NONNEGATIVE CODE DIVISION MULTIPLE ACCESS TECHNIQUES IN MOLECULAR COMMUNICATION

LINCHEN WANG

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

MASTER OF APPLIED SCIENCE

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

Abstract

In molecular communication, two types of multiple access have been studied: time division and molecule division. In this work, we consider code division multiple access. However, unlike code division multiple access that has been used for electromagnetic signals, we investigate optical code division multiple access: since molecular signals have the same non-negativity feature as optical signals, this scheme is a promising solution for molecular communication.

In this thesis, we perform experiments and set up simulation models which match these experiments. Moreover, using simulations, we find the features of optical code division multiple access for molecular communication. Our results include an optimal information transmission scheme, and an algorithm to decode molecular information signals. Finally, we demonstrate reliable communication with multiple access by using this scheme.

Acknowledgements

This thesis could not be completed without help, support, and encouragement from many people.

I wish to express my gratitude to my supervisor, Professor Andrew Eckford. You are a knowledgeable professor, I always be inspired from our meeting. You are a patient professor, you taught me how to find out and solve the problems step by step. And you are an amiable professor, you always treat me like a father. Thank you, for everything.

To my committee members, Professor Sebastian Magierowski and Professor John Moores, thank you to read my thesis, attend my oral examination, and give me valuable comments.

To my department member, graduate program assistant Stefanie Caputo, thank you for your encouragement before my oral examination.

Finally, I want to thank my fellow graduate student, Nariman Farsad, you always encourage me, and give me your advices. You are like my big brother.

Table of Contents

List of Tables

List of Figures

1 Introduction

1.1 Motivation

Molecular communication is a new field of science which uses chemical signals to propagate information. Using molecular communication, molecules are used as the message media. These molecules propagate from transmitter to receiver using methods such as diffusion without flow [1–3], diffusion with flow [4–6], and molecular motors [7–9].

The primary motivation of molecular communication is to communicate at the microscale or nanoscale. For example, cells are a kind of microscale device, which can exchange information by molecular transmission. This process is called intercellular signal transduction, which is investigated and used in modern biotechnology.

Aside from the microscale and nanoscale, molecular communication was also demonstrated at the macroscale. For example, a recent paper [10] used alcohol vapour to transmit short text messages by encoding those messages into alcohol concentration in the air. The authors established a communication link by vaporiz-

ing alcohol at a sprayer and adjusting the concentration based on the transmitted sequence, and measuring the alcohol concentration at a distant sensor. Thus, it was shown that molecular communication can be used to send text messages. This idea has useful applications in areas where electromagnetic communication can't be used, such as underground or underwater. For instance, in the case of collapsed building, traditional radio communication would be restricted while attempting to transmit radio signal through concrete and steel bars. However, molecular communication is feasible in transmitting information because diffusion works even when electromagnetic propagation does not. It has been shown that molecular communication has better performance than traditional radio communication in some applications [11]. In this thesis, we show that optical code division multiple access is a useful technique in macroscale molecular communication. Moreover, since the signals in both microscale and macroscale molecular communication propagate using diffusion, we believe our results are also applicable to microscale molecular communication.

Molecular communication also has features which are not available in electromagnetic communication. For example, one feature of this method is that unlike a radio signal which is not persistent, a chemical tag can stay on a surface for a period of time. To show this, paper [12] employed a mobile platform to read chemical bar codes which were left ahead of it on a surface. The mobile platform would collect

the bar codes it read, and process them into a binary sequence.

Multiple access is a channel access method which allows multiple users to share an allotted medium to transmit over it. Multiple access is important because users share a limited medium, and the key objective of multiple access is to share the limited medium in a way that allows every user to reliably transmit information. Thus, we need multiple access techniques in molecular communication. We will review these techniques in this chapter.

Our experiment is based on the work in [10], and we want to improve this system from single pair transmitting to multiple pairs transmitting. Therefore, the main problem is interference between different pairs since they share the same communication medium, i.e. air. And alcohol molecules would be detected by both sensors during transmission process. That may cause missed detection (i.e. detect as "0" but truth is "1") and false alarm (i.e. detect as "1" but truth is "0"). Thus, we should consider multiple access techniques in molecular communication.

In this thesis, we study optical code division multiple access techniques in molecular communication, which is the first time such a scheme is proposed. The remainder of the thesis is organized as follows. We briefly review some work in terms of multiple access for molecular communication in the rest of this chapter. Chapter II discusses why we chose Optical Code Division Multiple Access (OCDMA), and how OCDMA works. Chapter III describes the experiment set up and simulation

model, and Chapter IV evaluates the performance of the simulation system and bit detection. In Chapter V, we adapt the results of the previous experiments to test our OCDMA algorithm. A summary of this thesis is provided in Chapter VI.

1.2   How does molecular communication work?
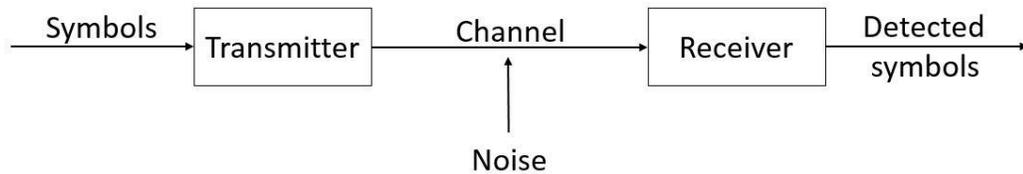
Figure 1.1: Molecular communication model

In molecular communication, chemical symbols have been transmitted by transmitter though channel such as liquid or air. There are noises in the channel, and receiver would detect symbols. Here we introduce how to modulate a binary sequence into molecular media, in order to perform molecular communication.

There are various types of modulation in molecular communication that have been published recent years. Based on [13] and [14], there can be considered to

be three main modulation techniques. The first technique is Concentration Shift Keying (CSK), which uses one type of molecule to communicate, and the receiver detects information by evaluating the concentration of that molecule. The second technique is Pulse Position Modulation (PPM) which also uses only one type of molecule, but it encodes and transmits messages in a specific time shift. The third technique is Molecule Shift Keying (MoSK), which uses different types of molecules to represent each different modulation symbol.

1.2.1 Concentration Shift Keying (CSK)

Papers [15] and [16] consider the use of CSK. In [15], using binary modulation, each transmitter releases one molecule when sending "1", and zero molecules when sending "0" (like Fig 1.2). All the molecules propagate by Brownian motion, and the receiver determines which bit was sent by counting the number of molecules it gathers. If it receive zero molecules, it can conclude a "0" was sent; however, if it gathers more than one molecule, a sequence of bits was sent.

Paper [16] considers a strength-based signal detection model called Concentration-Encoded Molecular Communication (CEMC). Here, simulations were performed in terms of diffusion-based propagation for one type molecule, which was sent from a transmitting nanomachine to a receiving nanomachine. One feature of this work is that it considers the effect of intersymbol interference (ISI). ISI is caused by resid-
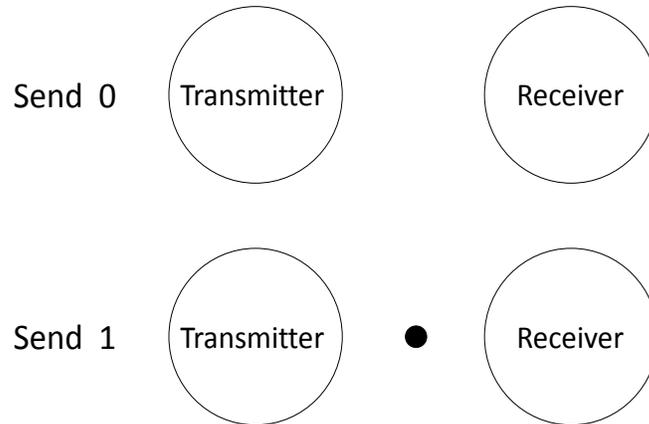
Figure 1.2: Send zero molecule means "0", whereas send one molecule means "1"

ual molecules from previous transmissions, and those residual molecules affect the detection of current or future bits. The problem of ISI is important in molecular communication, and we would face this challenge in our own experiments; thus, this paper can help us to understand CEMC. Based on the performance of the simulation, the bit error rate could become very low if the author chose a short communication range or a low propagation rate, because these reduce the ISI.

1.2.2 Pulse Position Modulation

In pulse position modulation, time is divided into frames, and the transmitter would convey its molecules in specific time frame to represent "1" or "0". Fig 1.3 gives us an example. Using binary modulation, say the transmitter conveys a molecule at

time 0 to send "0", and conveys a molecule at time $t > 0$ that means sending "1". Thus, the receiver can distinguish the symbol by measuring the arrival time of the molecule.
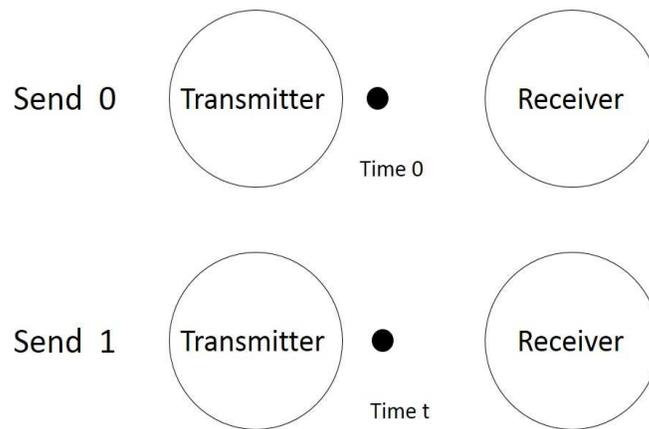
Figure 1.3: "0" by send at time 0, whereas sending at time t means "1"

For example, [17] considered particles propagating with Brownian motion, where the authors encoded information in time, and the result was that a capacity was achieved of more than one bit per particle.

1.2.3 Molecule Shift Keying

Paper [18] considers communication by using different types of molecules. Based on the paper, if we propose to transit x bits of information at a time, we would need $M = 2^x$ types of molecules. There are approximately 38,000 specific trisaccharides

[18] if carbohydrates are chosen as information molecules.

For example, using binary modulation, in Fig 1.4, the transmitter sends a molecule (represented by a circle) meaning "0", and sends a different molecule (represented by a square) meaning "1".
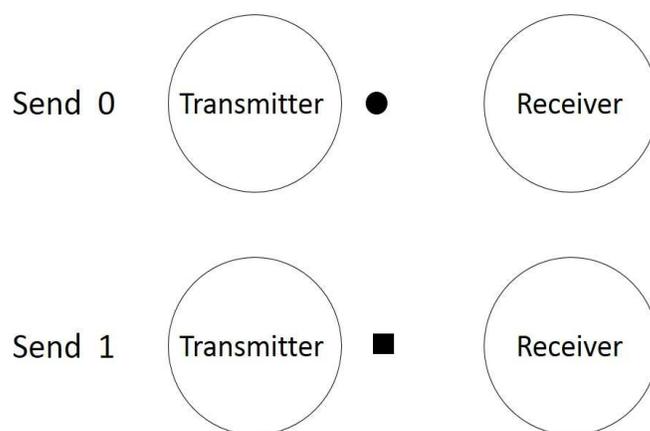
Figure 1.4: A "0" is sent by a "circle" molecule, whereas a "1" is sent by a "square" molecule.

1.3   Multiple Access In Molecular Communication

There are few existing multiple access techniques that have been used in molecular communication. These include time division multiple access (TDMA) and molecule division multiple access (MDMA). We describe these briefly below.

1.3.1 TDMA

Paper [19] presents a TDMA based neural network transmission from some sources to an unique receiver with sharing a common channel. The construction is shown in Fig 1.5.
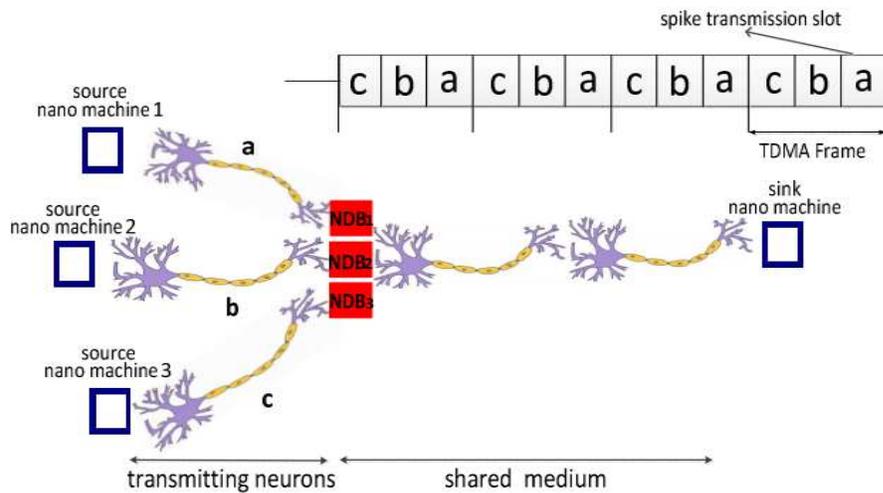
Figure 1.5: TDMA based neural network [19]

In this figure, there are three source nano-machines which have their transmitting neurons a, b, and c. Time is divided into frames, and each frame has three spike transmission slots. The sources would convey the order in which they transmit (e.g. c, b, a). There are Neural Delay Boxes (NDBs) connecting between transmitting neurons and the shared medium. The NDBs act as a buffer to store the information temporally while other sources are using the shared medium. It is important

that performance with NDBs is better than without delay lines, especially when the number of sources increases. There is another paper [21] considering a group of bio-nanomachines which multiplex their transmission using TDMA to prevent interference among different sources. Paper [22] also uses a TDMA transmission scheme with a genetic algorithm to simulate two different sizes of neural network.

1.3.2 MDMA

Paper [20] used different types of molecules as different symbols. This paper uses pheromone diversity to achieve multiple access among nano-machines. Each nano-machine is equipped with a pheromone receptor that can detect a specific type of pheromone. Based on that, the channels are separated, and the author called it molecular division multiple access (MDMA).

