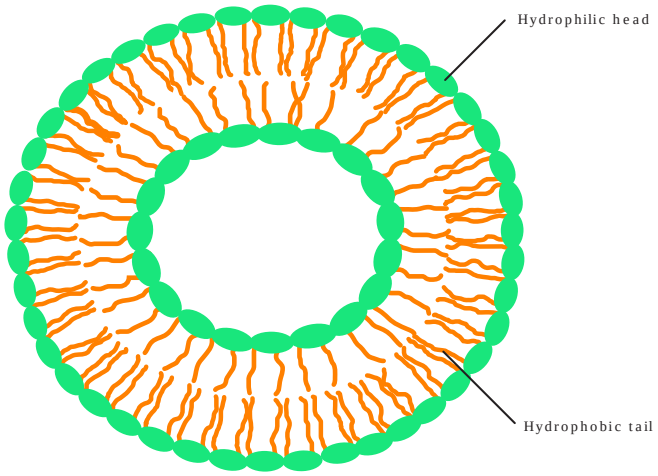


Figure 2.3: Liposome's lipid bilayer structure.

of lipid molecules, where one side of the lipid is hydrophobic (repelled from water) and the other side is hydrophilic (attracted towards water). It separates the aqueous environment inside the liposome from the outside of the liposome. In [58], the relationship between the number of lipid molecules and the liposome's wall thickness and radius is derived and tested experimentally. In [59], a chemical assembly line for generation of lipid vesicles and encapsulation process is presented.

Some examples of information particles used in nature by biological systems include hormones, pheromones, neurotransmitters, intracellular messengers, and deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) molecules. Information particles can also be synthesized for specific purposes as demonstrated in drug delivery. For example, in [60] nanoparticles are used to target particular tissue

types.

In molecular communication, information can be encoded in to the structure (type of) the information particles, in the number of information particles, or in the release timing. The encoding and modulation schemes will be discussed in more details in Chapter 3.

Another issue to consider when selecting the information particles is the chemical bandwidth, which is the amount of information that can be encoded in the structure of the information particle. For example, small chemical molecules such as hormones or pheromones or carbohydrates can be sent at short or long ranges, are stable in most channels, and can provide 10^2-10^6 unique structural combinations [61]. However, the problem of generating and detecting these different combinations can be challenging. Dial-a-Molecule is project that tackles this problem with the goal of rapidly synthesizing compounds [62].

2.1.2 Channel and Propagation

For microscale molecular communication, many different propagation schemes are possible for transporting information particles. They are: diffusion based propagation [63–65], flow assisted propagation [66], active transport using molecular motors and cytoskeletal filaments [67], bacterial assisted propagation [18, 28, 68], and kinesin molecular motors moving over immobilized microtubule (MT) tracks [69, 70].

Each propagation method is described in detail over the next few subsections.

2.1.2.1 Free Diffusion

Diffusion, also known as Brownian motion, refers to the random motion of a particle as it collides with other molecules in its vicinity. Through this random motion, the information-carrying particles can propagate from the transmitter to the receiver by utilizing the thermal energy that is already present in the channel environment. Therefore, no source of external energy is required for diffusion-based propagation. There are many examples of this class of propagation in biology. An example is the neurotransmitter molecules Acetylcholine (ACh) emitted from nerve cell to neuromuscular junction for conveying motor actions in muscle cells [71]. When muscles in a specific part of the body need to contract, the nerve cells in that region send a signal to the muscle tissue through these junctions to trigger contraction [9]. Another example is DNA binding molecules (e.g., repressors) propagating over a DNA segment to search for a binding site [72].

Diffusion can be accurately simulated using Monte Carlo simulation [73]. In particular, the motion of the information particles is considered for a discrete time period of duration Δt. Given some initial position (x_0, y_0, z_0) at time $t = 0$, for any integer $k > 0$, the motion of the particles is given by the sequence of coordinates (x_i, y_i, z_i) for $i = 1, 2, \ldots, k$. Each coordinate (x_i, y_i, z_i) represents the position of

the particle at the time $t = i\Delta t$

$$
\begin{aligned}
(x_i, y_i, z_i) &= (x_{i-1}, y_{i-1}, z_{i-1}) + (\Delta x_i, \Delta y_i, \Delta z_i) && (2.1)\\
\Delta x_i &\sim \mathcal{N}(0, 2D\Delta t), && (2.2)\\
\Delta y_i &\sim \mathcal{N}(0, 2D\Delta t), && (2.3)\\
\Delta z_i &\sim \mathcal{N}(0, 2D\Delta t), && (2.4)
\end{aligned}
$$

where Δx_i, Δy_i, Δz_i are independent random displacements over each spatial axis during Δt, D is the diffusion coefficient, and $\mathcal{N}(\mu, \sigma^2)$ is the normal distribution with mean μ and the variance σ^2. The diffusion movement in these equations is a Martingale process, where subsequent motions are unrelated to previous motions and positions [74]. Alternatively diffusion can be simulated via the following method:

$$
\begin{aligned}
(x_i, y_i, z_i) &= (x_{i-1}, y_{i-1}, z_{i-1}) + (\Delta x_i, \Delta y_i, \Delta z_i) && (2.5)\\
\Delta x_i &= \Delta r \cos\theta_i \sin\phi_i, && (2.6)\\
\Delta y_i &= \Delta r \sin\theta_i \sin\phi_i, && (2.7)\\
\Delta z_i &= \Delta r \cos\phi_i, && (2.8)
\end{aligned}
$$

where Δr is the particle's total displacement over each time interval of Δt, and

both θ_i and ϕ_i define the angle of displacement in 3-dimensional space over the time interval Δt. Furthermore, over each time interval of Δt, the particle's displacement Δr is given by

$$\Delta r = \sqrt{6D\Delta t}\ . \tag{2.9}$$

The angle θ_i is an independent, identically distributed (iid) random variable for all i, uniformly distributed on $[0, 2\pi]$, and the angle ϕ_i is also an iid random variable given by $\phi_i = \arccos(1 - 2U)$, where U is a uniformly distributed random variable over $[0, 1]$.

The comparative sizes between the propagating molecule S_m and the molecules of the fluid S_{fluid} affect the diffusion coefficient [75]. For a given particle and fluid environment, D is given by

$$D = \begin{cases} \dfrac{k_B T}{6\pi\eta R_H} & \text{if } S_m \gg S_{\text{fluid}} \\ \\ \dfrac{k_B T}{4\pi\eta R_H} & \text{if } S_m \approx S_{\text{fluid}} \end{cases}, \tag{2.10}$$

where $k_B = 1.38 \cdot 10^{-23}$ J/K is the Boltzman constant, T is the temperature (in K), η is the dynamic viscosity of the fluid, and R_H is the hydrodynamic radius (also known as Stoke's radius) of the information particle. For most applications, it can be assumed that D is stationary throughout the medium, and that collisions with

Table 2.1: Diffusion coefficients of selected molecules in water at 25° C.

Molecule	D (μm^2/s)
DNA	0.81 to 53
Human serum albumin	61
Insuline	150
Sucrose	520
Glucose	600
Glycerol	930
Nitrate	1700
Water molecule	2100

the boundaries are elastic. Example values of D in μm^2/s in water at 25 °C are obtained from [76] and listed in Table 2.1.

2.1.2.2 Diffusion with First Hitting

In nature, most receptors remove the information molecules from the environment through binding, absorbing, or chemical reactions [77]. For example, acetylcholinesterase breaks down the molecules in neuronal junctions [71]. This process is modeled by first hitting processes since information particles are assumed to be removed from the environment by the receptors or the detection mechanisms.

The first generalized model for molecular communication which considered the first hitting process in 1-D environment was derived in [20] as

$$f_{\text{hit}}^{\text{1D}}(t) = \frac{d}{\sqrt{4\pi D t^3}}\, e^{-d^2/4Dt} \tag{2.11}$$

where d and D correspond to distance and the diffusion coefficient, respectively. This equation can be interpreted as the impulse response of the diffusion channel in 1-D environment with an absorbing receiver. To find the probability of hitting an absorbing receiver until time t, (2.11) is integrated with respect to time

$$\begin{aligned} F_{\text{hit}}^{\text{1D}}(t) = \int_0^t f_{\text{hit}}(t')\, dt' &= \text{erfc}\left(\frac{d}{\sqrt{4Dt}}\right) \\ &= 2\Phi\left(\frac{-d}{\sqrt{2Dt}}\right) \end{aligned} \tag{2.12}$$

where erfc(.) and $\Phi(.)$ represent the complementary error function and the standard Gaussian cdf, respectively.

Similarly, the fraction of hitting molecules to a perfectly absorbing spherical receiver in a 3-D environment is derived in [78]. The hitting rate of molecules to a spherical receiver in a 3-D environment is formulated as

$$f_{\text{hit}}^{\text{3D}}(t) = \frac{r_r}{d + r_r} \frac{d}{\sqrt{4\pi D t^3}}\, e^{-d^2/4Dt} \tag{2.13}$$

where r_r denotes the radius of the spherical receiver [78]. One can obtain the fraction of hitting molecules until time t by integrating $f_{\text{hit}}^{\text{3D}}(t)$ in (2.13) with respect to time, which yields similar results with the 1-D case. Note that, there is a positive probability of not hitting to the absorbing boundary for a diffusing particle in a

3-D environment when time goes to infinity. The survival probability depends on the radius of the receiver and the distance. First hitting time formulations in 2-D environments with infinite boundaries do not exist. However, in [79], an analytical expression is derived for planar wedge boundaries over specific wedge angles of $\pi/2$ and π/k where k is an odd integer.

2.1.2.3 Flow Assisted Propagation

Although diffusion is advantageous because no external energy is required for propagation, it can be very slow at transporting information particles from the transmitter to the receiver over large separation distances. One way to assist the speed of propagation in pure diffusion is to introduce flow into the environment. The most effective flow is the one that is in the direction of the the transmitter to the receiver. An example of this propagation scheme is found in biology. Certain cells can secrete hormonal substances that propagate using the flow present in the blood stream to distant target cells.

The flow assisted propagation can be simulated using Monte Carlo simulations by modifying the equation presented in the previous section. Introducing flow will

change Equations (2.5)-(2.8) into,

$$\Delta x_i = v_{x,i-1}(x_{i-1}, y_{i-1}, z_{i-1})\Delta t + \Delta r \cos\theta_i \sin\phi_i, \tag{2.14}$$

$$\Delta y_i = v_{y,i-1}(x_{i-1}, y_{i-1}, z_{i-1})\Delta t + \Delta r \sin\theta_i \sin\phi_i, \tag{2.15}$$

$$\Delta z_i = v_{z,i-1}(x_{i-1}, y_{i-1}, z_{i-1})\Delta t + \Delta r \cos\phi_i, \tag{2.16}$$

where $v_{x,i-1}(x_{i-1}, y_{i-1}, z_{i-1})$, $v_{y,i-1}(x_{i-1}, y_{i-1}, z_{i-1})$, and $v_{z,i-1}(x_{i-1}, y_{i-1}, z_{i-1})$ are flow velocities in the x, y, and z directions [73]. These velocities are a function of time and space because in practice flows can change with respect these variables.

2.1.2.4 Kinesin Moving Over Microtubule Tracks

It is also possible to transport the information particles actively from the transmitter to the receiver using motor proteins. One way to achieve this is using microtubules and kinesin motor proteins [67, 70, 80]. Microtubules are hollow tube-like structures, 24 nm in diameter, whose walls are formed by adjacent protofilaments. They are made up of dimeric subunits composed of α- and β-tubulin that polymerize into microtubules. In biology, microtubules are a component of the cytoskeleton, found throughout the cytoplasm. They are involved in maintaining the structure of the cell, providing platforms for intracellular transport, and a number of other intracellular processes.

One of the examples of the molecular motor proteins is Kinesin that walks on microtubule tracks. Kinesin contains a head domain, a neck domain, a stalk, and a tail as shown in Figure 2.4. The head domain binds to the microtubule, the neck domain provides the needed flexibility for the act of walking, the stalk connects the neck to the tail, and the tail domain attaches to a cargo. Hydrolysis of adenosine triphosphate (ATP) in the head domain causes the head domain to walk along the microtubule track in one direction by repeating cycles of conformational changes. ATP hydrolysis detaches the phosphate from an ATP to produce adenosine diphosphate (ADP), which releases energy. This energy is then used by the motor protein to change its conformation and take a single step. During each step one head is detached and moves forward, while the other head stays stationary. Through repeating this cycle, kinesin moves in one direction. In biology, kinesin can carry different cargoes over microtubule tracks from one location in the cell to another location. Therefore, kinesin acts as a locomotive in cells.

Inspired by this, it is possible to engineer a microtubule track between a transmitter and a receiver, where kinesin motor proteins would carry information particles from the transmitter to the receiver. For example in [70], it was shown that it is possible to create microtubule tracks *in vitro* in a self-organizing manner, using polymerization and depolymerization. It was also shown that a second approach based on reorganization of microtubules using motor proteins can also be used to

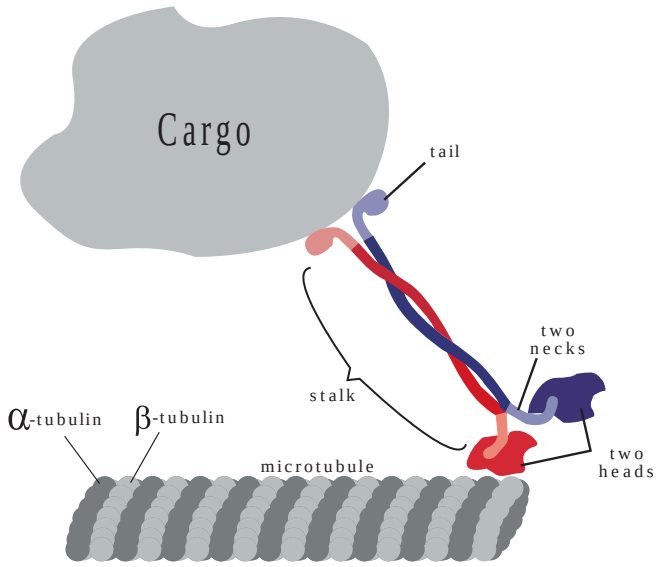


Figure 2.4: Structure of the microtubule and the kinesin motor.

create tracks that connect a transmitter to a receiver.

2.1.2.5 Microtubule Filament Motility Over Stationary Kinesin

Although in biological cells kinesin moves over microtubule tracks, in [81] it was shown that stationary kinesin attached to a substrate can mobilize microtubule filaments. This scheme is very suitable for on-chip molecular communication applications [80], where the transmitter and receiver are located in a lab-on-a-chip device. For example, recently it was shown that electrical currents can be used to control the speed and direction of the microtubules [82,83]. However, unlike kinesin which can naturally carry cargoes, a technique for carrying information particles

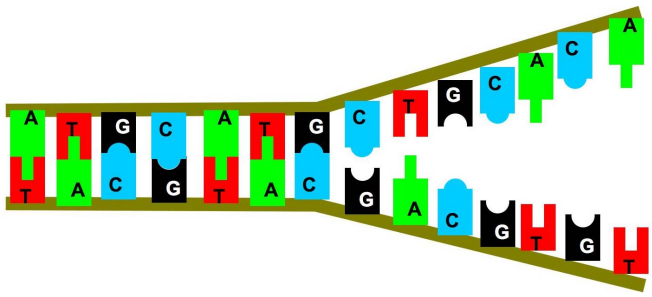


Figure 2.5: An example of ssDNAs hybridization bonds.

by microtubule filaments is necessary.

In [67, 80], DNA hybridization is proposed for carrying cargoes such as vesicles. A single strand of DNA is made up of many nucleotides. Each nucleotide is made up of one of the four possible nucleobase compounds: guanine (G), adenine (A), thymine (T), or cytosine (C). A DNA strand is made up of two single stranded DNAs (ssDNA)s connected together using intermolecular bonds, with adenine bonding only to thymine with two hydrogen bonds, and cytosine bonding only to guanine with three hydrogen bonds. The bases are said to be complimentary to one another, resulting in complimentary base pairing (AT and GC), also referred to as "canonical" or "Watson-Crick" base pairing. The bonding occurs when two ssDNAs with complementary base pairs come close together. Figure 2.5 shows an example of two ssDNAs hybridizing.

In [67,80], cargo vesicles, transmission area, receiver area, and microtubules are covered with ssDNAs. Microtubules moving over a kinesin covered glass substrate

are covered with 15 base ssDNAs, and the information particles or the vesicles are also covered with 23 base ssDNAs which are complementary to that of the microtubules' ssDNAs. When the microtubule glides close to an information particle or a vesicle, the two ssDNA sequences hybridize and the microtubule carries the cargo until it gets close to the receiver. The receptor module at the receiver is covered with 23 base ssDNAs, which are complementary to that of the cargo. When a loaded microtubule filament glides close to the receptor module, it will unload the information particle or the vesicle through hybridization bound with the complementary 23 base ssDNA at the receptor module. Since 23 base hybridization bound is stronger than the 15 base hybridization bonds between the cargo and the microtubules, the particles or the vesicles are unloaded at the receiver. Figure 2.6 depicts the microtubule motility mechanism on a DNA microarray.

In [84, 85], it was shown that the motion of the microtubule filaments over stationary kinesin can be simulated using Monte Carlo simulations. The motion of the microtubules is largely regular, with Brownian motion effects causing random fluctuations. Moreover, the microtubules move only in the x and y directions, and do not move in the z direction (along the height of the channel) because they move right on top of kinesins. In simulations, a two-dimensional motion of a microtubule can be simulated over discrete time intervals of Δt seconds. Given some initial position (x_0, y_0) at time $t = 0$, for any integer $k > 0$, the motion of the microtubule

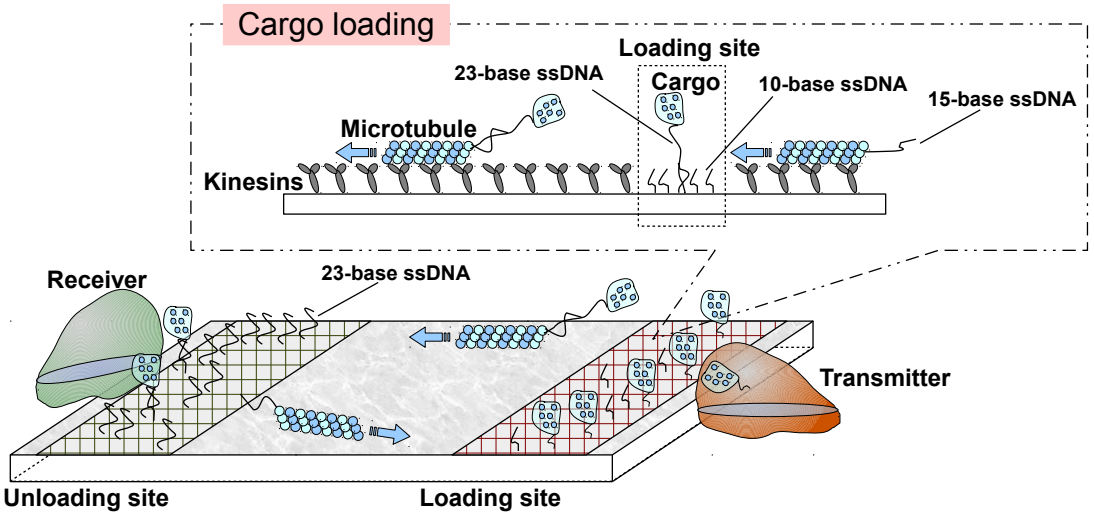


Figure 2.6: Cargo transport mechanism for microtubule filaments gliding over stationary kinesin substrate using DNA hybridization bonds.

is given by the sequence of coordinates (x_i, y_i) for $i = 1, 2, \ldots, k$. Each coordinate (x_i, y_i) represents the position of the microtubule's head at the end of the time $t = i\Delta t$, where

$$x_i = x_{i-1} + \Delta r \cos\theta_i, \tag{2.17}$$

$$y_i = y_{i-1} + \Delta r \sin\theta_i. \tag{2.18}$$

This is similar to the equations for Brownian motion, because the motion of the microtubules are effected by the same physical processes.

In this case, the step size Δr at each step is an iid Gaussian random variable

with mean and variance

$$E[\Delta r] = v_{\text{avg}}\Delta t, \tag{2.19}$$

$$\text{Var}[\Delta r] = 2D\Delta t, \tag{2.20}$$

where v_{avg} is the average velocity of the microtubule, and D is the microtubule's diffusion coefficient. The angle θ_i is no longer independent from step to step: instead, for some step-to-step angular change $\Delta\theta$, the new θ_i is given by

$$\theta_i = \Delta\theta + \theta_{i-1}. \tag{2.21}$$

Now, for each step, $\Delta\theta$ is an iid Gaussian-distributed random variable with mean and variance

$$E[\Delta\theta] = 0, \tag{2.22}$$

$$\text{Var}[\Delta\theta] = \frac{v_{\text{avg}}\Delta t}{L_p}, \tag{2.23}$$

where L_p is the persistence length of the microtubule's trajectory. In [84], these values were given as $v_{\text{avg}} = 0.85$ μm/s, $D = 2.0 \cdot 10^{-3}$ μm^2/s, and $L_p = 111$ μm. In case of a collision with a boundary, it is assumed that the microtubule does not reflect off the boundary, as in an elastic collision, but instead sets θ_i so as to follow

the boundary [84].

It may also be possible to use actin filaments and myosin in place of microtubule filaments and kinesin [86]. Myosin is a motor protein that is used for muscle contraction and actin filaments are rod-like proteins. The motion which results in muscle contraction is generated using ATP hydrolysis. In [87], it was shown that silicon nanowires covered with myosin can be used as tracks for actin filaments. Dynein molecular motors and microtubule filaments are another example of motor protein-filament motility [88]. It is also possible to use active transport in conjunction with flow based propagation in microfluidic lab-on-chip devices as was shown in [89].

2.1.2.6 Bacteria Assisted Propagation

In [18,28], bacteria-based communication was proposed for transferring information particles from a transmitter to a receiver. In this scheme, the information particles are placed inside bacteria at the transmitter and the loaded bacteria are released into the channel. The bacteria then propagate in the channel until they reach the receiver and deliver the information particles. In order to assist this propagation, self-propelling flagellated bacteria could be used. Figure 2.7 depicts a flagellated bacteria and the structure of flagella. A flagellum consists of a filament and a motor section. The motor section rotates the filament, which in turn propels the

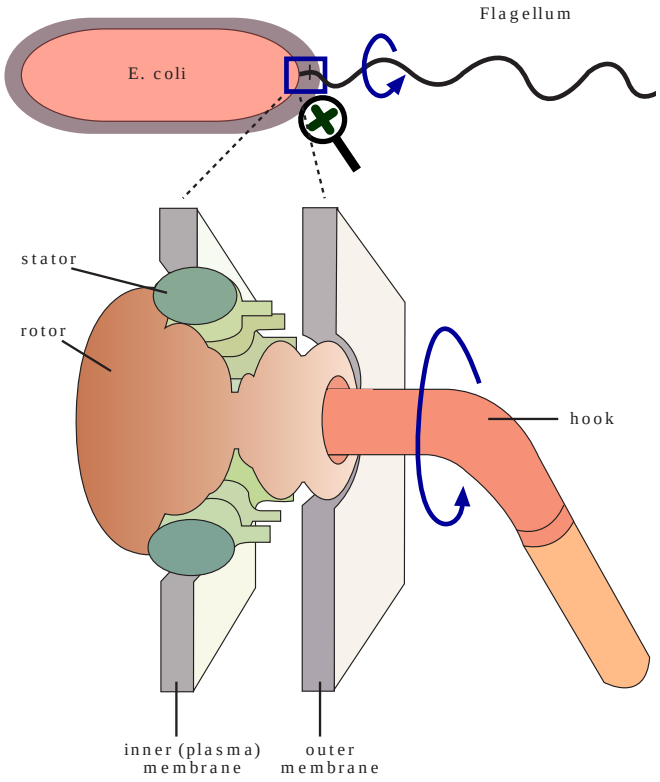


Figure 2.7: Flagellated bacteria and the structure of flagella.

bacteria in a certain direction. To guide the bacteria towards the receiver, attractant molecules can be released by the receiver. The bacteria are then attracted towards the receiver by following the concentration gradient and moving towards the increasing concentration of attractant molecules. It is possible to simulate the motion of bacteria as shown in [72, 90, 91].

2.1.2.7 Propagation Through Gap Junction

Gap junctions are intercellular connections between neighboring cells located at the cell membrane. They are gaps that are formed out of two aligned connexon structures which connect the cytoplasm of two adjacent cells. These gap junctions allow the free diffusion of selected molecules between two neighboring cells. Their permeability can vary from time to time, allowing different molecules to pass between the cells. Also, they can be closed and reopened by the cell during the lifetime of the cell.

Intercellular calcium wave (ICW) is one of the main intercellular communication systems observed among biological cells [9]. In this system, after a cell is stimulated, it generates a response to this stimulus with an increased cytosolic Ca^{2+} concentration through the usage of the secondary messenger molecule inositol 1,4,5-triphosphate (IP_3). The Ca^{2+} concentration wave passes to the neighboring cells via usage of IP_3 or ATP. There are two main pathways: the internal pathway and the external pathway. In the internal pathway, IP_3 molecules diffuse through gap junctions, while in the latter, ATP molecules are released to the extracellular space and diffuse freely to trigger nearby cells. According to experimental studies, it has been shown that these pathways complement each other rather than being alternative approaches [92]. The frequency and effective range of ICW heavily de-

pends on which pathway is used. Using internal pathway enables frequent fast communication between nearby cells, however reaching distant cells requires the usage of external pathway [92]. ICW can reach to $200-350$ μm in one dimension (assuming that the cells are aligned linearly) at a speed of $15-27$ μm/s and oscillate in the $10^{-3} - 1$ Hz range [93, 94]. It has been shown experimentally that cells encode the information on the frequency and amplitude components of ICWs [94].

A molecular communication system based on ICW was proposed in [95]. The system is composed of two types of devices, the transmitter-receiver pair (i.e., source and destination devices), and the ICW-capable intermediary cells. The source device modulates the information onto an ICW wave by generating a stimulus to the closest intermediary cell in the ICW channel. Then the calcium signal propagates through the intermediary cells via the internal and external pathways until the destination device is reached. In [26], authors presented the capabilities, limitations, and deployment scenarios of calcium signaling system.

2.1.2.8 Neurochemical Propagation

Neurotransmitters are endogenous chemicals that transmit signals across a synapse from a neuron to a target cell. The target cell can be another neuron or another junction cell as in the case of neuromuscular junctions (NMJ). Neurotransmitters are packaged into synaptic vesicles on the presynaptic side of a synapse and then

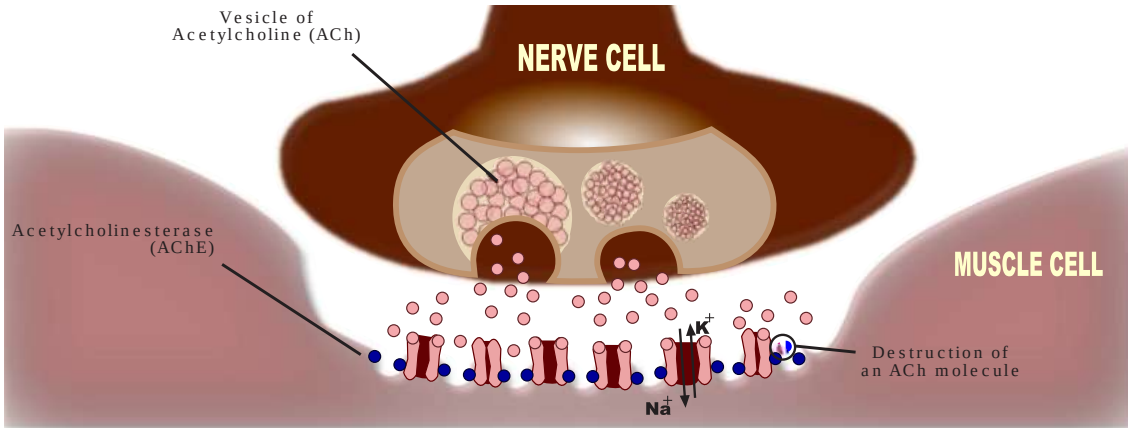


Figure 2.8: Representation of neuromuscular junction. Motor neurons release ACh molecules and they diffuses through the synapse and bind ACh receptors. To keep the communication between the nerve and muscle cell couple, the ACh molecules in the environment are cleaned via AChE molecules.

they are released into and diffuse across the synaptic cleft, where they bind to specific receptors in the membrane on target cell [96].

NMJs are one of the many occurrences in biological systems where two cells communicate with each other using an intermediary molecule that diffuse in the extracellular environment. A NMJ is depicted in Figure 2.8 and is a semi-closed environment between a pair of nerve and muscle cells, with a typical length of 10 to 100 nm [1, 97]. When muscles in a specific part of the body need to be contracted, the nerve cells in that region send a signal via neurotransmitters to the muscle tissue through these junctions to trigger the contraction [9].

As an action potential reaches the end of a motor neuron, neurotransmitters are

released into the synaptic cleft. In vertebrates, motor neurons release acetylcholine (ACh), a small molecular neurotransmitter, which diffuses through the synapse and binds to ACh receptors (AChRs) on the plasma membrane of the muscle fiber. This causes the ion channels at the cell membrane to open, which in turn allows the passage of Na^+ and K^+ ions. The increased cytosolic concentration of these ions causes the muscle cell to be contracted. The neurotransmitters stay in the bounded state for some time after which the bond degrades and the ACh molecules are again set free to the NMJ. The degradation of this bond is crucial to the muscle contraction procedure, allowing the muscle to relax and gradually revert back to its original resting position, awaiting further contraction signals. After the degradation of the bonds between ACh and AChR, the neurotransmitter molecules are very likely to re-bond with the receptors. Such an occurrence causes further unwanted muscle contractions and after a few muscle contraction signals, the NMJ will be filled with ACh molecules. This causes the ion channels linked to the AChRs to become inactive, which in turn blocks all further contraction signals. To resolve this issue, the ACh molecules in the environment should be removed from the NMJ after the muscle cell is successfully contracted. This is achieved by using a secondary type of molecule, called acetylcholinesterase (AChE), which resides in the muscle cell. AChE is a special kind of enzyme that is capable of attracting and hydrolyzing the ACh molecules into their two building blocks, Acetate and Choline. Since both

of these substructures are incapable of forming bonds with AChR, it is possible to say that in practice AChE molecules remove (or destroy) the ACh molecules from the NMJ.

2.1.3 Transmitter/Receiver Mechanisms and Components

The transmitter and the receiver of microscale molecular communication can be any machine in the microscale to nanoscale dimensions. For example, the transmitter and the receiver can be generated by modifying cells genetically [98–100], or by creating artificial cells [101–103]. Regardless of the method used, the transmitter requires at least the three components mentioned in the previous section: a unit for generation or storing information particles, a unit for controlling the release of information particles, and a central processing unit. Similarly, the receiver needs at least two components: a detection unit for sensing the information particles, and a processing unit for decoding and deciphering the intended message from the detection unit's measurements.

Physically, the central processing unit of the transmitter and the receiver may be implemented by synthesizing logic gates and memory into cells as shown in [104, 105]. The information particles can be generated by modifying a metabolic pathway of a biological cell, which then synthesizes and releases specific signaling molecules [99, 100]. Transfection, transfer via viral vectors, direct injection to germ line, and

transfer via embryonic stem cells are some of the methods for gene transfer for modifying a metabolic pathway [106–108]. Among these methods, viral vectors are a commonly used tool by molecular biologists to deliver genetic material into cells. This process is used for manipulating a living cell to engineer regulatory networks that can be used for communication. For example, in [109] a platform for biological system designers to express desired system functions via high-level programming language was introduced. After forming the genetic regulatory network code, living cells are manipulated via viral vectors. This process can be performed inside a living organism or in cell culture. Viruses efficiently transport their genomes inside the cells they infect for desired function. Main types of viral vectors are retroviruses, lentiviruses, adenoviruses, adeno-associated viruses, and nanoengineered substances [110].

To control the release timing, a synthetic oscillator can be introduced into a cell [111, 112], which with the help of the central processing unit acts as the release control module. Therefore, it is possible to have all the components of a transmitter synthesized into cells. Similarly, it is possible to synthesize receptors for a specific type of molecules into cells [113, 114]. Therefore, it is possible to have all the components required for the receiver inside a synthetic cell. More sophisticated processing units can also be designed as shown in [115], using novel materials and spin waves (spin waves are propagating disturbances in the ordering of magnetic

materials). It is also possible to shrink the current electronic technology to nano scales to create nanoscale processing units [116].

In nature, signals are received via protein structures called receptors. Therefore, these protein structures can be seen as receiver antennas. Receptors are the special protein structures that can bind to specific ligand structures. The binding occurs by intermolecular forces, such as ionic bonds, hydrogen bonds and van der Waals forces. The docking (association) is usually reversible (dissociation). Ligand binding to a receptor alters the receptor's chemical conformation and the tendency of binding is called affinity. The conformational state of a receptor determines its functional state [77].

Almost all ligand structures in nature capture and remove the information particles from the propagation environment during the detection process [77]. In some cases the information particles are destroyed after dissociation (e.g., acetylcholinesterase breaks down the molecules in neuromuscular junctions) [71, 77].

Most of the works in receptor theory focus on ligand-gated ion channels and G-protein coupled receptors [117]. Ligand-gated ion channels are a group of transmembrane ion channel proteins which open to allow ions such as Na^+, K^+, or Ca^{2+} to pass through the membrane in response to the binding of a chemical messenger (i.e., a ligand). These proteins are typically composed of a transmembrane domain which includes the ion pore, and an extracellular domain which includes the ligand

binding location (an allosteric binding site). G-protein coupled receptors constitute a large protein family of receptors that sense molecules outside the cell and activate inside signal transduction pathways and, ultimately, cellular responses.

Cells can also be created artificially by using liposome vesicles as the membrane encapsulating different functional proteins that together carry the task of the central processing unit, the particle generation and storage unit, receptors, and the particle release control unit for the transmitter and the receiver. Finally, it is possible to develop the transmitter and the receivers by synthesizing novel materials [118].

2.1.4 Power Source

In molecular communication, the transmitter, the receiver, and the propagation process may require energy. There are many different techniques for powering these components [119]. Sometimes the power is already present in the environment, or can be harvested from the environment through chemical processes. For example, diffusion relies on the thermal energy already present in the channel, and the transmitter and the receiver could be synthetic cells that run on ATP molecules already present in the environment. It is also possible to use other types of chemical reactions to create motion for propagation [120].

Different components can be powered using an external energy source (a power source that is not present or cannot be harvested form the environment). For

example, external magnets could be used to power and propel different components [121,122], while syringe pumps could be used to generate flows that would assist the transport of information particles. In [11], an energy model is derived for molecular communication that is based on diffusion. This model can be used to calculate the power required to generate, encapsulate, and release the information particles at the transmitter. By the help of the proposed model, energy requirement per bit is evaluated with different parameters. Results show that the energy budget affects data rate significantly. It is also shown that selecting appropriate threshold and symbol duration parameters are crucial to the performance of the system.

2.1.5 Potential Applications

There are many potential applications for molecular communication at microscales, such as medical applications [1, 123], control and detection of chemical reactions [124], computational biology [125, 126], better understanding of biology [127], environmental control and preservation [128], and communication among nanorobots [129]. The main driving force for engineering molecular communication is the medical field [130] with applications in lab-on-a-chip devices [131], cell-on-chips devices [132], point-of-care diagnostic chips [133], and targeted drug delivery [134]. In many of these applications, communication between different components or devices is the key to unlocking their true potential.

Another driving force behind molecular communication is recent advancements in the field of nanotechnology, which is making nanoscale devices such as nanorobots a reality [129]. Limited by their size, nanorobots can perform only simple tasks. Communication and cooperation among nanorobots can result in performing complex tasks [135]. Communicating nanorobots can be used for biomedical engineering applications, where nanorobots inside the body provide significant improvements in diagnosis and treatment of diseases [118, 136, 137]. In [138, 139], it is shown that communicating nanorobots can be much more effective at targeted drug delivery than uncommunicative nanorobots. In [140], nanorobots are proposed for detection of brain aneurysm. Nanorobots can also be used for transporting molecular payloads to cells [141], and altering the cell's behavior [142].

Ongoing development of nanoscale electronics, sensors and motors has further advanced the field of nanorobots [143]. For example, programmable bacteria can create computation capability for nanodevices [144]. Moreover, different techniques can be employed to power nanorobots [119]. For example, flagellated magnetotactic bacteria along with magnetic resonance imaging (MRI) can be used as medical nanorobots. Tracking and controlling of swarms of these bacterias are demonstrated in [145].

2.2 Macroscale Molecular Communication

Although use of molecular communication for long range communication between nanomachines has been proposed in [146], engineering a truly macroscale molecular communication system has not been considered in the past. In this dissertation, macroscale molecular communication is referred to a system where the dimensions of the transmitter/receiver, and the distance between them is a few centimeters or more. These systems have not been considered in the past partly because of the availability of wireless radio technology at these scales, which is very fast and reliable. Nevertheless, there are application for which use of radio technology is not possible or desirable, and molecular communication may be a suitable solution. For example, use of the radio based sensor networks for infrastructure monitoring in urban areas can be inefficient and unreliable [8]. Therefore, in this section the physical components required for macroscale molecular communication are discussed.

2.2.1 Information Particles

The information particle at macroscales can be any volatile chemical or gas, for over the air applications, or liquids for aqueous environments. For most practical applications, these chemicals must have low toxicity and be safe for humans and the environment. At macroscales, detecting individual molecules would be very

difficult. Therefore, in practice concentration of chemicals and chemical mixtures would be used as carriers of information.

2.2.2 Channel and Propagation

There are two main forms of propagation at macroscales: diffusion and flow based propagation. In the previous section, it was shown that the diffusion and flow based propagation of a single information particle can be simulated using Monte Carlo techniques. At macroscales, instead of a single or few information particles, large number of particles are used to transfer information. Therefore, a more general formulation, namely the diffusion equations, could be effectively used to model the propagation of chemicals. The diffusion equation allows us to talk about the statistical movements of randomly moving particles. It is important to know that, unlike Monte Carlo simulation equations presented in the previous section, the movement of each individual particle does not follow the diffusion equation. However, many identical particles each obeying the same boundary and initial conditions share some statistical properties and the diffusion equation deals with the average spatial and temporal evolution of these particles.

The diffusion equation is given by partial differential equation [147] as:

$$\frac{\partial c}{\partial t} = D\nabla^2 c, \tag{2.24}$$

where D is the diffusion coefficient, and c is the concentration at a particular spatiotemporal location. Therefore, c is a function of x, y, z, and t, where the first three variables represent a Cartesian coordinate and t represents time.

This partial differential equation can be solved using different initial conditions. One of the most useful initial conditions corresponds to a point source release of molecules at the transmitter. Assuming M_0 molecules are released by the point source suddenly at time $t = 0$, the initial conditions for 1-D, 2-D, and 3-D diffusion can be mathematically represented as

$$c(x, t = 0) = M_0\delta(x), \tag{2.25}$$

$$c(x, y, t = 0) = M_0\delta(x)\delta(y), \tag{2.26}$$

$$c(x, y, z, t = 0) = M_0\delta(x)\delta(y)\delta(z), \tag{2.27}$$

where $\delta(.)$ is the Dirac delta function defined by: $\delta(x) = 0$ for $x \neq 0$ and

$$\int_{-\infty}^{\infty} \delta(x)\ dx = 1. \tag{2.28}$$

Solving the diffusion equation with these initial conditions, the solution for 1-D,

2-D, and 3-D diffusion are obtained as:

$$c(x,t) = M_0(4\pi Dt)^{-1/2} \exp\left[-\frac{x^2}{4Dt}\right], \tag{2.29}$$

$$c(x,y,t) = M_0(4\pi Dt)^{-1} \exp\left[-\frac{x^2+y^2}{4Dt}\right], \tag{2.30}$$

$$c(x,y,z,t) = M_0(4\pi Dt)^{-3/2} \exp\left[-\frac{x^2+y^2+z^2}{4Dt}\right]. \tag{2.31}$$

The diffusion equation can also be solved for other initial conditions [72]. The channel response can be derived by adjusting the boundary conditions when there is an absorbing receiver [148]. The capture function can be obtained by integrating the channel response obtained from the differential equation system with these boundary conditions [78]. One can find the peak concentration and the peak time, by differentiating the impulse response with respect to time [78, 149]. Moreover, it is possible to find channel inversion in complex s-domain [150].

At macroscales, molecular diffusion itself can be a very slow propagation mechanism. For example, the diffusion coefficient of water vapor in air is about 0.3 cm^2/s. Therefore, assuming pure diffusion, on average, in 1 second a water molecule propagates about 0.5 cm, in 1 minute 4 cm, and in 1 hour 30 cm. However, it is possible to speed up the propagation by introducing flow.

Other forms of mass transport assist the propagation of particles at macroscales. These include:

- *Advection* – Advection refers to transport with the bulk fluid flow [151]. For example, information particles released inside a duct with air flows, are moved by bulk air flows.

- *Mechanical dispersion* – Mechanical dispersion, is the result of (a) variations in the flow pathways taken by different fluid parcels that originate in the nearby locations near one another, or (b) variations in the speed at which fluid travels in different regions [152]. For example, information particles moving through liquid flows in porous materials such as soil follow mechanical dispersion because of different paths they could travel. Similarly, in a pipe where there is a liquid flow, at the boundaries the flow speed could be slower than at the center of the pipe. This effect will disperse the information particles in the pipe.

- *Convection* – In thermodynamics convection is the fluid flows generated because of difference in temperature [153]. For example, in a room cold air, which is dense, moves downward while warm air moves upward. It must be noted that convection is also used to specify the combined advection-diffusion process in fluid mechanics. However in this dissertation, convection is fluid flows created due to temperature differences.

- *Turbulent flows* – Turbulent flows are random movements within a fluidic

flow [154]. Because this random movement is similar to molecular diffusion, the motion of particles inside turbulent flows are sometimes called turbulent diffusion. This is fundamentally different from the processes which determine molecular diffusion: in turbulent diffusion, it is the random motion of the fluid that does the transport, while in molecular diffusion it is the random motion of the information particles themselves. To indicate this difference in causation, the diffusion coefficient for turbulent diffusion is often referred to as the eddy diffusion coefficient. The value of the eddy diffusion coefficient depends on the properties of the fluid flow, rather than on the properties of the information particles. Most important is the flow velocity. Turbulence is only present at flow velocities above a critical level, and the degree of turbulence is correlated with velocity. More precisely, the presence or absence of turbulence depends on the Reynolds number, a unitless number which depends on velocity, width of the pipe, and the viscosity of the fluid. In addition, the degree of turbulence depends on the material over which the flow is occurring, so that flow over bumpy surfaces will be more turbulent than flow over a smooth surface.

When flow is present, the diffusion equation in (2.24) becomes the advection-diffusion equation (also known as convection-diffusion equation or diffusion equation with drift)

$$\frac{\partial c}{\partial t} + \nabla.(\mathbf{v}c) = D\nabla^2 c, \tag{2.32}$$

where $\mathbf{v}$ is the velocity vector and it is a function of spatial location and time. The D in (2.32) can be the diffusion coefficient for pure molecular diffusion or a combined molecular diffusion and turbulent diffusion coefficients [155]. In (2.24) and (2.32), it was assumed that the diffusion coefficient is the same in all spatial dimensions. However, it is also possible to derive analytical solutions when the diffusion coefficient is spatially variable [156]. For example, in oceans, the horizontal diffusivity along the water can be 10^7 times larger than the vertical diffusivity [157].

The solution for this partial differential equation in (2.32) can be obtained using different initial conditions. For example, assume that the space is 1D, and the flow is constant in the direction of positive x axis only. Moreover, assume that the boundaries are infinite, and M_0 information particles are spontaneously released at the origin at time zero. Then the solution of advection-diffusion equation becomes

$$c(x,t) = M_0(4\pi Dt)^{-1/2} \exp\left[-\frac{(x-vt)^2}{4Dt}\right], \tag{2.33}$$

where v is the speed in the positive x direction. The advection-diffusion equation can be solved for some other initial conditions as well.

2.2.3 Transmitter/Receiver Mechanisms and Components

At macroscales, the transmitter and the receiver require the same set of components depicted in Figure 2.2 at the beginning of the chapter. For the transmitter, a storage container is required for holding the information particles. It is also possible to generate the information particles using different processes. A mechanism must be set in place for controlled release of chemicals. For example, sprays could be used for controlling the release of information particles. In [158, 159], a technique for releasing complex blends of compounds in specific ratios, which mimics insect pheromones, is developed.

For detection at the receiver, chemical sensors could be used to detect the information particles. For example, metal-oxide gas sensors [160] are typically inexpensive sensors, that are capable of detecting the concentration of various types of volatile chemicals and gases. It is also possible to create more sophisticated sensors for detecting mixtures of chemicals as shown in [161].

The processing unit at macroscales can be a computer or a microcontroller, depending on the application. The power source could be electrical, solar, or any other source. At macroscales, different power sources and processing unit have already been well studied and developed.

2.2.4 Potential Applications

There are some potential applications for macroscale molecular communication. However, because radio communication is a well developed and established technology, most applications must be in areas where radio communication is not possible or not desirable. For example in [8], it is shown that current sensor network technology based on radio signals is not very reliable for some infrastructure monitoring applications. Moreover, there are still numerous problems that must be overcome for radio based communication in underground environments such as mines [162]. Finally, molecular communication system can be used for sending information in pipes and ducting systems, which could be beneficial in different industries such as oil and gas.

Macroscale molecular communication can also be used as a tool for studying animals and animal behavior. In nature, animals rely on pheromones, which are essentially chemical signals, for simple communication [10]. For example, ants use chemical trails for navigation and tracking [163]. Inspired by nature, a number of works have used molecular communication to mimic pheromonal communication in insects [158, 159, 161, 164]. Not only could this be a valuable tool to better understand animal behavior, it could also potentially be used to control these behaviors.

Another potential application for macroscale molecular communication is in

robotics, which was initially inspired from pheromonal communication and olfaction from nature [14, 165, 166]. Generally, previous works in this area can be divided into two main streams: pheromone based communication inspired by nature [15, 166–168], and plume or chemical tracking robots [165, 169–176]. Besides these applications, molecular communication can also be used for robotics search and rescue operations, and robot communication in harsh environments such as sewer systems.

3 Communication Engineering Aspect of Molecular Communication

Although molecular communication has been present in nature for billions of years, it was only recently that engineering molecular communication systems has been proposed [13]. Therefore, compared to modern radio based telecommunication systems, molecular communication is still in its infancy. In this section, some of the recent work on molecular communication through a communication-engineering lens are reviewed.

In this regard, the molecular communication literature can be categorized based on five partially overlapping topics:

- *Modulation techniques*: how information is encoded on the information carrying particles and chemical signals.
- *Channel models*: Because molecular communication channel can be very different from radio based communication channels, new channel models are

required for molecular communication.

- *Coding techniques*: Since the channel is different, there may be a need for new error-correcting codes for molecular communication systems.

- *Network architectures and protocols*: The network architecture and protocols used in molecular communication.

- *Simulation tools*: Because laboratory experimentation can be time consuming, laborious, and expensive, many different simulation environments have been developed for studying molecular communication.

Over the next sections each category is surveyed and the main concepts from each category are presented.

3.1 Modulation Techniques

Modulation is the process of varying one or more properties of the carrier signal according to the transmission symbol. In essence, transmission symbols are encoded in changes of the properties of the carrier signal. In traditional radio based wireless communication systems, the carrier of information is electromagnetic waves. Radio waves are sinusoidal signals that can be presented mathematically as

$$s(t) = A\sin(2\pi ft + \phi), \tag{3.1}$$

where A is the amplitude of the sinusoidal, f is the frequency, and ϕ is the phase. The amplitude affects the peak to peak height of the sinusoidal, the frequency affects the number of cycles per seconds, and the phase affects the amount of shift from the origin. Information can be modulated on the amplitude, the phase, the frequency, or any combination of these parameters. Figure 3.1(a) demonstrates the modulation of binary data (0 and 1) using amplitude, frequency, and phase.

In molecular communication, the carriers of information are very tiny information particles (e.g., molecules). In this scheme, information can be modulated on the following properties of the information particles.

- *Number of Particles*: Information can be encoded in the number of information particles released. If the number of information particles that are released is very large (e.g., at macroscales), concentration (i.e., the number of particles per unit volume) could be used instead of the number of particles.

- *Type/Structure of Particles*: It is possible to encode information in the type (or structure) of information particles released. For example, two different types of molecules could be used to encode one bit of information, or the structure of the DNA could be used to encode a large amount of data.

- *Time of Release*: Information can be encoded in the time of release of information particles. Another timing based approach is the pulse position

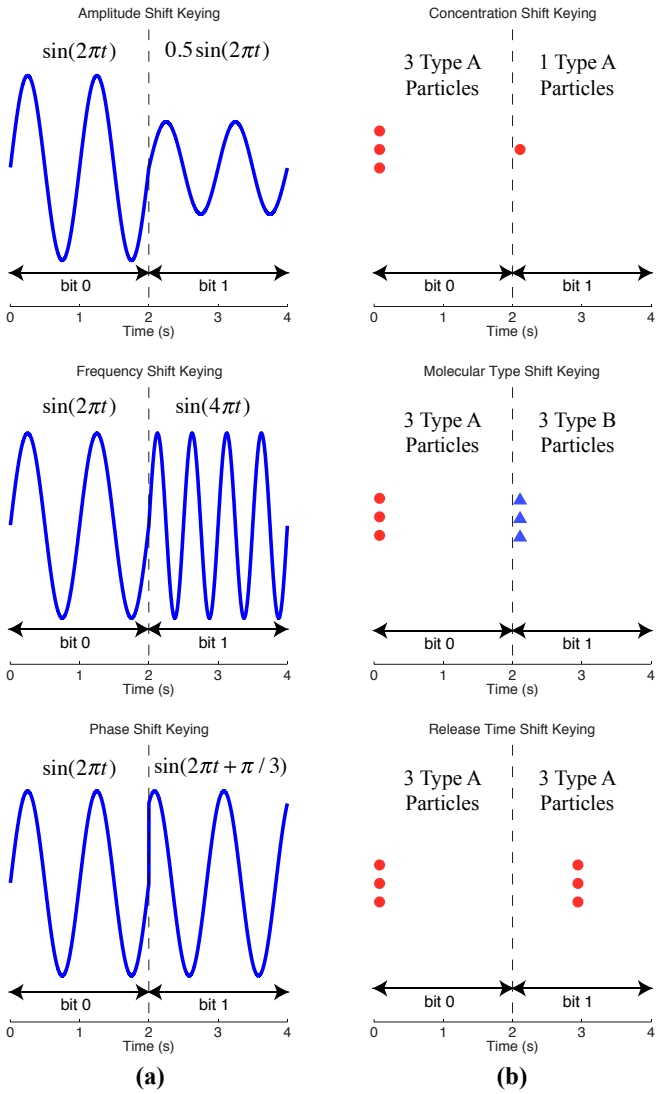


Figure 3.1: Modulation techniques in (a) the traditional radio communication (b) molecular communication.

modulation (PPM), which is similar to PPM in optical communication. To generate a pulse, large number of information particles are released almost instantaneously by the transmitter at a specific time.

