BIOLOGY-INFORMED COMMUNICATION PROTOCOLS FOR BIO-MOLECULAR NETWORKS

BY

JIAMING WANG

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Electrical and Computer Engineering
in the Graduate College of the
University of Illinois Urbana-Champaign, 2024

Urbana, Illinois

Doctoral Committee:

Adjunct Professor Haitham Al-Hassanieh, Chair and Director of Research
Professor Romit Roy Choudhury
Professor Rayadurgam Srikant
Assistant Professor Bhuvana Krishnaswamy, UW-Madison

Abstract

The field of synthetic biology and bioengineering has made significant strides in the development of micro-scale and nano-scale bio-implants that can operate inside the human body, which has led to an expansion of research interest in developing new networks for these in-body devices. Due to the small factor and bio-compatibility requirements of bio-implants, traditional wireless technologies are unsuitable. Instead, molecular communication (MC) has emerged as a promising alternative to fill this gap.

Molecular communication is a scheme that uses particles as information carriers to exchange messages between devices, which can be easily implemented by manipulating the release of these information-bearing particles–the most straightforward method is using the presence or absence of these molecules to deliver "0" or "1" bits. Molecular communication stems mainly from chemistry and biology, making it bio-compatible and highly energy-efficient for operation with bio-implants inside the human body. While molecular communication has been studied in synthetic biology and information theory, there is little research on building molecular communication systems that account for the unique challenges of communication inside the human body.

This thesis aims to design and implement efficient biology-informed bio-molecular communication protocols that account for the unique challenges of in-body networks.

We start by building an understanding of the bio-molecular communication channel and study how the molecular signals propagate inside the blood vessels. We highlight new properties of the MC channel that differ from traditional communication channels and that were not accounted for by past work. By understanding the biological properties and constraints, we identify the challenges and opportunities to bridge the gap between theory and practice. In particular, our design is based on experimental setups that emulate the blood flow inside the body. Leveraging experimental setups allows us to highlight the following challenges that fall into three categories:

- **Challenges for general molecular channels.** Based on the theory of particle propagation under diffusion and advection, two known challenges for molecular communication: (1) *Long delay spread.* The signals released by the transmitter at the same time can arrive at the receiver through different propagation delays. Such spreading is not only longer than the expected propagation delay, but also positively related to it. This leads to the inter-symbol interference (ISI) in communication. (2) *Signal-dependent noise.* Due to diffusion and unpredictable flows, the propagation of information-bearing particles is highly random. Thus, the number of particles detected by the receiver follows a probability distribution that the variance (channel noise), instead of some constant, is highly related to the expectation (channel response).

  This dissertation not only verifies the above two challenges in a blood-vessel-like testbed but also observes two other general challenges in a molecular channel: (1) *Non-causal channel.* The randomness in the propagation delay indicates that a non-negligible proportion of particles will arrive earlier than

expected. Moreover, particles released in one symbol can arrive earlier than the majority of particles from previous symbols, which is more prominent under high transmission rate. (2) *Short coherence time.* The variation of molecular channels is around tens of seconds. Although this number is significantly larger than conventional wireless channels, the ratio to the delay spread is however much smaller, which highly restricts the efficiency of data packets with conventional techniques.

- **Challenges for multiple access.** This dissertation also highlights two challenges that cause trouble in designing fair and scalable communication protocols among multiple transmitters when interference from other devices comes into play. (1) *Lack of synchronization.* Particles are more inclined to propagate along the direction of environmental flow. Because of this, the molecular channel is asymmetric in that the downstream device can receive the signal from upstream, but not the opposite. This unique property decides that most conventional multiple access protocols are either infeasible (if they require synchronization or feedback) or compromised in fairness and scalability (such as carrier sensing). (2) *Non-negative signal.* Molecular signals are some representations of the number of information-bearing particles. Thus, in a molecular network with a single type of particle, the signals are non-negative, which cannot be cancelled like wireless signals with phases. Multi-User Interference (MUI) remains a problem in molecular channels.

- **Challenges for blood flow.** The unique phenomenon in the blood vessel also raises a challenge for communication, while no similar issues can be found in conventional wireless systems: *Heartbeat-induced varying channel.* Heartbeat leads to the periodic variation of the blood flow in the vessels, especially in the arteries that are close to the heart where the maximum flow can be 3× more than the minimum. The normal heartbeat period has roughly the same order of magnitude as the packet symbol interval, which leads to the evident channel variation across every symbol.

To address these biology-informed communication challenges, this dissertation proposes new protocols for each of the categories, which are not contradictory and can be stacked up together as a whole:

- **μ-Link for general molecular channel.** Since the received signal for each symbol in a molecular packet is highly influenced by the interference from other symbols, μ-Link uses Hidden Markov Model (HMM) to model the molecular channel as the basic solution to track the long ISI–both causal and non-causal. Viterbi algorithm is adopted to solve the problem, while a channel tracking module is added to periodically compensate the divergence of the channel response beyond the coherence time. μ-Link is able to achieve comparable performance to a well-trained deep learning solution.

- **MoMA for multiple access.** Because of asymmetry, protocols that rely on feedback or sensing are not suitable for molecular channels. Taking this into consideration, this dissertation extends the CDMA (Code-Division Multiple Access) and proposes MoMA (Molecular Multiple Access). Since collision is inevitable in molecular network, MoMA carefully chooses the codebook, symbol representation and packet construction to facilitate the detection and decoding of collided packets on the receiver end. Empirical knowledge of molecular channels is introduced to further reduce the error and extend scalability. MoMA is able to achieve a 70% higher throughput than existing CDMA solutions in a network of four transmitters with frequent collisions.

- **FlowLink for blood flow.** In order to track the channel variation per symbol, we need to stick to the invariable in the molecular channel. FlowLink's core idea is an observation that the total amount of

flow required for the propagation, i.e. the integration of the flow rate over the propagation period, is a constant. Instead of modeling the molecular signal as a function of time, it models the signal as a function of flow. By transforming the molecular signal from time samples to flow samples, FlowLink unifies the channel response in the new representation, which can then be fed into a modified decoder with respect to flow interval. FlowLink demonstrates obvious improvement on more than 90% of molecular packets, while shows little negative influence on the rest data.

This dissertation also proposes an in-vitro (outside a living organism) testbed over which the proposed protocols are evaluated. It simulates part of the human circulatory system, which holds similar properties as the blood vessels, including material, diameter, topology, and blood flow. The dataset comprises extensive samples over varying datarate, TX-RX distance, flow cycle and other parameters that will help the illustration of the contributions.

To father and mother.

Acknowledgments

I am thrilled to share this achievement with everyone who has supported me throughout this journey.

Firstly, I would like to express my profound gratitude to my advisor, Prof. Haitham Al-Hassanieh. His unwavering support and guidance over the past seven years have been invaluable. Despite molecular communication being a new field for him, he entrusted me with the freedom to explore and innovate within this exciting domain. His faith in my abilities has been instrumental in the completion of this dissertation. Beyond research, his concern for my well-being, especially during times when we were geographically separated, has been deeply appreciated. His mentorship has not only enriched my academic journey at UIUC and EPFL but also provided me with invaluable industry exposure through collaborations with Microsoft and Facebook.

I would also like to extend my deepest gratitude to my committee–Prof. Romit Roy Choudhury, Prof. Rayadurgam Srikant and Prof. Bhuvana Krishnaswamy–for their valuable feedback and insights to make the dissertation complete and comprehensive. Haitham was a second-year professor when I became his student, and Romit offered a lot of useful advice on how to become a good research group. We two groups used to have weekly meetings including various out activities and holiday celebrations. Romit taught me a lot about how to make good plans as a PhD student, as well as future plans after graduation. Besides, my first course at UIUC was ECE439 Wireless Networks, which was taught by Romit, and that was a very intuitive review and extension on what I have learnt during undergraduate from a different perspective. Bhuvana is like a second advisor to me on this dissertation. Her PhD dissertation was also on the topic of molecular communication but had more coverage on the biology aspect in this field. Her experience with wet labs has filled in the gaps of related knowledge, which is essential to the feasibility and practicability of the protocols proposed in this dissertation.

I am also thankful to Prof. Deepak Vasisht, with whom I had the opportunity to work during my internship at Microsoft Research. Despite the challenges posed by the COVID-19 pandemic, the experience deepened my understanding of the rapidly evolving field of neural networks and millimeter-wave radar in human sensing.

I appreciate the contributions of Dongyin Hu, Chirag Shetty, Sevda Öğüt, and Samin Beheshti Zavareh to the molecular communication project. Their dedication to setting up experiments, constructing, testing, and improving the dataset has been instrumental in the success of this project. Paticularly, Samin has done an excellent job on renovating the testbed for various purposes at EPFL, while Sevda was brave enough to answer all questions at the conference in her first year, both when I was not able to be present on site.

I am grateful for the camaraderie and support of my colleagues at the SyNRG group at UIUC and the SENS group at EPFL. The diverse projects and experiences have shaped my understanding of research and its conduct. During my first year, I joined two different projects which were led by Suraj Jog and Anadi Chaman. Those projects were about conventional wireless medium, different from this dissertation, but showed me what research is and how research should be conducted. Junfeng Guan started PhD the same

year as me, but he focused on a different topic and was more familiar with the devices to keep the whole lab organized. The rest of my lab mates are Mahanth, Nirupam, Ashutosh, Sheng, Zhijian, Wally, Sohrab, Ishani, Waleed, Jitian, Hailan, Samah, Arman, Raphael, Haoxin, Aotu, Camilla.

Lastly, but most importantly, I owe a debt of gratitude to my parents. Their endless love, care, and unwavering support have been my pillar of strength throughout this journey. Their belief in me has empowered me to make my own decisions and persevere through the challenges of living and studying in a foreign country.

Table of contents

List of Abbreviations

ADMA	Amplitude Division Multiple Access.
AWGN	Additive White Gaussian Noise.
BAN	Body Area Network.
BER	Bit Error Rate.
BPM	Beats Per Minute.
BPSK	Binary Phase Shift Keying.
CIR	Channel Impulse Response.
CDMA	Code Division Multiple Access.
CSK	Concentration Shift Keying.
EC	Electric Conductivity.
EM	Electromagnetic.
FDMA	Frequency Division Multiple Access.
HMM	Hidden Markov Model.
IoBNT	Internet of Bio-Nano Things.
ISI	Inter-Symbol Interference.
LS	Least Squares.
LSTM	Long Short-Term Memory.
MAC	Medium Access Control.
MAP	Maximum-A-Posteriori.
MC	Molecular Communication.
MDMA	Molecule Division Multiple Access.
ML	Maximum-Likelihood.
MSE	Mean Squared Error.
MoSK	Molecule Shift Keying.
MUI	Multi-User Interference.

NSK	Nucleotide Shift Keying.
OOC	Optical Orthogonal Code.
OOK	On-Off Keying.
PAM	Pulse Amplitude Modulation.
TDMA	Time Division Multiple Access.
TX	Transmitter.
RNN	Recurrent Neural Network.
RTSK	Release Time Shift Keying.
RX	Receiver.
SNR	Signal-to-Noise Ratio.
SINR	Signal-to-Interference-and-Noise Ratio.

List of Symbols

x	Position on x-axis.
y	Position on y-axis.
z	Position on z-axis.
t	Time.
Δt	Difference in time.
$\vec{v}$	The vector field of flow velocity.
v_x	The component of $\vec{v}$ along x-axis.
v_y	The component of $\vec{v}$ along y-axis.
v_z	The component of $\vec{v}$ along z-axis.
C	Concentration of particles. $C(x,t)$ in 1-D model, while $C(x,y,z,t)$ in 3-D model.
∇C	Gradient of C.
$\nabla \cdot \vec{v}$	Divergence of $\vec{v}$.
$\delta(t)$	The continuous Dirac delta function, i.e. the unit pulse.
$\delta[j]$	The discrete Dirac delta function.
D	Diffusion coefficient.
N	Number of particles.
V	Volume of the receiver detection space.
$Q(t)$	Volumetric flow rate. In 3-D model, it is defined as the volume of fluid passing a cross-section area per time.
$V(t)$	The volume of fluid passing a cross-section since time $t = 0$, which equals to the integral of $Q(t)$ over $[0, t]$.
$\mathcal{N}(\mu, \sigma^2)$	Gaussian distribution with expectation μ and variance σ^2.
$f_X(x)$	The probability distribution function of random variable X.
Re	Reynolds number.

Chapter 1

Introduction

The Internet of Bio-Nano Things (IoBNT) promises to revolutionize medicine and healthcare [1]. It is composed of biological computing machines such as micro and nano-implants that are capable of monitoring body environmental change, analyzing the cause and performing treatment. The past decade has witnessed huge leaps towards enabling this vision. Advances in bio-engineering, synthetic biology, and nanotechnology have led to biosensors that collect and process data [2]–[4], nano-scale Lab-on-a-Chip that run medical tests inside the body [5]–[7], the use of bacteria to design biological nano-machines that can detect pathogens [8]–[10], and all the way to nano-robots that can swim through the bloodstream to perform targeted drug delivery and treatment [5], [11], [12].

Apart from the numerous and fascinating tasks that can be conducted by the bio implants, another highlight and fundamental component of IoBNT is networking. Although plenty of research has been pushing the limit of the functionalities of the bio-medical devices inside the body, it is still treated as a difficult problem to integrate multiple functions on a single micro device, which is mainly due to the space limit on the device. Thus, compared to a single micro device that can conduct healthcare tasks independently, it is more practical to envision a system with devices of different functionalities that can cooperate towards a certain purpose. Communication between the bio-medical devices is essential to the implementation of IoBNT.

However, the past experience has limited involvement in networking such bio-implants to deliver data to the labs-on-chip and communicate with nano-robots. Conventional techniques with electromagnetic communication (EM), which work well for large implants such as pacemakers, defibrillators, and pill cams [13], cannot be scaled in form factor to micro and nano-dimensions [14]–[16]. Moreover, wireless signals do not propagate well in body fluids. This amplifies the need for power storage but contradicts with the minute size of the device. Thus, a long-term in-body network consisting of micro- or nano-robots that can frequently exchange data could not be implemented with the current wireless techniques.

As a result, molecular communication (MC) has emerged as the most suitable solution for networking nano-implants [17]–[22]. The idea behind MC is to encode “1” and “0” bits by releasing molecular particles into the bloodstream. In its simplest form, one can encode “1” bit by releasing particles and a “0” bit by releasing nothing similar to ON-OFF-Keying (OOK) in wireless networks. The small form factor and bio-compatibility of MC make it the most promising approach which led researchers to design bio-transceivers that can send and receive particles using synthetic cells or genetically engineered bacteria as well as biological circuits that use cells to emulate NAND and NOR logic gates [23]–[30].

However, molecular communication is a huge family, where various research directions aim towards

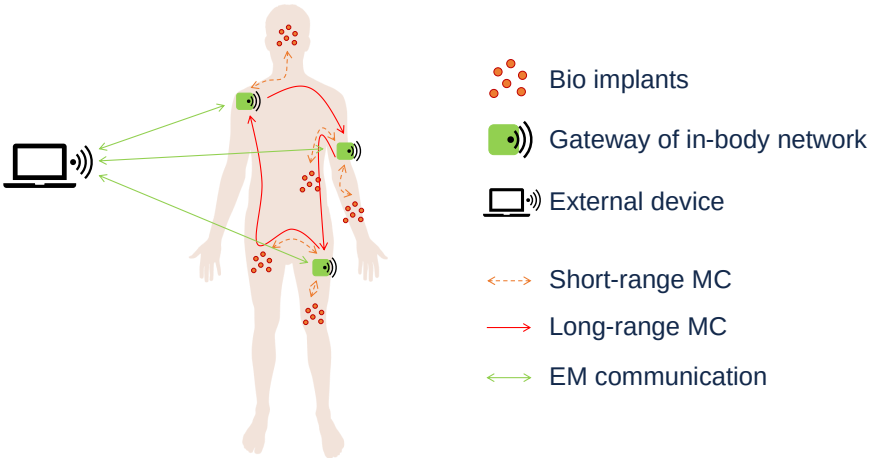


Figure 1.1: A bio-implant network model with in-body subnet (molecular communication between implants and gateways) and outside subnet (EM communication between gateways and external devices).

different applications in different environments and these directions usually have huge gaps between each other. Thus, the discussion of this work is based on one promising in-body bio network architecture with diffusion-based communication (illustrated in Figure. 1.1), which consists of three different types of devices:

- **Bio-implants**. These micro-scale or nano-scale sensors are implanted in different areas of the body, monitoring the variation of the environment. Such data can be used for medical purposes, which is much more complete than a one-time physical examination. However, considering the size of these bio-implants, it would be difficult to include multiple functionalities in the design. Thus, the sensor devices have little storage and periodically send the collected data to the gateways through molecular communication.

- **Gateways**. The gateways also distribute all over the body but have much smaller amounts than bio-implants. One gateway typically serves multiple biosensors, storing the data that can be used during the treatment. Thus, the gateways are composed of two parts: one with molecular module that receives data from the bio-implants, and the other with EM module that sends data to external devices when requested.

- **External stations.** The external station can provide readable bio information collected by the implants, which can be either a medical instrument or a smart phone.

These devices can form three different subnets:

- **Between gateway and external stations**. This subnet exchanges information between the in-body network and the outside network, where the doctors can use the sensor data for diagnosis and send new commands to the bio-medical devices if any necessary changes to the treatment are to be made. Theoretically, one such subnet is adequate for one person.

- **Between gateway and bio-implants**. Each of these subnets is in charge of a certain area in the body, where the bio-medical devices have a very limited diffusion-based communication range in the tissue environment.

- **Between gateways**. In order to exchange information among the gateway-to-implants subnets, a bridge subnet is compulsory. As we will explain in Chapter. 2, the flow of the fluid can greatly extend

the coverage of diffusion-molecular communication. Thus, a network that employs the advantages of the blood flow is a suitable choice to establish connections among gateways, which is named the "blood network."

This work focuses on the molecular communication in the *blood network, where both the transmitter and the receiver are immobile*. Although the architecture of molecular transceivers is not the highlight of the work, we do recognize their importance to the work. A detailed discussion of past and ongoing work is presented in Chapter. 3, which have a firm influence on our protocol design and testbed implementation. We will dive into the opportunities and challenges of the blood network from the perspective of communication, and the necessary assumptions for the work will be listed in Chapter. 2. Until this point, we have identified and proposed our solutions for biology-informed challenges in three categories–challenges for general molecular channel, challenges for multiple access and challenges for blood flow. We highlight the differences between MC and conventional wireless carriers, which prohibit the direct application of conventional protocols, and leverage the unique properties of MC to make efficient adaptations.

- **Challenges for general molecular channels.** This dissertation first discusses the challenges for general molecular channels, which are applicable to all covered scenarios. Based on the theory of particle propagation under diffusion and advection, past work has demonstrated two outstanding challenges for molecular communication.
  - *Long delay spread.* The signals released by the transmitter at the same time can arrive at the receiver through different propagation delays. This is similar to the multipath phenomenon in conventional wireless channels but is more prominent because every single particle could follow on a different path. This leads to the large deviation of propagation delay, which is named long delay spread in the field of communication and leads to the ISI (Inter-Symbol Interference) problem that affects decoding. The delay spread of the molecular channel can span multiple symbol intervals, so guard period is not an efficient solution while the sequence of symbols needs to be decoded as a whole.
  - *Signal-dependent noise.* Diffusion and unpredictable advection are the fundamental causes of the distinct propagation path of each particle, and the direct consequence is that the number of particles detected by the receiver follows a certain probability distribution, the parameters of which are highly dependent on the time since transmission. From the perspective of communication, although the expectation of random distribution can be taken as the channel response, the variance would follow as a separate source of molecular channel noise. Unlike AWGN (Additive White Gaussian Noise) in conventional wireless systems, this variance is highly related to the expectation, and thus is named signal-dependent noise. Because of this, decoders based on LS (Least-Squares) metric are no longer equivalent to ML (Maximum-Likelihood) or MAP (Maximum-A-Posteriori).

  Apart from the above two challenges that can be derived from the theory of particle propagation, this dissertation highlights two other challenges that are common under practical scenarios:

  - *Non-causal channel.* The randomness in the propagation delay indicates that a non-negligible proportion of particles will arrive earlier than expected. Moreover, particles released in one symbol can arrive earlier than the majority of particles from previous symbols, which is more prominent under high transmission rate. Not only will this influence the modulation based on the transmission time of symbols, but also those based on concentration of particles. Unlike wireless channel where

the strongest multipath signal is also the one that has the shortest propagation delay in most cases, the peak of the molecular signal always lies in the middle of the delay spread. The packet detection algorithm will possibly report the peak location with the best metric, but this will also ignore the non-negligible proportion of particles arriving before the peak and lead to erroneous decoding.

  - *Short coherence time.* In communication, information is encapsulated in packets that are composed of two parts–preamble used for estimating channel, and data used for delivering information. The coherence time refers to the period that the channel can be considered as invariant, such that the preamble and the data share similar channel response. The variation of in-body molecular channels is around tens of seconds, which is a consequence of various human activities. Although this number is significantly larger than conventional wireless channels, the ratio to the delay spread is however much smaller, which highly restricts the efficiency of data packets with conventional techniques.

- **Challenges for multiple access.** How to share the medium among multiple devices is a significant component of wireless communication network. This dissertation highlights two challenges that cause trouble in designing fair and scalable communication protocols among multiple molecular transmitters, when interference from other devices comes into play.

  - *Lack of synchronization.* Environment flow (advection) is an important factor that affects the propagation of molecular signals, where particles are more inclined to propagate along the direction of environmental flow. Because of this, the molecular channel is asymmetric in that the downstream device have a higher probability to receive signals from upstream than the opposite direction. The faster the flow, the more unbalance it could be. Unfortunately, the blood vessel is an environment with obvious flow. This has led to three major consequences: first, information can only be transmitted to downstream; second, feedback is not practical because signal can only go in one direction; third, carrier sensing is unfair for downstream devices because they cannot be heard by upstream. All of these are important factors for the implementation of multiple access.

  - *Non-negative signal.* Molecular signals are some representations of the number of information-bearing particles. Thus, in a molecular network with a single type of particle, the signals are non-negative, which cannot cancel out like wireless signals with phases. Thus, MUI (Multi-User Interference) cannot be easily addressed in molecular channels. Nor can it be ignored, since the lack of synchronization will increase the possibility of packet collisions among multiple transmitters, when decoding is greatly compromised.

- **Challenges for blood flow.** The above challenges are all related to the concepts that have been studied more or less in conventional wireless systems. However, the unique phenomenon in the blood vessel also raises a challenge for communication, while no similar issues can be found in conventional wireless systems.

  - *Heartbeat-induced varying channel.* Heartbeat leads to periodic variation of the blood flow in the vessels, where the maximum flow rate can be 3× more than the minimum. The normal heartbeat period has roughly the same order of magnitude as the packet symbol interval, which leads to the evident channel variation across every symbol. Although it sounds similar to the well-studied fast-varying channel problem in conventional wireless system, heartbeat has a much worse influence

on the channel because the coherence time is even shorter than the delay spread, which is not considered before.

This dissertation takes the following steps to identify, understand and address the above challenges:

- We propose a low-cost tabletop testbed with pumps and tubes that mimics molecular signal propagation in blood vessels. It uses edible salt (sodium chloride, $NaCl$) and baking soda (sodium bicarbonate, $NaCO_3$) as information molecule and electrical conductivity (EC) reader as the receiver. Unlike using acid and base with pH value, our testbed maintains the linearity of the molecular signal as a representation of particle concentration variation.

- We target two key characteristics of the MC channel–signal non-causality and short coherence time. Together with two previously studied properties–long delay spread and signal-dependent noise, they significantly limit the achievable data rate between a TX-RX pair. We propose μ-Link, a decoder that altogether addresses these challenges. μ-Link uses HMM (Hidden Markov Model) to explicitly estimate the ISI (Inter-Symbol Interference) caused by the molecular channel properties. It estimates the channel response and directly computes the probability of each possible transmission sequence, which is optimal under signal-dependent noise. The intermediate decoding results are used to compensate for the channel variation beyond the coherence time, while the usage of empirical channel knowledge can tolerate some errors in the temporal decoding sequence. Experiments show that μ-Link can improve the data rate and reduce the bit error, which is comparable to a well-trained deep learning decoder.

- Apart from the above general challenges in molecular channel, we add into consideration two challenges to implementing medium access in molecular networks: non-negative signal and lack of synchronization. We propose MoMA (Molecular Multiple Access), a medium access protocol for molecular networks that enables multiple unsynchronized transmitters to transmit their packets at any time to a receiver that is able to accurately decode colliding packets. MoMA is a CDMA protocol customized to molecular channels, including codebook construction, packet encoding, packet detection, channel estimation and packet decoding. It balances network throughput, scalability and fairness, which achieves 1.7$\times$ throughput than a baseline scheme in a four-transmitter network with frequent collisions. Moreover, by leveraging the unique MC channel properties, MoMA potentially has a much larger scalability than what is currently seen in conventional CDMA systems.

- We target the per-symbol varying channel challenge due to the human heartbeat. As the first in the community to address such symbol-scale channel variation, we propose FlowLink. The design of FlowLink is based on the key identity–although the propagation delay of each symbol is different, the total amount of flow passing the receiver during the propagation is a constant. Thus, FlowLink no longer treats the received signal as a function of time. Instead, it intuitively introduces the re-sampling phase that transforms the RX signal from time-axis to flow-axis, which equals to the integral of the flow rate starting from the time origin. This addresses the hardest problem because the flow-axis channel response is almost identical for every symbol, but also leads to some modifications to the HMM because the symbols are not transmitted at a constant rate on the flow-axis. With corresponding amendments, FlowLink can improve the decoding for more than 90% of the experimental data, as well as little negative influence on the rest.

The road map of the dissertation is as follows: Chapter. 2 reviews the basic knowledge of molecular communication and the necessary assumptions for the whole work in this dissertation. Chapter. 3 provides a

general view of the huge molecular communication family, which is a promising field with rich opportunities and possibilities together with exciting challenges. Chapter. 4 proposes μ-Link, which discusses the channel between one transmitter and one receiver and addresses the channel properties of non-causal channel, short coherence time and long delay spread. Chapter. 5 explains MoMA, which is a new MAC protocol between multiple transmitters and one receiver. MoMA modifies conventional CDMA schemes to overcome the challenges of non-negative molecular signal and lack of feedback under uni-directional propagation. Chapter. 6 introduces FlowLink, which handles the fast-varying channel with symbol-scale period caused by human heartbeat and periodic propagation flow. In the end, Chapter. 7 concludes the work in this dissertation and presents the directions of future research.

Chapter 2

Background

Considering that molecular communication is an emerging paradigm, we will provide a brief introduction on the relevant theory in this chapter. We are concerned with the properties of the molecular signals, which will provide us with guidelines for designing appropriate and efficient communication protocols for molecular networks. On the one hand, molecular signal is distinguished from conventional wireless signals–both electromagnetic and acoustic waves. Some developed communication theories can be generalized to this new signal, while others can no longer be applied. On the other hand, molecular communication itself is a big family. Systems proposed for different environments, particle releasing, propagation and sensing mechanisms, information coding and decoding methods are usually based on different models, and thus are incomparable and even exclusive to a certain application scenario. As a result, it is essential to understand the system concerned in the first step before we explain our contributions.

In the rest of this chapter, we will first explain the phenomenons and corresponding mathematical abstractions that are related to the propagation of molecular signals, and how they decide the properties of the molecular channel. Next, we will introduce the communication system model adopted by this dissertation.

2.1 Propagation of molecular signals

As was mentioned in the previous chapters, this dissertation focuses on the molecular signals in the blood network. Thus, we will explain a few major factors that influence the propagation of molecular signals in the blood vessel. But in the first place, we have to point out that the formation of molecular signals is completely different from conventional communication carriers. Communication over acoustic or electromagnetic waves utilizes the wave properties of the information carrier, which is in essence the propagation of energy and has little to do with the matter. On the contrary, molecular signal is carried over particles and their propagation are a movement of matter itself. Such difference is the fundamental cause that inspires wide molecular communication research.

Coming back to the dissertation, the propagation of particles is a well-studied problem in physics, chemistry and biology. But from the perspective of molecular signal in the blood vessel, the propagation of particles is mainly attributed to two factors–diffusion and advection.

2.1.1 Diffusion

Diffusion refers to the phenomenon that particles move from a region of high concentration to a region of low concentration, which has no net movement. Such phenomenon is attributed to Brownian motion, referring to the movement caused by the collisions between particles. Considering the extreme difficulty describing the behavior of every particle of a countless quantity, Brownian motion is usually treated as a random process and described with a statistical model.

The Wiener process [31], [32] is a simple yet effective physical model to describe the Brownian motion under the condition of minimal friction. For the purpose of simple explanation, we divide the concerned particles into two categories–the information-bearing particles and other particles in the medium–and make the following assumptions:

- The particles are immutable (i.e. they do not change identity, emerge or disappear during propagation).
- The information-bearing particles are of a comparably lower concentration than the medium particles. It means the collisions between these information-bearing particles are negligible, so their motions are independent and identically distributed (i.i.d.).

Under the above conditions, the motion of each individual information-bearing particle can be modeled by the Wiener process. Let $(x(t), y(t), z(t))$ denotes the position of the particle in a 3-D space as a function of time. Temporarily ignoring the obstacles in the environment (i.e. the wall of blood vessels), the displacement of the particle over a certain amount of time Δt follows the Gaussian distribution, i.e.

$$\begin{aligned} x(t+\Delta t) - x(t) &\sim \mathcal{N}(0, 2D\Delta t) \\ y(t+\Delta t) - y(t) &\sim \mathcal{N}(0, 2D\Delta t) \\ z(t+\Delta t) - z(t) &\sim \mathcal{N}(0, 2D\Delta t) \end{aligned} \tag{2.1}$$

The notations in Equation. 2.1 are given as follows:

Gaussian distribution

Suppose X is a random variable, then $X \sim \mathcal{N}(\mu, \sigma^2)$ signifies that x follows the Gaussian distribution with mean μ and variance σ^2, i.e. the probability density function $f_X(x)$ is given by

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \tag{2.2}$$

where $\exp(t) = e^t$.

Diffusion coefficient

The diffusion coefficient D, in the units of m^2/s, is a parameter that describes the speed of the Brownian motion, which is given by

$$D = \frac{k_B T}{6\pi\eta R_H} \tag{2.3}$$

where $k_B = 1.38 \times 10^{-23}$ J/K is the Boltzmann constant, T is the Kelvin temperature [K], η is the dynamic viscosity of the fluid [kg/(m·s)], and R_H is the hydraulic radius of the molecule [m]. Typically, the

diffusion coefficient of a molecule in the gas phase is in the range of $10^{-6} - 10^{-5}$ m^2/s, while in an aqueous solution $10^{-10} - 10^{-9}$ m^2/s.

2.1.2 Advection

Advection refers to bulk motion of an entire body as a consequence of pressure gradient, such as blood flow (liquid) and wind (gas), which can be described by a vector field $\vec{v}$, in the units of m/s. It is obvious to identify the non-negligible influence of advection over the propagation of molecular signals in the blood network. A simple modification to Equation. 2.1 is able to model advection as follows:

$$\begin{aligned} x(t+\Delta t) - x(t) &\sim \mathcal{N}(v_x \Delta t, 2D\Delta t) \\ y(t+\Delta t) - y(t) &\sim \mathcal{N}(v_y \Delta t, 2D\Delta t) \\ z(t+\Delta t) - z(t) &\sim \mathcal{N}(v_z \Delta t, 2D\Delta t) \end{aligned} \tag{2.4}$$

where (v_x, v_y, v_z) is the value of the advection vector field.

Although Equation. 2.4 presents a simple modeling of the flow in environment, it should be noted that uniform flow (i.e. v_x, v_y and v_z as constants) is rarely viewed under a practical scenario. For instance, the blood flow speed in the vessel is decided by many factors including the total flow amount per time, the shape of the vessel, etc. Taking a step back, even if we abstract the space in the blood vessel as a cylinder through which passes a stable blood flow, the flow speed is not identical everywhere in the vessel due to the existence of viscosity.

Laminar flow

Viscosity quantifies the internal frictional force between adjacent layers of fluid that are in relative motion, which is the reason why the fluid can only pass through a tube with pressure difference between the two ends. For example, the heartbeat provides the pressure difference and forces blood to circulate over the body. And because of the internal friction, it is observed that the viscous fluid flows slower when it is closer to the wall of the tube, which is termed as laminar flow–the fluid particles follow smooth path in layers, with each layer moving smoothly past the adjacent layers with little or no mixing [33].

Suppose the blood vessel delivers a cylinder flow with a constant radius of R, the speed profile is given as a function of radius of the layer r as

$$v(r) = v_m \left(1 - \frac{r^2}{R^2}\right) \tag{2.5}$$

where v_m is the maximum flow speed that corresponds to the layer on the axis of the cylinder.

As a result, the propagation distance of each information-bearing particle is not simply decided by its diffusion along the flow direction. It is more influenced by how it diffuses along the lateral direction between layers and takes over different flow speeds across layers.

Turbulent flow

Although laminar flow is the widely adopted assumption for the human circulatory system, there are cases where such an assumption is broken. Empirically, the laminar flow becomes unstable when an obstacle suddenly appears in the flow or when the pressure at the two ends of the tube suddenly bears a huge difference.

Under such conditions, the flow is no longer organized in layers and becomes irregular, which is termed as turbulence.

The Reynolds number (Re) is a dimensionless quantity that helps predict the flow pattern. It is defined as

$$Re = \frac{\rho v L}{\eta} \tag{2.6}$$

where ρ is the density of the fluid [kg/m^3], v is the mean velocity of the fluid [m/s], L is the characteristic length [m] and η is the dynamic viscosity of the fluid [kg/(m·s)].

At low Reynolds number, the flow tends to be dominated by laminar flow. For example, the common Reynolds number of blood flow is in the range of 100 to 1000, which varies in different parts of the body. On the contrary, at high Reynolds number, the flow tends to be turbulent. There does not exist a clear division in the Reynolds number that tells laminar flow and turbulence apart. Even with co-existence of the two flows, it is possible that some space is dominated by laminar flow while others by turbulence, which is decided by the boundaries of the fluid.

Due to the irregularity of the turbulent flow, it is not always feasible to accurately describe the vector field. But fortunately, it is observed that turbulence tends to accelerate the mixing of fluid mixtures, which is very similar to the molecular diffusion (see Section. 2.1.1). Thus, people introduce the turbulent diffusion coefficient to approximately quantize the influence of turbulence.

2.2 Advection-Diffusion Equation

Although Equation. 2.4 provides a good guideline to simulate the propagation of information-bearing particles, it is required to predict the motion of each individual particle, which is termed as Monte-Carlo method. This is a suitable and accurate method to track the motion of a limited number of particles but is time-consuming when the transmitter releases a large number of particles (more than 10^{10}) to guarantee the basic communication reliability in the blood network.

To this end, the advection-diffusion equation offers a macro view over the molecular propagation, which describes the concentration of information particle C as a function of time and position,

$$\frac{\partial C}{\partial t} = \nabla \cdot (D\nabla C - \vec{v}C) + P \tag{2.7}$$

where P represents the sink/source of the particles and other reactions that can lead to the formation or deformation of the particles, which is also a function of position and time. ∇ and $\nabla\cdot$ are respectively the gradient and divergence operators. Note that turbulence can either be accurately described as part of the vector field $\vec{v}$, or approximately as part of the diffusion coefficient D.

Equation. 2.7 is combined with different boundary conditions to model the propagation of particles in various environments. For example, in an infinite 1-D environment with uniform flow v_x, no initial information-bearing particles and no sinks. Suppose a unit pulse of immutable information-bearing particles is released at time $t = 0$ at origin $x = 0$. The expected concentration $C(x, t)$ can be computed by solving with the following equation system:

$$\begin{cases} \dfrac{\partial C}{\partial t} = D\dfrac{\partial^2 C}{\partial x^2} - v_x \dfrac{\partial C}{\partial x} \\ C(x,t)|_{t=0} = \delta(0) \end{cases} \tag{2.8}$$

which has an analytical express as

$$C(x,t) = \frac{1}{\sqrt{4\pi Dt}} \exp\left(-\frac{(x-v_x t)^2}{4Dt}\right) \tag{2.9}$$

2.3 Signal-Dependent Noise

Since Equation. 2.9 is a deterministic expression when the transmitter releases a *single* molecule, it is more suitable to interpret $C(x,t)$ as the probability distribution function of the particle with regards to time. Assuming the transmitter releases a pulse of N particles and each of them propagates independently and identically, The number of particles $n(x,t)$ detected by a passive receiver of volume V then follows a binary distribution.

$$n(x,t) \sim \mathcal{B}\left\{N, VC(x,t)\right\} \tag{2.10}$$

The expectation and the variance of $n(x,t)$ is

$$\begin{aligned} E[n(x,t)] &= NVC(x,t) \\ Var[n(x,t)] &= NVC(x,t)\left(1 - VC(x,t)\right) \end{aligned} \tag{2.11}$$

In the blood vessel channel, the distance x between the transmitter and the receiver is usually large, such that $C(x,t) \approx 0$. Moreover, the transmitter releases a large amount of molecules N at a time, so $n(x,t)$ can be approximated to a normal distribution.

$$n(x,t) \sim \mathcal{N}\left(NVC(x,t), NVC(x,t)\right) \tag{2.12}$$

If the receiver samples the concentration of the particles as the received signal $r(x,t)$, we can see that

$$r(x,t) = \frac{n(x,t)}{V} \sim \mathcal{N}\left(NC(x,t), \frac{NC(x,t)}{V}\right) \tag{2.13}$$

where the variance of the received signal is a function of the expectation

$$Var[r(x,t)] = \frac{E[r(x,t)]}{V} \tag{2.14}$$

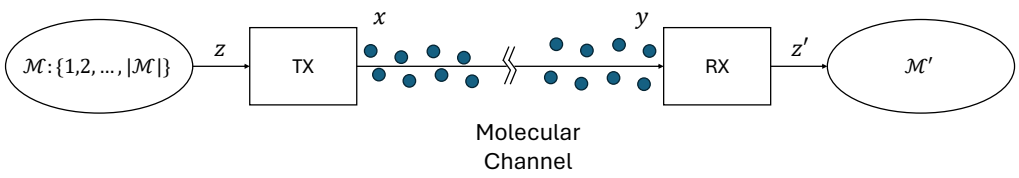
This is named signal-dependent noise, which is the inherent property of the molecular channel.