All the above-mentioned works are attractive in terms of both modulation or multiple access techniques in molecular communication. According to our previous experimental experience [10, 12], concentration shift keying may be suitable for our current work. Moreover, we would use the same shared medium, that is, air. The concentration of molecules would change if we spray streams of molecules into the air. Moreover, receivers can determine either "1" or "0" by measuring the alcohol concentration around them like in [10, 12]. We will introduce our system in detail in chapter III.

1.4 CDMA and OCDMA in molecular communication

In terms of multiple access, CDMA is widely used in radio communication, but there is not much work on CDMA in molecular communication. However, we think CDMA, especially Optical Code Division Multiple Access (OCDMA) may help us to achieve multiple access in our work.

In CDMA, each user is assigned an unique binary sequence as their identifying code, and this code represents that user's own "1" (in binary modulation). By using CDMA, each transmitter can communicate with their paired receiver reliably. We propose to employ CDMA in our work. However, the conventional code division multiple access techniques might not work well in molecular communication, since molecular communication has only non-negative signals (i.e. ≥ 0). Since there is no negative signal in molecular communication, we can't get a sum of zero or nearly zero when we add the identifying code of each transmitter together; that is, the interference between users can't be cancelled. That means we can't achieve quasi-orthogonal codes, recalling that CDMA is not strictly orthogonal. Therefore, conventional CDMA is not applicable to molecular communication.

This challenge has been addressed in optical communication, which has a similar non-negative constraint. Optical signals are noncoherent and can only illuminate or extinguish the light source, which may only produce a nonnegative signal. This lim-

itation is exactly same with a molecular signal. Therefore, we propose to adapt the research of OCDMA to help us achieve multiple access, and improve performance in molecular communication.

In OCDMA, each user has a unique signature sequence by sending short optical pulses in several chip intervals. Moreover, in binary modulation, each encoder uses its own signature sequence to represent "1", and the all-zero (blank) sequence to represent "0". These signature sequences might not be strictly orthogonal; however, they could be quasi-orthogonal, as we will explain in the next chapter. At the transmitter, the data would be converted into a spread spectrum signal, representing one user's signature sequence; it would then be converted to a light signal at the optical encoder (e.g. a light-emitting diode). This optical signal would be transmitted through an optical star coupler to every optical decoder (e.g. photo diode) like in Fig 1.6. The data is then extracted using code selection logic.

1.5   Contributions

The main idea of this thesis is to achieve multiple access in molecular communication using OCDMA. The original contributions of this thesis are the following:

- We set up our simulation models. We show that they match the results from our experimental apparatus used in [10, 12]. Based on these simulations, we
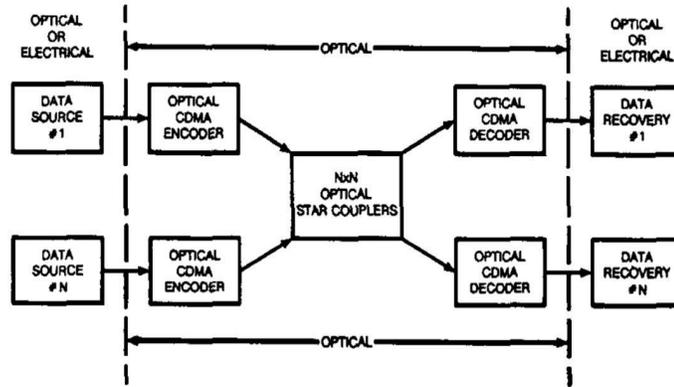
Figure 1.6: A star configuration of Optical Code Division Multiple-Access [23]

can get our results faster and easier than using experiments.

- We set up simulations involving: one transmitter and one receiver; and two transmitters to two receivers. In both cases, we find the best transmission distance range and chip time to achieve optimal performance.

- We simulate the OCDMA scheme using two transmitters and two receivers. We find the performance is related not only to the distance and chip time, but also to the chip sequence length. We give optimal values for these parameters.

2   OCDMA

Molecular signals have the same limitations as optical signals, since there is no negative signal. Optical CDMA (OCDMA) has been introduced as a multiple-access solution in optical communications. By applying OCDMA to molecular communication, we believe we can achieve good performance in molecular communication, similar to optical communication. This is the primary motivation for our work.

2.1   What is Optical Code Division Multiple Access

In paper [23], the role of OCDMA in access networks was investigated. Researchers tried to combine the large bandwidth of the fiber medium with the flexibility of CDMA in terms of decentralized control. Therefore, they used the excess bandwidth to achieve random asynchronous communication in a fiber medium.

The two main problems in OCDMA are, firstly, interference from pairs of transmitting users; and secondly, the non-negativity of the channel. Therefore, it is important to design an optimal signature sequence that works with a non-negative

channel. The properties of a good optimal signature sequence are that each sequence can easily be distinguished from its own shifted version sequence, and any version of every other sequence. The solution to this problem is to use optical orthogonal codes (OOCs) in OCDMA – note that OOCs are actually quasi-orthogonal, as is typical in CDMA. We will introduce OOCs in this section. Another solution is to spread the OOCs in both the temporal and wavelength domain at the same time, by using this approach, the time chips and wavelength channels can be viewed as the axes of a two dimensional codeword.

OCDMA is a known technology for application in multiple access networks. The users of OCDMA could be provided a fair division to share the optical bandwidth. Furthermore, OCDMA is a flexible system, since two-dimensional OCDMA codes can use both the time and wavelength domain as we mentioned before. Moreover, it is easy to control and manage the network in OCDMA. For instance, any additional user would only need a new OOC different from any existing OOC.

Paper [24] gives us the fundamental principles of OCDMA. In its communication system, optical signals are transmitted from OCDMA encoder to OCDMA decoder; however, in our molecular communication system, we will use appropriate chemical hardware. The air is a transmission medium in our experiment whereas optical star couplers are used in optical communication systems (a star coupler is a device that spilts an input signal to several outputs [23]).

15

Paper [24] designed a new class of signature sequences which is called "optical orthogonal codes" for optical signal processing. The codes should follow the rules below:

1. Each sequence can easily be distinguished from its shifted version.

2. Each sequence can easily be distinguished from any version of any other sequence.

These rules are the key of OOCs. Asynchronous transmission is achieved since any shift of any sequence can be distinguished. The asynchronous transmission scheme allows any user to join the channel and transmit their sequence at any time with minimal interference. Furthermore, every user has their own unique identifying code.

The following two equations satisfy the above two conditions:

1. Let x be a periodic signature sequence with period F. For one period of the sequence $x = [x_0, x_1, \ldots, x_{F-1}]$

$$|Z_{x,x}(l)| = \left| \sum_{n=0}^{F-1} x_n x_{n+l} \right| = \begin{cases} K & l = 0 \\ \\ \leq \lambda_a & 1 \leq l \leq F - 1 \end{cases} \tag{2.1}$$

2. For any pair of periodic signature sequences $x = [x_0, x_1, \ldots, x_{F-1}]$ and $y =$

$$[y_0, y_1, \ldots, y_{F-1}],$$

$$|Z_{x,y}(l)| = \left| \sum_{n=0}^{F-1} x_n y_{n+l} \right| \leq \lambda_c \qquad\qquad 0 \leq l \leq F - 1 \qquad (2.2)$$

The results here are the sum of any two sequences following previous conditions: l is any integer value representing time shift, F is the period of the signature sequence, and K represents the number of "1" in the sequence (i.e., the Hamming weight). Furthermore λ_a is autocorrelation constraint, and λ_c is the cross-correlation constraint. In the optimal situation, we can get $\lambda_a = \lambda_c = 1$, and such codes are called quasi-orthogonal. All OOCs are quasi-orthogonal.
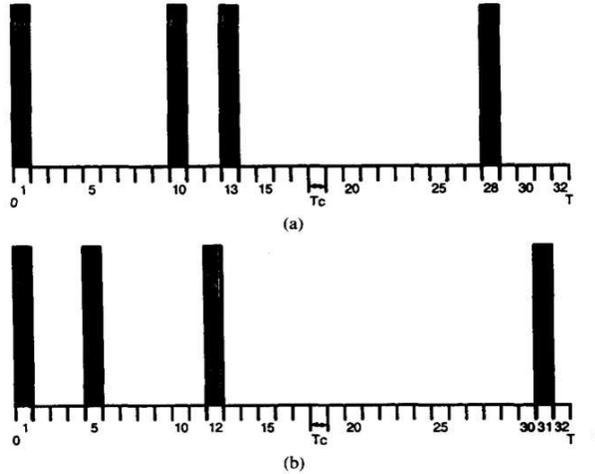
Figure 2.1: Two optical orthogonal codes [24]

In Fig 2.1, two codes are shown which achieve the requirements of OOCs. In these two examples, T is the time of the entire signature sequence period for one bit of information; and T_c, the time of one symbol in the OOC, is called the chip

17

time. Given the period F, $T = FT_c$. For example, in the figure,

$$F = T/T_c = 32 \tag{2.3}$$

Moreover, the number of "1" chips in each sequence is K, and in this example, $K = 4$. The two sets below are the sets of distances between "1" chips in Fig 2.1. The set for (a) is

$$A = \{9, 3, 15, 5\} \tag{2.4}$$

and the set for (b) is

$$B = \{4, 7, 19, 2\}. \tag{2.5}$$

We can extend these sets by considering the set of distances between adjacent *and non-adjacent* "1" chips. For sequence (a), this is

$$A_{EXT} = \{9, 3, 15, 5, 12, 18, 20, 14, 27, 23, 29, 17\} \tag{2.6}$$

and the extended set for sequence (b) is

$$B_{EXT} = \{4, 7, 19, 2, 11, 26, 21, 6, 30, 28, 25, 13\} \tag{2.7}$$

For the extended sets A_{EXT} and B_{EXT}, there are no two elements that are equal. This satisfies the autocorrelation property which is condition (2.1) with $\lambda_a = 1$. Furthermore,

$$A_{EXT} \cap B_{EXT} = \emptyset \tag{2.8}$$

18

i.e. there is no intersection of the extended set of A and B, which satisfies the cross-correlation condition (2.2) with $\lambda_c = 1$. That is, such sequences are OOCs.

In paper [25], the author utilized OOCs operating in fiber-optical-code division a multiple-access communications system to get the probability of error per bit (P_e). Based on their performance analysis, the actual error rate is in the range of two extreme cases: chip synchronous interference (Case A) and ideal chip asynchronous interference (Case B). Their relationship as below:

$$P_e(\text{Case B}) \leq P_e(\text{exact}) \leq P_e(\text{Case A}) \tag{2.9}$$

Ideal chip asynchronous interference is the best case, whereas chip synchronous interference may cause the worst performance. Moreover, length (F), weight (K), number of users (N), and other receiver parameters also affect P_e. In terms of the length F, the system has better performance as F increases. In terms of weight K, when we keep the weight K fixed, P_e would decrease in an optimal receiver. Furthermore, the system would have good performance with a small number of users N. In [24] the following relationship is obtained:

$$N \leq \left\lfloor \frac{F-1}{K(K-1)} \right\rfloor \tag{2.10}$$

Here, the symbol $\lfloor x \rfloor$ means to get the integer portion of the value x. Moreover, the author also mentioned using hard-limiter at the front end of optical correlator

in paper [25]. An ideal optical hard-limiter is defined as

$$g(x) = \begin{cases} 1 & x \geq 1 \\ 0 & 0 \leq x < 1 \end{cases} \qquad (2.11)$$

which would reduce the effect of the interference.

Paper [26] analyses the behaviors and characteristics of OCDMA base on OOCs. That paper considers generalized OOCs (i.e. the cross-correlation constraint of optical signature sequences would not bigger than weight K). The main results compared the values of cross-correlation (λ_c). If our goal is to accommodate the maximum number of interfering users, we can achieve the optimal operation by setting $\lambda_c = 2, 3$; however, if our main concern is minimum error rate, $\lambda_c = 1$ could help us to get best performance. For our experiment, providing our main purpose is the minimum error rate, we will set $\lambda_c = 1$.

2.2 An example of OCDMA

In terms of applying OCDMA, say we have two users A and B, and we assign them two OOCs a and b (Fig 2.2). In these codes, length (F) is 15, weight (K) is 3, and number of users (N) is 2. Thus, these codes conform the equation (2.10). A's code is

$$a = [1, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0] \qquad (2.12)$$

and B's code is

$$b = [1, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0]. \tag{2.13}$$
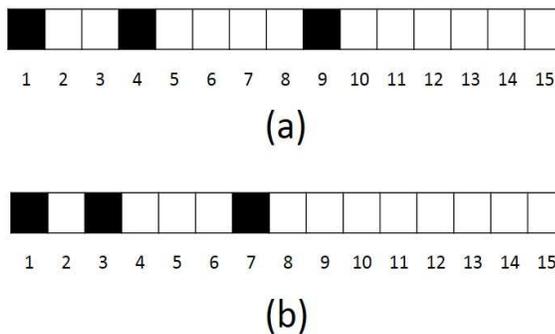
Figure 2.2: We sample two codes with K=3 and F=15

Say the time chip (T_c) is 1 second. Based on equation (2.3),

$$T = F/T_c = 15 \tag{2.14}$$

so that the entire sequence period T is 15 seconds. Thus, user A sends his signature sequence by sending a "1" pulse at the time chips of 1, 4, and 9 seconds, and keeps silent (sending "0") at rest of the time chips. User B sends his signature sequence by sending a "1" pulse at 1, 3, and 7 seconds, and keeps silent at rest of the time chips.

Suppose the users send binary data: a "1" bit is transmitted using the signature sequence, while a "0" bit is transmitted using an all-zero sequence (i.e. silence).

To demonstrate detection of OCDMA, say users A and B have their corresponding receivers A_r and B_r, respectively. Both receivers listen to the medium. Once A_r receives a "1" pulse (this "1" pulse could from anyone at any time), it will turn to the next stage of listening. In this next stage, if A_r receives a "1" pulse at both the 4th and 9th time chips (counting from the first "1" pulse it received), we say that A_r received its signature sequence. Similarly, at B_r, it will turn to the second stage once it receives any "1" (this "1" pulse could also from anyone at any time). The receiver B_r can decide that it has received its signature sequence if it receives a "1" pulse at both 3rd and 7th time chip (counting from its first "1" pulse).