Figure 3.1(b) demonstrates the modulation of binary data (0 and 1) using the number of molecules (concentration), the type of molecules, and their release timing. Just like radio based communication, it is also possible to use any combination of these modulation schemes.

In the rest of this section some of the recent work on modulation techniques for molecular communication are surveyed. Generally, these works can be broken down into three main categories: modulation schemes, inter-symbol-interference (ISI) mitigation schemes, and optimal detection at the receiver. Table 3.1 summarizes the different modulation schemes proposed in the literature.

One of the first works that considered modulation in molecular communication was [17]. In this work the authors considered two different modulation schemes for the diffusion-based propagation channels. In the first modulation scheme, a binary bit-0 was represented by concentration zero and the bit-1 by concentration Q. This modulation scheme is similar to on-off-keying from traditional telecommunication. In the second scheme, the concentration of information particles was varied according to a sinusoidal signal with a given frequency, and it is assumed information can be encoded in the amplitude and the frequency of this sinusoidal.

In [177] and [178], two new modulation schemes were proposed for diffusion based propagation channels. One was based on the number of molecules, where transmission symbols were encoded in the number of information particles re-

Table 3.1: Comparison matrix of modulation schemes.

Name	Abbreviation	Reference	Based on	ISI Reduction	Molecule Types
On-Off Keying	OOK	[17]	Concentration	No	1
Concentration Shift Keying	n-CSK	[177, 178]	Concentration	No	1
Molecular Shift Keying	n-MoSK	[177, 178]	Molecule Type	Moderate	n
Isomer-based Ratio Shift Keying	n-IRSK	[19]	Molecule Type & Ratio	Moderate	n
Pulse Amplitude Modulation	PAM	[179]	Concentration	No	1
Pulse Position Modulation	PPM	[179]	Time of Emission	No	1
Emission Time-based Modulation	-	[20, 180]	Time of Emission	No	1
Time-elapse Communication	TEC	[22]	Time of Emission	No	1
Molecular Transition Shift Keying	n-MTSK	[181]	Molecule Type & Concentration	Yes	2n
Molecular Array-based Communication	n-MARCO	[182]	Molecule Order	Yes	n
Molecular Concentration Shift Keying	n-MCSK	[183]	Molecule Type & Concentration	Yes	2n

leased by the transmitter. This modulation scheme was termed concentration shift-keying (CSK). The second modulation scheme was molecular type-based modulation, where the transmission symbols were encoded in the type of the information particles released. The authors named this modulation scheme molecular shift keying (MoSK), and they proposed hydrofluorocarbon based information particles as an example for this type of modulation. The effects of co-channel inference in a two-sender, two-receiver system was also considered in the work.

The choice of information particles for in-body molecular communication is very important. In [19], the authors proposed aldohexose isomers as an effective information particle for this use. As they pointed out the information-carrying isomers must be selected carefully such that they would not be harmful to the body as the isomers of hydrofluorocarbon proposed in [178]. The authors considered isomer based CSK, and isomer based MoSK. They also presented a new modulation scheme, which they called isomer-based ratio shift keying (IRSK). In this scheme, the information was encoded in the ratio of two isomers (i.e., the ratio of the concentration of two information particles).

It is also possible to encode information on the system response pulses of continuous diffusion models. Two different pulse based approaches were presented in [179]: pulse amplitude modulation (PAM), and pulse position modulation (PPM). In PAM, a transmission of bit-1 was represented by a pulse (a short burst of in-

formation particles) at the beginning of the bit interval, and transmission of bit-0 was represented with no pulse. In PPM, a bit interval was divided into two equal halves, and transmission of bits-1 and 0 were represented by a pulse in the first or the second half, respectively. Another work that considered modulation on the system response of continuous diffusion was [184], where information was encoded in features of the system response. The authors provided expressions for the peak's max, peak's width at half max, and peak's max delay as well as expressions for the energy of the impulse response. The information could then be encoded and detected on these features of the pulses.

Besides using the the number, type and continuous concentration pulse features for encoding information, the release timing of information particles could also be used for encoding information [20]. In [180], the authors considered diffusion based molecular communication, where a hybrid modulation scheme based on type and time of release of particles was proposed. This type of modulation is asynchronous, and could achieve higher information rates than the type based or timing based approaches. To derive their models, the authors presented their modulation scheme as an event-driven system.

Another release timing based modulation scheme called time-elapse communication (TEC) was proposed in [22] for very slow networks such as on-chip bacterial communication. In TEC, information was encoded in the time interval between

two consecutive pulses (i.e., two consecutive pulses of information particle transmission). The authors showed the feasibility of this technique for communication through bacterial fluorescence. In this setup information particles were released from the transmitter in an on-chip device, where they propagated using microflows until they arrived at the reception chamber. The reception chamber contained bacteria that would produce fluorescence light in proportion to the concentration of information particles. The authors showed that TEC can outperform on-off-keying when techniques such as differential coding was used.

Molecular communication channels typically have memory since some of the information particles that are released may remain in the channel and arrive in future time slots. Therefore, another important problem in molecular communication is the ISI due to delayed arrivals of molecules. Hence, modulation techniques should also consider ISI mitigation. One way for minimizing ISI is to combine CSK and MoSK modulation techniques together, which was proposed and named molecular transition shift keying (MTSK) in [181]. Binary version of this modulation technique utilized two types of information particles A and B. The modulator decided which molecule to send depending on the current bit and previously sent transmission bit. Using this scheme, the ISI can be reduced. It is also possible to reduce the ISI using enzymes [185]. In this scheme, enzymes that are present in the channel environment slowly degrade the information particles as they move through the

environment. It was shown that this could reduce the ISI and probability of error at the receiver.

Another ISI-reducing modulation technique based on the order of information particles released by the transmitter was presented in [182]. In this scheme, which the authors named Molecular ARray-based COmmunication (MARCO), bit-0 was encoded by releasing information particle a followed by information particle b, and bit-1 was encoded by release of b followed by a. This reduced the ISI. In a similar work [183], two types of molecules are used in an alternating fashion to reduce ISI. For example, the transmitter uses type-a molecules in odd time slots and type-b molecules in even time slots. As the molecule types are different in two subsequent time slots, the ISI is significantly reduced.

The demodulation and detection problem at the receiver are an important part of molecular communication systems. In [186], a receiver circuit was proposed for frequency shift-keying-modulation presented in [17]. For frequency detection, an enzymatic circuit consisting of a number of interconnected chemical reaction cycles was proposed. The circuit was then modeled using reaction-diffusion equations and reaction kinetics equations, which were all partial or ordinary differential equations. The feasibility of the detection scheme was then verified using simulations.

Optimal detection for molecular communication receivers was considered in a number of works. Optimal receiver design for MoSK modulation scheme was con-

sidered in [187]. The receiver was optimal in the sense that it minimized the probability of error according to the maximum a-posteriori (MAP) criterion. The authors first showed that for diffusion based propagation, the channel is linear and time-invariant. They then used this assumption to derive the optimal receiver design. Similarly, in [188,189] optimal receiver detection model was derived for binary concentration modulation scheme using MAP criterion. In [190], the authors considered receiver design for on-off-keying modulation scheme based on MAP, and maximum likelihood criterion. They used the continuous diffusion equation and considered the random propagation as noise. Using some simplifying assumptions it was shown that this noise is Gaussian. The sequential detection in the presence of ISI was then achieved by considering the equivalent problem of estimating the state of a finite-state machine. Finally, an optimal detector for the molecular communication channels with enzymes that help reduce the ISI was designed in [191].

3.2 Channel Models

Information theory is the mathematical foundation of any communication system [192]. One of the most basic theoretical models for a communication system is the channel model. Physically, the channel is the environment over which the transmission signal propagates. Mathematically, the channel relates the transmis-

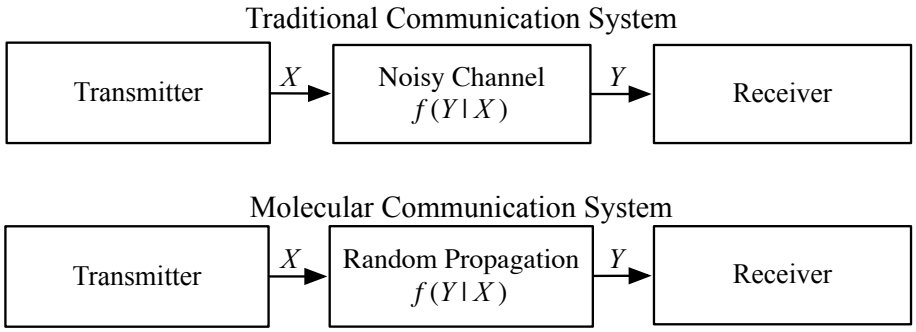


Figure 3.2: Block diagram representation of communication channels.

sion information, which would be the channel input, to the received information, which would be the channel output. There is an uncertainty associated with a communication channel. The received signal at the receiver can be different from the transmitted signal because of the noise introduced by the channel. In molecular communication one source of noise is the random propagation of information particles. Figure 3.2 shows a block diagram representation of the channel.

The mathematical model of the channel can be described as follows. Let $X \in \mathcal{X}$ be the symbol (i.e., message) that a transmitter wants to send to the receiver, where $\mathcal{X}$ is the set of all possible symbols (i.e., messages) that the transmitter could send. In typical radio based communication, the signal sent by the transmitter gets attenuated and corrupted before it reaches the receiver. For example, the source of this corruption could be thermal noise from the electronic equipment, interference with other signals, and signal fading. Therefore, at the receiver the detected symbol

(i.e., message) could be incorrect. Let $Y \in \mathcal{Y}$ be the detected symbol at the receiver, where $\mathcal{Y}$ is the set of all possible detection symbols. Because the receiver has no prior knowledge of the original transmission symbol, the actual transmitted symbol can be considered a random variable X. Also, because the original transmitted symbol can be corrupted by noise, the detected symbol can be represented by a random variable Y. The communication channel can therefore be characterized by the conditional probability mass function (PMF) $P(Y|X)$.

In molecular communication channels, the main source of noise is the random motion of the information-carrying particles. All propagation schemes described in the previous sections are random in nature; although some exhibit more variance (pure molecular diffusion), while some are more directional and exhibit lower variance (active transport).

There are two important measures associated with any noisy channel in a communication system: the achievable information rate, and the channel capacity [192]. Achievable information rate is a theoretical measure of how many bits can be transmitted across the channel in a given period of time. For a discrete memoryless channel, where the current received symbol is independent of all the previous transmission symbols, achievable information rate can be computed from the PMFs

$P(Y|X)$ and $P(X)$ as

$$I(X;Y) = E\left[\log_2 \frac{P(y|x)}{\sum_{x\in\mathcal{X}} P(y|x)P(x)}\right], \tag{3.2}$$

where $I(X;Y)$ is also called the mutual information between X and Y.

Channel capacity is the maximum achievable information rate, which is the maximum number of transmission bits that can be transferred over the channel in a given period of time. Mathematically channel capacity is represented as

$$\mathcal{C} = \max_{P(x)} I(X;Y), \tag{3.3}$$

where $\mathcal{C}$ is the channel capacity. For a discrete memoryless channel the mutual information is concave with respect to $P(x)$ [192]. Therefore, if $P(Y|X)$ is known, Blahut-Arimoto algorithm [193,194] could be used to find the maximizing $P(x)$ and the channel capacity for the channel.

Unfortunately, most practical molecular communication channels have memory and suffer from inter-symbol interference (ISI). In this case, channel capacity is calculated differently [195]. Let $\mathbf{x}^n$ be a sequence of n consecutive transmission symbols, and $\mathbf{y}^n$ be the corresponding sequence of n consecutive received symbols.

If the channel is information stable then the channel capacity is given by [196]

$$\mathcal{C} = \liminf_{n\to\infty} \sup_{\mathbf{x}^n} I(\mathbf{X}^n; \mathbf{Y}^n). \tag{3.4}$$

The makes calculation of channel capacity much more difficult compared to the discrete memoryless channels.

Because channel models play a central role in designing communication systems, one of the first problems that communication engineers tackled was modeling the molecular communication channel [20, 197]. As mentioned in Section 2.1.2 different propagation schemes are possible for molecular communication channels. Therefore, for each propagation scheme and modulation technique, a different channel model can be derived. Table 3.2 shows a comparison of most channel models proposed in the literature.

3.2.1 Diffusion Channel Models

In the previous chapter, Sections 2.1 and 2.2, the diffusion propagation scheme for both microscale and macroscale molecular communication were described. The channel models for diffusion-based molecular communication can be divided into three broad categories: models based on continuous diffusion equations, discrete models, and diffusion models with flow. Besides the propagation, depending on the

Table 3.2: Comparison matrix of channel models.

Description and Modulation	Propagation	Reference	Based on	Considers ISI	Symbol Set
On-off keying (binary-CSK)	Diffusion	[197–200]	Number of Particles	No	$\{0, 1\}$
On-off keying (binary-CSK)	Diffusion	[201–204]	Number of Particles	Yes	$\{0, 1\}$
On-off keying (binary-CSK) with particle degradation	Diffusion	[205]	Number of Particles	Yes	$\{0, 1\}$
On-off keying (binary-CSK) with energy constraints	Diffusion	[11]	Number of Particles	Yes	$\{0, 1\}$
Continuous Diffusion Modulation and Ligand Receptors Demodulation	Diffusion	[64, 65, 206, 207]	Continuous Diffusion	Yes	$\mathbb{R}^+$
CSK using bacterial colonies as Tx and Rx	Diffusion	[208–210]	Concentration	No	$\mathbb{R}^+$
Contact-Based Neurochemical	Diffusion	[211]	Concentration	No	$\{0, 1\}$
Reaction Diffusion Channel	Reaction Diffusion	[212]	Number of Particles	No	$\mathbb{N}$
Channels with spontaneous emission, absorption and particle generation	Diffusion	[213]	Continuous Diffusion	No	-
Timing Channels - Time of Release Modulation	Diffusion	[20]	Hitting Time	No	$\mathbb{R}^+$
Timing Channels - Time of Release Modulation	Diffusion with Flow	[66, 214]	Hitting Time	No	$\mathbb{R}^+$
CSK for Diffusion Channels with Flow - Noise Models	Diffusion with Flow	[215]	Continuous Diffusion	No	$\mathbb{N}$
Kinesin Moving Over Stationary Microtubule Tracks	Molecular Motors	[69]	Number of Particles	No	$\{0, 1\}$
CSK Gap Junction Channels	Gap Junction Diffusion	[216–218]	Number of Particles	No	$\{0, 1\}$
DNA Propagation Using Bacteria	Bacteria	[28, 91]	Type/Structure of Particles	No	$\mathbb{N}$
Particle Flow Inside Arteries	Diffusion with Flow	[219]	Continuous Diffusion	No	$\mathbb{R}^+$
Intercellular Signal Transduction Channels	-	[220]	Ligand Receptors	Yes	$\mathbb{R}^+$
Microfluidic Channel Models	Diffusion with Flow	[221]	Continuous Diffusion	No	$\mathbb{R}^+$

modulation scheme the channel formulations would be different.

The most popular molecular communication channel in the literature is the time slotted binary-CSK channel employing pure diffusion-based propagation. In this scheme, if the transmission bit is 0, no molecule is transmitted and if the transmission bit is 1, A molecules are transmitted. In its simplest form $A = 1$ (i.e. 1 particle is released to represent bit-1 and 0 to represent bit-0).

In [197], with using some simplifying assumptions, it was shown that the time-slotted binary-CSK channel can be represented as a binary symmetric channel when ligand-receptors are used for detection at the receiver. The maximizing input probability distribution for this model was presented in [198]. The model was then extended to broadcast channels, relay channels and multiple-access channels in [199, 200].

Because in diffusion molecules move randomly, there could be ISI between consecutive transmission symbols. In order to reduce the ISI effect, a symbol dependent symbol duration could be used, where the transmission time for the current symbol is selected based on the current and the previous symbols. This type of channel was modeled as a Markov chain in [201], and achievable rates were derived. The authors then completed their analysis in [202], where they modeled the ligand receptor channel as a Markov Chain and derived the capacity and achievable rates for this channel.

For the binary-CSK channel with $A = 1$, if there is no interference with the other molecules from previous transmissions, the channel could be represented as a z-channel [203] . Moreover, the authors in [203] considered a very simplified ISI model, where the interference could affect only the next transmission (i.e. the channel memory length is 1). The achievable information rates were then calculated using computer simulations and the results were presented. In [204], a more general ISI model for this channel was considered. It was shown that this communication scheme can be represented as an additive noise channel, where the noise has a Poisson-Binomial probability mass function (PMF). Using this channel model maximizing input probability distributions was estimated.

In practice, information particles may degrade over time while propagating in the channel. In [205], binary-CSK channels where the molecules degrade over time according to the exponential probability distribution were considered. Then the channel capacity was compared for different degradation rates and different binary-CSK modulation strategies. Another practical problem in molecular communication is the limited energy and resources available at the transmitter and the receiver. Energy models for the end-to-end binary-CSK communication channel, which could be used to calculate the maximum number of information particles that can be generated and transported in a particular system was presented in [11]. Using this model, an optimal transmission strategy that considered the limited energy

and resources of the system, and maximized the channel capacity was presented.

Another channel model for diffusion-based molecular communication is based on the continuous solution of the diffusion equations presented in Section 2.2. In this case, the transmission signal is an analog concentration signal $x(t)$, and the received signal is $x(t)$ convolved with the impulse response of diffusion equation plus some noise. One of the first works that considered this approach was [64], where end-to-end gain (attenuation factor) and delay function were derived for continuous molecular diffusion channels with ligand receptors for detection. To derive these terms the authors used techniques from circuit theory to calculate the transfer function for the transmission, diffusion, and reception process. Using these transfer functions they calculated the end-to-end gains and delays.

In practice, solutions for continuous diffusion channels are just the average behaviour across many trials. Moreover, the number of particles that are released by the transmitter are continuous rather than discrete. Therefore, in [206] a more realistic model of continuous diffusion based molecular communication channel was proposed by considering the discrete nature of the number of information particles and randomness of the diffusion propagation. It was assumed that discreteness can be modeled as quantization noise and the random propagation as an additive noise which are added to the solution of continuous diffusion equations. Similarly, in [207] the authors captured the inherent randomness in ligand-binding reception

as noise. Simplifying assumptions and Markov chains were used in the derivation of a closed-form expression of random ligand binding process.

In [65], the authors derived a closed-form expression for the lower bound on channel capacity of continuous diffusion-based molecular communication channel in gaseous environments. They showed that this lower bound capacity increases linearly with the bandwidth of the transmission signal. In this derivation, it was assumed that the receiver and the transmitter are perfect, and the derivation was purely based on the diffusion transport process.

Microbial colonies are one of the prominent transmitter and receiver entities in molecular communication. In this scheme, the transmitter and the receiver are made from colonies of genetically modified bacteria that would collectively transmit and receive information particles. The mathematical models for such a receiver were presented in [208], where the sensing capacity of bacterial nodes was developed. To derive the capacity, it was assumed that the bacteria react to the concentration of information particles by emitting light. The higher was the concentration of the information particles at the receiver node, the higher was the luminance of the bacterial colony. The optimal input distribution and the corresponding sensing capacity were obtained based on the received power. Models for the transmitter and the communication channel between the transmitter and the receiver was considered in [209], where it was assumed that the transmitter was

made of colony of bacteria and the transmission of information particles was a two stage process. It was assumed that signalling molecules (type A molecules) were used to control the number of information particles (type B molecules) released by the transmitting colony. The released information particles then diffused until they arrived at the receiver colony, where they were detected by ligand receptors, which then initiated a fluorescent light generation response that would be dependent on the concentration of the received molecules. The end-to-end channel model and some expressions for the capacity were then presented.

Another work that considered the transmitter and the receiver in diffusion-based molecular communication is [211], where the transmitter and the receiver were assumed to be genetically modified cells that are moving themselves. The information was transported through a three phase process: collision, adhesion, and neurospike transmission. In this scheme, two cells (the transmitter and the receiver) moved according to the laws of diffusion, until they collided and stuck to one another. The information was then transported from one cell to another using molecular neurospike communication. Through some simplifying assumptions, the authors derived the channel models for this communication scheme.

In practice, information particles may chemically react with other molecules present in the channel. In [212], a model based on reaction-diffusion master equation, which is a well known model in physics and chemistry for jointly modeling

diffusion and chemical reactions, was proposed. The model which is called reaction-diffusion master equation with exogenous input (RDMEX), discretizes time and space and uses a Markov model, where the states of the model are the number of information particles in discrete locations of space. During each discrete time step, the state of the Markov model changes according to the laws of diffusion and chemical reactions. Another work that considers the effects of the channel on the information particles is [213], where a stochastic model for three different environmental effects are considered: particle absorption (where the particles are absorbed in the environment), particle generation (where a single information particle becomes two information particles through reactions in the medium), and spontaneous emission (where information particles are spontaneously generated through chemical reactions). Through simplifying assumptions, a model was derived using birth-and-death processes. The model was then applied to diffusion based molecular communication with and without flow.

As mentioned in the previous section, it is possible to modulate information on the time of release of information particles. One of the first works that considered diffusion-based timing channel was [20], which was based on the 1-dimensional first hitting time distribution. The channel capacity for this model was derived and it was shown that for this simplified channel, more than one bit of information can be transmitted per information particle. As mentioned in Section 2.1.2, flow could be

used to increase the speed of propagation. In [66], it was shown that such a channel, employing time of release modulation, could be modeled with an additive inverse Gaussian noise (AIGN) term. Lower and upper bounds for the channel capacity of AIGN channel was then derived. It was shown that these bounds were tight for low flow rates, but the bounds became separated as the flow rate increased or as it approached zero. The derived channel model was then used to design a receiver based on a maximum likelihood estimator. Tighter bounds for AIGN channel were presented in [214].

3.2.2 Other Channel Models

Many different channel models are proposed for diffusion-based molecular communication. However, for other propagation schemes presented in the previous chapter, there are not many channel models in the literature. In this section, some of the works on these propagation schemes are surveyed.

One of the most promising propagation schemes besides diffusion is active transport using molecular motors. Kinesin-microtubule motility is the most widely used form of active transport. As it was discussed in Section 2.1 in the previous chapter, two forms of transport are possible: kinesin moving on stationary microtubule tracks, or microtubule filaments gliding over kinesin-covered substrate. Generally, active transport is a suitable propagation scheme for on-chip applications.

Previous works on molecular communication have only considered models for kinesin moving over microtubule tracks. One of the early works on modelling this channel is [69], where simulations were used to compare the arrival probabilities of diffusion channels, and active transport channels based on a microtubule track that connects the transmitter to the receiver. In this setup, kinesin motor proteins carry the information particles on the microtubule tracks. A hybrid approach based on combination of these propagation schemes was also considered. It was shown that diffusion had low rate of delivery and lower channel capacity for separation distances in the order of micrometers compared to the active transport and the hybrid schemes.

Gap junction channel is another propagation scheme that has been considered in a number of previous works. In this scheme, it is assumed that the transmitter and the receiver are cells or artificial lattice structures that are connected together through a series of similar cells or artificial lattice structures. Each cell or lattice structure is connected to its neighbors through gap junction channels. Calcium ions are generated by the transmitter and propagate through the gap junction channels, traveling from lattice to lattice until they arrive at the receiver. Binary-CSK modulation scheme in such a molecular communication system was considered in [216]. The channel was then simulated and achievable information rates were presented. In [217], the channel capacity of the calcium signaling system based

on an inter-cellular calcium wave model for astrocytes was investigated. Calcium waves are formed by the cytosolic oscillations of Ca^{+2} ion concentration, which propagates through neighboring cells via secondary messenger molecules. Channel capacity of the investigated system was analyzed under different noise levels and symbol durations. In [218], the gap junction communication between cardiac muscle cells, called cardiomyocytes, was modeled. In particular the propagation of action potential, which helps to regulate the heart beat was modeled as a communication channel and then simulated to calculate the channel capacity.

Bacteria are another effective tool for actively transporting information particles from the transmitter to the receiver. A molecular communication channel in which information particles such as DNA strands are embedded in flagellated bacteria that then transport the particles was modeled in [28, 91]. It was assumed that the receiver uses attractant molecules to attract these bacteria. The motion of the bacteria from the transmitter to the receiver was then modeled and simulated.

Communication theoretic channel modelling has been used in a number of works to represent interesting biological processes. In [219], a channel model for blood flow through arteries was developed using harmonic transfer matrix and some simplifying assumptions. Two different models were used for small and large arteries, and the models were then combined to form a complete model of blood flow through the arteries in the body. A model for drug injection propagation in the arteries was then

presented based on this model. In [220], simple intercellular signal transduction channels were considered, where information particles were detected using ligand receptors. The paper abstracted the concept of propagation and assumed the input to the channel is the concentration level at the vicinity of the receiver, which was equipped with ligand receptors. The detection process was then represented as a discrete-time Markov model. It was shown how the capacity can be obtained for this channel. It was also shown that the capacity achieving input distribution is iid, which is unusual for channels with memory. Furthermore, it was demonstrated that feedback does not increase the capacity of this channel.

Microfluidics is another important area where molecular communication can have a high impact. Channel models for flow-based molecular communication in microfluidic channels was presented in [221]. First, the transfer function of the straight channel, the turning channel, bifurcation channel, and combining channel was derived. Then it was shown that the overall transfer function of any channel configuration (i.e. any combination of straight, turning, bifurcation, and combining channels) can be calculated using the transfer function of these channels. Using this model, the authors demonstrated that finite impulse response (FIR) filters could be implemented on a microfluidic device.

3.3 Error Correction Codes

In traditional communication systems, channel codes are used to mitigate the effects of noise and fading that is introduced into the system by the channel and electronic components. In essence, channel coding introduces redundancy which can then be used to detect and correct errors. For example, in a simple repetition code, a binary bit-0 is encoded as 000, and a binary bit-1 is encoded as 111. At the receiver the majority number of bits are used to decode the bit. For example, 011 is decoded as bit-1, and 010 is decoded as bit-0. It is fairly trivial that such a simple repetition code can correct one bit errors.

Capacity-approaching codes are codes that allow information transmission rates that are very close to the maximum theoretical limits of the channel (i.e., channel capacity). Many different capacity-approaching codes such as low density parity check (LDPC) codes and turbo codes have been developed for the traditional communication systems [222]. However, the encoder and the decoder for these channel codes are computationally complex, and as such they may not be practical for many molecular communication systems, especially at microscales. Furthermore, because the nature of the noise is different in molecular communication channels, it is not clear if better error correcting codes exist for these channels. For example, molecular communication channels typically have memory. Only a limited amount

of work has been done on channel codes for molecular communication, and there are still many open problems in this area.

The early works on coding schemes for molecular communication considered applying the previously devised coding schemes for radio communications. One of the early works that considered channel codes for molecular communication systems was [223], where Hamming codes were proposed as simple error correcting codes for on-off-keying in diffusion based molecular communication. First, it was shown that these error correcting codes can improve the probability of bit error rate when a large number of information particles were used for representing each bit. Uncoded transmission outperformed coded transmission, however, when a smaller number of molecules were used to represent each bit. This occurs because of the extra ISI introduced by the extra parity bits. The authors then modeled the extra energy needed for transmitting the extra parity bits and showed that coding was not energy-efficient when the separation distance between the transmitter and the receiver was small. In [224], convolutional coding schemes were applied to the same channel and it was shown that convolutional codes could improve the bit error rate of these molecular communication channels.

A few works have developed new coding schemes tailored for molecular communication channels. In [225], a new code family called ISI-free code family was introduced for diffusion based molecular communication channels with drift. It was

assumed that the type of the information particles is used to encode information. In this coding scheme, the crossovers up to level-ℓ between consecutive codewords and within the current codeword were eliminated. It was shown that these code families can improve the bit-error rate performance of molecular communication channels compared to uncoded transmission, under similar throughputs. Another work that developed a new coding scheme for molecular communication timing channels was [226]. In this work, zero-error codes were developed which guaranteed successful recovery of information bits at the receiver, assuming the channel has finite memory (which is a valid assumption for practical channels).

3.4 Architectures, Protocols, and Optimal Design

Because molecular communication is still in its infancy, most of the research has been focused on modeling the channel as well as the modulation and coding schemes necessary for setting up a reliable communication link. However, once the link is established other important aspects of communication networks are the network architecture and protocols. The Open Systems Interconnection (OSI) model is a conceptual model that characterizes and standardizes the internal functions of a communication system by partitioning it into abstraction layers. For example, TCP/IP model of the Internet partitions the communication system into 4 layers: link layer, Internet layer, transport layer, and the application layer.

When considering molecular communication, one might think that it may be possible to use the same layered architecture over a molecular communication physical link. However, one issue is that the goal is not to connect electronic devices. Instead, the goal is to connect biological entities. Therefore, it is not clear if using the same layering structure would be appropriate for molecular communication networks. Moreover, even if the conventional layering architectures are used, it is not clear how each layer should communicate with the layer above or below. A detailed survey on how molecular communication networks could be layered was presented in [227].

To solve some of these issues, IEEE Standards Association has started the IEEE P1906.1 project to provide recommended practice for nanoscale and molecular communication framework [228]. Two of the goals of this project is as follows: 1) To provide a common framework for developing molecular communication simulators. This includes interconnecting systems of multiple types of simulators. 2) To provide a common abstract model that would enable theoretical progress to proceed from different disciplines with a common language. This framework serves as a recommended practice for additional nanoscale networking standards as industry becomes more involved. In the rest of this section, some of the important works on network architecture, protocols, and optimal design are highlighted.

One of the early works that looked at network architectures for molecular com-

munication was [229]. The authors assumed that the intended transmission information can be divided into two categories: sensor data, and command data. They then assumed that the command data has higher priority and as such would be encoded with error correction codes, where as sensor data are uncoded. To solve the computing problem required for carrying out the network protocol, the authors proposed DNA and Enzyme based computations. In particular, DNA based computations was proposed for application interface stack, which encodes messages and addresses, network stack which routes messages, and error correction stack which encodes and decodes error correction codes. Enzyme based computation was proposed for link switching stack.

Another work that explored the network architecture for molecular communication was [28], where each nano-node could communication with other nano-nodes over the short range (nm to μm) using diffusion, the medium range (μm to mm) using bacteria transport and catalytic nanomotors, and the long range (mm to m) using pheromones. In the proposed architecture, it was assumed that each nanodevice was connected to a gateway over the short range and the gateways are distributed over the medium to long range. It was assumed that instead of bits, DNA base pairs are used for addressing and carrying information.

Genetically engineered bacteria can play an important role in molecular communication systems. A number of works explore network architectures based on

genetically modified bacteria. In [68], the authors proposed a routing mechanism for bacterial based molecular communication. In this scheme, it was assumed that each node in the network released a chemoattractant to attract the bacteria from neighboring nodes. It was also assumed that information particles are DNA sequences that are embedded in bacteria. Node A embedded a DNA sequence inside bacteria that pass through neighboring node (node B) until they arrive at the destination (node C). At the relay node (node B), it was assumed that the bacteria transfer their information to a new set of bacteria that are sensitive to node B's neighboring chemoattractants. The DNA sequences were transferred from bacteria to bacteria using bacterial conjugation, a process that involves transfer of single stranded DNA from one cell to another. Since there may be errors during conjugation, in [230] use of antibiotics were proposed to kill bacteria with incomplete DNA sequences. In this scheme, bacteria resistance genes were inserted at the end in the DNA sequence carrying the information. If bacterial conjugation results in an incomplete DNA sequence, the bacteria carrying the information particles was assumed to be killed by the antibiotics.

In a system with multiple sources and multiple receivers, it is important for the bacteria to deliver the DNA sequence to the correct destination. In [231], the authors proposed establishing a coordinate system using beacon nodes. This idea was analogous to the way global positioning system (GPS) works. It was

assumed that each beacon node releases a specific type of chemical at a specific rate. Therefore, there is a concentration gradient generated by each beacon. Distances from a receiver to a beacon was measured using this concentration gradient and the fading of concentration over a distance. The information particles (DNA sequences) was then assumed to be carried by genetically engineered bacteria to the correct destination using chemotaxis.

Another work that considers networks with multiple transmitters and receivers was [232], where the authors considered optimizing the transmission rate of each transmitter such that the overall throughput and efficiency was maximized. In their model, it was assumed that there are N transmitters and M receivers. At the receiver, it was assumed that the information particles were detected through Michaelis Menten enzyme kinetics. The throughput was defined as the number of information particles processes at the receiver per unit time. The efficiency was defined as the throughput divided by number of information particles transmitted per unit time. A simplified optimization model where all transmitters are at origin and all receivers are at distance r was presented, and the solution was provided with an upper bound on throughput and efficiency. A feedback mechanism was also presented that would adjust the optimal transmission rate in dynamic environments.

Another important area of interest in molecular communication is on-chip ap-

plication. Generally, works in this field can be divided into two categories: those based on active transport propagation schemes, and those based on microfluidics. In [70], an architecture for design of self-organizing microtubule tracks for on-chip molecular communication was proposed. In this scheme, microtubule tracks act as railways or wires connecting the transmitter to the receiver. In particular, two different approaches were considered: one based on polymerization and depolymerization of microtubules, and the other based on molecular motors for reorganizing the tracks. Preliminary *in vitro* experiments were conducted then to investigate the feasibility of the proposed techniques.

Droplet microfluidics is another promising form of communication for on-chip applications. One of the first works that considered the problem of networking in droplet microfluidics was [233]. The authors first gave an overview of this field, and then discussed some of the issues and problems in creating network architecture for hydrodynamic based droplet microfluidics. It was shown that a ring topology, where each processing module is connected to a main ring through a microfluidic network interface (MNI) circuit could be an effective architecture for networking different components in these systems. In this scheme, the main ring starts and ends at the central processing unit, and the network interface circuits control the flow of droplets through and from each processing module. A general introduction to hydrodynamic microfluidics was also provided in [234].

The similarity between Ohm's law and Hagen-Poiseuille's law was used in [235] to design a switch circuit for the MNI. In this scheme, the circuit was designed in such a way that depending on the distance between a header droplet and a payload droplet, the payload droplet was either delivered to the corresponding processing unit or it continued flowing in the ring. A medium access control circuit was also proposed for releasing a payload into the ring such that the released droplet would not collide with the other droplets flowing in the ring. In [236], channel capacity expressions for the different droplet encoding schemes were provided.

For many practical molecular communication systems, the number of information particles may be limited. In [237], molecular communication in confined environments was considered, where it was assumed that the number of information particles within the environment was constant. Therefore, different nodes needed to harvest the information particles that were within the environment for communication. The authors then simulate this particular architecture and present some preliminary results based on different harvesting and communication protocols.

It is possible to exploit the properties of molecular communication channels to develop clever protocols. In [238], a protocol for measuring the distance between two nanomachines by exchanging information particles and measuring their concentration was proposed. In this scheme, first node A released N_a particles of type-a into the environment, where they diffused. At node B the change in concentration

of type-a particles was measured. Then, N_b particles of type-b were released into the environment according to some property of the measured concentration of type-a particles. By measuring the concentration of type-b particles back at node A, the distance was estimated. Different detection processes such as round trip time and signal attenuation were considered.

Another interesting protocol for blind synchronization between different nodes in a diffusion-based molecular communication system was proposed in [239]. In particular, an MoSK channel was assumed, where during each symbol duration different molecules are used to encode different messages. Then a non-decision directed maximum likelihood criterion was used to estimate the channel delay, which would be used for synchronization. The Cramer-Rao lower bound for the estimation was also derived, and the results were compared with simulations.

3.5 Simulation Tools

Experimental study of molecular communication systems is very difficult because of the expensive nature of wet-lab experimentation, and the fact that performing these experiments can be very time-consuming and laborious. Therefore, previous research activities on molecular communication have heavily relied on simulations to verify and evaluate the new communication solutions. These simulators are required to precisely track the behavior of information-carrying particles within a realistic

Table 3.3: Comparison matrix of molecular communication simulators.

	dMCS	**N3Sim**	**MUCIN**	**NanoNS**	**BINS**	**BNSim**
Development language	Java	Java	MATLAB	NS-2, C++	Java	Java
Parallelization	Yes	No	No	No	No	No
Open source	No	Yes	Yes	No	No	Yes
Propagation	Diffusion	Diffusion	Diffusion	Diffusion	Diffusion	Bacteria
Track each carrier	Yes	Yes	Yes	No	Yes	Yes
Reception	Absorption	Sampling	Absorption	Berg, Gillespie	Receptors	Receptors
Imperfect reception	No	No	Yes	No	Yes	No
Support unbounded medium	No	Yes[a]	Yes	Yes	Yes	No
Environment dimensions	3-D	2-D, 3-D[a]	1-D $\sim$ 3-D	3-D	3-D	3-D
Molecule interactions	No	No	No	No	Yes	No
Sending consecutive symbols	No	Yes[a]	Yes	No	No	Yes

[a] Possible only under specific conditions.

environment. In this section, an overview of some of the most recent molecular communication simulators in the literature are provided. Table 3.3 provides a comparison matrix for these simulators.

In [240], distributed simulations based on high level architecture (HLA) were proposed and analyzed for molecular communication. The authors mostly focused on confined space and parallelism gain, and did not consider the effect of consecutive symbols in their simulator, dMCS. In [241], NanoNS simulation platform, written in C++ and Tcl, was introduced based on the well-known NS-2 simulation platform. The multi-particle lattice automata algorithm was implemented in NanoNS, which divided the propagation medium into a mesh grid. Thus, the exact particle position was not considered, and both time and space were discretized.

In [242], amplitude-based modulations were analyzed using a custom simulator, with non-absorbing receivers (i.e. information particles continued free diffusion after they reacted the receiver). In [243], a simulation framework called N3Sim was introduced, where the reception process at the receiver site was similar to the model in [242] (i.e. non-absorbing receivers). In both of these simulators, it was assumed that the transmitters are a point source. In reality, the transmitter nodes may have a volume and they could reflect the emitted molecules. N3Sim is a Java package that simulates particle diffusion in a 2-D environment, with an ongoing expansion to 3-D diffusion models (currently 3-D simulations are possible only under specific

conditions [244]). In these works, the reception process was simulated by counting the particles located within a given area close to a receiver node during each time step. N3Sim also considered the laws for handling electrostatic forces, inertial forces, and particle collisions.

A simulation tool, BINS, for nanoscale biological networks was developed for simulating information exchange at nanoscale [245]. BINS is adaptable to any type of information particle. In [245], the authors provided a comprehensive description of their simulation libraries, and demonstrated the capabilities of their simulator by modeling a section of a lymph node and the information transfer within it. Their simulator supports 3-D space, receptor dynamics including affinity, different carriers, lifetime for molecules, tracking each molecule, collisions, and mobile receiver nodes. BINS was upgraded to BINS2, which offers new features to simulate bounded environments. Using the BINS2 simulator, propagation in blood vessels was analyzed in [246] and *in vitro* experiments were simulated in [247]. The authors simulated an experiment that focused on the communication between platelets and endothelium through the diffusion of nanoparticles [247]. They verified their simulation results with experimental data.

In [248], the MolecUlar CommunicatIoN (MUCIN) simulator for diffusion-based molecular communication systems was presented. The MUCIN simulator is an end-to-end simulator that considers first hitting process for the signal reception. It

supports 1-D to 3-D environments, sending consecutive symbols, imperfect molecule reception, extendable modulation, and filtering modules. The MUCIN simulator was developed in MATLAB and it is available in MATLAB file exchange central.

In [249], an open source bacteria network simulator BNSim was introduced. BNSim supported parallel multi-scale stochastic modeling, chemical signaling pathways inside each bacterium, movement in 3-D environments, and chemotaxis. Similarly, in [91] the authors proposed medium-range point-to-point molecular communication that was based on the transport of DNA-encoded information between transmitters and receivers using bacterium. The authors presented the channel characterization and a simulator for the communication channel. The proposed communication scheme was as follows: first, a DNA message was embedded inside the bacterium cytoplasm. Then, the bacterium was released into the environment that in turn followed its natural instincts and propelled itself to a particular receiver, which was continuously releasing attractant particles. Finally, in the last step, the reception and decoding of the DNA message was preformed through exchanging genetic material.

3.6 Overview of Research Chapters

In this chapter as well as the previous one, a comprehensive survey of the field of molecular communication was provided. Although information particles can travel

from the transmitter to the receiver using many different propagation schemes, most previous works have focused only on diffusion-based propagation (see Table 3.2). This is partly because in practice diffusion is simple to implement: the transmitter simply releases the information particles and they diffuse away until they reach the receiver. Moreover, there exists closed-form expressions for certain diffusion-based propagation channels (e.g. under infinite boundary conditions). However, diffusion can be slow as the separation distance between the transmitter and the receiver increases. Moreover, practical channels may be confined, and therefore the infinite boundary conditions used in previous works may not hold.

To fill in this gap, I propose to simulate and model on-chip applications using active transport with molecular motors, diffusion, and flow-assisted diffusion. In particular, I consider a special form of active transport where stationary kinesin and mobile microtubule filaments are used for information particle transport. This propagation scheme has not been considered previously by other researchers, and in [67, 80] it is shown that mass transport is possible using this scheme. Moreover, recently it is shown that electrical currents can be used to control the speed and direction of the microtubules [82, 83, 250], and in [89] it was shown that this approach could be used for molecular transport and assembly in microfluidics. Therefore, kinesin-microtubule motility can be a suitable choice for on-chip molecular communication.

My research goals in this direction are as follows.

- Simulate an on-chip molecular communication environment (Chapter 4): since laboratory experimentations can be very laborious and time consuming, computer simulations must be employed. Therefore, creating a simulation environment based on kinesin-microtubule transport, diffusion transport, and flow-assisted diffusion transport can be very beneficial. Most previous works on diffusion-based propagation have assumed infinite boundary conditions, which may not hold for small and confined on-chip environments. Therefore, the benefits of this simulation environment are twofold.

- Compare the achievable information rates of different propagation schemes (Chapter 4): By comparing the performance of different propagation schemes in on-chip channels, I can provide a guidepost as to when each transport mechanism would be suitable. I show that kinesin-microtubule motility can be very effective for on-chip molecular communication applications.

- Find mathematical models for kinesin-microtubule based molecular communication (Chapter 5): deriving mathematical models for active transport propagation can help with design and development of these systems.

- Find optimal design choices for kinesin-microtubule based molecular communication (Chapter 6): different channel parameters such as shape of the

channel, or the shape of transmission area can have an effect on the achievable information rates in on-chip molecular communication. I provide optimal design choices for kinein-microtubule based molecular communication.

There have been many advancements in theoretical molecular communication over the past several years. All these advancements, however, have not translated into practical and experimental systems. Perhaps, one of the culprits is the multidisciplinary nature of molecular communication that requires a strong collaboration between telecommunication engineers, mechanical engineers, chemical engineers and biological engineers. Another possible hurdle is the cost associated with running wet labs. These labs are typically very expensive to run, which makes progress slower. Despite all these issues, there have been a few elementary experiments that have explored the feasibility of molecular communication.

One of the first works on engineering an experimental demonstration of molecular communication at microscales is [144]. With the prevalence of synthetic biology and applications such as biological computation, simple intercellular communication has been used for multicellular computation [251]. In [252], artificial cells are used as relays (translates) to send a message to *Escherichia coli* (*E. coli*) cells. In this scheme, the intended message is first detected and decoded by the artificial cells, and then relayed to the *E. coli* cells. A short survey on cell-to-cell communication in synthetic biology is presented in [253].

Most work in synthetic cell-to-cell communication consider sending a single message from a transmitter to a receiver. However, recently in [42] it was demonstrated that a message sequence encoded in DNA can be transmitted between cells by packaging the DNA massage inside a bacteriophage. I believe that through more collaborative work between communication theorist and biomedical engineers working on synthetic biology, more advanced systems could be developed in the future.

Replicating many of the experiments explained above requires access to wet labs, which many researchers in the field of molecular communication do not have access to. Moreover, the experimental systems discussed above cannot be used to transport digital sequential data. Therefore, in Chapter 7 of the dissertation, I develop a simple and inexpensive platform that demonstrates the practicality of molecular communication and can be used for experimentation. This system, which can be used in a simple lab environment, is capable of transferring sequential digital data. This platform is different from the on-chip molecular communication channels considered in Chapters 4-6, and is not intended to be an experimental demonstrator for those particular channels. In Chapter 8, I use this platform to show that the models used in many previous works cannot be applied to my platform. Therefore, I propose corrections to these models, and also explore the inherent nonlinearity of my experimental setup. I present the conclusion and future research directions in Chapter 9 of the dissertation.

4 On-Chip Molecular Communication

In this chapter, on-chip molecular communication systems are considered. As was discussed in Chapter 2, on-chip molecular communication have many potential applications in lab-on-chip devices, point-of-care diagnostic devices, and microfluidic devices. These devices are typically made from different components that may measure different chemical properties. In some cases, converting these chemical properties into electrical signals can be very challenging. In this case, these components on the device can be networked together using molecular communication. Since studying these systems in a laboratory setting can be very expensive and time consuming, a computer simulation environment for on-chip molecular communication is developed for different transport mechanisms. Different propagation schemes are then compared to determine the most suitable form of transport, that can achieve the highest information transmission rate.

4.1 Overview of On-Chip Molecular Communication

In on-chip molecular communication, a transmitter generates information carrying particles such as molecules or lipid vesicles, and then releases them for transmission to the receiver, in an aqueous environment. Therefore, the transmitter consists of a particle generation module where the information particles are generated, and a release mechanism which oversees the timing, the number, and the type of the particles released. Together the particle generation and release modules can encode information into the information carrying particles. For example, the information can be encoded in the release time, type, number, or the concentration of the particles released. Using this scheme, the design of the particle generation and release mechanisms can be treated as separate problems. Similarly, at the receiver the information particles are captured by a reception module and then decoded using a decoder module.

In this chapter, the particle transport mechanism is abstracted and is considered to be completely separated from the transmitter's release module and the receiver's reception module, as shown in Figure 4.1. There are three major propagation schemes: passive transport, active transport using molecular motors, and flow-assisted transport using an external device such as a syringe pump. When passive transport is employed, the particles diffuse in the fluidic environment and

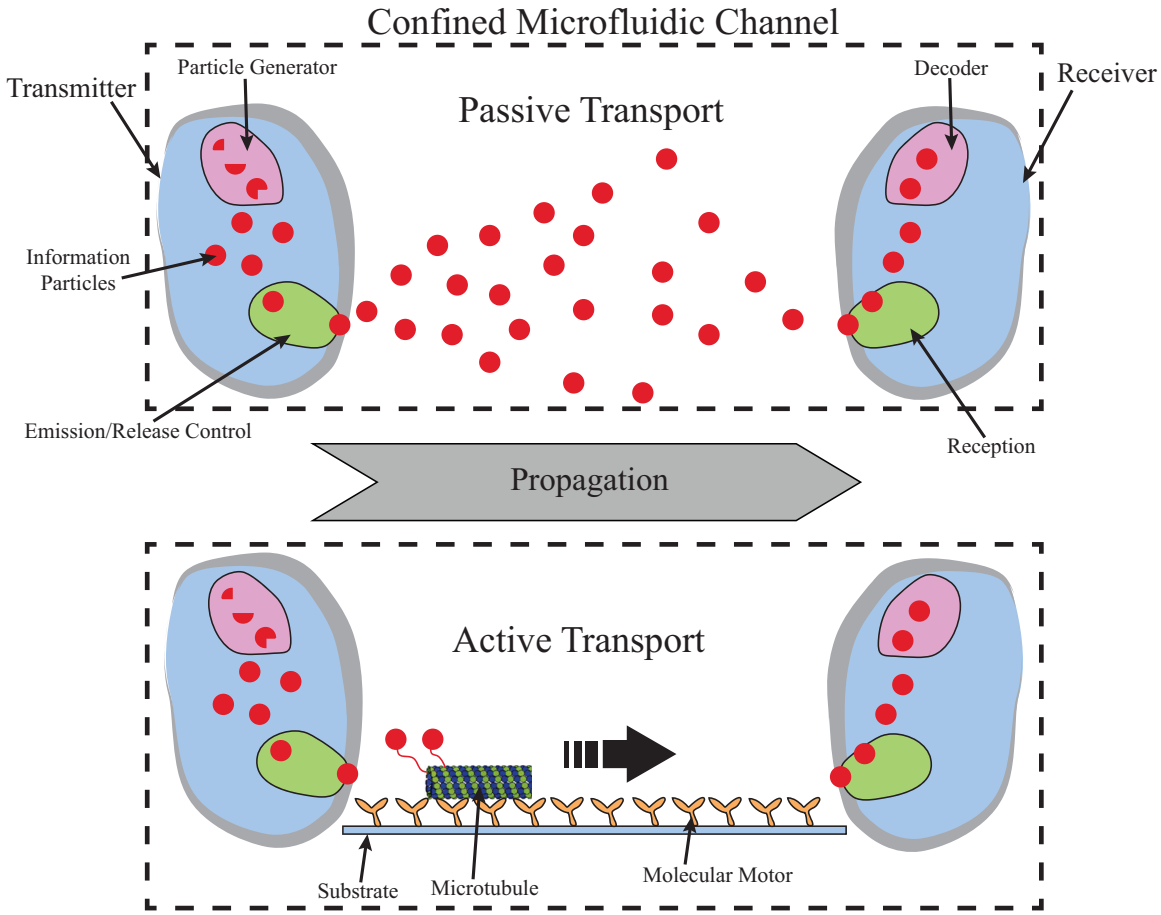


Figure 4.1: Two molecular communication systems depicting the transmitters, receivers, the confined microfluidic channel (dashed lines), and different propagation schemes. *TOP*: Passive transport is employed, where the information carrying molecules diffuse in the confined microfluidic environment and follow a Brownian motion path from the transmitter to the receiver. *BOTTOM*: Active transport using stationary molecular motors attached to a glass substrate and microtubule filaments are employed to carry the information particles from the transmitter to the receiver.

follow a random Brownian motion until they arrive at the receiver. When active transport using molecular motors is employed, a molecular motor system consisting of kinesin and microtubule filaments can be used to transport the particles from the transmitter to the receiver [80]. In flow-assisted transport, a syringe pump is used to create a flow that would assist the simple Brownian motion of the information particles. In the rest of this chapter flow-assisted transport is sometimes referred to as *Brownian motion with flow*.

In active transport using a molecular motor, a loading and an unloading mechanism for picking up the information particles from the transmitter and dropping them off at the receiver is necessary. In [67, 80], single stranded deoxyribonucleic acids (ssDNA) and the corresponding hybridization bonds between complementary ssDNA pairs is proposed for the loading and the unloading mechanisms. Microtubules moving over a kinesin covered glass substrate are covered with 15 base ssDNAs, and the information particles are also covered with 23 base ssDNAs which are complementary to that of the microtubules' ssDNAs. When the microtubule glides close to an information particle, the two ssDNA sequences hybridize and the microtubule carries the information particle until it gets close to the receiver. The receptor module at the receiver is covered with 23 base ssDNAs, which are complementary to that of the information particles. When a loaded microtubule filament glides close to the receptor module, it will unload the information particle through

hybridization bound with the complementary 23 base ssDNA at the receptor module. Since 23 base hybridization bound is stronger than the 15 base hybridization bound between the particles and the microtubules, the particles are unloaded at the receiver. Figure 2.6 in Chapter 2 summarizes this process.

