MOLECULAR COMMUNICATION IN DIFFUSION BASED CHANNEL

by

Bayram Cevdet Akdeniz

B.S., Electronics and Communication Engineering, Yıldız Technical University, 2010

M.S., Electrical and Electronics Engineering, Boğaziçi University, 2013

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

Graduate Program in Electrical and Electronics Engineering

Boğaziçi University

2018

MOLECULAR COMMUNICATION IN DIFFUSION BASED CHANNEL

APPROVED BY:

Assoc. Prof. Ali Emre Pusane (Thesis Supervisor)	.....
Prof. Emin Anarım	.....
Prof. Tuna Tuğcu	.....
Assoc. Prof. Ertuğrul Başar	.....
Asst. Prof. Tunçer Baykaş	.....

DATE OF APPROVAL: 22.10.2018

ACKNOWLEDGEMENTS

First and foremost, I would like to express my sincere gratitude to my advisor Assoc. Prof. Ali Emre Pusane for accepting me as his student for the continuous support of my Ph.D study for his kindness, motivation, and immense knowledge. His guidance helped me in all the time of this research and he always behaves me as a friend rather than a student. Obviously no one can imagine a better advisor and mentor for a Ph.D study.

My sincere thanks also goes to Prof. Tuna Tuğcu who provided me an opportunity to join his Nanonetworking Research Group, and who gave access to the laboratory and research facilities. Without his technical and emotional support it would not be possible to conduct this research.

I would like to thank Prof. Emin Anarım for serving as a member of my thesis committee and their academic guidance. I am grateful to Assoc. Prof. Ertuğrul Başar and Assoc. Prof. Tuncer Baykaş for serving on my jury.

I am also grateful to thank Prof. Chan-Byoung Chae and his research team who invited me Yonsei University and give the opportunity to use their testbed. He and his team were great host and his guidance leads to improve my ideas.

My special thank goes to Dr. H.Birkan Yılmaz who has made significant contributions to my research. Being one of the most important researchers on the field, he always pointed me to the right direction and shared his invaluable ideas which have been extremely important for me and for my development. He will always be a important researcher that I would like to follow.

I should also thank Nanonetworking Research Group members who have proposed many great ideas that direct my research according to their suggestions. Our research meetings were not only efficient technically but also quite funny.

I thank my labmates in WCL lab for stimulating technical discussions and non-technical but very beneficial conversations. Their friendship leads me to stay in the lab as much as I can.

I would like to thank my girlfriend Müjde Aktaş for her continuous support in thesis and in life. I should also greatful to her family who regard me as a member of them.

Last but not the least, I would like to thank my family for supporting me spiritually throughout writing this thesis and my my life in general.

This thesis is partly supported by Bogazici University Research Fund (BAP) under project number 17A02D3, and by Scientic and Technical Research Council of Turkey (TUBITAK) under Grant number 116E916.

ABSTRACT

MOLECULAR COMMUNICATION IN DIFFUSION BASED CHANNEL

This thesis deals with diffusion based molecular communication. Unlike conventional communication systems, the information is encoded with molecules and these molecules are emitted by nanodevices in a diffusive environment. Since the molecules diffuse through the environment, their movement is governed by Brownian Motion, resulting in very slow and random movement that leads to excessive interference compared to conventional communication channels. In this thesis, various modulation, equalization and coding schemes are proposed to combat with this interference issue. The main motivation for the determination of these schemes is proposing computationally simple and/or sparse communication methods suitable for nanodevices. All proposed methods diminish the interference at the received signal and improve the performance of the molecular communication channels compared to the other existing methods in the literature. In particular, pulse position modulation is adopted for molecular communication channel to diminish interference without any additional complexity and less channel information compared to other proposed modulations. A specific sparse channel code for molecular communication is also proposed by presenting its improved performance. For equalization, the received signal is equalized by analytically determining the optimum reception delay analytically that minimizes interference and maximizes the signal power. With the same goal, optimum aperture region of the receiver is determined by deriving the joint angle-time distribution of the absorbed molecules at the receiver. Finally, network coding methods are proposed for two-way one-hop and multi-hop nanonetworks.

ÖZET

DİFÜZYON BAZLI KANALDA MOLEKÜLER HABERLEŞME

Bu çalışma difüzyon bazlı moleküler haberleşme ile ilgilidir. Moleküler haberleşmede bilgi, klasik haberleşme sistemlerinden farklı olarak, moleküller ile kodlanır ve bu moleküller nanomakineler aracılığı ile difüzyonlu ortama gönderilir. Moleküller ortamda dağılırken yaptıkları hareketler Brown hareketi ile modellenir. Yavaş ve rastgele olan bu hareket nedeniyle geleneksel haberleşme sistemlerine oranla çok daha fazla girişim meydana gelmektedir. Bu tezde, simgelerarası girişimi önlemek amacıyla çeşitli modülasyon, eşitleme ve kodlama teknikleri önerilmiştir. Önerilen yöntemlerdeki ana motivasyon nanomakineler için uygun olabilecek, düşük hesaplama işlemleri gerektiren ve olabildiğince az molekül gönderilmesini gerektiren çözümler sunmaktır. Sunulan bütün yöntemler girişimi azaltmakla birlikte literatürde önerilmiş olan diğer yöntemlerden daha başarılı performanslar vermektedir. İlk olarak girişimi engellemek için darbe konum modülasyonunun moleküler haberleşmeye uyarlanması ele alınmış ve bu uyarlama yapılırken diğer önerilmiş modülasyon tekniklerinden daha az kanal bilgisi kullanarak daha yüksek performans sergilediği gözlemlenmiştir. Buna ek olarak, moleküler haberleşme için düşük yoğunluklu kod tasarımı yapılmış ve litaratürdeki diğer yöntemlerden daha başarılı olduğu benzetimlerle gösterilmiştir. Eşitleme yöntemi olarak, alınan sinyalin en uygun gecikme süresi, sinyalin gücünü arttıracak, girişiminkini ise azaltacak şekilde belirlenmiştir. Benzer mantıkla alıcıdaki en uygun alma açıklığı, alıcıda alınan moleküllerin ortak açı-zaman olasılık yoğunluk fonksiyonu türetilerek bulunmuştur. Son olarak, ağ kodlaması yöntemi bir atlamalı ve çok atlamalı nanoağlara uygulanmıştır.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

B	Binomial Distribution
b_a	analytical parameter
b^i	bit sequence
b_{LS}	LS parameter
CW_n	Codeword Matrix whose codeword length is n
D	Diffusion coefficient
d_0	slot with minimum molecules
d_1	slot with maximum molecules
$F_{hit}(t)$	Total probability of a single molecule's arrival to the receiver until time t
$F_{\text{hit}}^{A^*}(t)$	Total probability of a single molecule's arrival to point A until time t
$F(\alpha, t)$	Cumulative joint angle-time distribution of the absorbed molecules
N	Number of transmitted molecules per bit
$\aleph$	Normal Distribution
P_E	Probability of error
p_k	channel taps
$p(\theta)$	Marginal angle distribution of the absorbed molecules
$p(\theta, t)$	Joint angle-time distribution of the absorbed molecules
r_0	Receiver-transmitter distance
r_r	Receiver radius
s	Transmitted symbol
t_s	Time slot
T_s	Time slot for nanonetworks
W(t)	Wiener Process
α	Cumulative reception angle variable
α^*	Optimum cumulative reception angle

θ	Reception angle variable
τ	optimum reception delay
τ_1	threshold
τ_a	adaptive threshold

LIST OF ACRONYMS/ABBREVIATIONS

1-D	one dimensional
3-D	three dimensional
BER	bit error rate
CSI	channel state information
CSK	Concentration Shift Keying
CW	Codeword
D-MoSK	Depleted molecular shift keying
ISI	inter-symbol interference
LS	Least Squares
MC	Molecular Communication
MCSK	Molecular concentration shift keying
MCvD	Molecular Communication via Diffusion
MIMO	Multiple input multiple output
MMSE	Minimum Mean Square Error
PPM	Pulse Position Modulation
Rx	Receiver
SER	symbol error rate
SID	Signal to Interference Difference
SIR	Signal to interference ratio
Tx	Transmitter

1. INTRODUCTION

Through billions of years of producing communication at small scales (i.e., distances of up to a few micro/nano meters), nature has provided, tested, and improved molecular communication (MC). Humans, on the other hand, struggle at this scale to utilize electromagnetic waves due to the constraints imposed by the ratio of the antenna size to the wavelength of the electromagnetic signal [2, 3]. As an alternative to electromagnetic signal, molecular signals have been proposed for nanonetworks in order to overcome the hurdles imposed by antenna constraint. There are other advantages to molecular communication– molecular signals are typically more bio-compatible and can reach an intended receiver within challenging environments even on the macro-scale such as pipelines, tunnels, and saline water environments [4]. Therefore, researchers direct their attentions to molecular communication via diffusion (MCvD) to achieve communication in nanonetworks.

Most of the existing research on MC has focused on channel modeling, interference mitigation, and modulation issues [5–8]. To address the challenges in a methodological and inclusive manner, IEEE has established the standardization group IEEE P1906.1 for MC.