OCDMA in chip synchronous transmission is shown in Fig 2.3. Say both A and B are transmitting "1" bits at the same time, and A is sending his signature sequence continuously (i.e., all "1" bits). B has a transmission sequence in one of fifteen possible time shifts (from b_1 to b_{15}, as in the figure). There is no more than one "1" chip overlap between A and B, no matter which time shift B used. That means these two sequences would have minimal interference on each other's detector.
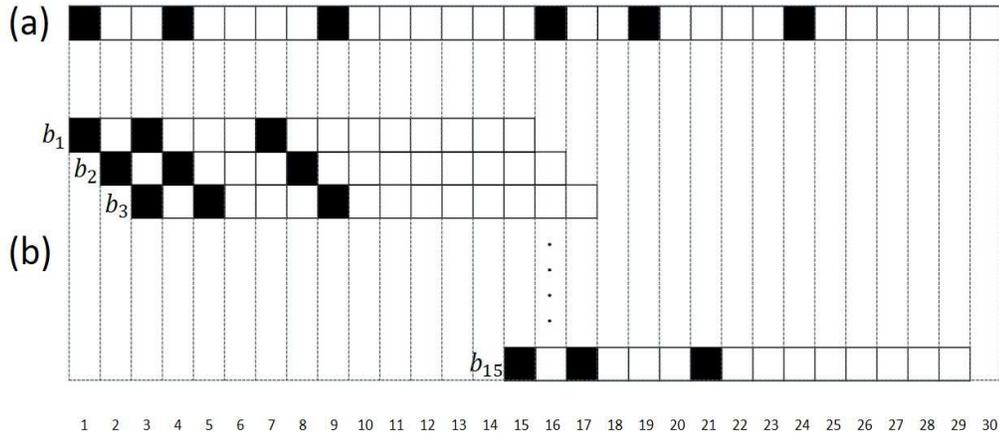
Figure 2.3: OCDMA in ideal chip synchronous transmission

2.3 A visualization of OCDMA

An intuitive way to visualize OCDMA is to consider it like playing a children's "fortune-telling" game called Gypsy card (Fig 2.4). Players choose their own mask card (the one on the left in Fig 2.4), and they pick an information card randomly (the one on the right in Fig 2.4). By covering the information card with their own mask card, they can get their unique "fortune".

In multiple access, users share a communication channel, therefore, they get unordered information (this process is similar to all players choosing the same information card in the game). For example, Fig 2.5 is an unordered message that
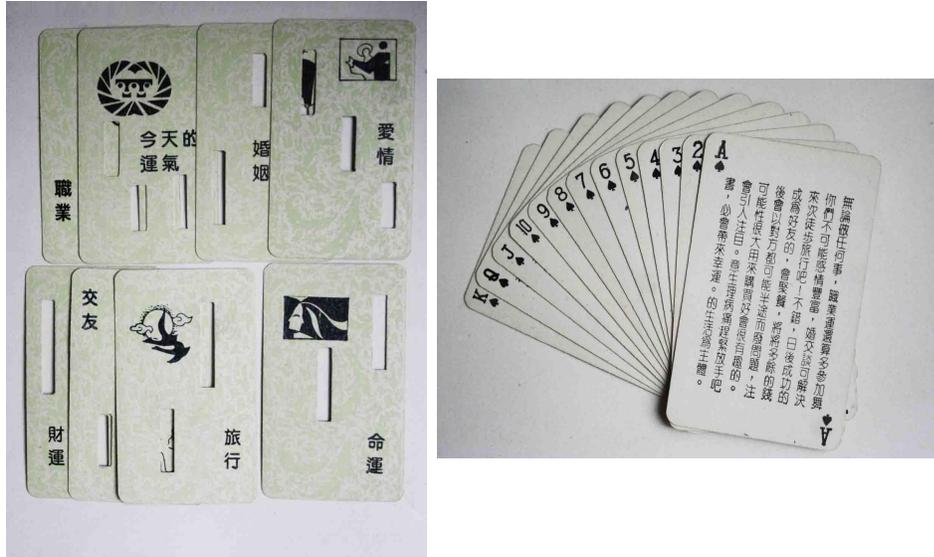
Figure 2.4: One kind of Gypsy card divination [27]
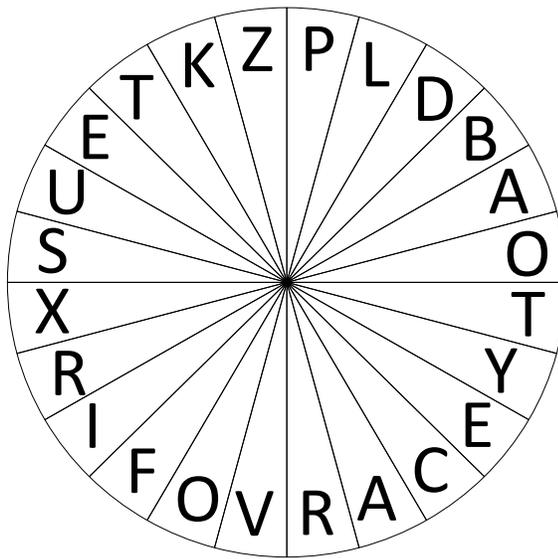
Figure 2.5: Unordered message

all players get. Moreover, we assign two mask cards a and b, one to player A, and one to player B.

Player A obtains his information as "LOVE" (Fig 2.6) once he covers his mask card a on the information card (Fig 2.5); furthermore, player B receives his information as "YORK" (Fig 2.7) by covering his mask card b on the information card (Fig 2.5) as well.
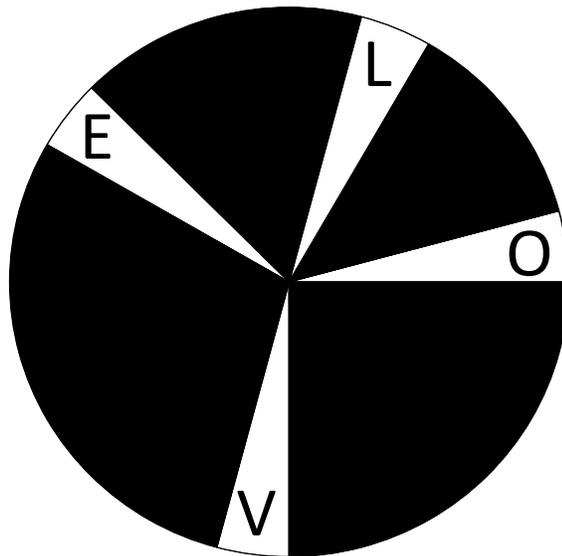
Figure 2.6: Place player A's mask card a on the unordered message, gets "LOVE"

Translating this example to OCDMA, we can say that player A receives a "1" bit once he receives his signature sequence "LOVE", otherwise, he receives a "0" bit. It is similar with player B, say data "1" can be obtained if he receives his signature sequence "YORK", and data "0" is received otherwise.
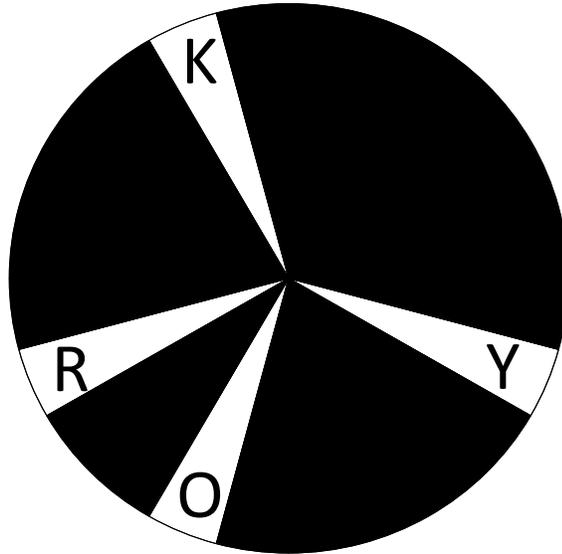
Figure 2.7: Place player B's mask card b on the unordered message, gets "YORK"

In addition, there may be a shifted mask card with some players, and the information under that mask card might get lost; this emphasizes the importance of synchronization. We say player C gets the information with his shifted mask card c shown below in Fig 2.8:

The information for player C should be "BEST", and he can get three different signature sequences "BEST", "ACOK", and "ACUK" from his shifted mask card. However, we can still estimate that this word is "BEST" since this is the only word in the list that is a valid signature sequence (there are also player A's "LOVE", player B's "YORK", and players C's "BEST" in the signature sequence list), while the other two sequences are meaningless. Moreover, even though we totally lose
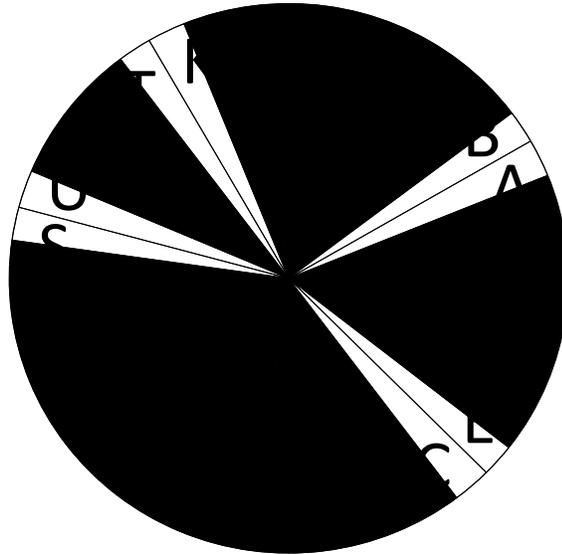
Figure 2.8: Place player C's mask card on the unordered message, and its information should be "BEST"

this signature sequence with the mask card shifted to next four letters, we can still get player C's signature sequence by continuing to rotate the mask card clockwise until we see the signature sequence.

In OCDMA, there is one signature sequence for each player, and there is only one overlap block with every mask card (i.e. $\lambda_c = 1$, and we do not consider this restriction in this visualization section). OCDMA works in that case because the mask card with this overlap block will be meaningless for other players. For example, we shift player B's mask card b clockwise by $45°$, and let one block of this mask card overlap one block of player A's information "L" as shown in Fig 2.9
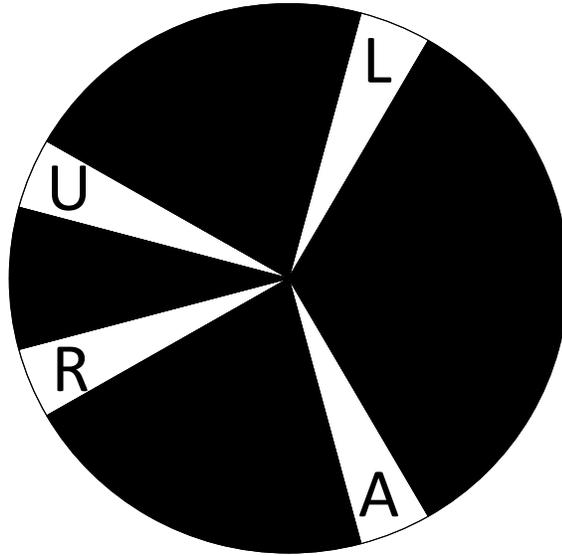
Figure 2.9: We shift player B clockwise with 45°, and overlap one block with player A's information

We get the sequence "LARU", and it is not a valid signature sequence for any player A, B, or C. This meaningless sequence will not affect our detection, and we can keep turning this player B's mask card clockwise until it sees its signature sequence "YORK".

3 Experiment set up and simulation model

We do both simulations and experiments in our work. Experiments help us to estimate whether our scheme could be achieved in practice; however, simulations give us our scheme's performance faster and easier. In this chapter, we introduce our experimental set up and simulation model; furthermore, we will explain how our simulation models match our experiments.

3.1 Experimental apparatus

Our experiment design is related to earlier work [10]. However, in this thesis, our goal is to transmit information using OCDMA. Thus, our experiment has three key components: *the transmitter*, which encodes the information into an OCDMA code and broadcasts that information to the medium; *the receiver*, which senses the medium and decodes the transmission using their assigned signature code; and *the testing environment*, which provides the medium for our system to achieve multiple access in molecular communication. In the remainder of this section, we describe

all three components in detail.

3.1.1   The transmitter

The transmitter includes a microcontroller, two sprayers, and a fan, as shown in Fig 3.1.

Dyson AM01 fan

Electronic sprayers

Arduino microcontroller & Adafruit LCD shield
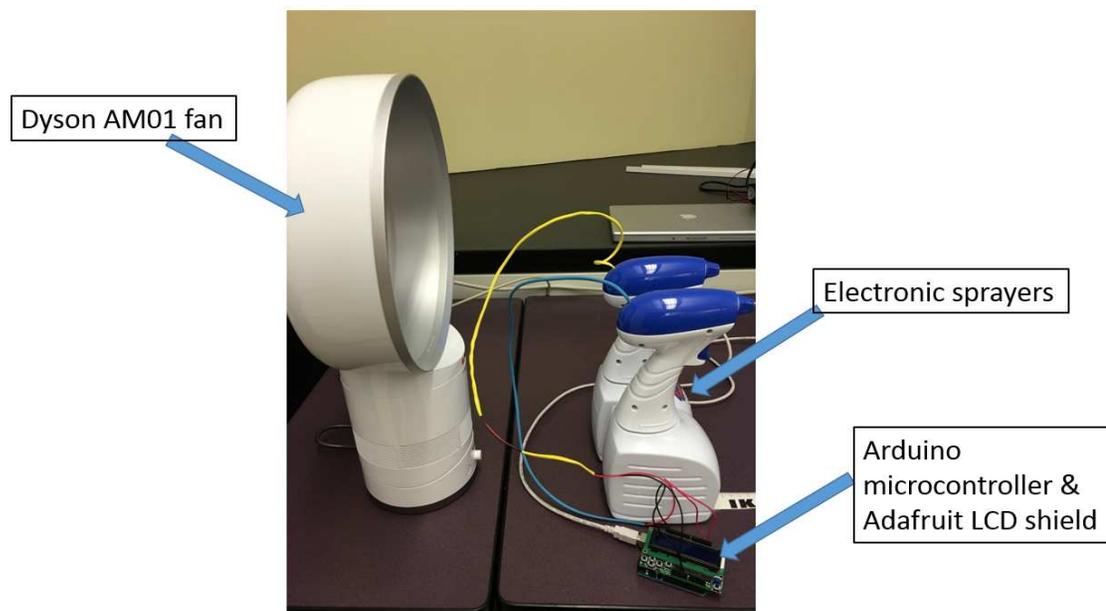
Figure 3.1: The transmitter

The system is controlled by an Arduino microcontroller. The microcontroller takes input information from either a computer or an Adafruit LCD shield, where data can be entered directly with push buttons. The microcontroller converts information to an OOC, and activates the sprayers. Our electronic sprayers have a battery inside them to operate their electrical nozzles, and they have reservoirs at

their bottom which can store alcoholic liquid. (We use ethanol as our chemical transmission symbol). For each OOC transmitted from the microcontroller, the electronic sprayers will be controlled, and will operate their electrical nozzles. The chips in the OOC are represented by sprays of ethanol: one chip is implemented by a spray for "1", and no spray for "0". Furthermore, a Dyson AM01 fan is placed 30 cm behind the sprayers. We turn it on and set it to its maximum level during the experimental transmission. It helps alcohol molecules to propagate more quickly to the receiver through the airborne medium.