2.4 System model

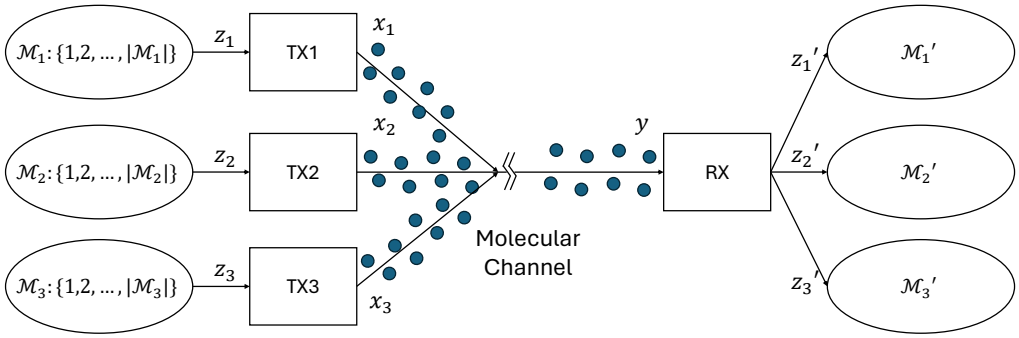
For a better understanding of the work in this dissertation, we will first explain the related communication model for the new information carrier, so the readers can match the tasks in conventional theory. In the second part of this section, we will explain the assumptions from the molecular perspective, which will provide a clear view of the molecular system and a better understanding of the contributions.

2.4.1 Communication Perspective

Referring to the layered structure of conventional communication systems, this dissertation focuses on the physical layer and medium access control (MAC) sub-layer in molecular communication. These layers can be



(a) Caption



(b) Caption

Figure 2.1: The description of the communication model used in this paper for: (a) single-TX-single-RX pair; and (b) multiple-TX-single-RX.

abstracted as a mathematical model as depicted in Figure. 2.1, including two scenarios of a single transmitter and multiple transmitters.

Figure. 2.1a presents the mathematical model with one transmitter (TX) and one receiver (RX) A message z is selected from a discrete and finite message set $\mathcal{M}$, which is then encoded for transmission. The transmitter (TX) maps the message z into the channel input x, which is transformed by the channel and the receiver into output y. The receiver infers the message z' based on the received signal y, If $z' = z$, the receiver gets the correct message; otherwise, an error occurs. This model is used to study the physical layer of the molecular channel and is referred to in Chapter. 4 and Chapter. 6.

Taking a step further, Figure. 2.1b presents the model with multiple transmitters and one receiver, which is referred to in Chapter. 5. The only difference from the single-transmitter-signal-receiver model is that the channel inputs of each transmitter $x_1, x_2, ...$ are transformed by the molecular channel all together into a signal output y, given which the receiver has to recover the message of each transmitter separately. This model is used to study the medium access control (MAC) of the molecular channel and is referred to in Chapter. 5.

2.4.2 Molecular Perspective

Although this dissertation aims for practical and efficient communication protocols in molecular networks, it is not sufficient to study the whole family of molecular communication and address all issues in this topic. Thus, in order to focus on the molecular channel, this dissertation adopts a few ideal but necessary assumptions, which are listed as follows.

Information-bearing particles

The dissertation assumes that the basic communication protocol employs a single type of information-bearing molecule. This is the simplest system from the hardware perspective, but it can be directly generalized to multiple molecules with minor modifications in protocol by treating them independently. However, when the dissertation raises the point of using multiple molecules, it is not simply related to the natural channel capacity gain with more information carriers. Instead, it highlights the extra advantages of the interconnection between different molecules (details in Chapter. 5).

Moreover, individual particles of this molecule are indistinguishable. With a large number of information-bearing particles, very little information can be directly encoded in the transmission time of the particle. The receiver relies more on the counting (or concentration) to guess the original message.

Finally, the concentration of the information-bearing particles is comparably lower than the concentration of the medium particles. This is the same assumption as Equation. 2.4. With negligible interaction between information-bearing particles, the counting (or concentration) is a linear summation. This is the fundamental prerequisite of the contributions to this dissertation.

Architecture of TX and RX

Until now, the architecture of molecular devices remains an unsettled question with rich opportunities in various fields. This dissertation does not cover such problems. Instead, it makes a few ideal but necessary assumptions to facilitate research in communication.

- *The transmitter and the receiver have the perfect clock, which enables them to send or sample the molecular signal at perfect timing.* Constructing the circuit for a micro-scale bio-implant is an open problem and there does not exist a widely adopted solution to the clock module–either using classical oscillator or using bio-engineering methods. This is an interesting and crucial problem for the bio-implant network, but the performance of the clock is unknown at this point. Thus, this dissertation adopts the "perfect clock" assumption to reduce the factors that is not very relevant to the molecular channel itself. However, the devices are not necessarily synchronized, which means the receiver needs to detect the start of the molecular signal before sampling the channel output and guessing the message.

- *The transmitter has a sufficient source of information-bearing particles to sustain the transmission of molecular signal.* Considering the size of the micro-implants, it is impractical for them to carry an infinite-size particle container for long-term service. However, the possibility exists that the bio-implants can harvest material and energy from the environment and construct the required particles for communication, which is a competent research direction. With this assumption, the transmitter adopts OOK (On-Off Keying) as the modulation scheme and can release a pulse of molecules in a negligible amount of time.

- *The transmitter and the receiver are immobile.* This dissertation considers the molecular communication in the blood vessel, where the information-bearing particles propagate under the influence of blood flow. However, if the transmitter or the receiver is mobile in the vessel, it is under the same influence of blood flow and has similar propagation speed as the molecular signals. Under such case, it is more suitable to use another model that studies communication happens in the near field within a very short period of time during which the devices meet, but the drawback is that it could be possible that the

memory of the transmitter runs out before communication requirement is met. Devices with navigation are a possible solution, but they are beyond the content of this dissertation at the moment.

We have to emphasize that all of the above assumptions are applicable to the work in this dissertation but may not generalize to all possible molecular network architectures which are not covered. Besides, due to the vast varying molecular environments, quantitatively modeling every possible channel is impossible. Thus, the protocols in this dissertation all follow the conventional procedure including detection, channel estimation and decoding. The above assumptions serve as qualitative guidelines in designing new components that are suitable for the molecular carrier.

2.4.3 Metrics for evaluation

This dissertation mainly adopts the BER (Bit Error Rate) as the metric for evaluation, which is defined as the ratio of the proportion of wrongly decoded bits in the whole data bits of a packet. The BER is reported under various controlled conditions for thorough evaluation, such as symbol interval, TX-RX distance, etc. However, SNR (Signal-to-Noise Ratio) or SINR (Signal-to-Interference-and-Noise Ratio) is a more universal description for all these controlled conditions, where the "BER versus SNR" plot is a widely adopted evaluation in conventional wireless research field. This dissertation cannot present this evaluation, because we are not sure about the definition of the power of molecular signal, and we cannot find any reference to this problem in existing work. We are faced with a conflict in defining the power: on the one hand, the particles should be measured in quantity, so the power should be proportional to the number of particles; on the other hand, the received signals in experiments are measured by electric devices, and the noise in the circuit gets involved. We believe that a good definition can only be achieved under a large number of experiments, but it is beyond the reach of this dissertation. Thus, BER under different controlled conditions is chosen to present and compare the performance among different protocols.

Chapter 3

Literature Survey

As discussed in Chapter. 1, the emergence of research interest in molecular networks is attributed to the advancements in synthetic biology and bioengineering, which have allowed us to construct biosensors in micro and nano scales. In this chapter, we will first briefly review a few efforts in the implementation of these bio devices to show the promising future of the in-body sensor network. We will then present past work in understanding the molecular channel and designing medium access control protocols.

3.1 Bio-implants

Bio-implants are the most outstanding part that distinguishes in-body networks from traditional networks. The most important component of a bio-implant is to monitor a biological signal or activity or process. Researchers have put significant effort into extending the coverage of sensing techniques on various signals. With synthetic biology and bio-engineering, biosensors are now able to monitor water quality [34], [35], distinguish cancerous cells from others [36]–[38], and detect metallic ion level [39].

Another important component of a bio-implant is the communication module, which mainly includes processor, oscillator and transceiver. The researchers have implemented biological components of the same functionality in the past work.

- **Processor:** The processor is in charge of the whole biosensor. The fundamental building blocks of a digital circuit are NAND and NOR. [26], [27], [40] have synthesized the biological NAND and NOR circuits, which reveals the promising future of biological processors. [41] presents a robust, general, and scalable system (named BLADE) to engineer genetic circuits with multiple inputs and outputs in mammalian cells, enabling sophisticated cellular computation.

- **Oscillator:** Oscillator is responsible for tracking the transmitted or received samples. [8], [42], [43] presented genetically engineered circuits that have the functionality of a "clock."

- **Transceiver:** Molecular transceiver is the most diverse part, where researchers propose various information carrier molecules and corresponding transceiver architecture. These can be classified as nucleic acids [44]–[48], elemental ions [49], [50], neurotransmitters [51], [52], proteins [53], and others [54], [55].

3.2 Molecular Modulation

Based on the type of information carrier molecule, various modulation schemes are proposed and tested:

- **OOK:** On-Off Keying (OOK) is the widely adopted modulation scheme in MC for its simplicity. The transmitter releases a pulse signal or a rectangular signal to transmit a bit "1" and no signal for a bit "0". A more detailed survey is presented later in Section 3.3.
- **CSK:** [54], [56], [57] Concentration Shift Keying (CSK) is a general version of OOK. The transmitter releases different amounts of molecules to transmit different symbols, while for OOK the number of symbols is limited to two. Although CSK improves the data rate by encoding more bits in one data symbol, the scheme is more vulnerable to inter-symbol interference (ISI) and signal-dependent noise in MC channel.
- **MoSK:** [54], [56], [57] Molecule Shift Keying (MoSK) encodes the information in the type of molecule transmitted. Since the same molecule is not usually transmitted consecutively for the data bits, MoSK is not as influenced by ISI as CSK. However, MoSK is fundamentally limited by the design of the system, especially by the number of available molecules that are suitable in a certain environment. Further, [54] proposes a method encoding information in the ratio of two released molecules.
- **RTSK:** [58]–[60] Instead of changing the pulse amplitude released by the transmitter, Release Time Shift Keying (RTSK) encodes information in the release time of molecules. Some examples can be pulse position modulation (early pulse for bit "1" and late pulse for bit "0") [61] and time elapse modulation (encoding information in the period between two pulses) [60].
- **NSK:** [62], [63] The above modulations usually offer only a limited number of symbols, so they typically suffer from a low data rate considering the long propagation delay. Nucleotide Shift Keying (NSK) proposes that the information can be encoded in the base sequences of DNA, which can fundamentally address the defect of low data rate. However, there is yet no practical system to construct such DNA packets in micro-scale devices.

Among all these schemes, we choose OOK due to its simplicity, which introduces the least interfering factors against our efforts on understanding the molecular channel.

3.3 Molecular Channel for OOK

There has been a lot of efforts in understanding the molecular channel from both theoretical and experimental perspectives:

- **MC Channel Modeling and Estimation:** There has been a significant amount of work on modeling the MC channel taking into account different factors such as propagation environment, dimensions (1D, 2D, 3D), external influences like turbulence and advection, mobility, and type of receiver (destructive vs non-destructive) [64]–[69]. Time varying channel due to mobility is also modeled in [70], [71]. However, the models do not account for other time varying factors such as unstable advection that can be caused by pumping. A number of papers also focus on estimating the channel parameters such as distance, diffusion coefficient, and fitting them to a known model or closed-form expression [72]–[74]. Unfortunately, all these works are based on simulations and none of them are validated in practice.

While modeling provides valuable insights, it cannot account for all factors and parameters that can influence the channel in practice as shown in [75].

- **MC Channel Capacity:** There has been theoretical work on studying the capacity of molecular communication [76]–[80]. However, to derive the capacity of the channel, these works must make simplifying assumptions that do not hold in practice. [79], [80] derive the MC channel with mobile transmitter or receiver but completely ignore ISI. [77], [78] consider signal dependent noise but assume a static channel. [76] considers signal dependent noise and ISI but assumes propagation through diffusion only. None account for non-causal ISI. Deriving the capacity is, indeed, valuable for understanding the data rate limits of MC. However, the channels tend to be significantly more complex in practice.

- **MC Decoders (Simulations):** Several MC decoders have been introduced and tested using simulations. Some exploit specific features of the MC channel response like monotonicity and local convexity [81], [82] but assume minimal ISI. Others introduce optimal decoders [83]–[85] but assume the channel impulse response is perfectly known a priori. [86] introduced a non-coherent decoder using blind channel estimation but it treats ISI as noise. These decoders are tested only in simulations and underestimate the severity of ISI in practice especially as we go to shorter symbols (higher data rates). Furthermore, they do not address non-causal ISI and the large delay-spread-to-coherence-time ratio. Some theoretical work tries to address the severity of forward-ISI in MC channel using mechanisms to reduce the tail of the channel impulse response. Some propose using enzymes or photolysis reactions to remove interfering molecules retained in the channel [87], [88] while others propose using magnetic molecules and adding an external magnetic field [89], [90]. These methods, however, have not been tested in practice and can only reduce forward-ISI but not eliminate it.

- **MC Experimental Work:** Unlike theoretical and simulation based past work, there are very few experimental MC testbeds especially in liquids. Most MC testbeds are for airborne molecular communication where the molecules are released in air [91]–[96]. The dynamics of molecular propagation in such gaseous mediums tend to be significantly different than in liquids. The applications of airborne MC also tend to be limited since wireless RF works really well in air. In this paper, we focus on liquid-based molecular communication for which very few testbeds exist [75], [97], [98]. All testbeds use water as a liquid environment. For molecules, they use magnetic particles [75], RNA [97] and acid/base [98]. [75], [97] focus on channel modeling and fitting and do provide an evaluation of data rate or BER. [75] uses a symbol interval of 4 sec (data rate of 0.25 bits/sec) whereas [97] does not report symbol interval or data rate. [98] uses deep neural network techniques including LSTM RNN to decode without explicitly modeling or estimating the channel. While it achieves good performance, it relies heavily on tedious training for every distance, channel, and bit rate.

3.4 Medium Access Control

Enabling multiple access in molecular networks has been studied by past work. Except for a few [99], [100], most work has been theoretical or simulation-based:

- **TDMA:** [101] proposes a TDMA (Time Division Multiple Access) scheme. However, TDMA requires synchronization which is difficult to implement in practice due to (a) the long propagation delay of molecular signals and (b) the added complexity of requiring a receiver on the micro-implant.

- **MDMA:** [102], [103] introduce MDMA (Molecular Division Multiple Access) where different molecules are used by different transmitters to avoid interference altogether. This scheme is similar to FDMA (Frequency Division Multiple Access) in wireless networks which avoids synchronization and interference. However, it is difficult to scale this approach to many transmitters as most practical bio-molecular systems are limited to 2-3 molecules [10], [104], [105].

- **ADMA:** [99] presents ADMA (Amplitude Division Multiple Access) that assigns the amplitude of the transmitted signal as the address of the source. This method reduces the error in resolving the source addresses under collisions, but the required signal power scales exponentially with the number of transmitters and the performance is sensitive to changes in the channel response.

- **CDMA:** [100], [106] propose a CDMA (Code Division Multiple Access) scheme that adopts Optical Orthogonal Codes (OOC) from fiber optic networks. [100] further tests the OOC codes in a gaseous molecular testbed with two transmitters. Despite the fact that, similar to molecular networks, the signal in fiber optic networks is also non-negative since it represents the intensity of light [107], OOC is not suitable for molecular networks for the following reasons: (a) Unlike molecular networks, fiber optic networks can deal with high ISI in hardware by compensating for the dispersion of light pulses. Techniques such as DCF (Dispersion Compensation Fiber) and FBG (Fiber Bragg Grating) [108] can change the physical properties of the channel to reduce ISI. However, in the presence of high ISI, one cannot simply correlate with the OOC code of each transmitter and decode them independently as proposed in [100], [106]. (b) Commercial fiber optic systems that use OOC typically use long OOC codes which can significantly reduce the data rate [109]. Unlike optical networks that operate at link budgets of 100 Gbps [110], molecular networks operate in the few bits per second regime and cannot afford a significant reduction in data rate. (c) Using OOC codes of short length (high data rate) leads to high Bit Error Rates (BER). Moreover, [100] itself reports poor performance where the error rate is low only at data rates below 0.051 bits per second. [111] also proposes a CDMA scheme that uses two molecules but is not based on OOC. However, the two molecules are used for modulation to encode different bits, i.e., releasing molecule A represents a "1" bit while releasing molecule B represents a "0". Since the two molecules cannot be released at the same time, the network capacity is cut by half.

- **Broadcast (No Multiple Access Scheme):** [112] computes the theoretical BER of a multi-transmitter network where the transmitters broadcast their data packets that collide at the receiver. However, in this work, each transmitter is decoded independently resulting in a drastic increase in BER as the number of transmitters increases. [113] also computes the theoretical capacity of a multiple-TX single-RX network when the network uses one molecule, and transmitters broadcast their packets. However, the work does not present any approach for decoding the colliding packets.

3.5 Blood network testbeds

The importance of blood network has received wide confirmation in the community. A lot of testbeds have been proposed to simulate blood vessel environment. [98], [114] used silicone tubes to simulate the vessels and evaluate the performance of decoders, but were not strict with the flow speed. [115] used a 2-mm diameter silicone tube to simulate the vascular system with volume rates between 50 mL/min and 90 mL/min. [116] used a 3-mm diameter tube which is close to the average diameter of coronary artery with volume rates between 50 mL/min and 350 mL/min. [117] used a narrower 1.52-mm diameter tube and lower volume

rates between 2 mL/min and 20 mL/min but tested the propagation of the bio-compatible SPION particles. However, all of the above work only provides some results under a constant flow rate and ignores the influence of flow variation due to heartbeat.

3.6 Fast-varying channel

This work, to our knowledge, is the first to address the fast-varying channel in molecular communication. However, there have been various works in the context of conventional wireless channels. Fast-varying channel was first officially considered in radio astronomy [118], [119], but has nowadays become a common phenomenon in most popular communication systems due to the rapidly growing mobility in different user scenarios. Besides, underwater acoustic communication views much higher relative Doppler shifts, since the sound speed in water is only around 1500 m/s [120]. [121] proposed a two-step approach to mitigate the Doppler influence in zero-padded OFDM for underwater acoustic communication, which includes a step of non-uniform resampling based on frequency-dependent Doppler drifts. [122] further demonstrated the effusiveness of Doppler-based resampling for underwater acoustic OFDM transmission without guard interval, but at a much lower-than-normal symbol rate. Despite the numerous solutions for conventional communication systems, it is possible that none of them can be applied to MC, because the delay spread can be much longer than the channel coherence time in the blood network.

Chapter 4

μ-Link

4.1 Introduction

Molecular communication (MC) has emerged as a promising technology for communication through fluids such as micro-implants communicating through the bloodstream or sensors communicating through industrial pipes [17], [21], [123]. In MC, a device can transmit data by releasing molecules into the fluid which are then transported and detected at a receiver [12]. For example, a device can release molecules to encode a "1" bit and release nothing to encode a "0" bit. The receiver can measure the concentration of molecules to determine whether the transmitted bit was a "1" or a "0".

Molecular communication has the potential to enable micro and nano-implants to communicate with each other inside the human body and coordinate sensing and actuation tasks. Recent advances in biomedical sciences have in fact led to the development of nano-implants that can sense human vitals from inside the body and even travel through the bloodstream to perform targeted drug delivery and treatment [5], [11], [12]. There is significant interest in enhancing the operation of such implants by connecting them using MC [17]–[22], [123]–[126]. MC presents a suitable alternative to other communication technologies such as wireless. In particular, RF signals do not propagate well in fluids and form factor constraints prevent scaling RF radios to micro and nano-dimensions [14]–[16]. In contrast, for MC, researchers can design synthetic cells that send and receive molecular signals [28]–[30], nano-scale Lab-on-a-Chip that monitor chemical content [5]–[7], and bio-implants that collect and process data [2]–[4].

While there is still a long way to realize the above vision, this work takes steps to achieve a better understanding of the characteristics of the MC channel from both theoretical and empirical perspectives. The MC channel tends to be more complex in practice than standard RF, optical, or copper wire channels [127]. Understanding and addressing the differences between these channels and MC allows us to improve the performance of molecular communication.

There has been a significant amount of work on theoretically modeling the MC channel [64]–[70], [72], [74]. However, these models tend to be overly simplified with assumptions that do not hold in practice (e.g. no inter-symbol-interference) or overly fitted to a closed form equation that does not capture practical constraints and imperfections (see Section. 4.7 for more details). Hence, these models are only evaluated through simulations. On the other hand, there has been little work on empirically validating the MC channel models in fluids [75], [97], [98]. Performance in these testbeds yields at best a data rate of 4 bits/sec with 3% bit error rate (BER) [128]. However, this is achieved by using deep neural networks to decode MC signals

without explicitly learning or modeling the channel. While promising, the approach requires the network to be trained for the specific channels being tested and hence, might not generalize to new channels that the neural network has not seen before, as we show in Section. 4.6.

In this work, we build on the above theoretical and empirical work to capture the characteristics of the MC channel. We introduce two new key properties that have been overlooked by past work on MC. Specifically, the molecular communication channel exhibits non-causal inter-symbol-interference, and a long delay spread that extends beyond the channel coherence time. We show how to leverage these new insights to improve decoding performance and achieve a higher bit rate with lower bit errors. We also empirically validate and incorporate insights introduced by past work, such as the presence of signal dependent noise and the channel response resulting from the statistical behavior of molecular diffusion.

Our work, in particular, reveals the following two insights:

- **Non-causal Inter-Symbol-Interference:** Due to the statistical nature of diffusion, molecules released at the transmitter can arrive at significantly different times at the receiver even if they follow the same path. As a result, the molecules of previous symbols can arrive late creating inter-symbol-interference (ISI). Molecules of future symbols can arrive earlier than molecules of the current symbol creating a non-causal ISI as well. In contrast, in standard communication systems, ISI is caused only by previously transmitted symbols which can be decoded first to eliminate their ISI as opposed to future unknown symbols that cannot be decoded before the current symbol.

- **Long Delay Spread vs. Coherence Time:** [1] In MC channels, the channel delay spread can be long (1 to 4 seconds), which is of a similar order to the channel coherence time (a few 10s of seconds) [129], [130]. Hence, the channel varies only slightly slower than the time it takes us to estimate it. Compare this to RF channels where the delay spread ($0.1\mu s - 10\mu s$) is orders of magnitude smaller than the coherence time ($10ms - 100ms$) and we can quickly estimate the channel and assume to be constant for the duration of packet transmission. However, in MC, simply estimating the channel impulse response and using it to decode the bits throughout the packet leads to poor performance.

To address the above issues, we design μ-Link, a molecular communication protocol and decoder. Due to non-causal ISI, μ-Link cannot sequentially decode bits. Instead, it jointly decodes the current, future, and past symbols. To do so, the receiver first estimates the channel impulse response using known preamble bits transmitted at the beginning of the packet. This allows it to know how many past and future symbols will interfere with the current symbol being decoded. We then design a sliding window block Viterbi decoder that jointly decodes bits by accounting for both future and past symbols. The decoder, however, suffers from high complexity. To address this, we incrementally build the decoding graph starting from the preamble and accounting only for valid transitions and the k highest probability states where k is a tuned parameter. As a result, we can maintain linear-time decoding with negligible degradation in performance.

The above μ-Link decoder, however, works under the assumption that the channel impulse response is static. As discussed earlier, the channel changes in the same order as the time it takes us to estimate the channel. Hence, constantly re-estimating the channel will result in a huge overhead that leads to a very low data rate. To address this, we modify the Viterbi decoder to jointly decode the bits and concurrently update the estimation of the channel. Specifically, the decoder uses the soft values of the decoded bits as a confidence

[1] Coherence time refers to the time for which the channel impulse response can be considered as static. Delay spread refers to the length of the channel impulse response, i.e., the time between the start and end of particles corresponding to the same symbol arriving at a receiver.

metric to decide which bits were correctly decoded. The correctly decoded bits are then used to update the estimate of the channel which is then used to improve the decoding.

We validate our MC channel model and the performance of our μ-Link decoder using an MC experimental testbed. The testbed follows a similar approach to past experimental testbeds on validating the molecular communication channel [75], [98]. Our testbed is shown in Figure. 4.6b. It consists of a network of narrow tubes and a pump that maintains a constant flow of water in the tubes. Another pump is used as a transmitter to release salt particles in the flowing liquid and a sensor is used as the receiver to measure the presence and concentration of these particles (See Section. 4.5 for more details). Our results confirm the MC channel characteristics we highlighted and show that by addressing these characteristics, we can improve the data rate and reduce the bit errors to achieve $5-10$ bits/sec with $0.2\%-5\%$ BER. This is $2.5\times$ the data rate for the same BER or $20\times$ lower BER for the same data rate as compared to [98] without the need to train and use neural networks which suffer from overfitting as we show in the results.

Contributions: The work has the following contributions:

- It highlights two new key characteristics of the MC channel. To the best of our knowledge, none of the past work addresses these two characteristics, which significantly limits the achievable data rates in practice.
- It presents a system that accounts for the characteristics of the MC channel and specifically addresses the issues of non-causality and long delay spread.
- It experimentally validates the findings and demonstrates improvement in achievable data rates and BER.

4.2 Background

MC is a communication paradigm inspired by chemical signaling between cells such as neurotransmitters and stimulating hormones. Similar to how neurotransmitters release molecules to signal an adjacent neuron, a device can transmit data by releasing molecules into the medium which are detected at a nearby receiver. While MC can work in both gaseous and liquid mediums, this paper focuses on MC in fluids due to its potential biomedical applications for communication between micro and nano-implants. In particular, we assume the device releases molecules into a flowing liquid in a micro-tube as shown in Figure. 4.6b.[2]

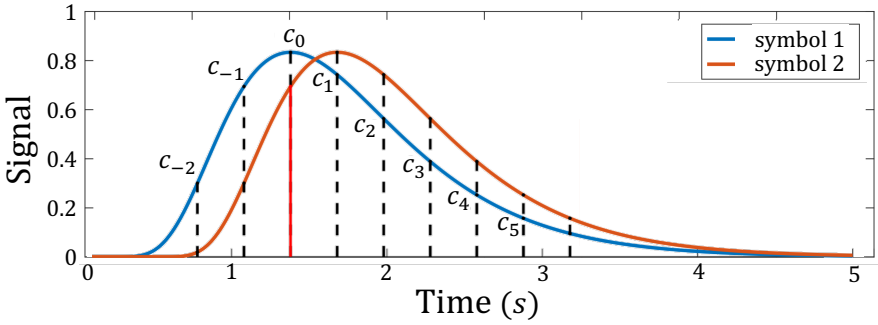
MC Transmitter

The transmitter can encode data bits by adjusting the amount of released molecules, the type of released molecules, or both [131], [132]. The simplest form of encoding is OOK (On-OFF Keying) where releasing molecules represents a "1" bit and not releasing molecules represents a "0" bit. Since the aim of this paper is to understand the characteristics of the MC channel, we adopt an OOK encoding. OOK helps simplify our analysis of the channel and is likely to be the most practical type of encoding for future applications like micro-implants.

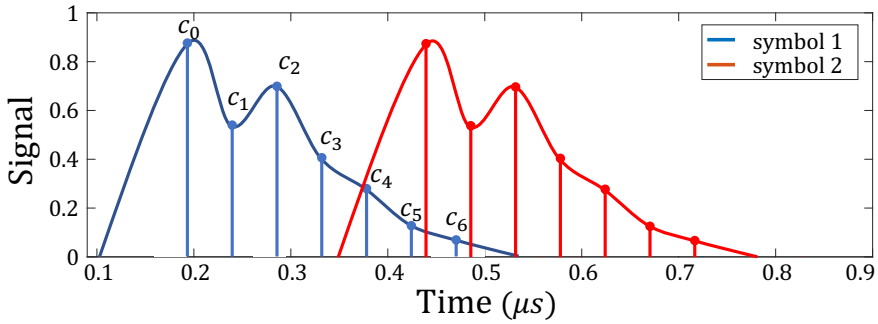
MC Receiver

The receiver can decode the transmitted bits by measuring the number of released molecules in the liquid. In theory, there are two types of receivers: destructive (active) or non-destructive (passive). Non-destructive receivers simply use a sensor to measure the concentration of molecules and decode the bits [79]. The

[2]Note that, in a gaseous medium like air, wireless RF is a far superior mode of communication as compared to MC. Hence, we focus on MC in fluids since RF signals do not propagate well in fluids.



(a) ISI in Molecular Communication



(b) ISI in Wireless RF Communication

Figure 4.1: Inter-Symbol-Interference (ISI) in MC vs. RF channels.

molecules are then eliminated through the bulk flow of the liquid. Destructive receivers, on the other hand, absorb the molecules from the medium in order to decode the bits [128], [133]. In this paper, we focus on non-destructive receivers since they are simpler to implement using sensors making them more practical for MC applications. We also assume the receiver response is linear, i.e., the measured value at the receiver is linearly related to the concentration/number of molecules at the receiver.

MC Channel

Section. 2.1 has provided a detailed description on the principles of signal propagation. In this chapter, we will refer to Equation.2.9 to help us understand the properties of molecular channel. However, unlike the majority of the work in the field which relies on accurate modeling and parameter estimation to simulate the CIR, we follow the procedure in conventional communication system and use the preamble to perform channel estimation.

4.3 MC Channel Characteristics

In this section, we highlight key unique characteristics of the molecular communication channel that differ from traditional wireless and wired communication. We use the theoretical model of the channel presented in the previous section to explain these characteristics. Later in section 4.6, we empirically validate that this model is linear and fits the channel impulse response (CIR) we see in practice. We also validate all the characteristics presented here. Note however that our μ-Link decoder does not use the theoretical CIR equation for decoding. Instead, it directly estimates and updates the CIR without trying to do parameter fitting.

We introduce two new characteristics which have been overlooked by past work: *Non-causal ISI* and *Long delay spread vs. Coherence time*. We also highlight a third one, *signal dependant noise*. While it has been covered by past work [73], [83], it is an important to account for it in μ-Link.

4.3.1 Non-causal Inter-Symbol Interference

In communication systems, data bits are encoded into data symbols and transmitted over the channel. Inter-symbol interference (ISI) occurs when signals corresponding to different data symbols arrive at the same time at the receiver, which creates interference. In wireless RF communication, for example, ISI is mainly caused by multipath where signals from previously transmitted symbols travel along longer reflected paths and arrive late at the receiver interfering with later symbols. In such cases, ISI is *causal* i.e. only previous symbols can interfere with current and future symbols. Future symbols, on the other hand, cannot interfere with current or previous symbols since they cannot arrive earlier.

This, however, is not the case in molecular communication where due to the randomness of diffusion, molecules from future symbols can arrive earlier. As a result, the ISI is non-causal and both future and previous symbols can interfere with the current symbol. In particular, molecules of previous symbols can linger and arrive late creating interference with the current symbol, and molecules of future symbols can arrive earlier than molecules of the current symbol creating interference as well. Such non-causal ISI does not typically occur in other communication systems. To the best of our knowledge, it has never been addressed by past work in MC which limits the decoding performance.

To better understand this problem, consider the example in Figure 4.1. Decoding achieves the highest confidence when the signal is sampled at its strongest, which is achieved around the peak of the channel impulse response. Let c_0 denote the peak of the unit pulse response as shown in Figure 4.1a. Let $c_1, c_2, ..., c_M$ denote samples on the tail corresponding to molecules that arrive later and influence the following symbols which we call *forward-ISI*. Let $c_{-1}, c_{-2}, ..., c_{-F}$ denote samples on the head that arrive early and influence previous symbols which we call *backward-ISI*. Let b_k be the transmitted symbol, the sampled concentration r_k is:

$$r_k = \sum_{i=-\infty}^{\infty} c_i b_{k-i} \approx \sum_{i=-F}^{M} c_i b_{k-i} \tag{4.1}$$

Thus, to decode symbol b_k, we need to know the M previously transmitted symbols as well as the F future symbols.

In RF communication, on the other hand, the following symbol can never arrive early before the current symbol as shown in Figure 4.1b. Hence, to decode symbol b_k, we only need to know the previously transmitted symbols. This significantly simplifies the decoder since we can decode symbols sequentially which will give us the previously transmitted symbols. Furthermore, since wireless signals travel at the speed of light, ISI occurs only between consecutive symbols.

4.3.2 Long Delay Spread vs. Coherence Time

In communication systems, in order to correctly decode the bits, the channel impulse response (CIR) must be estimated and corrected for. The CIR has two key properties: delay spread and coherence time.

• **Delay Spread:** represents the length of the channel impulse response in time. In wireless communication, delay spread is mainly caused by multipath where the signal travels along different paths and arrives at different times at the receiver. In this case, the delay spread represents the time between the arrival of the

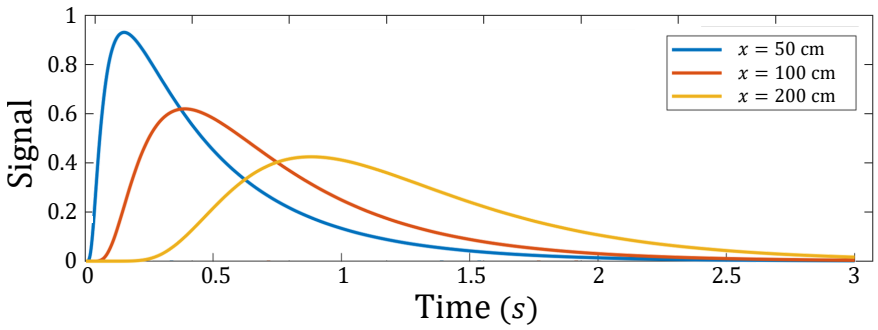


Figure 4.2: The molecular channel impulse response when the receiver is 50, 100, and 200 centimeters away from the transmitter.

first copy of the signal along the shortest path and the arrival of the last copy of the signal that is above the noise floor. However, since wireless signals travel at the speed of light, the difference in arrival time is typically of the order of nano to micro seconds. The delay spread also decreases as the distance between the transmitter and receiver increases since signals that travel along reflected paths start to arrive below the noise floor.

In contrast, the delay spread in molecular communication is mainly caused by diffusion of molecules where different molecules arrive at different times at the receiver. Since diffusion is slow, legacy molecules tend to remain in the channel creating a very long tail. As a result, the delay spread can easily extend to multiple seconds. Figure 4.2 shows an example of channel impulse response when the transmitter and receiver are 0.5, 1 and 2 meters apart. As can be seen, the delay spread can be as large as 2.5 seconds and expands as the distance increases, making the channel even harder to estimate.

• **Coherence Time:** corresponds to the time for which the channel impulse response can be considered fixed. In wireless communication, this is of the order of tens to hundreds of milliseconds which is orders of magnitude larger than the delay spread of the channel. Hence, it is possible to send a few training symbols at the beginning of a packet to estimate the CIR which can then be used to decode the bits throughout the packet. In molecular communication, however, the delay spread is of the same order as the channel coherence time. As a result, the channel changes as fast as the time it takes to estimate it. It is not possible to simply estimate the channel and use it to decode bits throughout the packet. Continuously estimating the channel, on the other hand, would result in a huge overhead that significantly reduces the data rate.