One of the main challenges in MC is to develop valid channel models capable of representing a time-dependent received signal. For the receiver and the reception process in diffusion-based MC models, there are mainly two types of models– the passive and absorbing receivers. The former assumes the molecules are unaffected by the receiver while the latter assumes the molecules are absorbed whenever they hit the receiver. In the passive receiver case, the molecules can pass through the receiver node surface multiple times without interaction [7,9,10]. Therefore, the molecules are allowed to contribute to the received signal multiple times when the receiver is passive. For the absorbing receiver case, the molecules contribute to the received signal only once and the molecules that hit the receiver are removed from the environment [11–14]. This process is modeled by the first-passage process and to model the received signal [15].

In [11] and [12], the received molecular signal is modeled in a one-dimensional (1-D) environment with an absorbing receiver and the system performance is analyzed by utilizing the received signal model at the physical layer. In [13], a received signal model is introduced for a point transmitter and a spherical absorbing receiver in a 3-D environment. Since then, researchers have focused on modeling the received signal for an absorbing receiver while relaxing some of the assumptions. Instead of using a fully absorbing receiver, the authors have incorporated the receptor effect instead of using a fully absorbing receiver [16]. Similarly, researchers have utilized machine learning techniques to model the received signal for a spherical reflecting transmitter with single absorbing receiver [17] or multiple point transmitters with multiple absorbing receivers [18, 19]. In [20], the communication between a spherical receiver and a spherical transmitter in which the surface is covered with evenly-spaced point transmitters has been modeled and the channel impulse response has been presented.

1.1. Diffusion Model

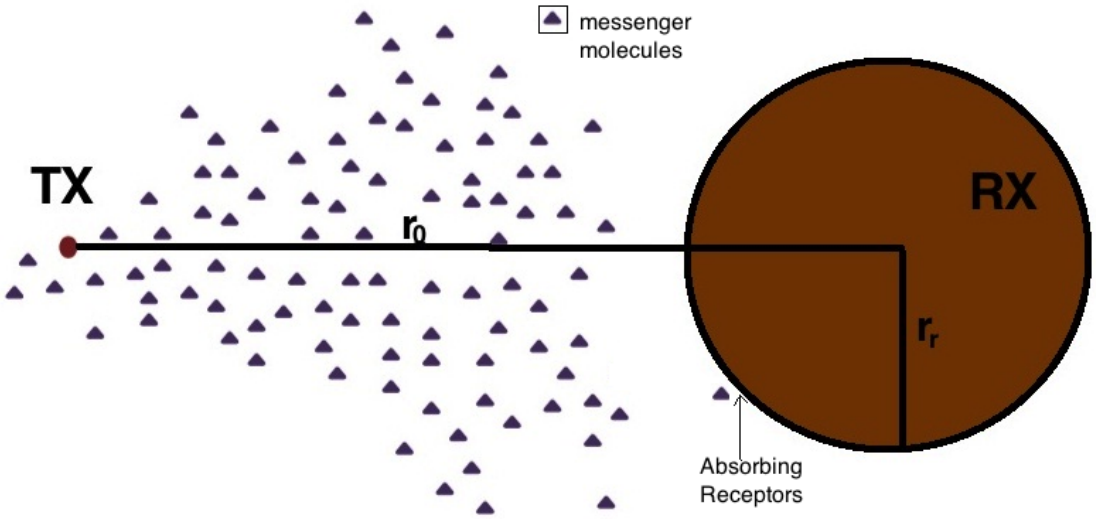


Figure 1.1. Channel model

One of the basic channel models for Molecular Communication via Diffusion (MCvD) is depicted in Figure 1.1. In this figure, r_0 denotes the distance from the transmitter to the center of the receiver, whose radius is r_r. The point transmitter emits molecules to convey its intended symbol to the receiver surrounded by absorbing receptors. The emitted molecules propagate subject to Brownian Motion, which is

described by the Wiener process [11]. The Wiener process $W(t)$ is characterized as follows:

- $W(0) = 0$,
- $W(t)$ is almost surely continuous,
- $W(t)$ has independent increments,
- $W(t_2) - W(t_1) \sim \mathcal{N}(0,\, c(t_2 - t_1))$ for $0 \leq t_1 \leq t_2$,

where $\mathcal{N}(\mu, \sigma^2)$ is the Gaussian distribution with mean μ and variance σ^2, which means the variance of random step size is a constant (c) times the time difference. Simulating the Brownian Motion includes consecutive steps in an n-dimensional space that obeys Wiener process dynamics. For an accurate simulation, time is divided into sufficiently small time intervals (Δt), and at each time interval the molecules take random steps in all dimensions. In an n-dimensional space, a random step is given as

$$\begin{aligned} \Delta\zeta &= (\Delta\zeta_1, ..., \Delta\zeta_n), \\ \Delta\zeta_i &\sim \mathcal{N}(0,\, 2D\Delta t) \;\; \forall i \in \{1, .., n\}, \end{aligned} \tag{1.1}$$

where $\Delta\zeta$, $\Delta\zeta_i$, and D correspond to the random displacement vector, the displacement at the i^{th} dimension, and the diffusion coefficient, respectively.

1.2. Major Challenges in Molecular Communications

In order to illustrate the major challenges in molecular communication, it is beneficial to examine first hitting time distribution ($f_{hit}(t)$) of a molecule. Let p_k denote the hitting probabilities that are obtained by integrating $f_{hit}(t)$ shown in Figure 1.2. In particular, p_1 is the hitting probability in the current symbol duration and p_k for $k \geq 2$ denote the hitting probabilities in the consecutive symbol durations that corresponds to ISI. Note that, hitting probabilities are sufficient to describe the characteristics of the diffusion channel completely, since they can be used to model the finite impulse channel response.

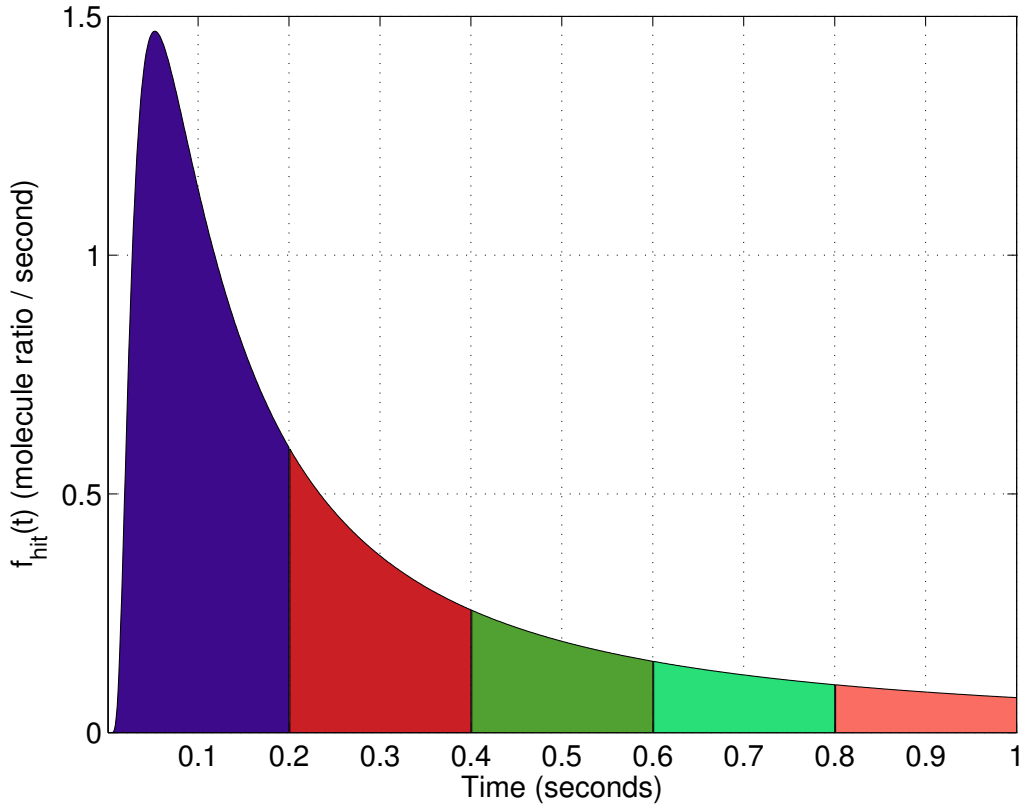


Figure 1.2. First hitting time of a molecule to the receiver

Let $\mathbf{b}^n = \{b_1, b_2, ..., b_n\}$ denote the binary message sequence of length n, and let b_i and N_i denote the message symbols and number of molecules sent from the transmitter in the i^{th} time slot for $i = 1, 2, ..., n$, respectively. Assuming $N_i = N$ for $b_i = 1$, and $N_i = 0$ for $b_i = 0$, when N is large enough, although the actual model is Binomial, the number of molecules induced at the receiver for a given time slot can be modelled as a Gaussian random variable as

$$\mathcal{B}(M, p) \sim \mathcal{N}(Np, Np(1-p)). \tag{1.2}$$

Let C_i denote the number of molecules induced at the receiver in the i^{th} time slot due to the transmission of $\mathbf{b}^i$. The probability model for C_i can be defined as

$$b_i \sim \mathcal{BE}(P[b_i = 1]), \tag{1.3}$$

$$C_i \mid \mathbf{b}^i \sim \mathcal{N}(\mu^{\{i\}}, \sigma^{2\{i\}}). \tag{1.4}$$

where $P[b_i = 1]$, and $P[b_i = 0]$ denote probability of occurrence for bit-1 and probability of occurrence for bit-0 in the message sequence, respectively.