3.1.2   The receiver

The receiver contains two MQ-3 alcohol sensors, and each sensor is connected to Arduino microcontroller (see Fig 3.2). These microcontrollers are separate from the transmitter microcontrollers.

In our experiments, alcohol molecules propagate through the medium. The MQ-3 sensors measure the concentration of alcohol, convert it to a voltage, and convey that voltage to their Arduino board. Thereafter, they convert this voltage to digital values and apply OCDMA techniques to detect the signature sequence.

We used MQ-3 sensors (as in our previous paper [12]) to detect alcohol concentrations which propagated by fan. Here, we describe briefly how the MQ-3 sensor works. The MQ-3 sensor is based on tin oxide (SnO_2): after heating to 350°C, the
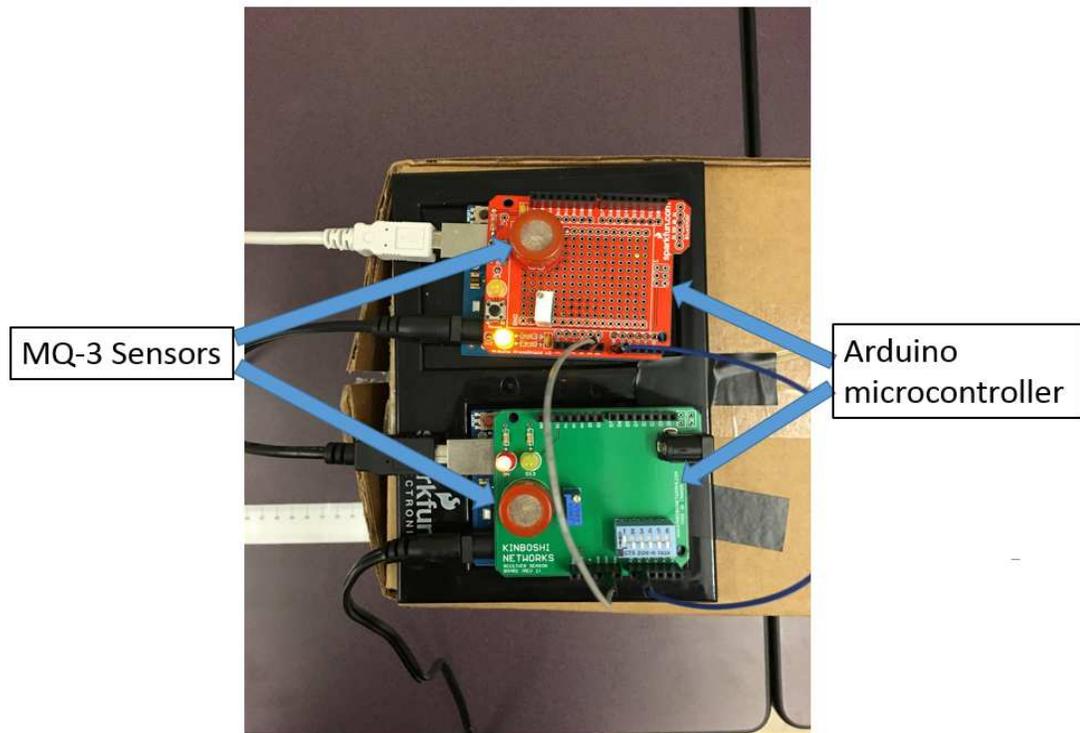
Figure 3.2: The receiver

SnO$_2$ sensor exhibits a drop in electrical resistance in the presence of flammable gases, such as ethanol or propanol [28]. As mentioned previously, we use ethanol as the chemical that is used to transmit information; thus, the MQ-3 can measure its concentration.
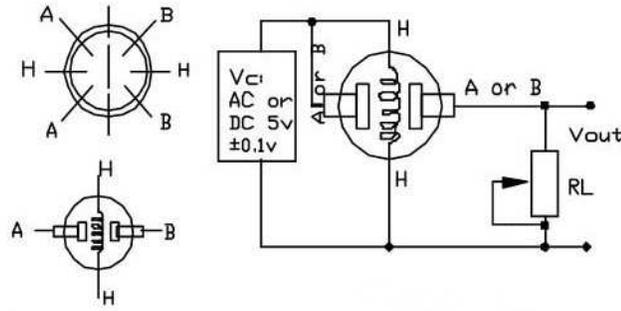
Figure 3.3: MQ-3 sensor. [29].

Fig 3.3 gives a sensor schematic: the MQ-3, illustrated with a circle, has six pins. Both A pins and B pins are used to take measurements, while the H pins are used for providing heating current. As depicted in Fig 3.3, the sensor is a voltage divider on the range from 0V to +5V, with the output across a load resistor R_L. Suppose the resistance across the sensor is R_S, which is a function of concentration; then the measured voltage V_{out} is given by

$$V_{out} = V_{in} \frac{R_L}{R_S + R_L} = 5 \frac{R_L}{R_S + R_L}. \tag{3.1}$$

For example, the output would return a +5V signal as $R_S \to 0$ in the presence of alcohol saturation, and 0V as $R_S \to \infty$ in clean air. The response curves relating

R_S to concentration are given in the sensor data sheet [29].

3.1.3   The testing environment

Our goal is to show that our experimental apparatus can be used to test OCDMA. We set the receiver 225 cm away from the transmitter, and we utilize an OCDMA scheme in our experiments from one sprayer to one sensor. Assume there are 15 time chips in the signature sequence, and we set the first pulse at the first time chip to represent the start of detection. The sensor would start to detect molecular pulse for each time chip; moreover, if each pulse appears at a specific time chip as this sensor's assigned signature sequence, we say this sensor receives a "1" signal. Otherwise, it receives a "0" signal. Since we only have two sets of communication equipment (i.e., two transmitters and two receivers), a short-length OOC may satisfy the relation in (2.10).

We want to achieve ideal asynchronous operation to reduce the probability of error as mentioned before in (2.9). Since every signature sequence starts with a chip "1", the sensor would treat every received pulse as the potential first pulse of its signature sequence (i.e. the sensor would monitor the remaining 14 time chips after each pulse is detected). Thereafter, it can execute its regular detection as mentioned in last paragraph (and described in detail in the last chapter). By using this method, other users can start their communication at any time without

synchronization, and we can achieve asynchronous operation. This is a key benefit of OCDMA.

Now, the complete process is that each sensor would detect the rest of the time chips to match its assigned signature sequence after each pulse is detected. Once it finds its "1" signal, it would start its regular detection (i.e. not the full detection process for each pulse as before). If the sequence detected matches its signature sequence, then "1" is detected, otherwise "0" is detected. With this communication logic, we apply OCDMA to molecular communication.

3.2   Simulation system model

Although experiments give the most reliable data, simulations would be more efficient and less time consuming compared with experiments. We want to get the optimal performance of our communication system, thus, we want to evaluate and improve the performance in a variety of circumstances. The performance relies on the raw sequence (i.e. the original sequence without any processing). In order to evaluate the performance under many conditions, it is far more efficient to evaluate it in simulation. In this section, we show that the simulation is an accurate subtitute for experimental results.

We do simulations in Matlab based on the model described in the last section. Our communication system is comprised of the transmitter part and the receiver

part as we mentioned above. Both of these can be simulated.

Paper [30] proposed two models.

$$M_1(t) = \frac{a}{(\sqrt{t})} \exp\left(-b\frac{(d-ct)^2}{t}\right) \tag{3.2}$$

$$M_2(t) = \frac{a}{(\sqrt{t^3})} \exp\left(-b\frac{(ct-d)^2}{t}\right) \tag{3.3}$$

Here a, b, and c are constants, which represent three main factors that affect system response. Constant a is a scaling factor related to the "respond and recover" times of the sensor; the tin oxide sensor needs time to respond a sudden concentration change, and it also needs time to recover to its original voltage after this sudden concentration change (see also [12], where the sensor needed some time to "wake up" after it detected an alcohol concentration wave). Constant b is related to the diffusion coefficient. Since a fan is used to generate air flow, the flow is not perfectly laminar and uniform: the fan's blades would create streams of air pressure, and that would cause turbulent flows. Therefore, b plays a role as a correction to the diffusion coefficient to account for the turbulence. Constant c is related to flow speed. Alcohol molecules have a certain weight, and experience friction, so a stream of alcohol molecules would propagate slower than the wind speed. Thus, c is correction to the average flow speed. In addition, d is the distance between sprayer and sensor, and t is time. The output is the voltage from sensor.

The difference between models (3.2) and (3.3) is that (3.3) assumes that there

Table 3.1: Paper [30]'s value versus our value

Parameters	Paper's values	Our values
a	15.3909	15.3909
b	1.6035×10^{-4}	2.1×10^{-4}
c	35.3137	0.08

are residual alcohol molecules around the sensor after detection, while (3.2) does not (i.e. no residual alcohol molecules stay around the sensor). Thus, (3.2) is ideal in the sense that there is no interference. However, we choose (3.3) for our simulation because our experiment is in a confined space (as we keep all doors and windows closed), so there are alcohol molecules that stay in the environment after detection. This is a typical application scenario for our system. Thus, (3.3) helps us to accurately simulate the response from receiver.

We want the simulation model to represent the performance our of receiver accurately. We use parameter values from paper [30] (Table 3.1) in simulation model (3.3), and check the performance of the simulation versus the performance of the receiver in real-world experiments. As mentioned, our experiment set up in a closed room. The distance between the sprayer and the sensor is 225 cm (i.e. $d = 225$ cm in (3.3)). The duration of the spray within each chip is 100 ms, and

the duration between each chip is 5 s (i.e. the chip time $T_c = 5$ s). A Dyson AM01 fan is set on its "high" setting at 30 cm behind the sprayer. Moreover, the sprayer is controlled by an Arduino board with a program that we upload in advance.

For example, where the signature sequence is "1010000", and the data we want to send is "10110", then the whole sequence that would be sent out is "10100000000001010001010000000000000". This is a 35-chip sequence, and the duration of the experiment is

$$T = T_c * 35 = 175 \text{ s}. \tag{3.4}$$
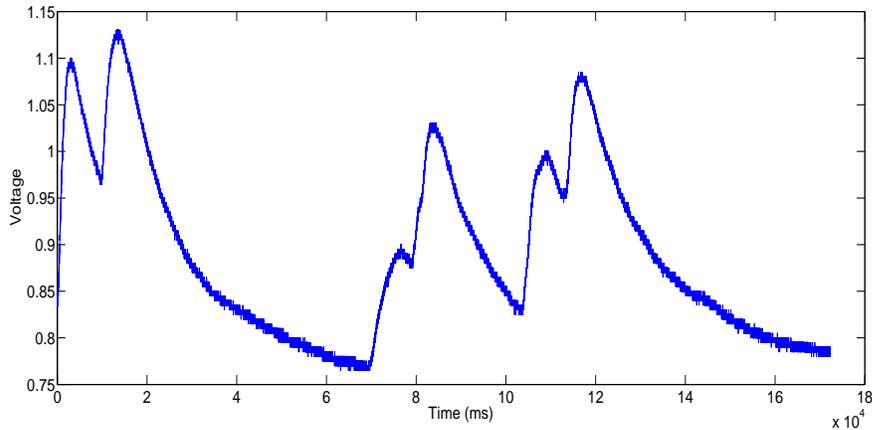
Figure 3.4: The performance of receiver

Notice that the weight of the signature sequence is 2, and the weight of the transmitted sequence is 6; that is, there are 6 "1" chips.

In Fig 3.4 we show the performance of receiver, illustrating the measurement

at the sensor versus time. The sensor output voltage has a range from 0V to +5V (based on (3.1)). Moreover, we can observe six clear peaks at the times corresponding to the "1" chips in Fig 3.4.
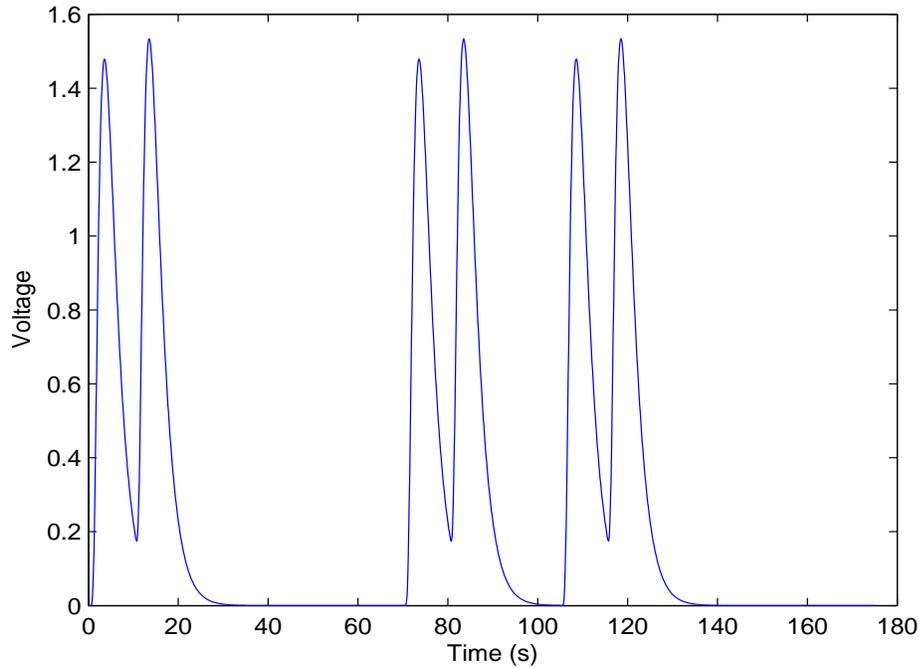
Figure 3.5: The performance of simulation by using values from paper [30]

Fig 3.5 gives simulation results using the parameter values from paper [30], plotting voltage versus time. This figure also has six sharp pulses, one possible reason for which is that turbulence may mix molecules and air better, leading to sharper pulses; however, there are two significant differences compared to the experimental performance in Fig 3.4. In that figure, the difference of voltage of

the first pulse (i.e. the voltage of the peak 1.1 minus its initial voltage 0.835) is

0.265 V. This difference is significantly lower than the voltage change of the first

pulse from Fig 3.5. The second difference is that the pulses in Fig 3.5 are sharper

than those in Fig 3.4, and the voltages between two "1" chips in all three signature

sequences have dropped to the almost initial level. However, this does not happen

in Fig 3.4. So, using the parameter values of paper [30] gives a poor representation

of the performance of the receiver.