Although these propagation schemes are stochastic, it is not clear which performs better and can achieve a higher information transmission rates. Brownian motion is more "random" than the microtubule's motion over a kinesin covered glass substrate. However in Brownian motion, information particles start propagating as soon as they are released into the fluidic channel by the release control module at the transmitter, while in molecular motor based active transport they remain at the release control module until they are picked up by the microtubules. Furthermore, it is unclear how the shape of the transmitter, receiver, and the confined microfluidic channel effects the information rate. Moreover, if the shape of these components effects the information rate, it is not clear if there exists an optimal shape that would maximize the information rate.

In this chapter, the channel capacity (maximum achievable information rate) of different propagation schemes are considered. Therefore, the transmitter and the receiver design are ignored. Moreover, isolated pulse transmissions are considered and the effects of inter-symbol interference (ISI) is ignored. As was demonstrated in Chapter 3 this could be a valid assumption since there are ISI mitigation techniques

such as use of enzymes to destroy the information particles that contribute to ISI. In the rest of this chapter, it is assumed that the transmitter and the receiver are perfect and can encode, release, receive, and decode the information carrying particles flawlessly.

4.2 Information Theory and Achievable Information Rate

Previous work has considered molecular communication either as a *timing channel* problem (i.e., where information is encoded in the times when molecules are released) [20, 21]; as an *inscribed matter* problem (i.e., where information is encoded by transmitting custom-made particles, such as specific strands of DNA) [18]; or as a *mass transfer* problem (i.e., a message is transmitted by moving a number of particles from the transmitter to the receiver) [17].

In this chapter, information transmission is considered as a mass transfer problem. In the simplest possible conception of this scheme, the particles themselves are not information-bearing, and a message is conveyed in the *number* of particles released by the transmitter. For example, if a maximum of three particles may be used, two bits long messages could be formed (i.e., $\log_2 4$): "00" for 0 particle, "01" for 1 particle, "10" for 2 particles, and "11" for 3 particles. However, this message might not be perfectly conveyed to the receiver: given a time limit τ for the communication session, it is possible that some of the particles will not arrive

at the receiver after τ has elapsed. This time limit τ is called *time per channel use*. In other words, time per channel use is a predefined amount of time representing the time duration for a single message transmission session, and it is one of the parameters of the molecular communication system.

Let $\mathcal{X} = \{0, 1, 2, \cdots, x_{max}\}$ be the set of possible transmission symbols, x_{max} be the maximum number of particles the transmitter can release per channel use, and $X \in \mathcal{X}$ be the number of information particles released into the medium by the transmitter. In a traditional communication system, the received symbols at the receiver are corrupted with noise from the environment, while in the molecular communication system, the received symbols are corrupted because of the random propagation of particles. Let $Y_\tau \in \mathcal{X}$ represent the number that arrive at the destination after time per channel use τ. From the channel's perspective, $X \in \mathcal{X}$ is a discrete random variable given by probability mass function (PMF) $P(x)$, and $Y_\tau \in \mathcal{X}$ is also a discrete random variable given by PMF $P(y_\tau)$,

The maximum rate at which any communication system can *reliably transmit information* over a noisy channel is bounded by a limit called *channel capacity* [192]. The channel capacity can be calculated as

$$C = \max_{P(x)} I(X; Y_\tau), \tag{4.1}$$

where $I(X;Y_\tau)$ is the mutual information between X and Y_τ. Mutual information is defined as

$$I(X;Y_\tau) = E\left[\log_2 \frac{P(y_\tau|x)}{\sum_{x\in\mathcal{X}} P(y_\tau|x)P(x)}\right], \tag{4.2}$$

where $P(y_\tau|x)$ represents the probability of receiving symbol y_τ at the destination, given that symbol x was transmitted by the source; $P(x)$ represents the probability of transmitting symbol x; and $E[\cdot]$ represents expectation.

Clearly, there exists some PMF $P(y_\tau|x)$ of the number of arrived particles given the number of transmitted particles. If this PMF is known, it is possible to calculate mutual information for any $P(x)$. However, in order to calculate the channel capacity the PMF $P(x)$ that maximizes mutual information must be found. The Blahut-Arimoto algorithm [193, 194] can be used to find the PMF $P(x)$ such that, given $P(y_\tau|x)$, mutual information is maximized. This is because mutual information is convex with respect to $P(x)$, and Blahut-Arimoto algorithm is a convex optimization algorithm. Therefore, if PMF $P(y_\tau|x)$ is known, the channel capacity of the molecular communication system can be calculated in a straightforward manner.

Finding the PMF $P(y_\tau|x)$ is non-trivial, however, because of the shape and the geometry of the molecular communication channels, which generally rules out closed-form solutions. For example, mathematically it is well known that Brownian

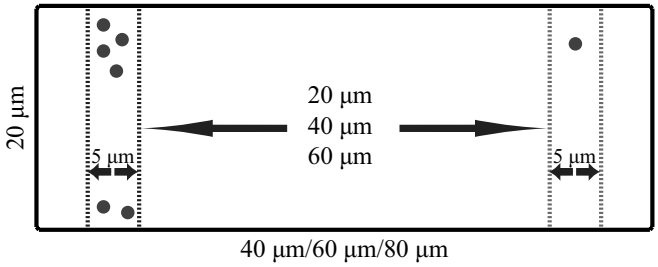


Figure 4.2: Depiction of the simulation environment. In this figure, dots represent information particles. The transmission zone is on the left (dashed strip on the left), and the receiver is on the right (dashed strip on the right). The width of the channel is constant at 20μm. The height of the channel, not shown here, is also constant at 10μm. The distance between the transmission area and the receiver area and hence the length of the channel is variable.

motion can be described by stochastic differential equations, and properties of the motion (such as first arrival time distributions) can be derived from solutions of these equations. However, in confined spaces, solutions are generally not available in closed-form. To overcome this issue, the PMF can be estimated using Monte Carlo simulations. In order to do this, first a computer simulation environment was created where the actual random motion of the particles, for both Brownian motion and molecular motor based active transport, is modeled. Then, the propagation from a hypothetical transmitter to a hypothetical receiver is simulated at least 100,000 times to obtain an estimate for the PMF $P(y_\tau|x)$. Using this estimate, the channel capacity and therefore, the maximum achievable information rate is then calculated. In the next section, this simulation environment is described in detail.

4.3 Simulation Environment and Modelling Propagation

My molecular communication setup shown in Figure 4.2 is a rectangular propagation environment, representing a channel on a microfluidic chip, with a fixed width and height of 20μm and 10μm, respectively, consisting of a *transmission zone* on the left and a *receiver zone* on the right. Please note that Figure 4.2 is the top view of the simulation environment and does not show the height of the channel. Regardless of the propagation model, message-bearing particles originate at the transmission zone, and propagate until they arrive at the receiver zone. The separation between the transmission zone and the receiver zone and hence the length of the environment is variable with different separation values of 20μm, 40μm and 60μm. In the rest of this section, this particular setup is considered and is used as the basis for the simulation and comparison of different propagation schemes.

4.3.1 Simulating Brownian Motion

Brownian motion refers to the random motion of a particle as it collides with other molecules in its vicinity [73]. Through this random motion the information carrying particles propagate from the transmission zone to the receiver zone. Following the method described in Section 2.1.2.1, Monte Carlo simulations are performed on particles in order to obtain the needed properties of the motion. In particular,

a three-dimensional discrete-time simulation of information carrying particles are performed, using Equations (2.5)-(2.10), which are shown again in Algorithm 4.1 below. It is assumed that the size of the information particles are larger than the size of the molecules of water. Therefore, the diffusion coefficient D is given by

$$D = \frac{k_B T}{6\pi\eta R_H}, \tag{4.3}$$

where $k_B = 1.38 \cdot 10^{-23}$ J/K is the Boltzman constant, T is the temperature (in K), η is the dynamic viscosity of the fluid, and R_H is the hydraulic radius of the molecule [73]. It is assumed that D is the same throughout the medium, and that collisions with the boundaries are elastic. Moreover, it is assumed that particles do not get stuck to the walls.

The function presented in Algorithm 4.1 simulates the Brownian motion of the information particle during a single time step Δt. In the algorithm, the function RAND() return a random number uniformly distributed between 0 and 1. The function REFELECTBACK(), returns a point inside the channel according to the reflected path of the information particle during a collision with the channel boundaries.

In the simulations, it is assumed that the particles initially start propagating randomly and uniformly on the (x, y) plane of the transmission zone, but z is

Algorithm 4.1 Simulation of the Brownian motion of the information particle during one time step Δt.

function SIMONESTEPBM($(x_{i-1}, y_{i-1}, z_{i-1})$, D, Δt, $ChanBoundaries$)
  $\phi \leftarrow \arccos(1 - 2\ \text{RAND}())$
  $\theta \leftarrow 2\pi\ \text{RAND}()$
  $\Delta r \leftarrow \sqrt{6D\Delta t}$
  $x_i \leftarrow x_{i-1} + \Delta r \cos\theta \sin\phi$
  $y_i \leftarrow y_{i-1} + \Delta r \sin\theta \sin\phi$
  $z_i \leftarrow z_{i-1} + \Delta r \cos\phi$
  if the point (x_i, y_i, z_i) is outside $ChanBoundaries$ **then**
    $(x_i, y_i, z_i) \leftarrow$ REFELECTBACK($(x_{i-1}, y_{i-1}, z_{i-1})$, (x_i, y_i, z_i), $ChanBoundaries$)
  end if
  return (x_i, y_i, z_i)
end function

set to the maximum vertical height (i.e., the particles are "dropped" onto the microchannel at the transmission zone). Furthermore, the particles start moving as soon as they are released into the microchannel by the transmitter. The propagation halts as soon as the particle arrives at the receiver zone. The receiver zone is on the bottom right side and it is a box with a small height of 23nm representing the height of the 23 base ssDNAs that are used to capture the information particles.

Using these equations, to calculate the PMF $P(y_\tau|x)$, the Brownian motion of a single particle is simulated from the transmitter to the receiver, as shown in Figure 4.3. By repeating these simulation trials, the probability that a single particle arrives at the destination in τ seconds can be calculated. Letting p_a represent this

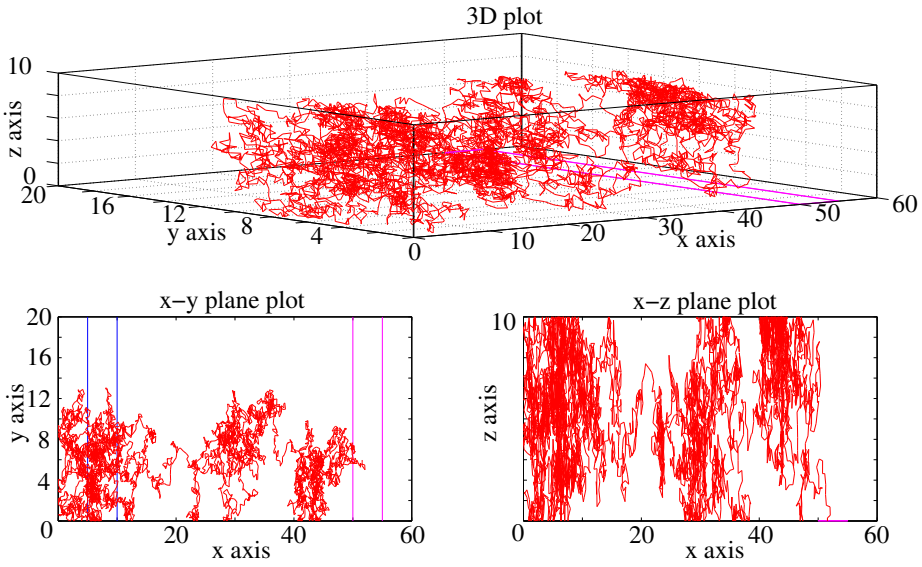


Figure 4.3: A sample Brownian movement of a single particle in three dimensions, from the transmitter area to the receiver area.

probability, the PMF $P(y_\tau|x)$ has the binomial distribution, given by

$$P(y_\tau|x) = \begin{cases} \binom{x}{y_\tau} p_a^{y_\tau}(1-p_a)^{x-y_\tau}, & 0 \le y_\tau \le x \\ 0, & \text{otherwise} \end{cases}. \tag{4.4}$$

Thus, p_a is found by simulating many trials of a single particle's motion over τ seconds, and counting the fraction that arrive, as shown in Algorithm 4.2. In the algorithm, the function RandomStartBM() return a random point uniformly distributed inside the transmission area, and the function SimOneStepBM() is given in Algorithm 4.1. Note that different simulations are necessary for different

Algorithm 4.2 Simulation algorithm to calculate p_a for Brownian motion.

```
1: function CalcPa(D, τ, Δt, NumbTrials, ChanBoundaries)
2:     NumbArrivals ← 0
3:     for k = 1, 2, ..., NumbTrials do
4:         (x_i, y_i, z_i) ←RandomStartBM(ChanBoundaries)
5:         for j = 1, 2, ..., τ/Δt do
6:             (x_i, y_i, z_i) ←SimOneStepBM((x_i, y_i, z_i), D, Δt, ChanBoundaries)
7:             if (x_i, y_i, z_i) inside receiver zone then
8:                 NumbArrivals ← NumbArrivals + 1
9:                 break
10:            end if
11:        end for
12:    end for
13:    p_a ← NumbArrivals/NumbTrials
14:    return p_a
15: end function
```

values of τ. Furthermore, because the motion of each particle is independent of the other particles released by the transmitter, $P(y_\tau|x)$ has a binomial distribution.

4.3.2 Simulating Molecular-Motor-Based Active Transport

When active transport using molecular motors is employed, it is assumed that the microchannel is lined with static kinesin motors, which cause microtubule filaments to propagate along their surface, carrying information-bearing particles from the transmitter to the receiver. It is also assumed that the particles are anchored to the transmission area until they are picked up by the microtubules for delivery to the destination. Therefore, instead of simulating the motion of particles, the motion of the microtubules moving over kinesin covered substrate must be simulated.

Algorithm 4.3 Simulation of the microtubule motion during one time step Δt.

1: **function** SIMONESTEPMT((x_{i-1}, y_{i-1}), θ_{i-1}, v_{avg}, L_p, D, Δt, $ChanBoundaries$)
2: $\theta_i \leftarrow \theta_{i-1}+$ RANDN$(0, \frac{v_{\text{avg}}\Delta t}{L_p})$
3: $\Delta r \leftarrow$ RANDN$(v_{\text{avg}}\Delta t,\ 2D\Delta t)$
4: $x_i \leftarrow x_{i-1} + \Delta r \cos\theta_i$
5: $y_i \leftarrow y_{i-1} + \Delta r \sin\theta_i$
6: **if** the point (x_i, y_i) is outside $ChanBoundaries$ **then**
7: $(x_i, y_i), \theta_i \leftarrow$ FOLLOWWALL((x_{i-1}, y_{i-1}), (x_i, y_i), $ChanBoundaries$)
8: **end if**
9: **return** (x_i, y_i), θ_i
10: **end function**

The motion of the microtubule is largely regular, although the effects of Brownian motion cause random fluctuations. Again, Monte Carlo simulations are used to obtain the needed properties of the motion, using the Equations (2.17)-(2.23), where the equations are shown again in Algorithm 4.3 below. Please note that the microtubules move only in the x–y directions, and do not move in the z direction (along the height of the channel).

In Algorithm 4.3, the motion of a single microtubule filament over a kinesin covered substrate during a single time step Δt is simulated. The function RANDN() in the algorithm returns a normally distributed random variable according to the input mean and variance. The first input parameter to the function is the mean, while the second is the variance. In the simulation, it is assumed that the microtubules follow a path along the channel wall when they collide with it. This property was

shown to hold experimentally in [84, 85]. Therefore, the function FollowWall() returns a point inside and along the channel wall when the microtubule collides with the wall. It also returns an angle for the final path's direction of motion.

For Monte Carlo simulations, the starting location of the microtubule is assumed to be random and uniformly distributed across the entire propagation area. Moreover, the initial directional angle θ_0 is selected uniformly at random from the range $[0, 2\pi]$, and microtubules are assumed to be initially unloaded (without any cargoes). It is also assumed, as demonstrated in [80], DNA hybridization bond is used to anchor the information particles to the transmission area. Similarly, ssDNAs on the surface of the microtubules hybridize with the ssDNAs on the surface of the information particles when a microtubule passes in close proximity, thereby loading the particles onto the microtubule. From experimental observations it is evident that microtubules can load multiple information particles [80]. Therefore, the loading of a particle happens only if a microtubule with available cargo spots passes in close proximity of an information particle.

In order to capture this loading effect in the simulations, the transmission zone is divided into a square grid, where the length of each square in the grid is the same as the diameter of the information particle. Then, information particles are distributed randomly and uniformly among the squares in the grid. If a microtubule enters a square which is occupied by a particle, and it has an empty cargo slot available,

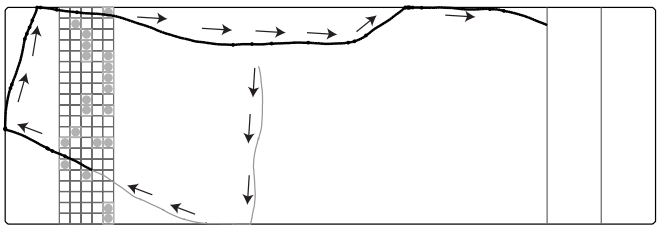


Figure 4.4: A sample trajectory of active transport. The microtubule initially starts to the left of the receiver zone (strip on the right side of the microchannel), and moves down and then left (Lighter thinner line). It passes through the transmission zone (grid with empty and particle-bearing tiles) loads a particle at which point the line turns darker and thicker indicating a loaded mictorubule. Then the microtubule passes through again, loading another particle (maximum load is assumed to be 5). The loaded microtubule then travels toward the receiver zone, where it delivers the two particles and the trajectory terminates.

the microtubule loads that particle. In general, it is observed that the microtubules can load multiple particles, and the maximum number of particles a microtubule can load is given by half of its length divided by the diameter of the particles. This assumption is based on experimental observations in [67]. For unloading at the receiver, it is assumed that all the loaded particles are dropped off as soon as a loaded microtubule enters the receiver zone. This assumption is made to assist the speed of simulations. Figure 4.4 shows a sample trajectory with the loading and unloading mechanism.

To calculate the PMF $P(y_\tau|x)$, for a given number of particles released by the transmitter (x), the motion of a microtubule is simulated for τ seconds and the number of particles delivered to the receiver (y_τ) is measured at the end of τ seconds.

Algorithm 4.4 Monte Carlo simulation algorithm for calculating $P(y_\tau \mid x)$.

```
1: function CalcPyGx(X, τ, v_avg, L_p, D, Δt, MaxLoad, ChanBoundaries,NumbTrials)
2:     Y[0 : X] ← 0
3:     for k = 1, 2, ..., NumbTrials do
4:         NumbDelivered, NumbLoaded ← 0
5:         (x_i, y_i) ←RandomStartMT(ChanBoundaries)
6:         θ_i ← 2π Rand()
7:         Grid ← RandPlaceParticles(X, ChanBoundaries)
8:         for j = 1, 2, ..., τ/Δt do
9:             (x_i, y_i), θ_i ←SimOneStepMT((x_i, y_i), θ_i, v_avg, L_p, D, Δt, ChanBoundaries)
10:            if (x_i, y_i) is in Grid with a particle and NumbLoaded < MaxLoad then
11:                Grid ← RemoveParticle(Grid, (x_i, y_i))
12:                NumbLoaded ← NumbLoaded + 1
13:            end if
14:            if (x_i, y_i) is in receiver zone then
15:                NumbDelivered ← NumbDelivered + NumbLoaded
16:                NumbLoaded ← 0
17:            end if
18:        end for
19:        Y[NumbDelivered] ← Y[NumbDelivered] + 1
20:    end for
21:    PyGx ← Y[0 : X]/NumbTrials
22:    return PyGx
23: end function
```

By repeating this process the PMF $P(y_\tau|x)$ can be estimated for the time duration τ as shown in Algorithm 4.4. In the algorithm, the function RandomStartMT() set the starting position of the microtubule randomly inside the channel environment, and the function RandPlaceParticles() places the information particles randomly inside the grid transmission zone. The whole simulation process can be repeated for different values of x. More generally, for each value of x and τ a set of simulations is necessary.

4.4 Improving Information Rate

In the previous section, a simple molecular communication system employing either active transport using molecular motors or passive transport using Brownian motion was described. Also, simulation schemes for calculating the channel capacity based on each propagation technique was presented. In this section, number of methods for further improving the channel capacity is proposed.

4.4.1 Brownian Motion with Flow

In general, one way to improve the channel capacity is to transport information particles from the transmitter to the receiver more quickly. This is because information is carried using these particles and naturally if they arrive faster, the information will be transported over a shorter time duration. In Brownian motion, one way to achieve this is by introducing flow in the direction from the transmitter to the receiver. However, Brownian motion with flow requires an external device such as a syringe pump to produce the flow. In its simplest form, introducing flow will change Equations (2.6)-(2.8) into Equations (2.14)-(2.16) presented in Chapter 2. In this chapter, the flow velocities are assumed to be constant throughout the molecular communication channel, and in the positive x direction. This assumption is realistic for channels with laminar flow, and liquid flow in microfluidics is typi-

cally laminar. To implement this in the algorithm, line 5 of Algorithm 4.1 changes to

$$x_i \leftarrow x_{i-1} + v\Delta t + \Delta r \cos\theta \sin\phi, \tag{4.5}$$

where v is the constant velocity along the positive x direction. Note that the velocity components along all the other directions are zero and therefore lines 6-7 of the algorithm remain unchanged.

4.4.2 Improving Molecular Motor Based Active Transport

Again, channel capacity can be improved if information particles are transfered from the transmitter to the receiver quicker on average. Two possible schemes are proposed: increasing the number of microtubules and optimizing the transmission zone.

In the first scheme, by increasing the number of microtubules, more information particles could be picked up and transported to the receiver area during the time interval τ. From experimental results, it is evident that microtubules generally move independently of each other, and therefore the same assumption is made in the simulations of multiple microtubules.

In the second scheme, the fact that the microtubules mostly move along the

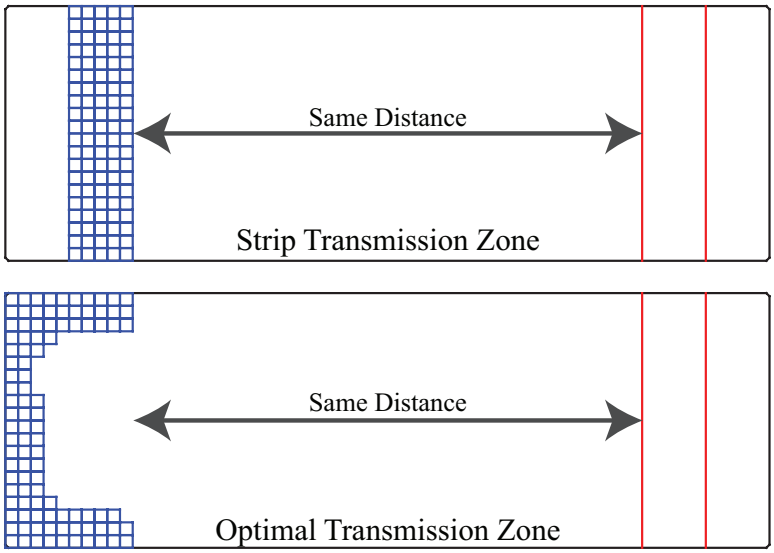


Figure 4.5: Depiction of the strip transmission zone and the optimal transmission zone. The optimal transmission zone is along the walls of the channel. The shortest distance between the transmission zone and the receiver zone are the same in both cases. The number of squares in both grid structures are 100.

walls of the molecular communication channel is used, to optimize the transmission area. This fact has been observed experimentally, and verified using computer simulations [80, 85]. An information particle is picked up from the transmission zone, and delivered to the receiver zone, if the corresponding square is visited. Therefore, it is desirable to find squares with the highest probability of being visited during one trip. If information particles are placed there, they have the highest probability of being loaded and therefore of being delivered. Therefore, placing the transmission zone along the walls of the channel would increase the chances of the particles being picked up. This is called the optimal transmission zone.

Figure 4.5 depicts the strip transmission zone and the optimal transmission zone. In Chapter 6, it will be shown how this optimal transmission zone can be found using mathematical models presented in Chapter 5 for information particles of different size.

Unlike the improvement technique proposed for Brownian motion (i.e. flow), these improvements are internal to the molecular communication system and do not require any external device.

4.5 Comparison of Simulation Results

In this section, the simulation results for the molecular communication system proposed in Section 4.3 is presented, with variable separation distance between the transmitter and the receiver, employing different propagation schemes.

When Brownian motion is used for propagation, the following simulation parameters are used: simulation time steps of $\Delta T = 0.1$ seconds, and the free diffusion coefficient $D = 1\ \mu\text{m}^2/\text{s}$. Therefore, from Equation 2.9, at each simulated time step the information particles moves $\Delta r = \sqrt{0.4}\ \mu\text{m}$ in a random direction. The same values are also used when Brownian motion with flow is considered. The flow is assumed to be always in the direction from the transmitter to the receiver (i.e. the flow velocities in the y and z direction are zero and positive in the x direction.)

For simulating the motion of the microtubule, when molecular motor based

active transport is employed, the following parameters are used: simulation time steps of $\Delta T = 0.1$ seconds, microtubule diffusion coefficient $D = 2.0 \cdot 10^{-3}$ μm^2/s, average speed of the microtubule $v_{\text{avg}} = 0.5$ μm/s, and persistence length of the microtubules trajectory $L_p = 111$ μm. It is also assumed that the size of the information particles (which is typically a liposome covered with ssDNA stand for microtubule loading and unloating)is 1μm, the average length of the microtubules is 10μm, and each microtubule can load up to 5 information particles in one trip from the transmission zone to the receiver zone. These parameters are all selected based on experimental observations of ssDNA covered microtubules moving over a kinesin covered substrate [80, 84].

It is assumed that the set of possible transmission symbols are $\mathcal{X} = \{0, 1, 2, \cdots, x_{max}\}$, for some value of x_{max}, where a transmission symbol $X \in \mathcal{X}$ is represented by release of X information particles into the medium. In the case of active transport using molecular motors, all the released particles will be randomly distributed and will remain stationary at the transmission zone until they are picked up for delivery by a microtubule. By simulating the motion of the particles or the motion of the microtubules many times (100,000 plus), the PMF $P(y_\tau|x)$ can be estimated for each propagation scheme. Then, Blahut-Arimoto algorithm [193,194] is used to find the PMF $P(x)$, that would maximize the mutual information, and hence calculate the channel capacity for each propagation scheme.

4.5.1 Brownian Motion Versus Molecular Motor Based Active Transport

First, simple diffusion, flow assisted diffusion, and molecular motor based active transport are compared over constant channel dimensions. For this set of simulations a channel length of 60μm is used, which results in a 40μm separation between the transmitter and the receiver. For Brownian motion with flow, flow values of 0.2 and 0.3 μm/s are considered, while for active transport using molecular motors, 1 and 15 microtubules as well as the strip and optimal transmission area are considered.

Figure 4.6, shows the channel capacity in bits versus the maximum possible number of transmission particles (x_{max}), for the time duration of 1000 seconds per single channel use. The value of 1000 seconds is presented since the plots are easily distinguishable for this time duration. From the graph it can be seen that as the transmission symbol set is increased, the channel capacity increases for all propagation schemes. Also, Brownian motion with flow achieves a much higher information rate than simple Brownian motion without flow, and the optimal molecular motor based active transport with 15 microtubules and optimal transmission zone has a much higher channel capacity compared to the non-optimal one (a single microtubule and strip loading area).

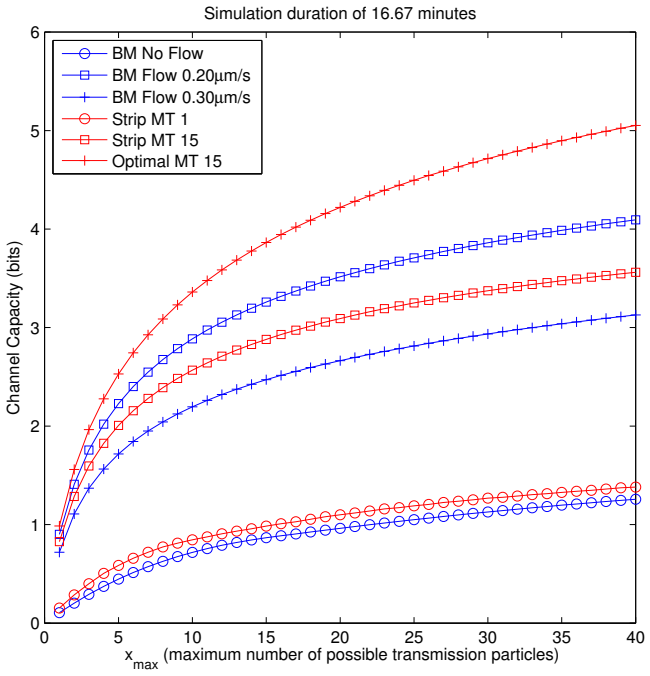


Figure 4.6: Channel capacity in bits versus the maximum number of possible transmission particles for Brownian motion (with and without flow) and molecular motor based active transport. Time per channel use is 1000 s.

From Figure 4.6 it can be concluded that simple Brownian motion has the lowest channel capacity even slightly lower than molecular motor based active transport using a single microtubule and strip transmission area. Molecular motor based active transport with the optimal transmission area and multiple microtubules achieves the highest channel capacity.

The channel capacity in bits per seconds versus the value of time per single channel use is presented in Figure 4.7. Different values of time duration per single channel use are considered, ranging from 50 seconds to 2000 seconds. The transmis-

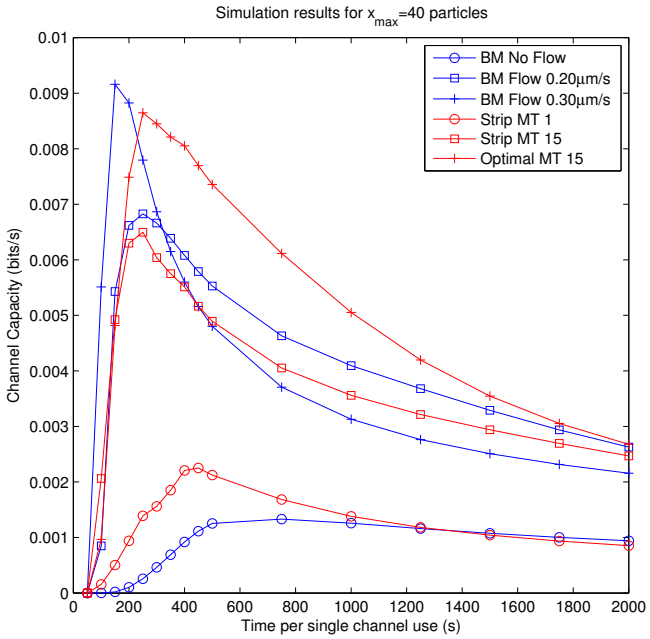


Figure 4.7: Channel capacity in bits per seconds versus time per channel use for Brownian motion (with and without flow) and molecular motor based active transport. ($x_{\max} = 40$).

sion symbol set is fixed at $\mathcal{X} = \{0, 1, 2, \cdots, 40\}$. The channel capacity in bits per seconds is then calculated by dividing the capacity in bits with the corresponding time duration per single channel use, τ. As can be seen from the graph the channel capacity initially increases as the τ increases, reaches a peak, and then start to decrease. This peak value represents the optimal value of time per single channel use τ.

Again it can be observed that simple Brownian motion has the lowest channel capacity (i.e peak value), even lower than unoptimized molecular motor based active

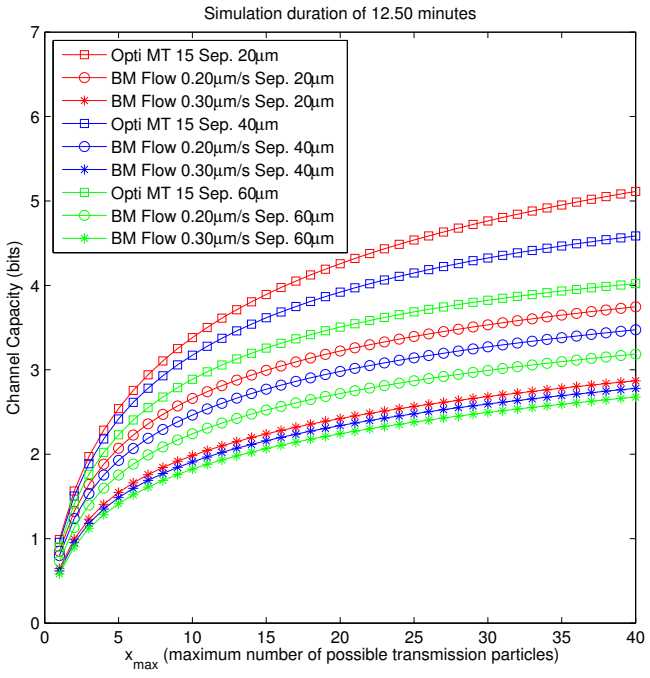


Figure 4.8: Channel capacity in bits versus the maximum number of possible transmission particles for Brownian motion with flow and molecular motor based active transport. Time per channel use is 750 seconds and different separations between the transmitter and receiver are considered.

transport using a single microtubule. However, Brownian motion with 0.3μm/s flow achieves the highest channel capacity. Similarly, in the case of the molecular motor based active transport, increasing the number of microtubules, and using the optimal transmission area, will increase the peak values and shifts the peak location to lower τ values.

4.5.2 Effects of Separation Distance

In these simulations, the separation between the transmitter and the receiver is changed to determine its effects on each propagation scheme. Since it has been shown that Brownian motion with flow (active transport using an external device) and molecular motor based active transport using multiple microtubules with optimal transmission area achieve better information rates, only these schemes are considered in the comparison.

In Figure 4.8, the channel capacity versus x_{max} is considered for separation distances of 20μm (red plots), 40μm (blue plots) and 60μm (green plots). The time per channel use is fixed at 750 seconds since the plots are more easily distinguishable at this duration. It can be seen that regardless of the separation distance molecular motor based active transport with optimal transmission zone and 15 microtubules achieves higher channel capacity. Also for a given propagation scheme, as the separation distance increases channel capacity decreases.

The channel capacity in bits per seconds versus time per single channel use is presented in Figure 4.9. Different values of time duration per single channel use are considered, ranging from 50 seconds to 2000 seconds. The transmission symbol set is fixed at $\mathcal{X} = \{0, 1, 2, \cdots, 40\}$. In general, it can be observed that regardless of the propagation scheme, as the separation distance increases the in-

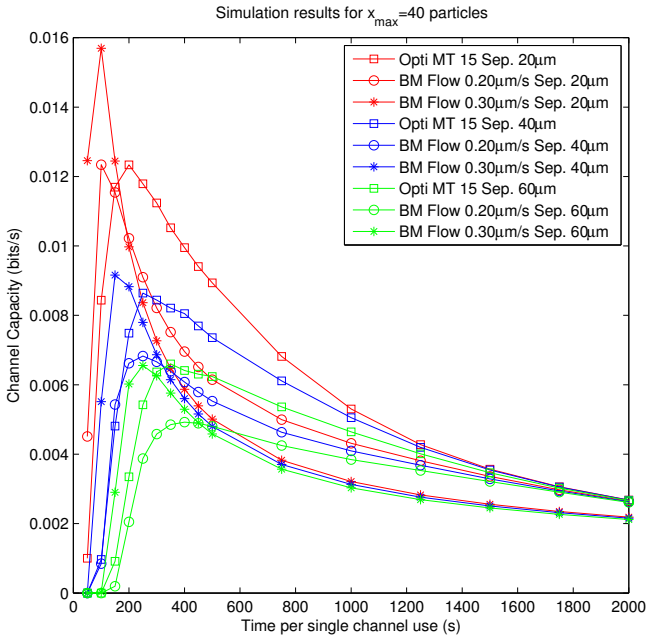


Figure 4.9: Channel capacity in bits per seconds versus time per channel use for Brownian motion with flow and molecular motor based active transport. ($x_{\max} = 40$). Different separations between the transmitter and receiver are considered.

formation rate decreases. It can also be seen that for separation distance of 20μm (red plots), Brownian motion with flow reaches a higher peak than molecular motor based active transport. However as the separation distance is increased to 60μm (green plots), the difference between the peaks lessens with active transport using molecular motors achieves a slightly higher information rate.

It is concluded that Brownian motion with flow is slightly better at smaller separation distances, while optimized molecular motor based active transport is slightly better at larger distances. However, it must be noted that introducing flow would

require an external device such as a syringe pump and therefore the system would not be as self-contained and compact as molecular motor based active transport. Therefore, molecular motor based active transport are a suitable choice for lab-on-chip applications. As shown in [89], this technique can be used in microfluidic devices as well, where there is also a flow present in the channel.

5 Mathematical Models for Active Transport Molecular Communication

In the previous chapter, a computer simulation environment was developed that could be used to study on-chip molecular communication systems. It was also shown that a novel transport mechanism [67,80], where molecular motors and microtubule filaments are used to actively transport information particles from the transmitter to the receiver, can be very effective at achieving high information rates.

As was pointed out in Chapter 3, there are no mathematical models for this type of active transport, and the simulation environment developed in Chapter 4 is not efficient enough to solve design and optimization problems. Therefore, in this chapter, mathematical models are derived that could be used to speed up the simulations significantly. Then, these models are used in the next chapter to provide optimal design guidelines for on-chip active transport molecular communication.

The models that are developed in this chapter can be divided into two categories: a simple channel model, and a Markov chain channel model. The simple channel

model, presented first, provides insights into the transport mechanism. However, this model is very simplistic, and although it improves the speed of the simulation environment developed in the previous chapter significantly, it is not very accurate. Moreover, it works only for channels that use a single microtubule filament for transporting information particles. To improve this model, more accurate Markov chain channel models for both single microtubule and multi-microtubule channels are developed in the second half of this chapter.

5.1 A Simple Mathematical Model

To derive a simple mathematical model, it is assumed that there is a single microtubule in the channel transporting information particles from the transmitter to the receiver.

Assume that the grid transmission zone described in Chapter 4 has n squares (i.e. the maximum number of particles that can be anchored to the loading zone is n). Let $X \leq n$ be the number of particles at the transmission zone in the beginning, and let $Y_\tau \leq X$ be the number of particles delivered to the receiver zone after time duration τ. Let X_i be a Bernoulli random variable representing the event where a particle is placed in the ith square for $i = 1, 2, \cdots, n$. Therefore, if the simplifying assumption is made that X_i are independent of each other, the probability that an

information particle is placed in the ith square is given by

$$P(X_i = 1|X = x) = \frac{x}{n}, \tag{5.1}$$

where particles are distributed uniformly among squares. Note that the independence assumption here is an approximation because it does not satisfy the constraint

$$x = \sum_{i=1}^{n} X_i, \tag{5.2}$$

Let V_i be a Bernoulli random variable representing the event that the ith square is visited by the microtubule in a single trip from the receiver zone, to the transmission zone, and back. Therefore, $P(V_i = 1)$ represents the probability that the ith square is visited and $P(V_i = 0)$ the probability that it is not visited. This probability distribution can quickly be calculated using simple simulations for any molecular communication channel. For example, the top part of the Figure 5.1 shows this probability distribution for squares of size 1μm covering the left side of the microchannel. From the probability distribution, it can be seen that the squares close to the walls are visited the most, which is a property of the motion of the microtubules [84].

Let K be another random variable representing the number of microtubule trips between the transmission and the receiver zone in time duration τ. The

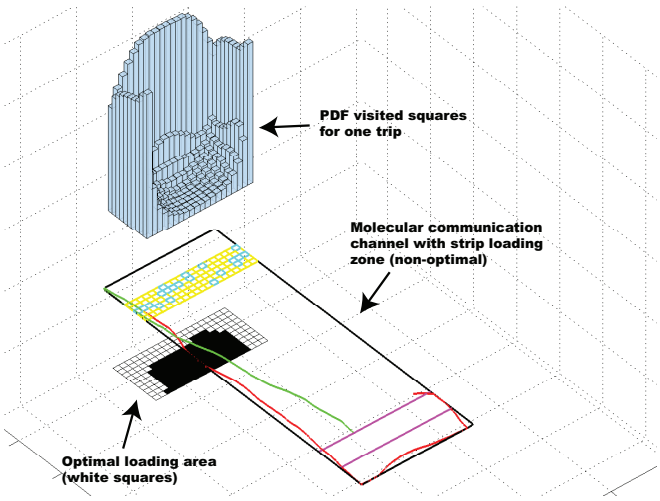


Figure 5.1: *(Top)*: Probability distribution of $P(V_i = 1)$ for squares of size 1μm to the left side of the loading area. *(Middle)*: Strip transmission area (yellow and cyan squares) for $n = 100$ squares. *(Bottom)*: Projection of the probability distribution $P(V_i = 1)$ on top. The top 100 values of $P(V_i = 1)$ are shown in as white squares and they represent the optimal loading area.

probability mass function (PMF) for K can be quickly calculated for any molecular communication channel using simple simulations. Let $V_i^{(k)}$ be a Bernoulli random variable representing the event that the ith square is visited at least once by the microtubule during k trips. Therefore,

$$P(V_i^{(k)} = 1) = 1 - (1 - P(V_i = 1))^k, \tag{5.3}$$

represents the corresponding PMF.

Let $D_i^{(k)}$ be a Bernoulli random variable representing the event that a particle

from the ith square is delivered to the destination after k trips. Then, the PMF of $D_i^{(k)}$ is given by

$$P(D_i^{(k)} = 1) = P(V_i^{(k)} = 1)P(X_i = 1), \tag{5.4}$$

using a simplifying assumption that $P(V_i^{(k)} = 1)$ and $P(X_i = 1)$ are independent. This independence assumption is not accurate since $P(X_i = 1)$ changes depending on the number of particles already delivered in previous trips. In general, this assumption becomes less accurate as the number of trips k increases or in other words the time per channel use τ increases. Let $Y^{(k)}$ be the total number of particles delivered to the receiver zone during k trips. Then, $Y^{(k)}$ is given by

$$Y^{(k)} = \min(\sum_{i=1}^{n} D_i^{(k)}, X), \tag{5.5}$$

for any given X. Since $\sum_{i=1}^{n} D_i^{(k)}$ represents a Poisson-Binomial distribution, its corresponding probability distribution can be calculated using many different techniques [254]. Finally, PMF $P(y_\tau|x)$ can be calculated as

$$P(y_\tau|x) = \sum_{k \in K} P(Y^{(k)} \mid X)P(k), \tag{5.6}$$

where $P(Y^{(k)} \mid X)$ is the PMF of $Y^{(k)}$ given in Equation (5.5), and $P(k)$ is the PMF of K, the number of trips between the transmitter and receiver during the

time duration τ.

The benefits of this model are twofold. First, it can be employed to quickly estimate the information rate of any molecular communication system, and although it relies on simulations for calculating the PMFs of V_i and K, these simulations are very simple and can be performed on an average computer quickly. Second, because of the model's simplicity, it provides insights into the transport mechanism. Therefore, the model could be used to solve various design problems. For example, in the next chapter this model is used to generate an optimal transmission zone for different sized information particles. The only drawback of this model is that the resulting PMF, $P(y_\tau|x)$, is not as accurate as the one estimated using Monte Carlo simulation.

5.1.1 Computational Comparison

In this section, the information rates obtained through simulation are compared to the information rates obtained using the simple mathematical model developed. In order to make this comparison, a rectangular propagation environment with the dimensions 20μm by 60μm is considered, as presented in Figure 4.4. The transmission area is the strip on the left, while the receiver area is on the strip right and the separation between the transmission zone and the receiver zone is 40μm. It is assumed that the diameter of the information particles is 1μm and there are

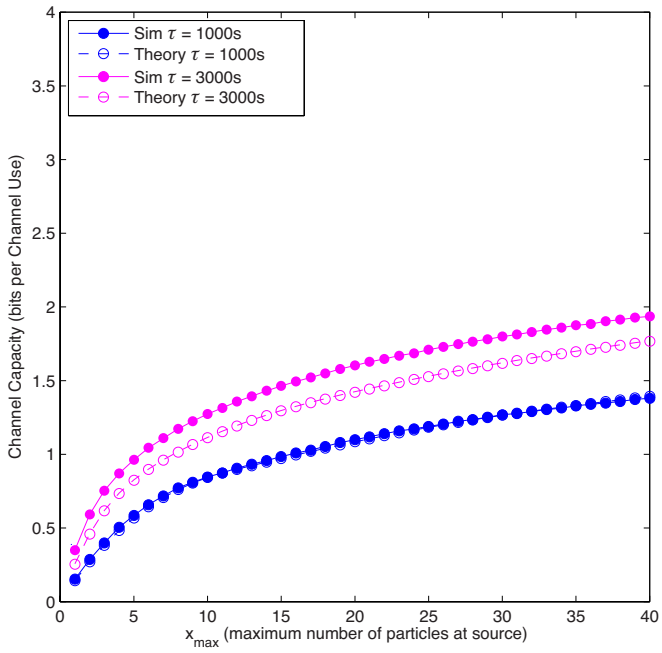


Figure 5.2: Channel capacity plot. The solid markers indicate the results obtained from full Monte Carlo simulations and the empty markers present the information rate calculated using the simple model. The blue plots present the time per channel use interval of 1000s and the magenta plots represent a time per channel use duration of 3000s.

100 squares in the grid loading area (i.e. there could be as much as 100 squares in the transmission strip). The parameters used for Monte Carlo simulations are similar to those used in Chapter 4.

Figure 5.2 shows the channel capacity (i.e. maximum achievable information rate) in bits per channel use. The channel capacity is plotted against the x_{max} which represents the maximum number of possible transmission symbols. The capacity is calculated by using both simulation (solid markers) and the simple

Table 5.1: Number of simulations required to estimate $P(y_\tau|x)$

Method of PMF Estimation	Number of Computer Simulations
Monte Carlo Simulation	$2 \times 40 = 80$
Mathematical Model	3

mathematical model (empty markers). As can be observed for small time durations, the mathematical model estimates the channel capacity closely. However, as time duration increases, the difference between the mutual information obtained through simulations and the mathematical model increases. This error is a by product of the independence assumptions made during the derivation of the mathematical model.

Next, the number of computer simulations required to estimate PMF $P(y_\tau|x)$ are compared for the case where full Monte Carlo simulations are used and the case where the proposed mathematical model is employed. It is assumed that the message set is given by $\mathcal{X} = \{0, 1, 2, \cdots, 40\}$, and the results are calculated for two time per channel use durations of 1000 s and 3000 s. The results are summarized in Table 5.1. Since $x_{max} = 40$ and there are two time durations(of 1000s and 3000s), when Monte Carlo simulation is used to estimate the PMF $P(y_\tau|x)$, there are 80 sets of simulations required to generate the simulation plots in Figure 5.2. However, using the mathematical model, two sets of simulation are necessary to calculate the PMF for K, the number of microtubule trips from the transmission zone to the receiver zone in times 1000s and 3000s. Also, one simulation is needed to calculate

the PMF for V_i, the probability that the ith grid square is visited in one trip.

When Monte Carlo simulation is employed, as the number of information particles released by the transmitter increases, the actual simulation times increases asymptotically linearly. Moreover, simulating one particle released by the source take much less time than simulating 40 particles released by the source. In this case, when one particle is assumed to be released, the simulation runs until time duration τ is simulated or until the single particle is delivered. On the other hand, the simulation must run until time duration τ is simulated or 40 particle are delivered when 40 particles are assumed to be released. However, using the proposed mathematical model, the simulation times are constant regardless of the number of particles released by the transmitter. Therefore, using the proposed mathematical model the channel capacity can be calculated much more quickly than the Monte Carlo simulations.

For example, Monte Carlo simulations in MATLAB [255] were executed on average desktop computers equipped with Intel Core Due processors with different speeds ranging from 2.5GHz to 3.0GHz, to generate the channel capacity plots shown in Figure 5.2. Specifically, 40 CPU cores were used for one week to estimate the $P(y_\tau|x)$ using Monte Carlo simulations. However, the plots based on the mathematical model were generated using three CPU cores in less than a day.

In the next section, more accurate but more complex models using Markov

chains are presented.

5.2 Markov Chain Channel Models

Deriving a channel model for molecular communication systems with stationary kinesin and a single moving microtubule filament is very difficult because of the complex motion of the microtubule and the shape of the channel. Moreover, for channels with multiple microtubules the problem becomes even more difficult because of the dependencies between all the microtubules. In this case, although the movement of the microtubules themselves can safely be considered to be independent [80, 85], the delivery of an information particle depends on whether it has already been picked up by other microtubules. Therefore, many dependencies will be introduced into the system, which makes the derivation of mathematical models extremely difficult. To overcome these issues, the channel models are derived in two steps. First, the case where a single microtubule is inside the channel is considered. Although the problem of modeling the channel with a single microtubule is still quite complex, by focusing on a single microtubule, the dependencies between different microtubules is eliminated and channel models that resemble closed-form expressions can be derived. Then the results are extended to channels with multiple microtubules and a Markov chain model is proposed, where transition probabilities can be calculated using simpler simulations compared to full simulations of the

channels.

5.2.1 Single Microtubule Channels

For the rest of this section it is assumed that there is a single microtubule inside the channel, delivering the information particles. Therefore, an arbitrary information particle is received at the destination if it is picked up by this microtubule filament from the transmission zone, and then delivered to the receiver zone. Considering this fact, a *single microtubule trip* is defined as the movement of the microtubule from anywhere in the channel to the transmission zone and then the receiver zone. For example, a single microtubule trip is shown in Figure 4.2. After the microtubule completes its first trip, subsequent trips are defined as the movement of the microtubule from the receiver zone to the transmission zone and back. During any trip, a microtubule can deliver zero or more information particles, up to its maximum load capacity.

Let K_τ be the total number of microtubule trips during the time per channel use duration τ, and l_{max} be the maximum load capacity of the microtubule. Since the motion of the microtubule filament is random in nature, for a given value of time duration τ, the number of trips is random and represented by the PMF $P(k_\tau)$.

Therefore, the PMF $P(y_\tau \mid x)$ can be written as

$$P(y_\tau \mid x) = \sum_{k_\tau \in K} P(y \mid x, k_\tau) P(k_\tau), \tag{5.7}$$

where x is the number of information particles transmitted, and y_τ is the number of information particles received. This follows because the number of particles that arrive at the destination during the time duration τ, Y_τ, depends on the number of trips the microtubule takes during that time, K_τ.

From Equation (5.7) $P(y_\tau \mid x)$ can be calculate if the PMFs $P(k_\tau)$ and $P(y \mid x, k_\tau)$ are known. The first PMF $P(k_\tau)$ can be estimated using a simple Monte Carlo simulation of the motion of the microtubule inside the channel. This Monte Carlo simulation is very simple compared to the full simulations, because it does not include simulation of the loading and unloading processes. To estimate the PMF $P(y \mid x, k_\tau)$, a Markov chain model is derived.

Let X_i be the number of information particles at the transmission zone at the *end* of the ith microtubule trip, $Y^{(i)}$ be the total number of information particles already delivered to the receiver at the *end* of the ith microtubule trip, and D_i be the number of information particles delivered during the ith microtubule trip, for $i = 1, 2, 3, \ldots$. Therefore, $D_i \in \{0, 1, 2, \cdots, l_{max}\}$ is a random variable representing the number of particles delivered during the ith trip, where l_{max} is the maximum

load the microtubule could have. Each trip is related to the next trip by the following properties

$$X_i = X_{i-1} - D_i,$$
$$Y^{(i)} = Y^{(i-1)} + D_i, \tag{5.8}$$

where it is assumed that $X_0 = X$ is the number of particles released by the transmitter at the beginning of the communication session, and $Y^{(0)} = 0$ is the number of particles initially at the receiver.