Due to ISI, the expected number of molecules arriving at the receiver in the i^{th} time slot can be given as

$$E[C_i \mid \mathbf{b}^i] = \mu^{\{i\}} = N \sum_{k=1}^{i} p_k \, b_{i-k+1}, \tag{1.5}$$

which is the mean of the Gaussian distributed molecule count at the receiver. The variance of C_i is similarly given by

$$Var[C_i \mid \mathbf{b}^i] = \sigma^{2\{i\}} = N \sum_{k=1}^{i} p_k \, (1 - p_k) \, b_{i-k+1}. \tag{1.6}$$

Considering Equation (1.5) and (1.6), one can conclude that, the received signal is affected by previous symbols and due to the nature of diffusion, the mean and variance parameters of the received signal are also random. In order to reduce this randomness, we need to increase p_1 as much as possible and reduce the other p_ks. One trivial solution is keeping symbol duration as high as possible, but eventually this will reduce data rate further. Consequently, many recent works have focused on overcoming this issue by proposing either modulation, equalization or coding methods which will be

mentioned in the following chapters.

One of the other problems in molecular communication is communication range. As the distance between the receiver and the transmitter increases p_1 decreases drastically hence the other p_ks increases. Another main issue is, nanomachines that will be used for molecular communication may not be as capable as conventional communication sytems. Therefore, solutions to these challanges with lower computational complexity are preferred.

1.3. Contribution of the Thesis

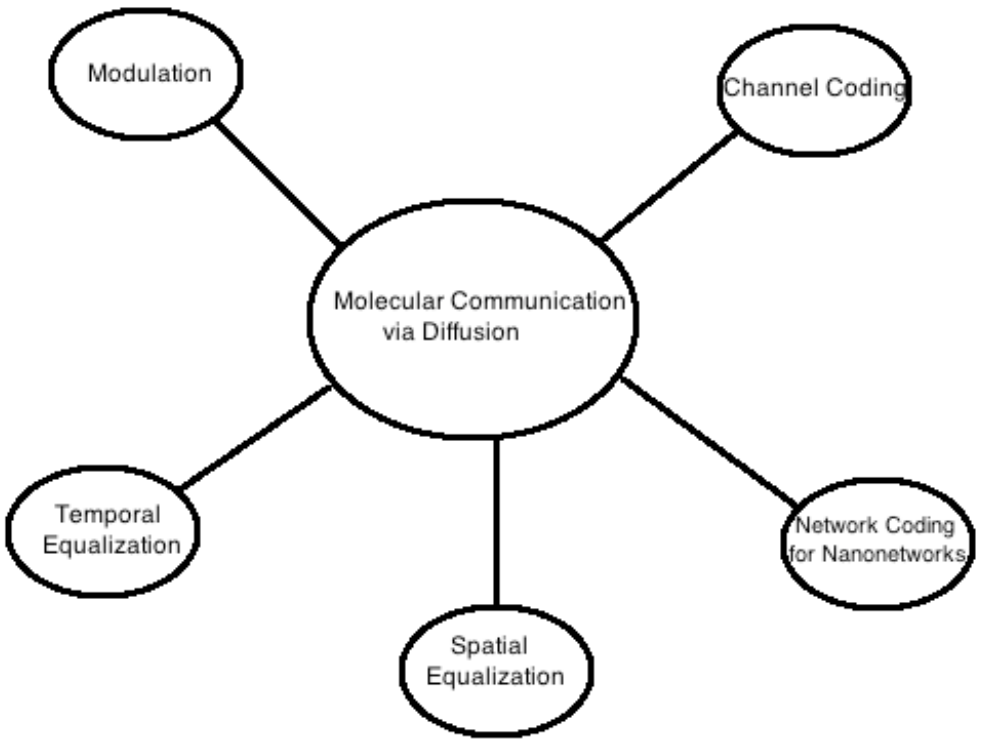


Figure 1.3. Contributions of the thesis

Considering the challenges in molecular communication we have proposed various solutions. The main contribution of the thesis can be summarized in Figure 1.3. As can be seen from this figure, we have proposed modulation, coding, equalization and network coding methods to increase the performance in molecular communication. In particular, we have proposed pulse position based modulation scheme for molecular

communication. We have built encoding and decoding schemes, derived analytical error bounds on probability of error and presented the improved performance. Afterwards, we have proposed a code family for molecular communication channel. The code family is constructed considering the properties of the molecular communication and due to the deconding and correcting stages and adaptive threshold the channel performance is significantly improved. We have also present analytical optimum reception delays at the receiver to shift the detected signal to reduce ISI and increase signal power optimally. In order to do that, we have proposed an objective function, signal to interference difference, which is a good and tractable alternative for bit error rate function in molecular communication systems. Thereafter, we have derived the joint angle-time distribution of the absorbed molecules by the receiver, which was not available in the literature before. Using this distribution, optimum aperture angle at the receiver is determined using the signal to interference difference objective function. Finally, network coding approaches have been developed for one-hop and multi-hop molecular nanonetworks.

1.4. Organization of the Thesis

Throughout the thesis, modulation, coding and equalization methods are proposed for diffusive molecular communication channels.

In particular, in Chapter 2, modulation problems in the literature for MCvD are presented by emphasizing the cons of these methods. Accordingly, we propose to use Pulse Position Modulation (PPM) in MCvD channel to overcome interference, complexity, and channel state information issues. It is justified with the Monte Carlo simulations that, the proposed method outperforms the other modulation methods in the literature by requiting less channel information.

Accordingly, we have proposed a sparse coding method for molecular communication channel in Chapter 3. The proposed coding method adopts sparsity and adaptive threshold, hence, outperforms the other proposed methods in the literature.

In Chapter 4, the optimal reception delay for the receiver is analytically derived by considering the received signal. It is shown that proposed method improve the performance of the any communication scheme in the literature..

In Chapter 5, a partially counting receiver is proposed by deriving the joint angle-time distribution of molecules at the receiver. Using this derivation the optimum aperture angle is determined and improved performance of the system is presented.

Chapter 6 examines the communication of nanomachines for longer ranges by proposing network coding approaches for one-hop and multi-hop networks and finally Chapter 7 concludes the dissertation.

2. MODULATION METHODS IN MOLECULAR COMMUNICATION

2.1. Related Literature

Although there are many different channel models in molecular communication, all of them have some common problems that need to be considered. Firstly, nanomachines are tiny devices in nano-scale or micro-scale and have limited computational capability. Therefore methods that has been proposed for EM communications can not be directly applied to these channels due to this limitation. Furthermore, the receiver can only absorb or observe a limited portion of the released molecules from the transmitter that represent a symbol due to diffusive property of the channel. Thus, other molecules will be taken after the period of the decoding of this symbol and this leads to Inter Symbol Interference (ISI) which is a fundamental problem that needs to be considered. Therefore, molecular communications schemes should be designed by taking into account ISI problem and computational complexity.

One of the most frequently used modulation techniques in molecular communications is concentration shift keying (CSK), where each information bit is encoded by a definite number of molecules and decoded based on a decision threshold [21]. In this method, due to the diffusive channel, the molecules are delayed in the channel, which results in ISI that leads to an error floor. There are many modulation and equalization techniques in the recent literature to cope with ISI in molecular communications.

In [22], using two different molecule types for two consecutive transmission symbols is proposed to eliminate ISI partially. This approach is known as molecular concentration shift keying (MCSK). Similarly, in [23], a modultion scheme called depleted Molecular Shift keying (D-MoSK) is proposed by using different types of molecules as orthogonal channels to reduce ISI. In [24], the authors propose to encode information with difference of the concentration of two molecules. In particular transmission

is achieved by sending A-type molecules followed by B-type molecules after a predetermined delay. The receiver decodes the symbol by the difference of the numbers of A-type and B-type molecules that are received in the symbol slot. The release-time and number of B-type molecules are optimized to maximize the signal to interference ratio (SIR).

On the other hand, using different types of molecules can be considered as an additional cost, and the receiver needs to have two different types of receptors, which increases overall system complexity. Furthermore, all these works assume that channel state information (CSI) is known by the receiver and the transmitter hence optimum demodulation threshold can be easily determined. However, CSI may not always possible can be time-varying. It these cases, proposed methods in the literature do not give effective solutions.