4.3.3 Signal-Dependent Noise

The randomness of diffusion creates signal-dependent noise. In particular, the receiver will sample the concentration of molecules. Let r_k be the sampled concentration at the receiver, then $E[r_k] = \mu_k = C(x_R, k\Delta t)$ where x_R is the location of the receiver. In case of multiple symbols, μ_k is given by Equation 4.1. Let N_k be the number of molecules at the receiver, then $N_k = V\mu_k$ where V is the detection volume of the receiver. The arrival of particles at the receiver follows a Poisson distribution with the same mean and variance. For tractability, the distribution can be approximated as Gaussian, i.e. $N_k \sim \mathcal{N}(V\mu_k, V\mu_k)$ [83]. Hence, the distribution of the sample concentration r_k is:

$$r_k \sim \mathcal{N}\left(\mu_k, \frac{\mu_k}{V}\right) \tag{4.2}$$

The variance of the above distribution is proportional to the concentration. As a result, the noise in the

signal is dependent on the signal itself and releasing more molecules to increase the signal strength increases the noise as well.

Compare this to wireless or wired communication where the noise is independent of the transmitted signal. There, the expectation of the distribution represents the signal whereas the variance represents the noise. The independence of noise has two important consequences. First, increasing signal strength significantly increases the Signal-to-Noise Ratio (SNR). Second, maximum-likelihood decoders can be mapped into simple decision rules for demodulation (e.g. in BPSK, it is sufficient to check whether the real part of the symbol is positive or negative to decode "1" or "0" bit.) . However, such decoders designed for Additive White Gaussian Noise (AWGN) channels do not work well over the MC channel [83].

4.4 μ-Link

This section describes the μ-Link communication protocol and decoder. μ-Link starts by sending a short preamble of known bits which are used for packet detection, channel estimation, and synchronization. The preamble is followed by data bits which are decoded using Hidden Markov Models and a Viterbi decoder customized to account for the specific characteristics of MC channels. μ-Link uses On-Off Keying and encodes a "1" bit by releasing molecules into the flow and a "0" by releasing nothing. Algorithm. 1 provides a high-level pseudo-code for μ-Link's overall protocol and decoder. In the rest of this section, we describe μ-Link in detail.

4.4.1 Channel Estimation & Synchronization

Every μ-Link packet starts with a "1" bit, i.e. a pulse of molecules released into the liquid. The receiver constantly monitors the concentration of molecules and when it detects an increase in the signal strength i.e. concentration level, it triggers packet detection. However, due to the molecular channel, a transmitted pulse will result in a slow rise in concentration of molecules at the receiver followed by a slow drop with a very long tail as we have shown in Figure. 4.2. Hence, packet detection is likely to be triggered ahead of the peak signal strength of the transmitted symbol. It is important, however, to synchronize our decoder to the peak of the signal in order to maximize the SINR of each symbol.

Consider the example of the received signal, shown in blue, in Figure. 4.3. The channel impulse response (CIR) is shown in red which spans multiple symbols due to its long delay spread. If we synchronize our symbols to the first rise in signal strength, then the signal value will not properly represent the bit transmitted in each symbol. However, if we synchronize our symbols to the peak of the channel impulse response, then the signal value will be aligned with the bit transmitted in each symbol. This allows the decoder to achieve the highest confidence in decoding each bit since the signal is sampled at its strongest value with the minimal ISI, i.e. maximum SINR, as can be seen from the dotted red lines in Figure. 4.3.

In order to synchronize the decoder, we must first estimate the channel impulse response (CIR) to detect the peak of the signal. A simple approach is to start the preamble by transmitting a "1" bit followed by a lot of "0" bits. In principle, this would allow us to detect the peak. However, sending a single "1" bit results in a weak signal that is very susceptible to noise which compromises the accuracy of our channel estimation. Transmitting additional "1" bits in the preamble would increase the concentration of molecules making channel estimation more robust to noise which is needed for decoding data bits. Hence, μ-Link's preamble is formed of a "1" bit followed by a pseudo-random sequence of bits similar to the one shown in Figure. 4.3. The pseudo-random sequence improves the accuracy of channel estimation.

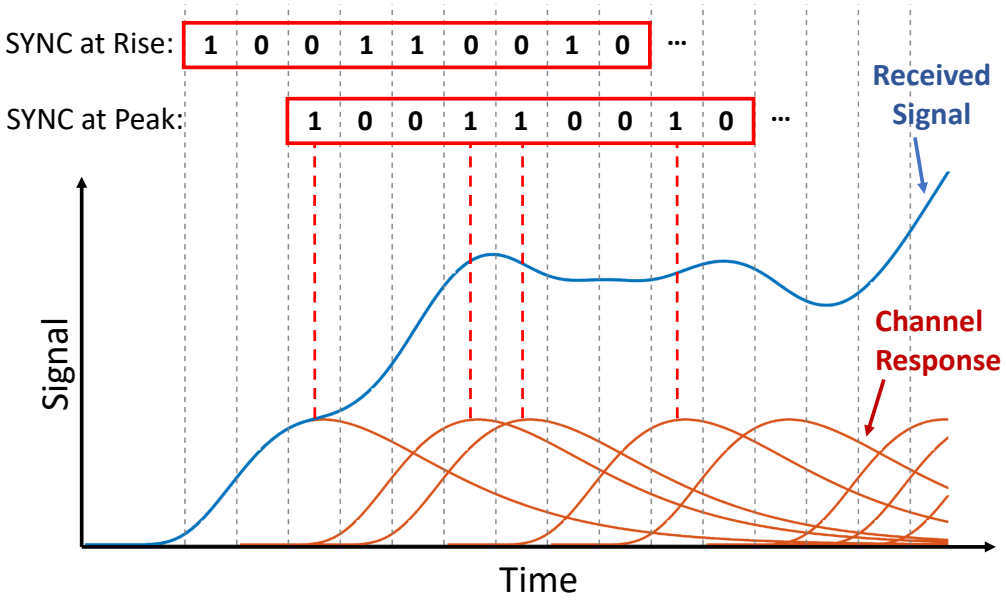


Figure 4.3: Blue curve depicts received signal without noise, while orange curve separates contribution of each bit. Synchronizing at the CIR peak will give higher SINR than the CIR start.

While μ-Link's design is driven by the theoretical channel model introduced in Equation. 2.9, μ-Link's channel estimation does not try to fit the estimated CIR to that model as it might not capture all practical constraints of the environment. Instead, μ-Link explicitly estimates the CIR using a Least Squares (LS) method. Specifically, the received signal can be written as $r(t) = c(t) * b(t)$ where $c(t)$ is the channel impulse response shown in Figure. 4.2 and $b(t)$ is the sequence transmitted bits. For the preamble, we can rewrite the received samples in vector format as: $\mathbf{r} = \mathbf{Bc}$ where $\mathbf{B}$ is the Toeplitz matrix of the preamble bits, and each column of $\mathbf{B}$ is a zero-padded version of the preamble bits. Thus, the LS estimator of $\mathbf{c}$ is:

$$\tilde{\mathbf{c}} = (\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T\mathbf{r} \tag{4.3}$$

where $(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T$ is the pseudo-inverse of $\mathbf{B}$.

After we estimate the channel $\mathbf{c}$, we can find the peak value of the CIR and align our decoder to the peak sample in $\mathbf{c}$. It is worth noting that the receiver can oversample $r(t)$ to improve decoding performance. In this case, $\mathbf{B}$ and $\mathbf{c}$ will represent oversampled versions of the preamble and the CIR.

4.4.2 Decoder

Decoding in molecular communication tends to be more complex than traditional wireless and wired systems. In particular, due to the long delay spread and non-causal ISI, each sampled value is a combination of many consecutive symbols including previous as well as future symbols which must all be jointly decoded. Moreover, widely adopted decoders for white Gaussian noise channels are non-optimal in molecular channels due to signal dependent noise. As a result, μ-Link adopts a Hidden Markov Model (HMM) to describe the molecular communication process and decodes using a sliding window block Viterbi algorithm.

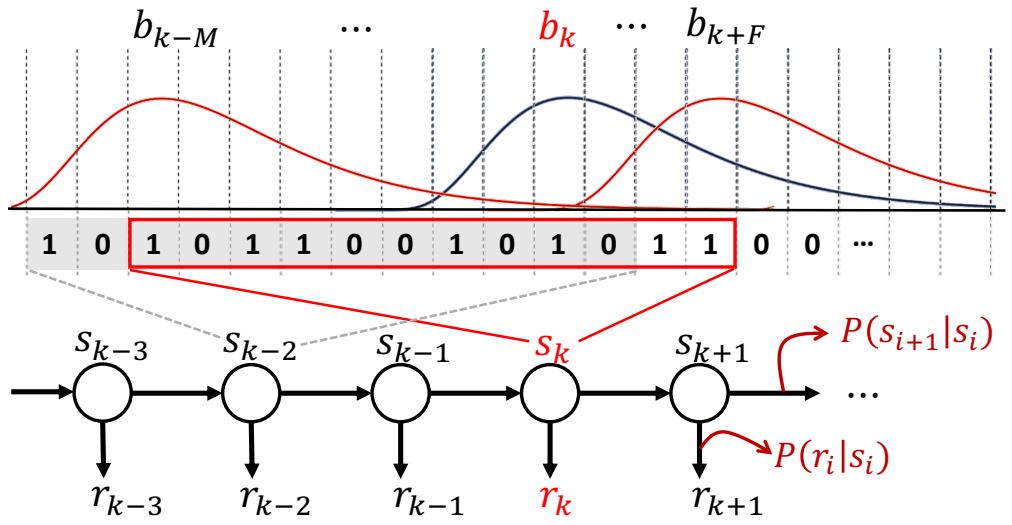


Figure 4.4: μ-Link's HMM where states represent a sliding window of bits accounting for both causal and non-causal ISI.

Hidden Markov Model

Recall that in HMMs, a hidden Markov chain consists of an unknown sequence of states, each of which generates (emits) an observation under certain probability. The observations are then used to infer the hidden states based on known transition and emission probabilities. In the context of μ-Link, the hidden states represent transmitted bits, and the observations represent received signals.

μ-Link defines an HMM state s_k as a sequence of transmitted bits that contribute to the observation of the received signal r_k as shown in Figure. 4.4. As described in section. 4.3.1, r_k is a combination of the current bit, the M previous bits, and the F future bits weighted by the channel impulse response. Thus,

$$s_k = \{b_{k-M}, b_{k-M+1}, \cdots, b_k, \cdots b_{k+F}\}, \tag{4.4}$$

i.e., the HMM states represent a sliding window of $W = M + F + 1$ bits that account for both causal and non-causal ISI where W is the delay spread. M and F are computed based on the estimated channel impulse response. For example, the channel shown in Figure. 4.4 has $M = 8$ and $F = 3$.

The transition probability between states depends on the modulation and coding scheme being used. In case of On-OFF keying, each Markov state has only two valid transitions depending on whether the next bit to be added to the sliding window is a "0" or a "1". Thus, give an arbitrary state $s_i = \{s_i[1], \cdots, s_i[W]\}$, the state transition probability is:

$$P(s_{i+1}|s_i) = \begin{cases} \frac{1}{2} & s_{i+1} = \{s_i[2], \cdots, s_i[W], 0\} \\ \frac{1}{2} & s_{i+1} = \{s_i[2], \cdots, s_i[W], 1\} \\ 0 & otherwise \end{cases} \tag{4.5}$$

Moreover, given a channel impulse response $\mathbf{c}$ and an arbitrary state $s_i = \{s_i[1], \cdots, s_i[W]\}$, the emission or observation probability is defined as:

$$P(r_i|s_i) = \mathcal{N}\left(\mu_i, \frac{\mu_i}{V}\right) \tag{4.6}$$

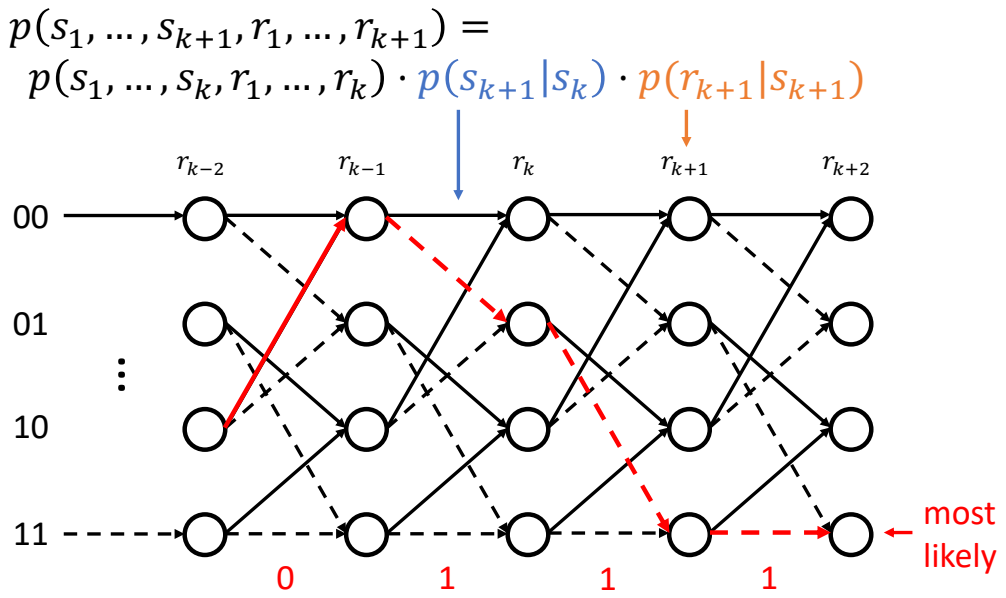


Figure 4.5: A simple example of a Viterbi decoding graph. Blue arrow points to transmission probability (hidden state), the orange points to emission probability (signal-dependent noise), and the red indicate the most likely path.

which follows from Equation. 4.2 for signal dependent noise in MC channels. $\mathcal{N}$ is the normal (Gaussian) distribution, μ_i is the expected value of r_i given s_i and $\mathbf{c}$, and V is the detection volume at the receiver. μ_i can be written as:

$$\mu_i = \sum_{j=-F}^{M} c_j s_i[M+1-j] \tag{4.7}$$

Block Viterbi Algorithm

The above HMM is solved using a block Viterbi decoder which returns the maximum-a-posteriori (MAP) solution of the Markov state transition sequence, i.e. the solution that maximizes the joint probability $P(s_1, \cdots, s_L, r_1, \cdots, r_L)$ where L is the total number of bits being decoded. The Viterbi decoder uses the following dynamic programming equation to solve the problem:

$$\begin{aligned} &P(s_1, \cdots, s_{k+1}, r_1, \cdots, r_{k+1}) \\ &= \max\{P(s_1, \cdots, s_k, r_1, \cdots, r_k) P(s_{k+1}|s_k) P(r_{k+1}|s_{k+1})\} \end{aligned}$$

Figure. 4.5 shows a simplified example of the Viterbi decoder for $W = 2$ where the most likely path is highlighted in red. However, for practical MC channels, W can easily exceed 10, similar to Figure. 4.4 where $W = 12$. The number of states in the Viterbi graph is 2^W. Thus, the Viterbi decoder requires $O(L2^W)$ memory and computation which for a 100-bit packet with $W = 12$ results in $> 400,000$ operations. To address this, we incrementally build the Viterbi decoding graph starting from the preamble and accounting only for valid transitions and the K highest probability states where K is a tuned parameter. Hence, only the top K states are considered at every stage of the Viterbi decoder which reduces the number of operations to $O(LK)$ resulting in a practical decoder complexity. For example, in our evaluation, when $W = 12$ and we

Algorithm 1: μ-Link Protocol & Decoder Pseudo-code

```
1  r ← sampled signal at receiver.;
2  b_p ← known preamble bits.;
3  t_start ← PacketDetection(r);
4  c ← ChannelEstimate(r, b_p, t_start);
5  M, F ← GetISI(c);
6  W ← M + F + 1;
7  HMM ← Viterbi.init(M, F, r, b_p, K);
8  t_sync ← Sync(r, c, t_start);
9  i ← t_sync + length(b_p);
10 while i < L do
11 |   // Viterbi Forward Pass;
12 |   HMM ← Viterbi.updateProb(HMM, r[i], c, K);
13 |   // Viterbi Backward Pass;
14 |   b̃[i − F] ← Viterbi.decode(HMM);
15 |   if i mod T = 0 then
16 |   |   c ← Viterbi.channelTrack(c, b̃, r);
17 |   end
18 |   i ← i + 1;
19 end
```

have 4096 states, we set $K = 128$. Furthermore, due to the use of a sliding window, the decoder does not need to memorize all the columns in the Viterbi graph shown in Figure. 4.5. Instead, it is sufficient to store the F previous columns which reduces the memory requirement to $O(FK)$. This allows μ-Link to efficiently adapt to longer packet lengths (i.e. larger L) and MC channels with longer delay spread (i.e. larger W).

Channel Tracking

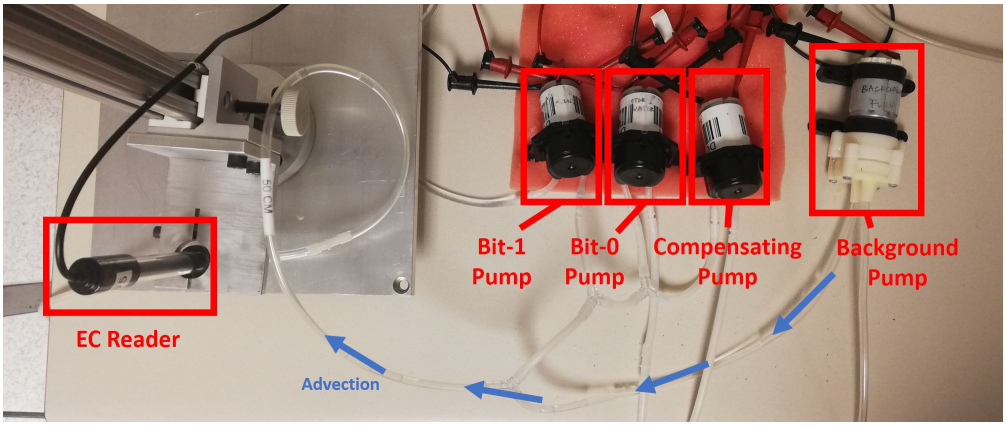
The above μ-Link decoder accounts for non-causal ISI, long delay spread, and signal dependent noise. However, it assumes that the channel impulse response is known and fixed. This is not true in MC channels where the channel coherence time is of the same order as the delay spread. Thus, the time it takes to estimate the channel is only slightly longer than the time it takes for the channel to change. Re-estimating the channel, by sending more preamble or pilot bits, would result in a huge overhead since the CIR is very long and we can only send a few data bits before we need to estimate the channel again.

μ-Link tracks the CIR throughout the packet without sending any additional preamble or pilot bits. To do so, μ-Link outputs the decoding results of the Viterbi dynamic programming algorithm without waiting till the end of the sequence as shown in line 14 of Algorithm. 1. After every T iterations of the Viterbi algorithm, μ-Link uses the decoded bits as a ground truth to re-estimate the channel impulse response. In order to minimize errors, μ-Link adopts the following:

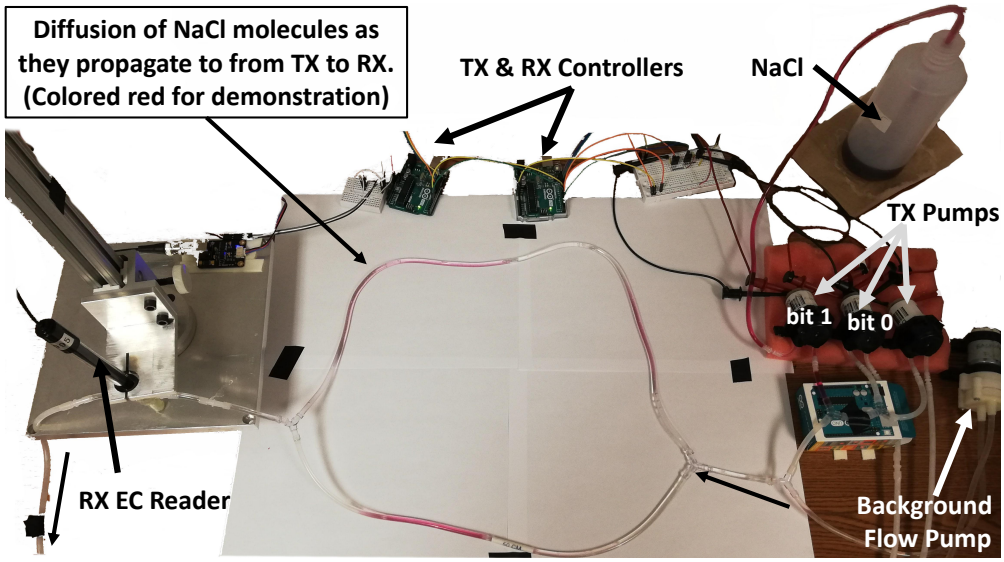
1. It leverages the soft values (probabilities) of the window blocks of decoded bits as a decoding confidence metric to decide which bits should be used for re-estimation.

2. It avoids the last F bits for which the decoded values cannot yet be trusted due to non-causal ISI from the current and futures bits that have not been decoded.

3. It rejects new CIR estimates that vary significantly from the most recent CIR estimate. In particular, if wrongly decoded bits are used, the CIR will exhibit huge changes that do not match the diffusion



(a) Single path.



(b) Forked path.

Figure 4.6: Molecular communication testbed using pumps and tubes.

model of MC.

The above criteria ensures that μ-Link generates reliable estimates of the channel throughout the packet.

4.5 Testbed

In order to evaluate μ-Link and our insights into the MC channel, we build an MC experimental platform shown in Figure. 4.6b and Figure. 4.6a. The testbed follows a similar approach to past experimental testbeds on validating the molecular communication channel [75], [98]. The testbed is built up with common and non-harmful materials, and can be divided into three parts: channel, transmitter, and receiver.

• **Channel**: The channel is mainly formed of a network of tubes with flowing water. A background pump continuously delivers water into the tubes to maintain a stable flow from the transmitter to the receiver and propel the propagation of molecules. The path between TX and RX is not limited to the single tube shown in Figure. 4.6a, but can be a combination of tubes/paths as shown in Figure. 4.6b.

• **Transmitter**: The transmitter is composed of three peristaltic pumps: two data pumps and one compensation pump. The data pumps either inject saline solution to encode a "1" bit or pure water to encode "0" bit. During a symbol interval, the data pumps only inject liquid for 50 ms. The compensation pump injects water for the remaining time in the symbol interval n order to compensate for flow variation.
• **Receiver**: The receiver uses an EC (electrical conductivity) reader, controlled by an Arduino board. The value output by the ADC is linearly related to the EC value which is in turn linearly related to the concentration of salt.

4.6 Evaluation

In this section, we start by validating the characteristics of the molecular communication channel presented in this paper. We then evaluate the performance of μ-Link and compare it with past work. Finally, we provide some micro-benchmarks that give insights into μ-Link's performance.

4.6.1 Channel Validation

We start by providing results and illustrative examples that validate the diffusion channel model presented in Equation. 2.9, the linearity of the channel, the signal dependent noise, the delay spread and the non-causal ISI. [3]

Diffusion Channel Model

Figure. 4.7 provides a model fitting of the channel model in Equation. 2.9 to a channel impulse response measured on our testbed in which the receiver is at a distance $x = 0.25m$ from the transmitter. We fit the CIR parameters as $v = 0.33m/s$ and $D = 0.025m^2/s$. The figure shows that the testbed provides a good platform for molecular communication that incorporates the advection-diffusion properties and closely matches the theoretical model. However, the equation cannot fully capture some of the hard-to-model properties in practical testbeds, such as the non-uniformity of diffusion and turbulence throughout the testbed.

Linearity

Our design as well as Equation. 4.1 assume that the receiver is linear, i.e. the measured concentration is a sum of the concentrations resulting from individually transmitted bits. To verify this assumption, we transmit 3 sequences of bits: S_1, S_2, and S_3 and measure the received signals R_1, R_2, and R_3 respectively. S_1 and S_2 are two random sequences with non-overlapping "1"s shown in red and blue in Figure. 4.8. $S_3 = S_1 + S_2$ and is shown in black. If the model is linear, then we should get $R_3 = R_1 + R_2$. Figure. 4.8 plots R_1, R_2, R_3, and $R_1 + R_2$ after averaging them over multiple runs to eliminate noise. The figure shows that R_3 closely matches $R_1 + R_2$. While the curves do not perfectly match, the approximation is sufficient to yield good performance in practice.

Signal-Dependent Noise

To measure the noise in the channel, we send a pulse followed by a long sequence of zeros to directly compute the channel shown on the top of Figure. 4.9. We repeat the measurement many times and compute the

[3] The channel coherence time property will be evaluated in Section. 4.6.3 by showing that if we do not track the changing channel frequently enough, our decoding performance will suffer.

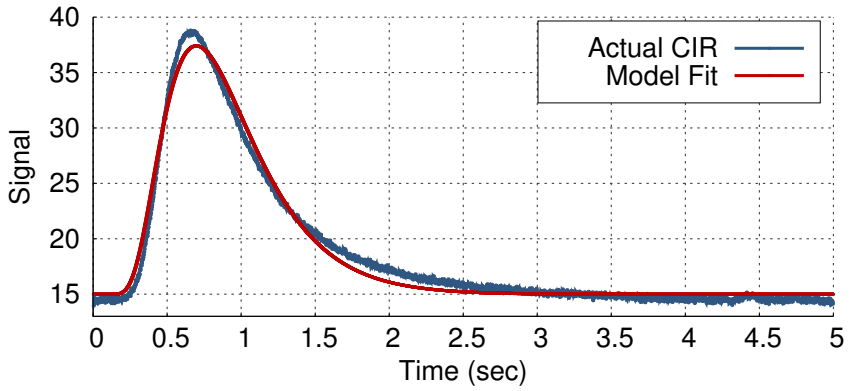


Figure 4.7: Curve fitting for single-tube MC channels to Equation. 2.9.

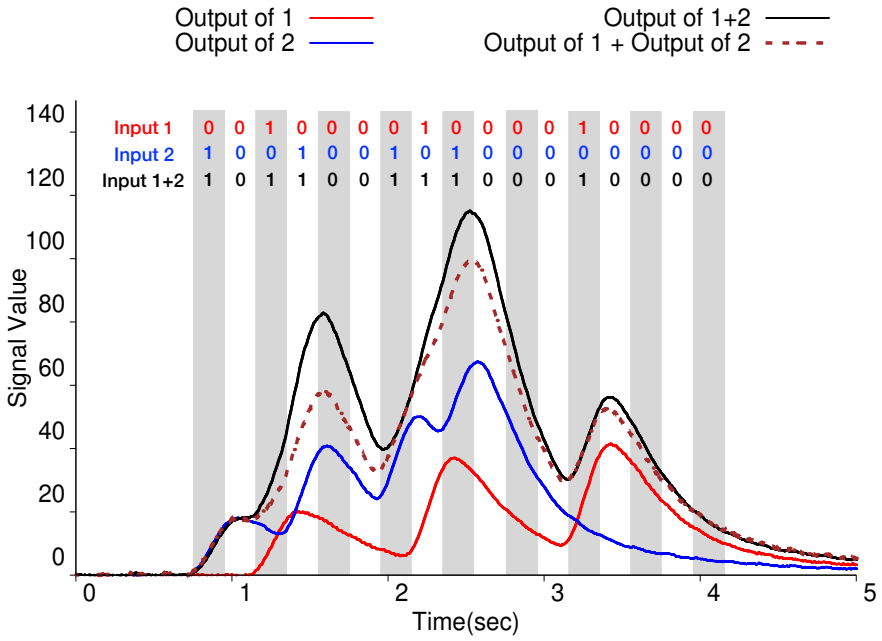


Figure 4.8: Channel linearity validation experiment. The sum of signals by two complementary sequences roughly matches the signal by the sum of these two sequences.

variance at each point of the channel across the measurements which we denote as the noise. Figure. 4.9 plots the noise power as a function of the signal power. While we cannot separate thermal Gaussian noise from signal-dependent noise, the graph clearly shows that higher signal power leads to a higher noise variance indicating that the noise is dependent on the signal. Note that the signal dependent noise is higher in the samples corresponding to the rising head of the CIR, likely due to imperfections in the transmitter pumps.

Long Delay Spread and Non-causality ISI

Figure. 4.10 empirically validates the long delay spread and non-causality properties of the molecular channel. After estimating each channel with the preamble, the delay spread is computed as the period from when the CIR first reaches 10% of the peak value until it drops below 10%. Similarly, the non-causal ISI duration is computed as the period from when CIR first reaches 10% of the peak value until it achieves the maximum

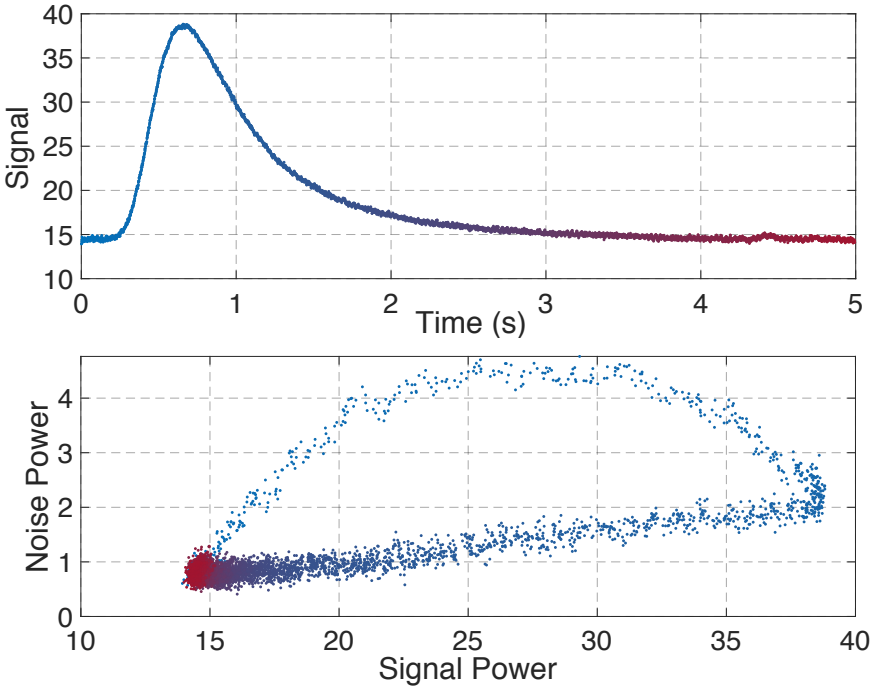


Figure 4.9: SDN validation experiment. CIR is computed from data between manually synchronized transmitter and receiver. Noise of each CIR sample is plotted in the same color as the sample.

peak. Figure. 4.10 plots the CDF of the delay spread and the non-causal ISI duration for when the receiver is 50 cm and 100 cm away from the transmitter. The median delay spread is 1.2s at 50 cm and increases to 1.5s at 100 cm which spans 12 and 15 symbols at a bit rate of 10 bps. The maximum, however, can reach 2.4s which would create ISI for 24 symbols at 10 bps. The median non-causal ISI duration is 0.27 sec and 0.36 sec for 50 cm and 100 cm respectively which affects 2 - 4 symbols at 10 bps. As we increase the bit rate, the influence of ISI extends to more symbols which presents a major obstacle for improving the data rate in MC.

4.6.2 μ-Link's Performance

In this section, we evaluate the performance of μ-Link.

Baselines

We compare μ-Link to two state-of-the-art decoding techniques:

- **Least Square (LS) Decoder:** The LS decoder is very similar to the LS channel estimation introduced in Section. 4.4. The main difference is that now we know the CIR and wish to estimate the bits. Hence, we write the CIR into Toeplitz matrix $\mathbf{C}$ and multiply its pseudo-inverse with the received samples to get the LS solution $\tilde{\mathbf{b}} = (\mathbf{C}^T\mathbf{C})^{-1}\mathbf{C}^T r$.
- **RNN decoder:** Recent work in [128] leveraged Recurrent Neural Networks (RNN) to decode MC. Specifically, [128] uses Long Short-Term Memory (LSTM) which is especially good at processing time sequences. The LSTM uses a sliding window for decoding and has a bi-directional structure, i.e. the sequence is analyzed in both forward and backward directions. We implement and train the bi-directional LSTM using the same neural network architecture and parameters described in [128].

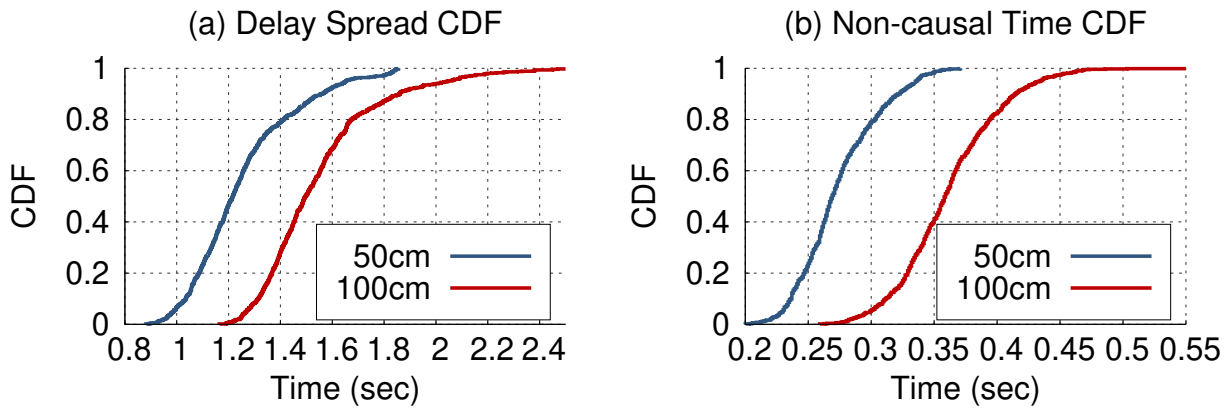


Figure 4.10: Delay spread and non-causality statistics. For 100 cm molecular channel, information particles take longer time to propagate from transmitter to receiver, so diffusion takes more proportion in CIR and makes more ISI than 50 cm channel.

Dataset

We collect a large dataset of molecular signals. Our dataset includes sequences with varying symbol interval (i.e. bit rate) and distance between transmitter and receiver. The dataset is divided into two parts. The first is for single-path channel shown in Figure. 4.6a. We test for distances of 25, 50 and 100 cm. The second part of the dataset is for a two-path channel shown in Figure. 4.6b. The paths have distances of 50 and 70 cm. In both parts, we vary the symbol interval between 100 and 500 ms which translates to bit rates ranging from 2 to 10 bps. For each experiment, we transmit 100 different sequences, each containing 100 bits. For the RNN decoder, 70% of the dataset is used for training.

Results

Figure. 4.11 summarizes the average BER performance as a function of bit rate for all three decoders: μ-Link, LS, and RNN. As expected, the BER degrades as the bit rate increases, i.e. the symbol length decreases. However, the LS decoder shows a major increase in BER with an increasing bit rate where at 10 bps the BER is above 20% for the single-path channels and 37% for the double-path channel. The BER degradation in case of μ-Link and the RNN is significantly less. For the single-path channels at 10 bps, μ-Link and the RNN achieve a BER of around 5%. However, for the double-path channel at 10 bps, μ-Linkachieves a BER of 5% whereas the RNN achievers a BER of 10%. Compared with single-path, a two-path channel incurs longer delay spread and more ISI. This is because the same amount of flow splits into two tubes, decreasing the flow velocity and increasing propagation time. In such a case, μ-Link's explicit CIR estimation outperforms the RNN's ability to implicitly account for the channel.

The advantage of μ-Link over the RNN is further emphasized once we consider the RNN's need for offline training on a large dataset specific to the channel. Training for one kind of channel is unlikely to generalize to other and possibly more complicated channels. To verify this, we run an experiment where we train the RNN for specific channels and test it on different channels in the dataset. Specifically, we omit data for certain distances or bit rates from the training set. Figure. 4.12 shows the performance of the RNN using different training sets but the same testing set on single-path channel. For the training set:

- "all dist" refers to distances of 25, 50, and 100 cm.