It is assumed that the number of information particles at the transmission zone and the receiver zone at the end of each trip iteration depends only on the previous iteration. Mathematically, this can be written as

$$P(Y^{(i)} \mid X, Y^{(i-1)}) = P(Y^{(i)} \mid X, Y^{(i-1)}, Y^{(i-2)}, \cdots, Y^{(0)}), \tag{5.9}$$

which satisfies the Markov property. Therefore, given the number of information particles released by the transmitter X, $Y^{(i)}$s form a Markov Chain.

Figure 5.3 shows the graphical representation of this Markov chain. Each state represents the number of information particles delivered to the receiver with state s indicating the starting state. The transition probabilities can be calculated from the

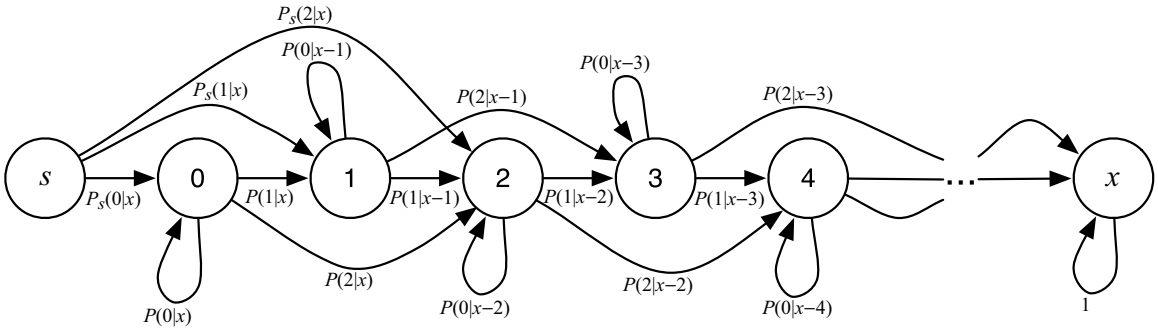


Figure 5.3: The Markov chain representing the number of information particles received at the destination. State s is the starting state, and each of the other states represents the number of information particles that are delivered to the destination. it is assumed that the maximum number of particles a single microtubule can load is $l_{max} = 2$ to generate a simplified and comprehensible figure. It is assumed that x information particles are released by the transmitter. The transition probabilities are given by $P(D_i = d \mid X_{i-1} = x_{i-1})$. The transition probabilities are different for the starting state s, since for the first trip the microtubule can start its trip from anywhere in the channel, while for subsequent trips the microtubule starts from the receiver.

PMF $P(D_i|X_{i-1})$ (the probability of the number of particles delivered during the ith trip given there were X_{i-1} particles at the transmission zone at the beginning of the trip) as follows

$$P(Y^{(i)} = y^{(i)} \mid X = x, Y^{(i-1)} = y^{(i-1)}) =$$

$$\begin{cases} P(D_i = y^{(i)} - y^{(i-1)} | X_{i-1} = x - y^{(i-1)}) & y^{(i-1)} \leq y^{(i)} \leq y^{(i-1)} + l_{max} \\ 0 & \text{otherwise} \end{cases} . \quad (5.10)$$

For the first trip, the microtubule can start from anywhere in the channel uni-

formly at random. However, for all the subsequent trips, the microtubule starts from the receiver zone. Therefore, to distinguish between these two cases, the notation $P_s(D_1|X_0)$ is used for the first trip, and $P(D_j|X_{j-1})$, with $j = 2, 3, \ldots$, for all the subsequent trips. Moreover, it is assumed that

$$\begin{aligned} P(D_2 = d|X_1 = x) =& P(D_3 = d|X_2 = x) = \cdots \\ &= P(D_i = d|X_{i-1} = x) = \cdots . \end{aligned} \tag{5.11}$$

This assumption is valid since the number of information particles delivered at trip iterations greater than one, are dependent only on the number of information particles in the transmission zone at the end of the previous iteration. The equality follows because after the first trip, each subsequent trip starts from the receiver zone.

The probability transition matrix given x particles are released by the trans-

mitter is defined as

$$\mathbf{P}(x) = \begin{bmatrix} 0 & P_s(0|x) & P_s(1|x) & P_s(2|x) & \cdots & P_s(l_{max}|x) & 0 & 0 & 0 & \cdots \\ 0 & P(0|x) & P(1|x) & P(2|x) & \cdots & P(l_{max}|x) & 0 & 0 & 0 & \cdots \\ 0 & 0 & P(0|x-1) & P(1|x-1) & \cdots & P(l_{max}|x-1) & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & P(0|x-2) & \cdots & P(l_{max}|x-2) & 0 & 0 & 0 & \cdots \\ \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ \cdots & 0 & 0 & 0 & 0 & 0 & 0 & P(0|2) & P(1|2) & P(2|2) \\ \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 & P(0|1) & P(1|1) \\ \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \tag{5.12}$$

where $\mathbf{P}(x)$ is an $(x+2) \times (x+2)$ matrix. Each element in the matrix is equal to 0, $P_s(D_1|X_0)$, or $P(D_j|X_{j-1})$. The PMFs $P_s(D_1|X_0)$ and $P(D_j|X_{j-1})$ can be estimated from Monte Carlo simulations that are much more efficient than the full simulations. Instead of simulating multiple microtubule trips, only single trip simulations are needed for estimation of $P_s(D_1|X_0)$ and $P(D_j|X_{j-1})$. Therefore, instead of simulating the motion of the microtubule for the time per channel use interval τ, a single microtubule trip must be simulated which is typically much smaller than τ.

In Equation (5.7), it was shown that when PMFs $P(k_\tau)$ and $P(y \mid x, k_\tau)$ are known, the PMF $P(y_\tau \mid x)$ can be calculated. The PMF $P(k_\tau)$ can be estimated using a simple Monte Carlo simulation, while the PMF $P(y \mid x, k_\tau)$ can be calculated

using the probability transition matrix $\mathbf{P}(x)$ as

$$P(y \mid x, k_\tau) = \mathbf{s}(x)\mathbf{P}(x)^{k_\tau}, \tag{5.13}$$

where $\mathbf{s}(x)$ is the initial state distributions represented as a row vector by

$$\mathbf{s}(x) = (1, \underbrace{0, 0, \ldots, 0}_{x+1}). \tag{5.14}$$

This follows because it is always assumed that the Markov chain starts at state s. Substituting Equation (5.13) into Equation (5.7),

$$P(y_\tau \mid x) = \sum_{k_\tau \in K} P(k_\tau)\mathbf{s}(x)\mathbf{P}(x)^{k_\tau}. \tag{5.15}$$

Therefore, the PMF $P(y_\tau \mid x)$ can be estimated if $P(k_\tau)$, $P_s(D_1|X_0)$ and $P(D_j|X_{j-1})$ are known. All these three PMFs can be estimated using simple Monte Carlo simulations, where their combined computational time would be less than that of the full simulations used in previous works.

To further reduce the simulation time required for calculating the PMF $P(y_\tau \mid x)$, an estimated mathematical expression for the transition probabilities is derived. For simplicity, it is assumed that both PMFs $P_s(D_1|X_0)$ and $P(D_j|X_{j-1})$ are the same (i.e. there is no difference between the first initial trip or any subsequent

trips). Therefore, for this case the state s in Figure 5.3 is removed and the transition probability matrix becomes an $(x+1) \times (x+1)$ matrix. The initial state is also changed to state 0 with probability one. To derive this model, it is also assumed that a grid loading structure is used to capture the loading process with n squares in the grid, where $n \geq x_{max}$ represents the number of places (squares) where an information particle could be released from, and x_{max} represents the maximum number of information particles that could be released by the transmitter.

Let p_{D_i} be the probability that an information particle from any of the squares is delivered to the destination during the ith microtubule trip. This probability can be calculated as

$$p_{D_i} = p_V \times \frac{x_{i-1}}{n}, \tag{5.16}$$

where x_{i-1} is the number of information particles in the transmission zone at the end of the previous trip, and p_V is the probability that a square in the grid is visited during a single microtubule trip. For simplicity, it is assumed that this probability is the same for all the squares in the grid. For example, as will be shown in the next chapter, the optimal transmission zone tends to be close to the walls of the channel, and the squares would have similar probability of being visited. The same Monte Carlo simulation which is used to estimate $P(k_\tau)$, can be used to estimate p_V through saving an extra parameter. Therefore, the computational complexity

of calculating $P(k_\tau)$ and $p_{_V}$ together is $O(\frac{\tau}{\Delta t}N)$, where N is the number of trials for obtaining the Monte Carlo estimation. Typically, the larger the value of N the more accurate the estimation.

The PMF $P(D_i|X_{i-1})$ can be estimated from $p_{_{D_i}}$ as

$$P(D_i = d \mid X_{i-1} = x_{i-1}) =$$
$$\begin{cases} \binom{n}{d} p_{_{D_i}}^d (1 - p_{_{D_i}})^{n-d} & 0 \leq d \leq l_{max} - 1 \\ \sum_{m=l_{max}}^{n} \binom{n}{m} p_{_{D_i}}^m (1 - p_{_{D_i}})^{n-m} & d = l_{max} \\ 0 & \text{otherwise} \end{cases} . \tag{5.17}$$

This follows since the probability that an information particle is delivered from a specific square in the grid is independent from other squares, and the maximum load a microtubule can have is l_{max}.

The PMF $P(y_\tau \mid x)$ can be calculated using this technique with a computational complexity of $O(\frac{\tau}{\Delta t}N)$. This is significantly smaller than the computational complexity of full simulations $O(\frac{\tau}{\Delta t}Nx_{max})$.

5.2.1.1 Evaluation of Single Microtubule Channel Model

The accuracy of the proposed Markov chain models is evaluated with respect to the full computer simulation environment developed in Chapter 4. For evalua-

tion, Kullback-Leibler (K-L) distance (also known as Kullback-Leibler divergence or relative entropy) is used [192]. K-L distance is calculated as

$$D\left(P_{\text{sim}} \parallel P_{\text{model}}\right) = \sum_{y_\tau \in \mathcal{X}} P_{\text{sim}} \log \frac{P_{\text{sim}}}{P_{\text{model}}}, \tag{5.18}$$

where P_{sim} and P_{model} represent the PMF $P(y_\tau \mid x)$, estimated using full Monte Carlo simulations or the Markov chain models, respectively. K-L distance signifies the amount of information lost if P_{model} is used instead of P_{sim}.

To better quantify this value, instead of using it directly the ratio of K-L distance over the entropy of the PMF $P_{\text{sim}}(y_\tau \mid x)$ is used, which is obtained through full Monte Carlo simulations. In this case, the entropy is calculated as

$$\begin{aligned} H(Y_\tau | X = x) = \\ - \sum_{y_\tau \in \mathcal{X}} P_{\text{sim}}(y_\tau \mid x) \log P_{\text{sim}}(y_\tau \mid x). \end{aligned} \tag{5.19}$$

Using Equations (5.18) and (5.19) the performance measure is given by

$$\mathcal{R} = \frac{D\left(P_{\text{sim}} \parallel P_{\text{model}}\right)}{H(Y_\tau | X = x)}. \tag{5.20}$$

This ratio gives a direct measure of normalized error for the estimation. In general the closer the ratio is to zero, the better the models estimate the PMFs. However,

it is still difficult to decide what would be a good value for this ratio (i.e. for what ratio the Markov chain calculated PMFs are acceptable compared to full Monte Carlo simulations).

To further investigate the accuracy of the models, an actual channel parameter, namely channel capacity, is calculated using both full simulation based PMFs and Markov chain model based PMFs.

In the rest of this section, the parameters used in all Monte Carlo simulations are as follows: simulation time steps of $\Delta T = 0.1$ seconds (this is different from the discretization time step Δt introduced in the previous section), MT diffusion coefficient $D = 2.0 \cdot 10^{-3}$ $\mu m^2/s$, average speed of the MT $v_{avg} = 0.5$ $\mu m/s$, and persistence length of the MT trajectory $L_p = 111$ μm. It is also assumed that the size of the information particles is $1\mu m$, the average length of the MTs is $10\mu m$, and each MT can load up to 5 information particles in one trip from the transmission zone to the receiver zone. These parameters are all selected based on experimental observations of DNA covered MTs moving over a kinesin covered substrate [80, 85].

Three different channels are considered: a square channel of 40 μm $\times$ 40 μm, an 8-sided regular polygon (octagon) channel with radius 23.78 μm, and a 20-sided regular polygon with radius 25.57 μm. The grid transmission zone structure is always assumed to be along the channel walls, since it is the optimal placement as will be shown in the next chapter.

The transition probabilities $P(D_i \mid X_{i-1})$ can be estimated using two different techniques: Monte Carlo simulations, and Equation (5.17). When Monte Carlo simulations are used to estimate the transition probabilities, two sets of simulations are performed to estimate $P_s(D_1 \mid X_0)$, the number of particles delivered during the first trip, and $P(D_j|X_{j-1})$, the number of particles delivered during subsequent trips ($j = 2, 3, \ldots$). When Equation (5.17) is used for calculation of transition probabilities, an estimate for p_v (the probability that a square in the grid loading structure is visited) is required. This probability can be obtained from the same simulation that is used to estimate the PMF of the number of microtubule trips $P(k_\tau)$.

Figure 5.4 shows the results for the time per channel use value of $\tau = 1000$ seconds. The ratio of K-L distance between the simulation and model based PMFs to the entropy of simulation PMFs is plotted against different values of x_{max} (maximum number of information particles that could be released). The term "Sim Tran" is used in the plot legend to indicate the case where the transition probabilities are calculated using Monte Carlo simulations, and the term "Eq. (5.17) Tran" to indicate the case where transition probabilities are calculated using Equation (5.17). In practice large number of information particles can be released by the transmitter. At the same time, channel capacity increases with the size of the symbol set. Therefore, the value of this ratio at $x_{max} = 34$ is considered. It can be seen

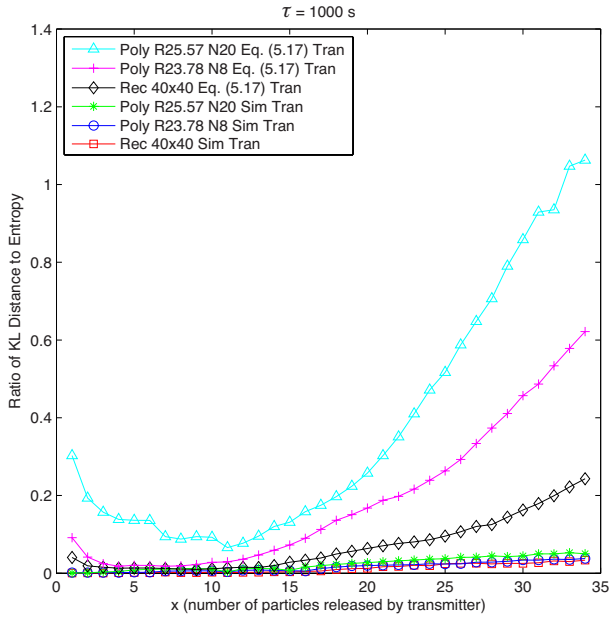


Figure 5.4: The ratio of K-L distance between the simulation and model based PMFs to the entropy of simulation based PMFs for different values of x (the number of particles released by the transmitter). Channels contain only a single microtubule.

that, as expected, when simulations are used to estimate the PMFs, lower ratio is achieved. However, transition probabilities and hence the PMFs can be estimated more quickly when Equation (5.17) is used.

Another property observed in Figure 5.4 is that for "Eq. (5.17) Tran" plots the $\mathcal{R}$ ratio increases as x_{max} increases. This follows because of some independence assumptions made in derivation of Equation (5.17), which simplifies the model. In particular, it is assumed that the probability that a particle is in a given square during the ith microtubule trip is x_{i-1}/n. However, this assumption does not take

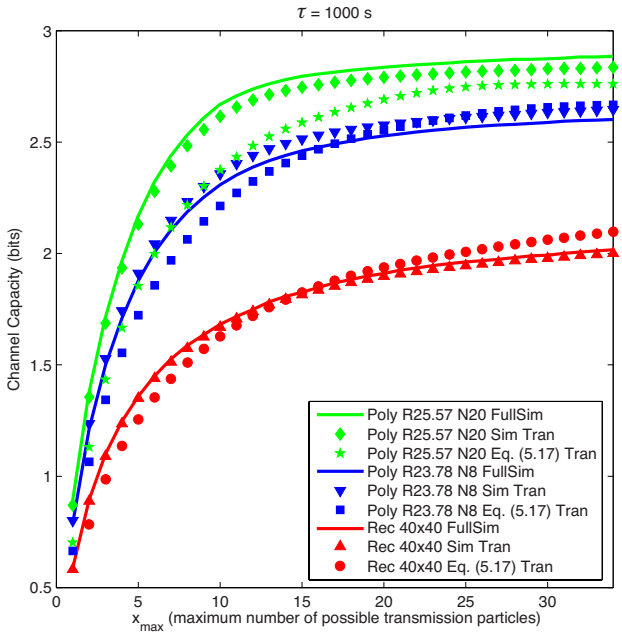


Figure 5.5: Channel capacity in bits versus $x_{\max}$, maximum number of information particles that can be released by the transmitter, calculated based on full simulations (solid lines), and Markov chain models (points). Channels contain only a single microtubule.

into consideration the fact that the particle may have already been picked up from this particular square. In other words, for each trip the remaining particles are redistributed across the squares. This error increases as the symbol size increases and hence the increasing $\mathcal{R}$ ratio.

To further investigate the estimated PMFs, and determine if they are accurate enough to characterize the channel, they are used for calculating channel capacity, which is one of the most important parameters of any communication channel. In particular, the channel capacity is calculated based on the PMF $P(y_\tau \mid x)$ obtained

using three different methods: full Monte Carlo simulations, the Markov chain model with full simulation-based transition probabilities, and the Markov chain model with transition probabilities estimated using Equation (5.17).

Figure 5.5 shows the calculated channel capacities versus x_{max}, the maximum number of particles the transmitter can transmit. The solid lines represent the channel capacities calculated using full Monte Carlo simulations, while the point plots show the channel capacities calculated using the Markov chain models. From the figure it is seen that when the Markov chain model with simulation estimation of transition probabilities is used, the estimated channel capacities are very close to the ones obtained from full simulations. Although, the channel capacities calculated using the model with transition probabilities based on Equation (5.17) are not as accurate, they are still relatively close to the ones obtained using full simulations, and can be calculated much faster. Comparing Figures 5.4 and 5.5, it can be seen that the estimated channel capacities are sufficiently accurate for design. Therefore, from a telecommunication perspective, the model's performance is satisfactory.

Finally, the computer simulation times required for calculating the PMF $P(y_\tau \mid x)$ are compared based on three different cases: full simulation, Markov chain model based on simulation estimation of $P(k_\tau)$, $P_s(D_1 \mid X_0)$, and $P(D_j|X_{j-1})$, and Markov chain model based on simulation estimation of $P(k_\tau)$ and p_v and using Equation (5.17). All the Monte Carlo simulations use the same propagation engine for the

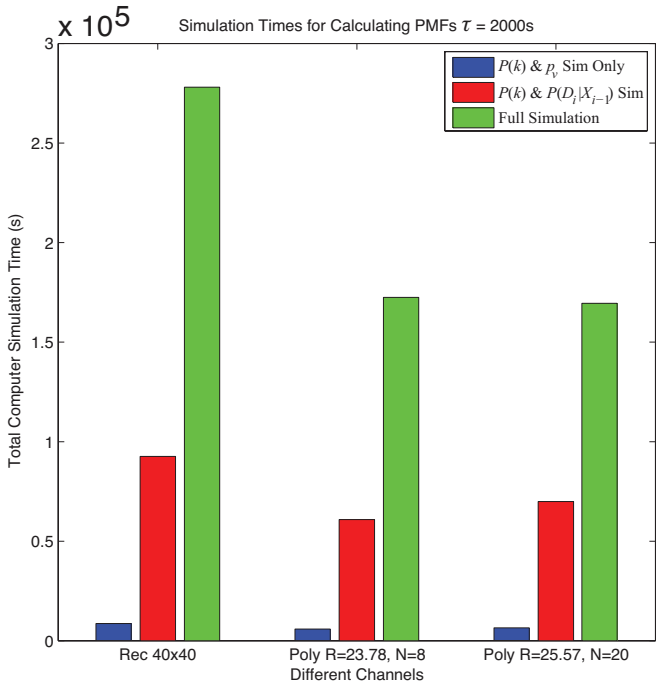


Figure 5.6: The simulation times required to calculate the PMFs $P(y_\tau \mid x)$ for $\tau = 2000$ s, $x = 1, 2, \ldots, 34$, and 5000 iteration per each Monte Carlo simulation. The number on the y-axis are multiplied by 10^5.

motion of the microtubule inside the channel. Therefore, although each simulation estimates a different PMF, the propagation engine is the same between them. Therefore, the algorithm used for computing different PMFs are computationally very similar.

The simulation times are all obtained by running the simulation code on the same notebook: Mac Book Pro (mid 2010) with 2.66 GHz Intel Core i7, 8GB 1067MHz DDR3 RAM, and Mac OS X version 10.7.4. The simulations were written in Matlab [255] for Mac OS X. Figure 5.6 shows the resulting simulation times for

calculating these PMFs, when time per channel use value is $\tau = 2000$ seconds, the number of information particles that are released by the transmitter ranges from 1 to 34, and the number of iteration used for each of the Monte Carlo simulations is 5000.

When instead of full simulations, the Markov model is used with simulated transition probabilities the simulation times are reduced by more than half. Moreover, when Equation (5.17) is used to estimate the transition probabilities, the simulations duration is reduced by more than 25 times compared to full simulations. For example, for the 40×40 square channel the simulation durations are about 77 hours, 26 hours and 2 hours for the full simulation, Markov chain model with simulation estimation of transition probabilities, and Markov chain model with transition probabilities calculated using Equation (5.17), respectively. Although in the latter case the estimated PMF is not as accurate as the other two methods, the gains in terms of simulation time are significant compared to the loss in accuracy. From these results, use of Equation (5.17) and the Markov model are proposed for initial system design and optimization, and use of simulation based calculation of transition probabilities for final stages of system design.

5.2.2 Multi-Microtubule Channel Models

In the previous section, a Markov chain model for active transport molecular communication systems employing a single microtubule was derived. To do this, a single microtubule trip was defined as the motion of the microtubule from the receiver zone to the transmission zone and back, and used a random variable K_τ to represent the number of trips during the time per channel use interval τ. When there is only one microtubule in the channel, the trips are ordered (i.e. each trip happens after the previous trip) and therefore the Markov property of Equation (5.9) holds. However, if there are multiple microtubules in the channel, two or more microtubules might be visiting the transmission zone simultaneously or be in between trips. Therefore, the number of particles delivered at a given time interval depends on all the microtubules and the Markov property of Equation (5.9) does not hold. Because of the interdependence of microtubules and the fact that the Markov property does not hold, modeling the communication system when multiple microtubules are inside the channel is extremely difficult. To overcome these issues, time is discretized instead of microtubule trips.

In particular, the time per channel use value τ is discretized into M equal time

intervals of $\Delta\tau$ as

$$\tau = M\Delta\tau. \tag{5.21}$$

Therefore, τ is broken down into M time steps. Let $m = 1, 2, \ldots M$ be each of these time steps. Furthermore, let X_m be the number of information particles at the transmission zone at the *end* of the mth time step, $Y^{(m)}$ be the total number of information particles already delivered to the receiver at the *end* of the mth time step, and D_m be the number of information particles delivered during the mth time step. Therefore, $D_m \in \{0, 1, 2, \cdots, X_{m-1}\}$ is a random variable representing the number of particles delivered during the mth time step of length Δt. Each time step is related to the next time step by the following properties

$$\begin{aligned} X_m &= X_{m-1} - D_m, \\ Y^{(m)} &= Y^{(m-1)} + D_m, \end{aligned} \tag{5.22}$$

where it is assumed that $X_0 = X$ is the number of particles released by the transmitter at the beginning of the communication session, and $Y^{(0)} = 0$ is the number of particles initially at the receiver.

With this new discretization, a Markov chain model can be obtained similar to

the one derived in the previous section for single microtubule channels. In this case the probability transition matrix given x particles are released by the transmitter is given by

$$\mathbf{P}(x) = \begin{bmatrix} P(0|x) & P(1|x) & P(2|x) & \cdots & P(x-2|x) & P(x-1|x) & P(x|x) \\ 0 & P(0|x-1) & P(1|x-1) & \cdots & P(x-3|x-1) & P(x-2|x-1) & P(x-1|x-1) \\ 0 & 0 & P(0|x-2) & \cdots & P(x-4|x-2) & P(x-3|x-2) & P(x-2|x-2) \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & P(0|2) & P(1|2) & P(2|2) \\ 0 & 0 & 0 & \cdots & 0 & P(0|1) & P(1|1) \\ 0 & 0 & 0 & \cdots & 0 & 0 & 1 \end{bmatrix}, \tag{5.23}$$

where transition probabilities are in the form $P(D_m = d|X_{m-1} = x)$, and the number of particles delivered during step m is $d = 0, 1, 2, \cdots, X_{m-1}$. Because of the dependencies between multiple microtubules discussed earlier, deriving a closed-form expression for the transition probabilities is not possible. However, the transition probabilities can be estimated using Monte Carlo simulations by using $\Delta\tau$ instead of τ, which would result in shorter simulations. Therefore, in general M-fold improvement in computation time of the simulations is expected compared to full simulations. Moreover, it would be expected that the estimated PMFs be more accurate for smaller M (i.e. larger time steps $\Delta\tau$) and less accurate for larger M (i.e. smaller time steps $\Delta\tau$). This is because smaller time steps capture the

motion of the microtubules for a shorter amount of time and therefore calculated transition probabilities would be less accurate. Similarly, larger time steps capture a longer duration of motion of the microtubules, which results in more accurate estimation of transition probabilities.

5.2.2.1 Evaluation of Multi-Microtubule Channel Model

In this section, the evaluation of the Markov chain model for channels with multiple microtubules is considered. The same three channel shapes used in the evaluation of the single microtubule model in the previous section is considered with similar simulation parameters. To get an accurate estimation of the transition probabilities, the discretization time steps $\Delta\tau$ must be such that at least a single microtubule trip is captured in each time step for most of the microtubules in the channel. Let P be the perimeter of the channel, and v_{avg} be the average speed of the microtubules. Because microtubules mostly follow the walls of the channel, the desirable values for $\Delta\tau$ are given by

$$\Delta\tau \geq \frac{P}{v_{\text{avg}}}. \tag{5.24}$$

In general the larger the value of $\Delta\tau$ is compared to the ratio on the right hand side of the equation, the better the transition probabilities are estimated. This follows because in this case more microtubule trips are captured in each time step, which

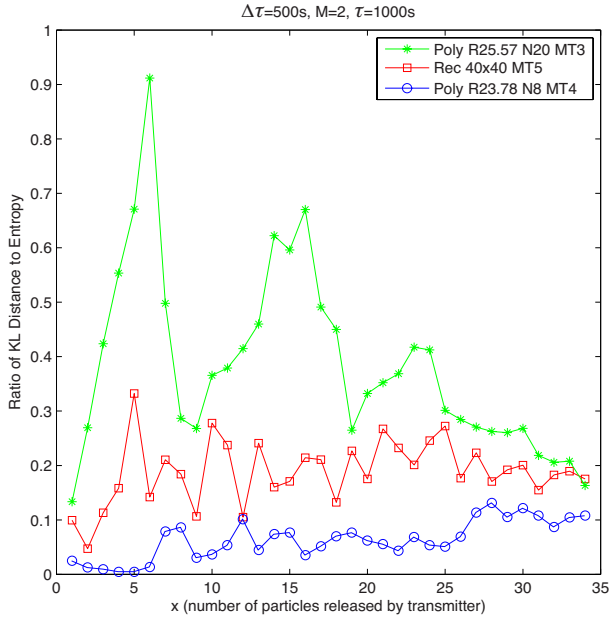


Figure 5.7: The ratio of K-L distance between the simulation and model based PMFs to the entropy of simulation PMFs for different values of x (the number of particles released by the transmitter). Channels contain multiple microtubules.

results in a more accurate estimation of the transition probabilities of the Markov chain model.

Figure 5.7 plots the ratio of K-L distance between the simulation and model based PMFs to the entropy of simulation PMFs for different values of x. It is assumed that there are 5 microtubules in the square channel, 4 microtubules in the octagon channel, and 3 microtubules in the 20-sided polygon channel. Different numbers haven been chosen to demonstrate that the model works for any number of microtubules. For the channels considered in this plot, it is assumed that $\Delta\tau = 500$

seconds (this value satisfies Equation (5.24) criterion), and the time per channel use is 1000 seconds (i.e. the number of steps $M = 2$).

In practice large number of information particles can be released by the transmitter. At the same time, channel capacity increases with the size of the symbol set. Therefore, the value of this ratio at $x_{\max} = 34$ is considered. For each channel the number of time steps is $M = 2$, and therefore the discretization intervals are one half the time per channel use duration. Therefore, the transition probabilities can be calculated in one half the time it takes for full simulations.

To further investigate the accuracy of the estimated PMFs, they are used to calculate the channel capacity, which is one of the most important parameters of any communication channel. Figures 5.8–5.10 compare the channel capacity of each channel shape obtained through full simulations (line plots), to the ones calculated using the proposed Markov chain model (point plots). For each channel, the number microtubules are taken to be 3, 4, and 5. For both polygon channels $\Delta\tau = 400$ seconds and $\tau = 800$ seconds, and for the square channel $\Delta\tau = 550$ seconds and $\tau = 1100$ seconds. These values have been chosen since the plots for different number of microtubules are all visible on a single plot at these durations, and to show that the model works for different values of τ. As can be seen the channel capacities can be estimated fairly accurately at half the time it takes for full simulations. Moreover, by increasing the value of M (i.e. decreasing the discretization time

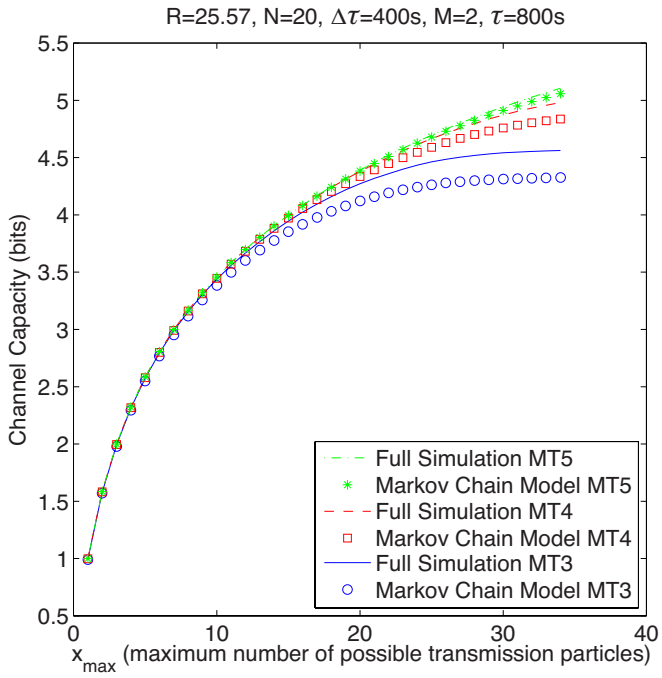


Figure 5.8: Channel capacity in bits per channel use versus $x_{\max}$, maximum number of information particles that can be released by the transmitter, for both full simulations and Markov chain model for multiple microtubules. The channel here is a 20-sided regular polygon channel with radius of 25.57 μm.

intervals) channel capacities can be estimated more quickly at the cost of a loss in accuracy.

The PMFs $P(y_\tau \mid x)$ estimated using the Markov chain model are sufficiently accurate for design, and become more accurate as the time step durations $\Delta\tau$ increases compared to the term on the right hand side of Equation (5.24). This effect can be seen in the figure; the octagon channel has a better estimated channel capacity, since it has a smaller perimeter compared to the 20-sided polygon channel

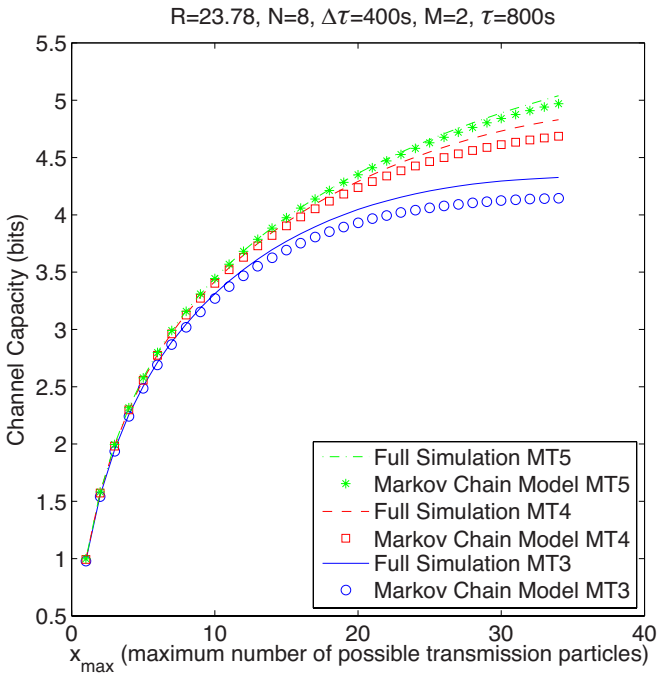


Figure 5.9: Channel capacity in bits per channel use versus $x_{\max}$, maximum number of information particles that can be released by the transmitter, for both full simulations and Markov chain model for multiple microtubules. The channel here is an octagon channel with radius of 23.78 μm.

for different number of microtubules. Therefore, for $\Delta\tau = 400$ more microtubule trips are captured in estimated tradition probabilities of the Markov chain model, which results in a slightly more accurate estimation of channel capacity.

Finally, as can be seen channel capacity increases with the number of microtubules for all channels. In Figure 5.11, the channel capacity is plotted for different number of microtubules in the channel. The channel is a square channel with 40 μm sides, and the size of the message set is 34. As can be seen capacity increases

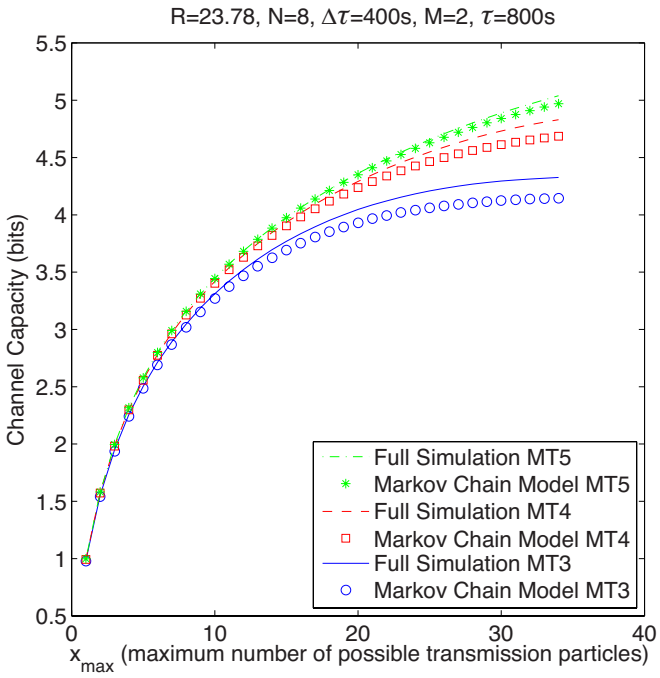


Figure 5.10: Channel capacity in bits per channel use versus $x_{\max}$, maximum number of information particles that can be released by the transmitter, for both full simulations and Markov chain model for multiple microtubules. The channel here is a square channel of length 40 μm.

linearly with number of microtubules for each value of τ. Moreover, the slope increases as time per channel use increases.

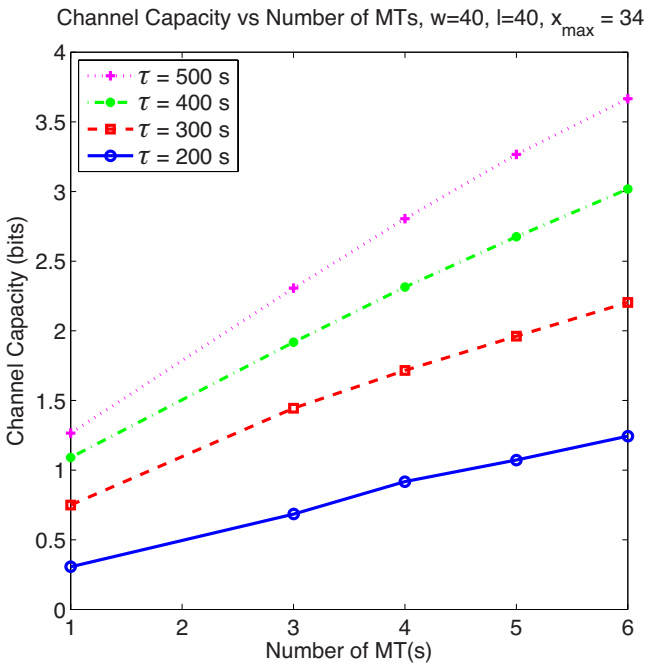


Figure 5.11: Channel capacity in bits versus the number of microtubules for a square channel with 40 μm sides and $x_{\max} = 34$. Capacity increases linearly with the number of microtubules.

6 Optimal Design Strategies for On-Chip Active Transport Molecular Communication

In Chapter 4, a simulation environment for on-chip molecular communication based on different propagation schemes was developed, and it was shown that active transport can be a suitable choice for on-chip molecular communication. Then, in Chapter 5, different mathematical models for on-chip active transport was derived to reduce the computational complexity of simulating these channels, and to gain some insight into this propagation scheme.

In this chapter, some of the models developed in Chapter 5 is used to find optimal design strategies for on-chip active transport. Two important design parameters are considered. First, the problem of finding the optimal size for the information particles is discussed, and the optimal shape for the transmission region in the channel is obtained. Next, the problem of finding the optimized channel shape that would maximize the rate or transport and information transmission rate is considered.

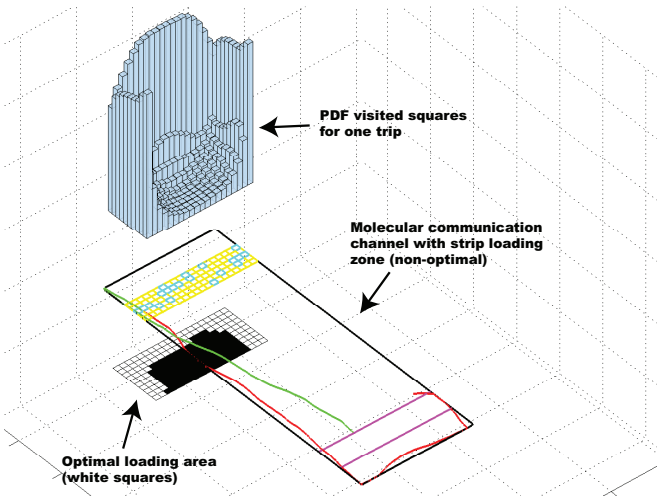


Figure 6.1: *(Top)*: Probability distribution of $P(V_i = 1)$ for squares of size 1μm to the left side of the loading area. *(Middle)*: Strip transmission area (yellow and cyan squares) for $n = 100$ squares. *(Bottom)*: Projection of the probability distribution $P(V_i = 1)$ on top. The top 100 values of $P(V_i = 1)$ are shown in as white squares and they represent the optimal loading area.

6.1 Optimizing the Transmission Zone

The transport models presented in the previous chapter, and the fact that the microtubules mostly move along the walls of the molecular communication channel, are used to optimize the transmission area. For the rest of this section the analysis will be based on rectangular-shaped channels, but the same technique can be applied to any channel shape.

Recall from the top part of the Figure 5.1, which is depicted here again as Figure 6.1, that microtubules mostly move along the walls, and therefore $P(V_i = 1)$ (probability that a square is visited in one microtubule trip) is higher for squares

close to the walls of the molecular communication channel. An information particle is picked up from the transmission zone, and delivered to the receiver zone, if the corresponding square is visited. Therefore, it is desirable to find squares with maximum $P(V_i = 1)$, which are squares that have the highest probability of being visited during one trip. In Figure 6.1, the probability distribution of all the squares of length 1μm to the left of transmission area are plotted (the bar plot on the top). The middle plot in Figure 6.1 shows the original *strip transmission area* with 100 squares. Notice that the squares in the middle of the strip transmission zone have low probability of being visited in a single microtubule trip. The bottom plot shows the projection of the probability distribution function of $P(V_i = 1)$. The first 100 squares with highest probabilities are shown as white squares, and the rest of the squares are shown in black. Note that in this projection, the minimum distance between the transmission area and the receiver area is still the same and the transmitter and the receiver are the same distance apart. Finally, according to the transport models in Chapter 5, this white area is the optimal transmission zone that will give us the highest information rate because probability of visiting and picking up particles is highest.

6.1.1 Vesicular Encapsulation

As mentioned in the Chapter 2, information-carrying particles could be encapsulated inside lipid vesicles to keep them isolated from the propagation environment, thereby preventing destructive chemical reactions that could result in unsuccessful detection at the receiver. The process of on-chip vesicular encapsulation was demonstrated in [59].

In this section, optimal strategies for encapsulating information particles are presented. For example, is it better to encapsulate all the information molecules inside one giant vesicle or inside several smaller ones? Is there an optimal vesicle size for a given molecular communication channel? To answer these questions, a model for vesicular encapsulation is proposed and is used along with the models developed in Chapter 5 to provide optimal design strategies.

It is assumed that egg phosphatidylcholine vesicles (a type of liposome) are used for encapsulation. Furthermore, it is assumed that lipid vesicles are perfect spheres, and the diameter of all the vesicles are exactly the same. Moreover, assume the number of information molecules per vesicle is deterministic and known. These assumptions are made to facilitate the analysis.

Let R be the inner radius of the lipid vesicle. Then the number of lipid molecules

required to create such a vesicle is given by [58]:

$$m_R = \frac{4}{3}\pi N \left((R + 3.7 \times 10^{-7})^3 - R^3\right) / (768\bar{v}), \tag{6.1}$$

where m_R is the number of lipid molecules required to create a vesicle with radius R, N is the Avogadro's number (6.022×10^{23}), R is the inner radius of the vesicle in cm, and $\bar{v}$ is the vesicle partial specific volume (0.9848 mL/g). For example, using this equation 3.72×10^7 molecules are required to create a single vesicle with diameter of 2μm.

Let M be the total number of lipid molecules available at the transmitter for creation of lipid vesicles. Then the total number of lipid vesicles of size R that can be created by the transmitter is given by

$$n_R = \frac{M}{m_R}. \tag{6.2}$$

Let x be the number of information particles to be sent by the transmitter. Then assuming there are n_R vesicles available for encapsulation, there will be $(n_R - x \bmod n_R)$ vesicles with $\lfloor x/n_R \rfloor$ information molecules and $(x \bmod n_R)$ vesicles with $\lceil x/n_R \rceil$ information particles, where (mod) is the modulo operator (i.e. the remainder of the integer division), $\lfloor . \rfloor$ the floor function and $\lceil . \rceil$ the ceiling function. For example, if there are $x = 18$ information particles and $n_R = 5$ vesicles, after the

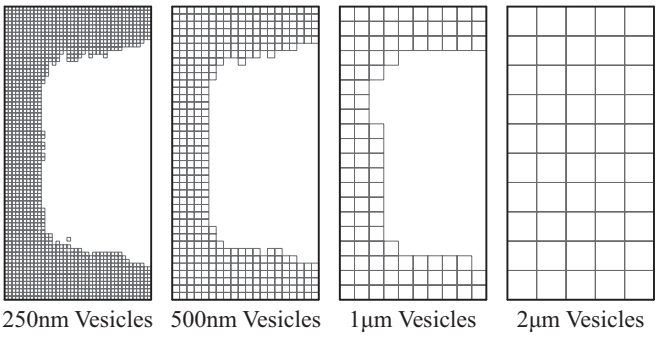


Figure 6.2: The shape of the optimal transmission area of different vesicle diameters. Each of the four rectangles is the total area that could be dedicated as the transmission area. The dimension of this area is 20 μm by 10 μm. The squares inside each transmission area represent the optimal grid for each vesicle size.

encapsulation process, 2 vesicles encapsulate 3 information particles and 3 vesicles encapsulate 4 information particles.

6.1.2 Vesicular Encapsulation Analysis and Optimization

In this section, the transport model presented in Chapter 5 and the vesicular encapsulation models presented in the previous subsection are used to analyze and optimize the vesicular encapsulation process. To show the optimization scheme, four different diameter values for vesicles are considered: 250nm, 500nm, 1μm, and 2μm. Assume that the area that could be used by the transmitter is 20 μm by 10 μm. The goal is to find the optimal vesicle size. The optimization scheme presented in Section 6.1 is used to optimize the shape of the grid transmission area for each vesicle size. Figure 6.2 shows the shape of the optimal transmission area for each

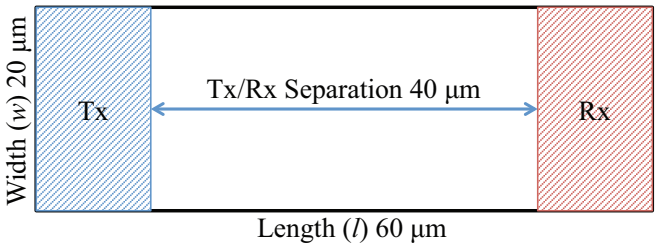


Figure 6.3: Rectangular channel used as a sample for optimization. Transmission area is the region on the left (blue area), and the receiver area is the strip on the right (red area).

diameter size. Notice that for the 2μm vesicle size, since the square grid occupies the whole transmission area, it contains squares that have a small probability of being visited during a microtubule trip.

To make the comparison Monte Carlo simulation of active transport propagation presented in Chapter 4 is employed with the following parameters: simulation time steps $\Delta t = 0.1$ seconds, microtubule diffusion coefficient $D = 2.0 \cdot 10^{-3}$ µm^2/s, average speed of the microtubule $v_{\text{avg}} = 0.5$ µm/s, and persistence length of the microtubules trajectory $L_p = 111$ µm. These parameters are all selected based on experimental observations of ssDNA covered microtubules moving over a kinesin covered substrate [80]. Furthermore, for all the subsections that follow, a rectangular propagation environment is considered, with the dimensions 20µm by 60µm presented in Figure 6.3. The transmission area is the strip on the left, while the receiver area is the strip on the right, and the separation between the transmission

Table 6.1: Simulation parameters for vesicular encapsulation

Vesicle's Diameter	Number of Squares in Grid	Microtubule's Maximum Load
250 nm	1600	20
500 nm	400	10
1 μm	100	5
2 μm	50	2

zone and the receiver zone is 40μm. In the simulations, the corresponding optimized grid transmission area are used for each vesicle size as shown in Figure 6.2. Based on experimental results, it is assumed that the length of the microtubules is 10μm, and that the number of vesicles a microtubule can load is half its length divided by the diameter of the vesicles [67]. These parameters are summarized in Table 6.1.

Since there are many different parameters effecting the encapsulation process, three distinct set of results are presented in the next three subsections. First, the amount of liquid volume transported by each vesicle size is considered (i.e. the liquid volume encapsulated inside vesicles). Since the molecular communication environment is aqueous, if it is assumed that the concentration of the information particles is constant inside vesicles, the more liquid volume is transported the more information particles are transported. Second, the size of the vesicles are kept constant and the effect of different concentrations of information particles on the number of transported information particles is studied. Third, it is assumed that a limited amount of lipid molecules are available at the transmitter for generation

of lipid vesicles, and I study whether it is better to generate small number of large vesicles or numerous small vesicles.

6.1.2.1 Effects of Vesicle Size

Table 6.2: Number of vesicles to keep the liquid volume constant at 4.19fL

Vesicle's Diameter	Number of Vesicles
250 nm	512
500 nm	64
1 μm	8
2 μm	1

The effects of vesicle size on the amount of encapsulated liquid transported from the source to the destination is considered. It is assumed that the concentration of information particles in the liquid encapsulated by vesicles is constant. Since the number of information particles inside each vesicle is related to its concentration in the encapsulated liquid and the concentration is constant, the more liquid volume is transported the more information particle is transported.

Let D be the inner diameter of the lipid vesicle. The volumetric capacity of the vesicle is therefore given by,

$$V = \frac{4}{3}\pi \left(\frac{D}{2}\right)^3 = \frac{1}{6}\pi D^3. \tag{6.3}$$

For example, assuming the diameter of the vesicle is 2μm, the volume of the liquid

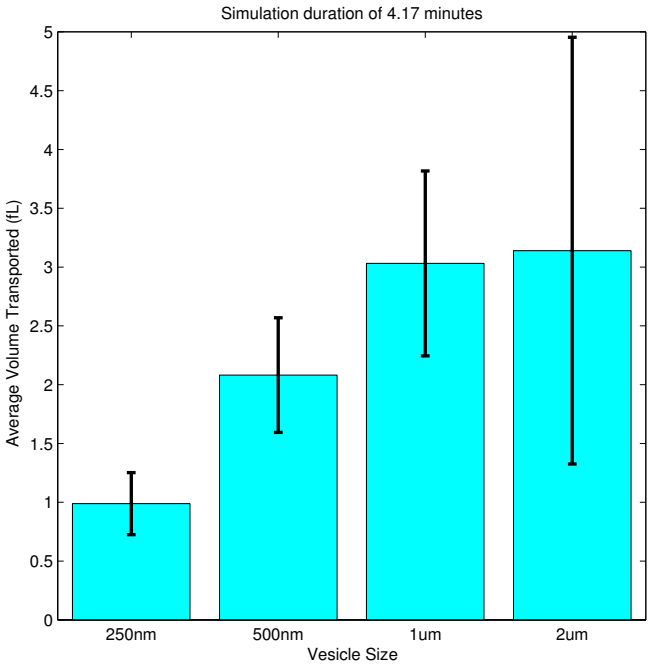


Figure 6.4: Average liquid volume transported for each vesicle size. The total liquid volume transmitted by the source is 4.19fL. The lines show the standard deviation of the received volume. (T=250s).

inside a single vesicle is 4.19fL. If the vesicle's diameter is decreased to 1μm, the liquid volume per single vesicle is 0.52fL. Therefore, liquid volume inside a single 2μm vesicle is equivalent to liquid volume inside eight 1μm vesicles. In the simulations, to keep the total liquid volume at the transmitter constant at 4.19fL, different total number of vesicles is used for each vesicle size as shown in Table 6.2.