Figure 3.6: Fan used in [30] versus Dyson fan without fan blades

In order to accurately simulate our experiment, we should modify the parameter

values. The parameter a is related to the the response and recovery time, and these

are two inherent features of the sensor. Because we are using the same sensor, we

Table 3.2: Wind speed at test point

Fan	Wind speed
Fan from [30]	2.7 m/s
Dyson	0.95 m/s

can keep the same value of a. The parameter b is related to the effect of flow. Even though the diffusion coefficient is the same since we are using the same alcohol molecule, the effects of turbulent flows are different. The reason is that [30] used a different type of fan; however, in our experiment, we use a Dyson fan (see as Fig 3.6) which does not have any blades. As we mentioned before, blades create turbulent flows; therefore, we should modify parameter b since the flow characteristics have changed compared with [30]. In terms of the parameter c, as we said before, we use a new fan to generate wind, and that can result in a difference in wind speed. Thus, although the models (3.2) and (3.3) are valid, both parameters b and c should be modified when compared with [30], and a should remain the same.

We measured the speed difference of the two fans. We used a Pyle PMA82 digital anemometer to measure the wind speed for each fan, setting the anemometer 30 cm in front of the fan, and at the same height as the sprayer nozzle. As we can see in Table 3.2, the Dyson fan generates lower wind speed than the previously-used fan from [30]. Thus, we can change the parameter value c to reflect this wind
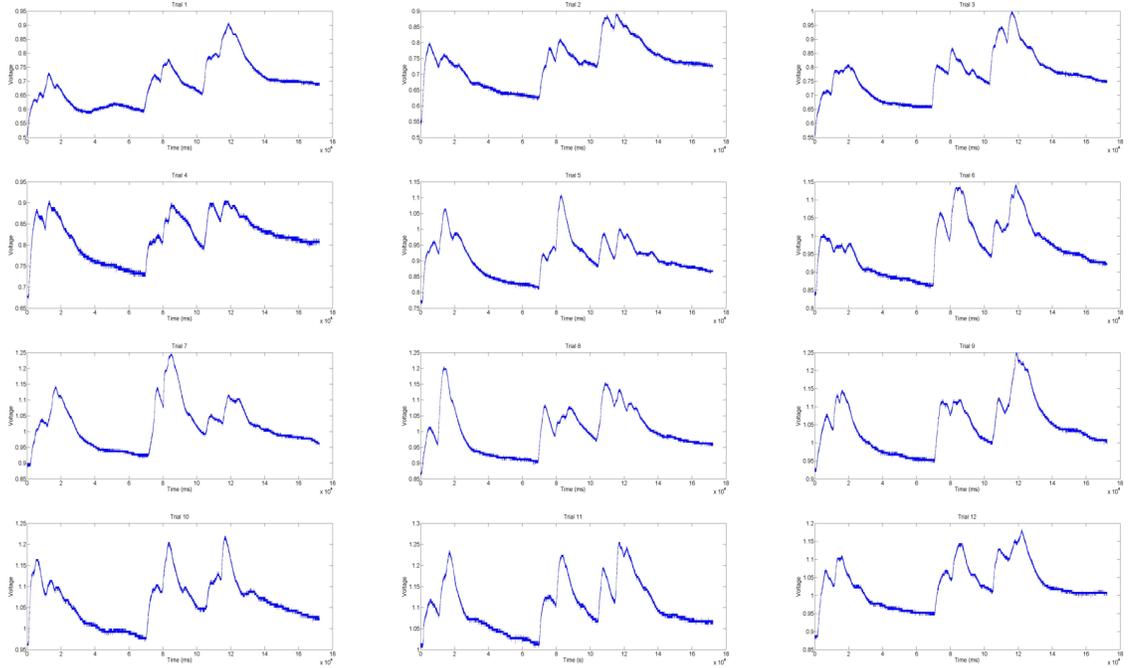
Figure 3.7: Sensor performance for 12 trials

speed; however, there are differences between wind speed and flow speed, so c is not exactly given by the value in the table.

In order to find the appropriate values of b and c in our modelling equations, we collect 12 experimental trials of the system performance (Fig 3.7). We obtain the average difference of the first chip, which is 0.18 V. We modified the parameters b and c to match this difference in voltage. The matching values are given in Table 3.1, which gives the accurate parameter values representing our experimental apparatus. Fig 3.8 shows a simulation using these parameter values: the difference in the first chip is 0.18V, and the performance more closely matches that in Fig

3.4.
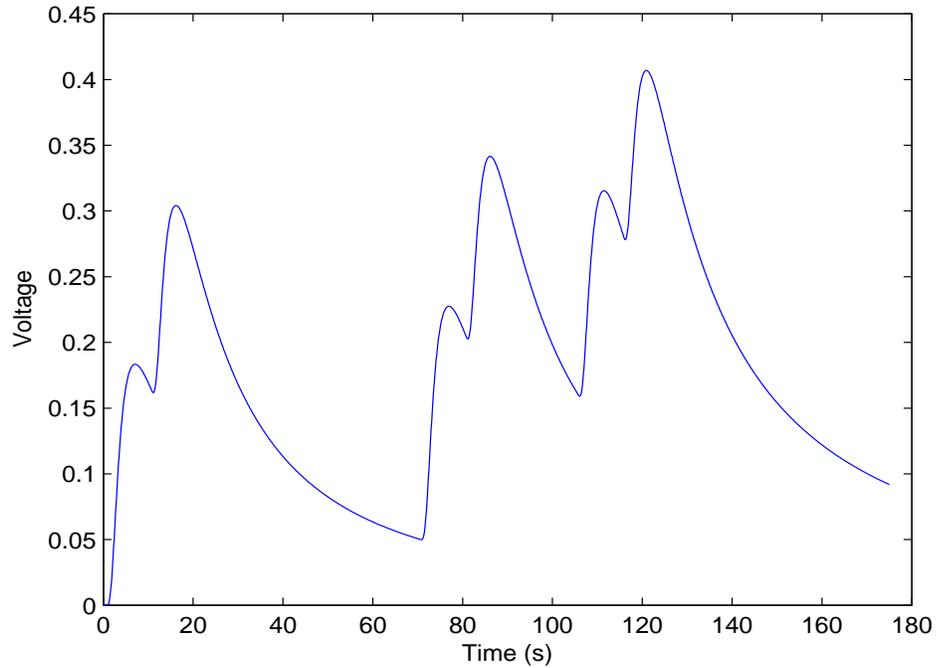
Figure 3.8: Simulation performance with modified value

3.3   Conclusion

By applying the simulation model with new parameter values, we can see our simulation model (Fig 3.8) matches our experiments (Fig 3.7). Even though the performances of experiment are not consistent from trial to trial due to noise and imperfections in the experimental apparatus (we will explain this in next chapter), our simulation model is still good enough to represent our experiments.

4 Simulation result and performance of simulation system

In this chapter, we describe how we simulate and obtain the performance of simulation for both one and two pairs of transmitting users. Moreover, we only consider chips sequences transmission, and we don't consider signature sequences. This means we don't apply OCDMA scheme in this chapter.

4.1 Simulation set up

We set up simulations in both the transmitter and the receiver. In the transmitter, we sent a 100 chip sequence starting with "1". The transmitted sequence is a random binary sequence, with $\Pr(1) = 0.14$. Although in this chapter we evaluate the performance of the system using raw chips (OOCs will be used in the next chapter), this value for $\Pr(1)$ is chosen to reflect real signature sequences: for example, a signature sequence with length $F = 7$ and weight $K = 2$, while data bits "0" and "1" equiprobable (i.e., the transmission is blank with probability 0.5).

Thus, we obtain

$$\Pr(1) \approx \frac{2}{7}(0.5) = 0.14 \tag{4.1}$$

Moreover, we set a "1" chip at the beginning of the transmitted sequence, since we use this "1" as a sign to inform the receiver to start detection.

We now describe the detection algorithm used at the receiver. For each chip in the sequence, we measure the voltage at three points: at the beginning, in the middle, and at the end of the chip duration. Furthermore, we calculate the differences between middle point and start point, and between end point and middle point. If one of these differences is higher than a threshold, we say we get a "1" bit, otherwise we say "0" is obtained.

Chips last T_C seconds. For the measurement points at the start, middle, and end of the chip, we choose measurement times of

$$T_{\text{start}} = 0.2T_C \tag{4.2}$$

$$T_{\text{middle}} = 0.5T_C \tag{4.3}$$

$$T_{\text{end}} = 0.9T_C \tag{4.4}$$

After obtaining the measurements, and letting $V(\cdot)$ represent the voltage of the

45

measurement, we calculate the differences

$$D_a = V(T_{\text{middle}}) - V(T_{\text{start}}) \tag{4.5}$$

$$= V(0.5T_C) - V(0.2T_C) \tag{4.6}$$

$$D_b = V(T_{\text{end}}) - V(T_{\text{middle}}) \tag{4.7}$$

$$= V(0.9T_C) - V(0.5T_C) \tag{4.8}$$

Finally, we compare these differences with the threshold V_{th} (the threshold comes from previous works [10] and [12]): we decide that the chip is "1" if one of our differences is equal to greater than the threshold, and we decide "0" otherwise. The decision function $O(D_a, D_b)$ is given by

$$O(D_a, D_b) = \begin{cases} 1 & D_a \geq V_{\text{th}} \text{ or } D_b \geq V_{\text{th}} \\ 0 & \text{otherwise} \end{cases} \tag{4.9}$$

This decision algorithm is based on the one used in [10].

In addition, we will add noise in our simulation to reflect that our experimental apparatus is imperfect. (For example, see Fig 3.7.) One main source of noise is the difference in the amount of alcohol in each spray. Even though the spraying time is fixed at 100 ms for each "1" chip, we can't ensure that the amount of alcohol in each spray is exactly the same. This may cause differences in voltage. Another source of noise is variations in flow speed. We perform our experiment in a closed room, however, we can not guarantee that there are not any variations in wind other

than from our fan during the experiment. Thus, these variations can be considered as noise. Furthermore, when we evaluate OCDMA, there will be interference, and adding noise at this stage will help us understand the effect of the interference.

Recall that one simulation consists of 100 chips. We run each simulation scenario 1000 times. Using the decision algorithm in (4.9), we calculate the number of chips detected correctly, and the number detected incorrectly; these values are used to calculate the probability of error. We focus on the probability of error given that a "1" chip was sent. The main cause of decoding error is missing a "1" bit in a signature sequence: say we have a signature sequence "1,0,1,0,0,0,0", then the key to determine that this sequence is the correct signature sequence is to observe a "1" pulse at the first and third bits. On the other hand, if we observe "0", we may decide that this sequence is not a signature sequence. Since a user's correct signature sequence is used as a mask, the user ignores the channel in positions where a "0" appears in its signature sequence. Therefore, we only consider the probability of error given that a "1" was sent in our simulations.

4.2 Simulation results

4.2.1 One to one simulation

We first test the performance from one sprayer to one sensor. To account for noise, we add random independent, identically distributed Gaussian noise at every sample point; the noise has mean of zero and variance of σ^2. To account for the noise in our system (see Fig 3.7), we use $\sigma^2 = 1$. This gives a strong noise compared to our voltage level (see Fig 3.8). In this result, we set the chip duration $T_C = 5$ s, and we run simulations at distances (from sprayer to sensor) of $d = 175$ cm, 225 cm, and 275 cm.

The $randn()$ function gives us a Gaussian distribution with mean = 0 and variance $\sigma^2 = 1$. In order to reduce the noise $randn()$ level, we reduce the range of $randn()$ by multiplying $randn()$ by standard deviation $1/x$. Therefore, we reduce the range of noise by increasing x, and noise level is thereby reduced. In Fig 4.1, we plot probability of error versus x, where $\sigma^2 = 1/x^2$.

At a distance of $d = 175$ cm (blue solid line), the top error rate is 0.025 when $x = 1$, and the error rate drops below 10^{-6} when x reaches 200. This is reasonable because the noise here has already dropped to a small value, and the sensor is closer to the sprayer compared to the other two distances.