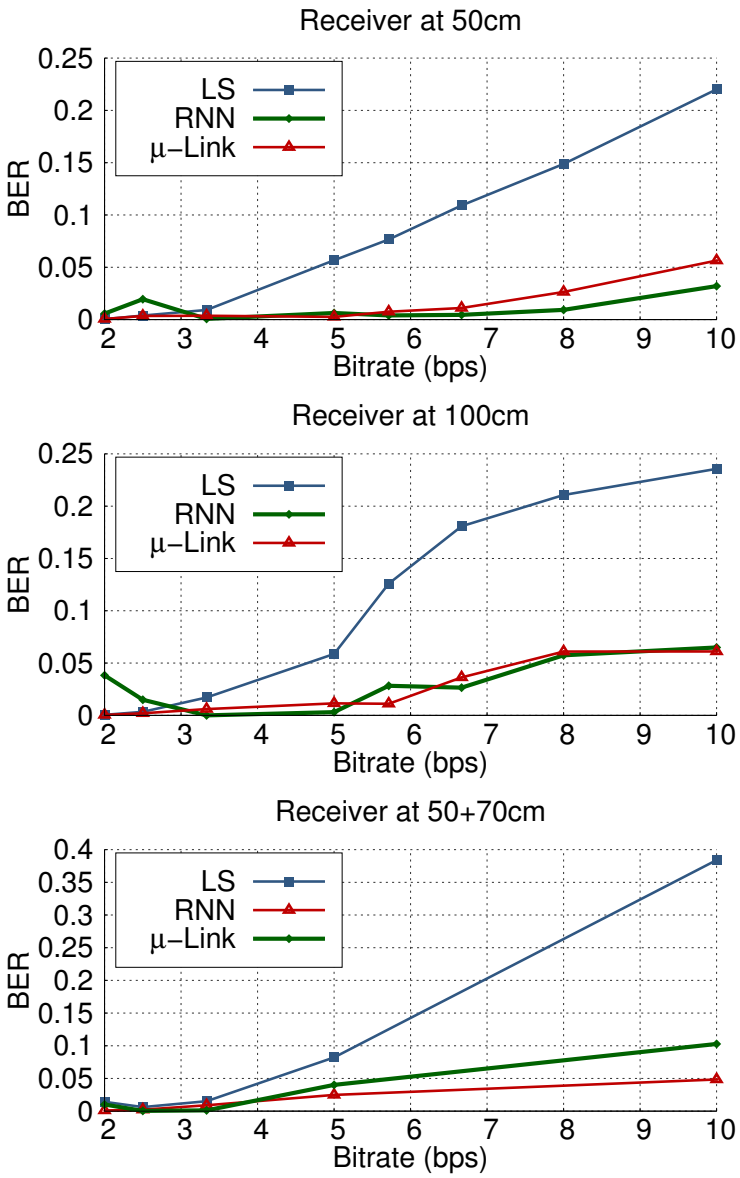


Figure 4.11: μ-Link, LS and RNN performance on single-path channel of 50 cm (top) and 100 cm (middle), and on double-path channel of 50 cm and 70 cm (bottom). μ-Link and RNN achieves comparable performance and which show great advantage over LS decoder.

- "all rate" refers to bit rates: 2, 2.5, 3.3, 5, 5.7, 6.7, 8 & 10 bps.
- "some dist" refers to 25 and 50 cm only.
- "some rate" refers to bit rates: 2, 2.5, 3.3, 5, & 10 bps.

Figure. 4.12 shows that the performance of the RNN is good only if the training set contains "all dist + all rate," covering all possible channels in the data set. However, as soon as, "some dist" or "some rate" are not in the training set, the RNN performance is extremely poor. For example, in the case "some dist + all rate," the RNN constantly performs worse with BER larger than 15% at bit rates above 5 bps. In the case of "all dist + some rate," the figure shows that the RNN performs well for rates that are included in the training

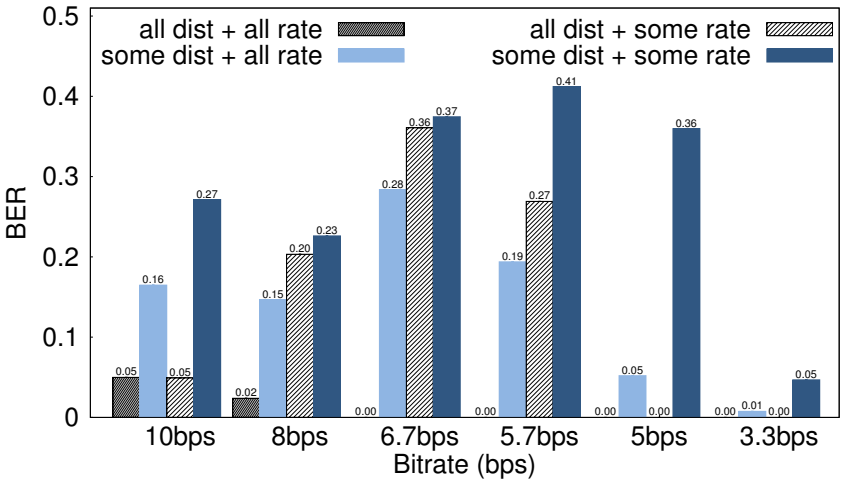


Figure 4.12: RNN overfitting issue. RNN model with four partially (or no) missing datasets are tested on the same sequences.

set such as 3.3, 5, and 10 bps and performs very poorly for rates that are not in the training set such as 5.7, 6.7, and 8 bps where the BER is larger than 20%. Thus, while RNNs can be useful for decoding MC, they require training for all possible channels and bit rates which is not practical.

Finally, μ-Link is able to significantly improve the BER and the achievable bit rates compared to LS decoding. It is able to achieve comparable performance to the RNN when the RNN's training set contains all channels and bit rates. However, significantly outperforms the RNN if the RNN is not trained for the specific channel and bit rate. Specifically, μ-Link can improve the data rate and reduce the bit errors to achieve 5 - 10 bps with 0.2% - 5% BER. Compared to the state-of-the-art experimental results reported in the RNN paper [128], μ-Link achieves 2.5× higher bit rate for the same BER or 20× lower BER for the same bit rate without the need to train and use deep neural networks.

4.6.3 Microbenchmarks

In this section, we present two microbenchmarks that give us insight into the importance of addressing the two new characteristics of MC channels presented in this paper. Figure. 4.13 shows the performance of μ-Link with and without its design components for addressing non-causal ISI and channel coherence time. The figure shows three curves: μ-Link, μ-Link without CU (Channel Update), μ-Link without NC (Non-causal ISI) and CU. The figure shows the average BER versus the bit rate at 50 cm distance between the transmitter and receiver. At a bit rate of 10 bps, μ-Link achieves a BER of 5%. However, this BER degrades to 22% if we do not track and update the channel and further to 29% if we do not account for non-causal ISI. This shows why past work that does not incorporate such insights remains limited in achievable bit rate to less than 4 bps.

Figure. 4.14 shows the performance of μ-Link as function of the frequency of tracking and updating the channel, i.e. parameter T in Alg. 1. The figure shows that the more frequently we update the channel (smaller update interval T), the lower the BER. This is particularly important for higher bit rates where the symbol interval is smaller. There, the contribution of ISI to many symbols becomes more prominent and having good estimates of the channel impulse response is necessary to ensure reliable decoding. For example, at 10 bps, where the delay spread (ISI) spans 12 symbols, updating the channel every 10 bits yields a BER of 5% which increases to 18.6% if we update it every 50 bits. Hence, channel coherence time is of the same order as

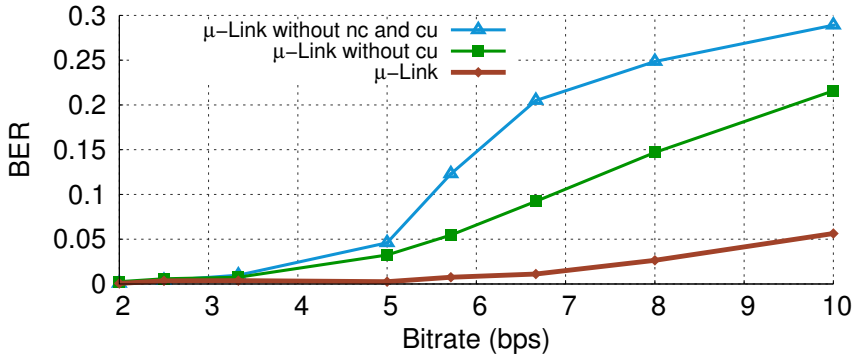


Figure 4.13: Importance of accounting for non-causal ISI (nc) and tracking the channel (cu) in μ-Link.

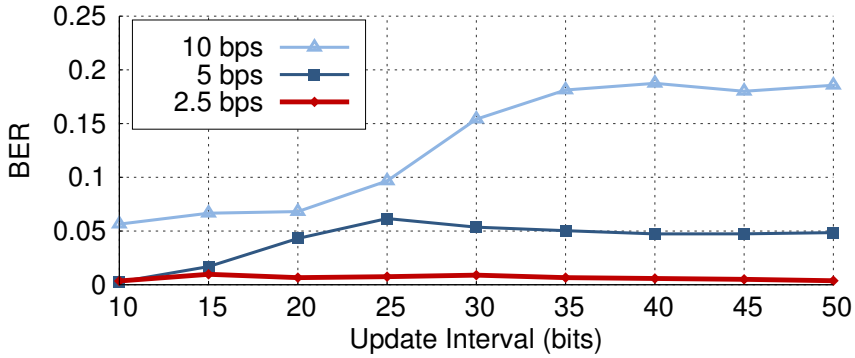


Figure 4.14: Importance of quickly tracking the channel in μ-Link.

the delay spread which presents a bottleneck for improving the bit rate in MC channels. By tracking the channel, μ-Link is able to significantly improve the BER for higher bit rates.

4.7 Related Work

In this section, we present further background and related work on molecular communication. As mentioned before, for a thorough literature review, we refer the reader to the following surveys [17], [21], [123].

4.7.1 MC Channel Modeling and Estimation

There has been a significant amount of work on modeling the MC channel taking into account different factors such as propagation environment, dimensions (1D, 2D, 3D), external influences like turbulence and advection, mobility, and type of receiver (destructive vs non-destructive) [64]–[69]. Time varying channel due to mobility is also modeled in [70], [71]. However, the models do not account for other time varying factors such as unstable advection that can be caused by pumping. A number of papers also focus on estimating the channel parameters such as distance, diffusion coefficient and fitting them to to a known model or closed-form expression [72]–[74]. Unfortunately, all these works are based on simulations and none of them are validated in practice. While modeling provides valuable insights, it cannot account for all factors and parameters that can influence the channel in practice as shown in [75]. In this work, we use theoretical models and empirical experimentation to guide our understanding of the channel and introduce new characteristics. However, for

decoding, we do not attempt to parameter fit our estimate to a closed form model but rather directly estimate the channel impulse response.

4.7.2 Channel Capacity

There has been theoretical work on studying the capacity of molecular communication [76]–[80]. However, to derive the capacity of the channel, these works must make simplifying assumptions that do not hold in practice. [79], [80] derive the MC channel with mobile transmitter or receiver but completely ignore ISI. [77], [78] consider signal dependent noise but assume a static channel. [76] considers signal dependent noise and ISI but assumes propagation through diffusion only. None account for non-causal ISI. Deriving the capacity is, indeed, valuable for understanding the data rate limits of MC. However, the channels tend to be significantly more complex in practice.

4.7.3 MC Decoders (Simulations)

Several MC decoders have been introduced and tested using simulations. Some exploit specific features of the MC channel response like monotonicity and local convexity [81], [82] but assume minimal ISI. Others introduce optimal decoders [83]–[85] but assume the channel impulse response is perfectly known a priori. [86] introduced a non-coherent decoder using blind channel estimation but it treats ISI as noise. These decoders were tested only in simulations and underestimated the severity of ISI in practice especially as we go to shorter symbols (higher data rates). Furthermore, they do not address non-causal ISI and the long delay spread versus coherence time.

Some theoretical work tries to address the severity of forward-ISI in MC channel using mechanisms to reduce the tail of the channel impulse response. Some propose using enzymes or photolysis reactions to remove interfering molecules retained in the channel [87], [88] while others propose using magnetic molecules and adding an external magnetic field [89], [90]. Unfortunately, these methods have not been tested in practice and can only reduce forward-ISI but not eliminate it. μ-Link, however, aims to decode in the presence of both forward and backward ISI.

4.7.4 MC Experimental Work

Unlike theoretical and simulation based past work, there are very few experimental MC testbeds especially in liquids. Most MC testbeds are for airborne molecular communication where the molecules are released in air [91]–[96]. The dynamics of molecular propagation in such gaseous mediums tend to be significantly different than liquids. The applications of airborne MC also tend to be limited since wireless RF works really well in air. In this work, we focus on liquid-based molecular communication for which very few testbeds exist [75], [97], [98]. All testbeds use water as a liquid environment. For molecules, they use magnetic particles [75], RNA [97] and acid/base [98]. [75], [97] focus on channel modeling and fitting and do provide an evaluation of data rate or BER. [75] uses a symbol interval 4 sec corresponding to 0.25 bps data rate, whereas [97] does not report symbol interval or data rate.

The closest to our work is [98] which uses deep neural network techniques including LSTM RNN to decode without explicitly modeling or estimating the channel. While it achieves good performance, it relies heavily on tedious training for every distance, channel, and bit rate. Hence, it suffers from overfitting as we have shown in section 4.6 and cannot deal with new channels and new bit rates that it has not seen before. μ-Link,

however, is able to achieve better performance by properly modeling the characteristics of the MC channel without relying on neural networks.

4.8 Limitations & Future Work

This paper only takes the first steps towards better understanding of the molecular communication channel. We believe the paper provides valuable insights that are driven by theoretical models, simulations, and experimentation. However, the empirical validation of these insights is limited to our testbed. Further experimental research is required which will enable more accurate channel models and new insights. In this section, we outline some of the limitations of the current system and future directions for this project.

- **In-Vitro Testbed:** The current testbed is limited. In particular, a more elaborate network of tubes, connections, and pumps that better resemble the circulatory system should be adopted. Moreover, testing on different types of molecules, receivers, and fluids, other than water, with different diffusion coefficients should be conducted.
- **Linearity:** While both theoretical and empirical results show that the MC channel is linear, the linearity of the sampled MC signal itself largely depends on the choice of receiver. In our testbed, we make sure that all our received signals fall into the linear range of the EC meter. However, this linearity might not be guaranteed for all types of molecules and corresponding receivers.
- **Modulation & Coding:** In this paper, we only consider OOK modulation without any coding since our main goal is to understand the MC channel. Further research is needed to explore different modulation and coding schemes that could enhance the performance of MC.
- **In-Vivo Experiments:** A future goal of this work is to validate the channel model in-vivo thought animal testing for micro-implants and micro-fluids in wet-labs. Such experiments might reveal new and different challenges in MC that would require further research.
- **Mobility:** The insights and system developed in this paper only applies to static transceivers. However, in some applications like targeted drug delivery, the device might move through the bloodstream resulting in abrupt or hard to track channel variations. Such changes are not captured by our system and are left for future work.

Chapter 5

MoMA

Chapter. 4 has studied the molecular channel properties that are challenging for the communication between a single transmitter and a single receiver. In this chapter, we will focus on identifying the practical challenges in scaling molecular networks beyond a single transmitter and designing protocols that enable multiple molecular transmitters to send data to a receiver, as shown in Figure. 5.1. However, such a task is not simple due to the following factors, which also explains why we consider the multi-TX-single-RX model (because it is no different from the multi-TX-multi-RX model).

First, it is hard to synchronize molecular transmitters due to the large propagation delay of molecules [134] and the added complexity of implementing a receiver in addition to the transmitter on the nano-implant as we explain in more detail in Section. 5.2. Moreover, molecular signals are likely to propagate in one direction in the blood vessels. As a result, schemes like TDMA (Time Division Multiple Access) cannot be implemented and packets can be transmitted at any time, i.e., while the receiver is decoding a packet, new packets might arrive.

Second, the number of available distinct molecules that can be released by a biological nano-machine is typically limited, and designing a receiver that can distinguish many molecules is a challenging problem [10], [104], [105]. Hence, techniques like MDMA (Molecule-Division Multiple-Access) [102] where each transmitter uses a different molecule (similar to FDMA where each wireless radio uses a different frequency) cannot scale.

Third, the molecular signal is non-negative as it corresponds to the concentration of released particles. This makes it hard to directly implement techniques CDMA (Code Division Multiple Access) as signals only add up constructively and correlating with the CDMA code cannot cancel interference from other transmitters. In contrast, CDMA in wireless networks like cellular and GPS use "+1" and "-1" which sum up destructively to cancel interference. To address this, past work adopts Optical Orthogonal codes (OOC) [106], [135] used in fiber optic networks where the signal is also non-negative. However, as we explain in Sec. 5.7 and show in our results in Section. 5.6, OOC does not work well in molecular networks. It requires very long codes which can significantly reduce the data rate, and its highly unbalanced code creates large fluctuations in concentration making it hard to detect and decode new packets. Finally, due to the diffusion of molecular particles, the molecular communication channel has a very long tail which results in high inter-symbol-interference (ISI) [136]. Unlike, fiber optic networks which can compensate for ISI in hardware [108], [110] and still use CDMA albeit with long codes, high ISI is unavoidable in molecular networks which necessitates accurate estimation and compensation of the molecular channel.

In this chapter, we present MoMA (Molecular Multiple Access), a protocol that enables multiple molecular

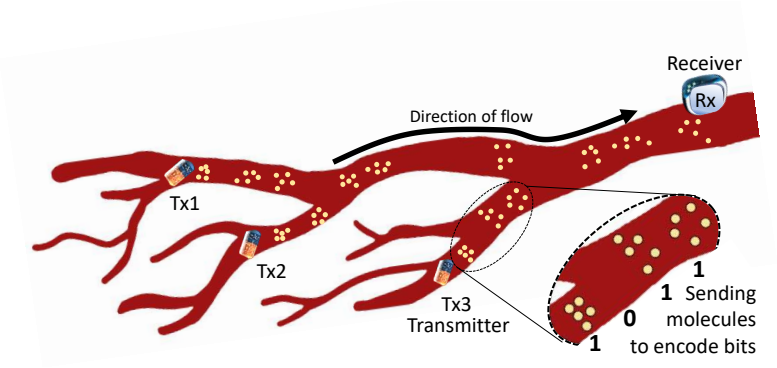


Figure 5.1: Molecular network with multiple transmitters.

transmitters to send their data to a central receiver that is capable of identifying and accurately decoding the data packets. MoMA opts for a CDMA-based scheme that uses Gold codes as it does not require synchronization and enables easy addressing of the different transmitters. However, MoMA modifies the code and packet structure and intertwines packet detection, channel estimation, and decoding of data bits to address the above challenges. Due to the non-negativity of the molecular signal and high ISI, failing to detect a single packet or inaccurately estimating a single transmitter's channel will bias the entire concentration of the received signal and completely corrupt the decoding. Hence, while decoding, MoMA constantly checks for new packets and re-estimates the channel.

To enable accurate packet detection, MoMA ensures that the signal power (concentration) throughout a data packet is stable and balanced while the power in the preamble fluctuates quickly. To do so, we modify the CDMA code to include only balanced codes i.e., the number of "1"s and the number of "0"s differ by at most 1. Moreover, as opposed to the standard approach [106], [135] of multiplying the code with the data bit which leads to no transmissions for "0" bits, MoMA takes an element-wise XOR of the code and the complement of the data bit, i.e., it sends the code as is for bit "1" and sends its complement for bit "0" which balances the power across the packet. For the preamble, MoMA repeats each element of the code a few times leading to a large sequence of pulse transmissions or a large sequence of no transmissions. Such fluctuations make the preamble easily distinguishable and allow the receiver to detect new incoming packets that start in the middle of previous packets being decoded.

To enable accurate channel estimation, MoMA jointly estimates the channels of all detected transmitters using adaptive filters. It also incorporates the unique properties of the molecular channel such as non-negativity and the long tail into the optimization function. Channel estimation is also used to further improve packet detection by rejecting falsely detected packets if their channel deviates too much within consecutive samples of the preamble or if deviates too far from the statistical model of the channel presented in Section. 5.1, i.e., the channel cannot look random and cannot drastically change within the preamble.

Finally, MoMA introduces a second molecule to each transmitter i.e., each transmitter sends two molecules. Unlike, MDMA which requires N distinct molecules to support N transmitters, MoMA only requires two which is still practical [104], [105]. Using two molecules helps eliminate false positives and false negatives in packet detection, improves the accuracy of channel estimation, and increases throughput by sending different

data streams on different molecules.

We implement and evaluate MoMA using a synthetic experimental testbed adapted from prior work [98], [114] and extended it to support multiple transmitters. The testbed, described in detail in Sec. 5.5, uses a network of tubes and pumps to emulate a constant flow of liquid with transmitters releasing particles into the liquid. Our results reveal the following: MoMA can scale up to four transmitters even in cases when the four packets from the transmitters always collide with random offsets. It can achieve a per transmitter throughput of around 0.89 bps which is 20× larger than prior work that uses OOC with CDMA for molecular networks [100] and 1.7× larger compared to a baseline that combines CDMA with MDMA. It also improves the BER (bit error rate) by 200× compared to [100] and by 10× compared to using OOC codes instead of MoMA's modified Gold codes. Moreover, MoMA's channel estimation improves BER by 10×, and the use of an additional molecule improves the packet detection rate by 20%.

Contributions: The chapter has the following contributions

- We present MoMA, a medium access protocol for molecular networks that enable multiple unsynchronized transmitters to transmit their packets at any time to a receiver that is able to accurately decode colliding packets.
- We introduce packet encoding, packet detection, channel estimation, and packet decoding techniques that are customized to molecular networks and leverage the unique properties of the molecular channel.
- We build a synthetic experimental testbed with four transmitters and evaluate the performance of MoMA to show its ability to support multiple transmitters and demonstrate significant improvement in performance.

5.1 Background

Apart from the challenges that will be explained immediately in the next section, it is important to keep in mind the challenges which have been explained in Chapter. 4–non-causal channel, long delay spread of similar scale as channel coherence time, and high ISI–still hold in this chapter. They will jointly influence the design in the molecular MAC protocol, but we will not repeat them in detail. Instead, we will next explain an important concept–CDMA (Code Division Multiple Access), which is borrowed from conventional communication systems and will be adapted to implement the medium access in molecular communication.

CDMA (Code Division Multiple Access) is a popular scheme used in wireless networks that allows multiple transmitters to transmit data at the same time by modulating their data bits with a unique binary code. The code should have good autocorrelation and cross-correlation properties, i.e., a transmitter's code c_i correlates very well with itself resulting in a peak if c_i is present in the signal. It also correlates very badly with the codes of other transmitters canceling out their signal. Ideally, the codes are orthogonal and the dot product $c_i \cdot c_j = 0$ if $i \neq j$. However, to achieve orthogonal codes, the length of the code L_c will be exponential in the number of needed codes G, i.e., $L_c = 2^G$. This exponentially reduces the data rate and is typically avoided in practice [137].

Gold codes are binary codes that are not perfectly orthogonal but still maintain good correlation properties while having a code length L_c that is linear in the number of codes G. As a result, they are widely used in CDMA systems, e.g. GPS [138]. Gold codes are generated using linear-feedback shift registers of size n that generate $G = 2^n + 1$ codes of length $L_c = 2^n - 1$ [139]. The codes have high autocorrelation: $c_i \cdot c_i = L_c$ and low cross-correlation:

$$\max_{i,j,i\neq j} c_i \cdot c_j = \begin{cases} 2^{(n+2)/2}+1 & \text{if } n \text{ even} \\ 2^{(n+1)/2}+1 & \text{if } n \text{ odd} \end{cases} \propto \sqrt{L_c} \tag{5.1}$$

In a set of Gold codes about half of the codes are balanced, i.e., the number of +1s and −1s in the code differ by at most 1. For example, the set of Gold codes with $n = 3$ are listed below where only the first 3 are balanced.

$$\begin{aligned} c_0 &= -1,+1,+1,+1,+1,-1,-1 \quad & c_1 &= -1,+1,-1,+1,-1,+1,+1 \\ c_2 &= +1,-1,+1,+1,-1,-1,+1 \quad & c_3 &= +1,+1,-1,-1,-1,-1,-1 \\ c_4 &= -1,-1,+1,-1,-1,+1,-1 \quad & c_5 &= +1,+1,+1,-1,+1,+1,+1 \\ c_6 &= -1,-1,-1,-1,+1,-1,+1 \end{aligned} \tag{5.2}$$

It is worth noting that Gold codes have poor performance for any n that is a multiple of 4. Hence, we are limited to using $n = 3, 5, 6, 7, 9, \cdots$ Finally, the individual bits of the code $c_i[m] = \pm 1$ for $m \in [0, L_c - 1]$ are referred to as chips.

5.2 Molecular Signal Informed Challenges

MoMA's goal is to enable multiple transmitters in molecular networks. We consider the most common case where multiple micro-implants transmit data to a central receiver [112], [113]. The receiver can potentially be a larger implant that is more powerful. For the purpose of this chapter, we only consider one-way communication from the micro-implants to the larger implant in which case the micro-implants can be as simple as possible or can even be implemented using biological cells [25], [27], [30]. We also assume that the receiver is placed downstream in the direction of blood flow making it easier for the particles to propagate from the micro-implants to the receiver. We adopt ON-OFF-Keying as the modulation scheme by releasing particles to encode a "1" bit and releasing nothing to encode a "0" bit, which also emphasizes our goal to keep the transmitters simple and push the complexity to the receiver.

Achieving multiple access in such molecular networks requires addressing the following challenges:

- **Non-negative signal:** Unlike typical wired and wireless signals, molecular communication signals are non-negative as they represent the concentration of particles. Hence, it is not possible to encode "+1" and "−1" chips of the CDMA code in a molecular signal and we must resort to using "1" and "0" instead. To understand why this is a problem, consider how CDMA works in today's wireless networks. Once we correlate the received signal with the code of a transmitter c_i, the "+1" and "−1" chips sum up destructively to cancel the signal from other transmitters. As a result, we can decode the packet of transmitter i while treating all the other signals as noise. This is no longer true when we transmit "1" and "0". The interference from other transmitters is always non-negative and will bias the decoder to think that the transmitter i is releasing particles even when it is not, leading to major errors in decoding. This problem is made worse when (1) the interfering transmitters have stronger signal strength (i.e. better CIR) and (2) when the decoder does not detect the packets of other interfering transmitters. Hence, it is not possible to decode packets independently and the decoder must detect all the transmitted packets as well as estimate their CIR in order to be able to decode the packets.

- **Lack of Synchronization:** Since the transmitters in a molecular network are not synchronized and

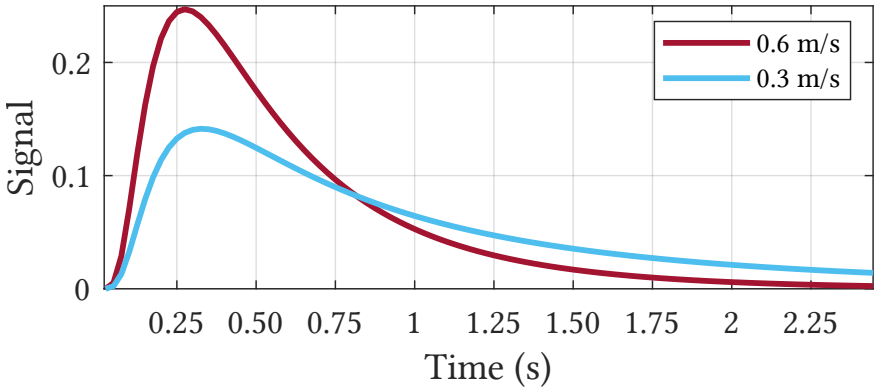


Figure 5.2: The channel impulse response in molecular communication for two different speeds of the flowing liquid.

cannot receive feedback from the receiver, we cannot implement typical multiple access schemes like TDMA. Furthermore, the transmitters are likely to transmit their packets at any time, i.e., a new packet might arrive while the receiver is already in the middle of decoding other packets. Simply discarding such a packet collision will lead to further degradation in what is already a relatively low data rate. Note that even if one were to afford the added complexity and equip micro-implants with receivers such that they can get feedback from a central implant, achieving synchronization is still difficult since (1) unlike wireless signals that travel at the speed of light, the propagation delay of molecular signals is very large, on the order of multiple transmitted bits and (2) it is not possible to diffuse data upstream in the opposite direction of blood flow. Having the particles loop around (through the heart, lungs, etc.) to reach the micro-implant will lead to an even larger propagation delay and very low signal strength.

- **High ISI:** Fig. 5.2 shows the simulation result based on Equation. 2.9 under two different flow speeds. It can be seen that the molecular channel has a very long tail that decays very slowly. As a result, molecular communication systems suffer from very high inter-symbol-interference (ISI) which further complicates the design and makes it hard to decode signals from different transmitters independently using CDMA. It also makes it difficult to decode without accurately estimating the channel of each transmitter.

The above challenges make accurate packet detection and channel estimation very difficult and at the same time essential for accurate decoding. In the following sections, we describe in detail how MoMA addresses these challenges to scale molecular networks to multiple transmitters.

5.3 MoMA Protocol

In this section, we describe how the MoMA transmitters encode their data packets to enable multiple access. As explained earlier, MoMA uses a CDMA-based multiple access scheme where each transmitter is assigned a CDMA code that is used to encode its data bits. The code also serves as a means to address (identify) different transmitters. Transmitters are not synchronized and can transmit at any time. Hence, their packets can collide with random offsets. To enable accurate packet detection at the receiver even when packets from many MoMA transmitters collide, MoMA ensures that the preamble demonstrates significant fluctuations in power whereas the data portion of the packet is balanced with a stable power across the data. MoMA also

uses two molecules per transmitter to increase the likelihood of correctly detecting the packet. Below we describe in detail how the MoMA transmitter encodes its packets.

5.3.1 Codebook

The codebook of MoMA is constructed from Gold codes described in Sec. 5.1. We only use Gold codes that are balanced, i.e., the number of "1" and "-1" in the code differ by at most one. In general, for parameter n, less than half of the $G = 2^n + 1$ generated codes are balanced. Hence, if we have N transmitters, we will generate Gold codes with parameter $n = \lceil \log_2(N+1) + 1 \rceil$. However, as mentioned in Sec. 5.1, Gold codes have poor performance when n is a multiple of 4 which is the case when $4 \leq N \leq 8$. In this case, using $n = 5$ would generate $G = 33$ codes of length $L_c = 31$ which will unnecessarily cut the data rate in half. Instead, we use $n = 3$ to generate codes of length $L_c = 7$ and append each code with a Manchester code such that every sequence becomes perfectly balanced. This results in codes of length $L_c = 14$ instead of 31. Finally, the chips of each code are represented as "1" and "0" instead of "+1" and "-1" to encode releasing molecules or releasing nothing.

5.3.2 Packet construction

The i-th transmitter is assigned a code $c_i = [c_i[0], \ldots, c_i[L_c - 1]]^T$. The preamble p_i of each packet has the length $L_p = R \times L_c$, which is generated by repeating each chip of the code R times. Let $\vec{1}_R^T$ be a vector of all "1" of length R. Then,

$$p_i = \left[c_i[0]\vec{1}_R^T, \ldots, c_i[L_c - 1]\vec{1}_R^T \right]^T \tag{5.3}$$

The consecutive pulses of "1"s or "0"s create large fluctuations in signal power making it easier to detect the preamble.

The data symbols d_i are encoded by performing elementwise XOR between the code and the complement of the data bit as follows:

$$d_i = \begin{cases} c_i \oplus \vec{0}_R^T & \text{if data bit is 1} \\ c_i \oplus \vec{1}_R^T & \text{if data bit is 0} \end{cases} \tag{5.4}$$

Effectively, we send the code if the data bit is 1 and we send the complement of the code if the bit is 0. Unlike past work that sends nothing if the bit is 0, MoMA's approach ensures that the power is balanced across the packet which improves both decoding and packet detection. Fig. 5.3 shows the power (concentration) of the received signal for both the preamble and the data portion of the packet when $R = 16$. As can be seen, the consecutive "1"s in the preamble lead to a buildup of concentration of molecules at the receiver, and the consecutive "0"s lead to a sharp drop creating large changes in power. On the other hand, the power in the data symbols is stable and does not vary drastically as the number of "1"s and "0"s are balanced, and due to high ISI, there is not enough buildup or drop in concentration.

Two points are worth noting. First, the total power of the preamble and the data symbols is the same, i.e., we are not sending the preamble at a higher power. We are simply rearranging the 1s and 0s to create a larger fluctuation in power. Second, one might assume that using the complement to send the 0 bit would hurt the decoding as sending nothing for L_c chips would cause a noticeable drop in concentration. However, note that this figure shows a single transmitter. When the signals from multiple transmitters combine, such

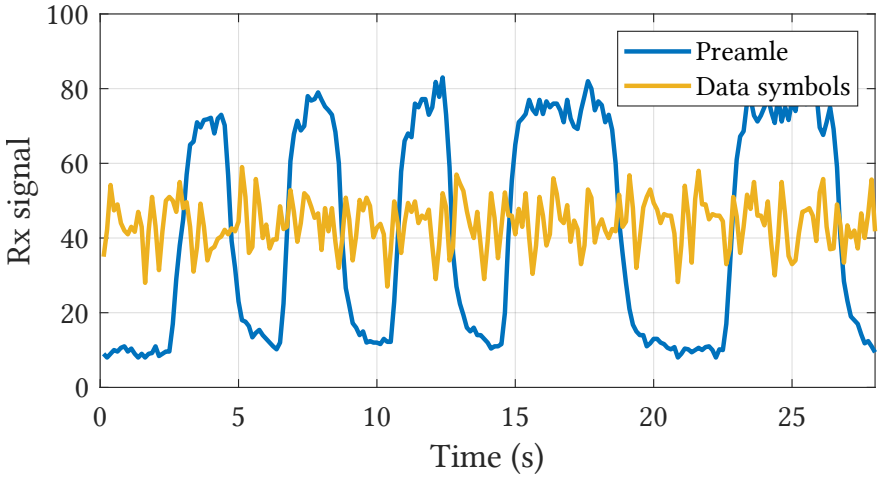


Figure 5.3: Fluctuation in power in the preamble vs. the data symbols.

large fluctuations in data symbols make it harder to decode as we show in our results section.

5.3.3 Multiple molecules

MoMA uses multiple molecules to help improve packet detection, decoding, and throughput. One option is to use multiple molecules, instead, to perform MDMA (Molecule Division Multiple Access) where each transmitter uses a different molecule and none of the packets interfere. However, as described earlier, the number of possible types of molecules is limited, and having a single receiver to decode all types is challenging [10], [104], [105]. Another approach is to combine MDMA with CDMA. Suppose we have N transmitters and use M molecules, we divide the transmitters into groups of N/M transmitters and use a CDMA code within each group. The packets from the different groups do not interfere. This is effectively similar to reducing the size of the network by $M\times$ which reduces the length of the code by $M\times$ and increases the data rate by $M\times$. We refer to these two approaches as MDMA and MDMA+CDMA. We compare with them our results in Sec. 5.6 to show that MoMA's approach yields better performance.

MoMA's approach is instead to have each transmitter use multiple molecules. Each transmitter is assigned a different code for each molecule and the assignment is legal as long as any two transmitters do not share the same code on the same molecule. An example of a legal assignment would be: transmitter i can use c_1 on molecule 1 and c_3 on molecule 2 whereas transmitter j uses c_6 on molecule 1 and c_1 on molecule 2. Besides, each transmitter can send different data streams on different molecules which also increases the data rate by $M\times$. While there is no gain in data rate compared to using MDMA+CDMA, MoMA is able to achieve higher overall throughput and better scaling for two reasons: (1) Using two molecules significantly improves packet detection as the probability of missing the packet on multiple molecules decreases exponentially with the number of molecules. (2) Different codes might have different performance, depending on the channel impulse response and the underlying data. Since the codes cannot be changed after deployment, having a bad code-channel combination can significantly harm the data rate of a transmitter. Using two codes on two molecules significantly reduces the probability of such a bad combination on both molecules.

5.3.4 Further scaling with multiple molecules

It is worth noting that MoMA's approach can yield further scaling gains if we allow some transmitters to share the same code on the same molecule. As long as, they do not share the same code on all molecules, it is still possible to distinguish and decode the data from the transmitters. For a codebook of size G, this allows us to scale the possible number of transmitters from $O(G)$ to $O(G^M)$. The methods are named code tuple and delayed transmission, which are both supported by the usage of multiple molecules.

5.3.5 Code Tuple

First, we propose the concept of code tuple, which refers to the code assignments of one transmitter on all molecules. It, allowing multiple transmitters to share the same code on some but not all molecules, greatly scales up the number of transmitters in an MC network. Theoretically this scheme is also applicable in conventional networks when the receiver jointly decodes all colliding packets with the knowledge of the CIRs. However, the benefits of decoding collisions do not beat the cost in hardware when the majority of collisions can be avoided by synchronization techniques. The more critical issue is that channel estimation is almost impossible when the two packets with the same code collide in the preamble.

As illustrated in Fig. 5.12, using the relation between CIRs is able to address collisions of multiple transmitters using the same code, as long as there is different coding on other molecules. Suppose the size of the codebook on each molecule is G, the total number of different code tuples for a system with M molecules is G^M. However, how to find the suitable number of transmitters in network is a complicated problem relating to multiple factors, such as the packet collision frequency, the various propagation loss, as well as the increasing signal dependent noise with more colliding packets. Besides, how to select the suitable ones among all the code tuples is also an interesting topic.