Figure 6.4 shows the average transported liquid volume with the standard deviation bars for time per channel use of 250 seconds. As the vesicle size increases, the average liquid volume transported increases. However, the standard deviation

Table 6.3: Number of vesicles and concentration of information particles (1μm vesicle)

Concentration (molecules per fL)	Info. Particles per Vesicle	Number of Vesicles
61.54	32	1
30.77	16	2
15.39	8	4
7.69	4	8
3.85	2	16
1.92	1	32

of the transported volume also increases as the liquid volume per single vesicle is greater. When the vesicle size is increased from the 250nm to 1μm, the average liquid volume transported increases consistently. However, the difference between the 1μm and 2μm vesicles is negligible. Moreover, the standard deviation of the 1μm vesicle is smaller. These results confirm that there exists an optimal vesicle size that can be found using the simulator and mathematical models developed in Chapter 4 and 5. In this particular system, the optimal vesicle size is 1μm.

6.1.2.2 Effects of Concentration of Information Particles

The effects of the concentration of information particles inside vesicles are considered in this section. It is assumed that the vesicle size is constant at 1μm, which is the optimal vesicle size based on the results in the previous section. It is also assumed that the total number of information particles transmitted by the source is constant. In particular, assume the transmitter releases 32 information parti-

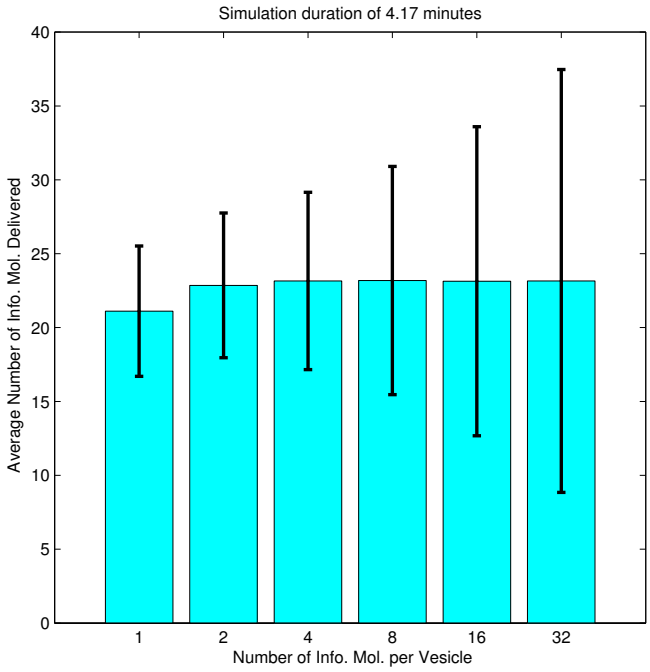


Figure 6.5: Average number of information molecules delivered for each concentration (i.e. number of information particles per vesicle). The total number of information particles transmitted is 32. The lines show the standard deviation of the number of information particles received. (T=250s).

cles. Let β be the concentration of information particles inside vesicles. Then the number of information particles inside each vesicle is given by

$$z = \beta V, \tag{6.4}$$

where z is the number of information particles per vesicle and V is the volume of the vesicle.

For example, assuming the diameter of the vesicle is 1μm and concentration of $\beta = 61.54$ molecules per fL, the number of information particles per vesicle is $z = 61.54 \times 0.52 = 32$. Therefore, assuming the number of information molecules to be transmitted by the source is 32, all the information particles can be encapsulated inside a single vesicle. Similarly, if the concentration is changed to $\beta = 30.77$ molecules per fL, the number of information particles per vesicle is $z = 30.77 \times 0.52 = 16$. Therefore, two vesicles are required to transmit 32 information particles. Following this approach, Table 6.3 summarizes different values used in the simulations.

Figure 6.5 shows the average number of information particles delivered with standard deviation bars based on each concentration for time durations per channel use of 250 seconds. One surprising outcome is that the average number of information particles delivered is very similar across different concentration values. Moreover, as the concentration of information particles per vesicle is increased the standard deviation of the number of information particles delivered is increased.

6.1.2.3 Effects of Number of Lipid Molecules

This section considers the effects of the number of lipid molecules on the number of information particles transported. The number of vesicles that can be generated based on the number of lipid molecules can be calculated using the equations pre-

Table 6.4: Number of vesicles that can be generated using 3.72×10^7 lipid molecules

Vesicle Size	Information Particles per Vesicle	Number of Vesicles
250 nm	1	64
500 nm	4	16
1 μm	16	4
2 μm	64	1

sented in Section 6.1.1. By keeping the number of lipid molecules constant, different number of vesicles can be generated for different vesicle sizes. For example, if the number of lipid molecules is about 3.72×10^7, applying Equation (6.1), a single 2μm vesicle can be generated. Similarly, using 3.72×10^7 lipid molecules, four 1μm vesicles can be generated. For the rest of this section it is assumed that the number of lipid molecules is constant at 3.72×10^7 and the number of information particles to be transmitted is also constant at 64. Table 6.4 shows the number of vesicles that can be generated as well as the number of information particles per vesicle.

Figure 6.6 shows the average number of information molecules delivered with standard deviation bars for time per channel use of 250 seconds. From the figure it can be seen that as the size of the vesicle is increased from 250nm to 1μm, the average number of information molecules delivered is increased. However, when the size of the vesicle is increased from 1μm to 2μm, the average number of molecules delivered does not change significantly. The standard deviation of the number of information particles delivered increases as the size of the vesicle increases. From these results it is concluded that there exists an optimal vesicle size for the channel

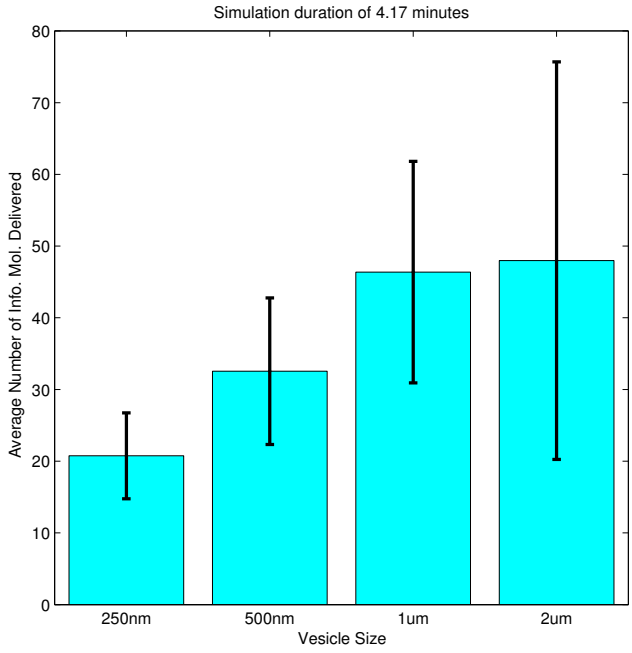


Figure 6.6: Average number of information molecules delivered for each vesicle size. The total number of information particles transmitted is 64 and the total number of lipid molecules is constant at 3.72×10^7. The lines show the standard deviation of the number of information particles received. (T=250s).

considered. Note that Figure 6.6 is very similar to Figure 6.4, which confirms that there is an optimal vesicle size for the channel (in this particular case 1μm).

6.1.2.4 Information Gains From Optimal Designs

This section evaluates the gains in achievable information rates from using the optimal transmission zone and optimal vesicle size. In the evaluation, it is assumed that the channel is the one shown in Figure 6.3, and that the simulation parameters

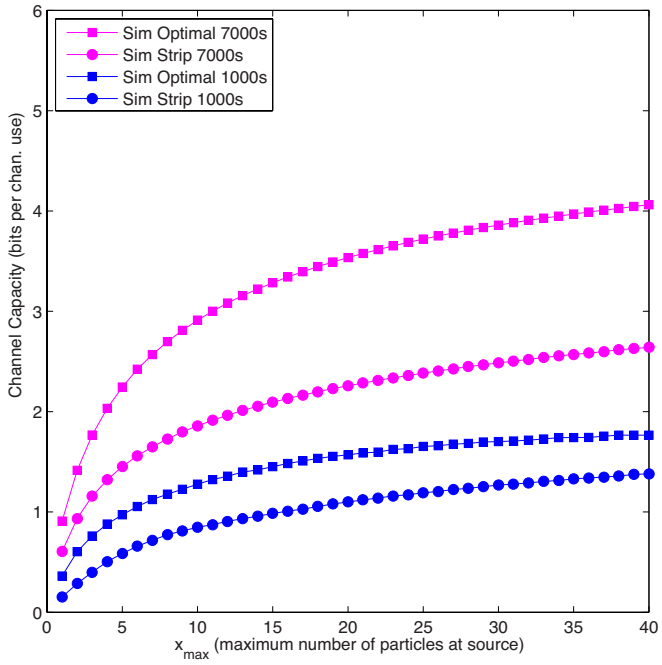


Figure 6.7: Channel capacity plot in bits per channel use. Strip transmission zone plots are presented as circles and optimal transmission zone plots are presented as squares. The blue plots present the communication time per channel use duration of 1000 s and the magenta plots represent a time per channel use duration of 7000 s. There is a significant gain in information rate by using the optimal design.

are the same as the ones used in previous chapters.

First, the information rate of the optimal transmission zone, generated using the transport model in Chapter 5, is compared to the strip transmission zone. It is assumed that the diameter of the information particles is 1μm, and there are 100 squares in the grid loading area (i.e. there could be as much as 100 information particles in the loading strip). Recall that this optimal transmission zone is represented by Figure 6.2, and the strip transmission zone is represented in the middle

part of Figure 6.1.

Figure 6.7 shows the channel capacity (i.e. maximum achievable information rate) in bits per channel use for two different transmission zones: strip transmission zone, shown using circles, and the optimal transmission zone, shown using squares. The channel capacity is plotted against the x_{max} which represents the maximum number of possible transmission symbols. As can be seen the optimal transmission zone achieves a much higher information rate compared to the strip transmission zone as proposed by the analysis. In particular, there is about 0.5 bit per channel use improvement when time per channel use is equal to 1000 seconds, and about 1.5 bits per channel use improvement when the time per channel use is equal to 7000.

Next, the effects of size of the vesicle encapsulation on channel capacity is considered. The same four vesicle sizes used in Section 6.1.2 are used, as well as the same simulation environment and parameters. Assume only a single particle is encapsulated inside the lipid vesicles regardless of the vesicle's diameter. Through this assumption, the number fo vesicles transported are compared. Also, assume the set of possible transmission symbols are $\mathcal{X} = \{0, 1, 2, \cdots, x_{max}\}$, for some value of x_{max}, where a transmission symbol $X \in \mathcal{X}$ is represented by release of X information particles into the medium. All the released particles will be randomly distributed over the transmission zone (i.e. grid transmission zone) and remain

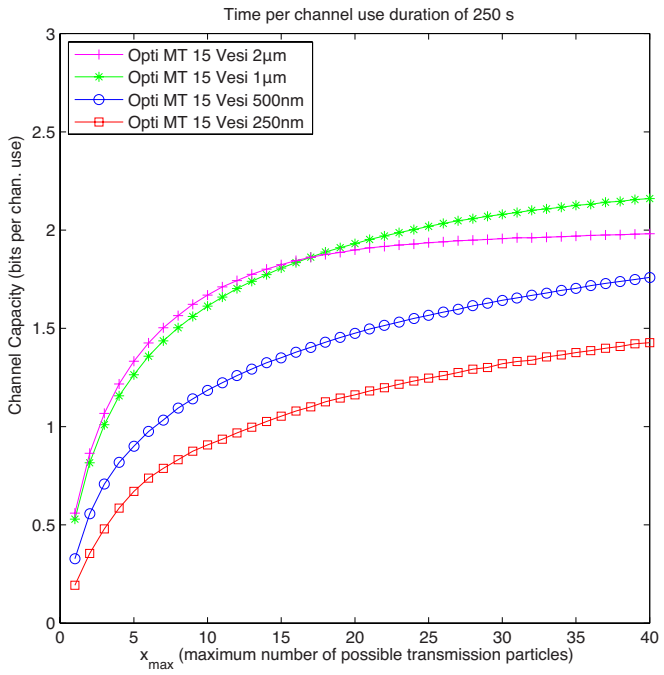


Figure 6.8: Channel capacity in bits per channel use for a single information particle encapsulated inside each lipid vesicle ($\tau = 250$s).

stationary until they are picked up by a microtubule for delivery. To make the comparison Monte Carlo simulations are used.

The results are shown in Figure 6.8 for time per channel use of 250 seconds. From the graph it can be seen that as the size of the vesicles increase from 250nm to 1μm the channel capacity of the microchannel increases. However, as the size increases from 1μm to 2μm the channel capacity does not increase. The effect is due to the difference in the shape of the transmission zone as shown in Figure 6.2. Because 2μm squares cover a larger area, they spread across the whole transmission

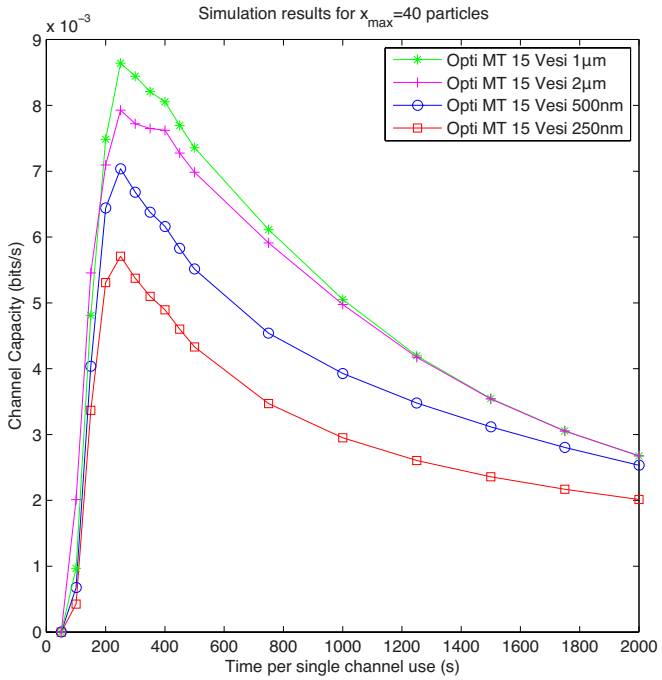


Figure 6.9: Channel capacity in bits per second versus time per channel use for a single information particle encapsulated inside each lipid vesicle.

area including the center parts where the probability of a microtubule visit is every low as shown in Figure 6.1.

In Figure 6.9, the channel capacity in bits per second is plotted against different time durations per channel use for $x_{max} = 40$. From the figure, again it can be observed that as the size of the vesicle is increased from 250nm to 1μm the curve's peak is increased. However, 1μm vesicles achieve a higher peak than the 2μm vesicles and hence have a higher channel capacity in bits per second. This is again because of the shape of the transmission area which in the case of 2μm vesicle is

not optimized.

Based on the results it is concluded that encapsulation of information vesicles using lipid vesicles improves the channel capacity. This effect is due to the small size of the information particles which are typically molecules. For example if no encapsulation is employed the size of a typical information particle would be 10nm. Therefore, based on the presented results it can be seen that 10nm squares in the grid loading area would have a much lower channel capacity compared to the larger squares of vesicular encapsulation. However, depending on the shape and the dimensions of the transmission area, there is an optimal vesicle size. This effect is particularly evident in the simulations as the vesicle size is increased from 1μm to 2μm. Therefore, it is concluded that the optimal vesicle size for this particular channel is 1μm, which is in accordance with the results in Section 6.1.2.

6.2 Channel Shape Optimization

In the previous section, a technique for finding the optimal transmission zone and the optimal vesicle size for transport was developed. In this sections, the optimal channel shape, among a large class of channel shapes is obtained. The shape of the channel is optimized with respect to the total number of information particles that could be transported during a time per channel use interval. Some of the models derived in Chapter 5 are used to formulate an optimization problem.

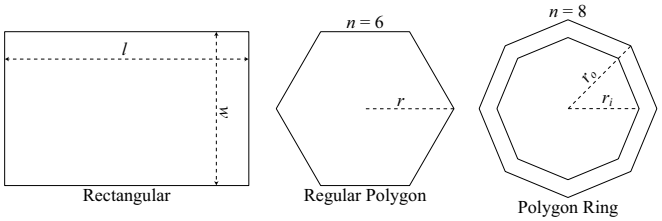


Figure 6.10: Different shape classes and their parameters.

6.2.1 Channel Shape Classes

Assume that channel shape can belong to any of the following three classes of shapes: rectangular, regular polygon, and regular polygon ring. Regular polygons, are equiangular (all angles are equal in measure) and equilateral (all sides have the same length), and they include a large class of geometric shapes, ranging from equilateral triangles, squares, pentagons, and hexagons, all the way up to a circle as number of sides approaches infinity. Figure 6.10 shows an example from each of the three different shape classes.

Depending on the class of the channel shape, different parameters can be used to further adjust the shape within the shape class. For rectangular channels two parameters can be used to adjust the dimensions and the shape: width w and length l. When both the width and the length are equal in value, the channel would be a square channel. Similarly, for regular polygons two parameters can be used to further adjust the shape: the number of sides n, and the radius of the circumscribed

circle r (this radius can also be defined as the distance from the center point of the regular polygon to one of the vertices of the polygon). For ring-shaped channels three parameters are used to define the shape: the number of sides n, the radius of the inner circumscribed circle r_i, and the radius of the outer circumscribed circle r_o. Figure 6.10 summarizes all these parameters for different classes of shapes.

Regardless of the shape of the channel, without loss of generality, assume the transmission zone is always on the left side of the channel and the receiver zone is along the right side of the channel. In the previous section, it was shown that optimal transmission zone is along the walls of the channel, since microtubules mostly glide very close to channel walls. This positioning increases the chance of a microtubule picking up an information particle during its trips. Therefore, it is always assumed that the transmission zone is along the walls of the channel regardless of the channel shape.

6.2.2 Channel Shape Optimization Model

Let the set $\mathcal{G}$ be the set of all possible cross-sectional geometric shapes the channel could have, where the cross-sectional shape is the shape of the channel when viewed from the top. First, it is assumed that this set contains all the rectangular, regular polygonal, and regular polygonal ring-shapes and then the result are extended to all two dimensional shapes. The goal is to find the optimal shape in these sets, given

system parameters such as time per channel use, concentration of microtubules, and average speed of the microtubules. To setup this optimization problem, the effects of channel shape on the number of microtubule trips is modeled.

As discussed in the previous chapter, the number of microtubule *trips* during the time per channel use interval has a direct effect on transportation of information particles. Let $g \in \mathcal{G}$ be the channel under consideration. Let $K_\tau(g)$ be the number of trips during the time per channel use interval τ inside the channel shape g, where a *single microtubule trip* is defined as the movement of the microtubule from anywhere in the channel to the transmission zone and then the receiver zone. As shown in Chapter 5, the number of microtubule trips $K_\tau(g)$ is a random variable. Therefore, to derive the optimization model, the expected value of $K_\tau(g)$ is used.

Let the random variable $K_\tau^{(S)}(g)$ be the number of trips for a single microtubule during the time per channel use interval, τ. Let v_{avg} be the average speed of the microtubules, $P(g)$ be the perimeter of the channel shape $g \in \mathcal{G}$. A good estimate for the average number of microtubule trips, when a single microtubule is inside the channel is given by

$$E[K_\tau^{(S)}(g)] \approx \frac{v_{\text{avg}}\tau}{P(g)}. \tag{6.5}$$

This approximation is based on the observation that on average from each trip, the microtubule travels a distance approximately equal to the perimeter of the

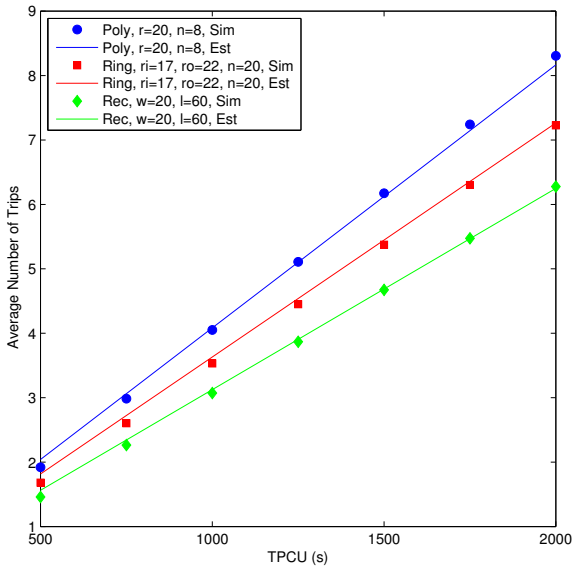


Figure 6.11: Average number of trips approximation compared with Monte Carlo simulations.

channel. This assumption is verified using the Monte Carlo simulation environment. The results of the simulations are shown in Figure 6.11. In this figure, a sample shape from each of three different shape classes have been considered, and it is shown that the approximations in Equation (6.5) (solid lines) are a good estimate of the actual average number of microtubule trips (point plots).

In practice, there are typically more than one microtubule inside the channel. Furthermore, the number of microtubules is not a constant, but is dependent on the volume of fluid in the channel: the DNA-covered microtubules are prepared in chemical solutions, and therefore have a constant concentration inside the solution.

Let C be the concentration of microtubules as number of microtubules per unit volume, and h be the height of the channel, and $A(g)$ be the cross sectional area of the channel with geometric shape $g \in \mathcal{G}$. Then the number of microtubules inside the channel is given by

$$M(g) = A(g)hC, \tag{6.6}$$

where $M(g)$ is the number of microtubules inside channel g.

For each microtubule in the channel let the random variables $K_\tau^{(i)}(g)$ be the number of trips during τ seconds inside the channel with shape g, where $i = 1, 2, \cdots, M$ indexes each microtubule. Then, the total number of microtubule trips during τ seconds is given by

$$K_\tau(g) = \sum_{i=1}^{M} K_\tau^{(i)}(g). \tag{6.7}$$

The average number of microtubule trips during τ seconds is therefore calculated as

$$E[K_\tau(g)] = \sum_{i=1}^{M} E[K_\tau^{(i)}(g)]. \tag{6.8}$$

It is assumed that the number of trips for individual microtubules are independent and identically distributed, because they are chemically similar and don't interact with each other. Generally, this assumption has been found to be valid in laboratory

experiments [67, 85]. Due to the identical distributions, the equation simplifies to

$$E[K_\tau(g)] = \sum_{i=1}^{M} E[K_\tau^{(S)}(g)] = M(g)E[K_\tau^{(S)}(g)]. \tag{6.9}$$

Using the approximation shown in Equation (6.5), and Equation (6.6), the total number of trips can be estimated as

$$E[K_\tau(g)] \approx \tau v_{\text{avg}} Ch \frac{A(g)}{P(g)}, \tag{6.10}$$

where $A(g)$ and $P(g)$ are the cross sectional area and perimeter of the channel with geometric shape $g \in \mathcal{G}$, respectively. The channel shape optimization problem can then be formulated as

$$\max_{g \in \mathcal{G}} E[K_\tau(g)]. \tag{6.11}$$

Assuming all the other parameters are constant (including the height of the channel) except the cross sectional shape, the optimization problem becomes

$$\max_{g \in \mathcal{G}} \left[\frac{A(g)}{P(g)} \right]. \tag{6.12}$$

Equation (6.12) is very interesting because it states that the optimal shape is the one with the largest area to perimeter ratio. However, there is an important con-

straint that must be satisfied. The perimeter of the channel must be small enough such that a single microtubule can complete at least a single trip. Therefore, assume the perimeter must be such that its length can be travelled by the microtubule, on average once, during the time per channel use interval τ. This constraint ensures that the perimeter is small enough such that on average enough information particles can be delivered during the given time per channel use duration. Finally, the optimization problem can be written in its complete form as

$$\max_{g \in \mathcal{G}} \left[\frac{A(g)}{P(g)} \right]$$
$$\text{subject to } P(g) \leq \tau v_{\text{avg}}. \tag{6.13}$$

6.2.3 Optimal Shape Analysis

Using the optimization formula derived in the previous section, the optimal channel design for each shape class is considered individually, and then the optimal overall design is presented.

6.2.3.1 Rectangular Channels

Let $\mathcal{G}_{\text{Rec}}$ be the set of all rectangular shapes. Because any rectangular shape can be characterized by the two parameters width w and length l, the Equation (6.13)

becomes

$$\max_{(w,l)\in\mathcal{G}_{\text{Rec}}} \left\lceil \frac{w \times l}{2w + 2l} \right\rceil$$
$$\text{subject to } 2w + 2l \leq \tau v_{\text{avg}}. \tag{6.14}$$

Solving this optimization problem, the optimal channel design given both time per channel use interval τ and average speed of the microtubules v_{avg}, is $w = l = 0.25\tau v_{\text{avg}}$. Therefore for rectangular channels, the optimal channel shape is always the square-shaped channel.

6.2.3.2 Regular Polygonal Channels

Let $\mathcal{G}_{\text{Poly}}$ be the set of all regular polygons. Regular polygons can be characterized by two parameters: the number of sides n, and the radius of the circumscribed circle r. Note that the set of all square shapes are also contained within the set $\mathcal{G}_{\text{Poly}}$ (i.e. when $n = 4$). Using these parameters Equation (6.13) becomes

$$\max_{(n,r)\in\mathcal{G}_{\text{Poly}}} \left\lceil \frac{0.5nr^2 \sin(2\pi/n)}{2nr \sin(\pi/n)} \right\rceil$$
$$\text{subject to } 2nr \sin(\pi/n) \leq \tau v_{\text{avg}}. \tag{6.15}$$

This equation can be simplified as

$$\max_{(n,r)\in\mathcal{G}_{\text{Poly}}} 0.5r\cos(\pi/n)$$
$$\text{subject to } nr\sin(\pi/n) \leq \frac{\tau v_{\text{avg}}}{2}, \tag{6.16}$$

where the fact that $\sin(2u) = 2\sin(u)\cos(u)$ is used. Based on this equation the optimal channel is the one with $n = \infty$ and $r = \tau v_{\text{avg}}/2\pi$. Therefore, when regular polygons are considered, the optimal shape is the circular-shaped channels. Because squares are also in the set of all regular polygons, it is also concluded that circular-shaped channels are also better than square-shaped channels.

6.2.3.3 Regular Polygonal Ring Channels

Let $\mathcal{G}_{\text{Ring}}$ be the set of all polygonal ring shapes. The elements of this set can be characterized using three parameters: the number sides n, the radius of the inner polygon's circumscribed circle r_i, and the radius of the outer polygon's circumscribed circle r_o. Note that $\mathcal{G}_{\text{Ring}} \supset \mathcal{G}_{\text{Poly}}$, which means that the set $\mathcal{G}_{\text{Ring}}$ contains all the regular polygonal shapes (this is achieved by setting $r_i = 0$). Using these

parameters the optimization problem becomes

$$\max_{(n,r_i,r_o)\in\mathcal{G}_{\text{Ring}}} \left[\frac{0.5n(r_o^2 - r_i^2)\sin(2\pi/n)}{2nr_o\sin(\pi/n)}\right]$$
$$\text{subject to } 2nr_o\sin(\pi/n) \leq \tau v_{\text{avg}}. \tag{6.17}$$

This equation can be simplified

$$\max_{(n,r_i,r_o)\in\mathcal{G}_{\text{Ring}}} \left[0.5\left(\frac{r_o^2 - r_i^2}{r_o}\right)\cos(\pi/n)\right]$$
$$\text{subject to } nr_o\sin(\pi/n) \leq \frac{\tau v_{\text{avg}}}{2}. \tag{6.18}$$

Solving Equation 6.18, the optimal channel has parameters $n = \infty$, $r_o = \tau v_{\text{avg}}/2\pi$, and $r_i = 0$. This means the optimal channel is the circular-shaped channel. This may seem surprising at first, since one would expect that a ring-shaped circular channel would be better than a solid circular-shaped channel at transporting information particles. However, because typically microtubules follow the walls of the channel, and that the number of microtubules in a channel is proportional to the volume of the channel [256], it becomes apparent that the solid circular channel would be better than a ring-shaped channel.

6.2.3.4 Overall Optimal Channel Design

From the solution to the optimization formulas for each of the three different shape classes, it is apparent that the optimal channel shape is the circular-shaped channel. Moreover, based on Equation (6.13) the optimal channel shape must have the largest area to perimeter ratio. Therefore, if the perimeter of the channel is fixed such that the constraint in the optimization is satisfied with an equality, the optimal channel shape would be a circle for all two dimensional geometric shapes.

6.2.4 Simulation Validation

To verify the optimization formulas and their results, Monte Carlo simulation are used. For these simulations the same parameters used in Chapter 4 and 5 are used. Moreover, it is assumed that the height of the channel is always a constant $h = 10$ μm regardless of the cross sectional shape of the channel. The concentration of the microtubules is also assumed to be $C = 0.001$ MT/fL unless specified otherwise. The number of microtubule in the channel is always assumed to be

$$MT = \lfloor A(g)hC \rfloor, \tag{6.19}$$

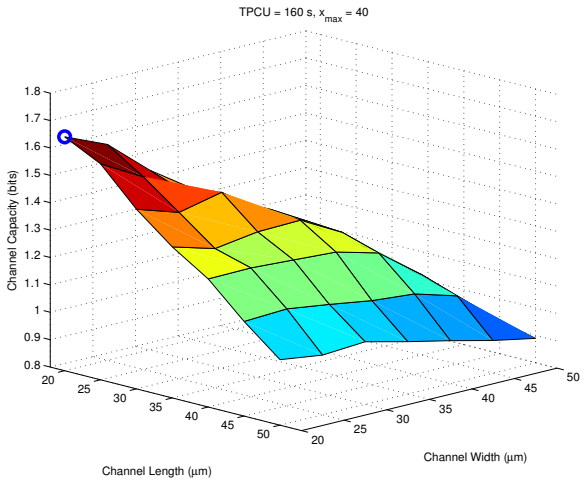
unless specified otherwise.

For the performance measure *channel capacity* [192] is used, which is the max-

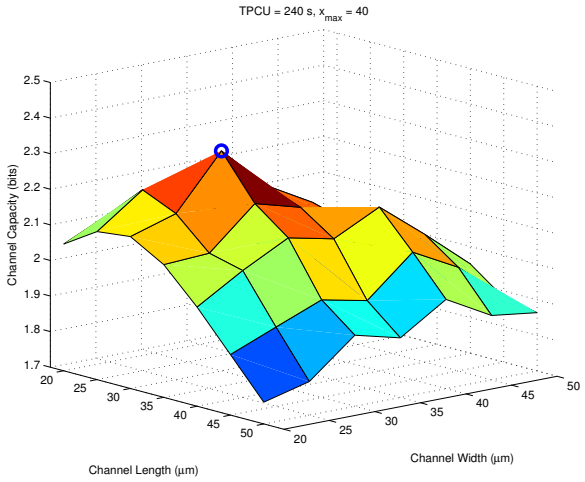
imum rate at which information can be transmitted reliably. Each of three shape classes discussed in the previous sections are considered, and Equations (6.14), (6.16), and (6.18) are used to find the optimal channel dimensions for each case. These results are then compared to the channel capacities obtained from Monte Carlo simulations.

6.2.4.1 Rectangular Channels

First, rectangular channels are discussed. Two values of time per channel use are considered: $\tau = 160$ seconds, and $\tau = 240$ seconds. Using the solution to Equation (6.14), the optimal channels are square channels with dimensions $w = l = 0.25 \times 160 \times 0.5 = 20$ μm and $w = l = 0.25 \times 240 \times 0.5 = 30$ μm for the $\tau = 160$ and $\tau = 240$, respectively. To verify this result, all the rectangular channels with widths and lengths raging from 20 μm to 50 μm are simulated in 5 μm steps. It is assumed that $x_{\max}$ (i.e. the maximum number of particles the transmitter can release) is 40. The channel capacity is then calculated based on the simulations. Figure 6.12 shows the results for $\tau = 160$ seconds (a), and $\tau = 240$ seconds (b). As can be seen the channels with the largest channel capacity are the 20 μm × 20 μm, and 30 μm × 30 μm, respectively. These results are in perfect agreement with the results obtained from the optimization model.



(a)



(b)

Figure 6.12: Channel capacities in bits per channel use of different rectangular channel shapes. The channel with the highest capacity is shown using the blue dot: (a) for the time per channel use of 160 s the optimal is 20 μm × 20 μm , (b) and for the time per channel use value of 240 s the optimal size is 30 μm × 30 μm.

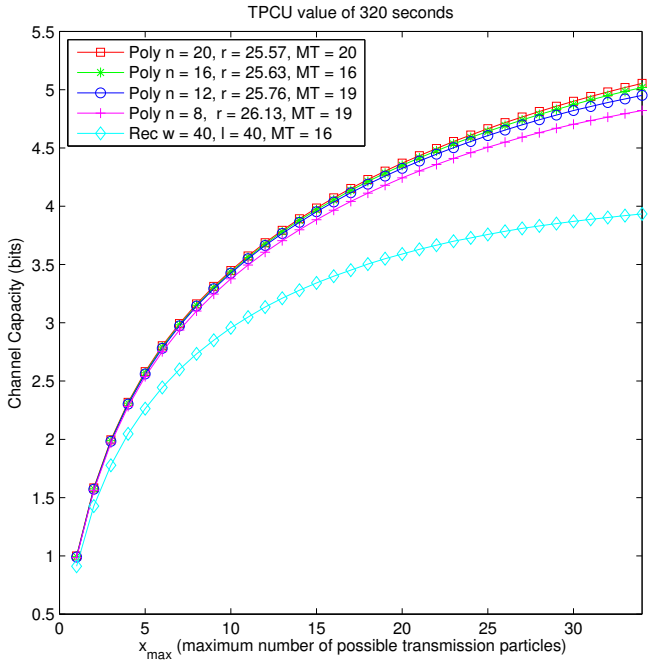


Figure 6.13: Channel capacity of regular polygonal channels with constant perimeter of 160 μm versus maximum number of particles that can be released at the transmitter.

6.2.4.2 Regular Polygonal Channels

In Section 6.2.3, it was shown that for regular polygonal channels the optimal shape is a circular-shaped channel. To verify this result, consider a constant time per channel use value of 320 seconds. From Equation (6.13), the maximum channel perimeter that satisfies the constraint is 160 μm. Therefore, it is assumed that the channel perimeter is fixed at this value and it is shown that the capacity increases as the channel shape become more circular by increasing the number of sides n.

In particular, the following channels are considered: square channel of length 40 μm (equivalent to a polygon with $n = 4$), and regular polygonal channels with parameters ($n = 8$, $r = 26.13$ μm), ($n = 12$, $r = 25.76$ μm), ($n = 16$, $r = 25.63$ μm), and ($n = 20$, $r = 25.57$ μm). Based on the optimization formula, it is expected for the 20-sided channel to be the optimal among all these channels since it is more circular. Figure 6.13 shows the channel capacity for each of these channels obtained through Monte Carlo simulations versus the maximum number of particles the transmitter can release $x_{\max}$. As can be seen, the optimal channel is indeed the 20-sided regular polygon. Moreover, from the obtained pattern it is obvious that as the number of sides increase the channel capacity increases. This supports the results obtained from the optimization formula that circular channels are optimal.

To evaluate the performance of Equation (6.16) consider the time per channel use interval of $\tau = 250$ seconds. The solution of this optimization problem is a channel with parameters $n = \infty$ and $r \approx 20$ μm. Because in the simulation environment the number of sides must be a finite number, and therefore the value of $n = 20$ is used in place of a perfect circle. Four different regular polygonal channels are simulated with parameters: ($n = 20$, $r = 17$ μm), ($n = 20$, $r = 20$ μm), ($n = 20$, $r = 22.75$ μm), and ($n = 20$, $r = 25.57$ μm). Figure 6.14 shows the channel capacity of each channel versus the maximum number of particles the transmitter could release. As can be seen in the figure, the channel with the radius

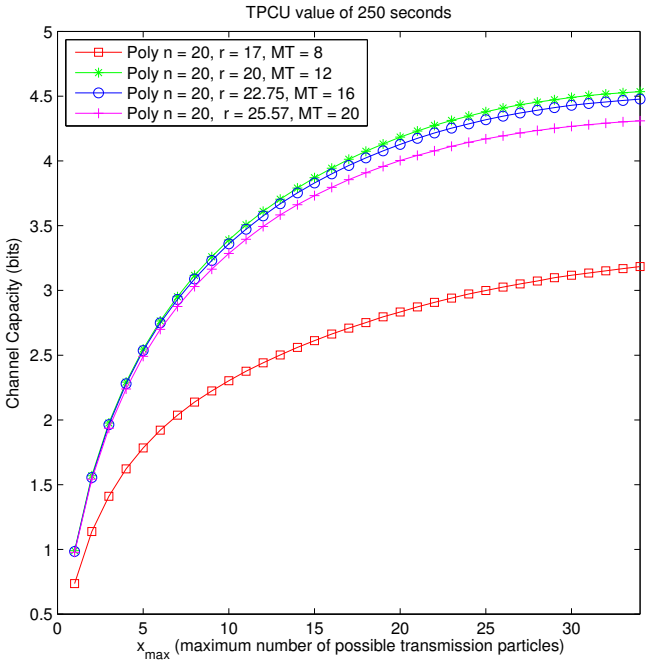


Figure 6.14: Channel capacity of four circular-shaped channels versus maximum number of particles that can be released at the transmitter.

of 20 μm achieves the highest capacity among the four channels. This result is in perfect agreement with the optimal solution derived from the optimization formula.

6.2.4.3 Regular Polygonal Ring Channels

In this section, the optimization formula is validated through simulations for polygonal ring-shaped channels presented in Equation (6.18). It is assumed that the time per channel use interval is fixed at 320 seconds. Using the solution to Equation

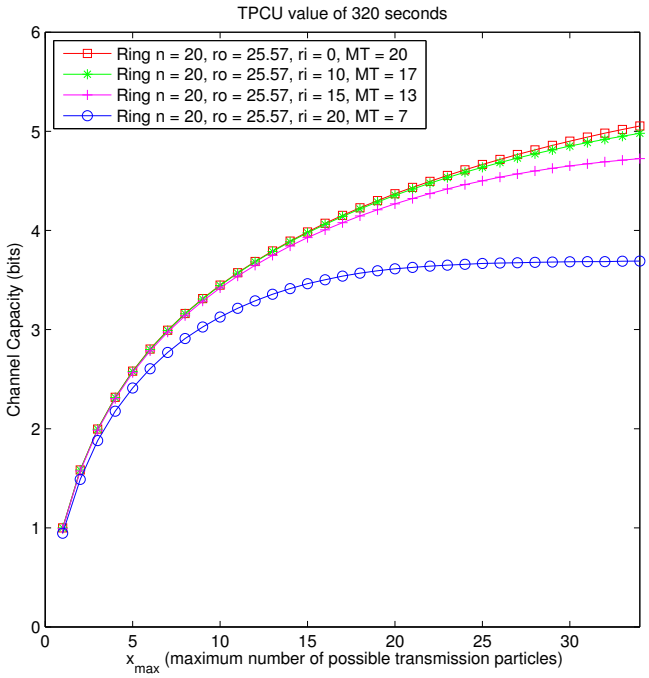


Figure 6.15: Channel capacity of ring-shaped channels versus maximum number of particles that can be released at the transmitter.

(6.18), the optimal channel dimensions are $n = \infty$, $r_i = 0$, and $r_o \approx 25.57$ μm. In Figure 6.13, it was already shown that the solid circular-shaped channel is optimal among regular polygons. Therefore, only circular ring-shaped channels are considered for validation. In particular, the following ring-shaped channels are considered: ($n = 20$, $r_o = 25.57$ μm, $r_i = 0$ μm), ($n = 20$, $r_o = 25.57$ μm, $r_i = 10$ μm), ($n = 20$, $r_o = 25.57$ μm, $r_i = 15$ μm), and ($n = 20$, $r_o = 25.57$ μm, $r_i = 20$ μm). Figure 6.15 shows the channel capacity of these channels versus the maximum number of particles the transmitter could release. As can be seen from the figure, a solid circular

channel achieves the highest capacity as predicted by the optimization formula.

Based on all these results, it can be concluded that a solid circular-shaped channel is the optimal shape that maximizes information rate, when stationary kinesin with mobile microtubule filaments are used for transportation in molecular communication systems. Moreover, as the number of sides increase and the channel becomes closer to circular, the performance gains of adding more sides decreases. Based on the obtained results, there is negligible performance difference between channels with the number of sides greater than 20. Therefore, these channels are sufficient representation of perfectly circular-shaped channels.

7 Tabletop Experimental Platform for Molecular Communication

In the previous three chapters, on molecular communication for on-chip applications were discussed. As was highlighted by the comprehensive survey in Chapters 2 and 3, however, another very important problem in molecular communication is the wide gap between theory and practice. Most previous works have focused on theoretical aspects of molecular communication. This is because of the expensive, laborious, and multidisciplinary nature of the laboratory experimentation. To overcome this wide gap, in this chapter the world's first tabletop experimental molecular communication platform that is capable of transporting sequential data such as text messages is developed. This system is purposefully designed to be built with inexpensive and readily available off-the-shelf parts, such that other researchers could replicate the platform with ease. Using this testbed, it is demonstrated that reliable communication is possible using chemical signaling. This platform is different from the on-chip molecular communication channels considered in the previous chap-

ters, and is not intended to be an experimental demonstrator for those particular channels.

7.1 Materials and Methods

In this section, a simple, robust, and cost-effective communication system is designed and implemented that uses chemical signals for carrying information from a transmitter to a receiver. To test this system, it is used to send a short text message: this is a familiar application, as billions of Short Message Service (SMS) text messages are sent daily by mobile users [257]. To design and develop the system, the following criteria are considered:

- The end product must be inexpensive to build. This would make the platform readily available for many different research and development projects with limited amount of funding.

- The designed system must be simple and robust, much the same as the telegraph, the ancestor of modern telecommunication systems. The simple and robust design would help in the adoption of the platform in different applications.

- The developed system must be easily modifiable and programmable. Again this is an important criterion for future expansions and adoption to different

applications.

As discussed in Chapter 2, any communication system can be broken down into three major parts: the transmitter, the receiver, and the channel. Over the next three subsections, the design for the transmitter, the receiver, and channel propagation are presented.

7.1.1 Transmitter Design

The transmitter takes an input text message from a user. It then converts the text message into a sequence of binary bits and modulates them on a chemical signal for propagation in the channel. To control all transmission operations, the Arduino Uno open-source electronics prototyping platform is used, which is an ATmega328 based microcontroller board. For text entry, the 16x2 character LCD Shield Kit from Adafruit is used. The LCD is an add-on module for the Arduino microcontroller board, which also has six push buttons. A program for the Arduino microcontroller is developed which employs the LCD and its buttons for text entry by the user.

To convert the text message to a binary sequence, the International Telegraph Alphabet No. 2 (ITA2) standard is used [258], where every letter is represented using five bits. For example, the letter "E" is represented by a five bit sequence "10000". Table 7.1 shows how each letter of the alphabet is encoded using a 5-bit

Table 7.1: International Telegraph Alphabet No. 2 (ITA2) source encoder.

Letter	**5-bit Binary Representation**	**Letter**	**5-bit Binary Representation**
A	11000	O	00011
B	10011	P	01101
C	01110	Q	11101
D	10010	R	01010
E	10000	S	10100
F	10110	T	00001
G	01011	U	11100
H	00101	V	01111
I	01100	W	11001
J	11010	X	10111
K	11110	Y	10101
L	01001	Z	10001
M	00111	‘ ’ (space)	00100
N	00110	null	00000

sequence. For simplicity, no error-correcting codes are used. Therefore, the five bit encoded letters are passed to the modulator block of the transmitter for modulation and transmission to the channel.

To modulate the channel symbols into chemical signals, an electronic spray called DuroBlast made by Durotech Industries is used. The DuroBlast electronic spray has a battery operated electrical pump that can spray a wide variety of liquid chemicals that can be stored inside its container. A custom electrical switch board is designed that can be used to control the spray from the Arduino microcontroller board. The circuit diagram for the switch board is shown in Figure 7.1. The parts used in the circuit include an 2K ohm resistor, an 1N4004 diaod, and an IRLZ44ZPBF MOSFET. The 5V port is connected to the positive port of the

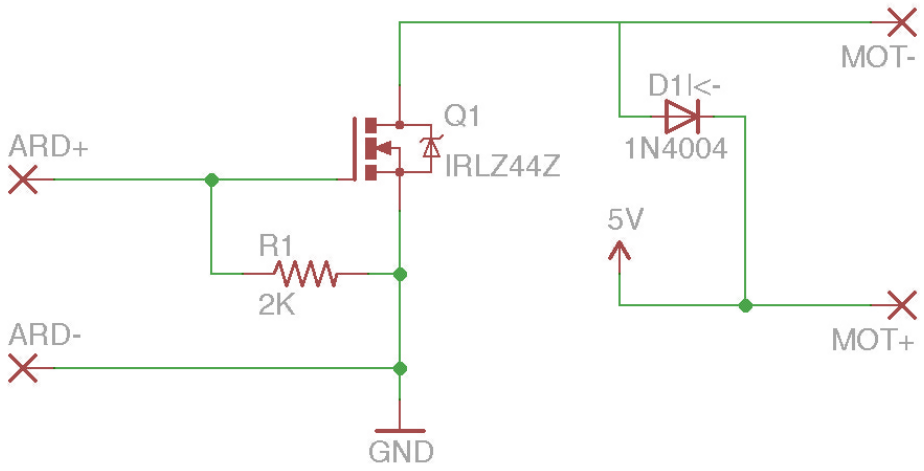


Figure 7.1: The circuit diagram for the custom electrical switch board for the spray. The parts include an 2K ohm resistor, an 1N4004 diaod, and an IRLZ44ZPBF MOSFET. The 5V port is connected to the positive port of the battery in the spray, and the GND port is connected to the negative port of the battery in the spray.

battery in the spray, and the GND port is connected to the negative port of the battery in the spray. By programming the Arduino microcontroller board, any type of modulation can be implemented through controlled set of sprays. Figure 7.2 shows the transmitter setup with all of its subcomponents.

7.1.2 Receiver Design

To design the receiver, a sensor is required that is capable of detecting a chemical signal. The data from the sensor is processed by the demodulation and detection algorithms, and finally decoded into text. Again, the Arduino Uno open-source

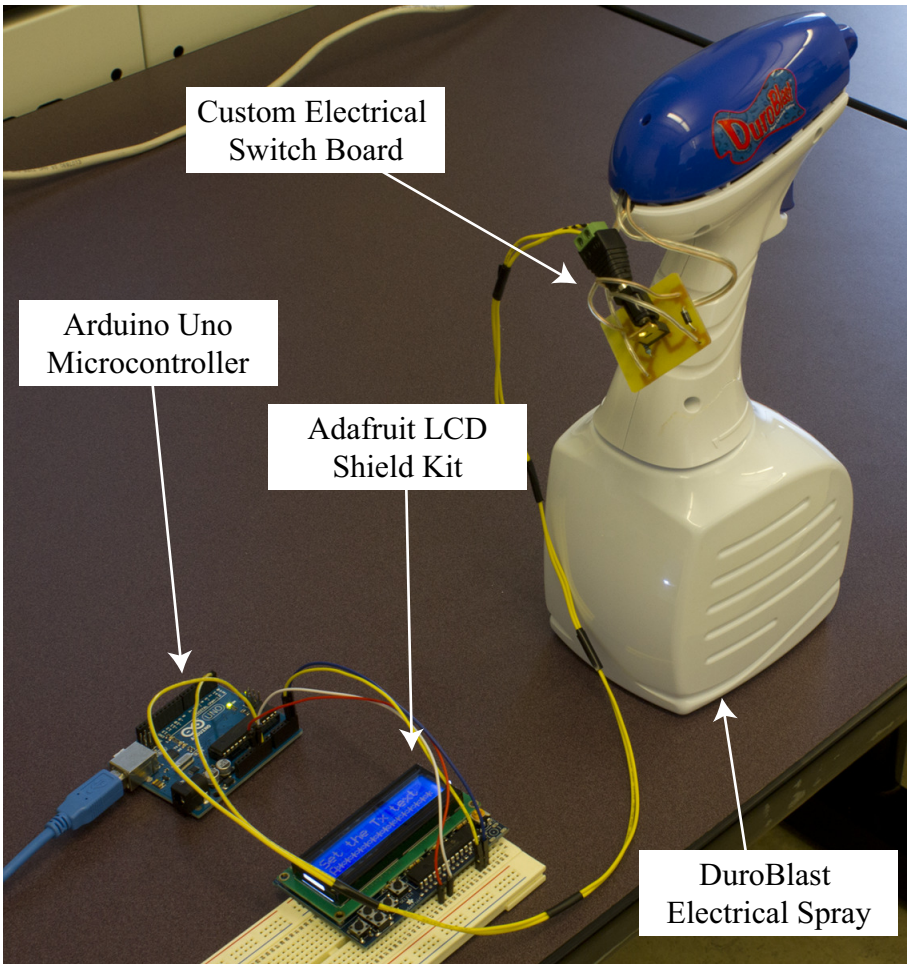


Figure 7.2: The transmitter components.

microcontroller is used for programming and controlling all the receiver operations. The Arduino Uno board has a 10-bit analog to digital converter that can be used to read the sensor data. The demodulation and detection block and the source decoder block can then be programmed into the microcontroller, and the resulting detected text message can be displayed on a computer screen using serial port.

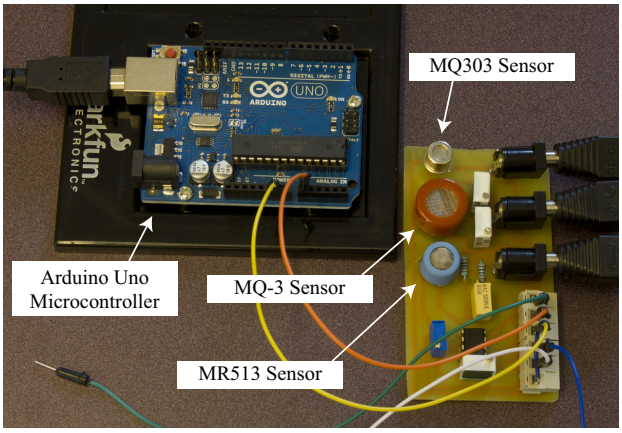


Figure 7.3: The receiver components.