In terms of $d = 225$ cm, we can see the error rate versus x (green dashed line) is
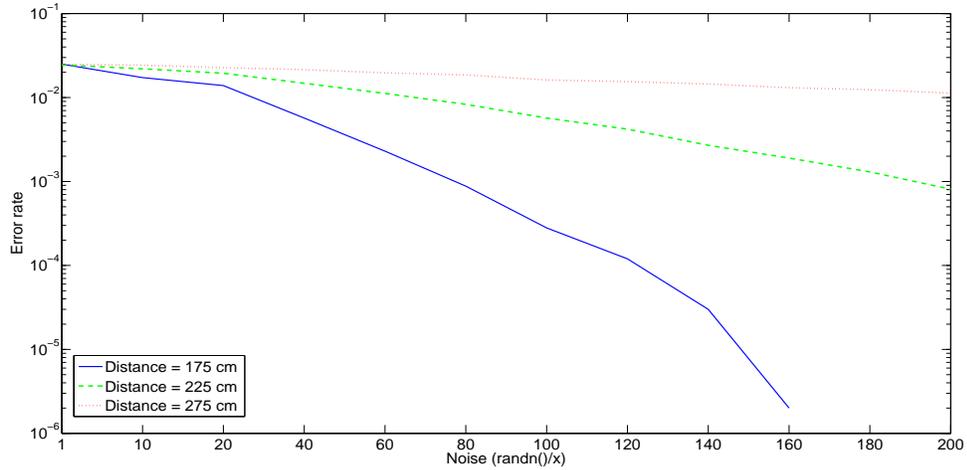
Figure 4.1: Error versus Noise at each sample point

also on a declining trend. It goes from an error rate of 0.0245 at $x = 1$ to 0.00081 when $x = 200$. The error rate doesn't change much at the maximum noise level $(x = 1)$ compared to $d = 175$ cm, however, the error rate remains significant even at $x = 200$. In this situation, the distance from the sprayer to the sensor can be consider as an important factor: the sensor is already out of "better performance" range when it reaches $d = 225$ cm.

We prove this assumption by increasing the distance to $d = 275$ cm, and the performance is even worse, as we see in the $d = 275$ cm curve (red dotted line). Even though the tendency of graph is going down, the error rates are higher compared to $d = 225$ cm.

Comparing these three lines, noise variance and distance are factors that af-

fect the error rate; further, the system can improve its error rate with smaller σ^2 and smaller d. All these characteristics are similar to conventional communication systems.
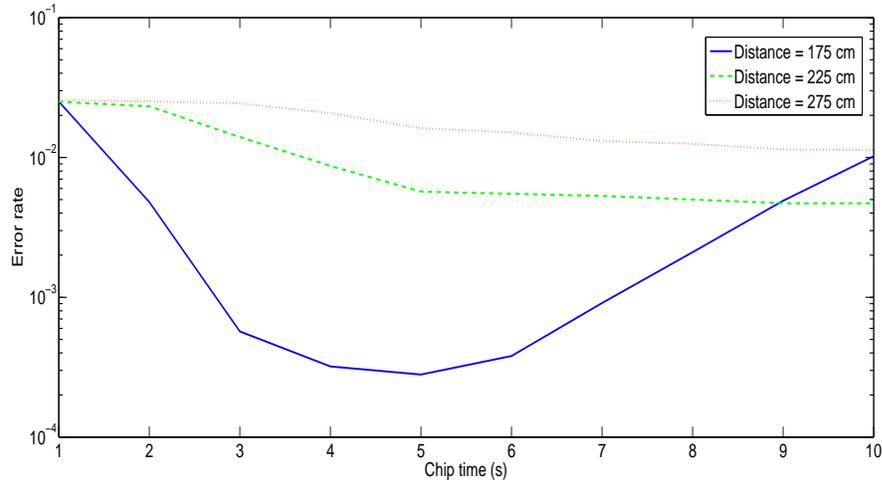
Figure 4.2: Error versus Chip time

We also tested the relationship between error rate and chip time T_C. We test T_C in three distance scenarios: $d = 175$ cm (blue solid line), 225 cm (green dashed line), and 275 cm (red dotted line). The noise variance is $\sigma^2 = 10^{-4}$. Results are shown in Fig 4.2, where the error rate has a complicated relationship with chip time T_C. The chip time affects not only the system efficiency (i.e. if we have a longer chip time, we need a longer time to represent signature sequence, and that would reduce efficiency), but also the error rate. The effect on error rate has different features compared with conventional communication, where lower bit rate generally means

50

lower error rate.

At distance $d = 175$ cm, it is clear that error rate has a sharp decrease from 0.025 at $T_C = 1$ s to 0.00028 at $T_C = 5$ s. However, the error rate increases to 0.0102 while $T_C = 10$ s. We can see it is a convex curve, and the error rate increases even though we lower the bit rate. And at distance $d = 225$ cm and $d = 275$ cm, they are gradual decrease compare to $d = 175$.

Based on Fig 4.2, it is not enough to get optimal performance by just reducing d, we also need to shrink T_C to the optimal time. Moreover, we tested two more simulations at $T_C = 20$ s for $d = 225$ cm and 275 cm, and we get the error rates are 0.0363 and 0.0203, respectively. It means that the graph of error rate versus T_C is a convex curve, and error rates would increase as T_C increases. The reason is that our detection algorithm samples three points out of the whole chip time, as we described in last section, and the voltage would already reduce to a low level before the middle of the chip time (T_{middle}). This may cause missed detection (i.e. detect 0 when a 1 was sent). Therefore, using out existing algorithm, we can not obtain a low error rate solely by increasing the chip time.

Distance is an another parameter in our simulation. From last two figures, we have demonstrated that the distance d between sprayer and sensor affects the performance of our system. In Fig 4.3, we still choose $\sigma^2 = 10^{-4}$ as the noise at each point, and we used $T_C = 3$ s (blue solid line), 5 s (green dashed line), and 7 s

(red dotted line) as chip time. We can see error rates increasing when we set the sensor either closer or farther, and we can find the lowest error rate of these three scenarios from the figure.
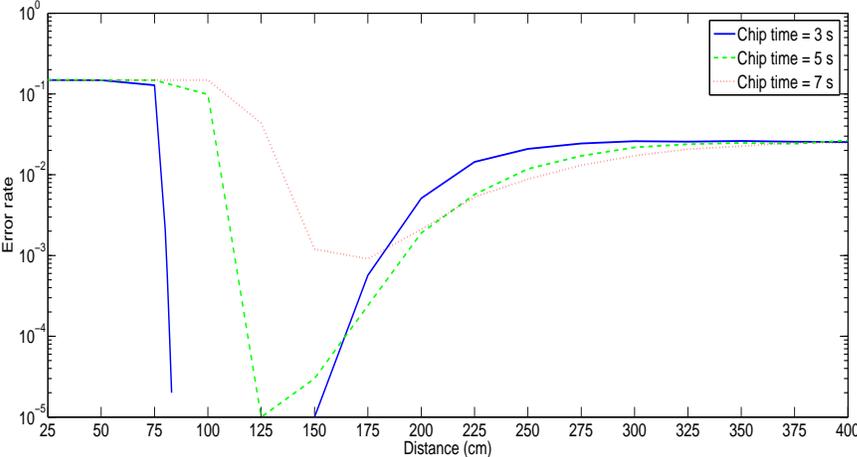
Figure 4.3: Error rate versus Distance between sprayer and sensor

For $T_C = 3$ s, we observed no errors from distance $d = 84$ cm to 149 cm. Thus, at these points, the error rate is below what the simulation can measure. For $T_C = 5$ s, the lowest point is located at $d = 125$ cm with error rate 0.00001. Moreover, for $T_C = 7$ s, the lowest error rate is 0.00091 when distance is $d = 175$ cm. These numbers represent proof our previous assumption: to achieve optimal performance at relatively small distance, we should shrink T_C. Thus, d should decrease as T_C decreases to maintain optimal performance. Furthermore, error rates increase after the optimal point, and rises to 0.026 when distance reaches

$d = 400$ cm. Meanwhile, when the distance is smaller than $d = 125$ cm, the error rate rises quickly, and almost reaches its highest point of 0.0148 when the distance is smaller than $d = 50$ cm.

We can get the lowest error rate of 0.00001 when we set $d = 125$ cm and $T_C = 5$ s. However, we can get the lowest error rate at $d = 200$ cm if we used the values from [29], which used a different fan. The reason for the difference is that the Dyson fan's wind speed is lower (as we show in Table 3.2). We believe the higher speed of flow can set the optimal point farther, because in the same condition ($T_C = 5$ s), the other fan's optimal point was $d = 200$ cm, whereas the Dyson optimal point is $d = 125$ cm.

There are two factors which govern the performance of the system as a function of distance. The sprayers would leak some alcohol liquid while spraying, and this liquid would increase the concentration of alcohol molecule around the sprayers. Therefore, detection is unreliable if the sensors are too close to the sprayers. Moreover, the high error rate at a long distance is the alcohol molecules is caused by the low concentration of alcohol, as the alcohol has already diffused in the air after propagating for a long distance. Therefore, there is a high error rate at both near and far distances.

4.2.2 Two to two simulation

To achieve multiple-access communications, we must have at least two users. Here we simulate two experimental apparatuses being used at the same time: that is, two sprayers and two sensors.

In terms of communication between two sprayers and two sensors, we set the two sensors close to each other with a separation of 8 cm (Fig 3.2), and we assume they have the same performances (i.e. both sensors will respond with the same voltage to an alcohol concentration) when they receive a chip. The two sprayers are also 8 cm apart (Fig 3.1), and we also assume they would spray the same quantity of alcohol molecules with same speed. However, the voltage response of the two sprayers spraying together is nonlinear: it is not the same as the addition of the two simulated voltages. That is,

$$v^{1,2}(t) \neq v^1(t) + v^2(t) \tag{4.10}$$

Here, $v^{1,2}(t)$ is the response of sensor when both sprayers spray together, while $v^1(t)$ and $v^2(t)$ are the sensor responses to the individual sprayer, if the other sprayer is silent. This non-linearity was demonstrated in paper [10]. The non-linearity is not well modelled, and introduces uncertainty. To account for the uncertainty, we use a noise process, defined by

$$nM_2(t) \approx \frac{N_{M_2}}{(\sqrt{t^3})} exp(-b_{M_2} \frac{(c_{M_2}t - d)^2}{t}) \tag{4.11}$$

54

from paper [30]. M_2 means this function considers there are residual alcohol molecules around sensor after detection, it is same with (3.3). N_{M_2} is a Gaussian random variable with mean is -0.7356, and we used this value. The rest value b_{M_2} and c_{M_2}, we used our value from table 3.1.
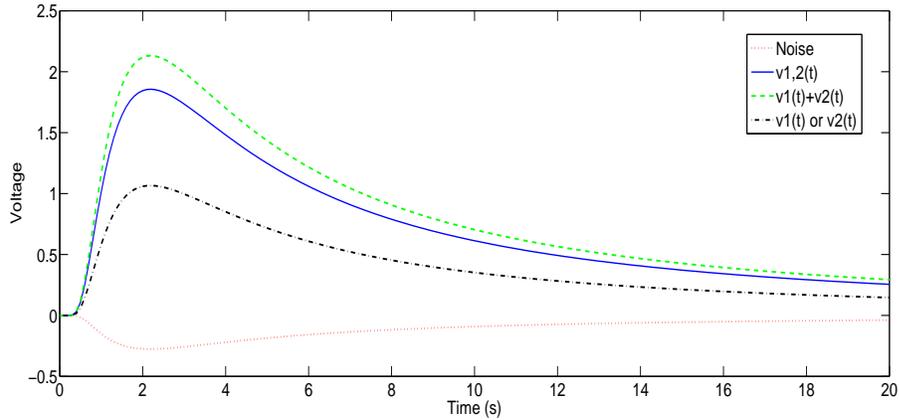
Figure 4.4: Sprayers response of nonlinearity

In Fig 4.4, we can see black dash-dot line is each responses from sensor 1 and 2, there are exactly same since we assume they receive from same performance sprayers and have same response. Green dashed line is the numeral voltage superposition of two sensors response, and blue solid line is the voltage response of two sprayers spray together. The last in red dot line is the noise function (4.11).

In a two-to-two simulation, we calculate average error rate of both sensor 1 and 2. We first test test the performance of the error rate versus distance between sprayers and sensors in Fig 4.5. In this simulation, we did not add any extra noise
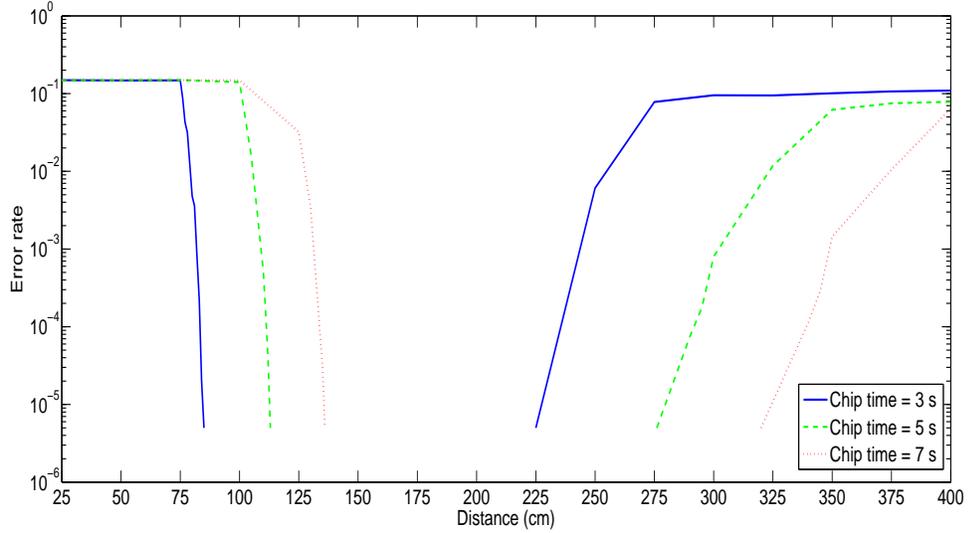
Figure 4.5: Error rate versus distance between sprayers and sensor in two sprayer to two sensors

except the noise function given above. Moreover, we have three scenarios here, which are: $T_C = 3$ s (blue solid line), 5 s (green dashed line), and 7 s (red dotted line). All the error rates of the three scenarios have approximately same maximum error rate as 0.0149 when sensors are closest to the sprayers at $d = 25$ cm. As we explain before, due to imperfections in the sprayer and there is a mass of molecules around the sprayers, and all these would affect the sensors' measurements. The distance of the optimal point is closer to the sprayer when chip time is reduced. Furthermore, the error rate seems grows after these optimal points. This figure also proves that both close and far distance would cause high error rate.