5.3.6 Delayed transmission

Another possible improvement of MoMA is to introduce delayed transmission. This delay does not refer to packets from different transmitters, which should be a common case without synchronization in MoMA. **Instead, one transmitter does not have to send packets on all molecules in the same time slot, but with specifically designed offsets.** For example, the packet on the second molecule starts one symbol later than the first, and the packet on the third molecule starts two symbols later than the second. With different delays introduced for transmitters using the same code tuple, the network can accommodate more transmitters. Even when the two transmitters share the same codes on all molecules, they can distinguish in the packet transmission order on molecules, such as the earliest packet of one transmitter is on the first molecule while another transmitter is on the second molecule. Besides, delayed transmission is expected with another advantage, which reduces the influence of sudden change in the channel especially for the preamble. The quality of packet detection and channel estimation could be greatly compromised due to the bursting noise when the packet arrives. By separating the preambles, the influence of such burst error could be reduced with the receiver signal on other molecules.

5.4 MoMA Decoder

In this section, we will explain how the MoMA receiver is able to decode the colliding packets from the transmitters.

Since packets can arrive at any time, the MoMA receiver must constantly keep checking for new packets while at the same time updating its estimates of each transmitter's CIR and decoding the data bits. To do so, MoMA uses a sliding window design. In each window, the decoder (1) performs packet detection to discover new packets; (2) estimates the CIR of new packets and updates CIR of already detected packets; (3) and decodes the data symbols. It does this for each molecule and combines measurements across molecules to improve packet detection. However, as previously mentioned, packet detection, channel estimation, and decoding are quite intertwined, as we will describe in detail in the following subsections.

5.4.1 Packet Detection

Packet detection is particularly important. As we will show in Section. 5.6, missing detection of a single packet will lead to an exploding BER. Hence, we opt for packet detection that favors false positives (detecting a packet that is not there) over false negatives (missing a packet). We can then use multiple molecules and channel estimation to eliminate detection errors. For simplicity, we will describe it under the case of a single molecule and then extend it to multiple molecules. The packet detection process works as follows. A full pseudocode can be found in Appendix. 2. We maintain a list of transmitters detected in the previous windows S_d. In each window:

Step 1. Initialize an empty list S_c of newly detected transmitters.

Step 2. Decode the data of transmitters in $S_d \cup S_c$ assuming no new packets will arrive in this window.

Step 3. Update the CIR of each transmitter and use it along with the decoded bit sequence to reconstruct the received signal y_d of these transmitters.

Step 4. Subtract the reconstructed signal y_d from the actual received signal y to get the residual signal y_r which might contain undetected transmitter packets.

Step 5. For each transmitter i that is not in $S_d \cup S_c$, compute the correlation of its preamble p_i with the residual signal y_r and find the peak of the correlation. If the peak is in the window of the preamble length, then a new packet might be present which indicates that the decoding in Step 2 was wrong.

Step 6. Iterate between (1) estimating the CIR of all TXs in $S_d \cup S_c \cup \{i\}$ and (2) decoding all these TXs' data. The iteration terminates when the decoding converges, i.e., we decode the same bits every time.

Step 7. Divide the preamble of the packet of transmitter i into two halves and estimate two CIRs of i from each half. If the two CIRs pass the similarity test described below, add transmitter i to S_c as a newly detected transmitter.

Step 8. Loop back to Step 5 for the next transmitter that exhibits a peak in correlation.

Step 9. If no more new transmitters were added to S_c, add S_c to S_d and move to the next window. Otherwise, loop back to step 2.

In the above packet detection scheme, the similarity test in step 7 aims to remove false positives. It is based on the fact that the CIR should not change drastically in a preamble period. After computing the CIR using the first and second half of the preamble, we compute the ratio of the total power of the two estimates. We then compute the correlation coefficient of the two estimates. The test fails if either the correlation or

the power ratio is below a threshold since the CIR should follow the model in Sec. 5.1 and should not look random.

To extend the above protocol to multiple molecules, we run the entire process in parallel on each molecule, but we average the peaks across molecules in step 5 and average the correlation coefficient in the similarity test in step 7. This reduces the probability of missing packets in step 5 and improves our ability to reject false positives.

Finally, as can be seen, the above packet detection is intertwined with channel estimation and decoding which we will describe in the following subsections.

Algorithm 2: MoMA Sliding Decoder

1 $S \leftarrow$ set of all transmitters;
2 $S_d \leftarrow \emptyset$;
3 **for** *each sliding window* **do**
4 $y \leftarrow$ receiver signal in the current window;
5 $S_c \leftarrow \emptyset$;
6 **while** $|S_d| + |S_c| < |S|$ **do**
7 **if** $|S_d| + |S_c| > 0$ **then**
8 $b \leftarrow$ decoded data for transmitters in $S_d \cup S_c$ with y;
9 $y_d \leftarrow \sum_{i \in S_d \cup S_c} b_i * h_i$;
10 $y_r \leftarrow y - y_d$;
11 **else**
12 $y_r \leftarrow y$;
13 **end**
14 $t \leftarrow$ peak location of correlating the preamble of each transmitter $i \in S/(S_d \cup S_c)$ with y_r;
15 $tx\_added \leftarrow false$;
16 **for** $i \in S/(S_d \cup S_c)$ *in the increasing order of* t **do**
17 $b \leftarrow$ decoded data for transmitters in $S_d \cup S_c \cup \{i\}$, which iterates between decoding and channel estimation until convergence;
18 $h_1 \leftarrow$ CIR of all transmitters in $S_c \cup \{i\}$, estimated with b and the samples of y which overlap with the 1st half of transmitter i's preamble;
19 $h_2 \leftarrow$ CIR of all transmitters in $S_c \cup \{i\}$, estimated with b and the samples of y which overlap with the 2nd half of transmitter i's preamble;
20 **if** h_{1i} *is similar to* h_{2i} **then**
21 $tx\_added \leftarrow true$;
22 $S_c \leftarrow S_c \cup \{i\}$;
23 **break for**;
24 **end**
25 **end**
26 **if** tx_added *is false* **then**
27 **break while**;
28 **end**
29 **end**
30 $S_d \leftarrow S_d \cup S_c$;
31 $h \leftarrow$ CIR estimation of all transmitters in S_d;
32 Decode all transmitters in S_d using h;
33 Remove all transmitters from S_d at end of packet;
34 Advance the sliding window;
35 **end**

5.4.2 Channel Estimation

In traditional systems like wireless, channel estimation typically takes place after detecting a new packet by using the preamble. However, past work has shown that the channel coherence time in MC is on the same order as the delay spread (the tail of the channel) [114]. Hence, the channel must be re-estimated and updated regularly throughout the packet to ensure accurate decoding. In the context of multiple access, this means that we cannot rely on CIRs estimated in the previous window and we need to re-estimate the CIR in every window as shown in the previous section. Moreover, since the received signal is the sum of packets from different transmitters, we cannot estimate the CIR of each transmitter independently, i.e., we must jointly estimate the CIR of all transmitters.

We estimate the channel for each molecule independently. In Section. 5.3.4, we showed how we can leverage multiple molecules to improve the channel estimation. Let y be the received signal, h_i be the channel impulse response, x_i be the transmitted signal of transmitter i, and n be the noise vector. For simplicity, we will sum in the below equations over all transmitters $i \in [1, \cdots, N]$. However, in practice, we should only sum over the transmitters that were detected during packet detection, i.e., transmitters in the set $S_d \cup S_c$ from the previous section. Then,

$$\begin{aligned} y &= \sum_{i=1}^{N} h_i * x_i + n = \sum_{i=1}^{N} X_i h_i + n = Xh + n \\ \text{where} \quad X &= [X_1, \ldots, X_N] \\ h &= \left[h_1^T, \ldots, h_N^T\right]^T \end{aligned} \tag{5.5}$$

Although linear matrix inversion can be directly used to compute the least squares channel estimate from Equation. 5.5, it does not account for the unique features of the molecular channel and leads to sub-optimal decoding performance as we show in Sec. 5.6. To account for the characteristics of the MC channel, we formulate channel estimation as an optimization problem whose loss term is composed of the following:

- **Least Squares loss.** This is the term that directly corresponds to the linear inversion of Equation. 5.5, which describes the distance between the actual received signal and the expected received signal constructed from the estimated CIR. Suppose the length of the signal y is L_y, then

$$\mathcal{L}_0 = \frac{1}{L_y} \|y - Xh\|^2 \tag{5.6}$$

- **Non-negativity loss.** This arises from the fact that the molecular signal is fundamentally the concentration of particles, which is strictly non-negative. Hence, we penalize the negative terms in the estimated CIR. Suppose the length of CIR h_i is L_h, then

$$\mathcal{L}_1 = \frac{1}{L_h} \sum_{i=1}^{N} \|\text{ReLu}(-h_i)\|^2 \tag{5.7}$$

- **Weak head-tail loss.** As can be seen from Fig. 5.2, the CIR of a molecular channel should start with a weak head and end with a weak tail. Thus, we penalized the non-zero taps that are far away from the

peak CIR sample. Let $a \odot b$ denote the element-wise product of vectors a and b, then

$$\begin{aligned} \mathcal{L}_2 &= \frac{1}{L_h^2} \sum_{i=1}^{N} \|g_i \odot h_i\|^2 \\ \text{where } g_i &= [1 - q_i, \ldots, L_h - q_i]^T \\ \text{and } q_i &\text{ is the index of the peak sample} \end{aligned} \tag{5.8}$$

- **CIR similarity loss.** CIRs of the molecular network reveal high correlation when they share the same molecule diffusion coefficient and/or the same transceiver distance. Defining $x' = x/D$ and $t' = t/D$, the diffusion model Equation. 2.8, can be transformed into

$$\begin{cases} \dfrac{\partial c}{\partial t'} + \dfrac{\partial}{\partial x'}(v_x c) = \dfrac{\partial^2 c}{\partial x'^2} \\ c(x', t')|_{t'=0} = D\delta(0) \end{cases} \tag{5.9}$$

  This indicates that the CIR $C(x,t)$ can be achieve from another one $C(x',t')$ by scaling in the amplitude and the time scale. Theoretically, the channel estimation can include a step of parameter fitting that unifies the common parameter in the estimated CIRs and improves the performance. However, an accurate channel model is hard to generalize to the various practical molecular environments. But, when the diffusion coefficients D of the molecules do not diverge much, the two CIRs can be compared directly with only scaling in amplitude. Thus, we only take a preliminary check among CIRs of the same transmitter and penalize those deviating from the average in this chapter. We introduce another loss term, similarity loss ($\mathcal{L}_3$), in the channel estimation algorithm. Note that this loss term, unlike the other loss terms mentioned above, is only applicable to multi-molecule channel and cannot be computed separately for each molecule.

$$\begin{aligned} \mathcal{L}_3 &= \frac{1}{L_h} \sum_{i=1}^{N} \sum_{j=1}^{M} \left\| \vec{h}_{ij} - a_{ij} \vec{h}_i^a \right\|^2 \\ \text{where } \vec{h}_i^a &= \frac{1}{M} \sum_{j=1}^{M} \frac{\vec{h}_{ij}}{a_{ij}} \text{ , and } a_{ij} = \|\vec{h}_{ij}\| \end{aligned} \tag{5.10}$$

By adding the above loss terms with different weights W_1, W_2, we formulate the molecular channel estimation problem as the below optimization. [1]

$$\min_{h_i, i=1,\ldots,N} \mathcal{L}_0 + \mathcal{L}_1 + \mathcal{L}_2 + \mathcal{L}_3 \tag{5.11}$$

MoMA solves the above optimization through an adaptive filtering algorithm using iterative gradient descent. The adaptive filter is initialized with the Least Squares solution to Eq. 5.5 which can also be used to initialize q_i in Eq. 5.8. Finally, once the adaptive filter converges, we can also compute noise power n by comparing the reconstructed Xh and the actual value y of the received signal. We can then use the noise power for decoding in the following section.

[1] The weights of the loss terms are not perfectly tuned in this chapter but may lead to potential performance improvement.

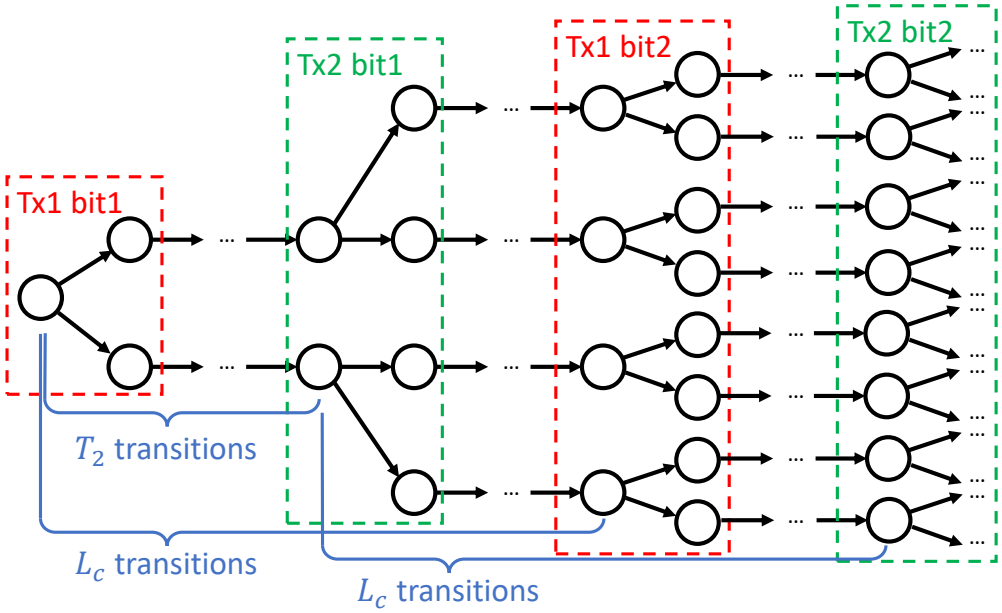


Figure 5.4: Chip-based Viterbi state transitions for two transmitters. L_c is the length of the coding, and T_2 is the time offset (in the unit of chips) between the two transmitters. The omitted transitions only have one outgoing state.

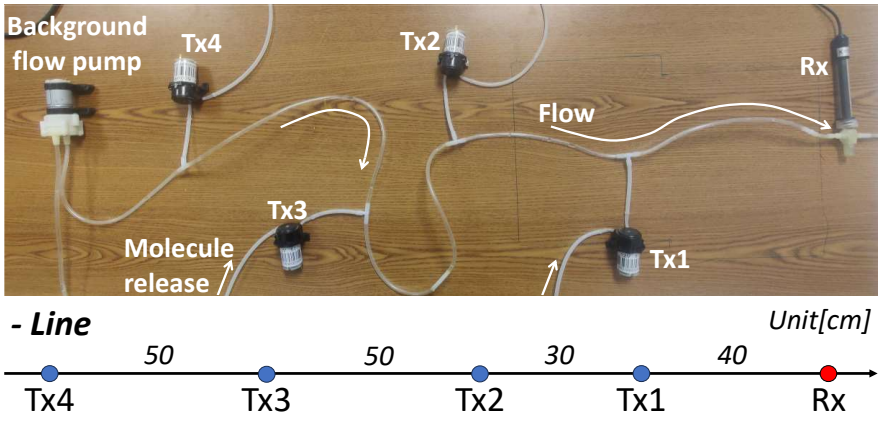
5.4.3 Viterbi Decoder

We use a Viterbi algorithm to decode the data bits which we modify to account for multiple transmitters, the CDMA codes, and the MC channel. The Viterbi algorithm [140] is a dynamic programming algorithm for obtaining the maximum a posteriori probability estimate of the most likely sequence of hidden states—-called the Viterbi path-—that results in a sequence of observed events, typically in the context of Hidden Markov Model (HMM). In general, the HMM for MC transmission has the following components:

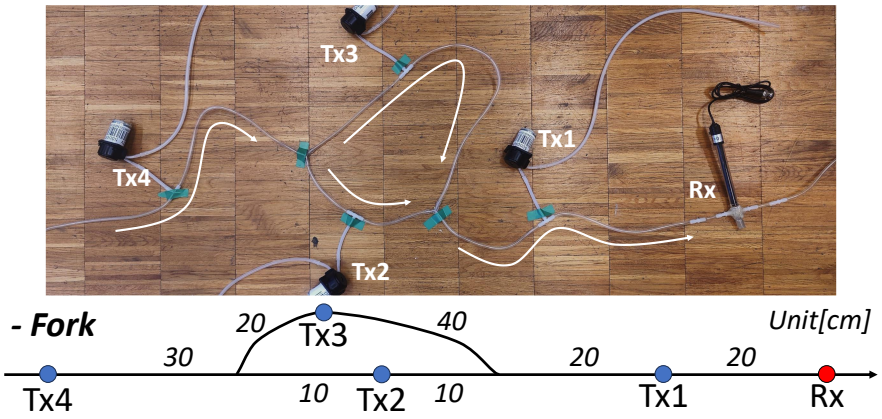
- Hidden Markov Chain: The state of the Markov Chain is a sequence of consecutively transmitted bits whose particles remain in the channel and influence the received signal. The length of the sequenced is decided by the length of ISI, and the transition between states happens when a new data bit is sent by the transmitter. All the possible transitions construct a large-scale Viterbi trellis shown in Fig. 5.4, and the algorithm aims to find the path with the highest probability.

- Observed events: The signals received serve as the observations of the HMM. First, expectation is computed from the current state of the HMM and the CIR estimated from the preamble. Then, the probability of the actual observation is computed with its divergence from the expectation and the noise power in the system. Such probability accumulates as the Viterbi path prolongs as more receiver signals are collected with time.

For the case with a single transmitter and a single receiver without coding or oversampling (e.g. [114]), the state of the HMM is a sequence of data bits each with one sample of the receiver signal as observation. With CDMA codes and multiple transmitters, we modify the model by changing the number of observations for each state to the length of the code. However, this HMM based on data bits is not accurate to describe the MoMA, due to the lack of synchronization among multiple transmitters. Considering that the packets



(a) Line Topology



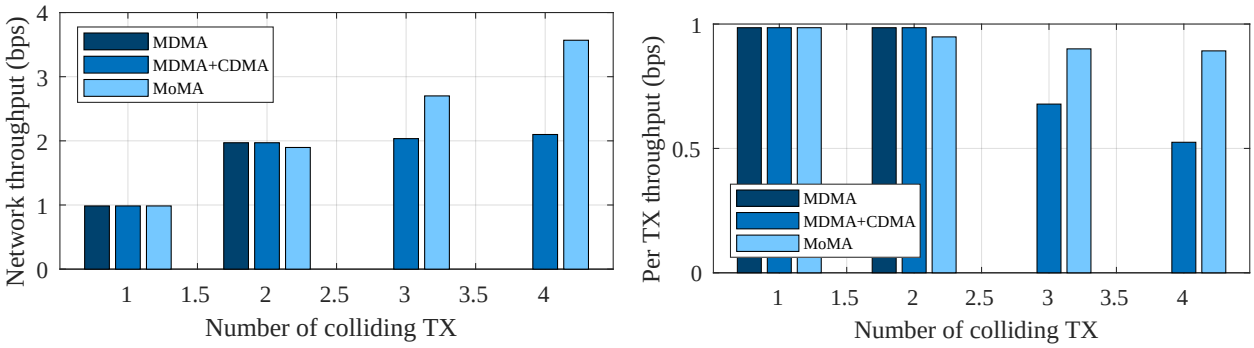
(b) Fork Topology

Figure 5.5: A synthetic liquid testbed with four transmitters and one receiver. Top is a line topology, while bottom is a fork topology. (Circuitry of the testbed is not shown for clarity.)

could arrive at the receiver with offsets, MoMA uses one or more sequences of chips (instead of data bits) to represent the HMM state, where each sequence represents one of the detected packets. With chip-rate sampling, each state still has one receiver sample as the observation. The major difference from single transmitter without CDMA is that not all HMM states will transition to two different states. From the perspective of each transmitter, such transition only happens when the first chip of the data symbol comes into the state sequence, while for the other states the transition is deterministic according to the CDMA code. From the perspective of all transmitters, there is the possibility that one HMM state could transit to more than 2 states (i.e. power of 2) when multiple transmitters are coincidentally synchronized at symbol level. In general, the multiplication of Viterbi states for each transmitter has a period equal to the CDMA code length, but with random delays.

5.5 Testbed

We built a synthetic experimental testbed, as shown in Figure. 5.5, to evaluate MoMA. This testbed is centered around the channel with either a single path or a fork in the middle, where on the one side a background flow pump continuously pumps water through the tube to the receiver on the other side. In



(a) Total network throughput.

(b) Per TX throughput.

Figure 5.6: Throughput of molecular network with different number of transmitters and multiple access protocols. The throughput is the lower bound for each transmission, as all transmitters are intentionally made to collide with different offsets.

addition, four other tubes are interconnected with the mainstream at various distances from the receiver, each of which has a pump that can inject bursts of information molecule solution into the mainstream. These four pumps are controlled electronically with transistor circuits and serve as the communication transmitters in the molecular network. Two different channels are evaluated (the line channel and the fork channel, which are presented in Figure. 5.5), while most of the results in Section. 5.6 are under the line channel. The fork-channel figures, as a support, will be marked out in the title.

In this testbed, we use *NaCl* as the information molecule, and we use an Electric Conductivity (EC) reader that indicates the *NaCl* concentration in the solution as the communication receiver. The TX pumps and the RX EC reader are controlled by an Arduino Mega 2560 REV3. Since most of the common soluble molecules will change the electric conductivity with high possibility, which means that introducing them to will infer the measurement of *NaCl*, we cannot conduct real-world experiments with two types of molecules in the current stage. However, we will discuss testbeds using other molecules as well as possible methods to implement a multi-molecule testbed in Section. 5.8. Alternatively, we use emulation to evaluate multiple molecules. For each data point reported in the paper, we repeated the experiment of one molecule 40 times with different data streams and code assignments. To emulate two molecules, we randomly pick two experiments of the same transmitters and concurrently process them, which assumes that the two molecules are not interfering. Each data point of the two molecules includes 500 such emulations.

5.6 Results

5.6.1 Main Results

We evaluate MoMA in a network with 4 transmitters, and each transmitter can emit 2 types of molecules. To demonstrate the scalability of MoMA, we compare the total network throughput and per transmitter throughput as we intentionally cause different number of transmitters to collide. We compare MoMA with the following two baselines.

- MDMA (Molecule-Division Multiple-Access): Each transmitter only uses a distinct type of molecule, and we assume that the receiver can independently measure each type of molecule. Under this scheme, OOK

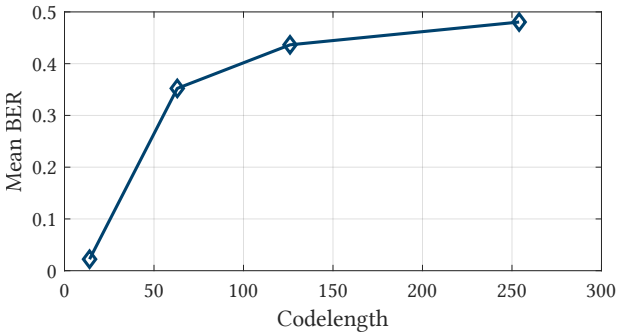


Figure 5.7: BER with different code lengths but the same data rate.

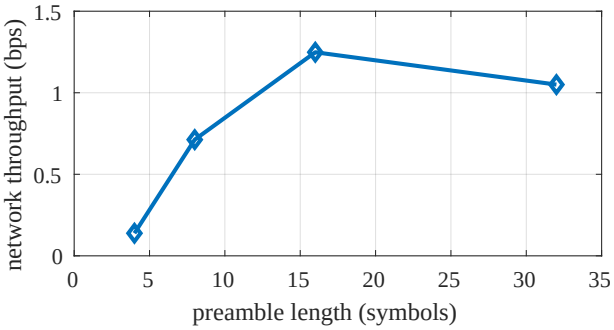


Figure 5.8: Network throughput with different preamble lengths.

(On-Off Keying) can be used to encode data symbols and pseudo-random sequences as the preambles. Note that MDMA requires the number of usable molecules to be greater than or equal to the number of transmitters.

- MDMA+CDMA: When there are more transmitters than the available types of molecules, one has to combine MDMA with CDMA. We first evenly divide all transmitters among the molecule categories and then assign distinct CDMA codes to different transmitters using the same molecule.

Since these two baselines can be viewed as special cases of MoMA, we use the same decoder, and we assume that the receiver drops packets with BERs greater than 0.1. For the fairness of throughput comparison, we normalize the transmission data rate to 2/1.75 bps for all schemes. In particular, (1) for MDMA, each transmitter uses 1 molecule and the symbol interval is 875 ms, (2) for MDMA+CDMA, each transmitter uses 1 molecule and the CDMA code length is 7 with a chip interval of 125 ms, (3) for MoMA, each transmitter uses 2 molecules and code length is 14 with a chip interval of 125 ms. Besides, we also ensure that the preambles introduce the same overhead in all schemes. To do so, we ensure that the length of the preamble is equal to 16 times the length of the data symbol and each packet encodes 100 data bits. One thing to note is that the reported throughput is NOT the capacity of the network, since the number of transmitters in the network does not reach the maximum size of the codebook.

Figure. 5.6 shows the overall and per transmitter throughput of the network as the number of actively transmitting and colliding transmitters varies from 1 to 4. When there are no more than two active transmitters, MDMA provides the highest throughput of 0.99 bps per transmitter. However, it cannot support more than 2 transmitters since the system is restricted to 2 different molecules. Even though MDMA+CDMA can support up to 4 transmitters, the per transmitter throughput drastically decreases when two transmitters share the same molecule. This is because MDMA+CDMA fails to detect colliding packets carried by the same molecules, which causes packet loss. On the contrary, the throughput of MoMA is able to scale with the number of transmitters with much less loss, and it still achieves 0.89 bps per transmission with 4 colliding transmitters, which is 1.7× the throughput of MDMA+CDMA.

5.6.2 Micro benchmarks

Impact of code length on decoding BER

Figure. 5.7 shows the average BER when the code length changes but the data rate is fixed. The BER increases along with the length of the code since longer code results in more inter-symbol interference. Taking

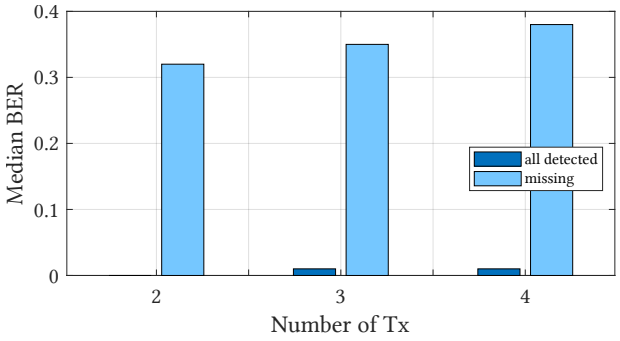


Figure 5.9: BER comparison with and without miss-detected packets.

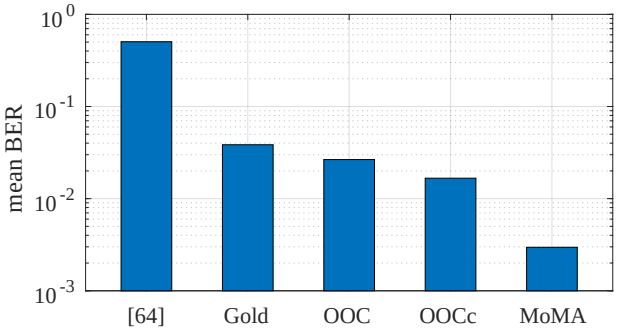


Figure 5.10: BER comparison against baseline coding schemes.

this into consideration, MoMA uses the shortest code possible when the codebook is large enough to support the required number of transmitters.

Impact of preamble length on throughput

Figure. 5.8 shows the total network throughput achieved by MoMA with different preamble lengths. All four TXs transmit and collide using only one molecule at a data rate of 1/1.75 bps. The network throughput first increases along with the preamble length until the preamble reaches 16× the symbol length. This is because the additional preamble sequences improve packet detection and channel estimation accuracy. However, these benefits saturate at 16× symbol length and the more overhead starts to harm the throughput.

Importance of detecting colliding packets

In Figure. 5.9, we compare the BER when all colliding packets are correctly detected and when some of them are missing, which uses the same experiments as MoMA with 2/3/4 transmitters in Figure. 5.6. The median BER only considers the transmissions that are still correctly detected. It is obvious that incorrect detection of any colliding packets results in a disastrous BER in the decoding of the other detected packets. As a consequence, almost all packets are dropped because the BERs are above 0.3. Based on this result, we prioritize packet detection when designing MoMA.

Comparison of coding schemes

We compare five decoding schemes of two categories: (1) individual threshold decoder on directly correlating the receiver signal with the code [100] and (2) MoMA's joint decoder. For MoMA's joint decoder, there are four possible combinations of two coding choices, i.e., OOC and MoMA's code, and two representations of bit -1, that are transmitting nothing and complementary code. The test case had an increasing number of colliding packets, each with 100 data bits, with code length 14 and chip interval 125 ms. In this comparison, we adopt a set of (14,4,2)-OOC codes as specified in [141], which also has code length 14. However, each code in this set has four +1s and the maximum cross-correlation is 2. To isolate the impacts of different coding schemes from all the other aspects, e.g., packet detection and channel estimation, we assume that the receiver knows the exact packet arrival time and the exact CIR of every packet. To obtain the ground truth CIR, we assume we know all the transmitted bits and use all of them to estimate the CIR as opposed to only using the preambles to estimate CIR. Figure. 5.10 shows the BER averaged over 40 different traces. Based

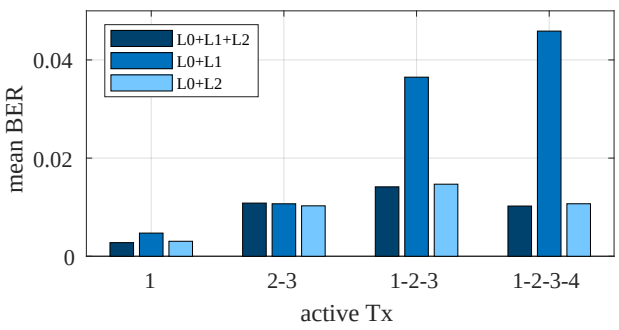


Figure 5.11: BER of using different combined loss terms for channel estimation. Single-molecule results are presented, so similarity loss $\mathcal{L}_3$ does not apply. ToA is assumed known.

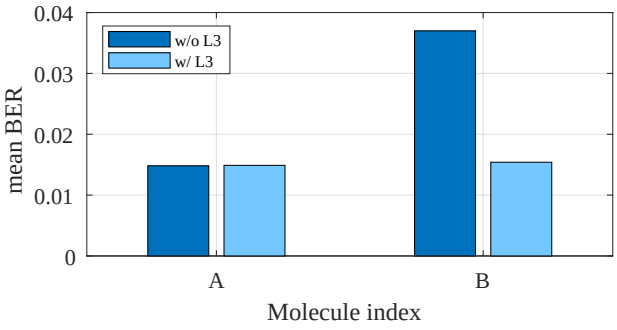


Figure 5.12: BER of decoding two colliding TXs under two-molecule emulations. Moreover, the two TXs share the same code on molecule B, but use different codes on molecule A, so $\mathcal{L}_3$ plays the role. ToA is assumed known.

on the first bar in the figure, the chosen OOC code cannot provide good correlation properties to address ISI and MAI (multiple access interference) issues. Using a longer OOC code set might handle the issue, but it cannot achieve a comparable data rate as MoMA. From the other bars in the figure, we can see that MoMA achieves the best possible decoding among all four coding schemes. Besides, using complementary code for bit -1 can also improve the performance of OOC.

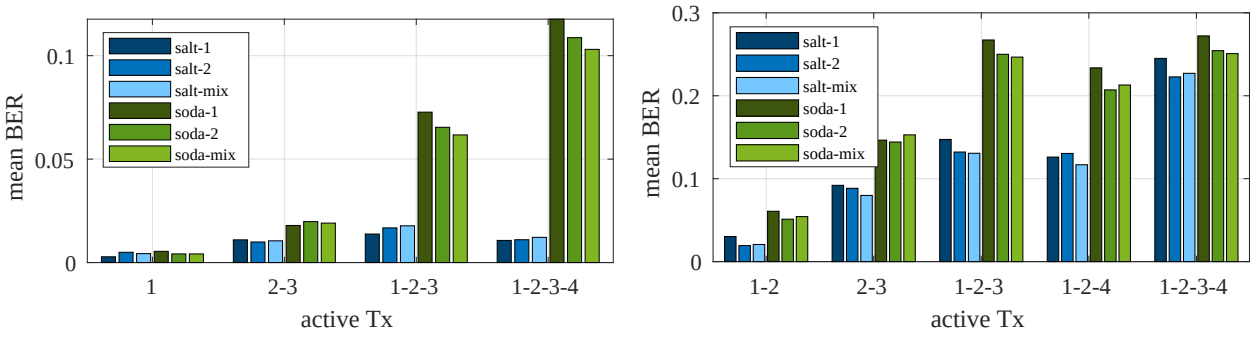
Benefits of empirical loss in channel estimation

We also validate the benefits introduced by the empirical loss terms in the adaptive filtering channel estimation algorithm in MoMA. The test case had an increasing number of colliding packets, each with 100 data bits, with code length 14 and chip interval 125 ms. First, we evaluate the improvement in channel estimation by assuming the ground truth time of arrival (ToA) and comparing the BER of different channel estimation losses. Figure. 5.11 shows the BER averaged over 40 different traces for different number of colliding TXs. These results are for one molecule, so $\mathcal{L}_3$ is not applicable and will be evaluated later. We compare the performance of the full loss with the cases of missing either $\mathcal{L}_1$ or $\mathcal{L}_2$. In Figure. 5.11, it can be observed that $\mathcal{L}_2$ greatly improves the mean BER as the second bar is much higher than the other two. We also note that although L1 offers some improvement (first and last bars), its contributions are not significant.

Benefits of multiple molecules in channel estimation

In Section. 5.3, we explain the advantages of multiple molecules. On the one hand, more molecules provide more samples to conduct better packet detection. On the other hand, more molecules make it possible for the decoder to retrieve the common information about the channel, which improves channel estimation accuracy and thus reduces decoding error.

In Figure. 5.13a (similar to Figure. 5.11), we extend our results to two molecules. Figure. 5.13a shows the improvement in channel estimation with two molecules due to $\mathcal{L}_3$. The "salt-1" bar denotes the original one molecule data with $NaCl$, and the "salt-2" bar denotes the two-molecule emulation with both molecules as $NaCl$. Comparing these two bars reveals no improvement for $NaCl$ for decoding collisions. This is because estimating $NaCl$ CIR is good enough with one molecule under the current preamble length. Thus, we also conducted the same experiments with baking soda $NaHCO_3$. Although we used different concentration of TX solution to achieve roughly equal number of molecules per volume (20 grams per liter for $NaCl$ and 40 grams



(a) Line channel

(b) Fork channel

Figure 5.13: BER of single-molecule experiments versus double-molecule emulations. ToA is assumed known.

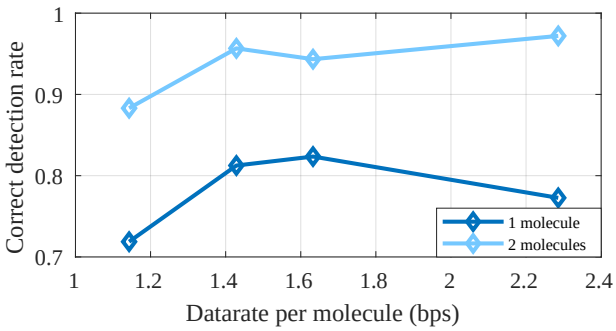


Figure 5.14: Benefits of multiple molecules in packet detection. Percentage of detecting all 4 colliding TXs correctly at different datarate.

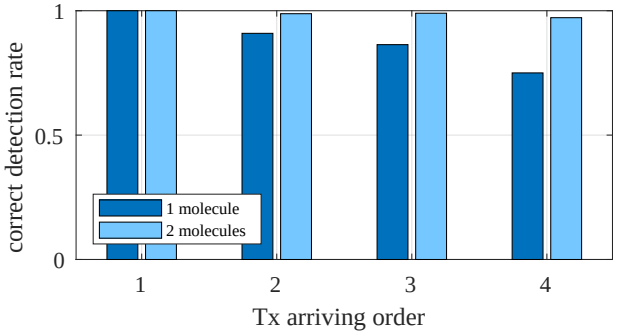


Figure 5.15: Benefits of multiple molecules in packet detection. Percentage of detecting each of the 4 colliding TXs correctly at a datarate of 2.29 bps per molecule.

per liter for $NaHCO_3$), we can still see that $NaHCO_3$ has worse performance by comparing "salt-1" with "soda-1". So, emulating two molecules with $NaHCO_3$ reveals higher improvement from "soda-1" to "soda-2". More intriguing results are revealed by the "salt-mix" and "soda-mix" bars, which represents the emulation of combining one $NaCl$ experiment with one $NaHCO_3$ experiment. The performance of each molecule is presented separately assuming that these are different molecules. By comparing "soda-1" with "soda-mix," we can clearly see the improvement in channel estimation for the worse molecule even under the case of combing two different molecules, while the better one is not much influenced ("salt-1" with "salt-mix"). This validates the idea of using the commonness between CIRs to improve both of them based on their similarity. By comparing "soda-2" with "soda-mix," we can see that such improvement could be better replicating each one of the molecules and more improvements are foreseen with a smarter channel estimation algorithm.