Pulse position modulation (PPM) is a widely used modulation method in conventional communications, where $\log_2 M$ bits are encoded by the position of a pulse in M candidate positions. In the molecular communications literature, PPM has already been utilized in [25] and [26]. In these papers, 1 bit is encoded and decoded using 2-PPM and its comparison with CSK is performed. However, it is concluded that CSK is more suitable than 2-PPM in molecular communications. For fair comparison, in 2-PPM, the time slot for each pulse should be the half of the time slot in CSK, which leads to more severe ISI. Furthermore, in 2-PPM, molecules are released for each symbol transmission, while in CSK, molecules are only released when the intended symbol is 1. Therefore, equal power consumption constraint, the number of molecules per symbol, should be half of the number of molecule used in CSK to convey bit-1. All of these studies conclude that 2-PPM is an inefficient alternative for CSK.

On the other hand considering the problems in MCvD, PPM can still be a good alternative to cope with the problems mentioned before since it does not require any additional computational complexity than CSK does, there is no need to know CSI and it is robust for time varying channels. We, therefore, propose to employ position based modulation for different M values together with a simple decoding algorithm

to determine the intended symbol. For M-PPM, the required number of time slots for a single transmission is reduced to $\frac{\log_2 M}{M}$ of the reference time slot used in CSK. On the other hand, the number of molecules transmitted per bit is also reduced, compared to that of CSK, as M increases. Therefore higher order PPMs may have better performance compared to 2-PPM and CSK due to their power advantages. The decoding algorithm is based on determining the time position, where maximum number of molecules is observed by the receiver. In particular, the receiver counts the number of absorbed molecules at each of the M candidate positions and determines the maximum one to decode the corresponding symbol. This receiver has two major advantages when compared with other receivers commonly employed in the molecular communications literature. Firstly, it is computationally simple to find the maximum among M integers. Secondly, since thresholding is not necessary for decoding, the receiver does not need to know the channel state information (CSI) except for the knowledge of the maximum channel tap. We derive the theoretical error bounds on PPM modulations in molecular communications for $M = 2, 4, 8, 16$, then compare it with CSK and other proposed and successful methods in the literature. The simulation results show that, PPM can outperform not only CSK but other proposed methods in the literature with less channel information usage.

2.2. System Model

In this work, the proposed modulation scheme is assumed to operate in a 3-dimensional (3-D) diffusive channel without drift that involves a point transmitter and a spherical receiver with fully absorbing capability as shown in Figure 1.1. Although the proposed modulations are tested on this system model, they do not impose any restriction on the channel model.

During transmission, N molecules are released to represent a pulse and no molecule is released to represent no pulse case. Due to the nature of the diffusive channel model, only a limited portion of the molecules reach the receiver. In fact, the fraction of

molecules that reach the receiver until time t is given by [27]

$$F_{hit}(t) = \frac{r_r}{r_0} \operatorname{erfc}\left[\frac{r_0 - r_r}{\sqrt{4Dt}}\right], \tag{2.1}$$

where D is the diffusion coefficient and is dependent on temperature, viscosity of the fluid, and Stoke's radius of the molecules.

Let t_{slot} be the slot duration that corresponds to the time required for the transmission and the decision of a pulse. Then, channel tap, p_k, can be computed as

$$p_k = F_{hit}\left(kt_{slot}\right) - F_{hit}\left(\left(k-1\right)t_{slot}\right). \tag{2.2}$$

Note that p_1 is the signal tap that determines the number of molecules arriving in the current symbol while the other taps correspond to molecules contributing to ISI.

2.3. Position-based Modulation

Let $\bar{a}_i$ be a $\log_2 M$-bits long bit stream that represents i^{th} transmitted symbol s_i where $s_i \in \{0, 1, 2...M-1\}$. In conventional binary CSK modulation, these bits are decoded one by one. In PPM, $\log_2 M$ bits are decoded together by determining the position of the pulse, i.e. released molecules. In particular, each symbol can be visualized as an M-bit length sequence, denoted by $\bar{x}_i$, that involves strictly one bit-1 and $M-1$ bit-0s. The bit-1 is located at position k, corresponding to the release of N molecules at the beginning of the corresponding time slot kt_{slot} where $k \in \{0, 1, 2...M-1\}$. At the receiver side, the receiver counts the number of the molecules arriving during each time slot and decides when these N molecules were released, i.e., what the value of k is.

The encoding scheme of the proposed modulation method for $M = 8$ is given in Table 2.1. Once the 3-bit binary information $\bar{a}_i$ is obtained from symbol s_i, the corresponding 8-bit sequence $\bar{x}_i$ is constructed by releasing molecules at corresponding time slot.

Table 2.1. Encoding and Modulation Scheme for 8-PPM (rightmost bit is the least significant bit)

Symbol(s_i)	Binary($\bar{a}_i$)	Encoding($\bar{x}_i$)	Release time(kt_{slot})
0	000	00000001	0
1	001	00000010	t_{slot}
2	010	00000100	$2t_{slot}$
3	011	00001000	$3t_{slot}$
4	100	00010000	$4t_{slot}$
5	101	00100000	$5t_{slot}$
6	110	01000000	$6t_{slot}$
7	111	10000000	$7t_{slot}$

2.4. Advantages of PPM in Molecular Communication

In [25], it is shown that CSK outperforms 2-PPM. There are two main reasons for this observation. Firstly, in 2-PPM, each bit is represented using two time slots while in CSK each bit is represented by one time slot. Therefore, for a fair comparison, the chosen time slot of CSK should be twice of the one used for 2-PPM. Another major reason is the average number of molecules per bit in 2-PPM, which is twice of that in CSK (In CSK, molecules are only released when symbol is bit-1 however in 2-PPM molecules are released for both cases). This can be also observed in Table 2.2. In this table, in order to avoid confusion, two different terms are defined: t_{slot} and t_s, which are the duration of a single slot for a pulse transmission and the average duration to transmit one-bit message to the receiver, respectively. In particular, in CSK t_{slot}=t_s while in 2-PPM, since 2 slots are used to transmit each bit, t_{slot}=0.5 t_s. Although the observations in [25] are reasonable, there are some additional points that need to be considered. The major problem in CSK is the accumulation of molecules that leads to an error floor. Especially, after the transmission of consecutive bit-1s, the decoding of a bit-0 can be potentially erroneous, due to the accumulated molecules, by exceeding the decision threshold. However in PPM, there is no threshold usage and ISI coming from the previous symbols are distributed to the slots of the PPM symbols, hence their

Table 2.2. Slot durations and number of molecules used for CSK and PPM modulations for equal power and data-rate constraints

Modulation:	CSK	2-PPM	4-PPM	8-PPM	16-PPM
bits per symbol	1	1	2	3	4
slot per symbol	1	2	4	8	16
slot duration (t_{slot})	t_s	$\frac{1}{2}\,t_s$	$\frac{1}{2}\,t_s$	$\frac{3}{8}\,t_s$	$\frac{1}{4}\,t_s$
bit duration (t_s)	t_s	t_s	t_s	t_s	t_s
Molecules per symbol	2N for bit-1 0 for bit-0	N	2N	3N	4N
Molecules per bit (N)	N	N	N	N	N

effects on decoding are reduced. Since the decoder aims to identify the slots that involve maximum number of molecules, depending on the previously transmitted symbols, ISI may even have constructive effect or relatively less destructive effect, compared to CSK.

Another advantage of PPM modulations can be deduced from Table 2.2. For higher-level PPM, since the number of transmitted bits per symbol increases and since each PPM symbol involves only one pulse, under the equal molecule per bit constraint, the number of allowed molecules to represent a pulse is increased compared to CSK. Therefore, for a fair comparison, higher order PPM modulations can use more molecules than CSK to represent a pulse.

The last but not least advantage of PPM is that the receiver does not need to know the CSI, except for the knowledge of the time of the strongest tap. Since the receiver needs to find the slot with the maximum number of molecules, no threshold calculations is necessary. Hence, CSI at the receiver is not needed like many other proposed methods in the literature, except for the location of the strongest channel tap. All of these reasons may give an insight to the use of PPM modulation in MCvD.

2.5. Decoding when the signal tap is maximum

We assume that the receiver fully absorbs and counts the number of molecules for each of the M-slot duration that corresponds to entries of $\bar{x}_i$. Assuming that p_1, the intended signal tap, is greater than all other taps, the receiver aims to find the slot that involves the maximum number of molecules and to decode the intended symbol s_i. In particular, let $\bar{y}_i$ be a vector of length M that represents the number of received molecules for the i^{th} intended symbol and let y_i^k be the received number of molecules at the k^{th} slot of the i^{th} intended symbol s_i hence k^{th} element of $\bar{y}_i$. When N molecules are released from the transmitter at k^{th} slot to transmit s_i, the average number of received molecules $E[y_i^k]$ at k^{th} slot of the i^{th} symbol x_i^k (i.e, k^{th} element of x_i) by the receiver can be obtained by considering molecules coming from the previous and current symbols as

$$E[y_i^k] = N \sum_{j=0}^{Mi+k} x_{i+*\frac{k-j}{M}}^{mod(k-j,M)} p_{j+1} = \mu_i^k, \tag{2.3}$$