In terms of error rate versus chip time in two-to-two communication, Fig 4.6 gives us a clear view. We use $\sigma^2 = 10^{-4}$ as the noise variance. There are three scenarios: $d = 175$ cm (blue solid line), 225 cm (green dashed line), and 275 cm (red dotted line).
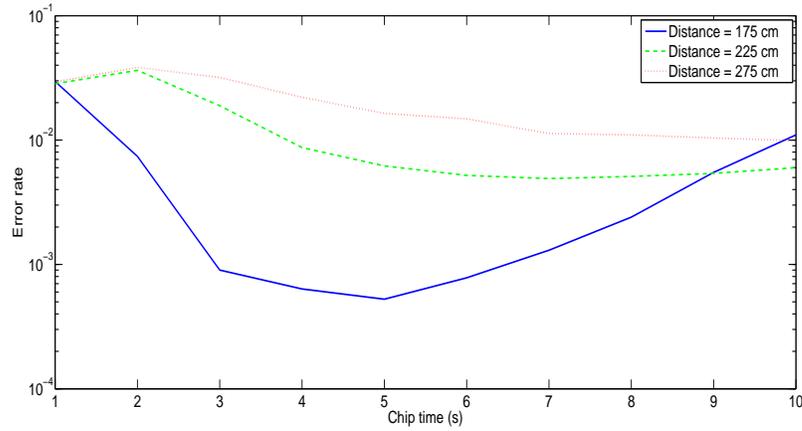
Figure 4.6: Error rate versus Chip time in two sprayers to two sensors

We can see the performance of two-to-two is similar to the performance of one-to-one (Fig 4.2). We see that the graph of error rate versus chip time is a convex curve, and we believe error rates would increase as chip time T_C increases.

Based on the performance of both one-to-one and two-to-two in simulation, we can see the error rate increases with increasing noise level σ^2. However, while increasing either chip time T_C or distance d between sprayers and sensors, the error rate will drop down and then increase until chip duration or distance reach a specific level, and this feature makes the figure look like a convex curve. This implies that

there are optimal levels for T_C and d in these scenarios. For example, in Fig 4.5, the low-error-rate range (the optimal range) for chip time $T_C = 7$ s is from $d = 125$ cm to $d = 350$ cm, while the low-error-rate range for chip time $T_C = 5$ s is from $d = 100$ cm to $d = 300$ cm, and the low-error-rate range for chip time $T_C = 3$ s is from $d = 75$ cm to $d = 225$ cm. We can see that the optimal range is moving closer to the sprayers when we reduce T_C. This feature also appears in Fig 4.2, Fig 4.3, Fig 4.5, and Fig 4.6. For example in Fig 4.3, the optimal range for chip time $T_C = 7$ s is from $d = 150$ cm to $d = 175$ cm, the optimal range for $T_C = 5$ is from $d = 125$ cm to $d = 150$ cm, and the optimal range for $T_C = 3$ s is from $d = 75$ cm to $d = 150$ cm. From Fig 4.2, even though we can only see a convex curve for $d = 175$ cm and the other two curves are still decreasing until $T_C = 10$ s, we tested the error rate for them at $T_C = 20$ s, and the relative error rates increase back to a high level. Thus, they also appear to be a convex curve. We can apply an OCDMA scheme in this optimal range to achieve the best performance.

5   Simulation result for applying OCDMA

In the previous chapter, we established how to conduct simulations in one-to-one and two-to-two communication scenarios. In this chapter, we use the same simulation setup to evaluate OCDMA. We encode data sequences using the OCDMA scheme described in Chapter 2: each user uses their own signature sequence to represent a "1" bit, and keeps silent (all "0" sequence) to represent a "0" bit. We send 100 bits of data in each simulation, and we run each simulation scenario 1000 times. Moreover, the error rate here is not the probability of missing a "1" bit as in the last chapter; since we are testing OCDMA data transmission, we will consider the error rate as both missed detection (i.e. detect "0" bit, when a "1" is sent) and false alarm (i.e. detect a "1" when a "0" is sent). In two-to-two communication, the error rate is the average error rate of both users.

5.1   Simulation result: Chip sequence length $F = 7$

For our first results, we use a 7-bit signature sequence (i.e. $F = 7$) where each signature sequence has a weight of $K = 2$. Since we have a pair of users (i.e. $N = 2$), it satisfies formula (2.10). We test 100 bits of data per simulation scenario, and $F = 7$ to represent 1 bit of data; therefore, we have a sequence of 700 (7×100) chips in total to send in one simulation scenario.

There are two users: a and b. We set data transmission from two sprayers (S_a and S_b) to two sensors (R_a and R_b) while applying the OCDMA scheme. S_a communicates with R_a, and S_b communicates with R_b. We set a signature sequence "1010000" for user a (from S_a to R_a), and a signature sequence "1001000" for user b (from S_b to R_b).

For simplicity, in our simulation scenarios, the transmissions of both users are synchronized, although this is not necessary in OCDMA. Moreover, for simplicity, we obtain the performance without noise, as the dominant source of errors in this scenario is interference from other users.

In terms of the detection scheme, we determine if it is a signature sequence by breaking the entire received sequence into windows of length $F = 7$. We set every seven chips from our receiver as one window, and for every "1" chip in each signature sequence, we determine whether there is a "1" chip at this position in the

window. For example, if we decide the first and third chips are "1" in a window, we say this window is a signature sequence for user a, and we decide that "1" bit was sent by user a. Moreover, if the first and fourth chips are "1", we say this window is a signature sequence for user b, and we decide "1" bit was sent by user b. Furthermore, if the first, third, and fourth chips are "1", then we say both users a and b get their signature sequence, and we decide "1" bit for both users. The last situation is where we observe neither first and third chips nor first and fourth chips as "1", which means we do not get any signature sequence, and we decide that both users sent "0". After this window detection, we will detect next window with the same method.
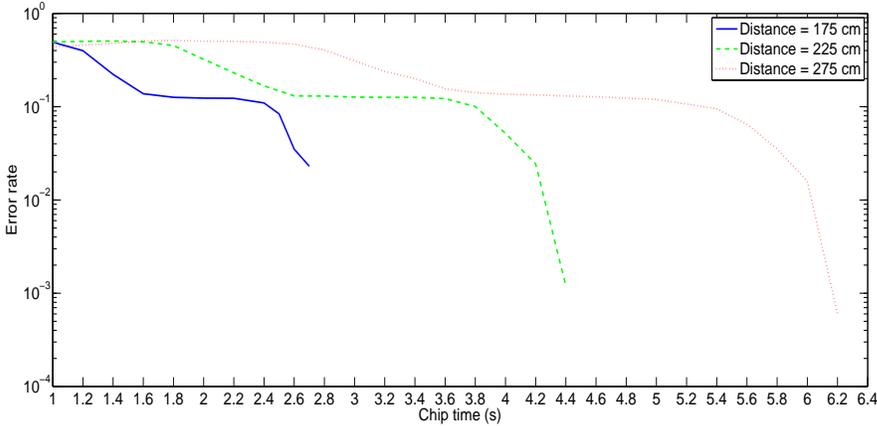

Figure 5.1: Error rate versus Chip time in data transmission with F = 7

In Fig 5.1, we test the error rate versus chip time from $T_C = 1$ s to $T_C = 10$ s in the detection scheme described above. In this figure, the error rate decreases

gradually with increasing chip time, up to $T_C = 6.4$. However, in terms of scenarios of $d = 225$ cm and 275 cm, the curves have a slight increase from $T_C = 1$ s to 1.4 s and 1.8 s respectively. This phenomenon exists in Fig 4.6 too, it shows that the worst performance of long-range transmission does not happen at the closest point; instead, it happens between $T_C = 1$ s and 2 s.

Note that the errors no longer increase when chip time T_C increases. In figures of error rate versus chip time T_C for one-to-one and two-to-two raw sequence transmission in Chapter 4 (i.e. Fig 4.2 and Fig 4.6), we can see the error rate would increase as T_c increases, after it reaches its optimal value. However, we do not observe this in our data transmission scenario (i.e. Fig 5.1). The reason is that we added noise our simulations from Chapter 4, whereas (as we mentioned) for simplicity we do not add any noise in the data transmission scenario here. The optimal T_C values are related to the addition of noise.

Error rate versus distance (for $F = 7$) is plotted in Fig 5.2. This figure is similar to the error rate versus distance figure in two-to-two raw sequence transmission from chapter 4 (Fig 4.5). The curve of $T_C = 3$ s, 5 s , and 7 s keep their flat and highest error rate until $d = 76$ cm, 101 cm, and 121 cm, and they then drop sharply until a low-error region at distances $d = 80$ cm, 106 cm , and 127 cm. After that, errors appear in the curve of $T_C = 3$ s at $d = 185$ cm, $T_C = 5$ s at $d = 246$ cm, and $T_C = 7$ s at $d = 298$ cm.
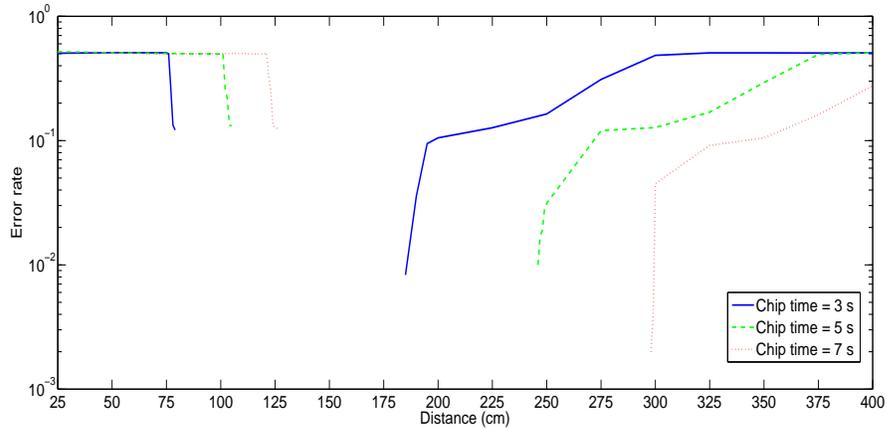
Figure 5.2: Error rate versus distance in data transmission with F = 7

Fig 5.2 gives us a same proof as Fig 4.5, that is the optimal range is approving to sprayers if we reduce the chip time, and we can achieve 0 error rate if we give it a ideal environment (without any noise).

5.2 Simulation result: Chip sequence length $F = 15$

In chapter 2.2, we introduced an additional pair of chip sequences with $F = 15$. Since we have two users here, to satisfy formula (2.10), this chip sequence has a weight of $K = 3$. Therefore, we assign the chip sequence "100100001000000" to user a (sprayer S_a to sensor R_a), and we assign the chip sequence "101000100000000" to user b (sprayer S_b to sensor R_b). Again, we say each sensor decides a "1" was sent for that user once it receive a user's own chip sequence, otherwise, it decides "0".

We test the error rate versus chip time from $T_C = 1$ s to 10 s with chip sequence length $F = 15$ in Fig 5.3. We expect the performance of $F = 15$ to be much better than $F = 7$, since the weight K is higher. All three curves achieve low error before $T_C = 4$ s, we now sample every 0.2 second as well to observe the shape of the curve, and to determine when low error is achieved.
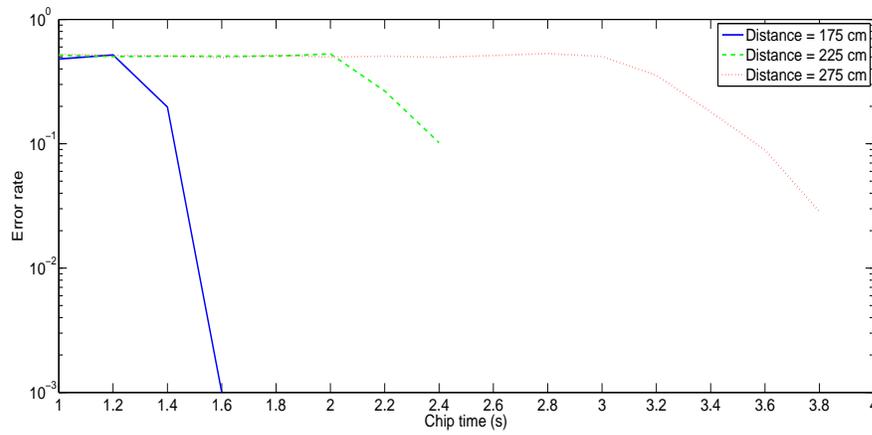
Figure 5.3: Error rate versus Chip time in data transmission with F = 15

From this figure, we can see the performance of $F = 15$ is better than $F = 7$. In the $d = 175$ cm curve with $F = 7$, the low error range is reached at $T_C = 2.8$ s; whereas with $F = 15$ the low error range is reached at $T_C = 1.8$ s, 1 s earlier than $F = 7$. Moreover, both the $d = 225$ cm curve and the $d = 275$ cm curve are 2 s and 2.4 s ahead of $F = 7$, respectively. It seems the chip sequence length $F = 15$ has much better performance around the distance from $d = 175$ cm to 275 cm. Furthermore, the low error rate is maintained until $T_C = 10$ s.
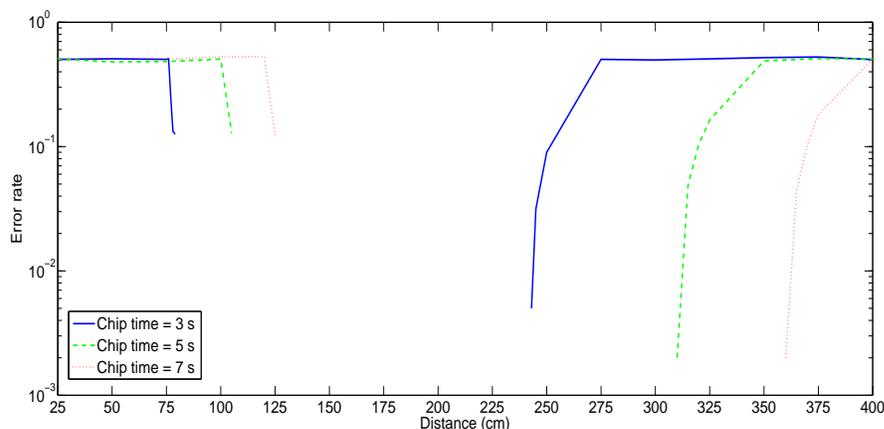
Figure 5.4: Error rate versus distance in data transmission with F = 15

We test the error rate versus distance for $F = 15$ in Fig 5.4. It seems the performance of $F = 15$ does not improve much over short distance ranges. As we can see, all the error rate curves for $T_C = 3$ s, 5 s, and 7 s are similar to the curves for $F = 7$ over short distance ranges (from $d = 25$ cm to 125 cm). However, the optimal range for $F = 15$ is longer than for $F = 7$. In the $F = 7$ curves for $T_C = 3$ s, 5 s, and 7 s, errors appear at $d = 185$ cm, 246 cm, and 298 cm respectively; whereas for $F = 15$, errors appear at $d = 243$ cm, 310 cm, and 360 cm. On the other hand, the error rates for $F = 15$ seem to have worse performance at long distance. Error rates reach high levels at $d = 275$ cm, 350 cm, and 400 cm respectively for $F = 15$. However, at these three points, the error rates with $F = 7$ are all lower. We believe chip sequences with length $F = 15$ are not suitable for data transmission over ranges longer than $d = 275$ cm. However, a chip length of

65

$F = 15$ has better performance than a chip length $F = 7$ over a short range.