Figure. 5.13b presents the similar experiments under the fork channel, which is more complicated than the line one. A preliminary conclusion can be achieved from Equation. 2.9 that slower background flow is equivalent to longer propagation distance. Assuming the flow splits equally into the forked tubes, which roughly doubles their length if under the same flow, the TX2 and TX3 in the fork channel are at an equivalent distance of 60cm and 120cm in the line channel. However, the BER for these two TXs is much higher. Although the benefits of multiple molecules can still be observed under such cases, the fork topology actually introduces more factors to the molecular channel, which will be considered in future.

A more extreme case is presented in Figure. 5.12. In this experiment, two colliding TXs used different

codes on molecule A but the same code on molecule B. Further, the packets are intentionally made to collide in the preamble, which is the worst-case scenario for channel estimation. Assuming ground truth ToA, we compare the mean BER when applying two different channel estimation losses, i.e. with and without the similarity loss $\mathcal{L}_3$. From Figure. 5.12, we can see that $\mathcal{L}_3$ merely affects decoding in molecule A since the two transmitters are distinguishable in coding, but the improvement is obvious when it comes to molecule B, where BER is reduced by more than half and becomes comparable to molecule A. This again validates the benefits of connecting signals from multiple molecules with $\mathcal{L}_3$.

Benefits of multiple molecules in packet detection

In the end of this part, the general improvement in packet detection is evaluated. Figure. 5.14 shows percentage of detecting all 4 colliding TXs correctly at different datarate. One can see that when using two information molecules, the correct detection rates are 10% higher than those with only one molecule and such improvement is consistent at all data rates. To better understand the improvements, we further examine whether each of the 4 packets is detected correctly, which is shown in Figure. 5.15. First, it can be seen that the later packets are more likely to be miss-detected because the detection happens concurrently with the decoding of all previous packets. The increased signal-dependent noise would lead to more decoding errors, making it harder to detect the newly arrived packets. Second, we can also see huge improvement in using multiple molecules in packet detection, especially for the last arriving packet.

5.7 Related Work

Enabling multiple access in molecular networks has been studied by past work. Except for a few [99], [100], most work has been theoretical or simulation-based:

- **TDMA:** [101] proposes a TDMA (Time Division Multiple Access) scheme. However, TDMA requires synchronization which is difficult to implement in practice due to (a) the long propagation delay of molecular signals and (b) the added complexity of requiring a receiver on the micro-implant.
- **MDMA:** [102], [103] introduce MDMA (Molecular Division Multiple Access) where different molecules are used by different transmitters to avoid interference altogether. This scheme is similar to FDMA (Frequency Division Multiple Access) in wireless networks which avoids synchronization and interference. However, it is difficult to scale this approach to many transmitters as most practical bio-molecular systems are limited to 2-3 molecules [10], [104], [105].
- **ADMA:** [99] presents ADMA (Amplitude Division Multiple Access) that assigns the amplitude of the transmitted signal as the address of the source. This method reduces the error in resolving the source addresses under collisions, but the required signal power scales exponentially with the number of transmitters and the performance is sensitive to changes in the channel response.
- **CDMA:** [100], [106] propose a CDMA (Code Division Multiple Access) scheme that adopts Optical Orthogonal Codes (OOC) from fiber optic networks. [100] further tests the OOC codes in a gaseous molecular testbed with two transmitters. Despite the fact that, similar to molecular networks, the signal in fiber optic networks is also non-negative since it represents the intensity of light [107], OOC is not suitable for molecular networks for the following reasons: (a) Unlike molecular networks, fiber optic networks can deal with high ISI in hardware by compensating for the dispersion of light pulses.

Techniques such as DCF (Dispersion Compensation Fiber) and FBG (Fiber Bragg Grating) [108] can change the physical properties of the channel to reduce ISI. However, in the presence of high ISI, one cannot simply correlate with the OOC code of each transmitter and decode them independently as proposed in [100], [106]; [2] (b) Commercial fiber optic systems that use OOC typically use long OOC codes which can significantly reduce the data rate [109]. Unlike optical networks that operate at link budgets of 100 Gbps [110], molecular networks operate in the few bits per second regime and cannot afford a significant reduction in data rate; (c) We show in our results in Section. 5.6 that using OOC codes of short length (high data rate) leads to high Bit Error Rates (BER). Moreover, [100] itself reports poor performance where the error rate is low only at data rates below 0.051 bits per second. MoMA, on the other hand, can achieve low error rates at data rates of 1 bit per second per transmitter. [111] also proposes a CDMA scheme that uses two molecules but is not based on OOC. However, the two molecules are used for modulation to encode different bits, i.e., releasing molecule A represents a "1" bit while releasing molecule B represents a "0". Since the two molecules cannot be released at the same time, the network capacity is cut by half. In contrast, MoMA uses the two molecules to send two data streams while at the same time improving packet detection and channel estimation.

- **Broadcast (No Multiple Access Scheme):** [112] computes the theoretical BER of a multi-transmitter network where the transmitters broadcast their data packets that collide at the receiver. However, in this work, each transmitter is decoded independently resulting in a drastic increase in BER as the number of transmitters increases. [113] also computes the theoretical capacity of a multiple-TX single-RX network when the network uses one molecule, and transmitters broadcast their packets. However, the work does not present any approach for decoding the colliding packets.

Finally, designing practical transceivers that can process multiple molecules is a requirement for MoMA. While this is outside the scope of this paper, it is important to note that there exist studies on both natural and artificial molecular transmitters that can leverage multiple information particles [142]. Researchers have created cells that can respond to and integrate multiple molecules. For example, [39] constructs an Escherichia coli consortium-based biosensor that can detect and integrate three environmental signals (arsenic, mercury, and copper ion levels). [41] presents a robust, general, and scalable system (named BLADE) to engineer genetic circuits with multiple inputs and outputs in mammalian cells, enabling sophisticated cellular computation.

5.8 Discussions

In this section, we shall explore the constraints of the study and the possible avenues for future research.

5.8.1 Alternative molecular testbed

In this chapter, we have used $NaCl$ to emulate the propagation of molecules in a liquid medium. In addition to $NaCl$, researchers have used a variety of particles to emulate molecular communication. This includes hydrogen ions with pH probe [98], super-paramagnetic iron oxide nano-particles with susceptometer [75], glucose with electrolyte-gated FET [143] and color pigments with color sensor [115]. A detailed survey on existing molecular testbeds can be found at [144]. It must be noted that our current testbed can be seamlessly

[2] Despite their name, OOC codes are not perfectly orthogonal but have good correlation properties when the code length is large.

combined with other existing particles, as well as biological systems to potentially implement a real-world molecular network system.

Although this chapter highlights the challenges to address the non-negative signal, some of the above-mentioned testbeds may have negative measurements. For example, in a pH-based testbed [98], acid (pH < 7) can be treated as negative signal while base (pH > 7) as positive. However, pH signal introduces its own challenge for CDMA. The variation in pH is in the log relation, instead of linear, which affects the cancellation of CDMA code because CDMA code assumes a linear relation between the measured signal and the transmitted code. But taking a step back, all these molecular signals reveal the variation in the concentration of information particles. Studying the relation between these two, such various signals can all be translated into the general concentration signal, so the MoMA decoder is generally applicable.

5.8.2 Multiple-molecule testbed

One major limitation of this chapter is that the testbed does not support concurrent transmission of multiple molecules. Due to the fact that our first measurement is electric conductivity indicating the concentration of $NaCl$, it is hard to introduce another molecule to the system such that it does not affect the EC value of the solution and at the same time can be measured with other methods. A direct improvement is to combine two existing testbeds (like the ones reported in [144]) with independent molecules and measurements. But targeting a more general solution, if we loosen the constraints on independence, there will be a much wider choice. For example, if we use some acid as the second information molecule and pH value as the second measurement, the two molecules can be distinguished easily since $NaCl$ only changes the EC value while the acid changes both. Such testbed requires the knowledge in the influence of the acid concentration on the EC and pH value respectively, so it can recover the concentration changes of each molecule from the received signal. Moreover, these two measurements could even support more than two molecules. The preliminary idea is to understand how one molecule changes the two measurements at a difference ratio. HCl dissolves in water and becomes H^+ and Cl^-, so EC and pH should change at a ratio of 1:1. Similarly, $NaCl$ is at a ratio of 1:0 and $NaOH$ of 1:-1. With such relation, the decoder is able to separate the signals of each molecule using a modified Viterbi algorithm that jointly decodes signals from all molecules.

5.8.3 Current and future MC research

Majority of research works on molecular communication focuses on studying and modeling the molecular channel, transceiver designs of various particles, and sensing mechanisms using a theoretical framework. This was followed by a shift towards experiments to evaluate molecular communication spanning a range of few millimeters all the way to a few meters. Our work falls in the longer range. A survey of the state of the art can be found in [144], [145]. The community seems to diverge in two directions. For the short range, experiments are typically involved with micro-scale transmitters that require a strong basis in chemistry or biology. The long-range experimental testbeds can emulate biological systems. Although such a testbed has similarities to wireless networks, due to slow propagation and particle diffusion, molecular signals endure high ISI that prohibits comparable datarate as existing wireless techniques in the same range.

In order to leverage the advancements in wireless and wired communication techniques in molecular systems, interdisciplinary collaboration is necessary. In this chapter, we have taken a step towards this collaboration by asking the question "what CAN a biological transmitter do" instead of "what SHOULD a transmitter do". Unlike existing works, our design incorporates the hardware constraints of a bio-transmitter, a

bottleneck posed by bioengineering. We then offload processing to the receiver, which is more computationally resourceful than the transmitters. With this offloading, MoMA eliminates the need for synchronization between the transmitter and the receiver. This is our attempt to bridge the gap between communication and biochemistry. Yet the community is not limited to these two fields. Based on our limited medical knowledge, we envision that such a system can be used to monitor the spread of inflammation or the metastasis of cancer cells. Nonetheless, the value as well as the practical implementation of such an in-body system should be evaluated from the medical perspective, which invites more interdisciplinary researchers to participate and work together towards an innovative community.

5.9 Conclusion

In this chapter, we introduce an innovative multiple access protocol, MoMA, for the molecular network with multiple transmitters and a single receiver. We highlight the unique properties of the molecular channel including non-negative signal, lack of synchronization and high ISI and explain the encoding and decoding schemes of data packets as well as the packet detection and channel estimation processes under our protocol. Our results demonstrate that using multiple transmitters along with multiple molecules with MoMA brings us closer to enabling more practical and scalable molecular networks.

Chapter 6

FlowLink

As we said, blood flow is the prominent feature in vessels, which is one of the factors that affect the propagation of molecular signal (see Chapter. 2). There has been various work studying the signal propagation under the influence of flow, including channel modeling, theoretical information capacity and decoding algorithms (see Chapter. 3). Moreover, some recent works also try to set up testbeds to emulate the vessel environment in the sense of channel diameter and average flow speed (see Section. 6.6.1). Although these works have made concrete contributions to the development of the blood network, one major feature of the vessel channel has been ignored–the blood flow speed is not a constant. Chapter. 4 has observed and addressed the potential gradual channel variation during packet transmission, but in this chapter, we emphasize that **the rapid channel variation due to heart beats can greatly degrade the performance of conventional decoders**.

The heart iterates between two stages–diastole, when the chambers relax and refill with blood, and systole, when the chambers contract and push blood into the artery. Figure. 6.1 shows the medical ultrasonic spectrum image that measures the blood speed variation. It can be seen that systole is accompanied by a sudden increase in the blood flow speed, which will gradually reduce during the diastole stage. Considering that (1) the normal heart rate of an adult is between 60 and 100 beats per minute ($\approx$ 1 to 1.6 beats per second) and (2) data rates of a few bits per second have been demonstrated in a vessel-like system (Chapter. 4, as well as [98]), with the blood flow cycle roughly as long as the MC symbol interval if not longer. Hence, the MC channel could change within symbols and the assumption for conventional wireless systems, that the channel does not change during the transmission of one packet, no longer holds. Even though the fast-varying channel problem has been studied in electromagnetic and acoustic communication systems, the propagation speed of the signal is typically constant, and the variation is usually limited to phase change of Channel Impulse Response (CIR) taps, which can be tracked with pilots.

MC networks, however, are faced with a more challenging varying channel where the propagation speed of the signal changes between symbols. This causes the boundaries of symbol periods to differ among consecutive symbols. Fig. 6.2 shows an example of the number of received particles when the speed of flow is constant and when it changes due to the heartbeat. As can be seen in the top figure, when the speed is constant, the particles arrive at equal intervals and channel impulse response is uniform. However, with heartbeat, the propagation time of particles can change from symbol to next and the channel impulse response changes. Hence, the boundaries where each symbol starts and ends change depending on the speed of the flow making it difficult to correctly decode the signal. The example shown in Fig. 6.2 is simulated for a relatively low data

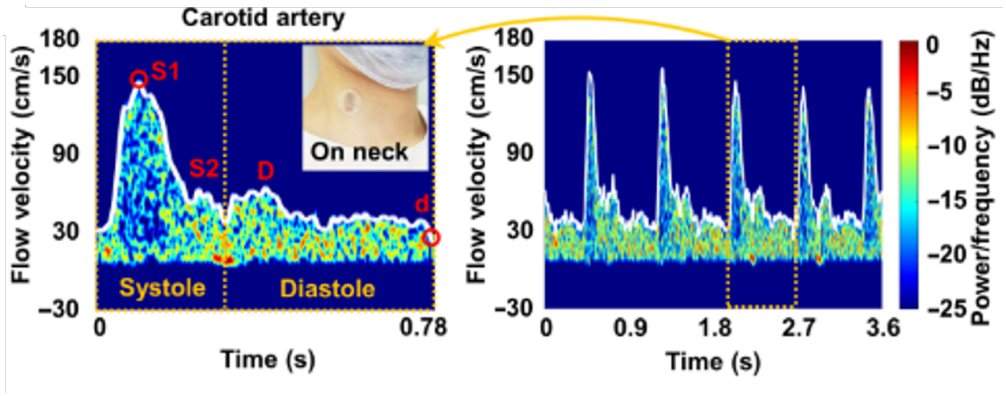


Figure 6.1: Blood flow velocity measurement in carotid artery [146].

rate of 1.4 bps i.e. symbol interval of 0.7s and small inter-symbol-interference (ISI) in order to demonstrate how the symbol boundaries shift. In practice, as we increase the data rate and with realistic ISI, it becomes increasingly harder to determine the symbol boundaries as particles from different symbols stack up.

In this chapter, we propose FlowLink, a heartbeat-aware MC protocol that compensates for the influence of heartbeats to correctly decode molecular signals. The core idea of FlowLink is to preprocess the received signal to extend the period during high flow speed and shrink it during low flow speed, such that all symbols can be equally sensed by the receiver and share the same CIR. It does so by projecting the received signal on the flow volume axis (i.e. integral of flow speed) which rescales the signal and unifies the CIR of each symbol as we prove in Section. 6.3.1.

While this rescaling allows all symbols to experience the same CIR in the flow volume axis, the symbols continue to be non-uniformly separated. The spacing between symbols is determined by the varying flow speed, not only when the receiver senses the symbol, but also when the transmitter sends the symbol. So, the dynamic symbol boundaries remain on the flow volume scale and prevent decoding. To address the problem, FlowLink begins with packet detection, including a coarse stage using energy detection and a finer stage using signal reconstruction. FlowLink scans over the potential time of transmission reported by the coarse stage–for each sample it infers the corresponding symbol boundaries with real-time flow rate, estimates the CIR with known preamble symbols, reconstructs the signal through convolution and computes its divergence from the received signal–until it finds the minimum error between two signals.

We built a testbed to evaluate the performance of FlowLink, which is modified from the artifact of [147] using a peristaltic pump with MODBUS speed control. The dataset contains various scenarios with varying TX-RX distance, flow cycle and channel diameter. By comparing FlowLink with the decoder [147] under the scenario of one TX and one RX, we show that FlowLink can distinguish the accurate symbol period and reduce the decoding BER by an average of 40%.

The chapter makes the following contributions. To the best of our knowledge, this is the first MC decoder that accounts for the heartbeat and flow speed variation and does not require any modeling of a channel with specific parameters. We also set up a testbed with varying flow speed to evaluate the performance of FlowLink, which is compared with a conventional decoder under different flow conditions.

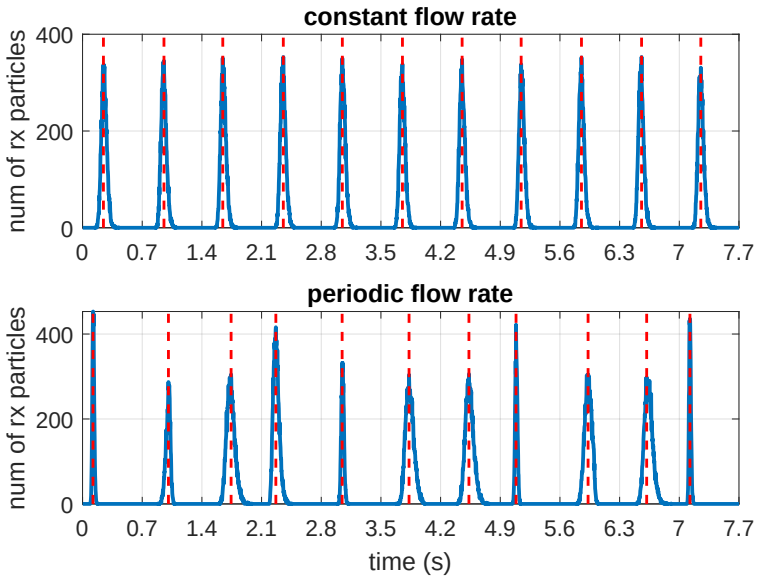


Figure 6.2: The number of particles detected at the receiver end under constant flow (above) and heartbeat-induced flow (below). In both cases, the transmitter sends a train of pulses with a constant interval of 0.7 seconds. It can be seen that under constant flow, the received signal reveals a consistent pattern with a period of 0.7 seconds, which is identical to the transmission rate. However, under the heartbeat-induced flow with a period of 1 second, the signal of each pulse no longer maintains constant spacing. Moreover, the response of each transmitted pulse also diverges from each other.

6.1 Background

In this chapter, we adopt the model that the blood network is composed of multiple gateways (Figure. 1.1). Here is a brief review of the assumptions, and details can be found in Section. 2.4.

The whole body is divided into multiple areas, and each is serviced by a sub molecular network and managed by a node. The blood network is the backbone connecting these separate areas, which also bridges the inside network with the outside through the gateways.

The gateways are immobile. Although uni-directional communication is the common case between one node and one gateway due to the blood flow, bi-directional communication for each node can be implemented by sending to the downstream gateway and receiving from the upstream. We assume that the transmitter adopts On-Off Keying (OOK) and releases a pulse ($\delta(t)$) of particles as a bit "1". The receiver passively measures the concentration of particles around it, which means it does not destroy the information particles.

The propagation of information particles is advection-dominant, while diffusion has negligible influence. However, this does not mean the Channel Impulse Response (CIR) is a single tap due to the following two major factors. On the one hand, the blood vessel topology is complicated with diameter variation and forking. On the other hand, the blood flow in the vessels is assumed to be mostly laminar (see Section. 2.1.2), where the flow speed is higher as it is further away from the wall of the vessel.

6.2 Blood Flow Informed Challenges

Fast time-varying channel is a well-studied topic in conventional wireless communication systems [120] because of the growing user scenarios, especially with the increased user mobility, data rates and carrier frequencies.

The movement of transmitter, receiver and reflectors features a significant Doppler effect, and there exist various models and solutions [120]. However, considering the transmission rate of multiple Mbps (i.e. short symbol interval), the object movement is mostly shown in the phase variation of each CIR tap, instead of change in tap delay [148]. This is usually true during the transmission of a standard-size packet.

However, for molecular signals, the propagation delay is multiple-order-of-magnitude longer than conventional wireless signals even with the aid of the environmental flow. Such difference determines that the molecular signal propagation delay in the blood network is much longer. Since diffusion of molecular signals will induce large Inter-Symbol Interference (ISI), which is proportional to the propagation delay, the data rate is limited and is generally not competitive with the convention wireless techniques. The long symbol interval provides more opportunity to observe the variation in the environmental flow speed between symbols.

To simplify the illustration, we assume the channel is 1-D and temporarily ignore the influence of diffusion. Moreover, With the general assumption that liquid is incompressible [1], the variation of speed could instantaneously reflect in every part of the channel. Equation. 6.1 depicts the relation between the transceiver distance Δx and the propagation delay Δt, where $v(t)$ is the flow speed and t_0 is the transmission time of the signal.

$$\Delta x = \int_{t_0}^{t_0+\Delta t} v(t)dt \tag{6.1}$$

The direct consequence of the varying flow speed is that the symbols in the same packet now bear different propagation delays, so the receiver can no longer use the predetermined symbol interval to mark the start of each symbol. Figure. 6.3 provides a simple example illustrating such influence. The channel is a narrow long tube (1-D), in which runs a periodic flow iterating between 80 cm/s for 0.2 s–mimicking the systole of heart–and 30 cm/s for 0.8 s–mimicking the diastole of heart (Figure. 6.3 top). The TX at $x = 0$ sends a pulse train with an interval of 0.7 s (Figure. 6.3 top second). However, none of the RXs at $x = 10$ cm and $x = 20$ cm (Figure. 6.3 blue plots) receives the same pulse train as transmitted. Moreover, the two RXs do not see the same pattern either, where the complexity lies in the various combinations of different flow profiles, as well as the symbol transmission time.

Conventional decoders are not ideal if the arrival time of each symbol is not predictable, especially when the high environmental flow speed could lead to some of the symbols arriving with little interval and compacting within one symbol period. Compensating for such flow variation is a unique challenge and is not the priority in conventional wireless systems.

6.3 Technical

The propagation delay of each symbol is not a constant. In other words, this means that the Channel Impulse Response (CIR) varies across the packet. A simple idea is that decoding is possible if we somehow know the CIR of each symbol. Then the solution is to compute the CIR based on some template CIR, which shrinks during high speed in the time scale and expands during low speed. This is not an incorrect direction, but is followed by a series of complicated and entwined problems: "what is this template CIR?", "what is the arrival time of each symbol?" and "how do we solve these two problems concurrently when the preamble is not in the regular order?". Instead, we found and focused on the non-varying factor among the altering CIRs

[1] Molecular communication is a wide area which also covers gas as medium, where compressing is much easier and may not be a common assumption. But the liquid testbed in this chapter has nothing typical in pressure, so we assume that the liquid is incompressible.

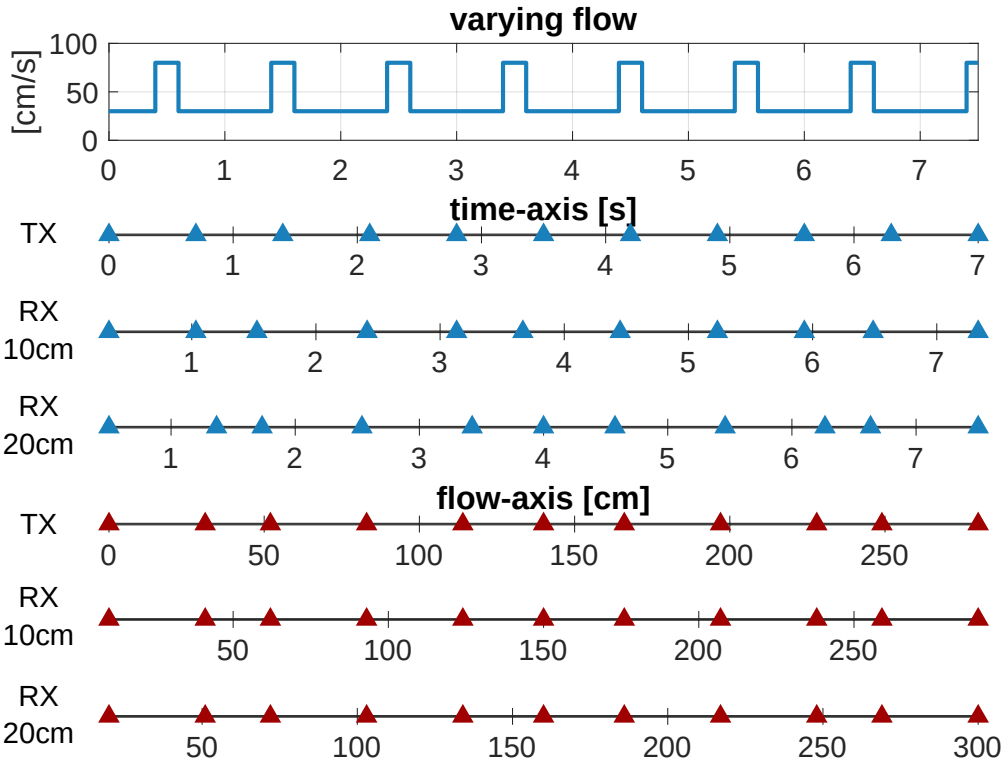


Figure 6.3: The arrival time of consecutive symbols of the same MC packet is messed up by the varying flow speed. From top to bottom: (1) the flow speed [cm/s] as a function of time [s]; (2) the pulse train of 0.7s interval sent by TX and received by two RX at 10cm and 20cm, where the blue plots use time-axis and the red plots use flow-axis. To compare the interval between pulses, the axes have the same scale but different range.

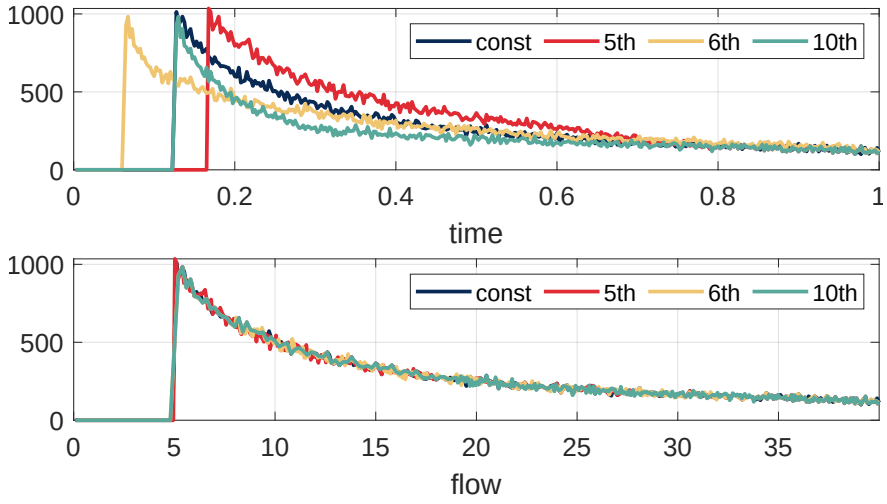


Figure 6.4: Monte-Carlo simulated CIR on 1-D channel model. The CIRs for different symbols with time x-axis (upper) are unified with flow amount x-axis (lower).

and developed the intuitive decoder FlowLink. In this section, we will sequentially explain the two major components of FlowLink which composes the solution to the varying flow.

6.3.1 Picking a new x-axis

As was discussed in Section. 6.2, in conventional wireless systems, the channel impulse response (CIR) is considered unvarying within the channel coherence time, i.e. the **propagation delay** of each path between transceivers is a constant. Thus, sampling by a fixed **time interval** can maintain the pattern of the original transmitted sequence. From this perspective, what is the constant value in the molecular channel that can inspire the corresponding sampling parameter? The blood network system model described in Section. 6.1 features a simple answer. Under the assumption of the immobile transceivers and the advection-dominant channel, the **total volume of blood flow** by molecular signal to propagate from the transmitter to the receiver is constant. Consequently, instead of treating the receiver signal as a function of time $y(t)$, FlowLink first converts it to a function of flow $y(V)$.

To understand the effectiveness of such x-axis replacement, let's start with an example assuming an advection-only 1-D channel model where the CIR is a single pulse. It is easier to understand in 1-D, because flow speed $Q(t)$ [volume per time (m^3/s)] can be reduced to $v(t)$ [distance per time (m/s)], which is directly related to the molecular propagation model described in Section. 6.1. Let's go back to Figure. 6.3 and focus on the red plots, which are the identical RX signal but using flow amount as the x-axis. It can be seen that the receivers can capture the same pulse train as it was sent by the transmitter on the flow-axis, regardless of the propagation distance.

But does the flow-axis solve the problem when the CIR is much more complicated like in a real blood channel? Figure. 6.4 provides a simple Monte-Carlo simulation for the receiver at x = 10 cm as in Figure. 6.3, but in a 3-D tube-like channel with laminar flow. Figure. 6.4(upper) first plots the CIR when the flow rate is a constant that is equal to the average speed of the varying flow (in blue), while the other three curves depict the CIR of the 5th, 6th and 10th pulse respectively. As can be seen, the varying flow not only affects the propagation delay and delay of each symbol, but also the width of the CIR, which is usually accompanied by an abrupt shape change between systole and diastole. On the contrary, in Figure. 6.4(lower), the CIR of

different pulses perfectly overlap with each other.

How do we get the receiver signal $y(V)$? From the algorithm's perspective, the easiest solution is to equip the device with a flow meter module, and this is also the solution we use to measure the ground truth flow in the testbed. It is reasonable to assume that the blood flow is uni-directional in most cases, so $Q(t) > 0$ and $V(t) = \int_0^t Q(t)$ is monotonically increasing and there exists the inverse function $V^{-1}(t)$. Thus, we can directly use the flow meter to preprocess the signal and achieve $y(V) = y(V^{-1}(t))$. This is also the solution we use to measure the ground truth in the testbed. To compensate for the non-uniform distribution of time samples on the flow-axis, FlowLink first adopts a higher sampling rate than the symbol rate and then computes the uniform flow samples through interpolation.

Here are a few remarks on using the flow-axis. First, it should be noticed that there is no one-to-one mapping between the time samples and the flow samples, because the blood flow is not a variable with discrete values. Thus, FlowLink makes a simple assumption that the CIR can be linearly interpolated under a high flow sampling rate, which enables the Viterbi algorithm with a few modifications. It maps one time pulse into no more than two flow pulses. Typically, a time pulse $\delta(t)$ will be approximated by

- If $j = \frac{V(t)}{\Delta V}$ is an integer, where ΔV is the flow sampling interval

$$\delta(t) \longmapsto \delta[j] \tag{6.2}$$

- Otherwise, $j = \left\lfloor \frac{V(t)}{\Delta V} \right\rfloor$

$$\delta(t) \longmapsto \left(j + 1 - \frac{V(t)}{\Delta V}\right) \delta[j] + \left(\frac{V(t)}{\Delta V} - j\right) \delta[j+1] \tag{6.3}$$

Second, although the above illustration assumes that the channel is a single tube of constant diameter, the FlowLink method can be generalized to other complicated channel topologies, with either varying cross sections or forking. In spite of the elasticity of the blood vessels, the topology is comparably stable when the person is static. Thus, it is reasonable to assume that the distribution of the blood flow in this network is also stable during the transmission of a packet, such that a large amount of information particles follows the same statistical distribution during propagation, which leads to an unvarying CIR. Besides, it should be noted that the receiver may sense a different amount of blood flow from the transmitter due to the forking in the blood vessels. But, since the stable topology indicates that the blood flow at different devices should maintain a constant ratio, the receiver only needs to resample the signal based on its own flow measurement $V_R(t)$, without any knowledge about the one at the transmitter end $V_T(t)$. The related benchmarks will be presented in Section. 6.5.2.

Third, in this illustration of a 1-D scenario, the flow rate has the unit of "m/s". But in a practical environment, the flow rate can either be described as the volume of fluid crossing through a certain area ("m^3/s"), or the average speed ("m/s"). Both are acceptable for μ-Link depending on the architecture of the flow meter. For laminar flow, the fluid maintains a stable relative behavior across different layers, so the above descriptions can be deducted from one another. More discussions are presented in Section. 6.7.2.

6.3.2 Detecting the symbol boundaries

Even though the CIR of each symbol is unified, another problem prevents the usage of conventional decoders. Although the symbols are sent by the transmitter with a fixed time interval, the propagation delay between

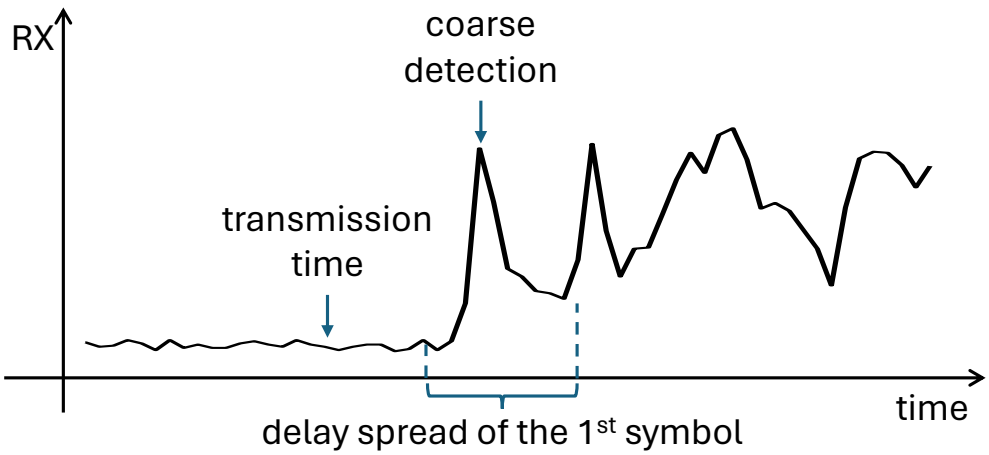


Figure 6.5: The relationship between coarse detection, delay spread and transmission time.

the transmitter and the receiver for each symbol is a variable, which is decided by the varying flow speed and leads to the uneven spacing between symbols at the receiver end (as depicted in Figure. 6.3). Failing to resolve the floating symbol boundaries can lead to improper handling of inter-symbol interference (ISI), which will further deteriorate the receiver performance.

It may seem very complicated at first glance that we need to decide the duration for each symbol, because there are usually hundreds of symbols in one data packet and the wrong detection of one symbol would also influence its neighbors. But with the knowledge of the blood flow $V_R(t)$, all symbol boundaries can be derived from one parameter, i.e. the **transmission time** of the first symbol. Note that this is fundamentally different from the received time of the first symbol in conventional wireless system, which also indicates another parameter–the TX-RX distance.

Returning to the above 1-D channel model with single-pulse CIR, where the TX-RX distance is δx and the flow rate is $v(t)$. Once the transmission time of the packet is known, the transmission time of each symbol can be derived by adding multiple symbol intervals, so the received time can be inferred by Equation. 6.1 accordingly. Besides, the TX-RX distance δx can be achieved by detecting the received time of the first symbol and integrating the flow rate since its transmission. However, in a real molecular network in the blood, the CIR is never a single tap but with a certain delay spread, so the above TX-RX "distance" is implicitly included in the CIR.

FlowLink's packet detection is composed of two stages:

1. **Coarse detection (time-axis).** This is identical to the detection method in conventional wireless systems. Specifically, FlowLink maintains two back-to-back sliding windows of time samples, computes the power ratio of the signals in the two windows and seeks the time sample with the peak ratio. As discussed above, the symbol boundaries are warped due to the varying flow, so this method does not provide an accurate signal received time, not to mention the transmitted time of the first symbol. But it does serve as a good starting point to proceed to the next finer detection step.