To achieve the design criteria, the receiver's sensor must be sensitive, widely available, and inexpensive. Moreover, it must be able to detect a volatile, widely available, and inexpensive signaling chemical that is safe at the low concentrations that are use. Therefore, isopropyl alcohol (rubbing alcohol) is chosen as the signaling chemical with three different candidate sensors for demodulation and detection at the receiver: MQ-3, MQ303A, MR513 alcohol sensors, all of which are manufactured by Henan Hanwei Electronics Co. Ltd. of China. All three sensors use a metal oxide semiconductor detection layer [160] for detecting the alcohol, but each has a different sensitivity, power and operation circuit diagrams. Detailed information about each sensor and its respected circuit can be found from the corresponding sensor data sheet [259–261]. Besides isopropyl alcohol, the sensors can detect other types of alcohol such as ethanol, which have been validated experimentally. How-

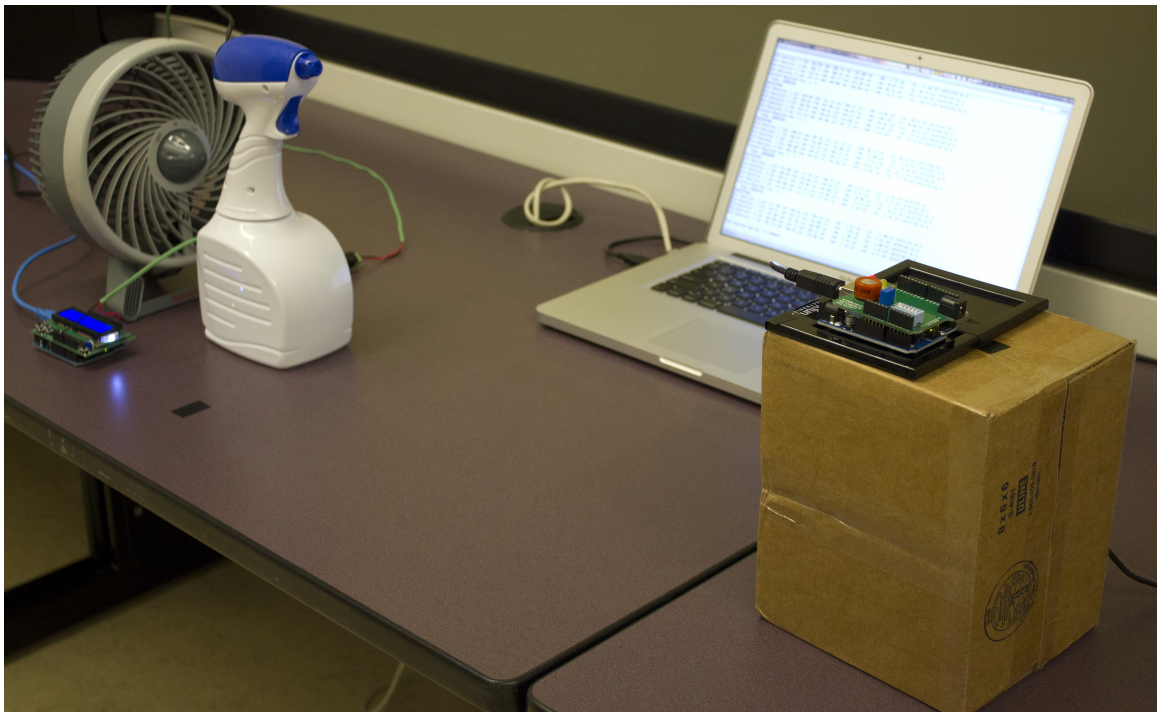


Figure 7.4: The system orientation.

ever, in this dissertation only used isopropyl is used for experimentation. All three sensors are implemented on a custom-made PCB board as shown in Figure 7.3.

7.1.3 Propagation Channel

Two different propagation schemes are considered for the channel: diffusion, and flow assisted propagation. For all experiments performed throughout this chapter, the transmitter and the receiver are oriented in such a way that the receiver is right in front of the transmitter. Moreover, the tip of the spray and the sensor have

the same approximate height. This is demonstrated in Figure 7.4. In the diffusion based propagation, after the initial spray the alcohol diffuses in the air until it reaches the receiver. In flow based propagation, a tabletop fan is used to guide the alcohol towards the receiver. Therefore, the diffusion propagation does not require external energy (beyond the energy required to release the chemical message), while the flow assisted propagation requires external power. Two different tabletop fans are used to generate flow:

- Honeywell 7 inch Personal Tech fan: This fan is an inexpensive bladed fan (approximately $16 USD) with two different fan speeds (low and high).
- Dyson AM01 10 inch bladeless fan: The Dyson fan is much more expensive (approximately $250 USD), but can generate more laminar flows and many different wind speeds by adjusting an analog nub.

When any of the two fans are used, they are placed 30 cm behind the spray.

To measure the flow speeds generated by each fan, the Pyle PMA82 digital anemometer is employed. The maximum flow speed is measured at distances of 10 cm, 50 cm, 100 cm, 150 cm, and 200 cm from the front of the spray (the fan is placed 30 cm behind the spray). For the Dyson fan 5 different nub positions are selected and these positions are labeled as very high, high, medium, low, and very low. Figure 7.5 shows the wind speed for each fan at each distance. Because there

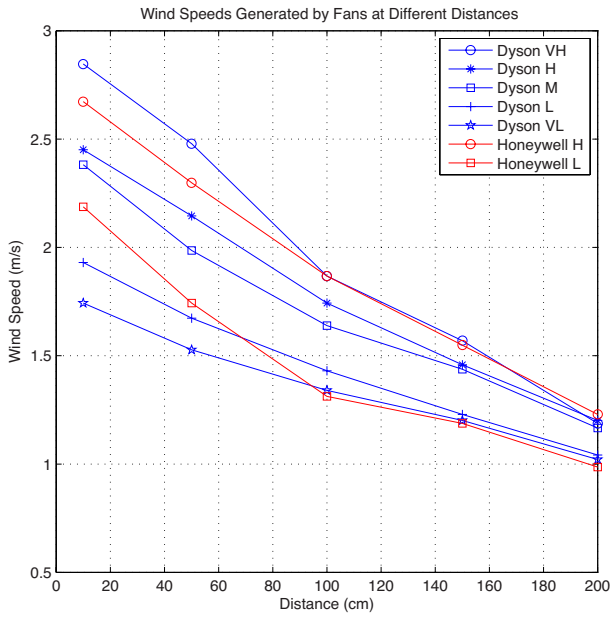


Figure 7.5: Wind speeds generated by each fan. The Dyson fan plots are in blue, and Honeywell plots are in red.

is $+/-3\%$ error associated with the digital anemometer, four different measurements are averaged to produce each plot point in Figure 7.5. Moreover, the digital anemometer is not rated for flow speeds below 1 m/s. Therefore, flow speeds below this range are not shown. The average flow velocities achieved over this distance is tabulated in Table 7.2 for each fan setting.

The system response (the output of the sensor for a single short spray between 100-250 ms) is compared under both propagation schemes (diffusion and flow assisted propagations). At short distances (up to 1 meter), the diffusion based propagation scheme performs well because the alcohol ejected from the spray reaches

Table 7.2: Average flow velocities. The average flow velocities over the distance of 200 centimeters generated using the Dyson and Honeywell fans.

Flow Generated By	**Average Flow Speed (m/s)**
Dyson on very high	1.99
Dyson on high	1.80
Dyson on medium	1.72
Dyson on low	1.46
Dyson on very low	1.37
Honeywell on high	1.92
Honeywell on low	1.48

the sensors almost instantaneously. However, if the spray is placed further away, diffusion based propagation would not be practical because of the extremely slow system response. This effect can be seen in Figure 7.6, where the system response to a very short spray of 250 ms in duration is plotted for both diffusion based and flow based propagations. The flow in this figure is generated using the Honeywell fan on the high setting, and the spray is placed at a distance of 2 meters from the detection sensor. As can be seen, the system has a quick and distinct response when flow based propagation is employed. Although the response is plotted for only one of the sensors (MQ-3 sensor), the same effect was observed for all the other sensors,

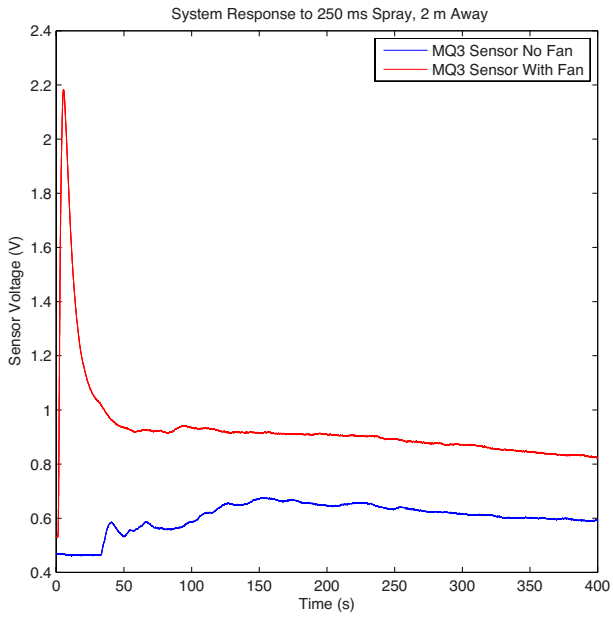


Figure 7.6: Diffusion based propagation versus flow based propagation. The system response to a 250 ms spray 2 meters away for diffusion based propagation (blue), and flow based propagation (red).

as well as when the Dyson fan is used in place of the Honeywell fan. Therefore, for the molecular communication setup flow based propagation is used.

7.1.4 Signal Modulation and Demodulation

Because communication is performed through chemical signals, and only a limited amount of signaling chemical can be stored in a container at the receiver, the modulation and demodulation scheme selected should minimize the amount of chemical used. The source coding scheme for encoding text messages, presented in previous

sections, uses the least amount of ones in the 5 bit sequence for characters that have a higher rate of occurrence in the English text. For example, letters "E" and "T" both have a single one in their 5 bit sequences. Therefore, the bit 1 is modulated using a single spray and the bit 0 is modulated with no spray. This modulation scheme, which is called on-off keying, effectively minimizes the amount of chemical used for communicating English text.

At the receiver the demodulation is performed by measuring the rate of change in concentration. If during a single bit's communication session the voltage reading from one of the sensors is increasing (i.e. the concentration of the chemical signal is increasing), then the signal is demodulated as the bit 1. Similarly, if the voltage reading from one of the sensors is decreasing (i.e. the concentration of the chemical signal is decreasing) the signal is demodulated as the bit 0. More details regarding the detection and demodulation process is provided later in the chapter.

7.1.5 Communication Protocol Design

In this section, the communication protocol between the transmitter and the receiver, and its implementation is discussed. In designing the protocol, the following criteria are used: the protocol must be simple, asynchronous (i.e. no synchronization is required between the transmitter and the receiver), and should work independent of the separation distance between the transmitter and the receiver (i.e. it

should not only work for a predefined fixed distance between the transmitter and the receiver).

At the transmitter, the output of the source encoder (i.e. the bit sequence representing the text message) is concatenated with a two bit sequence "10" at the beginning and the null character represented by "00000" at the end. The initial "10" indicates start of a text message, and the null character indicates the end of the text message. For example, if the text message that is being transmitted is the letter "A", the output of the source encoder is the five bit sequence "11000" (where the left most bit position is the first bit position), and transmission bit sequence is "101100000000". The transmission bit sequence is then modulated using the scheme discussed in the previous section, where 1 is modulated with a spray and 0 with no spray. Figure 7.7 represents the flowchart of the algorithm that runs at the transmitter, and summarizes this process.

At the receiver, there are two states: the *wait state*, and the *reception state*. In the wait state, the receiver uses its sensor to continuously monitor the concentration of alcohol. If there is a sudden increase in the concentration of alcohol (i.e. sudden increase in the sensor's voltage output), the receiver switches to the reception state. This sudden change is caused by the initial "10" bit sequence concatenated to the beginning of every text message sent by the transmitter. This sudden change can also be used as the reference time for synchronizing each bit interval for all the bits

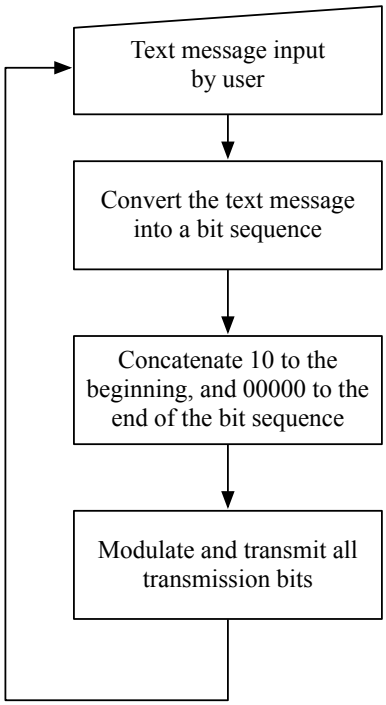


Figure 7.7: Flowchart representation of the algorithm that controls the transmitter.

that would follow. Therefore, no synchronization is required between the transmitter and the receiver in advance. Another factor that is taken into account in this scheme is the propagation delay. Because the receiver is triggered into the reception state as soon as the leading bit 1 is detected, the time delay caused by signal's propagation over the separation distance from the transmitter to the receiver is incorporated in the reference time. Therefore, the communication protocol is independent of the distance between the transmitter and the receiver, and it would work even when the distance is changed between communication sessions.

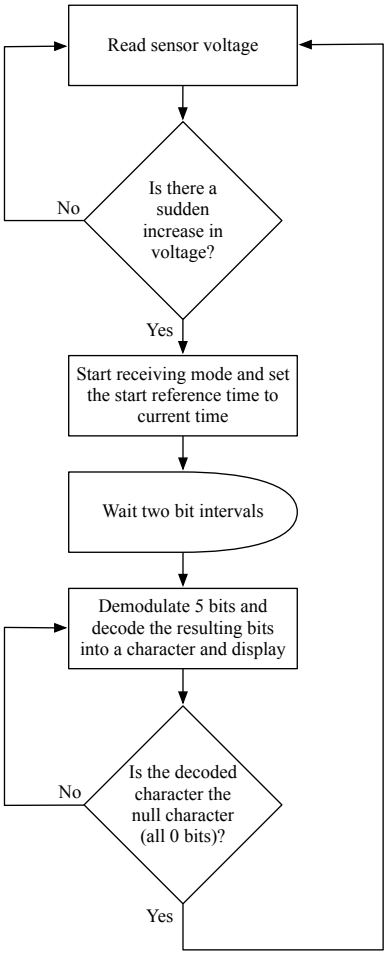


Figure 7.8: Flowchart representation of the algorithm that controls the receiver.

After the receiver enters the reception state, it waits for two bit intervals until the reception of the initial "10" bit sequence is finished. The receiver then demodulates and decodes the received signal 5 bits at a time. During each 5 bit interval, the source decoded character is displayed to the computer screen using serial port connection. This process continues until the null character represented

by all zero sequence "00000" is detected. Because the null character indicates the end of the text message, the receiver will go back to the wait state until another text message is sent by the transmitter. Figure 7.8 summarizes the algorithm that is implemented at the receiver.

7.2 Results and Discussions

In this section, first the impulse response of the overall system is presented and discussed, and then effects of different types of flow on the overall response is considered. Based on these results, the most suitable sensor is chosen to be used for the communication system. Different system parameters such as transmission rate, and the demodulation/detection algorithm are then fine tuned. Finally, some of the obtained results are presented and discussed.

7.2.1 Overall System Response

The overall system's impulse response is measured by using a very short spray (typically between 50-250 ms in duration) that resembles the delta function from signal processing. The terms system response and system's impulse response are used interchangeably. In the rest of this section, preliminary observations of the system response are presented to design the text messaging system, and in the next chapter the system response is modeled in more details.

Many parameters can effect the overall system impulse response. The most notable factors that have a major effect are:

- The sensor: Each sensor has its own response to a changing concentration. Three different sensors are used and the one that has the best overall response is chosen.

- The fan (flow type): Each fan has its own flow signature. An inexpensive bladed fan as well as a bladeless fan are used to generate different types of flow at different flow velocities. The Dyson fan can produce a more laminar flow at various velocities.

- The spray: Although the spray is electronically controlled with precise electrical signals, there are differences in the amount of particles that are released during each trial, and the size of the droplets in each spray stream. It is very difficult to precisely control these parameters within an inexpensive apparatus. Therefore, instead of precisely controlling the amount of chemical released by the spray, the overall response of the system is measured using a very short spray burst. By using similar burst durations, the amount of chemicals released is loosely controlled across different experimental trials.

There are other factors that could potentially affect the overall system response such as other flow patterns within the room, room temperature, and humidity. To lessen

the effects of these parameters, all the experiments are performed in a closed room with loosely regulated temperature and humidity. These precautions are enough, because other factors have a much greater effect on the overall response.

7.2.1.1 The Effects of the Sensor

To study the effects of the sensor on the overall response, the spray duration is set at 100 ms (i.e. the spray is switched on for 100 ms), and the system response is measured using each of the three sensors at various separation distances between the transmitter and the receiver. This scheme is used as it would be difficult to control and measure the actual volume of alcohol released during each burst. Figure 7.9 shows the system response for 2 meter separation distance (7.9a), and 4 meter separation distance (7.9b) for all three sensors. The Honeywell fan on the high setting is used to produce the flows for all these plots. As expected the amplitude of the peak decreases and the delay before the peak increases as the separation distance is doubled. The peaks also become wider and more flat as the separation distance increases. Similar effects are also observed when the Dyson fan is used to generate the flow.

From the overall system response, it is evident that there is a large amount of noise in the MR513's signal because of the operational amplifier used as part of its circuitry. Although the MQ303 has a high peak at 2 meters, the peak's full width

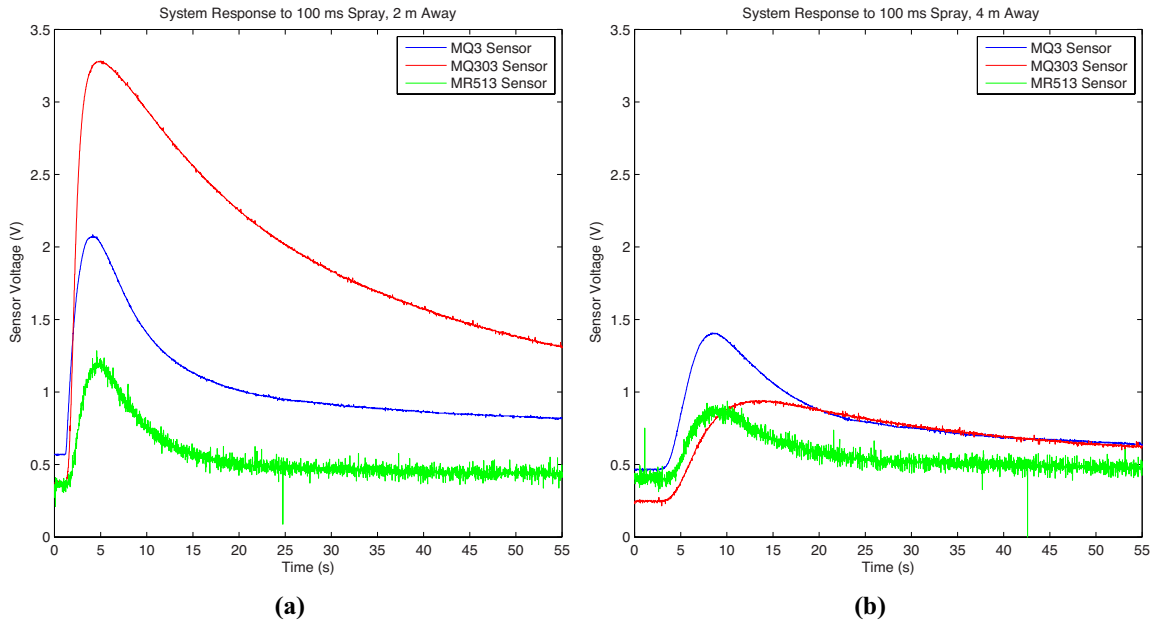


Figure 7.9: Comparing the three sensors. The system response to a short spray of 100 ms for all three sensors at (a) 2 m separation, and (b) 4 m separation between the transmitter and the receiver.

at half max (the width of the peak at half the value of its maximum shown in Figure 7.12) is much larger than the other two sensors. Ideally, this width must be as small as possible. Moreover, the height of the MQ303's peak drops significantly at 4 meters. The MQ-3 sensor has low noise and better system response over wider range of separation distances. Moreover, the MQ-3 has the simplest circuitry and can draw power directly from the Arduino microcontroller board. Therefore, the MQ-3 sensor is selected for the final implementation.

One of the major caveats of metal oxide gas sensors, including all three sensors

used, is the delays in responding to a changing concentration [160]. These delays core categorized as:

- The sensor's *response time*, i.e., the time it takes for the sensor to respond to a change in concentration; and
- The sensor's *resume time*, i.e., the time it takes for the sensor to be used reliably again after a change in concentration.

The change in system response based on the initial voltage reading (i.e., initial concentration at the sensor) is another factor affecting these sensors. This effect can be seen in Figure 7.10, where the system response to a single short spray of 100 ms at the distance of 4 meters away with different initial voltage readings (i.e. different initial concentrations) at the sensor is plotted. The flow in this figure is generated using the Honeywell fan on high setting (similar results are also observed when the Dyson fan is used for flow generation). This figure shows that the system response changes for different initial concentration levels at the sensor. To make sure that the sensor resume time is not effecting the readings, the sensor is brought up to a voltage level (i.e. concentration level) higher than the target initial voltage. Then, by waiting long enough the voltage reading drops to the target initial voltage level. This wait period is also long enough to eliminate the effects of sensor resume time. The impulse spray is then initiated and the corresponding system response

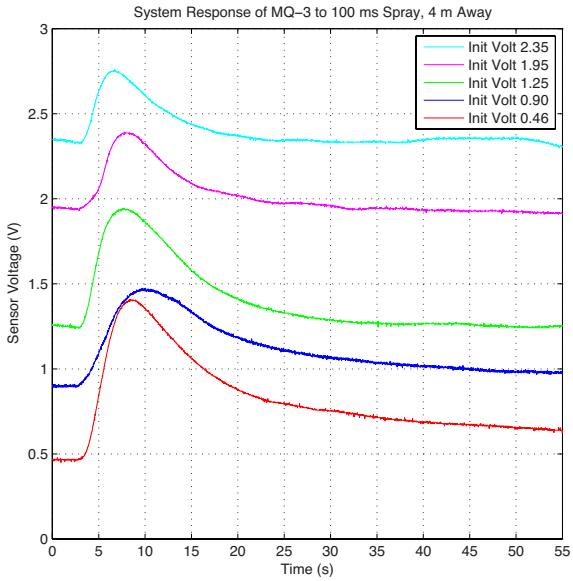


Figure 7.10: System response for different initial voltages. The system response changes based on the initial voltage.

is measured.

7.2.1.2 The Effects of Flow

Flow is an important part of the setup, because it carries the alcohol droplets from the transmitter to the receiver. Therefore, it has a significant effect on the overall system response. However, isolating the effects of flow can be very challenging. For example, the spray itself cannot release very precise amounts of alcohol with uniform droplet sizes between different experimental trials. Another factor that could potentially effect the results is random flows within the room. As a result,

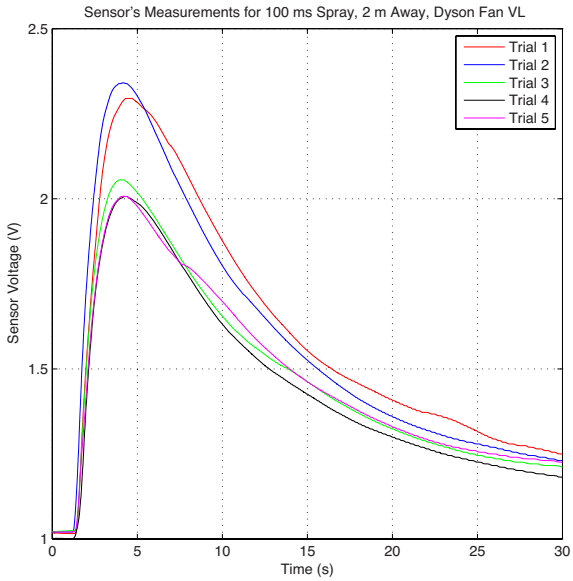


Figure 7.11: The system response for different experimental trials. The flow in these trials is generated using the Dyson fan on very low setting. The spray duration is 100 ms, and the separation distance is 2 m.

the overall impulse response of the system changes between different trials. This effect can be seen Figure 7.11, where the system response to a 100 ms spray 2 meters away is plotted for 5 different experimental trials. The initial voltage for each trial is kept constant at 1.02 volts, and the Dyson fan on the very low setting is used to generate the flow. Moreover, the sensor, the spray and the fans are kept at the same height. From the plot it is evident that there is some difference across the trials.

To mitigate this problem, and further isolate the effects of flow, multiple experimental trials are performed and average the results are presented. Two performance

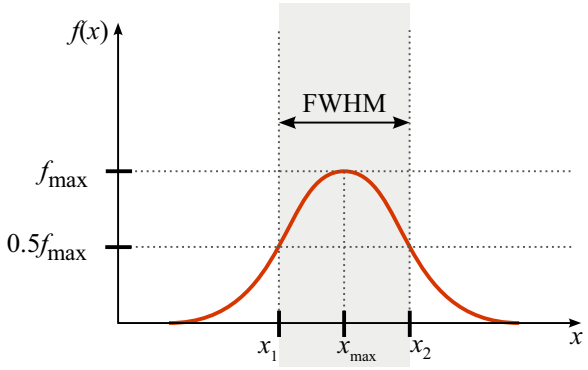


Figure 7.12: The peak's full-width at half maximum (FWHM) is shown in the plot. The peak's maximum (PM) is $f_{\max}$, and delay to peak's max (DPM) is $x_{\max}$.

metrics are used for comparing the system response generated using different fans and flow speeds: the peak's max to full-width at half max (PMFWHM), and delay to peak's max (DPM). The PMFWHM is the ratio of the peak's maximum voltage to the full width of the peak at half max. The larger this ratio the taller and narrower the peak shape will be. Ideally the peak must be as tall and as narrow as possible. Therefore, larger ratios are desirable. The DPM, is the time from the start of the spray to the time where the peak's maximum point is achieved. The smaller this time duration, the faster the peak's maximum is reached. It is desirable for this delay be as small as possible. Figure 7.12 shows FWHM, PM, and DPM.

As explained earlier, two different table fans are used, one made by Dyson and one by Honeywell. The Dyson fan is bladeless, more expensive, and can create more laminar flows. The Honeywell fan is inexpensive, but it is a bladed fan and

it creates more turbulent flows. Five different fan settings are considered for the Dyson fan, while the Honeywell fan has only two possible settings as explained in previous sections. The average flow velocities over a 2 meter distance are tabulated in Table 7.2. For each fan and each corresponding fan setting, the overall system response to a short spray of 100 ms, 2 meters away, is measured for 10 experimental trials. The initial sensor voltage reading for each trial is kept constant at about 1.02 volts (i.e. there is enough delay between trials such that the sensor voltage falls back to 1.02 volts).

To compare the fans and their corresponding fan settings, performance metrics, PMFWHM and DPM, are calculated for each of the 10 trials. The results are then averaged and presented in Figure 7.13. In Figure 7.13a, the DPM is plotted for different flow speeds generated by each fan and its corresponding fan setting. As can be seen, the Dyson fan has a shorter DPM for the same average flow speed because the flow generated by this fan is more laminar compared to the Honeywell fan. Moreover, the flow speed also decreases the delay. Finally, from the Dyson plot it can be seen that this delay decreases almost linearly with increasing flow speed.

The PMFWHM is shown in Figure 7.13b for different fans and setting. From the plots it can be observed that the PMFWHM ratio increases as the fan speed increases. Therefore, the impulse response becomes narrower and taller as the

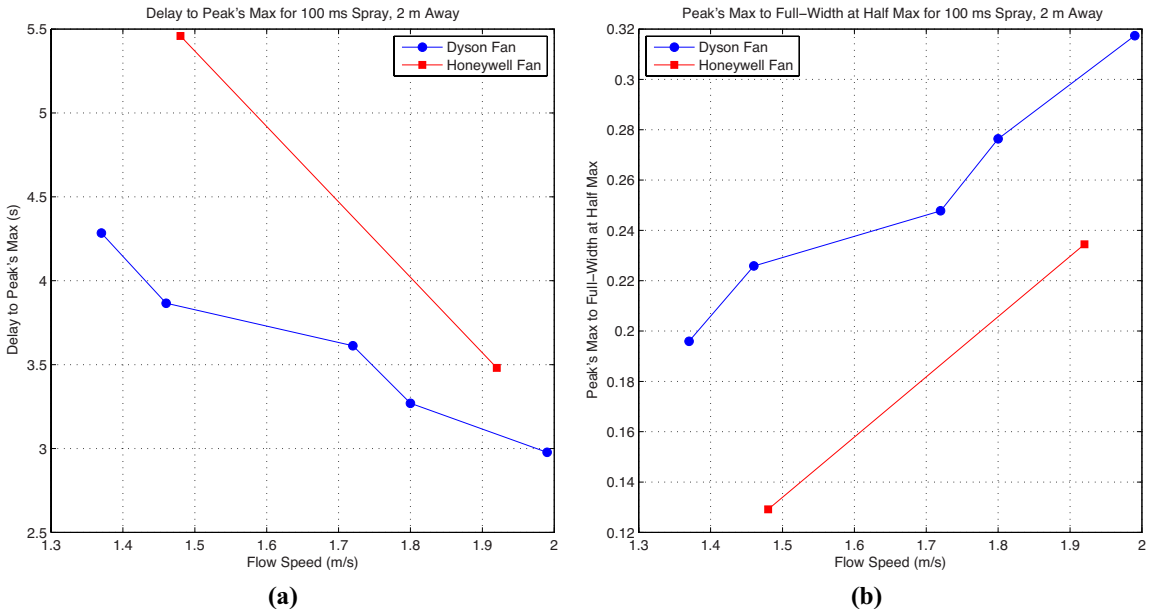


Figure 7.13: Analysis of the system response for different flows. The delay to peak's maximum is used as one performance metric (a), and the peak's maximum to full width at half max is used as a second performance metric (b). The results from 10 different trials are averaged to create each point. The spray duration is 100 ms, and the separation distance is 2 m.

fan speed increases. The Dyson fan also achieves higher ratios compared to the Honeywell fan. Therefore, the more laminar flows that the Dyson fan generates can create taller and narrower system response.

From these results, it is concluded that the Dyson fan is a better choice for generating flows. However, because it is more than 10 times expensive compared to the Honeywell fan, and one of the goals is to create a cost effective demonstration of macroscale molecular communication, for the final communication system the

Honeywell fan is used. Therefore, the achievable transmission rates can potentially be improved by simply using the Dyson fan in place of the Honeywell fan.

7.2.1.3 System Nonlinearity

In this section, it is shown that the overall system response of the setup is nonlinear. Although the exact cause for nonlinearity is not known, and more extensive research is required, this result by itself is very interesting. To show that the system is nonlinear, a set of periodic sprays of 100 ms are considered with a period of 2 seconds. The output sensor voltage is then measured and recorded as the system response. Figure 7.14 shows the results for both the Dyson fan on the very high setting and the Honeywell fan on the high setting.

As can be seen in Figure 7.14, the output does not follow that of a linear system. For example, between the 13 and 15 second marks (the arrows in the plot point to this time duration), where there should be another increase in concentration because of the seventh periodic spray, there is a sudden drop in voltage. This effect is observed for both the case where the Dyson fan is used and the case where the Honeywell fan is used. However, there are more clear peaks when the Dyson fan is used because of the narrower and taller system response explained in the previous section.

The nonlinear responses observed in the experiments are surprising, because

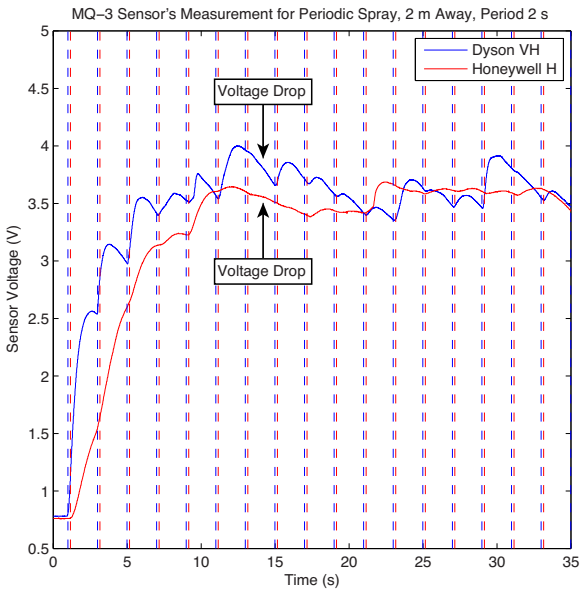


Figure 7.14: System response to a periodic spray. The dashed lines are 2 seconds apart and show each period. The arrows show the location where the sensor voltage decreases instead of increasing. The spray duration is 100 ms, and the separation distance is 2 m.

most molecular communication systems are normally assumed to be linear in the literature. Many of the mathematical tools used in the literature at microscales require the system to be linear, and these tools cannot be directly applied to a nonlinear communication system. Although the source of nonlinearity is not known, some likely candidates are: the sensor with its response and resume times, the flow generated by the fans which may be turbulent, the spray which is not precise enough to create uniform streams, and other environmental factors such as other flows within the room. It may be possible for the system response to become linear

with the use of more expensive and sensitive hardware, and within a precisely controlled environment. Nonetheless, the potential nonlinearity of the system is an issue, which is investigated further in the next chapter.

7.2.2 Final Implementation and Discussion

The final steps of implementation are discussed in this section. The Honeywell fan is used for demonstration despite the fact that the Dyson fan can create better system response, to make the final platform more cost effective. Moreover, experimenting using Honeywell fan provides a lower bound on reliability. In the final platform, MQ-3 sensor is used because it provides the best system response, and it has the simplest circuitry between all three sensors.

First, the noise issue is addressed. Although the MQ-3 sensor response is less corrupted by noise compared to the other two sensors, there is still some noise present in the signal. To further reduce this noise, 20 ms of sensor data is averaged to generate a single sensor reading. Because the Arduino sampling rate is observed to be about 8.33 kHz, 20 ms of sensor data contains 167 different readings which are then averaged. Therefore, in the wait state the receiver checks consecutive 20 ms of averaged sensor readings, and triggers a change to the reception state if the difference between the current reading and the previous reading is greater than 0.5 levels (because Arduino has a 10 bit analog to digital converter the sensor reading

would be an integer between 0 and 1023 representing 1024 different voltage levels, where 0 represents 0 volts and 1023 represents 5 volts).

An important communication parameter is the transmission rate. One of the major factors that affects reliable communication at a given transmission rate is the DPM. The DPM is in turn affected by the flow type and the flow speed. Therefore, for the platform the fan speed is always set to high. Another factor that effect the transmission rate is the sensor response and resume times discussed in the previous section. Finally, many other factors such as the environmental noise (e.g. random flow patterns in the room) can also effect the transmission rate.

Different experiments are carried out with various transmission rates from one bit per 5 seconds (a character per 25 seconds) to one bit per 2 seconds (a character per 10 seconds). To measure the reliability at each rate, multiple experiments are performed at different separation distances between the transmitter and the receiver. Each transmission rate at each separation distance is then classified according to the following ranking: very reliable (bit error rates of less than 0.01), reliable (bit error rates of 0.01 to 0.03), unreliable (bit error rates greater than 0.03). Table 7.3 summarizes the results.

At one bit per 2 seconds the transmission is unreliable at small distances of up to 2 meters, because of the sensor's resume time at higher concentrations is longer. Moreover, at larger distances (greater than 2 meters) successful communication is

Table 7.3: Different transmission rates and their reliability.

	Distance		
Transmission Rate (bits/s)	**2 m**	**3 m**	**4 m**
0.2	Very Reliable	Very Reliable	Very Reliable
0.33	Reliable	Very Reliable	Reliable
0.5	Unreliable	Unreliable	Unreliable

not possible at the rate of one bit per 2 seconds. At the rate of one bit per 5 seconds, the transmission is very reliable over various separation distances from 4 meters to 1 meters. Based on experiments, the smallest transmission rate that is reliable at distances up to 4 meters is one bit per 3 seconds. At this rate over the separation distance of 4 meters the communication session is reliable, over the separation distance of 3 meters the communication session is very reliable, and over the separation distance of 2 meters the communication session is reliable. The reason that the communication channel degrades slightly as the separation distance is decreased from 3 meters to 2 meters is because of the higher concentration levels at the sensor, and hence sensor saturation and longer sensor resume times.

In the rest of the section, this transmission rate is used (one bit per 3 seconds), and the demodulation and detection algorithm is described in details for this rate. Although the thresholds used in this algorithm may be slightly different for each

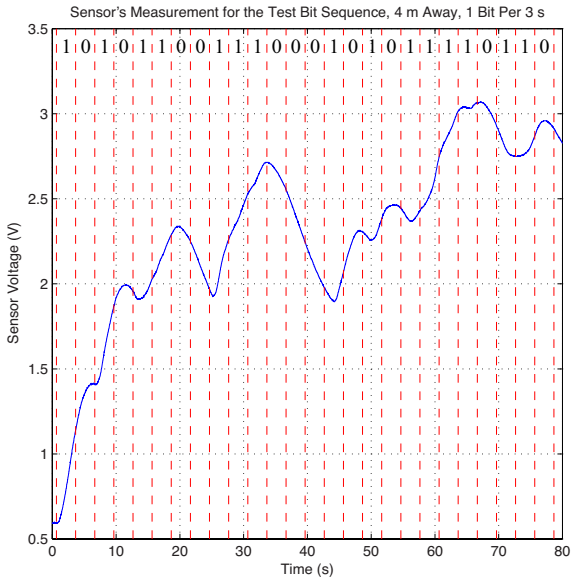


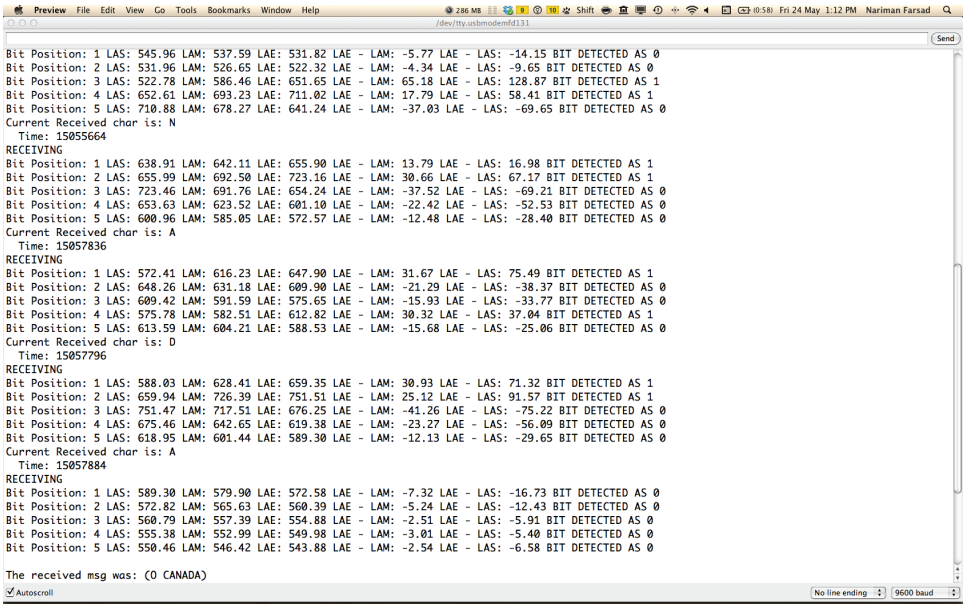
Figure 7.15: Received signal when a 26 bit test sequence is transmitted. Sensor reading for the 26 bit test sequence "10101100111000101011110110" transmitted 4 meters away at the rate of one bit per 3 seconds. The dashed red lines represent the start and the end of each bit. Honeywell fan is used to create the flow for this figure.

transmission rate, the same underlying principal is used for detection and demodulation at all rates: the rate of change in the concentration at the sensor. To fine tune this algorithm a 26 bit test sequence "10101100111000101011110110" is transmitted at the distance of 4 meters away, and the sensor reading is recorded. Figure 7.15 plots the sensor voltage reading during this transmission session. Dashed red lines are used to represent the start and the end of each bit. From this plot a simple detection and demodulation scheme is devised. The difference between the voltage level (there are 1024 levels in the Arduino's 10 bit analog to digital converter) at the

end of a bit interval and the middle of a bit interval is measured. If the difference is greater than 2.2 levels (this threshold is derived through experimentation), the bit is detected as 1; otherwise the bit is detected as 0.

Using this scheme, the test phrase "O CANADA" (the name of the national anthem of Canada) is successfully transfered from the transmitter to the receiver. Figure 7.16 shows this test phrase at the transmitter and received at the receiver.

(a)



(b)

Figure 7.16: Transmitted and received text message. Pictures from a communication session: (a) the text entered at the transmitter (b) the text received at the receiver.

8 Channel Models for the Tabletop Platform

In the previous chapter, an experimental platform for molecular communication was developed that is capable of transporting short text messages. The impulse response for the system was presented and it was shown that the system tends to be nonlinear. In this chapter, mathematical models are developed for the impulse response of the system, and this nonlinearity property is further investigated to show that it could be represented as additive noise.

8.1 Experimental Setup and Previous Theoretical Models

8.1.1 Tabletop Test bed

The macroscale tabletop test-bed which was presented in the previous chapter is used for the experiments performed in this chapter. As shown in the previous chapter, the end-to-end system impulse response for this platform can be obtained by using a very short spray (e.g. 100 ms) at the transmitter (which resembles the delta function from signal processing), and measuring the sensor output at the receiver.

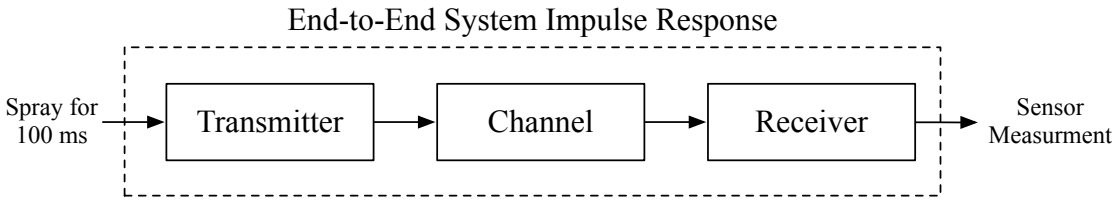


Figure 8.1: The end-to-end system impulse response is generated by using a very short spray.

This is demonstrated in Figure 8.1. The end-to-end system impulse response includes the transmitter block, the channel and the propagation mechanise, and the receiver block. Therefore, the effects of all these blocks are incorporated in the end-to-end system impulse response.

To perform measurements, the transmitter and the receiver are separated by 225 cm. This distance is selected as an example, and the separation can be any distance. At the sensor, the voltage output is measured and the observed data is recorded. Figure 8.2 shows the system responses for 5 different trials. Between each trial there is sufficient wait period until the initial voltage reading of the sensor drops to about 1 volts. Although it is extremely difficult to find the exact cause of deviations between trials, some likely causes are: the spray, which is not precise enough to spray the same amount of alcohol for each trial; the flow, which can be turbulent; the sensor, which can be noisy; and other environmental factors such as random flows within the room.

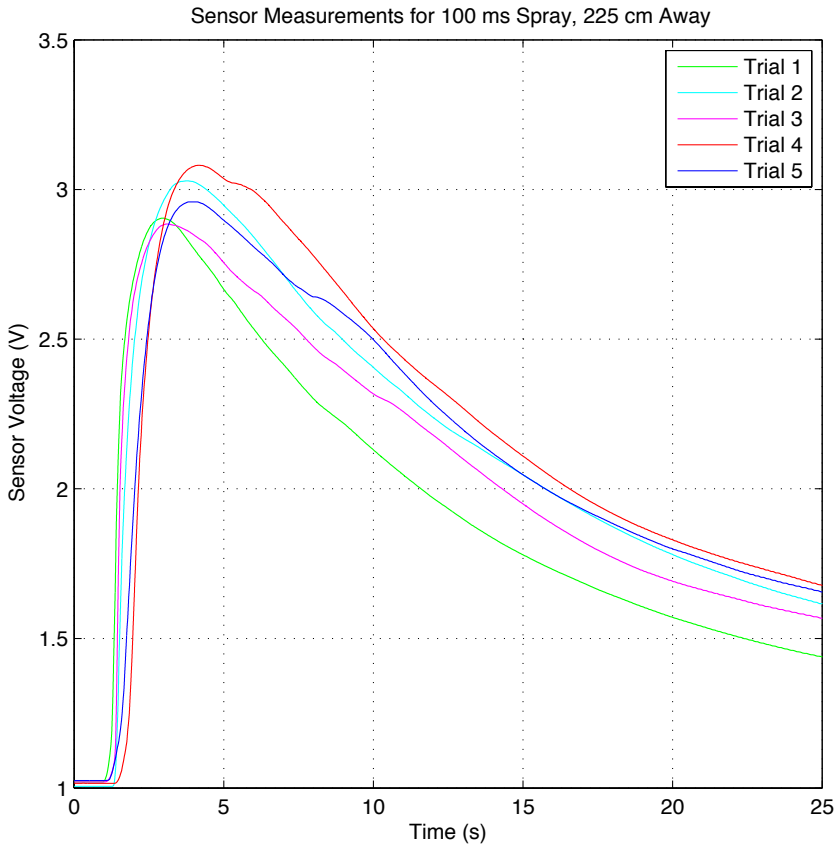


Figure 8.2: The end-to-end system impulse response obtained experimentally across five different trials.

8.1.2 Previous Theoretical Models

It is assumed that the spray and the sensor have the same height in the 3-D space, and that the fan's flow is perfectly aligned with the line connecting the transmitter to the receiver[1]. Therefore, the problem reduces to a one dimensional flow-assisted diffusion. If it is also assumed that the sensor and the transmitter are perfect, and

[1]These assumptions can be easily satisfied through careful placement of the transmitter and the receiver

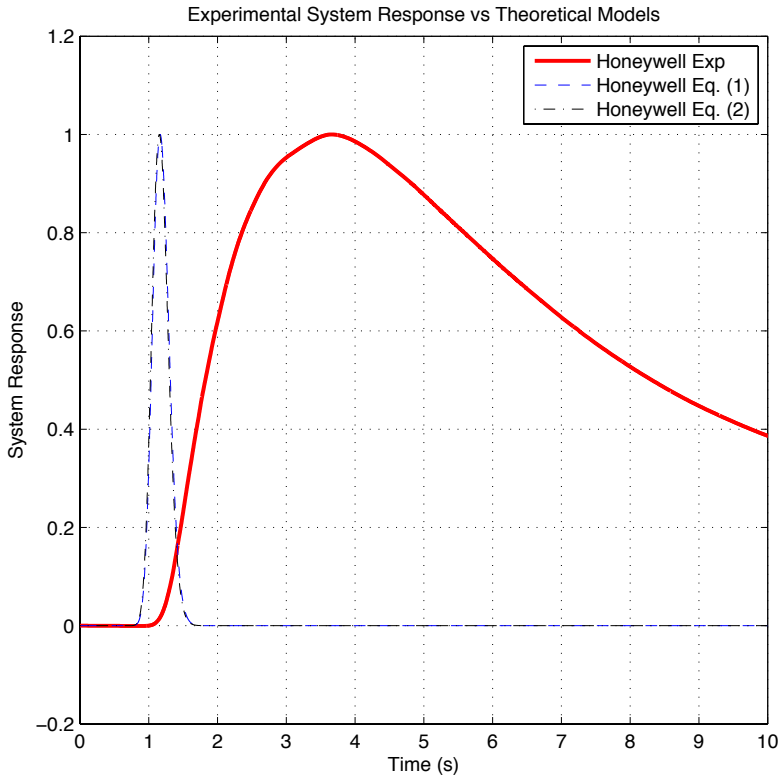


Figure 8.3: Comparison of the experimental data and theoretical models from previous works. The curves from the theoretical Equation (8.1) and (8.2) are very similar in this case and almost overlap. This follows because the speed of the flow is much greater than diffusion coefficient. Therefore, although the molecules are not absorbed by the sensor according to (8.1), they are moved away from the sensor by the flow.

that the sensor does not absorb (or adsorb) the alcohol molecules (i.e. the alcohol molecules stay in the environment after detection), the impulse response at the receiver should be well approximated by [66]:

$$h_1(t) = \frac{M}{\sqrt{4\pi Dt}} \exp\Big(- \frac{(d - vt)^2}{4Dt} \Big), \tag{8.1}$$

where M is the number of molecules released during the short burst, D is the diffusion coefficient, d is the separation distance between the transmitter and the receiver, v is the average flow speed from the transmitter to the receiver, and t is time. This is essentially the probability distribution of a Wiener process with drift, conditioned at a fixed distance d and multiplied by the number of molecules M. If it is assumed that the alcohol molecules are absorbed (or adsorbed) by the sensor upon detection, then the problem would be equivalent to the first arrival time, and the impulse response would have a similar shape to the inverse Gaussian distribution [66] given by

$$h_2 = \frac{Md}{\sqrt{4\pi Dt^3}} \exp\Big(- \frac{(vt-d)^2}{4Dt} \Big). \tag{8.2}$$

Although the number of molecules sprayed by the transmitter is not known (in fact it is random because each spray is not perfectly and precisely similar to previous sprays), based on theoretical results, it is expected that the sensor output should have a shape similar to the curves obtained from either (8.1) (in case the molecules are not absorbed by the sensor) or (8.2) (in case they are absorbed by the sensor).

Table 8.1: System Parameters.

Parameters	Values
Spraying duration for each bit	100 ms
Distance between a transmitter and a receiver	225 cm
Approximated fan speed Honeywell	190 cm/s
Diffusion coefficient of isopropyl alcohol	0.0959 cm^2/s
Temperature (room temperature)	25 °C = 298 K

8.1.3 Models versus Experimental Results

In this section, it is shown that the previously published theoretical models, described in the last section do not match the experimental results obtained using the tabletop platform. To demonstrate this, the transmitter and the receiver are separated by 225 cm. The Honeywell fan is set on the high setting to generate flows. Table 8.1 summarizes all the system parameters of this setup. The flow speed of the wind generated by the fan, which is tabulated in Table 8.1 is measured using Pyle PMA82 digital anemometer.

If these parameters are used in the theoretical Equations (8.1) and (8.2), the system response can be calculated. Because the number of particles released by the transmitter is not known, it is assumed that $M = 1$ and then the plots are normalized by dividing them by their respective maximums. Similarly, the system responses obtained from experimental results is normalized with its maximum. By

normalizing the plots, only the shape of the theoretical results are compared with the shape of the experimental results. For the experimental system response, the response from 12 different experimental trials are averaged to produce a single plot. Moreover, the initial voltage is subtracted from the system response to effectively zero the starting voltage.

As shown in Figure 8.3, the experimentally obtained response has a much wider peak width, and longer tail compared to theoretical predictions. The difference between the theoretical results and the observed system response is because of many assumptions made in the derivation of the theoretical results. For example, the flow is assumed to be perfectly laminar and the sensor are assumed to be perfect at detection of concentration. These assumptions do not hold for the experimental platform. Therefore, in the next section more realistic theoretical models are derived based on the observed experimental data.

8.2 Realistic Models

In this section, the experimental data are used to derive a more realistic theoretical model for the platform. First, likely causes of the deviation from the theoretical results are found. In particular, two system components can have a huge effect on the system response: the sensor, and the flow. The previously published channel models assume a perfectly laminar flow, as well as perfect and instantaneous detection

at the sensor. These assumptions do not hold for this experimental platform.

All metal-oxide sensors, have a response time and a recovery time [160]. The response time is the time it takes for the sensor to respond to a sudden change in concentration. The recovery time is the time it takes for the sensor to drop to its initial voltage after a sudden change in concentration. The concentration function with respect to time is therefore expanded because of the response and recovery times. To compensate for this effect, the system response function in (8.1) and (8.2) must be scaled in time by a factor of α as $h_1(\alpha t)$ and $h_2(\alpha t)$, where $0 < \alpha < 1$.

Another factor that affects the system response is the flow. Previous channel models have assumed the flow to be perfectly laminar and uniform. This is, however, not the case for the platform. The high wind speeds generates turbulence within the flow. Moreover, the Honeywell fan's blades can create pockets of air pressure that can result in more turbulent flows. Fortunately, Fick's law of diffusion can still be applied to turbulent flows with a correction term added to the diffusion coefficient [155]. Therefore, a correction must be made to the diffusion coefficient D in (8.1) and (8.2).