where $mod(.)$ is modulo operation, $k \in \{0, 1, 2...M-1\}$ and $i \in \mathbb{Z}_{\leq 0}$. When N is large enough, y_i^k can be approximated with a Gaussian distribution with mean μ_i^k and variance $\sigma_{i,k}^2 = N \sum_{j=0}^{Mi+k} x_{i+*\frac{k-j}{M}}^{mod(k-j,M)} p_{j+1}(1 - x_{i+*\frac{k-j}{M}}^{mod(k-j,M)} p_{j+1})$. Therefore, for each of symbol, the receiver observes an M-length y_i, which can be considered as a random vector. Once y_i is obtained by observing each y_i^k, $\hat{s}_i$ can be obtained as

$$\hat{s}_i = \arg \max_{k \in \{0,1,2,...,M-1\}} y_i^k. \tag{2.4}$$

2.6. Decoding when the signal tap is not the maximum

The decoding scheme mentioned in Section 2.5 is valid for the assumption that p_1 is the maximum among all channel taps. As the data rate increases (i.e., t_s, hence t_{slot}, decreases), the signal tap p_1 becomes lower and lower, resulting in one of the following taps to be the maximum. Therefore, the decoding method presented above does not work properly and needs modification. Let p_x be the highest tap of the channel where $x > 1$. Then, the corresponding decision is done using using z_i, which is $x-1$ times right shifted version of y_i. For example, assuming $x = 2$, the corresponding z_i is obtained as z_i=$\begin{bmatrix} y_i^1 & y_i^2 & y_i^3 & y_i^4 ... y_i^{M-1} & y_{i+1}^0 \end{bmatrix}$. Let z_i^k be the k^{th} element of z_i where $k \in \{0, 1, 2...M-1\}$. Then, the decoding of s_i can be obtained as

$$\hat{s_i} = \arg \max_{k \in \{0,1,2...M-1\}} z_i^k. \tag{2.5}$$

It is important to note that such a change in the decoding algorithm only requires the knowledge of p_x i.e the time of the maximum channel tap.

2.7. Channel Properties

In all molecular communications methods, the decrease in t_{slot} leads to a decrease in the performance of the system directly. As mentioned in Section 2.6, the decoding scheme can be adjusted even if the signal tap is not the maximum tap. Considering these adjustments, one can wonder about the worst-case performance in this channel for a given t_{slot} interval. The answer is found by considering the decoding scheme of this modulation presented in Equation (2.4). Since decoding depends on finding the slot that involves the maximum number of molecules, the worst case performance in a given t_{slot} interval occurs when the highest taps are equal to each other. Therefore, in PPM, unlike other modulations, a decrease in t_{slot} may lead to an incremental effect on the performance, if this decrement increases the difference of the maximum taps.

In order to demonstrate this idea, we heuristically chose objective function as the sum of the individual differences of the highest second and third tap from the maximum tap and this function is plotted in Figure 2.1. In other words, let p_x be the maximum tap while p_y and p_z are the second and third highest taps, respectively. Therefore, the objective function plotted in Figure 2.1 is p_x-p_y + p_x - p_z. In this figure, r_0 and r_r are fixed as 10×10^{-6}m and 5×10^{-6}m, respectively, to examine the effect of slot duration t_{slot} to the channel. Note that the objective function is chosen heuristically and may also be chosen to maximize the difference between the first highest two taps as p_x-p_y.

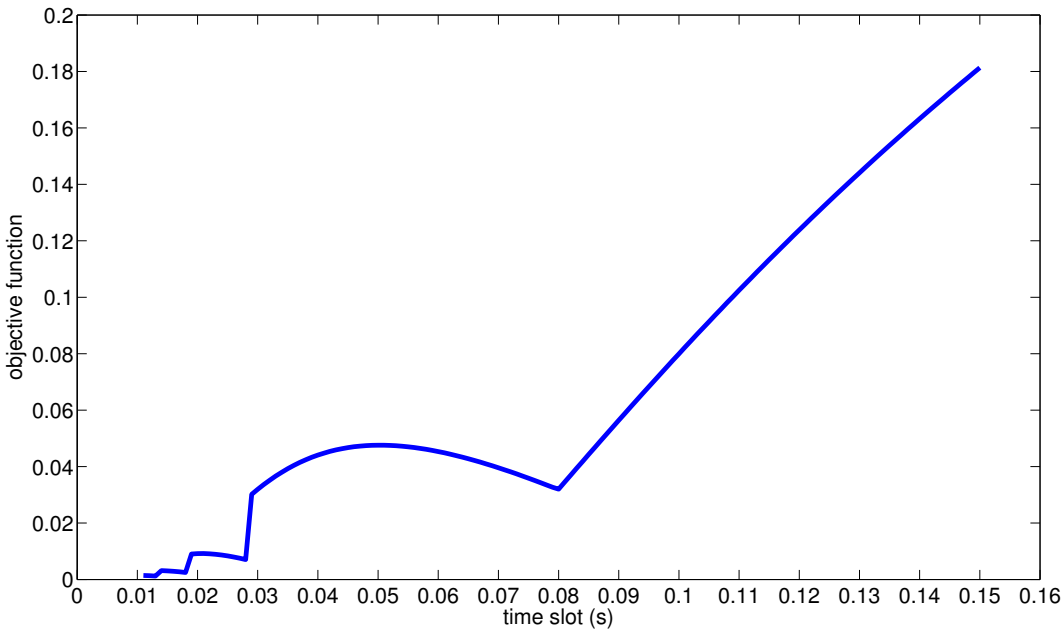


Figure 2.1. The objective function obtained by using the three highest taps for a given t_s

From Figure 2.1, it can be observed that there are local minima points for the channel (such as around t_{slot} =0.08s and t_{slot} =0.029s) that minimize the objective function. At t_{slot}=0.08s, it can be obtained from Equation (2.2) that the first two channel taps are highest and approximately equal and at t_{slot} =0.029s, the second and the third ones are the two highest and approximately equal taps. For the given specifications above, it is expected that the system will have greater performance around t_{slot} =0.05s than t_{slot} =0.08s, when PPM is employed. Furthermore, there is an optimum t_{slot} between any two local minima that makes the two highest taps equal. Therefore,

an optimization problem arises between any two local minima points. We only consider the interval between $t_{slot}^{(1,2)}$ and $t_{slot}^{(2,3)}$. $t_{slot}^{(1,2)}$ and $t_{slot}^{(2,3)}$ are t_{slot} values that make p_1 and p_2 the two highest and equal taps, and make p_2 and p_3 equal and the highest taps, respectively. In this interval, p_2 is the maximum tap while p_1 and p_3 are the second and third highest taps (in any order). Using these facts, the optimization problem can be written as

$$t_{slot}^{*} = \arg \max_{t_{slot}^{(2,3)} < t_{slot} < t_{slot}^{(1,2)}} p_2 - p_1 + p_2 - p_3. \tag{2.6}$$

Converting the channel taps into explicit form using Equation (2.2), we have

$$t_{slot}^{*} = \arg \max_{t_{slot}^{(2,3)} < t_{slot} < t_{slot}^{(1,2)}} 3F_{hit}(2t_{slot}) - 3F_{hit}(t_{slot}) - F_{hit}(3t_{slot}). \tag{2.7}$$

In order to find the optimum t_{slot}^{*} for given interval, the derivative of the objective function is taken and equated to 0 as

$$\frac{3 \quad e^{-\frac{m^2}{2t_{slot}^{*}}}}{\sqrt{2}} - \frac{e^{-\frac{m^2}{3t_{slot}^{*}}}}{\sqrt{3}} - 3e^{-\frac{m^2}{t_{slot}^{*}}} = 0, \tag{2.8}$$

where $m = \frac{r_0 - r_r}{\sqrt{4D}}$. The solution of the Equation 2.8 is obtained as

$$t_{slot}^{*} = 0.693m^2. \tag{2.9}$$

Therefore, for the given specifications of the channel, r_0, r_r, and D (or the m value which is obtained by these specifications), it is possible to obtain the optimum value

of t_{slot} between $t_{slot}^{(1,2)}$ and $t_{slot}^{(2,3)}$. Same optimization procedure can be done for other intervals.

2.8. Probability of Error

An upper bound on the probability of error for PPM is derived. Assuming that k is the correct index of the i^{th} symbol, y_i^k should be greater than the number of observed molecules in the other $M-1$ transmission slots for correct decoding. In order to derive a general formula, the distribution of y^k, which corresponds to the averaged version of the y_i^k, should be considered instead of y_i^k. The probability of error for 4-PPM is derived assuming that p_1 is the maximum tap, but these calculations can be easily adapted for other modulations and cases. The general bound on PPM modulation is also derived.