2. **Finer detection (flow-axis).** Considering the influence of laminar flow, a non-negligible proportion of particles in a pulse would arrive at the receiver earlier than average since they are far away from the wall of vessels. This indicates that the molecular CIR generally has a non-negligible rising stage

in power. So, the CIR delay spread should span both after and before the coarse detection, and the transmitted time of the first symbol is even earlier (Figure. 6.5). Understanding this, FlowLink searches the transmission time in the samples earlier than the coarse detection, by estimating the flow-axis CIR and finding the one that best fits the original signal. Specifically, for the packet preamble $\vec{p}$ and a tentative transmission time t, FlowLink first maps the time-axis preamble $\vec{p}_t$ to the flow-axis version as was described in Section. 6.3.1.

$$\vec{p}_t \longmapsto \vec{p}_V \tag{6.4}$$

It is then used to construct the Toeplitz matrix $P = T(\vec{p}_V)$. The one that results in the minimum distance from the flow-axis signal $\vec{y}_V$ to the column space of P will be the result of the finer detection.

$$\min_t \left\| \vec{y}_V - P\left(P^T P\right)^{-1} P^T \vec{y}_V \right\| \tag{6.5}$$

6.4 Testbed

We build a testbed to simulate the blood vessel environment (Figure. 6.6), with the following major components:

- **Channel**: Silicone tubes are used to simulate blood vessels. It is also a common material in medical research and physician training due to its similarity to biological tissues. These tubes have the advantages of low cost and high flexibility to be bent and connected to construct any expected topology.
- **Medium**: Water is used in the testbed as the propagation medium, which is shown to share similar properties with human blood [149]. We do not filter the water, so the ions in the water can serve as a source of noise in the channel, similar to the particles that are involved with human body activities but will influence the receiver measurement.
- **Flow**: Blood flow is the highlight feature in the blood network, so we use the Shenchen LabV6-III peristaltic pump [150] to provide a periodic flow. The pump can take commands from an RS485 interface and set the motor at a given rotating speed (Rounds Per Minute, RPM) to provide a specific flow. However, the command rate is limited by the RS485 protocol, which means that we cannot perfectly mimic real blood flow. Instead, the actual flow in the experiments is a two-step function, with a high RPM phase for systole and a low RPM phase for diastole.
- **Information-bearing particle:** Sodium chloride ($NaCl$) serves as the information-bearing particle in this testbed due to two major concerns: first, it is easy and inexpensive to buy $NaCl$; and second, $NaCl$ is friendly to the environment, so there is no need for dedicated process before disposal.
- **Transmitter (TX):** $NaCl$ is stored as the solution of high concentration, which will be pressed into the main channel through a Gikfun pump [151]. The Gikfun pump functions as the transmitter, which is powered on (through a transistor switch circuit) for a very short amount of time to send a bit "1" and stays silent for a bit "0". Its influence over the total amount of flow in the channel is negligible.
- **Receiver (RX):** The concentration of $NaCl$ is measured by the Gravity Analog Electrical Conductivity (EC) reader [152], which is at the downstream of the transmitter. It is connected to the tube with a specifically designed 3D printed cap, so the water can run through between the metal plates of the EC reader without leakage. Due to the spreading in the molecular channel, the high-concentration

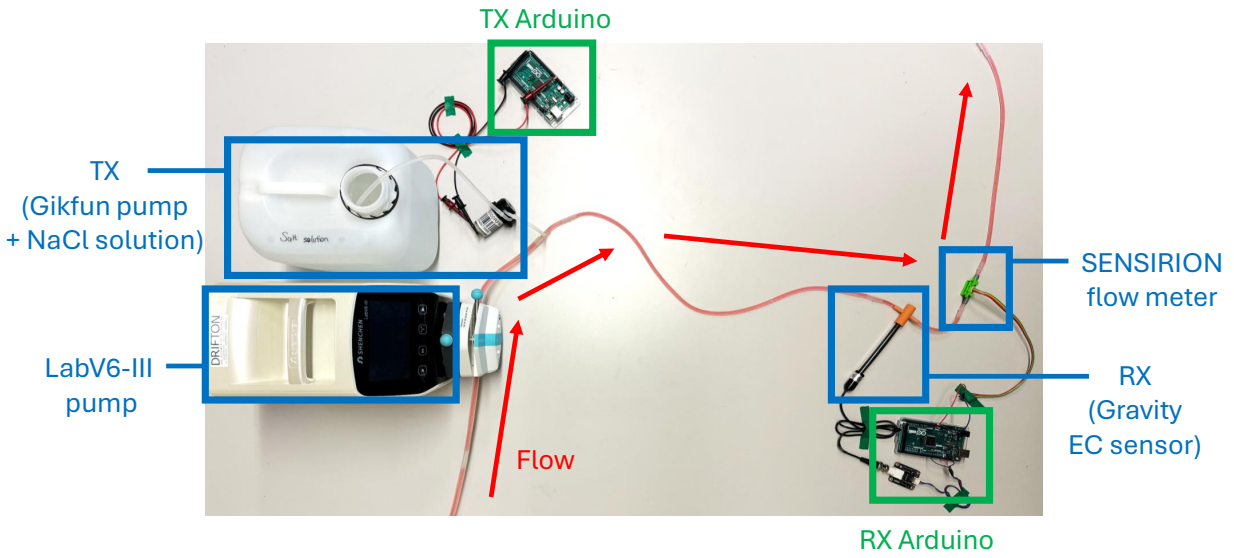


Figure 6.6: FlowLink testbed. It simulates the blood vessel in the silicone tube architecture, with the LabV6-III pump providing a heartbeat-like periodic flow.

pulse from the transmitter has been diluted by the time it arrives at the receiver. Thus, the EC is approximately linear to the concentration of $NaCl$.

- **Flow meter:** Due to the architecture of the peristaltic pump, it cannot offer a strictly constant flow when the motor rotates within one round. Thus, we cannot use the commanded RPM as the flow speed. Instead, SENSIRION SEK-SLF3S-4000B [153] is mounted after the receiver to provide ground truth flow measurement.

- **Control:** The transmitter and the receiver are controlled by two separate Arduino Mega 2560 rev3 [154] with SD modules, while the LabV6-III pump is controlled through MATLAB. All three are synchronized over the serial communication on one laptop.

We have conducted extensive experiments with various parameters, including the inner diameter of the tube, the packet symbol interval, the average rate and the HIGH-to-LOW duration ratio of the flow, the TX-RX distance and the channel topology, which will be presented immediately. Over 40 packets under each condition are collected, while each trace lasts for around two minutes. Packets of shorter symbol intervals have more bits.

6.5 Results

This chapter illustrates the performance of FlowLink with extensive experiments. Reviewing the key ideas of FlowLink, the most fundamental one is the re-sampling operation that transforms the signal from time scale into flow scale. Thus, we modify FlowLink's decoder by removing the resampling process and use it as the baseline, which is named decoder on "time-axis." Correspondingly, FlowLink is named decoder on "flow-axis."

This chapter adopts two metrics to evaluate the performance of FlowLink:

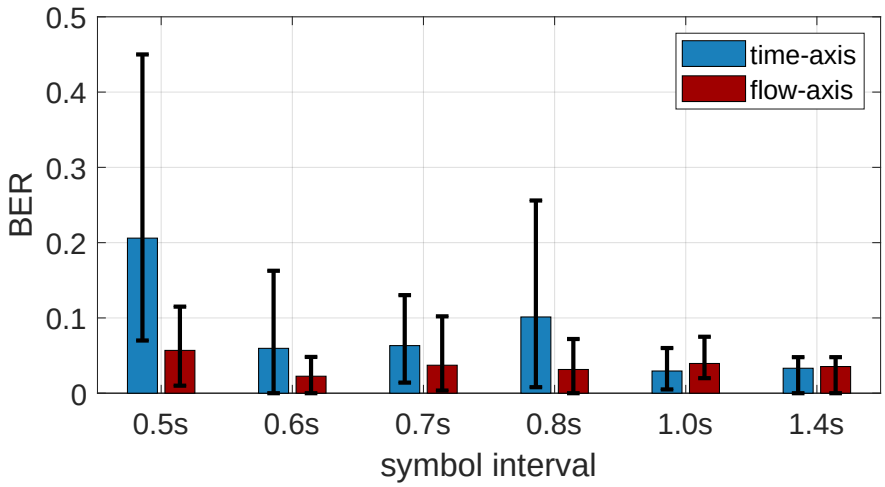


Figure 6.7: Normal. The flow simulates that in carotid artery. BER.

- Bit Error Ratio (BER)–the number of bit errors divided by the total number of transferred bits. This is the common metric used to evaluate a communication system. The bar plot presents the average BER, and the error bar shows the 10-th and 90-th percentile.
- Fitting error improvement. Fitting error is defined as the Mean Squared Error (MSE) between the received signal and the reconstructed signal, while the latter is computed through convolution between ground truth transmitted bits and a CIR estimated from the packet. Because large fitting errors indicate bad performance, the improvement is defined as the subtraction of the flow-axis error from the time-axis error.

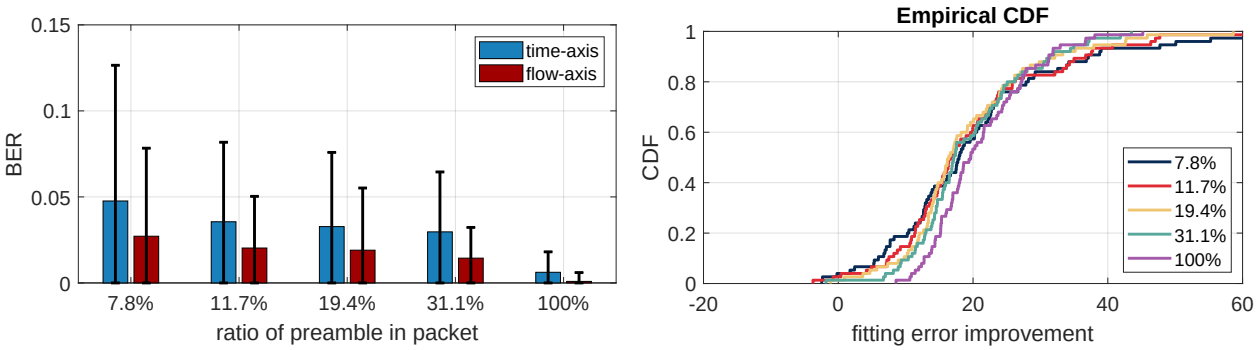
Besides, two different cases are considered:

- Normal, where the receiver performs packet detection, channel estimation and decoding to finally receive the transferred bits. This is used to evaluate the overall performance of FlowLink.
- GTPD, Ground-Truth Packet-detection, where the receiver is assumed to know the exact transmission time of a packet, but needs to perform the rest steps.

6.5.1 Main results

We first present FlowLink's overall performance in the setup simulating carotid artery. The channel is a single 10 cm tube of 4 mm ID, and the flow cycle is 12 mL/s (100 cm/s) for 0.3 seconds followed by 3 mL/s (25 cm/s) for 0.5 seconds. Under such conditions, we have collected data with different packet symbol intervals, i.e. 0.5/0.6/0.7/0.8/1.0/1.4 seconds.

Figure. 6.7 shows the overall decoder performance under the normal decoding procedure. It is clear to see that when the symbol interval is 0.5/0.6/0.7/0.8 seconds, there is a clear advantage of decoding on flow-axis over on time-axis, not only in average BER, but also the 10-th and 90-th percentile. However, such gain is not viewed when the symbol interval grows longer (1.0/1.4 seconds). The reason is that the BER under long symbol interval is too small to validate the advantage of FlowLink, so a larger dataset is needed for further validation.



(a) BER

(b) Fitting error improvement

Figure 6.8: Proportion of preamble used to estimate CIR.

Flow-axis advantage

To directly show the advantage of using flow-axis over time-axis, Figure. 6.8a compares the BER under time-axis and flow-axis when the transmission time is known (GTPD). As we mentioned in Section. 6.3.1, the flow-axis can present a coherent CIR for the whole packet while the time-axis cannot. To validate this claim, we show the decoding performance with regard to the proportion of the preamble to the whole packet. In fact, we process the same data (0.6 seconds symbol interval in Figure. 6.7, with 7.8% preamble), but intentionally treat part of the data symbols as preamble and decode the rest. The 100% data is a bit different–it decodes the original data symbols.

At first glance, we can observe that the BER goes down on both time-axis and flow-axis when the proportion of the preamble increases, which is a direct consequence of the increasing data samples for channel estimation. However, in the meanwhile, using time-axis can never achieve the same performance as flow-axis, because it only finds the best fitting CIR for the preamble, which cannot be generalized to the whole packet. Using the same data, Figure. 6.8b also validates the claim through the perspective of fitting error. As the proportion of preamble increases, the fitting error improvement is less deviated from the median (around 20), which means the flow-axis advantage is maintained.

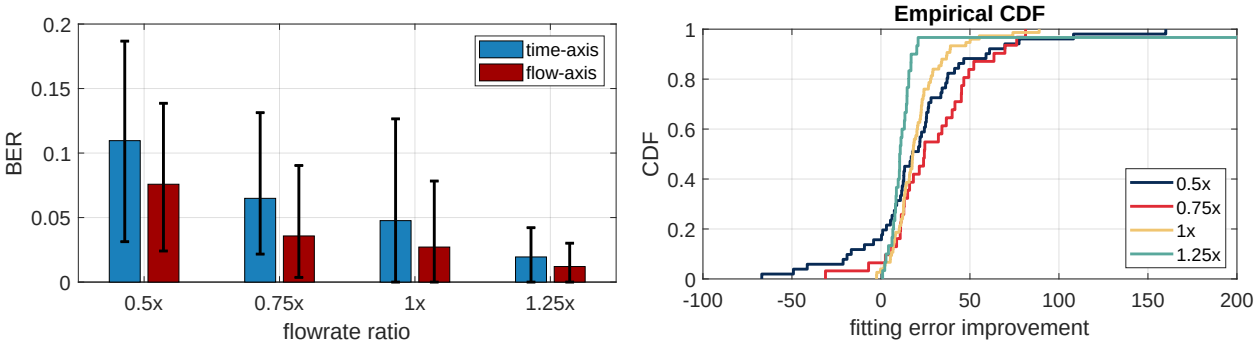
6.5.2 Micro Benchmarks

In the rest of this section, we will provide a few micro benchmarks showing the advantage of FlowLink under a few typical parameters, including the average flow rate, the background flow period, the TX-RX distance and the channel topology.

Average flow rate

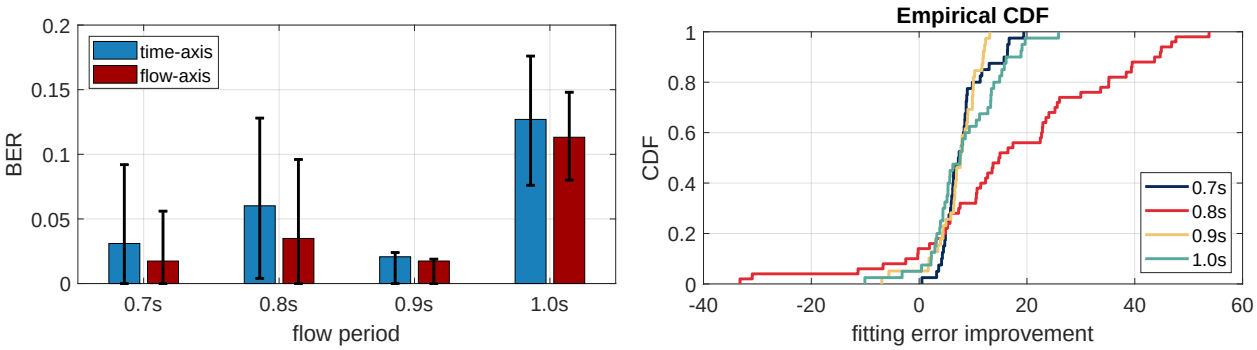
Next, we will show FlowLink's performance with regard to the average flow rate. In the following experiments, the channel is a single 10 cm tube of 4 mm ID, and the symbol interval is 0.6 seconds. The basic flow cycle is 12 mL/s (100 cm/s) for 0.3 seconds followed by 3 mL/s (25 cm/s) for 0.5 seconds. The flow rate is scaled by 0.5/0.75/1.0/1.25× respectively, but the duration of each rate is kept the same.

As shown in Figure. 6.9a, there is a clear decreasing trend in BER when the average flow rate is increasing. Faster flow reduces the propagation delay from the transmitter to the receiver. It also reduces the spreading of molecular signals due to laminar flow, which indicates decreasing Inter-Symbol Interference (ISI) and is the explanation for the decreasing BER. However, at all flow rates, the advantage of the flow-axis is clear.



(a) BER (b) Fitting error improvement

Figure 6.9: The performance of FlowLink under different average flow rates, with GTPD decoding process.



(a) BER (b) Fitting error improvement

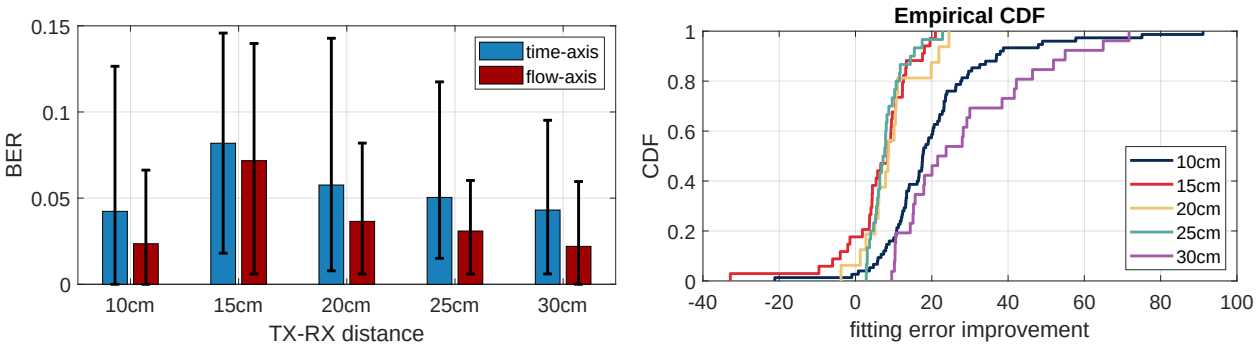
Figure 6.10: The performance of FlowLink under different total flow periods, with GTPD decoding process.

From another perspective, Figure. 6.9b shows that as the flow rate increases, the fitting error improvement is more clustered around a lower but still positive value. Finally, it can be inferred from these two figures that the time-axis BER will be good enough at a high flow rate such that flow-axis process is no longer required.

Flow period

Since each individual has a different cardiac cycle, we take a close look at FlowLink's performance over blood flows of different periods. In the following experiments, the channel is a single 10 cm tube of 4 mm ID, and the symbol interval is 0.8 seconds. The fast flow in the cycle is 12 mL/s (100 cm/s) for 0.3 seconds, but the slow flow is 3 mL/s (25 cm/s) for 0.4/0.5/0.6/0.7 seconds, so the total period is 0.7/0.8/0.9/1.0 seconds.

In Figure. 6.10b, we can see that for all flow periods, at least 15% of the data shows improvement in the fitting error when adopting FlowLink. Correspondingly, Figure. 6.10a shows the improvement in BER for all conditions. One thing to notice is that the growing flow period indicates a greater proportion of LOW flow in the cycle. This leads to the slower average flow speed and longer propagation delay. Intuitively, the overall BER is meant to increase because of the increasing ISI, like the trend of 0.7/0.8/1.0 seconds. However, 0.9 seconds wields an unusual trend, where the BER is much lower than its neighbors.



(a) BER

(b) Fitting error improvement

Figure 6.11: The performance of FlowLink under different TX-RX distance, with GTPD decoding process.

TX-RX distance

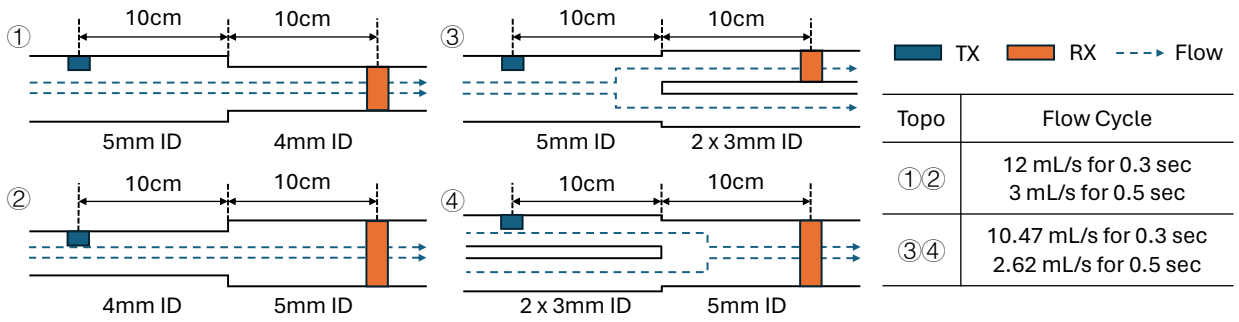
The TX-RX distance is another parameter that has an interesting relation with the flow. In the following experiments, the channel is a single tube of 4 mm ID, and the symbol interval is 0.6 seconds. The basic flow cycle is 12 mL/s (100 cm/s) for 0.3 seconds followed by 3 mL/s (25 cm/s) for 0.5 seconds. The distance between the transmitter and the receiver varies between 10/15/20/25/30 cm.

In Figure. 6.11b, we can see that for all TX-RX distances, at least 20% of the data shows improvement in the fitting error when adopting FlowLink. Correspondingly, In Figure. 6.11a shows the improvement in BER for all conditions. However, there is an unusual trend in BER as a function of growing TX-RX distance. Intuitively, the direct consequence of the increasing TX-RX distance is the longer propagation delay, which will lead to more ISI and higher BER. But except for the 10 cm condition, other conditions reveal the opposite trend, i.e. the BER of 30 cm is lower than the BER of 15 cm. This is an interesting joint result of ISI and varying flow–ISI can partially compensate for the error in the sampling time. For a simple case when a single pulse function passes a channel without ISI (i.e. the received signal is also a pulse function), a minor error in the sampling time will return value zero. On the contrary, a channel with ISI can tolerate such error under a certain threshold.

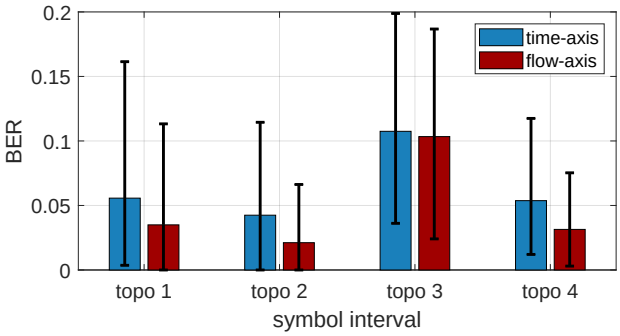
Various channel topology

As claimed in Section. 6.3.1, FlowLink not only works for a simple channel of a single tube with a constant diameter but can be generalized to a more complicated topology. The following experiments test FlowLink's performance in four different vessel topologies, which are presented in Figure. 6.12. They represent the basic channel variation in blood vessels, including narrowing (topo 1), widening (topo 2), forking (topo 3) and merging (topo 4). The tubes are connected with I-shape or Y-shape connectors. The dimension of the tubes and the distance between the transmitter/receiver and the connector are marked in the figure.

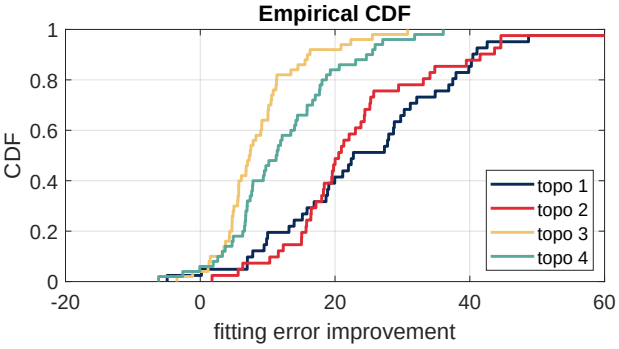
From Figure. 6.12b, we can see that FlowLink shows clear improvement in decoding under topo 1/2/4. However, topo 3 is an exception, where the BER for both FlowLink and baseline are much worse than the other three topologies. By analyzing all four topologies, we can see that under topo 1/2/4, all particles released by the transmitter will pass the receiver detection range, but under topo 3, only a proportion of the information-bearing particles will go to the receiver fork. It is possible that such proportion is not a constant under the experimental periodic flow, which experiences sudden velocity change and might lead to unstable distribution around the forking. And that is why the time-axis decoder also encounters a bad



(a) Four variants of the molecular channel: narrowing (topo 1), widening (topo 2), forking (topo 3) and merging (topo 4). And the simulated flow under each condition.



(b) BER



(c) Fitting error improvement

Figure 6.12: The performance of FlowLink under four variants of the molecular channel, with GTPD decoding process.

performance. However, the advantage of FlowLink can still be validated through the fitting error improvement in Figure. 6.12c. Further investigation is required on detailed flow measuring on both forks and better blood flow simulation.

6.6 Related Work

6.6.1 Blood network testbeds

There has been a lot of effort in simulating the blood network channel. [98], [114] used silicone tubes to simulate the vessels and evaluate the performance of decoders but were not strict with the flow speed. [115] used a 2-mm diameter silicone tube to simulate the vascular system with volume rates between 50 mL/min and 90 mL/min. [116] used a 3-mm diameter tube which is close to the average diameter of coronary artery with volume rates between 50 mL/min and 350 mL/min. [117] used a narrower 1.52-mm diameter tube and lower volume rates between 2 mL/min and 20 mL/min, but tested the propagation of the bio-compatible SPION particles. However, the above work only provided some results under constant flow rate and ignored the influence of flow variation due to heartbeat.

6.6.2 Fast-varying channel in conventional communication systems

This chapter, to our knowledge, is the first to address the fast-varying channel in molecular communication. But there have been various works for conventional wireless channels. Fast-varying channel was first officially

considered in radio astronomy [118], [119], but has nowadays become a common phenomenon in most popular communication systems due to the rapidly growing mobility in different user scenarios. Besides, underwater acoustic communication views much higher relative Doppler shifts, since the sound speed in water is only around 1500 m/s [120]. [121] proposed a two-step approach to mitigate the Doppler influence in zero-padded OFDM for underwater acoustic communication, which includes a step of non-uniform resampling based on frequency-dependent Doppler drifts. [122] further demonstrated the effusiveness of Doppler-based resampling for underwater acoustic OFDM transmission without guard interval, but at a much lower-than-normal symbol rate. Despite the numerous solutions for conventional communication systems, it is possible that none of them can be directly applied to MC, because the delay spread can be much longer than the channel coherence time in the blood network.

6.7 Discussions

6.7.1 Flow Sampling

Compared to conventional decoders, FlowLink employs a re-sampling operation in the first phase in order to recover linear assumptions in the molecular signal, such that conventional decoders can be applied. However, re-sampling is known to have outstanding drawbacks. On the one hand, from the communication perspective, re-sampling could lead to loss of information and aliasing because the molecular channel response is not band limited. On the other hand, from the hardware point, resampling consumes extra computation and memory resources, which are both limited for bio-nano devices in the current theory.

But, since we are still in a time when hundreds of molecular communication theories sprout, can we dream bigger and dive into the possibilities of designing flow sampling architectures to replace the conventional time sampling? The advantage is obvious that the sampled signals can be directly applied to a linear decoder. Nevertheless, it also comes with new challenges. For example, the sampling flow interval should not be a unanimous number, which could lead to information loss due to the flow difference in a complex molecular channel topology.

6.7.2 Complication of Blood Flow

The real blood vessel is a much more complicated system than what is studied in this chapter, so a few assumptions do not always hold.

First, laminar flow is not always true, and the contradictory is turbulence–unpredictable flow. Turbulence can lead to a similar influence as diffusion, but with a much larger coefficient. Whether a system is turbulent is usually described by the Reynolds number whose value in this testbed is given as:

$$Re = \frac{v}{\nu} L = \frac{0.3\ m/s}{10^{-6}\ m^2/s} \times \frac{4 \times 10^{-3}\ m}{4} = 300 \tag{6.6}$$

Usually, a system with Reynolds number $Re < 2000$ is assumed mostly laminar. However, it is reported that there exists turbulence in physiological blood flow [155]. It is not only a driving factor for numerous diseases such as atherosclerosis, stenosis and aneurysm (large Reynolds number), but also is observed under common cases (low Reynolds number), which is a challenge to FlowLink.

Second, the possibility of reverse blood flow is ignored, which is another possible symptom of blood disease. FlowLink adopts the one-to-one mapping when there is no reverse flow, which is not the optimum choice for

patients. How to use the flow-axis appropriately without losing any information is also a good topic.

Third, the diameter of blood vessels is not time-invariant. For example, the muscle pumps with more blood flow during workouts. This would influence the topology of the blood network, especially forking and merging. It is worthwhile to study the influence of such variation, whether it can be circumvented and how to address the issue before the actual deployment of the blood network.

6.7.3 Flow estimation

Using flow meter for real-time speed tracking complicates the architecture of a practical molecular device, because it is an extra sensor from original functionality. However, flow speed estimation is possible with pure communication module by using two receivers–one in the upstream and one in the downstream. Each sample of the downstream signal is a delayed version of one in the upstream signal. By finding the matched samples and the time difference between them, a linear equation system on flow speed can be constructed and solved–which falls in the reign of Dynamic Time Warping (DTW) [156]. Besides, the distance between the two receivers need not be known, because FlowLink only cares about the relative flow variation. Moreover, the periodicity of blood flow can also be employed to improve the estimation, together with either conventional methods or the uprising deep learning solutions (such as LSTM) that are proficient in time series processing.

6.8 Conclusion

This chapter looks into the blood network, which is the backbone that exchanges information between different parts of the body in a molecular-communication-enabled Internet of Bio-Nano Things–devices deployed along the blood vessels can release information-bearing particles which will travel along the blood flow, arrive at other devices and enable information delivery. Apart from the common challenges that arise from the advection-and-diffusion-based propagation mechanism, it the first to identify the destructive influence of the blood flow on state-of-the-art decoders–the heart pumping behavior results in the drastically varying blood flow rate in each heartbeat cycle. This causes complicated symbol boundary drift at the receiver end, especially when the symbol rate is in the same order of the heartbeat rate, because the propagation delay of molecular signals in the blood vessel is highly dependent on the blood flow rate. FlowLink is then proposed to compensate for such influence. Inspired by the observation that the total amount of flow during the propagation of any signal between two static devices is constant, FlowLink introduces the idea of flow sampling, which samples the signal every time when a certain amount of flow passes the receiver with the help of a flow meter. Using flow-axis, the molecular channel becomes coherent for a packet, but the symbol boundaries remain unequally spaced as a result of the constant-rate transmission, which is solved by signal interpolation and transmission time detection. Experiments show that FlowLink can greatly improve the symbol boundary tracking of molecular packets and, as a result, reduce the BER under various blood channel topologies.

Chapter 7

Conclusion and Future Work

In this dissertation, we have dived into a branch of the huge molecular communication family–diffusion-based communication in the blood circulatory system. Such network is one of the key components of the Internet of Bio-Nano Things (IoBNT), which will revolutionize healthcare by implanting various bio-compatible medical devices in every part of the body. These implants have various functionalities, but through networking, they can cooperate to continuously monitor the body environment, help with accurate diagnosis and provide long-term and pervasive treatment. In order to establish such a network, molecular communication is a strong challenger to conventional wireless techniques, because it wields unparalleled advantages in biocompatibility, long-term service and small form factors.

Molecular communication has multiple signal propagation mechanisms, and they each are suitable for providing different scales of network coverage. For example, molecular motors can guide proteins to move along cytoskeletal filaments, which is suitable for information exchange between cells and be used to enable communication in a bio-nano device, while bacteria can carry information-bearing molecules and swim across the medium to provide wider coverage and can be used to implement a subnet in the tissue. This dissertation focuses on diffusion-based molecular communication in the blood vessels, which leverages the blood flow to achieve a longer communication range and quickly delivers information between the central control of IoBNT and different parts of the body.

However, exploiting the vast potential of molecular communication are also accompanied by new challenges. Although rich theories have been developed for conventional wireless communication systems, a great proportion of them can not be directly applied to molecular signal, because it is a completely different carrier–the contrasting propagation mechanism (particles versus waves) leads to some unique channel properties that are unseen in conventional systems.

- **Non-negative signal.** Particles are the information carrier in molecular communication. Despite of the various modulation schemes–the concentration, the type, etc. of the information-bearing particles, they all can be attributed to the amount of the particles. Thus, any kind of molecular signal with regard to one information-bearing particle can be transformed into the amount of a certain particle, which is a non-negative value in nature. Unlike wireless signals with phase and amplitude as the properties of waves, there is no cancellation between positive and negative signals in molecular communication. Even if there exist plenty of particles that may react with others or naturally degrade in the environment, such phenomena are more than simply adding or subtracting, which are much more complicated than wireless signals.

- **Signal-dependent noise.** The RX signals are fundamentally the counting of information-bearing particles. It should be noted that the propagation of particles is Brownian motion (naming diffusion), which is considered as a random movement due to the difficulty to accurately track the collisions among countless particles. Thus, the number of particles counted by the receiver is also a random variable, which can be approximated to Gaussian distribution in the blood vessel channel. As a result, the variance of the RX signal, which can be treated as a type of noise, is strongly related to the expectation of the RX signal, i.e. MC channel has signal-dependent noise. This is the inherent property that comes before the RX sampling noise.

- **Non-causal channel.** The non-causality does not mean the receiver can sense the information-bearing particle before it is released by the transmitter. Instead, it means that particles from later symbols can arrive earlier than those from previous symbols. As was explained that the propagation of particles is random, it is always true that some particles will arrive later than expected, while some will arrive earlier. The shorter the symbol interval, the more particles will break the order. The receiver has to consider whether there are particles from future symbols and how many are they in order to achieve non-biased decoding.

- **Lack of synchronization.** Apart from the random movement caused by collision between particles, the propagation is also influenced by the flow in the environment, which is the mass transportation (naming advection). The particles will move along the flow direction, so the receiver in the downstream has higher probability of receiving the information-bearing particles than the receiver in the upstream. This bias is amplified when the flow increases, and until some point at which the upstream receiver can barely sense nothing from the transmitter. This uni-directional communication introduces challenges over conventional protocols that highly rely on the feedback between devices.

- **Long delay spread.** The interference between symbols is the direct consequence of the spreading in the propagation delay of signals. In molecular communication, one source of the spreading is diffusion. Another source is the complex components of the environmental flow. Due to the existence of fluid viscosity, the flow speed is not constant. Instead, the fluid near the wall of the channel moves slower than that in the middle. The information-bearing particles will propagate at different speeds in different areas of the fluid, which also causes the spreading in the propagation delay. Usually, the variance of the propagation delay is positively related to the expectation.

- **Short coherence time.** In the coherence time, the channel is considered as static so the channel response estimated from the preamble can be used to decode the following data symbols. Coherence time is decided by the changes in the environment, and the short coherence time usually indicates rapid changes. The blood vessel channel is not a stable environment, when the blood flow and vessel shape could change based on the activity and posture of the person. Such changes are in the scale of the channel delay spread, which complicates decoding.

- **Heartbeat-induced varying channel.** Apart from the above gradual channel variation, the blood channel is also fast-varying, which is caused by the heartbeat. Blood flow is the main character in the circulation system that delivers oxygen and nutrition to every part of the body and carries away the waste. Heart provides the momentum for blood circulation, and it periodically extracts to let the blood in from veins and squeezes to push the blood into arteries. This causes a periodic flow variation with a cycle normally within 1 second and over 2× max-to-min flow ratio. Considering the common

datarate in the blood channel, the symbol interval is at the same scale of the heartbeat cycle, so even the channel response of consecutive symbols deviates.

In this dissertation, we proposed μ-Link, MoMA and FlowLink to specifically analyze and address the above challenges to implement the molecular communication in the blood network.

- μ-Link focused on the channel between one TX and one RX and addressed the general challenges–signal-dependent noise, non-causal channel, high ISI and short coherence time. μ-Link uses the HMM (Hidden Markov Model) to model the received signal as a random observation of a Markov state, which is a sequence of consecutive transmitted symbols and is appropriate for the high ISI. The emission probability of the HMM takes into consideration the signal-dependent noise. The state sequence includes not only the previous symbols, but also a few incoming symbols, which handles the non-causal channel. Finally, Viterbi algorithm solves the HMM problem and outputs the decoded symbols, while a channel tracking process is introduced periodically to compensate for the short coherence time.

- MoMA continued to the channel between multiple TX and aimed at non-negative signal and uni-directional propagation, which caused trouble to the medium access. The uni-directional channel excludes the possibility of feedback synchronization. Collision is inevitable, so the receiver has to take full responsibility for decoding multiple packets, especially under collision. The CDMA scheme is adopted to ensure scalability and fairness among transmitters, but the addition of non-negative molecular signals causes trouble identifying a new packet from the slowly varying channel and noise. MoMA designed the new preamble structure, which has more consecutive "1"s and is distinguished from the CDMA data symbols. This greatly improves the detection of collisions. Moreover, MoMA points out that the similarity between the channel response of different particles can be leveraged–when the transmitters wield multiple information-bearing particles, the CDMA code can repeat on one particle. This extends the scalability of conventional wireless communication.

- FlowLink noticed the fast-varying channel challenge and introduced the preprocess operation such that the above molecular decoders still work. We noticed that although the propagation delay varies for different symbols, the total amount of flow required for the propagation is constant for any two fixed locations in the blood channel. Thus, FlowLink introduces the re-sampling phase to convert the time-axis signal to the flow-axis, while the flow is computed by integrating the measurement of the flow meter over time. It eliminates the fast-varying channel response in the symbol scale, but also causes the trouble that the symbols are no longer equally spaced on the flow-axis. Thus, the corresponding spacing between symbols are also computed, while interpolation is used accordingly to fit the packet detection, channel estimation and the Viterbi algorithm decoding.

Since the first testbed built for μ-Link, we have renovated it from multiple perspectives. Not only has it met the requirement of multiple transmitters for MoMA and heartbeat-like flow for FlowLink, but it has also become more similar to a practical blood channel. Extensive experiments with varying parameters have been conducted and validated the contributions of this dissertation.