The final factor considered is the flow speed. Although the wind speed generated by the fans is known, the alcohol droplets in the spray stream may be travelling at a slower average speed because of their weight and air friction. Therefore, a third correction is needed in (8.1) and (8.2) for the average flow velocity v.

Considering these three effects, two new models are proposed based on (8.1) and (8.2),

$$M_1(t) = \frac{a}{\sqrt{t}} \exp\Big(- b\frac{(d-ct)^2}{t}\Big), \tag{8.3}$$

$$M_2(t) = \frac{a}{\sqrt{t^3}} \exp\Big(- b\frac{(ct-d)^2}{t}\Big), \tag{8.4}$$

where a, b, and c are corrected constants. The corrected constant a contains the scaling factor α from the sensor respond and resume times, and the correction to the diffusion coefficient because of turbulent flow. The corrected constant b contains the correction to diffusion coefficient because of turbulent flow and scaling factor α. Finally, the corrected coefficient c contains the correction to the average flow speed as well as the scaling factor α.

8.2.1 Estimating the Coefficients

The experimental data from the platform are used to estimate the value of these corrections. To do this the transmitter and the receiver are placed 225 cm apart. The sensor, the spray and the fans are placed at the same height, with the fan blowing in the direction of the line connecting the spray to the sensor. The end-to-end system impulse response to a very short spray burst of 100 ms is measured and recorded during 12 different experimental trials.

To estimate the coefficients of models M_1 and M_2, let $M_i(t_k, \mathbf{p})$ $i \in \{1, 2\}$ be the corresponding model (i.e. model M_1 or M_2) at the sampled time instance t_k with parameter vector $\mathbf{p} = [a, b, c]^T$, where a, b, and c are the three coefficients for each model. Nonlinear least square curve fitting is used to estimate the coefficients. Assuming that there are N points in each sensor measurement and that each point in the measurement is represented by a function $m(t_k)$ $k \in \{1, 2, \cdots, N\}$, the coefficient estimation problem can be formulated as

$$\min_{\mathbf{p}} \sum_{k=1}^{N} \big(m(t_k) - M_i(t_k, \mathbf{p})\big)^2. \tag{8.5}$$

This problem can then be solved using iterative algorithms such as Levenberg-Marquardt algorithm [262].

To perform the nonlinear least square estimation using Levenberg-Marquardt algorithm, MATLAB's curve fitting function `fit()` is used for coefficients estimation of each experimental trial. Because in the experimental setup it was demonstrated that only the first few seconds of the impulse response is typically used in practice for information transmission, only the first 5 seconds of sensor measurements are used for curve fitting. Figs. 8.4 and 8.5 show the results where the Honeywell fan on the high setting is used for flow generation, and model M_1 is used for curve fitting. In Figure 8.4, it can be seen that the fitted model resembles the obtained

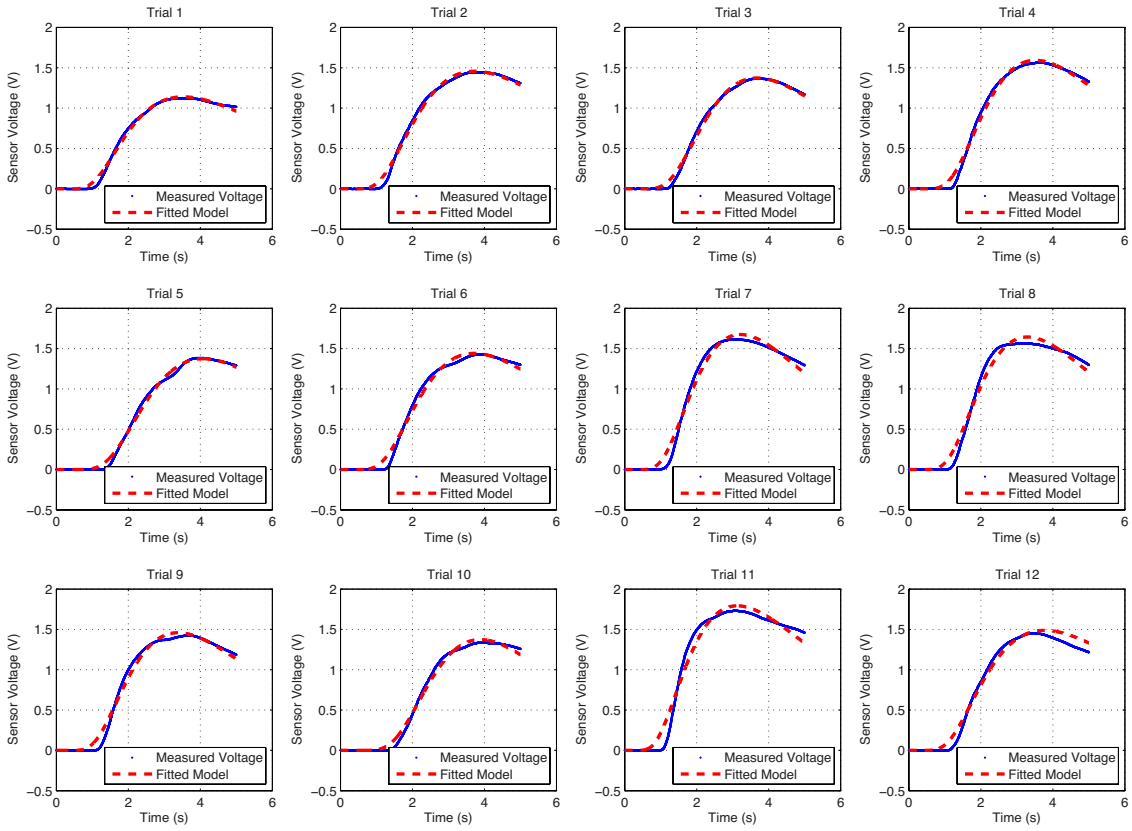


Figure 8.4: Sensor measurements and the fitted model for set of 12 trials. The measurements are fitted with model M_1.

results much more accurately compared to Figure 8.3. Figure 8.5 shows the plot of each coefficient value for each trial. The dashed red line indicates the mean value of each coefficient.

For the goodness of fit measure, the root mean square error (RMSE) between the fitted model and the experimentally observed system responses is used. The variance-to-mean ratio (VMR) is employed as a goodness of fit measure. If this ratio is greater than one, then the resulting coefficient is not a good fit. If this ratio

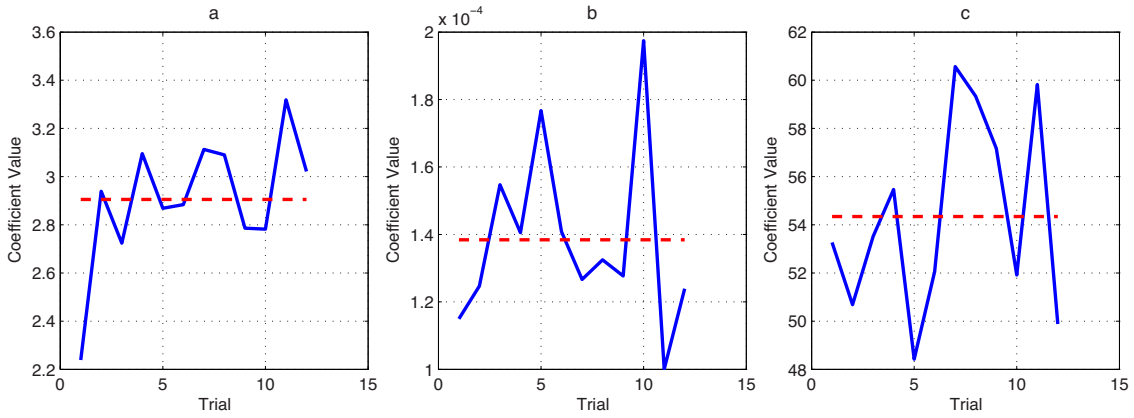


Figure 8.5: The coefficients' variation for 12 trials. The dashed red line is the mean value of each coefficient.

is less than 1, then the coefficient is a good fit. Table 8.2 summarize the result for both model M_1 and model M_2 given in (8.3) and (8.4), respectively. In the table, the mean RMSE is the average RMSE across all 12 experimental trials.

From the results it can be seen that model M_1 has a better VMR, while model M_2 has a slightly better RMSE. Generally, because model M_1 has a lower VMR, the coefficients are more consistent between different experimental trials. Therefore, model M_1 may be more effective at consistently modelling the end-to-end system impulse response.

To further compare the proposed models to the experimental results, the system response from all 12 trials are averaged to generate the final experimental system response. The mean of the coefficients across all 12 trials are used in each model to generate the corresponding system response. The results are shown in Figure 8.6.

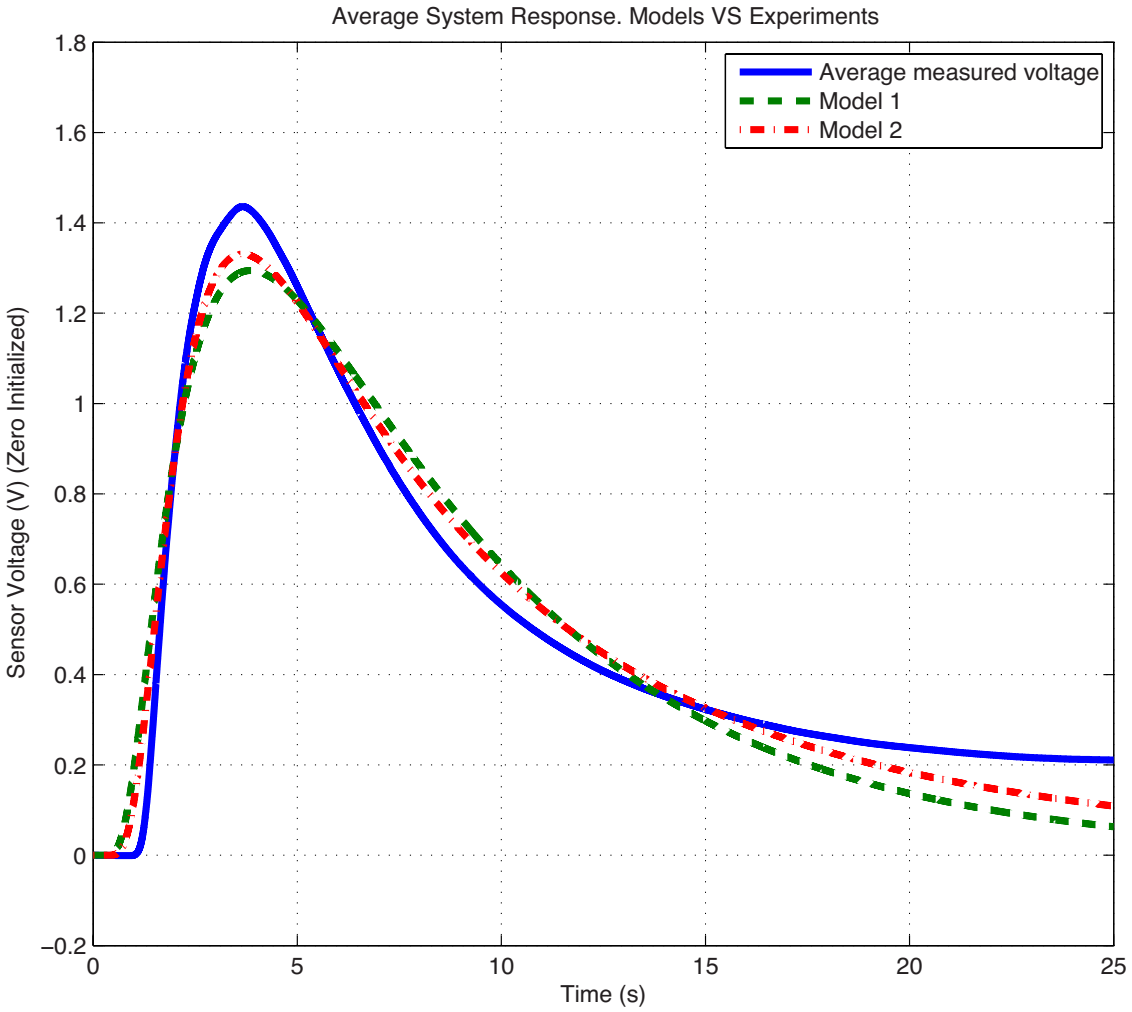


Figure 8.6: Average system response of experimental observations and fitted models.

Based on the results, it is evident that both new models capture the average system response of the test bed platform much more accurately compared to old models.

Table 8.2: The obtained coefficients for each model.

Model M_1				
	Mean	Variance	Variance/Mean	Mean RMSE
a	2.9050	0.0672	0.0231	
b	1.3839×10^{-4}	6.6117×10^{-10}	4.7775×10^{-6}	0.0539
c	54.3405	15.3455	0.2824	
Model M_2				
	Mean	Variance	Variance/Mean	Mean RMSE
a	15.3909	4.2150	0.2739	
b	1.6×10^{-4}	6.9109×10^{-10}	4.31×10^{-6}	0.0501
c	35.3136	29.5543	0.8369	

8.3 System Nonlinearity

In the previous chapter, it was demonstrated that the platform has a nonlinear system impulse response. This nonlinearity, however, was not investigated in great detail. Although it is extremely difficult to find the exact cause of nonlinearity (some likely causes are imperfect receiver, transmitter, and turbulent flows), in the previous chapter the nonlinearity property was demonstrated through measurements. In this section, the nonlinearity property is verified through systematic experimentation using two transmitters and a single receiver. This nonlinearity is then modeled as noise and the underlying distribution for this noise is found by employing the developed models in the previous section.

To study the nonlinearity, two transmitters and one receiver are used. Each

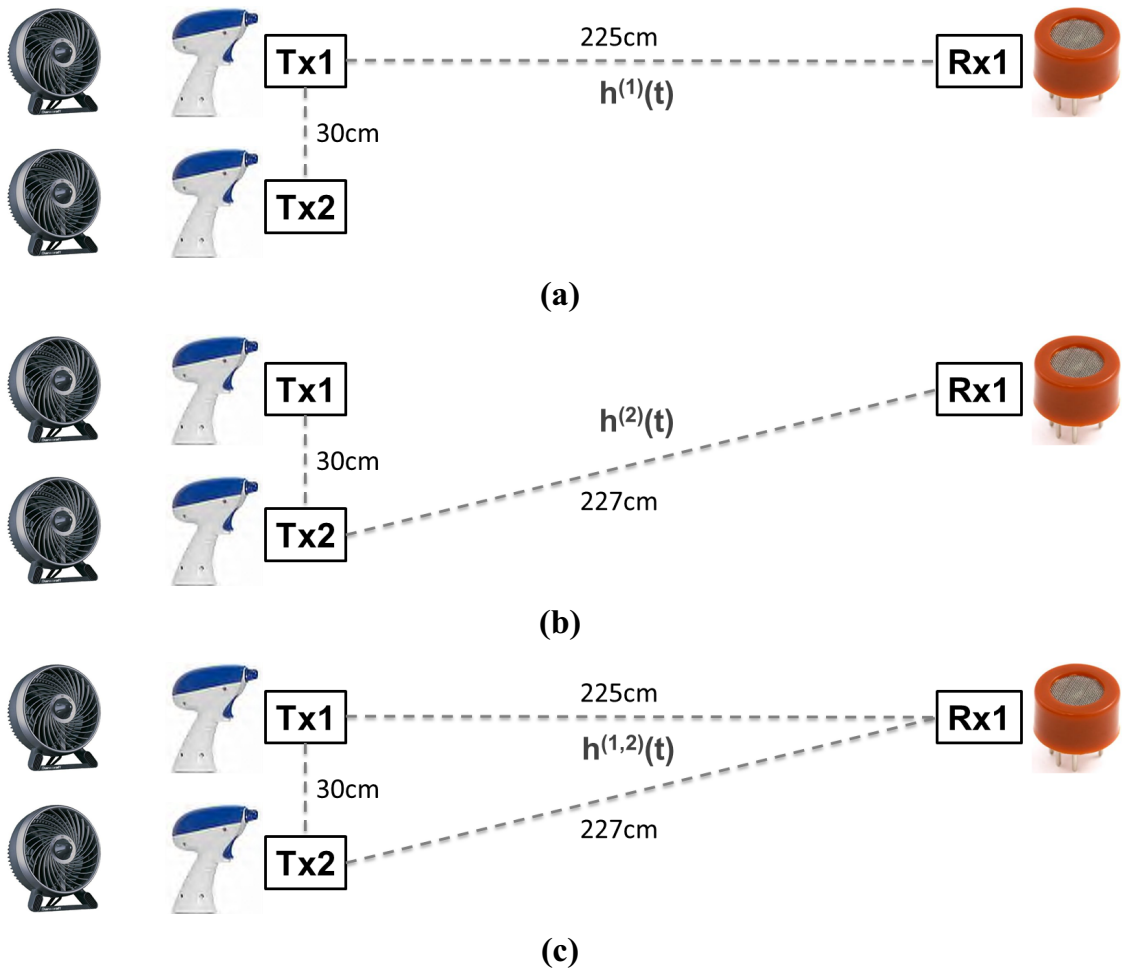


Figure 8.7: The setup used to demonstrate the nonlinearity.

transmitter has its own Honeywell fan. The transmitters are 30 cm apart, and the receiver is directly in front of the first transmitter separated by 225 cm. Figure 8.7 summarizes this setup. Let $h^{(1)}(t)$ and $h^{(2)}(t)$ be the end-to-end system impulse response, when transmitter 1 and 2 spray a short burst of 100 ms in duration, respectively. Moreover, let $h^{(1,2)}(t)$ be the end-to-end system impulse response

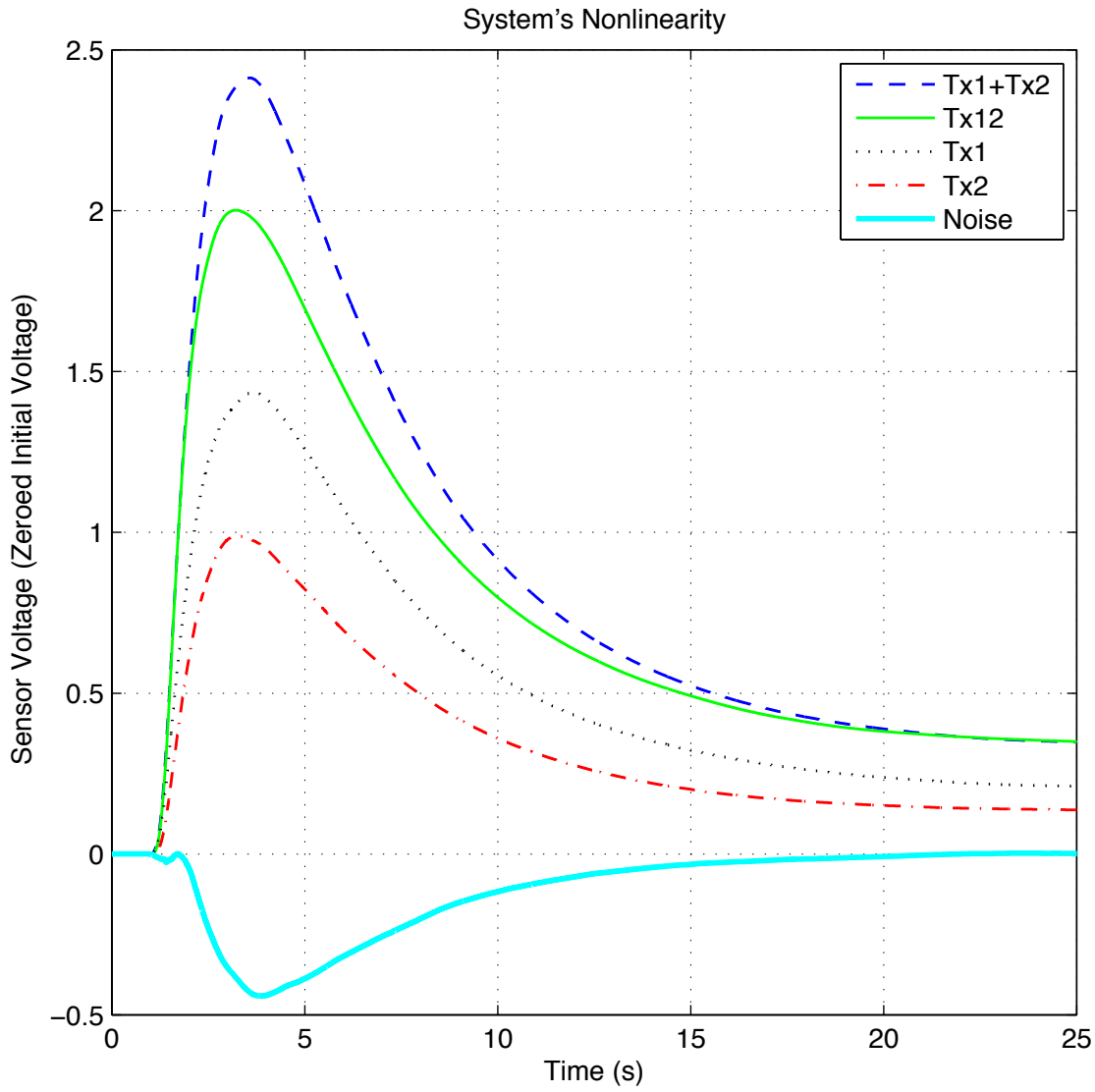


Figure 8.8: Representation of system's nonlinearity.

when both transmitters spray a short burst of 100 ms in duration simultaneously.

If the system is linear, then

$$h^{(1,2)}(t) = h^{(1)}(t) + h^{(2)}(t). \tag{8.6}$$

As shown in Figure 8.8 this property does not hold. The system responses for $h^{(1)}(t)$ (Tx1), $h^{(2)}(t)$ (Tx2), and $h^{(1,2)}(t)$ (Tx12) are generated by averaging 12 different trials to reduce the noise that may be introduced by other processes such as random flow patterns in the room. As can be seen the $h^{(1,2)}(t)$ (Tx12) plot and the $h^{(1)}(t) + h^{(2)}(t)$ (Tx1 + Tx2) plot are not equal.

Although there may be many contributors to the end-to-end system nonlinearity, based on these results, one of the most important contributors is the sensor. In essence the sensor may act as a nonlinear filter. In the next section, this nonlinearity is modeled as a noise process. Through this formulation the effects of this nonlinear may be counter by using a filter.

8.3.1 Modelling the Nonlinearity as Noise

The nonlinearity in the system can be modelled as noise. Because the functions $h^{(1)}(t)$, $h^{(2)}(t)$, and $h^{(1,2)}(t)$ are random processes, the nonlinearity noise can be

represented as a random process given by

$$h^{(1,2)}(t) = h^{(1)}(t) + h^{(2)}(t) + n(t), \tag{8.7}$$

where $n(t)$ is the noise process. To find the underlying model for the noise process using experimental measurements, (8.7) is rewritten as

$$n(t) = h^{(1,2)}(t) - h^{(1)}(t) - h^{(2)}(t). \tag{8.8}$$

The expected value of the noise process is then given by:

$$E[n(t)] = E[h^{(1,2)}(t)] - E[h^{(1)}(t)] - E[h^{(2)}(t)]. \tag{8.9}$$

In Figure 8.8 the expected value of the noise process is represented by the solid cyan plot. The noise process in this case is therefore nonstationary, because its expected value in not a constant in time.

To simplify the noise model, the mathematical models derived in the previous sections are used. To find the coefficients of this model $h^{(1)}(t)$, $h^{(2)}(t)$, and $h^{(1,2)}(t)$ are measured using 36 different trials (12 distinct trials for each case). Let $h_i^{(1)}(t)$, $h_j^{(2)}(t)$, and $h_k^{(1,2)}(t)$ $(i, j, k = 1, 2, \ldots, 12)$ be the results of each trial. Using least squares, the coefficients of model M_1 and M_2 are estimated based on the observed

Table 8.3: The obtained mean coefficients of model function.

Model M_1			
Coefficient	$h^{(1)}$	$h^{(2)}$	$h^{(1,2)}$
a	2.905	1.9815	3.9737
b	1.3839×10^{-4}	1.5605×10^{-4}	1.3474×10^{-4}
c	54.3405	59.4961	58.669
Model M_2			
Coefficient	$h^{(1)}$	$h^{(2)}$	$h^{(1,2)}$
a	15.3909	12.6246	18.7617
b	1.6035×10^{-4}	1.5119×10^{-4}	1.5885×10^{-4}
c	35.3137	28.9746	40.7749

data from each trial. Again, the first 5 seconds of sensor measurements are used for curve fitting, since in practice this would be the information carrying interval. Table 8.3 shows the average value of each coefficient across different trials.

From this table it can be seen that coefficients b and c do not differ by more than about 10 to 15 percent. Moreover, the separation distance between the transmitters 1 and 2, and the receiver are almost similar. However, coefficient a changes significantly, depending on which transmitters are spraying. This is consistent with previous theoretical works, where system linearity is assumed. Assuming that the coefficients b and c are similar across different trials and different transmitter sprays,

(8.8) becomes

$$
\begin{aligned}
n(t) &= h^{(1,2)}(t) - h^{(1)}(t) - h^{(2)}(t), \\
n_{M_1}(t) &\approx \frac{a_{M_1}^{(1,2)}}{\sqrt{t}} \exp\Big(\frac{-b(d-ct)^2}{t}\Big) \\
&\quad - \frac{a_{M_1}^{(1)}}{\sqrt{t}} \exp\Big(\frac{-b(d-ct)^2}{t}\Big) \\
&\quad - \frac{a_{M_1}^{(2)}}{\sqrt{t}} \exp\Big(\frac{-b(d-ct)^2}{t}\Big),
\end{aligned} \tag{8.10}
$$

$$
n_{M_1}(t) \approx \frac{N_{M_1}}{\sqrt{t}} \exp\Big(\frac{-b(d-ct)^2}{t}\Big), \text{ and} \tag{8.11}
$$

$$
N_{M_1} = a_{M_1}^{(1,2)} - a_{M_1}^{(1)} - a_{M_1}^{(2)}, \tag{8.12}
$$

where $a_{M_1}^{(1)}$, $a_{M_1}^{(2)}$, and $a_{M_1}^{(1,2)}$ are the first coefficients of model M_1 fitted to $h^{(1)}(t)$, $h^{(2)}(t)$, and $h^{(1,2)}(t)$, respectively, and N_{M_1} is the simplified noise model. Using the same procedure a simplified noise model can be generated based on model M_2 as

$$
n_{M_2}(t) \approx \frac{N_{M_2}}{\sqrt{t^3}} \exp\Big(\frac{-b(ct-d)^2}{t}\Big), \text{ and} \tag{8.13}
$$

$$
N_{M_2} = a_{M_2}^{(1,2)} - a_{M_2}^{(1)} - a_{M_2}^{(2)}, \tag{8.14}
$$

where $a_{M_2}^{(1)}$, $a_{M_2}^{(2)}$, and $a_{M_2}^{(1,2)}$ are the first coefficients of model M_2 fitted to $h^{(1)}(t)$, $h^{(2)}(t)$, and $h^{(1,2)}(t)$, respectively, and N_{M_2} is the simplified noise. Using this method the noises becomes N_{M_1} and N_{M_2} become random variables.

To find underlying probability distribution of N_{M_1} and N_{M_2}, the first coefficient (a coefficients) of each model is estimated for each $h_i^{(1)}(t)$, $h_j^{(2)}(t)$, and $h_k^{(1,2)}(t)$ $(i, j, k = 1, 2, \ldots, 12)$ from the experimental trials. In these estimations, it is assumed that the value of coefficients b, c and d are constant, and use the average value of $b_{M_1} = 1.4306 \times 10^{-4}$ and $c_{M_1} = 57.5018$ when model M_1 is used, and values of $b_{M_2} = 1.57 \times 10^{-}4$ and $c_{M_2} = 35.021$ when model M_2 is used. These values are obtained by averaging the corresponding row in Table 8.3. It is also assumed that the distance to the receiver for both transmitters is $d = 225$. From the obtained coefficients 1728 (12^3) different noise samples are generated using

$$N_{M_1}[i, j, k] = a_{M_1}^{(1,2)}[k] - a_{M_1}^{(1)}[i] - a_{M_1}^{(2)}[j], \tag{8.15}$$

$$N_{M_2}[i, j, k] = a_{M_2}^{(1,2)}[k] - a_{M_2}^{(1)}[i] - a_{M_2}^{(2)}[j], \tag{8.16}$$

Figs. 8.9 and 8.10 show the histogram of the of the sample noises generated using Equations 8.15 and 8.16, respectively. As can be seem the results are close to Gaussian. Therefore, although the exact distribution is not known Gaussian assumption is favourable [263, 264]. The Gaussian fit plot is generated using the mean and the variance of the sample. The mean of the sample for samples generated based on model M_1 is $\mu = -0.7356$ and the variance of the sample is $\sigma^2 = 0.5214$.

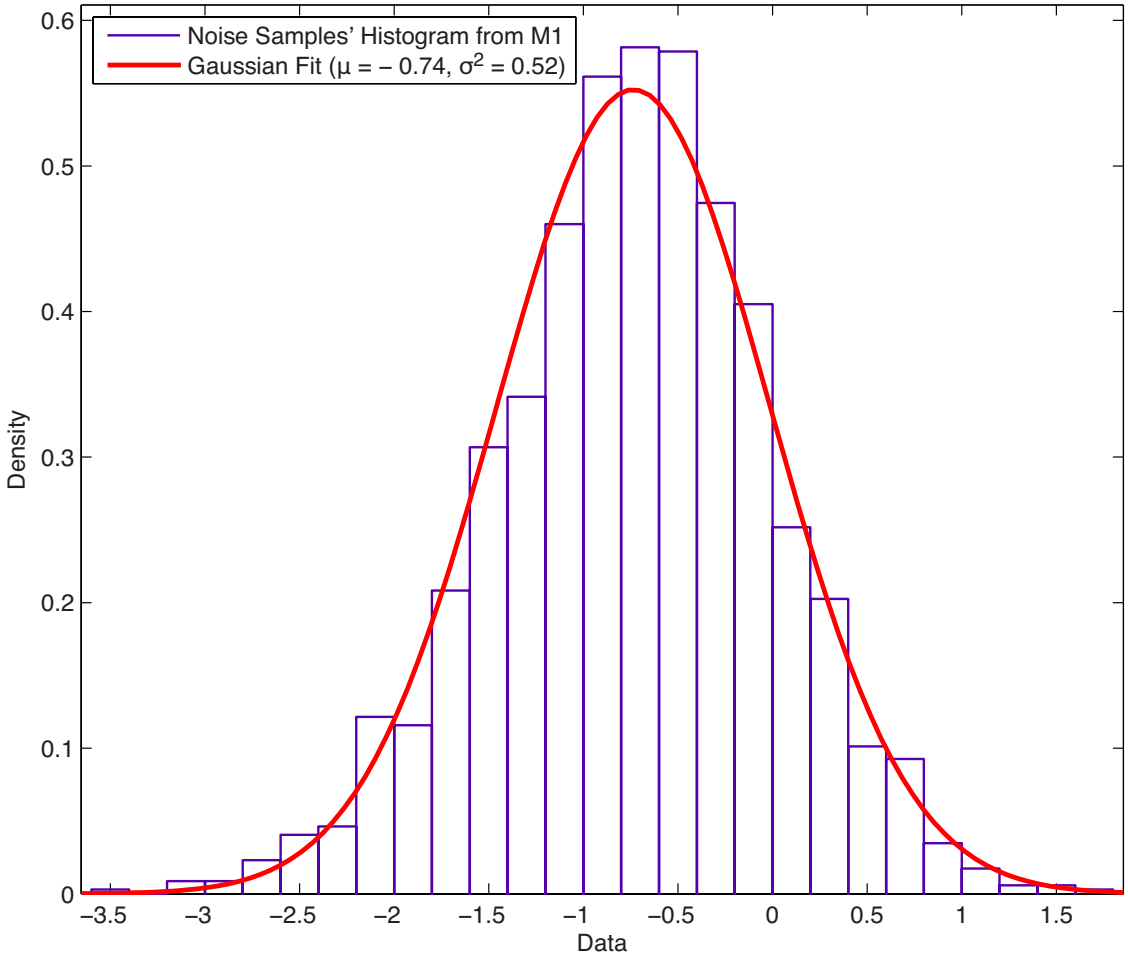


Figure 8.9: Histogram of the noise samples based on model M_1 and fitted Guassian probability density.

The mean and variance for the samples generated using model M_2 are $\mu = -3.9811$ and $\sigma^2 = 14.9589$. The samples generated based on model M_2 have a much larger variance because of the lager VMR of this model compared with model M_1.

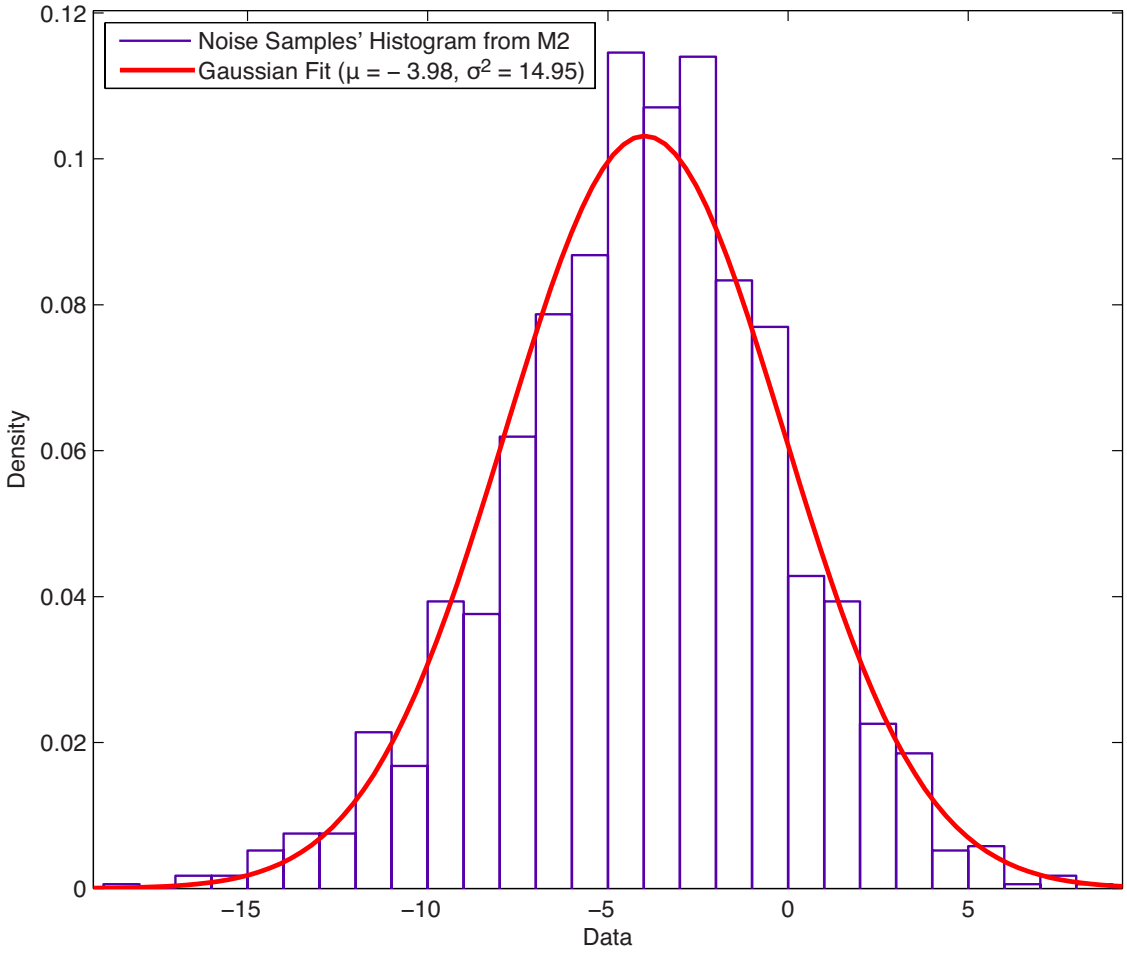


Figure 8.10: Histogram of the noise samples based on model M_2 and fitted Guassian probability density.

8.3.2 Noise Model Evaluation

To validate the noise estimation model of nonlinearity, two random noise processes are generated using

$$n_{M_1}(t) \approx \frac{N_{M_1}}{\sqrt{t}} \exp\Big(\frac{-b_{M_1}(d - c_{M_1}t)^2}{t}\Big), \tag{8.17}$$

$$n_{M_2}(t) \approx \frac{N_{M_2}}{\sqrt{t^3}} \exp\Big(\frac{-b_{M_2}(c_{M_2}t - d)^2}{t}\Big) \tag{8.18}$$

where N_{M_1} is the Gaussian random variable with mean $\mu = -0.7356$ and variance $\sigma^2 = 0.5214$, $b_{M_1} = 1.4306 \times 10^{-4}$ (average value of the corresponding row in Table 8.3), $c_{M_1} = 57.5018$ (average value of the corresponding row in Table 8.3), and $d = 225$ is the separation distance between the transmitter and the receiver. Similarly N_{M_2} is the Gaussian random variable with mean $\mu = -3.9811$ and variance $\sigma^2 = 14.9589$, $b_{M_2} = 1.57 \times 10^{-}4$ and $c_{M_2} = 35.021$. Using this noise process, 144 different samples for when both transmitters spray is generated using the sensor measurements from a single transmitter spray data using

$$\hat{h}_{i,j}^{(1,2)} = h_i^{(1)}(t) + h_j^{(2)}(t) + n(t), \tag{8.19}$$

where $\hat{h}_{i,j}^{(1,2)}$ is the estimated sample, $h_i^{(1)}(t)$, and $h_j^{(2)}(t)$, are sensor measurement from 12 different trials ($i, j = 1, 2, \ldots, 12$), and $n(t)$ is the noise process generated

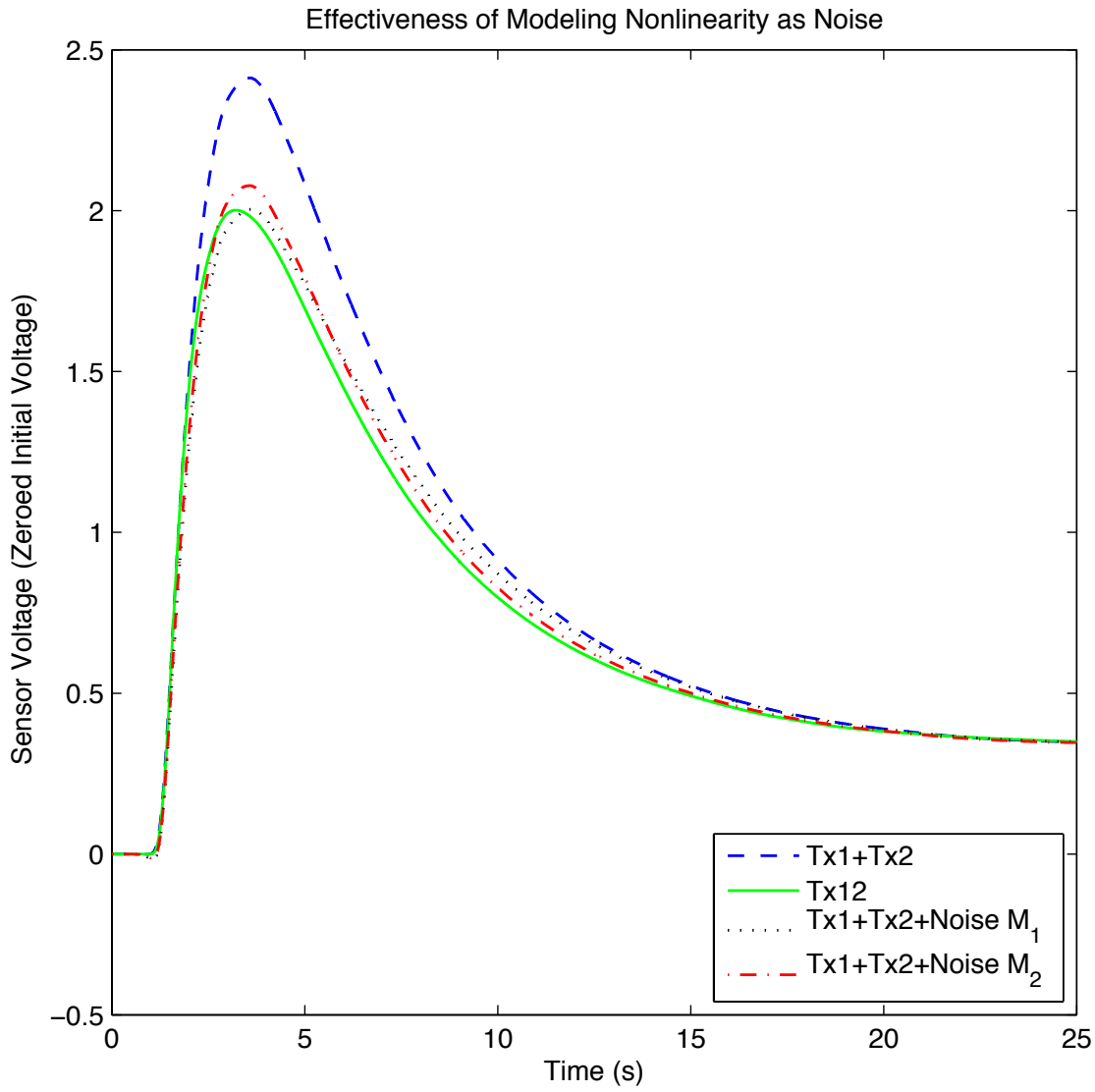


Figure 8.11: Effectiveness of modeling the nonlinearity as noise.

either using model M_1 (i.e. Equation 8.17) or model M_2 (i.e. Equation 8.18).

Figure 8.11 shows the results. The Tx12 plot shows the the average system response $h^{(1,2)}$ when both transmitters spray (averaged across 12 different trials).

The Tx1+Tx2 plot shows the average system response for $h^{(1)}(t) + h^{(2)}(t)$, the Tx1+Tx2+Noise M_1 plot shows the average $\hat{h}^{(1,2)}$ across the 144 samples when the noise model is based on M_1, and the Tx1+Tx2+Noise M_2 plot shows the red plot shows the average $\hat{h}^{(1,2)}$ across the 144 samples when the noise model is based on M_2. As can be seen from the plot, both noise models that are presented can effectively represent the nonlinearity which is present in the system. This is a significant result since the system can now be represented as a linear model with noise. The noise estimation is accurate.

9 Conclusions and Future Work

In this dissertation, I considered molecular communication systems from both theoretical and practical perspectives. I described that molecular communication is a new branch of communication engineering, where small information particles such as molecules are used to transfer data from a transmitter to a receiver.

In Chapter 2, I presented the physical, chemical, and biological processes that underlay molecular communication systems at both microscales and macroscales. I described what components are needed for the molecular communication transmitter and receiver. Moreover, I described different types of molecules and chemicals that could be used as carriers of information. I also discussed different types of transport mechanisms that could be employed by a molecular communication system, and presented the underlying mathematical models for these transport mechanisms. Finally, potential applications for molecular communication were discussed, where medical applications such as lab-on-a-chip devices and on-chip diagnostic devices were highlighted to be the main deriving force behind molecular commu-

nication. Chapter 3 discussed molecular communication from a communication engineering lens and a survey of recent works was presented.

Throughout the first two chapters, I showed that most previous works are theoretical in nature and have only focused on diffusion transport. Therefore, I identified two gaps in the literature: lack of theoretical work on active transport propagation schemes, and lack of any practical implementations of molecular communication systems.

To lessen these gaps, first, I focused on on-chip molecular communication. In Chapter 4, I developed a computer simulation environment for on-chip molecular communication using diffusion-based, flow-based, and active transport propagation schemes. I then used the simulator to compare the channel capacity of each transport mechanism. I showed that active transport can be very effective for on-chip applications. The simulation software that I developed in this chapter is the only simulator for molecular communication in confined environments, capable of modeling different propagation schemes. Previous simulators had only focused on diffusion-based molecular communication, mostly with infinite boundary conditions.

Chapter 5 presented different mathematical models for active transport molecular communication. First, a simple model was derived to gain a deeper understanding of on-chip active transport. Then, a more complicated, and more accurate

model was developed using Markov chains. It was shown that the derived models could speed up the simulation process significantly, and be used to solve optimal design problems. The models presented in this chapter are the first models developed for kinesin-microtubule-based molecular communication. These models could be used to solve various design problem, to calculate performance measures efficiently, and to improve the communication system.

Using these new mathematical models, in Chapter 6, I provided optimal design strategies for on-chip active transport molecular communication. First, I found the optimal location for the transmission area on the chip, and then I developed a strategy for finding the optimal size of the vesicles that encapsulate the information particles. Finally, I considered the problem of finding the optimal channel shape. In particular, I developed an optimization formulation that could be used to find the optimal channel shape and dimensions. Using the model, I showed that the optimal shape is the circular-shaped channel. Using the simulator, I showed that the optimal designs increase the channel capacity significantly.

In Chapter 7, I focused on the lack of any experimental implementations of molecular communication. I designed and built an inexpensive tabletop platform, from readily available and off-the-shelf parts, that is capable of transferring short text messages. The importance of this system is twofold. First, it proves that chemical signals could be used for reliable data transfer. Second, it provides re-

searchers with an experimental platform that they could use to test their theoretical models. Because the system is inexpensive, it could easily be replicated by other researchers. One of the important results that was presented in Chapter 7, was the nonlinearity of the system. Most theoretical works on molecular communication assume linearity. However, I show that in practical implementations, the system may be nonlinear.

Mathematical models for this experimental platform was presented in Chapter 8. To derive these models, I first showed that the experimental results obtained from the platform does not match the previously used models in the literature. I then proposed corrections to these models using the experimental data. The corrected models were then used to show that the nonlinearity of the system can be represented as additive noise. I showed that the nonlinearity could be then be removed. These were interesting and non-trivial results.

9.1 Future Research Directions

The field of molecular communication is still in its infancy. Therefore, there are many open problems for future research. Based on some of the results presented in this dissertation, I highlight some future research projects below.

1. In Chapter 4, I developed a simulation environment for on-chip molecular

communication. In the future, I would like to improve this simulation environment by including models for chemical reactions and chemical decay, non-deterministic loading mechanisms, and more realistic transmitter and receiver models. Also, other propagation schemes, such as bacterial shuttles could be added to the simulator.

2. As was shown in this dissertation, kinesin-microtubule active transport can be very effective for on-chip molecular communication. However, a laboratory demonstrator for such a communication system is non-existent. There are a number of works that use this form of transport for molecular assembly [89]. Therefore, it should be possible in the future to develop a laboratory demonstrator for this system.

3. There are similarities between optical channels and molecular communication channels. In optical communications photons are the information carrying particles, and there exists a wealth of work on modeling these channels. Therefore, the similarities and differences of these channels must be investigated in greater detail in the future.

4. A promising application for molecular communication, is chemical communication in robotics. Although there are a number of previous works that present some preliminary results on this topic [15, 166–168], a comprehensive

study of this area is lacking. For example, inspired by ants, small robots could use chemical trails and chemical tags to complete a predefined task very efficiently. Such a demonstrator would show the true potential of molecular communication in robotics.

5. In Chapter 7, I developed a general purpose molecular communication demonstrator that was capable of transmitting short text messages. This system achieved a data rate of 1/3 bits per second, over several meters. In the future, I would like to decrease the size of the system by one order of magnitude, and increase the data rate by one order of magnitude. One way this could be achieve is through miniaturizing the spray and the sensor, developing more sophisticated detection algorithms at the receiver, and using suitable error-correction codes. Ultimately, the goal is to shrink the system enough to demonstrate the feasibility of molecular communication in an environment analogous to human body.

6. As was shown in Chapter 8, practical molecular communication systems could be nonlinear in nature. However, most previous works have only focused on the linearity of the transport mechanisms such as diffusion. Therefore, nonlinear molecular communication systems must be studies in more details in the future. For example, as the information particles travel from the trans-

mitter to the receiver, chemical reactions in the environment may result in a nonlinear propagation. It may be possible to use the nonlinearity to improve system performance. For example, the nonlinearity present in some optical communication channels are used to enhance these systems.

Bibliography

[1] R. A. Freitas, *Nanomedicine, Volume I: Basic Capabilities*, 1st ed. Landes Bioscience, 1999.

[2] S. Sengupta, M. E. Ibele, and A. Sen, "Fantastic voyage: Designing self-powered nanorobots," *Angew. Chem. Int. Ed.*, vol. 51, no. 34, pp. 8434–8445, Aug. 2012.

[3] H. Wheeler, "Fundamental limitations of small antennas," *Proceedings of the IRE*, vol. 35, no. 12, pp. 1479–1484, Dec 1947.

[4] S. F. Bush, *Nanoscale Communication Networks*, 1st ed. Artech House, 2010.

[5] S. Fatikow and U. Rembold, *Microsystem technology and microrobotics*. Springer Science & Business Media, 2013.

[6] I. F. Akyildiz, F. Brunetti, and C. Blazquez, "Nanonetworks: A new communication paradigm," *Comput. Netw.*, vol. 52, no. 12, pp. 2260–2279, Aug. 2008.

[7] T. Nakano, A. W. Eckford, and T. Haraguchi, *Molecular communication*, 1st ed. Cambridge University Press, 2013.

[8] 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 Netw.*, vol. 8, no. 8, pp. 872–888, Nov. 2010.

[9] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter, *Molecular Biology of the Cell*, 5th ed. Garland Science, Nov. 2007.

[10] W. C. Agosta, *Chemical Communication: The Language of Pheromones*, 1st ed. W H Freeman & Co, Aug. 1992.

[11] M. S. Kuran, H. B. Yilmaz, T. Tugcu, and B. Ozerman, "Energy model for communication via diffusion in nanonetworks," *Nano Commun. Netw.*, vol. 1, no. 2, pp. 86–95, June 2010.

[12] I. S. M. Christopher Rose, "A fundamental framework for molecular communication channels: timing & payload," in *Proc. IEEE Int. Conf. on Commun. (ICC)*, 2015, to appear.

[13] S. Hiyama, Y. Moritani, T. Suda, R. Egashira, A. Enomoto, M. Moore, and T. Nakano, "Molecular communication," in *Proc. NSTI Nanotechnol. Conf. (NanoTech)*, 2005, pp. 391–394.

[14] R. A. Russell, *Odour detection by mobile robots*. Singapore; River Edge, NJ: World Scientific, 1999.

[15] M. Cole, Z. Racz, J. W. Gardner, and T. C. Pearce, "A novel biomimetic infochemical communication technology: From insects to robots," in *Proc. IEEE Sensors*, Oct. 2012, pp. 1–4.

[16] N. Farsad, A. W. Eckford, S. Hiyama, and Y. Moritani, "Quick system design of vesicle-based active transport molecular communication by using a simple transport model," *Nano Commun. Netw.*, vol. 2, no. 4, pp. 175–188, Dec. 2011.

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

[18] L. C. Cobo and I. F. Akyildiz, "Bacteria-based communication in nanonetworks," *Nano Commun. Netw.*, vol. 1, no. 4, pp. 244–256, Dec. 2010.

[19] N.-R. Kim and C.-B. 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.