In order to find the probability of error for 4-PPM, the probability of error for four different symbols should be considered. Let $P\left(e|s=0\right)$, $P\left(e|s=1\right)$, $P\left(e|s=2\right)$, and $P\left(e|s=3\right)$ be the probabilities of errors for 4 individual symbols in 4-PPM. Assuming equally likely transmission, overall probability of symbol error for 4-PPM (P_{4ppm}^{E}) is obtained as

$$P_{4ppm}^{E} = \frac{1}{4}(P\left(e|s=0\right) + P\left(e|s=1\right) + P\left(e|s=2\right) + P\left(e|s=3\right)). \tag{2.10}$$

2.8.1. Probability of Error when s=0

Probability of error when the intended symbol is $s=0$ is obtained by considering the distribution of the four slots in 4-PPM. Let y^k be the distribution of the slots when the intended symbol $s=0$ where $k=0,1,2,3$. Since the correct decoding is done when y^0 is the maximum, an error occurs when y^0 is not the highest value among four

random variable. Therefore, $P\left(e|s=0\right)$ can be obtained as

$$\begin{aligned} P\left(e|s=0\right) &= P\left(y^0 < y^1\right) + P\left(y^0 < y^2\right) + \\ &P\left(y^0 < y^3\right) - P\left(y^0 < y^1, y^2\right) - P\left(y^0 < y^1, y^3\right) \\ &- P\left(y^0 < y^2, y^3\right) + P\left(y^0 < y^1, y^2, y^3\right). \end{aligned} \tag{2.11}$$

Since in Equation (2.11), most of the contribution comes from the first three terms, $P\left(e|s=0\right)$ can be upper bounded as

$$P\left(e|s=0\right) < P\left(y^0 < y^1\right) + P\left(y^0 < y^2\right) + P\left(y^0 < y^3\right). \tag{2.12}$$

The distribution of y^k can be obtained by considering the molecules related to the current symbol in addition to the molecules coming from the previous symbols. For instance, when N molecules are released at the beginning of the first slot, the number of molecules at this slot y^0 can be modelled with a Gaussian random variable whose mean and variance are Np_1 and $Np_1(1-p_1)$, respectively as proposed in [28]. Similarly, y^1, y^2, and y^3 consist of Gaussian random variables whose means are Np_2, Np_3, and Np_4, and variances are $Np_2(1-p_2)$, $Np_3(1-p_3)$, and $Np_4(1-p_4)$, respectively.

In addition to the molecules coming from the current symbol, we need to take into account the molecules coming from the previous symbols. For example, the first slot y^0 involves an additional random variable due to previous the symbol with possible means Np_2, Np_3, Np_4, or Np_5, depending on the position of the released molecules of the previous symbols. Therefore, the mean of the number of molecules coming from the previous symbol can be averaged as $\frac{N}{4}\sum_{i=2}^{5} p_i$. Similarly, for one earlier previous sym-

bol, it can be written as $\frac{N}{4}\sum_{i=6}^{9} p_i$, and so on. To sum up, assuming that the channel finite impulse response has K taps, the distribution of ISI at the first slot is distributed as $\mathcal{N}(\frac{N}{4}\sum_{i=2}^{K} p_i, \frac{N}{4}\sum_{i=2}^{K} p_i(1-p_i))$. Therefore, using the independence assumption, the distribution of y^0 can be obtained as adding the two independent Gaussian distributions that come from the signal and ISI as $\mathcal{N}(Np_1 + \frac{N}{4}\sum_{i=2}^{K} p_i, Np_1(1-p_1) + \frac{N}{4}\sum_{i=2}^{K} p_i(1-p_i))$. In order to write the distributions of y^ks in a more compact form, let $y^0 \sim \mathcal{N}(Np_1 + \frac{N}{4}\sum_{i=2}^{K} p_i, Np_1(1-p_1) + \frac{N}{4}\sum_{i=2}^{K} p_i(1-p_i)) \equiv \mathcal{N}(\mu_0, \sigma_0^2)$.

Using the same approach, the distribution of the other slots can be obtained as

- $y^0 \sim \mathcal{N}\left(Np_1 + \frac{N}{4}\sum_{i=2}^{K} p_i, Np_1(1-p_1) + \frac{N}{4}\sum_{i=2}^{K} p_i(1-p_i)\right) \equiv \mathcal{N}\left(\mu_0, \sigma_0^2\right)$
- $y^1 \sim \mathcal{N}\left(Np_2 + \frac{N}{4}\sum_{i=3}^{K} p_i, Np_2(1-p_2) + \frac{N}{4}\sum_{i=3}^{K} p_i(1-p_i)\right) \equiv \mathcal{N}\left(\mu_1, \sigma_1^2\right)$
- $y^2 \sim \mathcal{N}\left(Np_3 + \frac{N}{4}\sum_{i=4}^{K} p_i, Np_3(1-p_3) + \frac{N}{4}\sum_{i=4}^{K} p_i(1-p_i)\right) \equiv \mathcal{N}(\mu_2, \sigma_2^2)$
- $y^3 \sim \mathcal{N}\left(Np_4 + \frac{N}{4}\sum_{i=5}^{K} p_i, Np_4(1-p_4) + \frac{N}{4}\sum_{i=5}^{K} p_i(1-p_i)\right) \equiv \mathcal{N}(\mu_3, \sigma_3^2)$.

Let $y^{0,k}$=y^0-y^k for $k \neq 0$. Since all y^ks are independent Gaussian random variables, $y^{0,k}$ is distributed as $y^{0,k} \sim \mathcal{N}(\mu_0 - \mu_k, \sigma_0^2 + \sigma_k^2) \equiv \mathcal{N}(\mu_{0,k}, \sigma_{0,k}^2)$. Using $y^{0,k}$s, Equation (2.12) can be rewritten as

$$P\left(e|s=0\right) < P\left(y^{0,1} < 0\right) + P\left(y^{0,2} < 0\right) + P\left(y^{0,3} < 0\right). \tag{2.13}$$

The expression in Equation (2.13) can be rewritten with Q functions as

$$P\left(e|s=0\right) < Q\left(\frac{\mu_{0,1}}{\sigma_{0,1}}\right) + Q\left(\frac{\mu_{0,2}}{\sigma_{0,2}}\right) + Q\left(\frac{\mu_{0,3}}{\sigma_{0,3}}\right). \tag{2.14}$$

2.8.2. Probability of Error when s=1

Following the same steps, the probability of error for $s = 1$ can be written as

$$P\left(e|s=1\right) < P\left(y^1 < y^0\right) + P\left(y^1 < y^2\right) + P\left(y^1 < y^3\right). \tag{2.15}$$

When the intended symbol is $s = 1$, y^ks are distributed as

- $y^0 \sim \mathcal{N}\left(\frac{N}{4}\sum_{i=2}^{K} p_i, \frac{N}{4}\sum_{i=2}^{K} p_i(1-p_i)\right) \equiv \mathcal{N}(\mu_0, \sigma_0^2)$
- $y^1 \sim \mathcal{N}\left(Np_1 + \frac{N}{4}\sum_{i=3}^{K} p_i, Np_1(1-p_1) + \frac{N}{4}\sum_{i=3}^{K} p_i(1-p_i)\right) \equiv \mathcal{N}(\mu_1, \sigma_1^2)$
- $y^2 \sim \mathcal{N}\left(Np_2 + \frac{N}{4}\sum_{i=4}^{K} p_i, Np_2(1-p_2) + \frac{N}{4}\sum_{i=4}^{K} p_i(1-p_i)\right) \equiv \mathcal{N}(\mu_2, \sigma_2^2)$
- $y^3 \sim \mathcal{N}\left(Np_3 + \frac{N}{4}\sum_{i=5}^{K} p_i, Np_3(1-p_3) + \frac{N}{4}\sum_{i=5}^{K} p_i(1-p_i)\right) \equiv \mathcal{N}(\mu_3, \sigma_3^2)$

As in the $s = 0$ case, we can define $y^{1,k}$=y^1-y^k for $k \neq 1$ and $y^{1,k} \sim \mathcal{N}(\mu_{1,k}, \sigma_{1,k}^2)$. Using the distribution of the $y^{1,k}$s, $P\left(e|s=1\right)$ can be written as

$$P\left(e|s=1\right) < Q\left(\frac{\mu_{1,0}}{\sigma_{1,0}}\right) + Q\left(\frac{\mu_{1,2}}{\sigma_{1,2}}\right) + Q\left(\frac{\mu_{1,3}}{\sigma_{1,3}}\right). \tag{2.16}$$

Comparing the distributions of the y^ks for cases $s = 0$ and $s = 1$, one can observe that, the distributions that come from the previous symbols remain same for each slot, and the only difference comes from the current symbol. In $s = 0$, y^k involves the distributions that come from the current symbol whose means are Np_1, Np_2, Np_3, and Np_4, respectively. In $s = 1$, since molecules are released in the second slot, the first slot does not involve any distribution from current symbol and starting from y^1, other slots involve Np_1, Np_2, and Np_3, respectively. The distributions for $s = 2$ and $s = 3$ can be easily derived using the same concept, hence they are skipped here and presented in the Appendix to avoid repetition. The probability of errors for $s = 2$ and $s = 3$ is implicitly written as

$$P\left(e|s=2\right) < P\left(y^2 < y^1\right) + P\left(y^2 < y^3\right) + P\left(y^2 < y^4\right) \tag{2.17}$$

and

$$P\left(e|s=3\right) < P\left(y^3 < y^0\right) + P\left(y^3 < y^1\right) + P\left(y^3 < y^2\right). \tag{2.18}$$

2.8.3. Probability of Error when s=2

When the intended symbol $s = 2$, the probability of error can be bounded as

$$P\left(e|s=2\right) < P\left(y^2 < y^1\right) + P\left(y^2 < y^3\right) + P\left(y^2 < y^4\right). \tag{2.19}$$

The corresponding distribution of y^ks can be listed as

- $y^0 \sim \mathcal{N}\left(\frac{N}{4}\sum_{i=2}^{K} p_i, \frac{N}{4}\sum_{i=2}^{K} p_i(1-p_i)\right) \equiv \mathcal{N}(\mu_0, \sigma_0^2)$
- $y^1 \sim \mathcal{N}\left(\frac{N}{4}\sum_{i=3}^{K} p_i, \frac{N}{4}\sum_{i=3}^{K} p_i(1-p_i)\right) \equiv \mathcal{N}(\mu_1, \sigma_1^2)$

- $y^2 \sim \mathcal{N}\left(Np_1 + \frac{N}{4}\sum_{i=4}^{K} p_i, Np_1(1-p_1) + \frac{N}{4}\sum_{i=4}^{K} p_i(1-p_i)\right) \equiv \mathcal{N}(\mu_2, \sigma_2^2)$
- $y^3 \sim \mathcal{N}\left(Np_2 + \frac{N}{4}\sum_{i=5}^{K} p_i, Np_2(1-p_2) + \frac{N}{4}\sum_{i=5}^{K} p_i(1-p_i)\right) \equiv \mathcal{N}(\mu_3, \sigma_3^2)$.