5.3 Summary

The length of the chip sequence (F) affects the weight of the chip sequence (K), and it also influences the performance of the simulation. Longer chip sequence length, while keeping the number of users the same, reduces the error rate (this was discussed in Chapter 2). Since our number of users here is fixed at 2, theoretically, we can get a better performance with $F = 15$ compared with $F = 7$.

To summarize these results, we compare $F = 7$ to $F = 15$ head-to-head. We compared the average error rate at different distances d between $F = 7$ and $F = 15$ in Fig 5.5. Since the longest chip time is $T_C = 6.2$ s, in the curve for $d = 275$ cm and $F = 7$, we calculate the average error rates from $T_C = 1$ s to 6.2 s for all curves for specific distances. As we can see, the error rate of all three distances are lower at $F = 15$ compared with $F = 7$. This proves longer chip sequence length F will reduce the error rate, as we mentioned above. Moreover, error rates increase as distance increases. This makes sense, because the alcohol vapour cloud would dissipate during longer distance transmission, and there might not enough voltage change at the sensor to detect chips at the longer distance.

In Fig 5.6, we compare average error rates versus chip time T_C between the two chip sequence lengths. The results are calculated from $d = 25$ cm to 400 cm in each
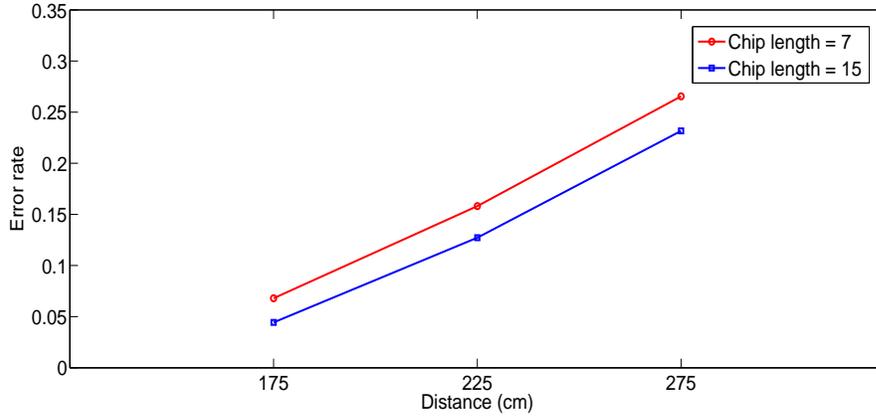
Figure 5.5: Error rate compare for chip sequence length versus distance

chip time. As we can see, longer chip times give us a lower error rate performance, which makes sense since longer chip times allow the sensor to go back to its original level prior to the next detection. This was shown in paper [12]. However, at chip time $T_C = 7$ s, the error rate at $F = 15$ is higher than $F = 7$. In addition, we find that the error rate differences between the two chip sequence lengths are not as clear as the last figure. The reason for this is due to the effect of longer distance at $F = 15$ compared with $F = 7$. We can find this in Fig 5.2 and Fig 5.4, and it shows that long chip sequence length does not obtain a low error rate at long distance (after $d = 275$ cm for chip time $T_C = 3$ s, $d = 350$ cm for $T_C = 5$ s, and $d = 400$ cm for $T_C = 7$ s).

In general, a larger F will lead to lower error rate, as has been shown in this chapter. Moreover, a longer distance from sprayer to sensor may cause an increase
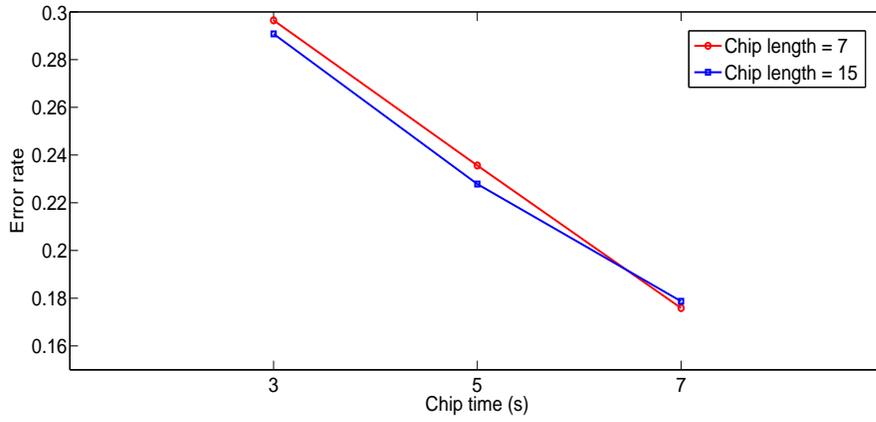
Figure 5.6: Error rate compare for chip sequence length versus Chip time

in error rate, while better performance comes with longer chip time. Overall, we

see that there are many choices of T_C and d that achieve very low error rates.

6 Conclusion and future work

In this thesis, we considered applying the OCDMA scheme to achieve multiple access in molecular communication. We used an OCDMA scheme using a two-to-two communication scenario (i.e., two sprayers to two sensors) to achieve multiple access. We also compared various parameters to obtain optimal performance.

Concerning our first contribution, we do both simulations and experiments involving one sprayer and one sensor, and involving two sprayers and two sensors. We show that the simulations and the experiments give similar performance, so that simulations can give us a faster way to reliably evaluate performance, compared with experiments.

Concerning our second contribution, where we analyze one-to-one and two-to-two communication scenarios using raw transmissions (not using OCDMA), and incorporating random noise in the transmission, we make two findings, one related to distance d, and one related to chip time T_C. First, we found that there is an optimal distance d, and we can get the best performance if we use this range. The

optimal range is roughly between $d = 100$ cm to 175 cm; moreover, this range moves closer to the sprayer if we reduce the chip time. Second, in terms of chip time, the error rates are all convex curves. This means it is not always true that the longer chip time is better. We showed that the error rate may increase as T_C increases beyond the optimal value.

Concerning our third contribution, where we applied the OCDMA scheme for multiple access data transmission for two users in the absence of noise, we show that longer chip sequence length F gives us better performance. However, this is also dependent on distance d and chip time T_c. Through simulation, we find values of d and T_c that lead to very low error rates.

We investigated the application of our OCDMA scheme from multiple sprayers to multiple sensors to achieve multiple access in molecular communication. Based on the results of our experiments and simulations, we show this scheme can achieve reliable communication and multiple access: we show that two users can communicate simultaneously with a low error rate.

For future work, we hope a platform for more simultaneously communicating users can be created and analyzed. Moreover, an improved algorithm could be obtained in the detector. As we mentioned in Fig 2.8, the algorithm can determine part of the signature sequence is missing, and we can still decide on a "1" bit successfully. Moreover, since flow is a significant factor in this work, we can improve

the communication environment to make the wind more smooth. We can try to seal the apparatus into a pipe to avoid uncertainty caused by wind. Furthermore, we can add more types of molecule to achieve molecule shift keying, which is will also improve the performance of communication because each sensor only needs to focus on its own molecules. We expect our work to advance the current state of molecular communication.

Bibliography

[1] M. Pierobon and I. F. Akyildiz, "Diffusion-based noise analysis for molecular communication in nanonetworks," *IEEE Transactions on Signal Processing*, vol. 59, no. 6, pp. 25322547, 2011.

[2] M. S. Kuran, H. B. Yilmaz, T. Tugcu, and I. F. Akyildiz, "Interference effects on modulation techniques in diffusion based nanonetworks," *Nano Communication Networks*, vol. 3, pp. 6573, Mar. 2012.

[3] H. B. Yilmaz, A. C. Heren, T. Tugcu, and C.-B. Chae, "Three-dimensional channel characteristics for molecular communications with an absorbing receiver," *IEEE Commuminications Letters*, vol. 18, no. 6, pp. 929932, 2014.

[4] K. V. Srinivas, A. Eckford, and R. Adve, "Molecular communication in fluid media: The additive inverse gaussian noise channel," *IEEE Transactions on Information Theory*, vol. 58, no. 7, pp. 46784692, 2012.

[5] H. ShahMohammadian, G. G. Messier, and S. Magierowski, "Nanomachine molecular communication over a moving propagation medium," *Nano Communication Networks*, vol. 4, no. 3, pp. 142153, 2013.

[6] A. Noel, K. Cheung, and R. Schober, "Optimal receiver design for diffusive molecular communication with flow and additive noise," *IEEE Transactions on NanoBioscience*, vol. 13, pp. 350362, Sept 2014.

[7] A. Enomoto, M. J. Moore, T. Suda, and K. Oiwa, "Design of selforganizing microtubule networks for molecular communication," *Nano Communication Networks*, vol. 2, no. 1, pp. 1624, 2011.

[8] N. Farsad, A. Eckford, S. Hiyama, and Y. Moritani, "On-Chip Molecular Communication: Analysis and Design," *IEEE Transactions on NanoBioscience*, 2012.

[9] N. Farsad, A. Eckford, and S. Hiyama, "A markov chain channel model for active transport molecular communication," *IEEE Transactions on Signal Processing*, vol. 62, pp. 24242436, May 2014.

[10] N. Farsad, W. Guo, A. W. Eckford, "Table-Top Molecular Communication: Text Messages Through Chemical Signals," *PLOS ONE, vol. 8, no. 12*, pp. 1-20 Dec, 2013.

[11] S. Qiu, W. Guo, S. Wang, N. Farsad, and A. Eckford, "A Molecular Communication Link for Monitoring in Confined Environments," *IEEE Conference on Computer Communications*, 2014.

[12] W. Linchen, N. Farsad, W. Guo, S. Magierowski, A. W. Eckford, "Molecular barcodes: Information transmission via persistent chemical tags," *IEEE International Conference on Communications*, 2015.

[13] H. ShahMohammadian, G. G. Messier, S. Magierowski, "Optimum receiver for molecule shift keying modulation in diffusion-based molecular communication channels," *Nano Communication Networks., Vol. 3, Issue. 3*, pp. 183195, Sep 2012.

[14] T. Nakano, A. W. Eckford, and T. Haraguchi, "Molecular Communication," *Cambridge University Press*, 2013.

[15] Y. Okaie, T. Nakano, M. J. Moore, and L. Jian-Qin, "Information transmission through a multiple access molecular communication channel," *IEEE ICC*, pp. 4030-4034, June, 2013.

[16] M. U. Mahfuz, D. Makrakis, and H. T. Mouftah, "A generalized strength-based signal detection model for concentration-encoded molecular communication," in*Proc. ICST*, pp. 461-467, 2013.

[17] A. W. Eckford, "Nanoscale Communication with Brownian Motion," *IEEE 41st Annual Conference on Information Sciences and Systems*, pp. 160-165, March, 2007.

[18] A. W. Eckford, "Achievable information rates for molecular communication with distinct molecules," in*Proc. BIONETICS-2007*, pp. 313-315, 2007.

[19] H. Tezcan, S. Oktug, and F. N. Kok, "Neural delay lines for TDMA based molecular communication in neural networks," *IEEE ICC*, pp. 6209-6213, June, 2012.

[20] L. P. Gin, and I. F. Akyildiz, "Molecular communication options for long range nanonetworks," *Computer Networks., vol. 53 Issue 16*, pp. 2753-2766, Nov, 2009.

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

[22] Y. Ruixiao, M. S. Leeson, and M. D. Higgins, "Multiple-access scheme optimisation for artificial neuronal networks," *IEEE CSNDSP*, pp. 428-433, July, 2014.

[23] A. Stok and E. H. Sargent, "The role of optical CDMA in access networks" *IEEE Communications Magazine., vol. 40*, pp. 83-87, Sept, 2002.

[24] J. A. Salehi, "Code division multiple-access techniques in optical fibernetworks. I. Fundamental principles," *IEEE Transactions Communications., vol. 37, no. 8*, pp. 824-833, Aug, 1989.

[25] J. A. Salehi and C. A. Brackett, "Code division multiple-access techniques in optical fibernetworks. II. Systems performance analysis," *IEEE Transactions Communications., vol. 37, no. 8*, pp. 834-842, Aug, 1989.

[26] J. A. Salehi and S. Mashhadi, "Code division multiple-access techniques in optical fibernetworks. part III: optical AND logic gate receiver structure with generalized optical orthogonal codes," *IEEE Transactions Communications., vol. 54, no. 8*, pp. 1457-1468, Aug, 2006.

[27] 7788 Shopping Online. Online: http://www.997788.com/21694/auction_560_11 -67875_0.html.

[28] J. Watson, "The tin oxide gas sensor and its applications," *Sensors and Actuators, vol. 5, no. 1*, pp. 29-42, 1984.

[29] Zhengzhou Winsen Electronics Technology Co. Ltd. of China, MQ3. Online: https://www.sparkfun.com/datasheets/Sensors/MQ-3.pdf.

[30] N. Farsad, N. R. Kim, A. W. Eckford, C. B. Chae, "Channel and Noise Models for Nonlinear Molecular Communication Systems," *IEEE, vol. 32, no. 12*, pp. 2392-2401, Dec, 2014.