[20] A. W. Eckford, "Nanoscale communication with brownian motion," in *Proc. Conf. on Inf. Sci. and Syst. (CISS)*, Baltimore, MD, 2007, pp. 160–165.

[21] ——, "Timing information rates for active transport molecular communication," in *Nano-Net*, ser. Lect. Notes Inst. Comput. Sci., Social Inf. and Telecommu. Eng., 2009, vol. 20, pp. 24–28.

[22] 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.

[23] S. Hiyama and Y. Moritani, "Molecular communication: harnessing biochemical materials to engineer biomimetic communication systems," *Nano Commun. Netw.*, vol. 1, no. 1, pp. 20–30, Mar. 2010.

[24] T. Nakano, M. J. Moore, F. Wei, A. V. Vasilakos, and J. Shuai, "Molecular communication and networking: Opportunities and challenges," *IEEE Trans. NanoBiosci.*, vol. 11, no. 2, pp. 135–148, June 2012.

[25] S. Kadloor and R. Adve, "A framework to study the molecular communication system," in *Proc. IEEE Int. Conf. on Comput. Commun. and Netw. (ICCCN)*, Aug. 2009, pp. 1–6.

[26] M. S. Kuran, T. Tugcu, and B. Edis Ozerman, "Calcium signaling: overview and research directions of a molecular communication paradigm," *IEEE Trans. Wireless Commun.*, vol. 19, no. 5, pp. 20–27, Oct. 2012.

[27] K. Darchini and A. S. Alfa, "Molecular communication via microtubules and physical contact in nanonetworks: A survey," *Nano Commun. Netw.*, vol. 4, no. 2, pp. 73–85, June 2013.

[28] M. Gregori and I. F. Akyildiz, "A new nanonetwork architecture using flagellated bacteria and catalytic nanomotors," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 4, pp. 612–619, May 2010.

[29] N. Farsad, H. B. Yilmaz, A. Eckford, C.-B. Chae, and W. Guo, "A comprehensive survey of recent advancements in molecular communication," *IEEE Commun. Surveys Tuts.*, pp. 1–29, 2015, in preparation for submission.

[30] N. Farsad, A. W. Eckford, S. Hiyama, and Y. Moritani, "Information rates of active propagation in microchannel molecular communication," in *Proc. Int. Conf. on Bio-Inspired Models of Netw., Inf. and Comput. Syst. (BIONETICS)*, Boston, MA, 2010, p. 7.

[31] A. W. Eckford, N. Farsad, S. Hiyama, and Y. Moritani, "Microchannel molecular communication with nanoscale carriers: Brownian motion versus active transport," in *Proc. IEEE Int. Conf. on Nanotechnol. (IEEE-NANO)*, Seoul, South Korea, 2010, pp. 854 – 858.

[32] N. Farsad, A. W. Eckford, S. Hiyama, and Y. Moritani, "On-chip molecular communication: Analysis and design," *IEEE Trans. NanoBiosci.*, vol. 11, no. 3, pp. 304–314, Sep. 2012.

[33] N. Farsad, A. W. Eckford, and S. Hiyama, "A Markov chain channel model for active transport molecular communication," *IEEE Trans. Signal Process.*, vol. 62, no. 9, pp. 2424–2436, May 2014.

[34] N. Farsad, A. W. Eckford, S. Hiyama, and Y. Moritani, "A simple mathematical model for information rate of active transport molecular communication," in *Proc. IEEE Int. Conf. on Comput. Commun. Workshops (INFOCOM WKSHPS)*, Shanghai, P. R. China, 2011, pp. 473–478.

[35] N. Farsad, A. W. Eckford, and S. Hiyama, "Modelling and design of polygon-shaped kinesin substrates for molecular communication," in *Proc. IEEE Int. Conf. on Nanotechnol. (IEEE-NANO)*, Aug. 2012, pp. 1–5.

[36] ——, "Channel design and optimization of active transport molecular communication," in *Proc. Int. Conf. on Bio-Inspired Models of Netw., Inf. and Comput. Syst. (BIONETICS)*, York, England, 2011.

[37] ——, "A mathematical channel optimization formula for active transport molecular communication," in *Proc. IEEE Int. Conf. on Commun. Workshops (ICC WKSHPS)*, Ottawa, Canada, 2012.

[38] N. Farsad, A. Eckford, and S. Hiyama, "Design and optimizing of on-chip kinesin substrates for molecular communication," *IEEE Trans. Nanotechnol.*, vol. 14, no. 4, pp. 699–708, July 2015.

[39] N. Farsad, W. Guo, and A. W. Eckford, "Tabletop molecular communication: Text messages through chemical signals," *PLoS ONE*, vol. 8, no. 12, p. e82935, Dec. 2013.

[40] 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 Commun. (ICC)*, 2014.

[41] N. Farsad, W. Guo, and A. W. Eckford, "Molecular communication link," in *Proc. IEEE Int. Conf. on Comput. Commun. Workshops (INFOCOM WKSHPS)*, 2014, pp. 107–108.

[42] M. E. Ortiz and D. Endy, "Engineered cell-cell communication via DNA messaging," *J. Biol. Eng.*, vol. 6, no. 1, p. 16, Sep. 2012.

[43] 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 Commun. (ICC)*, 2014.

[44] N. Farsad, N.-R. Kim, A. W. Eckford, and C.-B. Chae, "Channel and noise models for nonlinear molecular communication systems," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 12, pp. 2392–2401, Dec 2014.

[45] 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 Commun. (ICC)*, 2015, to appear.

[46] W. Guo, N. Farsad, J.-L. Wu, and A. Eckford, "Random walk through thin obstacles with gaps: Approximate hitting distribution," *IEEE Commun. Lett.*, pp. 1–4, 2015, submitted.

[47] W. Guo, C. Mias, N. Farsad, and J. Wu, "Molecular versus electromagnetic wave propagation loss in macro-scale environments," *IEEE Trans. Mol. Biol. Multi-Scale Commun.*, pp. 1–21, 2015, to appear.

[48] S. Qiu, W. Guo, M. Leeson, S. Wang, N. Farsad, and A. W. Eckford, "Nanoparticle communications: from chemical signals in nature to wireless sensor networks," *Nanotechnology Perceptions*, vol. 10, no. 1, pp. 1–13, 2014.

[49] N. Farsad, W. Guo, C.-B. Chae, and A. W. Eckford, "Stable Distributions as Noise Models for Molecular Communication," in *Proc. IEEE Glob. Commun. Conf. (GLOBECOM)*, 2015, to appear.

[50] L. Wang, N. Farsad, W. Guo, S. Magierowski, and A. W. Eckford, "Molecular Barcodes: Information Transmission via Persistent Chemical Tags," in *Proc. IEEE Int. Conf. on Commun. (ICC)*, 2015, to appear.

[51] S. Qiu, N. Farsad, T. Dong, A. W. Eckford, and W. Guo, "Under-Water Molecular Signalling: a Hidden Transmitter and Absent Receivers Problem," in *Proc. IEEE Int. Conf. on Commun. (ICC)*, 2015, to appear.

[52] C. Lee, B. Koo, N.-R. Kim, H. B. Yilmaz, N. Farsad, A. Eckford, and C.-B. Chae, "Molecular MIMO communication link," in *Proc. IEEE Int. Conf. on Comput. Commun. (INFOCOM)*, 2015.

[53] B. Alberts, D. Bray, K. Hopkin, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter, *Essential Cell Biology*, 3rd ed. Garland, 2009.

[54] T. D. Pollard, W. C. Earnshaw, and J. Lippincott-Schwartz, *Cell Biology*, 2nd ed. Saunders, 2007.

[55] R. P. Feynman, "Plenty of room at the bottom," *Engineering and Science (Caltech)*, vol. 23, no. 5, pp. 22–36, Dec. 1959. [Online]. Available: http://resolver.caltech.edu/CaltechES:23.5.1960Bottom

[56] K. Jensen, J. Weldon, H. Garcia, and A. Zettl, "Nanotube radio," *ACS Nano Lett.*, vol. 7, no. 11, pp. 3508–3511, Oct. 2007.

[57] J. F. Nagle and S. Tristram-Nagle, "Structure of lipid bilayers," *Biochim. Biophys. Acta*, vol. 1469, no. 3, pp. 159–195, Aug. 2000.

[58] C. Huang and J. T. Mason, "Geometric packing constraints in egg phosphatidylcholine vesicle," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 75, no. 1, pp. 308–310, Jan. 1978.

[59] S. Matosevic and B. M. Paegel, "Stepwise synthesis of giant unilamellar vesicles on a microfluidic assembly line," *J. Am. Chem. Soc.*, vol. 133, no. 9, pp. 2798–2800, Feb. 2011.

[60] D. A. LaVan, T. McGuire, and R. Langer, "Small-scale systems for in vivo drug delivery," *Nat. Biotechnol.*, vol. 21, no. 10, pp. 1184–1191, Oct. 2003.

[61] R. A. Laine, "Information capacity of the carbohydrate code," *Pure & Appl. Chem.*, vol. 69, no. 9, pp. 1867–1874, Jan. 1997.

[62] "Transforming synthesis, enabling science: a roadmap for synthesis in the 21st century," *UK Engineering and Physical Science Research Council (EPSRC)s Grand Challenge Network: Dial-a-Molecule, Technical Report*, Oct. 2012.

[63] T. Nakano, T. Suda, T. Koujin, T. Haraguchi, and Y. Hiraoka, "Molecular communication through gap junction channels," in *Trans. Comput. Syst. Biol. X*, ser. Lect. Notes Comput. Sci., C. Priami, F. Dressler, O. B. Akan, and A. Ngom, Eds., 2008, vol. 5410, pp. 81–99.

[64] M. Pierobon and I. F. Akyildiz, "A physical end-to-end model for molecular communication in nanonetworks," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 4, pp. 602–611, May 2010.

[65] ——, "Capacity of a diffusion-based molecular communication system with channel memory and molecular noise," *IEEE Trans. Inf. Theory*, vol. 59, no. 2, pp. 942–954, Feb. 2013.

[66] K. V. Srinivas, A. W. Eckford, and R. S. Adve, "Molecular communication in fluid media: The additive inverse Gaussian noise channel," *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4678–4692, July 2012.

[67] S. Hiyama, Y. Moritani, R. Gojo, S. Takeuchi, and K. Sutoh, "Biomolecular-motor-based autonomous delivery of lipid vesicles as nano- or microscale reactors on a chip," *RSC Lab on a Chip*, vol. 10, no. 20, pp. 2741–2748, 2010.

[68] P. Lio and S. Balasubramaniam, "Opportunistic routing through conjugation in bacteria communication nanonetwork," *Nano Commun. Netw.*, vol. 3, no. 1, pp. 36–45, Mar. 2012.

[69] M. J. Moore, T. Suda, and K. Oiwa, "Molecular communication: modeling noise effects on information rate," *IEEE Trans. NanoBiosci.*, vol. 8, no. 2, pp. 169–180, June 2009.

[70] A. Enomoto, M. J. Moore, T. Suda, and K. Oiwa, "Design of self-organizing microtubule networks for molecular communication," *Nano Commun. Netw.*, vol. 2, no. 1, pp. 16–24, Mar. 2011.

[71] M. S. Kuran, H. B. Yilmaz, and T. Tugcu, "A tunnel-based approach for signal shaping in molecular communication," in *Proc. IEEE Int. Conf. on Commun. Workshops (ICC WKSHPS)*, June 2013, pp. 776–781.

[72] H. C. Berg, *Random walks in biology*. Princeton, N.J.: Princeton University Press, 1993.

[73] J. Berthier and P. Silberzan, *Microfluidics for Biotechnology*, 2nd ed. Artech House, Dec. 2009.

[74] B. Oksendal, *Stochastic Differential Equations: An Introduction with Applications*, 6th ed. Springer, 2003.

[75] H. J. V. Tyrrell and K. R. Harris, *Diffusion in Liquids: A Theoretical and Experimental Study (Monographs in Chemistry)*. Butterworth-Heinemann, Mar. 1984.

[76] "Bionumbers: the database of useful biological numbers," http://bionumbers.hms.harvard.edu/search.aspx?log=y&task=searchbytrmorg&trm=diffusion+coefficient+in+water, accessed: 2014-08-18.

[77] P. Cuatrecasas, "Membrane receptors," *Annu. Rev. Biochem.*, vol. 43, no. 1, pp. 169–214, July 1974.

[78] 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, June 2014.

[79] D. L. L. Dy and J. P. Esguerra, "First-passage-time distribution for diffusion through a planar wedge," *Phys. Rev. E.*, vol. 78, no. 6, p. 062101, Dec. 2008.

[80] S. Hiyama, R. Gojo, T. Shima, S. Takeuchi, and K. Sutoh, "Biomolecular-motor-based nano- or microscale particle translocations on DNA microarrays," *ACS Nano Lett.*, vol. 9, no. 6, pp. 2407–2413, Apr. 2009.

[81] J. Howard, A. J. Hudspeth, and R. D. Vale, "Movement of microtubules by single kinesin molecules," *Nature*, vol. 342, no. 6246, pp. 154–158, Nov. 1989.

[82] I. Dujovne, M. Van Den Heuvel, Y. Shen, M. De Graaff, and C. Dekker, "Velocity modulation of microtubules in electric fields," *ACS Nano Lett.*, vol. 8, no. 12, pp. 4217–4220, Oct. 2008.

[83] E. Kim, K.-E. Byun, D. S. Choi, D. J. Lee, D. H. Cho, B. Y. Lee, H. Yang, J. Heo, H.-J. Chung, S. Seo, and S. Hong, "Electrical control of kinesin-microtubule motility using a transparent functionalized-graphene substrate," *Nanotechnology*, vol. 24, no. 19, p. 195102, May 2013.

[84] T. Nitta, A. Tanahashi, M. Hirano, and H. Hess, "Simulating molecular shuttle movements: Towards computer-aided design of nanoscale transport systems," *RSC Lab on a Chip*, vol. 6, no. 7, pp. 881–885, May 2006.

[85] T. Nitta, A. Tanahashi, and M. Hirano, "In silico design and testing of guiding tracks for molecular shuttles powered by kinesin motors," *RSC Lab on a Chip*, vol. 10, no. 11, pp. 1447–1453, June 2010.

[86] J. Howard, *Mechanics of motor proteins and the cytoskeleton.* Sunderland Mass. USA: Sinauer Associates, 2001.

[87] K.-E. Byun, K. Heo, S. Shim, H.-J. Choi, and S. Hong, "Functionalization of silicon nanowires with actomyosin motor protein for bioinspired nanomechanical applications," *Small*, vol. 5, no. 23, pp. 2659–2664, Dec. 2009.

[88] J. Ikuta, N. K. Kamisetty, H. Shintaku, H. Kotera, T. Kon, and R. Yokokawa, "Tug-of-war of microtubule filaments at the boundary of a kinesin- and dynein-patterned surface," *Nat. Sci. Rep.*, vol. 4, June 2014.

[89] D. Steuerwald, S. M. Fruh, R. Griss, R. D. Lovchik, and V. Vogel, "Nanoshuttles propelled by motor proteins sequentially assemble molecular cargo in a microfluidic device," *RSC Lab on a Chip*, vol. 14, pp. 3729–3738, June 2014.

[90] P. D. Frymier, R. M. Ford, and P. T. Cummings, "Cellular dynamics simulations of bacterial chemotaxis," *Chem. Eng. Sci.*, vol. 48, no. 4, pp. 687–699, Feb. 1993.

[91] M. Gregori, I. Llatser, A. Cabellos-Aparicio, and E. Alarcón, "Physical channel characterization for medium-range nanonetworks using flagellated bacteria," *Comput. Netw.*, vol. 55, no. 3, pp. 779–791, Feb. 2011.

[92] M. Kang and H. G. Othmer, "Spatiotemporal characteristics of calcium dynamics in astrocytes," *Chaos: Interdiscipl. J. Nonlinear Sci.*, vol. 19, no. 3, pp. 037 116–1 – 037 116–21, Sep. 2009.

[93] E. Scemes and C. Giaume, "Astrocyte calcium waves: what they are and what they do," *Glia*, vol. 54, no. 7, pp. 716–725, Nov. 2006.

[94] S. Schuster, M. Marhl, and T. Höfer, "Modelling of simple and complex calcium oscillations," *Eur. J. Biochem.*, vol. 269, no. 5, pp. 1333–1355, Mar. 2002.

[95] T. Nakano, T. Suda, M. Moore, and R. Egashira, "Molecular communication for nanomachines using intercellular calcium signalling," in *Proc. NSTI Nanotechnol. Conf. (NanoTech)*, 2005, pp. 478–481.

[96] L. J. Elias and D. M. Saucier, *Neuropsychology: Clinical and experimental foundations*. Pearson/Allyn & Bacon, 2006.

[97] R. D. Keynes and D. J. Aidley, *Nerve and Muscle*. Cambridge University Press, 2001.

[98] L. You, R. S. Cox, R. Weiss, and F. H. Arnold, "Programmed population control by cell-cell communication and regulated killing," *Nature*, vol. 428, no. 6985, pp. 868–871, Apr. 2004.

[99] S. Basu, Y. Gerchman, C. H. Collins, F. H. Arnold, and R. Weiss, "A synthetic multicellular system for programmed pattern formation," *Nature*, vol. 434, no. 7037, pp. 1130–1134, Apr. 2005.

[100] M.-T. Chen and R. Weiss, "Artificial cell-cell communication in yeast saccharomyces cerevisiae using signaling elements from arabidopsis thaliana," *Nat. Biotechnol.*, vol. 23, no. 12, pp. 1551–1555, Dec. 2005.

[101] M. J. Doktycz and M. L. Simpson, "Nano-enabled synthetic biology," *Mol. Syst. Biol.*, vol. 3, July 2007.

[102] Y. Sasaki, Y. Shioyama, W.-J. Tian, J.-I. Kikuchi, S. Hiyama, Y. Moritani, and T. Suda, "A nanosensory device fabricated on a liposome for detection of chemical signals," *Biotechnol. Bioeng.*, vol. 105, no. 1, pp. 37–43, Jan. 2010.

[103] M. Mukai, K. Maruo, J.-I. Kikuchi, Y. Sasaki, S. Hiyama, Y. Moritani, and T. Suda, "Propagation and amplification of molecular information using a photoresponsive molecular switch," *Supramol. Chem.*, vol. 21, no. 3/4, pp. 284–291, Apr. 2009.

[104] T. S. Moon, C. Lou, A. Tamsir, B. C. Stanton, and C. A. Voigt, "Genetic programs constructed from layered logic gates in single cells," *Nature*, vol. 491, no. 7423, pp. 249–253, Nov. 2012.

[105] P. Siuti, J. Yazbek, and T. K. Lu, "Synthetic circuits integrating logic and memory in living cells," *Nat. Biotechnol.*, vol. 31, no. 5, pp. 448–452, May 2013.

[106] G. M. Cooper and R. E. Hausman, *The cell*. Sinauer Associates Sunderland, 2000.

[107] S. Huang and M. Kamihira, "Development of hybrid viral vectors for gene therapy," *Biotechnol. Adv.*, vol. 31, no. 2, pp. 208–223, Apr. 2013.

[108] A. Gossler, T. Doetschman, R. Korn, E. Serfling, and R. Kemler, "Transgenesis by means of blastocyst-derived embryonic stem cell lines," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 83, no. 23, pp. 9065–9069, Dec. 1986.

[109] J. Beal, T. Lu, and R. Weiss, "Automatic compilation from high-level biologically-oriented programming language to genetic regulatory networks," *PLoS ONE*, vol. 6, no. 8, p. e22490, Aug. 2011.

[110] P. D. Robbins and S. C. Ghivizzani, "Viral vectors for gene therapy," *Pharmacol. & Ther.*, vol. 80, no. 1, pp. 35–47, Oct. 1998.

[111] J. E. Toettcher, C. Mock, E. Batchelor, A. Loewer, and G. Lahav, "A synthetic-natural hybrid oscillator in human cells," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 107, no. 39, pp. 17 047–17 052, Sep. 2010.

[112] O. Mondragon-Palomino, T. Danino, J. Selimkhanov, L. Tsimring, and J. Hasty, "Entrainment of a population of synthetic genetic oscillators," *Science*, vol. 333, no. 6047, pp. 1315–1319, Sep. 2011.

[113] D. Hymel and B. R. Peterson, "Synthetic cell surface receptors for delivery of therapeutics and probes," *Adv. Drug Deliv. Rev.*, vol. 64, no. 9, pp. 797–810, June 2012.

[114] H. Shankaran, H. Resat, and H. S. Wiley, "Cell surface receptors for signal transduction and ligand transport: A design principles study," *PLoS Comput. Biol.*, vol. 3, no. 6, p. e101, June 2007.

[115] M. Eshaghian-Wilner, *Bio-Inspired and Nanoscale Integrated Computing*, ser. Nature-Inspired Computing Series. Wiley, 2009.

[116] S. Das, A. J. Gates, H. A. Abdu, G. S. Rose, C. A. Picconatto, and J. C. Ellenbogen, "Designs for ultra-tiny, special-purpose nanoelectronic circuits," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 54, no. 11, pp. 2528–2540, Nov. 2007.

[117] H. P. Rang, "The receptor concept: pharmacology's big idea," *Br. J. Pharmacol.*, vol. 147, no. S1, pp. S9–S16, Jan. 2006.

[118] G. M. Patel, G. C. Patel, R. B. Patel, J. K. Patel, and M. Patel, "Nanorobot: A versatile tool in nanomedicine," *J. Drug Target.*, vol. 14, no. 2, pp. 63–67, 2006.

[119] T. E. Mallouk and A. Sen, "Powering nanorobots," *Nat. Sci. Am.*, vol. 300, no. 5, pp. 72–77, May 2009.

[120] A. A. Solovev, W. Xi, D. H. Gracias, S. M. Harazim, C. Deneke, S. Sanchez, and O. G. Schmidt, "Self-propelled nanotools," *Nano*, vol. 6, no. 2, pp. 1751–1756, Jan. 2012.

[121] W. Gao, S. Sattayasamitsathit, K. M. Manesh, D. Weihs, and J. Wang, "Magnetically powered flexible metal nanowire motors," *J. Am. Chem. Soc.*, vol. 132, no. 41, pp. 14 403–14 405, Sep. 2010.

[122] S. Sanchez, A. A. Solovev, S. M. Harazim, and O. G. Schmidt, "Microbots swimming in the flowing streams of microfluidic channels," *J. Am. Chem. Soc.*, vol. 133, no. 4, pp. 701–703, Dec. 2011.

[123] Y. Moritani, S. Hiyama, and T. Suda, "Molecular communication for health care applications," in *Proc. IEEE Int. Conf. on Pervasive Comput. and Commun. Workshops (PerCom WKSHPS)*, Pisa, Italy, 2006, p. 5.

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

[125] H. Kitano, "Computational systems biology," *Nature*, vol. 420, no. 6912, pp. 206–210, Nov. 2002.

[126] D. Noble, "The rise of computational biology," *Nat. Rev. Mol. Cell Biol.*, vol. 3, no. 6, pp. 459–463, June 2002.

[127] D. Malak and O. B. Akan, "Molecular communication nanonetworks inside human body," *Nano Commun. Netw.*, vol. 3, no. 1, pp. 19–35, Mar. 2012.

[128] 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 Nanotechnol. (IEEE-NANO)*, Aug. 2011, pp. 313–318.

[129] A. A. G. Requicha, "Nanorobots, NEMS, and nanoassembly," *Proc. IEEE*, vol. 91, no. 11, pp. 1922–1933, Nov. 2003.

[130] B. Atakan, O. B. Akan, and S. Balasubramaniam, "Body area nanonetworks with molecular communications in nanomedicine," *IEEE Commun. Mag.*, vol. 50, no. 1, pp. 28–34, Jan. 2012.

[131] H. Craighead, "Future lab-on-a-chip technologies for interrogating individual molecules," *Nature*, vol. 442, no. 7101, pp. 387–393, July 2006.

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

[133] 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, July 2006.

[134] O. Veiseh, J. W. Gunn, and M. Zhang, "Design and fabrication of magnetic nanoparticles for targeted drug delivery and imaging," *Adv. Drug Deliv. Rev.*, vol. 62, no. 3, pp. 284–304, Mar. 2010.

[135] K. Kostarelos, "Nanorobots for medicine: how close are we?" *Nanomedicine*, vol. 5, no. 3, pp. 341–342, 2010.

[136] S. P. Leary, C. Y. Liu, and M. L. J. Apuzzo, "Toward the emergence of nanoneurosurgery: Part III–Nanomedicine: Targeted nanotherapy, nanosurgery, and progress toward the realization of nanoneurosurgery," *Neurosurgery*, vol. 58, no. 6, pp. 1009–1026; discussion 1009–1026, June 2006.

[137] P. Couvreur and C. Vauthier, "Nanotechnology: Intelligent design to treat complex disease," *Pharmaceut. Res.*, vol. 23, no. 7, pp. 1417–1450, June 2006.

[138] A. Cavalcanti and T. Hogg, "Nanorobot communication techniques: A comprehensive tutorial," in *Proc. IEEE Int. Conf. on Control, Autom., Robot. and Vision (ICARCV)*, 2006, pp. 1–6.

[139] 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," *Nat. Mater.*, vol. 10, no. 7, pp. 545–552, July 2011.

[140] A. Cavalcanti, B. Shirinzadeh, T. Fukuda, and S. Ikeda, "Nanorobot for brain aneurysm," *SAGE Int. J. Robot. Res.*, vol. 28, no. 4, pp. 558–570, Apr. 2009.

[141] 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.

[142] J. Elbaz and I. Willner, "DNA origami: Nanorobots grab cellular control," *Nat. Mater.*, vol. 11, no. 4, pp. 276–277, Mar. 2012.

[143] A. Cavalcanti and R. A. Freitas-Jr., "Nanorobotics control design: a collective behavior approach for medicine," *IEEE Trans. NanoBiosci.*, vol. 4, no. 2, pp. 133–140, June 2005.

[144] R. Weiss and T. Knight, "Engineered communications for microbial robotics," in *DNA Comput.*, ser. Lect. Notes Comput. Sci., 2001, vol. 2054, pp. 1–16.

[145] S. Martel, M. Mohammadi, O. Felfoul, Z. Lu, and P. Pouponneau, "Flagellated magnetotactic bacteria as controlled mri-trackable propulsion and steering systems for medical nanorobots operating in the human microvasculature," *SAGE Int. J. Robot. Res.*, vol. 28, no. 4, pp. 571–582, Apr. 2009.

[146] L. Parcerisa Gine and I. F. Akyildiz, "Molecular communication options for long range nanonetworks," *Comput. Netw.*, vol. 53, no. 16, pp. 2753–2766, Nov. 2009.

[147] J. Crank, *The Mathematics of Diffusion*, 2nd ed. Oxford University Press, 1975.

[148] S. Redner, *A guide to first-passage processes.* Cambridge University Press, 2001.

[149] I. Llatser, E. Alarcon, and M. Pierobon, "Diffusion-based channel characterization in molecular nanonetworks," in *Proc. IEEE Int. Conf. on Comput. Commun. Workshops (INFOCOM WKSHPS)*, 2011, pp. 467–472.

[150] S. Wang, W. Guo, and M. D. McDonnell, "Transmit pulse shaping for molecular communications," in *Proc. IEEE Int. Conf. on Comput. Commun. (INFOCOM)*, 2014.

[151] W. Hundsdorfer and J. G. Verwer, *Numerical solution of time-dependent advection-diffusion-reaction equations*. Springer, 2003, vol. 33.

[152] I. Javandel, C. Doughty, and C. F. Tsang, *Groundwater transport: Handbook of mathematical models*. American Geophysical Union, 1984, vol. 10.

[153] E. L. Cussler, *Diffusion: mass transfer in fluid systems*. Cambridge University Press, 2009.

[154] G. T. Csanady, *Turbulent diffusion in the environment*, 3rd ed. Springer, 1973.

[155] A. Guha, "Transport and deposition of particles in turbulent and laminar flow," *Annu. Rev. Fluid Mech.*, vol. 40, no. 1, pp. 311–341, Jan. 2008.

[156] C. Zoppou and J. H. Knight, "Analytical solution of a spatially variable coefficient advection-diffusion equation in up to three dimensions," *Appl. Math. Model.*, vol. 23, no. 9, pp. 667–685, Sep. 1999.

[157] K. Sakai and W. R. Peltier, *The Influence of Deep Ocean Diffusivity on the Temporal Variability of the Thermohaline Circulation*. American Geophysical Union, 2013, pp. 227–242.

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

[159] L. Munoz, 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, p. e47751, Nov. 2012.

[160] V. E. Bochenkov and G. B. Sergeev, "Sensitivity, selectivity, and stability of gas-sensitive metal-oxide nanostructures," *Metal Oxide Nanostruct. Their Appl.*, vol. 3, pp. 31–52, 2010.

[161] M. Cole, J. W. Gardner, Z. Racz, S. Pathak, A. Guerrero, L. Muoz, G. Carot, T. C. Pearce, J. Challiss, D. Markovic, B. S. Hansson, S. Olsson, L. Kubler, J. G. E. Gardeniers, N. Dimov, and W. Bula, "Biomimetic insect infochemical communication system," in *Proc. IEEE Sensors*, 2009, pp. 1358–1361.

[162] A. E. Forooshani, S. Bashir, D. G. Michelson, and S. Noghanian, "A survey of wireless communications and propagation modeling in underground mines," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 4, pp. 1524–1545, Mar. 2013.

[163] R. K. Vander Meer, M. D. Breed, K. E. Espelie, and M. L. Winston, *Pheromone communication in social insects*. Westview Press Colorado, 1998.

[164] S. B. Olsson, L. S. Kuebler, D. Veit, K. Steck, A. Schmidt, M. Knaden, and B. S. Hansson, "A novel multicomponent stimulus device for use in olfactory experiments," *J. Neurosci. Meth.*, vol. 195, no. 1, pp. 1–9, Jan. 2011.

[165] R. A. Russell, "An odour sensing robot draws inspiration from the insect world," in *Proc. Int. Conf. on Bioelectromagnetism (ICBEM)*, Feb 1998, pp. 49–50.

[166] Y. Kuwana, S. Nagasawa, I. Shimoyama, and R. Kanzaki, "Synthesis of the pheromone-oriented behaviour of silkworm moths by a mobile robot with moth antennae as pheromone sensors," *Biosens. Bioelectron.*, vol. 14, no. 2, pp. 195–202, Feb. 1999.

[167] A. H. Purnamadjaja and R. A. Russell, "Pheromone communication in a robot swarm: Necrophoric bee behaviour and its replication," *Robotica*, vol. 23, no. 6, pp. 731–742, Nov. 2005.

[168] ——, "Bi-directional pheromone communication between robots," *Robotica*, vol. 28, no. 01, pp. 69–79, Apr. 2010.

[169] S. Kazadi, R. Goodman, D. Tsikata, D. Green, and H. Lin, "An autonomous water vapor plume tracking robot using passive resistive polymer sensors," *Auton. Robot.*, vol. 9, no. 2, pp. 175–188, Sep. 2000.

[170] H. Ishida, T. Nakamoto, T. Moriizumi, T. Kikas, and J. Janata, "Plume-tracking robots: A new application of chemical sensors," *Biol. Bull.*, vol. 200, no. 2, pp. 222–226, Apr. 2001.

[171] A. Lilienthal, A. Zell, M. Wandel, and U. Weimar, "Sensing odour sources in indoor environments without a constant airflow by a mobile robot," in *Proc. IEEE Int. Conf. on Robot. and Autom. (ICRA)*, vol. 4, 2001, pp. 4005–4010.

[172] S. Larionova, N. Almeida, L. Marques, and A. T. Almeida, "Olfactory coordinated area coverage," *Auton. Robot.*, vol. 20, no. 3, pp. 251–260, June 2006.

[173] W. Li, J. A. Farrell, S. Pang, and R. M. Arrieta, "Moth-inspired chemical plume tracing on an autonomous underwater vehicle," *IEEE Trans. Robot.*, vol. 22, no. 2, pp. 292–307, Apr. 2006.

[174] P. Sousa, L. Marques, and A. T. de Almeida, "Toward chemical-trail following robots," in *Proc. IEEE Int. Conf. on Mach. Learn. and Appl. (ICMLA)*, Dec 2008, pp. 489–494.

[175] A. R. G. Ramirez, A. B. Rodriguez, A. L. Lopez, D. W. Bertol, and A. D. C. de Albornoz, "Environmental odor perception: an evaluation of aplatform based on LabVIEW and the LEGO NXT," in *Proc. IEEE ISSNIP Biosignals and Biorobot. Conf.*, 2010, pp. 1–5.

[176] A. D. C. de Albornoz, A. B. Rodriguez, A. L. Lopez, and A. R. G. Ramirez, "A microcontroller-based mobile robotic platform for odor detection," in *Proc. IEEE ISSNIP Biosignals and Biorobot. Conf.*, Jan 2012, pp. 1–6.

[177] M. S. Kuran, H. B. Yilmaz, T. Tugcu, and I. F. Akyildiz, "Modulation techniques for communication via diffusion in nanonetworks," in *Proc. IEEE Int. Conf. on Commun. (ICC)*, June 2011, pp. 1–5.

[178] ——, "Interference effects on modulation techniques in diffusion based nanonetworks," *Nano Commun. Netw.*, vol. 3, no. 1, pp. 65–73, Mar. 2012.

[179] N. Garralda, I. Llatser, A. Cabellos-Aparicio, E. Alarcón, and M. Pierobon, "Diffusion-based physical channel identification in molecular nanonetworks," *Nano Commun. Netw.*, vol. 2, no. 4, pp. 196–204, Dec. 2011.

[180] Y.-P. Hsieh, Y.-C. Lee, P.-J. Shih, P.-C. Yeh, and K.-C. Chen, "On the asynchronous information embedding for event-driven systems in molecular communications," *Nano Commun. Netw.*, vol. 4, no. 1, pp. 2 – 13, Mar. 2013.

[181] B. Tepekule, A. E. Pusane, H. B. Yilmaz, and T. Tugcu, "Energy efficient ISI mitigation for communication via diffusion," in *Proc. IEEE Int. Black Sea Conf. on Commun. and Netw. (BlackSeaCom)*, 2014, pp. 33–37.

[182] B. Atakan, S. Galmes, and O. B. Akan, "Nanoscale communication with molecular arrays in nanonetworks," *IEEE Trans. NanoBiosci.*, vol. 11, no. 2, pp. 149–160, June 2012.

[183] H. Arjmandi, A. Gohari, M. N. Kenari, and F. Bateni, "Diffusion-based nanonetworking: A new modulation technique and performance analysis," *IEEE Commun. Lett.*, vol. 17, no. 4, pp. 645–648, Apr. 2013.

[184] I. Llatser, A. Cabellos-Aparicio, M. Pierobon, and E. Alarcon, "Detection techniques for diffusion-based molecular communication," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 12, pp. 726–734, Dec. 2013.

[185] A. Noel, K. Cheung, and R. Schober, "Improving receiver performance of diffusive molecular communication with enzymes," *IEEE Trans. NanoBiosci.*, vol. 13, no. 1, pp. 31–43, March 2014.

[186] C. T. Chou, "Molecular circuits for decoding frequency coded signals in nano-communication networks," *Nano Commun. Netw.*, vol. 3, no. 1, pp. 46–56, Mar. 2012.

[187] H. ShahMohammadian, G. G. Messier, and S. Magierowski, "Optimum receiver for molecule shift keying modulation in diffusion-based molecular communication channels," *Nano Commun. Netw.*, vol. 3, no. 3, pp. 183–195, Sep. 2012.

[188] M. U. Mahfuz, D. Makrakis, and H. T. Mouftah, "Sampling based optimum signal detection in concentration-encoded molecular communication: Receiver architecture and performance," in *Proc. Int. Conf. on Bio-inspired Syst. and Signal Process. (BIOSIGNALS)*, 2013, pp. 372–376.

[189] ——, "A generalized strength-based signal detection model for concentration-encoded molecular communication," in *Proc. ACM Int. Conf. on Body Area Netw. (BODYNETS)*, 2013, pp. 461–467.

[190] D. Kilinc and O. B. Akan, "Receiver design for molecular communication," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 12, pp. 705–714, Dec. 2013.

[191] A. Noel, K. Cheung, and R. Schober, "Optimal receiver design for diffusive molecular communication with flow and additive noise," *IEEE Trans. NanoBiosci.*, vol. 13, no. 3, pp. 350–362, Sept 2014.

[192] T. M. Cover and J. A. Thomas, *Elements of Information Theory 2nd Edition*, 2nd ed. Wiley-Interscience, July 2006.

[193] R. Blahut, "Computation of channel capacity and rate-distortion functions," *IEEE Trans. Inf. Theory*, vol. 18, no. 4, pp. 460–473, July 1972.

[194] S. Arimoto, "An algorithm for computing the capacity of arbitrary discrete memoryless channels," *IEEE Trans. Inf. Theory*, vol. 18, no. 1, pp. 14–20, Jan. 1972.

[195] R. G. Gallager, *Information Theory and Reliable Communication*. Wiley, 1968.

[196] S. Verdu and T. Han, "A general formula for channel capacity," *IEEE Trans. Inf. Theory*, vol. 40, no. 4, pp. 1147–1157, 1994.

[197] B. Atakan and O. B. Akan, "An information theoretical approach for molecular communication," in *Proc. Int. Conf. on Bio-Inspired Models of Netw., Inf. and Comput. Syst. (BIONETICS)*, Budapest, Hungary, 2007, pp. 33–40.

[198] ——, "On channel capacity and error compensation in molecular communication," in *Trans. Comput. Syst. Biol. X*, ser. Lect. Notes Comput. Sci., C. Priami, F. Dressler, O. B. Akan, and A. Ngom, Eds., 2008, vol. 5410, pp. 59–80.

[199] ——, "Single and multiple-access channel capacity in molecular nanonetworks," in *Nano-Net*, ser. Lect. Notes Inst. Comput. Sci., Social Inf. and Telecommu. Eng., A. Schmid, S. Goel, W. Wang, V. Beiu, and S. Carrara, Eds., 2009, vol. 20, pp. 14–23.

[200] ——, "Deterministic capacity of information flow in molecular nanonetworks," *Nano Commun. Netw.*, vol. 1, no. 1, pp. 31–42, Mar. 2010.

[201] A. Einolghozati, M. Sardari, A. Beirami, and F. Fekri, "Capacity of discrete molecular diffusion channels," in *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, July 2011, pp. 723–727.

[202] A. Einolghozati, M. Sardari, and F. Fekri, "Capacity of diffusion-based molecular communication with ligand receptors," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Oct 2011, pp. 85–89.

[203] D. Arifler, "Capacity analysis of a diffusion-based short-range molecular nanocommunication channel," *Comput. Netw.*, vol. 55, no. 6, pp. 1426–1434, Apr. 2011.

[204] B. Atakan, "Optimal transmission probability in binary molecular communication," *IEEE Commun. Lett.*, vol. 17, no. 6, pp. 1152–1155, June 2013.

[205] T. Nakano, Y. Okaie, and J.-Q. Liu, "Channel model and capacity analysis of molecular communication with brownian motion," *IEEE Commun. Lett.*, vol. 16, no. 6, pp. 797–800, June 2012.

[206] M. Pierobon and I. F. Akyildiz, "Diffusion-based noise analysis for molecular communication in nanonetworks," *IEEE Trans. Signal Process.*, vol. 59, no. 6, pp. 2532–2547, June 2011.

[207] ——, "Noise analysis in ligand-binding reception for molecular communication in nanonetworks," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4168–4182, Sep. 2011.

[208] A. Einolghozati, M. Sardari, and F. Fekri, "Collective sensing-capacity of bacteria populations," in *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, July 2012, pp. 2959–2963.

[209] ——, "Design and analysis of wireless communication systems using diffusion-based molecular communication among bacteria," *IEEE Trans. Wireless Commun.*, vol. 12, no. 12, pp. 6096–6105, Dec. 2013.

[210] ——, "Relaying in diffusion-based molecular communication," in *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, July 2013, pp. 1844–1848.

[211] A. Guney, B. Atakan, and O. B. Akan, "Mobile ad hoc nanonetworks with collision-based molecular communication," *IEEE Trans. Mobile Comput.*, vol. 11, no. 3, pp. 353–366, Mar. 2012.

[212] C. T. Chou, "Extended master equation models for molecular communication networks," *IEEE Trans. NanoBiosci.*, vol. 12, no. 2, pp. 79–92, June 2013.

[213] D. Miorandi, "A stochastic model for molecular communications," *Nano Commun. Netw.*, vol. 2, no. 4, pp. 205–212, Dec. 2011.

[214] H. Li, S. Moser, and D. Guo, "Capacity of the memoryless additive inverse Gaussian noise channel," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 12, pp. 2315–2329, Dec 2014.

[215] H. ShahMohammadian, G. G. Messier, and S. Magierowski, "Nano-machine molecular communication over a moving propagation medium," *Nano Commun. Netw.*, vol. 4, no. 3, pp. 142–153, Sep. 2013.

[216] T. Nakano and J.-Q. Liu, "Design and analysis of molecular relay channels: an information theoretic approach." *IEEE Trans. NanoBiosci.*, vol. 9, no. 3, pp. 213–221, Sep. 2010.

[217] A. C. Heren, M. S. Kuran, H. B. Yilmaz, and T. Tugcu, "Channel capacity of calcium signalling based on inter-cellular calcium waves in astrocytes," in *Proc. IEEE Int. Conf. on Commun. Workshops (ICC WKSHPS)*, 2013, pp. 792–797.

[218] D. Kilinc and O. B. Akan, "An information theoretical analysis of nanoscale molecular gap junction communication channel between cardiomyocytes," *IEEE Trans. Nanotechnol.*, vol. 12, no. 2, pp. 129–136, Mar. 2013.

[219] Y. Chahibi, M. Pierobon, S. O. Song, and I. F. Akyildiz, "A molecular communication system model for particulate drug delivery systems," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 12, pp. 3468–3483, Dec. 2013.

[220] A. Eckford and P. Thomas, "Capacity of a simple intercellular signal transduction channel," in *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, July 2013, pp. 1834–1838.

[221] A. O. Bicen and I. F. Akyildiz, "System-theoretic analysis and least-squares design of microfluidic channels for flow-induced molecular communication," *IEEE Trans. Signal Process.*, vol. 61, no. 20, pp. 5000–5013, Oct. 2013.

[222] T. K. Moon, *Error Correction Coding: Mathematical Methods and Algorithms*. Wiley-Interscience, 2005.

[223] M. S. Leeson and M. D. Higgins, "Forward error correction for molecular communications," *Nano Commun. Netw.*, vol. 3, no. 3, pp. 161–167, Sep. 2012.

[224] M. U. Mahfuz, D. Makrakis, and H. T. Mouftah, "Performance analysis of convolutional coding techniques in diffusion-based concentration-encoded PAM molecular communication systems," *BioNanoSci.*, vol. 3, no. 3, pp. 270–284, Sep. 2013.

[225] 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.

[226] M. Kovacevic and P. Popovski, "Zero-error capacity of a class of timing channels," *IEEE Trans. Inf. Theory*, vol. 60, no. 11, pp. 6796–6800, Nov 2014.

[227] 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, Sept 2014.

[228] *IEEE Draft Recommended Practice for Nanoscale and Molecular Communication Framework*, IEEE P1906.1 Std., 2011.

[229] F. Walsh, S. Balasubramaniam, D. Botvich, T. Suda, T. Nakano, S. F. Bush, and M. O. Foghlu, "Hybrid dna and enzyme based computing for address encoding, link switching and error correction in molecular communication," in *Nano-Net*, ser. Lect. Notes Inst. Comput. Sci., Social Inf. and Telecommu. Eng. Springer Berlin Heidelberg, 2009, vol. 3, pp. 28–38.

[230] S. Balasubramaniam and P. Lio', "Multi-hop conjugation based bacteria nanonetworks," *IEEE Trans. NanoBiosci.*, vol. 12, no. 1, pp. 47–59, Mar. 2013.

[231] M. J. Moore and T. Nakano, "Addressing by beacon distances using molecular communication," *Nano Commun. Netw.*, vol. 2, no. 2, pp. 161–173, Sep. 2011, biological Information and Communication Technology.

[232] T. Nakano, Y. Okaie, and A. V. Vasilakos, "Transmission rate control for molecular communication among biological nanomachines," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 12, pp. 835–846, Dec. 2013.

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

[234] A. Biral and A. Zanella, "Introducing purely hydrodynamic networking functionalities into microfluidic systems," *Nano Commun. Netw.*, vol. 4, no. 4, pp. 205–215, Dec. 2013.

[235] L. Donvito, L. Galluccio, A. Lombardo, and G. Morabito, "Microfluidic networks: Design and simulation of pure hydrodynamic switching and medium access control," *Nano Commun. Netw.*, vol. 4, no. 4, pp. 164–171, Dec. 2013.

[236] 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.

[237] D. Demiray, A. Cabellos-Aparicio, E. Alarcon, D. T. Altilar, I. Llatser, L. Felicetti, G. Reali, and M. Femminella, "Direct: A model for molecular communication nanonetworks based on discrete entities," *Nano Commun. Netw.*, vol. 4, no. 4, pp. 181–188, Dec. 2013.

[238] M. J. Moore, T. Nakano, A. Enomoto, and T. Suda, "Measuring distance from single spike feedback signals in molecular communication," *IEEE Trans. Signal Process.*, vol. 60, no. 7, pp. 3576–3587, July 2012.

[239] H. ShahMohammadian, G. G. Messier, and S. Magierowski, "Blind synchronization in diffusion-based molecular communication channels," *IEEE Commun. Lett.*, vol. 17, no. 11, pp. 2156–2159, Nov. 2013.

[240] A. Akkaya, G. Genc, and T. Tugcu, "HLA based architecture for molecular communication simulation," *Simul. Model. Pract. Theory*, vol. 42, pp. 163–177, Mar. 2014.

[241] E. Gul, B. Atakan, and O. B. Akan, "NanoNS: A nanoscale network simulator framework for molecular communications," *Nano Commun. Netw.*, vol. 1, no. 2, pp. 138–156, June 2010.

[242] M. U. Mahfuz, D. Makrakis, and H. T. Mouftah, "Strength-based optimum signal detection in concentration-encoded pulse-transmitted OOK molecular communication with stochastic ligand-receptor binding," *Simul. Model. Pract. Theory*, vol. 42, no. 0, pp. 189–209, Mar. 2014.

[243] I. Llatser, D. Demiray, A. Cabellos-Aparicio, D. T. Altilar, and E. Alarcón, "N3Sim: Simulation framework for diffusion-based molecular communication nanonetworks," *Simul. Model. Pract. Theory*, vol. 42, no. 0, pp. 210–222, Mar. 2014.

[244] "User guide of N3Sim: A simulation framework for diffusion-based molecular communication." [Online]. Available: http://www.n3cat.upc.edu/tools/n3sim/UserGuide

[245] L. Felicetti, M. Femminella, and G. Reali, "A simulation tool for nanoscale biological networks," *Nano Commun. Netw.*, vol. 3, no. 1, pp. 2–18, Mar. 2012.

[246] ——, "Simulation of molecular signaling in blood vessels: Software design and application to atherogenesis," *Nano Commun. Netw.*, vol. 4, no. 3, pp. 98–119, Sep. 2013.

[247] L. Felicetti, M. Femminella, G. Reali, P. Gresele, and M. Malvestiti, "Simulating an in vitro experiment on nanoscale communications by using bins2," *Nano Commun. Netw.*, vol. 4, no. 4, pp. 172–180, Dec. 2013.

[248] H. B. Yilmaz and C.-B. Chae, "Simulation study of molecular communication systems with an absorbing receiver: Modulation and ISI mitigation techniques," *Simul. Model. Pract. Theory*, vol. 49, pp. 136 – 150, Dec. 2014.

[249] G. Wei, P. Bogdan, and R. Marculescu, "Efficient modeling and simulation of bacteria-based nanonetworks with BNSim," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 12, pp. 868–878, Dec. 2013.

[250] T. Kim, M.-T. Kao, E. F. Hasselbrink, and E. Meyhöfer, "Active alignment of microtubules with electric fields," *ACS Nano Lett.*, vol. 7, no. 1, pp. 211–217, Dec. 2007.

[251] A. Tamsir, J. J. Tabor, and C. A. Voigt, "Robust multicellular computing using genetically encoded NOR gates and chemical 'wires'," *Nature*, vol. 469, no. 7329, pp. 212–215, Jan. 2011.

[252] R. Lentini, S. P. Santero, F. Chizzolini, D. Cecchi, J. Fontana, M. Marchioretto, C. Del Bianco, J. L. Terrell, A. C. Spencer, L. Martini, M. Forlin, M. Assfalg, M. D. Serra, W. E. Bentley, and S. S. Mansy, "Integrating artificial with natural cells to translate chemical messages that direct e. coli behaviour," *Nat. Comm.*, vol. 5, May 2014.

[253] W. Bacchus and M. Fussenegger, "Engineering of synthetic intercellular communication systems," *Metab. Eng.*, vol. 16, no. 0, pp. 33–41, Mar. 2013.

[254] S. X. Chen and J. S. Liu, "Statistical applications of the Poisson-Binomial and conditional bernoulli distributions," *Statistica Sinica*, vol. 7, pp. 875–892, Nov. 1997.

[255] MATLAB, *version 7.10.0 (R2010a)*. Natick, Massachusetts: The MathWorks Inc., 2010.

[256] J. Clemmens, H. Hess, J. Howard, and V. Vogel, "Analysis of microtubule guidance in open microfabricated channels coated with the motor protein kinesin," *Langmuir*, vol. 19, no. 5, pp. 1738–1744, 2003.

[257] M. O'Grady, "SMS usage remains strong in the US: 6 billion SMS messages are sent each day," *Forrester Blogs*, 2012. [Online]. Available: http://blogs.forrester.com/michael_ogrady/12-06-19-sms_usage_remains_strong_in_the_us_6_billion_sms_messages_are_sent_each_day

[258] The International Telegraph and Telephone Consultative Comittee (CCITT), "International Telegraph Alphabet No. 2," *Blue Book, Fascicle 7-1*, 1988.

[Online]. Available: http://www.itu.int/rec/dologin_pub.asp?lang=e&id=T-REC-S.1-198811-S!!PDF-E&type=items

[259] "MQ-3 data sheet," Henan Hanwei Electronics Co., Ltd., Zhengzhou, China. [Online]. Available: https://www.sparkfun.com/datasheets/Sensors/MQ-3.pdf

[260] "MQ303 data sheet," Henan Hanwei Electronics Co., Ltd., Zhengzhou, China. [Online]. Available: http://www.kosmodrom.com.ua/pdf/MQ303A.pdf

[261] "MR513 data sheet," Henan Hanwei Electronics Co., Ltd., Zhengzhou, China. [Online]. Available: https://www.futurlec.com/Datasheet/Sensor/MR513.pdf

[262] G. A. F. Seber and C. J. Wild, *Nonlinear regression*. New York: Wiley, 2003.

[263] P. Stoica and P. Babu, "The Gaussian data assumption leads to the largest Cramer-Rao bound," *IEEE Sig. Proc. Mag.*, vol. 28, no. 3, pp. 132–133, 2011.

[264] S. Park, E. Serpedin, and K. Qaraqe, "Gaussian assumption: The least favorable but the most useful," *IEEE Sig. Proc. Mag.*, vol. 30, no. 3, pp. 183–186, 2013.