Let $y^{2,k}$=y^2-y^k for $k \neq 2$ and $y^{2,k} \sim \mathcal{N}(\mu_{2,k}, \sigma_{2,k}^2)$. Using the distribution of the $y^{2,k}$s, $P\left(e|s=2\right)$ can be written as

$$P\left(e|s=2\right) < Q\left(\frac{\mu_{2,0}}{\sigma_{2,0}}\right) + Q\left(\frac{\mu_{2,1}}{\sigma_{2,1}}\right) + Q\left(\frac{\mu_{2,3}}{\sigma_{2,3}}\right). \tag{2.20}$$

2.8.4. Probability of Error when s=3

As in the other cases, the probability of error when $s=3$ is bounded as

$$P\left(e|s=3\right) < P\left(y^3 < y^0\right) + P\left(y^3 < y^1\right) + P\left(y^3 < y^2\right). \tag{2.21}$$

The distribution of y^ks for this case can be written as

- $y^0 \sim \mathcal{N}\left(\frac{N}{4}\sum_{i=2}^{K} p_i, \frac{N}{4}\sum_{i=2}^{K} p_i(1-p_i)\right) \equiv \mathcal{N}(\mu_0, \sigma_0^2)$
- $y^1 \sim \mathcal{N}\left(\frac{N}{4}\sum_{i=3}^{K} p_i, \frac{N}{4}\sum_{i=3}^{K} p_i(1-p_i)\right) \equiv \mathcal{N}(\mu_1, \sigma_1^2)$
- $y^2 \sim \mathcal{N}\left(\frac{N}{4}\sum_{i=4}^{K} p_i, \frac{N}{4}\sum_{i=4}^{K} p_i(1-p_i)\right) \equiv \mathcal{N}(\mu_2, \sigma_2^2)$
- $y^3 \sim \mathcal{N}\left(Np_1 + \frac{N}{4}\sum_{i=5}^{K} p_i, Np_1(1-p_1) + \frac{N}{4}\sum_{i=5}^{K} p_i(1-p_i)\right) \equiv \mathcal{N}(\mu_3, \sigma_3^2)$.

Let $y^{3,k}$=y^3-y^k for $k \neq 3$ and $y^{3,k} \sim \mathcal{N}(\mu_{3,k}, \sigma_{3,k}^2)$. Using the distribution of the $y^{3,k}$s, $P\left(e|s=3\right)$ can be written as

$$P\left(e|s=3\right) < Q\left(\frac{\mu_{3,0}}{\sigma_{3,0}}\right) + Q\left(\frac{\mu_{3,1}}{\sigma_{3,1}}\right) + Q\left(\frac{\mu_{3,2}}{\sigma_{3,2}}\right). \tag{2.22}$$

Then, using Equation (2.14) (2.16) (2.20) and (2.22), P_{4ppm}^{E} can be explicitly written as

$$\begin{aligned} P_{4ppm}^{E} < \frac{1}{4}\Bigg(& Q\left(\frac{\mu_{0,1}}{\sigma_{0,1}}\right) + Q\left(\frac{\mu_{0,2}}{\sigma_{0,2}}\right) + Q\left(\frac{\mu_{0,3}}{\sigma_{0,3}}\right) + \\ & Q\left(\frac{\mu_{1,0}}{\sigma_{1,0}}\right) + Q\left(\frac{\mu_{1,2}}{\sigma_{1,2}}\right) + Q\left(\frac{\mu_{1,3}}{\sigma_{1,3}}\right) + \\ & Q\left(\frac{\mu_{2,0}}{\sigma_{2,0}}\right) + Q\left(\frac{\mu_{2,1}}{\sigma_{2,1}}\right) + Q\left(\frac{\mu_{2,3}}{\sigma_{2,3}}\right) + \\ & Q\left(\frac{\mu_{3,0}}{\sigma_{3,0}}\right) + Q\left(\frac{\mu_{3,1}}{\sigma_{3,1}}\right) + Q\left(\frac{\mu_{3,2}}{\sigma_{3,2}}\right)\Bigg). \end{aligned} \tag{2.23}$$

2.8.5. General Probability of Error when for M-PPM

Considering Equation (2.23), the probability of error bound for M-PPM scheme can be written in general form as

$$P_{Mppm}^{E} < \frac{1}{M}\sum_{i=0}^{M-1}\sum_{j\neq i}^{M-1} Q\left(\frac{\mu_{i,j}}{\sigma_{i,j}}\right). \tag{2.24}$$

In Equation (2.24), the means and the variances can be explicitly written using the discussions used to find distributions for $M = 4$ case as,

$$\mu_{i,j} = Np_1 + \frac{N}{M}\sum_{k=i+2}^{K} p_k - \left(Np_{j-i+1} + \frac{N}{M}\sum_{k=j+2}^{K} p_k \right) \tag{2.25}$$

and

$$\sigma^2_{i,j} = Np_1(1-p_1) + \frac{N}{M}\left(\sum_{k=i+2}^{K} p_k(1-p_k) + Np_{j-i+1}(1-p_{j-i+1}) + \sum_{k=j+2}^{K} p_k(1-p_k) \right) \tag{2.26}$$

where p_k=0, when k$\leq$0. The error bound given in Equation (2.24) is compared with the actual PPM symbol error rate (SER) values obtained by Monte Carlo simulations in Figure 2.2 for each PPM. In this comparison, t_s is chosen as the time allocated for one bit transmission in CSK and slot duration of PPM modulations are adjusted using Table 2.2. Therefore, the duration of a single slot is different for different PPMs. As seen in this figure, the proposed error bound formula can be considered as a good upper bound.

2.9. Performance Evaluation

Performance comparisons of PPM with CSK and other modulation techniques are obtained using Monte Carlo simulations. As mentioned before, y_i^k can be approximated with Gaussian distribution. In simulations, this distribution is utilized to obtain the number of observed molecules at the receiver. Once a realization of y_i^k is obtained, it