However, we do recognize the limitations of the current work and there is still rich work to be done in this direction.

- **The testbed can be improved.** Although we have established a testbed to simulate the blood vessel channel, there are a lot of possible improvement that can inspire more qualitative and quantitative

finding. For example, microfluid chip techniques can be used to reduce the diameter of the tube from centimeters to micrometers, which is similar to the capillaries and can greatly push the variety in the testbed topology. Besides, more delicate flow control can reveal the propagation of molecular signals under various flow conditions. Beyond the blood network, nano-scale testbeds can be developed to study molecular communication in tissues and cells, which is another important component of the IoBNT network.

- **Metric is not universal.** At the current stage, the experimental results in the community (including this dissertation) cannot be directly compared with each other because the experiments are conducted under the environment with different parameters–for example, results reported on a liquid testbed cannot be compared to those on a gas testbed although the method are similar. Researchers cannot directly tell if there is improvement unless they replicate previous work on their own testbed. SNR is a universal description for such case in conventional wireless field, but the definition for SNR is uncertain in molecular field–the wireless definition is not applicable to molecular because one is wave but the other is matter. Unfortunately, this dissertation cannot provide a universal metric for easy comparison either for the community, but we hope extensive experiments can be done to find this metric, which can greatly facilitate molecular research.

- **The practicality of the work is limited.** Although we can declare the proposed protocols are practical in the perspective of communication without remorse, we have to admit that a great proportion or all of the assumptions made for the work are impractical in the perspective of chemistry and biology. Although these assumptions are appropriate for a dissertation in communication, there still exist numerous questions for the molecular system beyond the scope of the channel, e.g. "What is the information-bearing particle that is suitable for the in-body environment?", "What is the transmitter hardware to release the information-bearing particles and how can it store the particles?", "What is the receiver hardware to detect the information-bearing particles and what noise does it introduce?", etc. These questions have to be answered one by one towards the future of a real IoBNT system.

In summary, although still in its early stage, molecular communication has revealed its potential in body-range communication, which is promising to revolutionize the healthcare system. However, as a strong interdisciplinary topic, the strong cooperation among researchers from communication, physics, chemistry, biology, and many other fields is the right direction in the future.

References

[1] I. F. Akyildiz, M. Pierobon, S. Balasubramaniam, and Y. Koucheryavy, "The internet of bio-nano things," *IEEE Communications Magazine*, vol. 53, no. 3, pp. 32–40, 2015.

[2] M. Mimee, P. Nadeau, A. Hayward, *et al.*, "An ingestible bacterial-electronic system to monitor gastrointestinal health," *Science*, vol. 360, no. 6391, pp. 915–918, 2018.

[3] P. Nadeau, M. Mimee, S. Carim, T. K. Lu, and A. P. Chandrakasan, "21.1 nanowatt circuit interface to whole-cell bacterial sensors," in *2017 IEEE International Solid-State Circuits Conference (ISSCC)*, IEEE, 2017, pp. 352–353.

[4] Y. Rajavi, M. Taghivand, K. Aggarwal, A. Ma, and A. S. Poon, "An rf-powered fdd radio for neural microimplants," *IEEE Journal of Solid-State Circuits*, vol. 52, no. 5, pp. 1221–1229, 2017.

[5] O. Jonas, H. M. Landry, J. E. Fuller, *et al.*, "An implantable microdevice to perform high-throughput in vivo drug sensitivity testing in tumors," *Science translational medicine*, vol. 7, no. 284, 284ra57–284ra57, 2015.

[6] E. Samiei, M. Tabrizian, and M. Hoorfar, "A review of digital microfluidics as portable platforms for lab-on a-chip applications," *Lab on a Chip*, vol. 16, no. 13, pp. 2376–2396, 2016.

[7] A. Van Reenen, A. M. de Jong, J. M. den Toonder, and M. W. Prins, "Integrated lab-on-chip biosensing systems based on magnetic particle actuation–a comprehensive review," *Lab on a Chip*, vol. 14, no. 12, pp. 1966–1986, 2014.

[8] T. Danino, O. Mondragón-Palomino, L. Tsimring, and J. Hasty, "A synchronized quorum of genetic clocks," *Nature*, vol. 463, no. 7279, pp. 326–330, 2010.

[9] D. M. Shcherbakova, A. A. Shemetov, A. A. Kaberniuk, and V. V. Verkhusha, "Natural photoreceptors as a source of fluorescent proteins, biosensors, and optogenetic tools," *Annual review of biochemistry*, vol. 84, pp. 519–550, 2015.

[10] B. Krishnaswamy, "Algorithms for molecular communication networks," Ph.D. dissertation, Georgia Institute of Technology, 2018.

[11] J. E. Dahlman, K. J. Kauffman, Y. Xing, *et al.*, "Barcoded nanoparticles for high throughput in vivo discovery of targeted therapeutics," *Proceedings of the National Academy of Sciences*, vol. 114, no. 8, pp. 2060–2065, 2017.

[12] H. Shi, K. Nie, B. Dong, M. Long, H. Xu, and Z. Liu, "Recent progress of microfluidic reactors for biomedical applications," *Chemical Engineering Journal*, vol. 361, pp. 635–650, 2019.

[13] C. Steiger, A. Abramson, P. Nadeau, A. P. Chandrakasan, R. Langer, and G. Traverso, "Ingestible electronics for diagnostics and therapy," *Nature Reviews Materials*, vol. 4, no. 2, pp. 83–98, 2019.

[14] J. S. McLean, "A re-examination of the fundamental limits on the radiation q of electrically small antennas," *IEEE Transactions on antennas and propagation*, vol. 44, no. 5, p. 672, 1996.

[15] D. F. Sievenpiper, D. C. Dawson, M. M. Jacob, *et al.*, "Experimental validation of performance limits and design guidelines for small antennas," *IEEE Transactions on Antennas and Propagation*, vol. 60, no. 1, pp. 8–19, 2011.

[16] T. Yamada, T. Uezono, K. Okada, K. Masu, A. Oki, and Y. Horiike, "Rf attenuation characteristics for in vivo wireless healthcare chip," *Japanese journal of applied physics*, vol. 44, no. 7R, p. 5275, 2005.

[17] O. B. Akan, H. Ramezani, T. Khan, N. A. Abbasi, and M. Kuscu, "Fundamentals of molecular information and communication science," *Proceedings of the IEEE*, vol. 105, no. 2, pp. 306–318, 2016.

[18] Y. Chahibi, M. Pierobon, S. O. Song, and I. F. Akyildiz, "A molecular communication system model for particulate drug delivery systems," *IEEE Transactions on biomedical engineering*, vol. 60, no. 12, pp. 3468–3483, 2013.

[19] Y. Chahibi, I. F. Akyildiz, S. Balasubramaniam, and Y. Koucheryavy, "Molecular communication modeling of antibody-mediated drug delivery systems," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 7, pp. 1683–1695, 2015.

[20] Y. Chahibi, M. Pierobon, and I. F. Akyildiz, "Pharmacokinetic modeling and biodistribution estimation through the molecular communication paradigm," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 10, pp. 2410–2420, 2015.

[21] M. Kuscu, E. Dinc, B. A. Bilgin, H. Ramezani, and O. B. Akan, "Transmitter and receiver architectures for molecular communications: A survey on physical design with modulation, coding, and detection techniques," *Proceedings of the IEEE*, vol. 107, no. 7, pp. 1302–1341, 2019.

[22] S. Salehi, N. S. Moayedian, S. H. Javanmard, and E. Alarcón, "Lifetime improvement of a multiple transmitter local drug delivery system based on diffusive molecular communication," *IEEE transactions on nanobioscience*, vol. 17, no. 3, pp. 352–360, 2018.

[23] 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, 2012.

[24] B. D. Unluturk, A. O. Bicen, and I. F. Akyildiz, "Genetically engineered bacteria-based biotransceivers for molecular communication," *IEEE Transactions on Communications*, vol. 63, no. 4, pp. 1271–1281, 2015.

[25] R. Daniel, J. R. Rubens, R. Sarpeshkar, and T. K. Lu, "Synthetic analog computation in living cells," *Nature*, vol. 497, no. 7451, pp. 619–623, 2013.

[26] 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, 2011.

[27] P. Siuti, J. Yazbek, and T. K. Lu, "Synthetic circuits integrating logic and memory in living cells," *Nature biotechnology*, vol. 31, no. 5, pp. 448–452, 2013.

[28] M. Fakruddin, Z. Hossain, and H. Afroz, "Prospects and applications of nanobiotechnology: A medical perspective," *Journal of nanobiotechnology*, vol. 10, no. 1, pp. 1–8, 2012.

[29] G. Rampioni, L. Leoni, and P. Stano, "Molecular communications in the context of "synthetic cells" research," *IEEE Transactions on NanoBioscience*, vol. 18, no. 1, pp. 43–50, 2018.

[30] S. Slomovic, K. Pardee, and J. J. Collins, "Synthetic biology devices for in vitro and in vivo diagnostics," *Proceedings of the National Academy of Sciences*, vol. 112, no. 47, pp. 14 429–14 435, 2015.

[31] B. Motion, "Brownian motion and stochastic calculus," 2008.

[32] S. Goldstein, "Mechanical models of brownian motion," in *Mathematical Problems in Theoretical Physics: Proceedings of the VIth International Conference on Mathematical Physics Berlin (West), August 11–20, 1981*, Springer, 2005, pp. 21–24.

[33] V. L. Streeter, *Fluid mechanics streeter*, 1962.

[34] M. Di Lorenzo, A. R. Thomson, K. Schneider, P. J. Cameron, and I. Ieropoulos, "A small-scale air-cathode microbial fuel cell for on-line monitoring of water quality," *Biosensors and Bioelectronics*, vol. 62, pp. 182–188, 2014.

[35] W. Zhang, A. M. Asiri, D. Liu, D. Du, and Y. Lin, "Nanomaterial-based biosensors for environmental and biological monitoring of organophosphorus pesticides and nerve agents," *TrAC Trends in Analytical Chemistry*, vol. 54, pp. 1–10, 2014.

[36] J. Nemunaitis, C. Cunningham, N. Senzer, *et al.*, "Pilot trial of genetically modified, attenuated salmonella expressing the e. coli cytosine deaminase gene in refractory cancer patients," *Cancer gene therapy*, vol. 10, no. 10, pp. 737–744, 2003.

[37] A. Schroeder, D. A. Heller, M. M. Winslow, *et al.*, "Treating metastatic cancer with nanotechnology," *Nature Reviews Cancer*, vol. 12, no. 1, pp. 39–50, 2012.

[38] T. Danino, A. Prindle, G. A. Kwong, *et al.*, "Programmable probiotics for detection of cancer in urine," *Science translational medicine*, vol. 7, no. 289, 289ra84–289ra84, 2015.

[39] B. Wang, M. Barahona, and M. Buck, "A modular cell-based biosensor using engineered genetic logic circuits to detect and integrate multiple environmental signals," *Biosensors and Bioelectronics*, vol. 40, no. 1, pp. 368–376, 2013.

[40] B. Wang, R. I. Kitney, N. Joly, and M. Buck, "Engineering modular and orthogonal genetic logic gates for robust digital-like synthetic biology," *Nature communications*, vol. 2, no. 1, p. 508, 2011.

[41] B. H. Weinberg, N. Pham, L. D. Caraballo, *et al.*, "Large-scale design of robust genetic circuits with multiple inputs and outputs for mammalian cells," *Nature biotechnology*, vol. 35, no. 5, pp. 453–462, 2017.

[42] M. B. Elowitz and S. Leibler, "A synthetic oscillatory network of transcriptional regulators," *Nature*, vol. 403, no. 6767, pp. 335–338, 2000.

[43] A. E. Friedland, T. K. Lu, X. Wang, D. Shi, G. Church, and J. J. Collins, "Synthetic gene networks that count," *science*, vol. 324, no. 5931, pp. 1199–1202, 2009.

[44] S. Hernández-Ainsa and U. F. Keyser, "Dna origami nanopores: Developments, challenges and perspectives," *Nanoscale*, vol. 6, no. 23, pp. 14 121–14 132, 2014.

[45] C. Dekker, "Solid-state nanopores," *Nature nanotechnology*, vol. 2, no. 4, pp. 209–215, 2007.

[46] E. Slonkina and A. B. Kolomeisky, "Polymer translocation through a long nanopore," *The Journal of chemical physics*, vol. 118, no. 15, pp. 7112–7118, 2003.

[47] K. Chen, M. Juhasz, F. Gularek, *et al.*, "Ionic current-based mapping of short sequence motifs in single dna molecules using solid-state nanopores," *Nano letters*, vol. 17, no. 9, pp. 5199–5205, 2017.

[48] N. A. Bell and U. F. Keyser, "Digitally encoded dna nanostructures for multiplexed, single-molecule protein sensing with nanopores," *Nature nanotechnology*, vol. 11, no. 7, pp. 645–651, 2016.

[49] T. Nakano, T. Suda, M. Moore, R. Egashira, A. Enomoto, and K. Arima, "Molecular communication for nanomachines using intercellular calcium signaling," in *5th IEEE Conference on Nanotechnology, 2005.*, IEEE, 2005, pp. 478–481.

[50] J. Isaksson, P. Kjäll, D. Nilsson, N. Robinson, M. Berggren, and A. Richter-Dahlfors, "Electronic control of ca2+ signalling in neuronal cells using an organic electronic ion pump," *Nature materials*, vol. 6, no. 9, pp. 673–679, 2007.

[51] A. Jonsson, T. A. Sjöström, K. Tybrandt, M. Berggren, and D. T. Simon, "Chemical delivery array with millisecond neurotransmitter release," *Science advances*, vol. 2, no. 11, e1601340, 2016.

[52] I. Uguz, C. M. Proctor, V. F. Curto, *et al.*, "A microfluidic ion pump for in vivo drug delivery," *Advanced Materials*, vol. 29, no. 27, p. 1 701 217, 2017.

[53] S. R. Schwarze, A. Ho, A. Vocero-Akbani, and S. F. Dowdy, "In vivo protein transduction: Delivery of a biologically active protein into the mouse," *Science*, vol. 285, no. 5433, pp. 1569–1572, 1999.

[54] N.-R. Kim and C.-B. Chae, "Novel modulation techniques using isomers as messenger molecules for nano communication networks via diffusion," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 12, pp. 847–856, 2013.

[55] M. Chourasia and S. Jain, "Pharmaceutical approaches to colon targeted drug delivery systems," *J Pharm Pharm Sci*, vol. 6, no. 1, pp. 33–66, 2003.

[56] M. S. Kuran, H. B. Yilmaz, T. Tugcu, and I. F. Akyildiz, "Modulation techniques for communication via diffusion in nanonetworks," in *2011 IEEE international conference on communications (ICC)*, IEEE, 2011, pp. 1–5.

[57] M. U. Mahfuz, D. Makrakis, and H. Mouftah, "Spatiotemporal distribution and modulation schemes for concentration-encoded medium-to-long range molecular communication," in *2010 25th Biennial Symposium on Communications*, IEEE, 2010, pp. 100–105.

[58] Y. Murin, N. Farsad, M. Chowdhury, and A. Goldsmith, "On time-slotted communication over molecular timing channels," in *Proceedings of the 3rd ACM International Conference on Nanoscale Computing and Communication*, 2016, pp. 1–6.

[59] K. V. Srinivas, A. W. Eckford, and R. S. Adve, "Molecular communication in fluid media: The additive inverse gaussian noise channel," *IEEE transactions on information theory*, vol. 58, no. 7, pp. 4678–4692, 2012.

[60] B. Krishnaswamy, C. M. Austin, J. P. Bardill, *et al.*, "Time-elapse communication: Bacterial communication on a microfluidic chip," *IEEE Transactions on Communications*, vol. 61, no. 12, pp. 5139–5151, 2013.

[61] N. Garralda, I. Llatser, A. Cabellos-Aparicio, E. Alarcón, and M. Pierobon, "Diffusion-based physical channel identification in molecular nanonetworks," *Nano Communication Networks*, vol. 2, no. 4, pp. 196–204, 2011.

[62] M. Blawat, K. Gaedke, I. Huetter, *et al.*, "Forward error correction for dna data storage," *Procedia Computer Science*, vol. 80, pp. 1011–1022, 2016.

[63] L. Organick, S. D. Ang, Y.-J. Chen, *et al.*, "Random access in large-scale dna data storage," *Nature biotechnology*, vol. 36, no. 3, pp. 242–248, 2018.

[64] F. Dinc, B. C. Akdeniz, E. Erol, *et al.*, "Analytical derivation of the impulse response for the bounded 2-d diffusion channel," *Physics Letters A*, vol. 383, no. 14, pp. 1589–1600, 2019.

[65] F. Dinc, B. C. Akdeniz, A. E. Pusane, and T. Tugcu, "A general analytical solution to impulse response of 3-d microfluidic channels in molecular communication," *arXiv preprint arXiv:1804.10071*, 2018.

[66] W. Wicke, T. Schwering, A. Ahmadzadeh, V. Jamali, A. Noel, and R. Schober, "Modeling duct flow for molecular communication," in *2018 IEEE Global Communications Conference (GLOBECOM)*, IEEE, 2018, pp. 206–212.

[67] M. Zoofaghari and H. Arjmandi, "Diffusive molecular communication in biological cylindrical environment," *IEEE transactions on nanobioscience*, vol. 18, no. 1, pp. 74–83, 2018.

[68] M. Turan, M. Ş. Kuran, H. B. Yilmaz, I. Demirkol, and T. Tugcu, "Channel model of molecular communication via diffusion in a vessel-like environment considering a partially covering receiver," in *2018 IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom)*, IEEE, 2018, pp. 1–5.

[69] A. Noel, Y. Deng, D. Makrakis, and A. Hafid, "Active versus passive: Receiver model transforms for diffusive molecular communication," in *2016 IEEE Global Communications Conference (GLOBECOM)*, IEEE, 2016, pp. 1–6.

[70] N. Varshney, W. Haselmayr, and W. Guo, "On flow-induced diffusive mobile molecular communication: First hitting time and performance analysis," *IEEE Transactions on Molecular, Biological and Multi-Scale Communications*, vol. 4, no. 4, pp. 195–207, 2018.

[71] G. Chang, L. Lin, and H. Yan, "Adaptive detection and isi mitigation for mobile molecular communication," *IEEE Transactions on nanobioscience*, vol. 17, no. 1, pp. 21–35, 2017.

[72] M. Turan, B. C. Akdeniz, M. Ş. Kuran, *et al.*, "Transmitter localization in vessel-like diffusive channels using ring-shaped molecular receivers," *IEEE Communications Letters*, vol. 22, no. 12, pp. 2511–2514, 2018.

[73] V. Jamali, A. Ahmadzadeh, C. Jardin, H. Sticht, and R. Schober, "Channel estimation for diffusive molecular communications," *IEEE Transactions on Communications*, vol. 64, no. 10, pp. 4238–4252, 2016.

[74] A. Noel, K. C. Cheung, and R. Schober, "Joint channel parameter estimation via diffusive molecular communication," *IEEE Transactions on Molecular, Biological and Multi-Scale Communications*, vol. 1, no. 1, pp. 4–17, 2015.

[75] H. Unterweger, J. Kirchner, W. Wicke, *et al.*, "Experimental molecular communication testbed based on magnetic nanoparticles in duct flow," in *2018 IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, IEEE, 2018, pp. 1–5.

[76] M. Pierobon and I. F. Akyildiz, "Capacity of a diffusion-based molecular communication system with channel memory and molecular noise," *IEEE Transactions on Information Theory*, vol. 59, no. 2, pp. 942–954, 2012.

[77] G. Aminian, M. Mirmohseni, M. N. Kenari, and F. Fekri, "On the capacity of level and type modulations in molecular communication with ligand receptors," in *2015 IEEE International Symposium on Information Theory (ISIT)*, IEEE, 2015, pp. 1951–1955.

[78] G. Aminian, H. Ghourchian, A. Gohari, M. Mirmohseni, and M. Nasiri-Kenari, "On the capacity of signal dependent noise channels," in *2017 Iran Workshop on Communication and Information Theory (IWCIT)*, IEEE, 2017, pp. 1–6.

[79] L. Lin, Q. Wu, F. Liu, and H. Yan, "Mutual information and maximum achievable rate for mobile molecular communication systems," *IEEE transactions on nanobioscience*, vol. 17, no. 4, pp. 507–517, 2018.

[80] Z. Luo, L. Lin, Q. Fu, and H. Yan, "An effective distance measurement method for molecular communication systems," in *2018 IEEE International Conference on Sensing, Communication and Networking (SECON Workshops)*, IEEE, 2018, pp. 1–4.

[81] H. Zhai, T. Nakano, A. V. Vasilakos, K. Yang, and Q. Liu, "Increase detection algorithm for concentration-encoded diffusion-based molecular communication," in *Proceedings of the 4th ACM International Conference on Nanoscale Computing and Communication*, 2017, pp. 1–6.

[82] B. Li, M. Sun, S. Wang, W. Guo, and C. Zhao, "Local convexity inspired low-complexity noncoherent signal detector for nanoscale molecular communications," *IEEE Transactions on Communications*, vol. 64, no. 5, pp. 2079–2091, 2016.

[83] D. Kilinc and O. B. Akan, "Receiver design for molecular communication," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 12, pp. 705–714, 2013.

[84] A. Noel, K. C. Cheung, and R. Schober, "Optimal receiver design for diffusive molecular communication with flow and additive noise," *IEEE transactions on nanobioscience*, vol. 13, no. 3, pp. 350–362, 2014.

[85] R. Mosayebi, H. Arjmandi, A. Gohari, M. Nasiri-Kenari, and U. Mitra, "Receivers for diffusion-based molecular communication: Exploiting memory and sampling rate," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 12, pp. 2368–2380, 2014.

[86] V. Jamali, N. Farsad, R. Schober, and A. Goldsmith, "Non-coherent multiple-symbol detection for diffusive molecular communications," in *Proceedings of the 3rd ACM International Conference on Nanoscale Computing and Communication*, 2016, pp. 1–7.

[87] A. Noel, K. C. Cheung, and R. Schober, "Improving receiver performance of diffusive molecular communication with enzymes," *IEEE Transactions on NanoBioscience*, vol. 13, no. 1, pp. 31–43, 2014.

[88] O. A. Dambri and S. Cherkaoui, "Enhancing signal strength and isi-avoidance of diffusion-based molecular communication," in *2018 14th International Wireless Communications & Mobile Computing Conference (IWCMC)*, IEEE, 2018, pp. 1–6.

[89] W. Wicke, A. Ahmadzadeh, V. Jamali, H. Unterweger, C. Alexiou, and R. Schober, "Magnetic nanoparticle-based molecular communication in microfluidic environments," *IEEE transactions on nanobioscience*, vol. 18, no. 2, pp. 156–169, 2019.

[90] W. Wicke, A. Ahmadzadeh, V. Jamali, R. Schober, H. Unterweger, and C. Alexiou, "Molecular communication using magnetic nanoparticles," in *2018 IEEE Wireless Communications and Networking Conference (WCNC)*, IEEE, 2018, pp. 1–6.

[91] B.-H. Koo, C. Lee, H. B. Yilmaz, N. Farsad, A. Eckford, and C.-B. Chae, "Molecular mimo: From theory to prototype," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 3, pp. 600–614, 2016.

[92] P. N. Prasanth, K. P. Sumanth, V. K. Chakka, and G. Roy, "Experimental implementation of molecular communication system using sampling based adaptive threshold variation demodulation algorithm," in *2018 IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS)*, IEEE, 2018, pp. 1–5.

[93] N.-R. Kim, N. Farsad, C. Lee, A. W. Eckford, and C.-B. Chae, "An experimentally validated channel model for molecular communication systems," *IEEE Access*, vol. 7, pp. 81 849–81 858, 2019.

[94] D. T. McGuiness, S. Giannoukos, A. Marshall, and S. Taylor, "Parameter analysis in macro-scale molecular communications using advection-diffusion," *IEEE Access*, vol. 6, pp. 46 706–46 717, 2018.

[95] S. Giannoukos, D. T. McGuiness, A. Marshall, J. Smith, and S. Taylor, "A chemical alphabet for macromolecular communications," *Analytical chemistry*, vol. 90, no. 12, pp. 7739–7746, 2018.

[96] M. Ozmen, E. Kennedy, J. Rose, P. Shakya, J. K. Rosenstein, and C. Rose, "High speed chemical vapor communication using photoionization detectors in turbulent flow," *IEEE Transactions on Molecular, Biological and Multi-Scale Communications*, vol. 4, no. 3, pp. 160–170, 2018.

[97] T. Furubayashi, Y. Sakatani, T. Nakano, A. Eckford, and N. Ichihashi, "Design and wet-laboratory implementation of reliable end-to-end molecular communication," *Wireless Networks*, vol. 24, pp. 1809–1819, 2018.

[98] N. Farsad, D. Pan, and A. Goldsmith, "A novel experimental platform for in-vessel multi-chemical molecular communications," in *GLOBECOM 2017-2017 IEEE Global Communications Conference*, IEEE, 2017, pp. 1–6.

[99] B. Krishnaswamy, Y. Jian, C. M. Austin, *et al.*, "Adma: Amplitude-division multiple access for bacterial communication networks," *IEEE Transactions on Molecular, Biological and Multi-Scale Communications*, vol. 3, no. 3, pp. 134–149, 2017.

[100] L. Wang and A. W. Eckford, "Nonnegative code division multiple access techniques in molecular communication," in *2017 15th Canadian Workshop on Information Theory (CWIT)*, IEEE, 2017, pp. 1–5.

[101] J. Suzuki, S. Balasubramaniam, and A. Prina-Mello, "Multiobjective tdma optimization for neuron-based molecular communication," in *7th International Conference on Body Area Networks*, 2012.

[102] L. P. Giné and I. F. Akyildiz, "Molecular communication options for long range nanonetworks," *Computer Networks*, vol. 53, no. 16, pp. 2753–2766, 2009.

[103] X. Chen, M. Wen, C.-B. Chae, L.-L. Yang, F. Ji, and K. K. Igorevich, "Resource allocation for multiuser molecular communication systems oriented to the internet of medical things," *IEEE Internet of Things Journal*, vol. 8, no. 21, pp. 15 939–15 952, 2021.

[104] K. Stephens, M. Pozo, C.-Y. Tsao, P. Hauk, and W. E. Bentley, "Bacterial co-culture with cell signaling translator and growth controller modules for autonomously regulated culture composition," *Nature communications*, vol. 10, no. 1, p. 4129, 2019.

[105] J. L. Terrell, T. Tschirhart, J. P. Jahnke, *et al.*, "Bioelectronic control of a microbial community using surface-assembled electrogenetic cells to route signals," *Nature Nanotechnology*, vol. 16, no. 6, pp. 688–697, 2021.

[106] Y. Zamiri-Jafarian, S. Gazor, and H. Zamiri-Jafarian, "Molecular code division multiple access in nano communication systems," in *2016 IEEE Wireless Communications and Networking Conference*, IEEE, 2016, pp. 1–6.

[107] G.-C. Yang and W. C. Kwong, "Performance comparison of multiwavelength cdma and wdma+ cdma for fiber-optic networks," *IEEE Transactions on Communications*, vol. 45, no. 11, pp. 1426–1434, 1997.

[108] N. Kahlon and G. Kaur, "Various dispersion compensation techniques for optical system: A survey," *Open journal of communications and software*, vol. 1, no. 1, pp. 64–73, 2014.

[109] F. R. Chung, J. A. Salehi, and V. K. Wei, "Optical orthogonal codes: Design, analysis and applications," *IEEE Transactions on Information theory*, vol. 35, no. 3, pp. 595–604, 1989.

[110] L. Grüner-Nielsen, S. N. Knudsen, B. Edvold, *et al.*, "Dispersion compensating fibers," *Optical fiber technology*, vol. 6, no. 2, pp. 164–180, 2000.

[111] S. Korte, M. Damrath, and P. A. Hoeher, "Multiple channel access techniques for diffusion-based molecular communications," in *SCC 2017; 11th International ITG Conference on Systems, Communications and Coding*, VDE, 2017, pp. 1–6.

[112] Y. Deng, A. Noel, W. Guo, A. Nallanathan, and M. Elkashlan, "Analyzing large-scale multiuser molecular communication via 3-d stochastic geometry," *IEEE Transactions on Molecular, Biological and Multi-Scale Communications*, vol. 3, no. 2, pp. 118–133, 2017.

[113] B. Atakan and O. B. Akan, "On molecular multiple-access, broadcast, and relay channels in nanonetworks," in *Proceedings of the 3rd International Conference on Bio-Inspired Models of Network, Information and Computing Sytems*, 2008, pp. 1–8.

[114] J. Wang, D. Hu, C. Shetty, and H. Hassanieh, "Understanding and embracing the complexities of the molecular communication channel in liquids," in *Proceedings of the 26th Annual International Conference on Mobile Computing and Networking*, 2020, pp. 1–15.

[115] W. Pan, X. Chen, X. Yang, N. Zhao, L. Meng, and F. H. Shah, "A molecular communication platform based on body area nanonetwork," *Nanomaterials*, vol. 12, no. 4, p. 722, 2022.

[116] M. Bayat, M. Mostafavi, and A. Arabameri, "Towards practical implementation of molecular communication: A cost-effective experimental platform based on the human circulatory system," *IET Communications*, 2024.

[117] M. Bartunik, J. Teller, G. Fischer, and J. Kirchner, "Channel parameter studies of a molecular communication testbed with biocompatible information carriers: Methods and data," *IEEE Transactions on Molecular, Biological and Multi-Scale Communications*, 2023.

[118] J. D. Kraus, M. Tiuri, A. V. Räisänen, and T. D. Carr, *Radio astronomy*. McGraw-Hill New York, 1966, vol. 66.

[119] A. R. Thompson, J. M. Moran, and G. W. Swenson, *Interferometry and synthesis in radio astronomy*. Springer Nature, 2017.

[120] F. Hlawatsch and G. Matz, *Wireless communications over rapidly time-varying channels*. Academic press, 2011.

[121] B. Li, S. Zhou, M. Stojanovic, L. Freitag, and P. Willett, *Non-uniform Doppler compensation for zero-padded OFDM over fast-varying underwater acoustic channels*. IEEE, 2007.

[122] Y. V. Zakharov and A. K. Morozov, "Ofdm transmission without guard interval in fast-varying underwater acoustic channels," *IEEE Journal of Oceanic Engineering*, vol. 40, no. 1, pp. 144–158, 2014.

[123] N. Farsad, H. B. Yilmaz, A. Eckford, C.-B. Chae, and W. Guo, "A comprehensive survey of recent advancements in molecular communication," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 3, pp. 1887–1919, 2016.

[124] L. Felicetti, M. Femminella, and G. Reali, "Congestion control in molecular cyber-physical systems," *IEEE Access*, vol. 5, pp. 10 000–10 011, 2017.

[125] L. Felicetti, M. Femminella, G. Reali, T. Nakano, and A. V. Vasilakos, "Tcp-like molecular communications," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 12, pp. 2354–2367, 2014.

[126] L. Felicetti, M. Femminella, and G. Reali, "Smart antennas for diffusion-based molecular communications," in *Proceedings of the Second Annual International Conference on Nanoscale Computing and Communication*, 2015, pp. 1–6.

[127] V. Jamali, A. Ahmadzadeh, W. Wicke, A. Noel, and R. Schober, "Channel modeling for diffusive molecular communication—a tutorial review," *Proceedings of the IEEE*, vol. 107, no. 7, pp. 1256–1301, 2019.

[128] N. Farsad and A. Goldsmith, "Neural network detection of data sequences in communication systems," *IEEE Transactions on Signal Processing*, vol. 66, no. 21, pp. 5663–5678, 2018.

[129] W. Guo, T. Asyhari, N. Farsad, *et al.*, "Molecular communications: Channel model and physical layer techniques," *IEEE Wireless Communications*, vol. 23, no. 4, pp. 120–127, 2016.

[130] A. Ahmadzadeh, V. Jamali, and R. Schober, "Stochastic channel modeling for diffusive mobile molecular communication systems," *IEEE Transactions on Communications*, vol. 66, no. 12, pp. 6205–6220, 2018.

[131] N. Abadi, A. A. Gohari, M. Mirmohseni, and M. Nasiri-Kenari, "Zero-error codes for multi-type molecular communication in random delay channel," in *2018 Iran Workshop on Communication and Information Theory (IWCIT)*, IEEE, 2018, pp. 1–6.

[132] H. Arjmandi, A. Gohari, M. N. Kenari, and F. Bateni, "Diffusion-based nanonetworking: A new modulation technique and performance analysis," *IEEE Communications Letters*, vol. 17, no. 4, pp. 645–648, 2013.

[133] Y. Deng, A. Noel, W. Guo, A. Nallanathan, and M. Elkashlan, "3d stochastic geometry model for large-scale molecular communication systems," in *2016 IEEE Global Communications Conference (GLOBECOM)*, IEEE, 2016, pp. 1–6.

[134] I. Llatser, E. Alarcón, and M. Pierobony, "Diffusion-based channel characterization in molecular nanonetworks," in *2011 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, IEEE, 2011, pp. 467–472.

[135] A. Stok and E. H. Sargent, "The role of optical cdma in access networks," *IEEE Communications Magazine*, vol. 40, no. 9, pp. 83–87, 2002.

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

[137] A. Goldsmith, *Wireless communications*. Cambridge university press, 2005.

[138] J. K. Holmes, *Spread spectrum systems for GNSS and wireless communications*. Artech House Norwood, 2007.

[139] R. Gold, "Optimal binary sequences for spread spectrum multiplexing (corresp.)," *IEEE Transactions on information theory*, vol. 13, no. 4, pp. 619–621, 1967.

[140] A. Viterbi, "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm," *IEEE transactions on Information Theory*, vol. 13, no. 2, pp. 260–269, 1967.

[141] W. Chu and C. J. Colbourn, "Optimal (n, 4, 2)-ooc of small orders," *Discrete Mathematics*, vol. 279, no. 1-3, pp. 163–172, 2004.

[142] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, P. Walter, *et al.*, "Molecular biology of the cell," *Scandinavian Journal of Rheumatology*, vol. 32, no. 2, pp. 125–125, 2003.

[143] B.-H. Koo, H. J. Kim, J.-Y. Kwon, and C.-B. Chae, "Deep learning-based human implantable nano molecular communications," in *ICC 2020-2020 IEEE International Conference on Communications (ICC)*, IEEE, 2020, pp. 1–7.

[144] S. Lotter, L. Brand, V. Jamali, *et al.*, "Experimental research in synthetic molecular communications-part ii," *IEEE Nanotechnology Magazine*, 2023.

[145] S. Lotter, L. Brand, V. Jamali, *et al.*, "Experimental research in synthetic molecular communications–part i: Overview and short-range systems," *arXiv preprint arXiv:2301.06417*, 2023.

[146] F. Wang, P. Jin, Y. Feng, *et al.*, "Flexible doppler ultrasound device for the monitoring of blood flow velocity," *Science advances*, vol. 7, no. 44, eabi9283, 2021.

[147] J. Wang, S. Öğüt, H. Al Hassanieh, and B. Krishnaswamy, "Towards practical and scalable molecular networks," in *Proceedings of the ACM SIGCOMM 2023 Conference*, 2023, pp. 62–76.

[148] G. Matz and F. Hlawatsch, "Time-varying communication channels: Fundamentals, recent developments, and open problems," in *2006 14th European Signal Processing Conference*, IEEE, 2006, pp. 1–5.

[149] M. U. Mahfuz, D. Makrakis, and H. T. Mouftah, "On the characteristics of concentration-encoded multi-level amplitude modulated unicast molecular communication," in *2011 24th Canadian conference on electrical and computer engineering (CCECE)*, IEEE, 2011, pp. 000 312–000 316.

[150] [Online]. Available: https://www.drifton.eu/shop/9-flow-rate-peristaltic-pumps/1570-labv6-flow-rate-tube-pump/.

[151] [Online]. Available: https://www.amazon.com/Gikfun-Peristaltic-Connector-Aquarium-Analytic/dp/B01IUVHB8E?ref_=ast_slp_dp.

[152] [Online]. Available: https://store-usa.arduino.cc/products/gravity-analog-electrical-conductivity-sensor-meter-for-arduino.

[153] [Online]. Available: https://ch.farnell.com/sensirion/sek-slf3s-4000b/evaluationskit-durchflusssensor/dp/3977736?st=slf3s-4000b.

[154] [Online]. Available: https://store.arduino.cc/products/arduino-mega-2560-rev3.

[155] K. M. Saqr, S. Tupin, S. Rashad, *et al.*, "Physiologic blood flow is turbulent," *Scientific reports*, vol. 10, no. 1, p. 15 492, 2020.

[156] D. Dürrenmatt, D. Del Giudice, and J. Rieckermann, "Dynamic time warping improves sewer flow monitoring," *Water research*, vol. 47, no. 11, pp. 3803–3816, 2013.