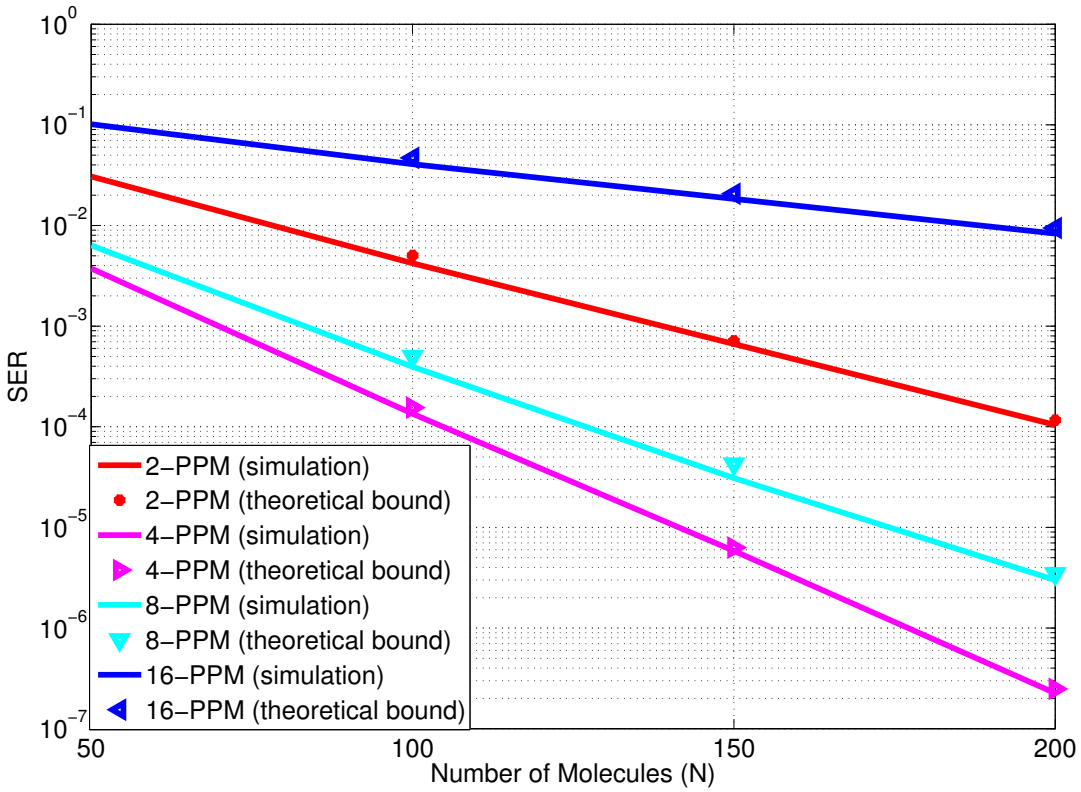
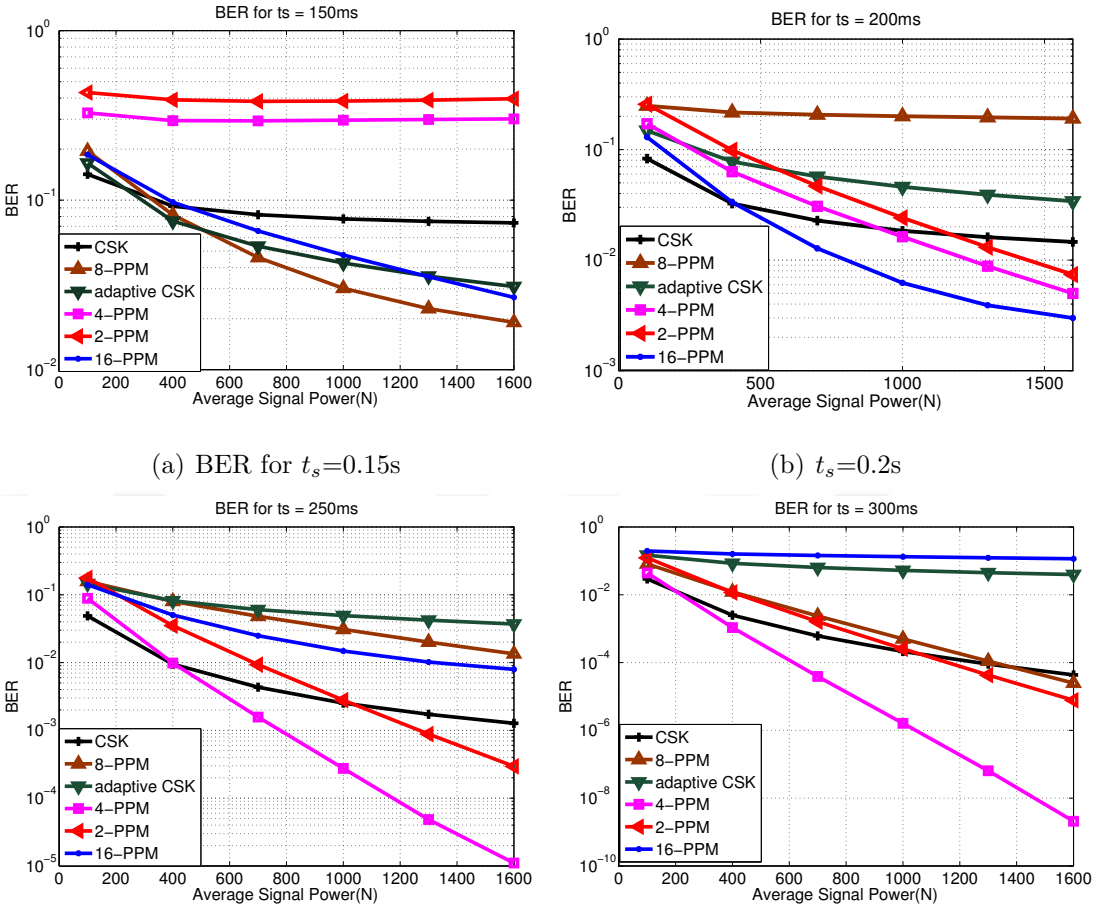


Figure 2.2. Error bounds for PPM modulations for t_s=0.4s D=79.4$\times$ 10^{-12} m^2/s, r_r=5 $\times$ 10^{-6} m and r_0=10 $\times$ 10^{-6} m

is used for decoding in the proposed communication schemes, CSK and other schemes in the literature.

In order to compare PPM with CSK and other modulations in the literature, their optimal thresholds are determined to achieve the lowest bit error rate (BER). In practice, this can be either achieved by knowing the channel characteristics or utilizing pilot symbols. In CSK with adaptive thresholding, the threshold is updated for each transmission and is set to the number of molecules observed at the previous transmission. Therefore, CSK with adaptive threshold is a special case of the CSK when the CSI is not available. In PPM, there is no need to determine a threshold, and decoding is done by finding the slot that has the maximum number of received molecules.

For all simulations, the diffusion coefficient is chosen to be D=79.4$\times$ 10^{-12} m^2/s. The radius of the receiver and the distance between the transmitter and the center of



(a) BER for t_s=0.15s

(b) t_s=0.2s

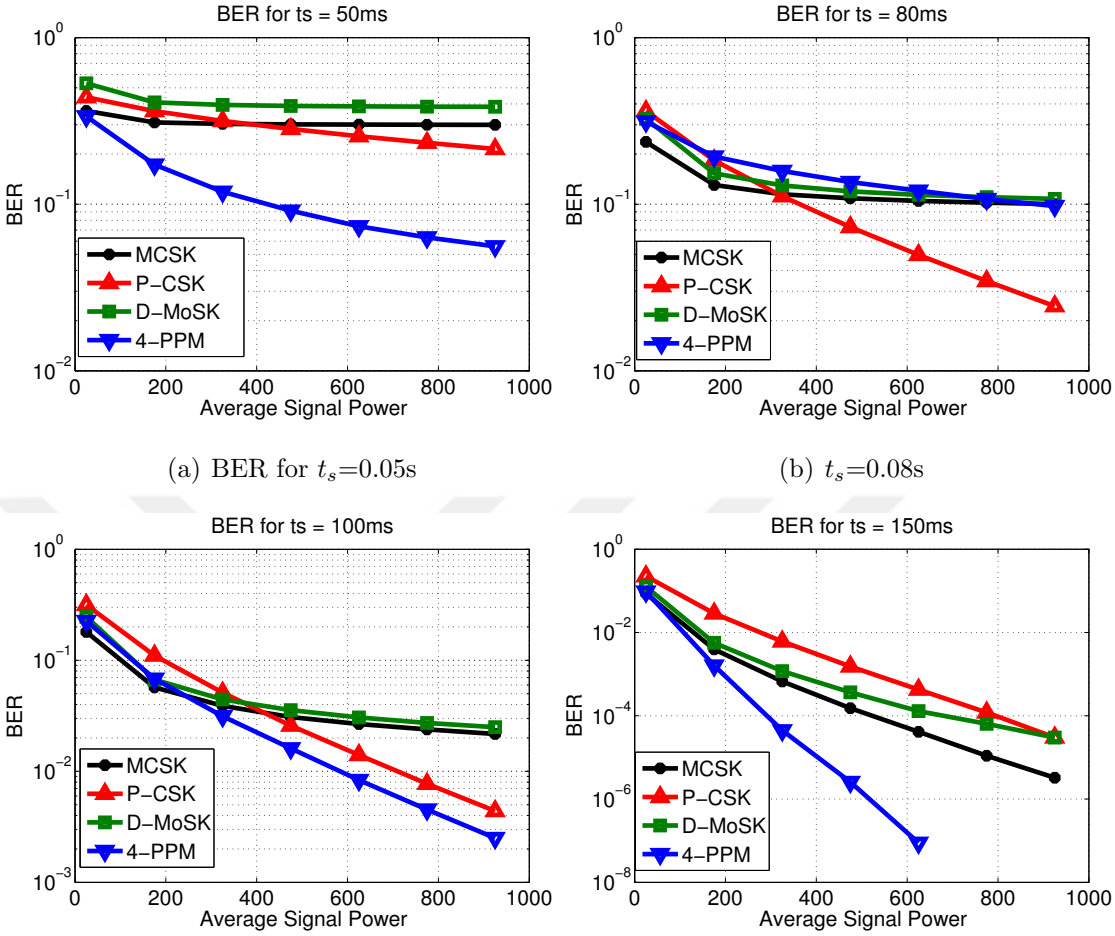
(c) t_s=0.25s

(d) t_s=0.3s

Figure 2.3. BER of the PPM modulations and CSK for different t_s. The t_s indicated on each figure corresponds to slot duration of CSK (hence bit duration for CSK) and it is adjusted for other modulations for fair comparison

the receiver are chosen as r_r=5 $\times$ 10^{-6}m and r_0=10 $\times$ 10^{-6}m, respectively. For the sake of a realistic simulation, the memory of the channel is chosen arbitrarily large as 1200, and 90000 consecutive bits are transmitted in a frame. In other words, the number of observed molecules at each slot consists of the molecules coming from the current and the previous 1199 slots. For a fair comparison, the time slot t_{slot} the number of molecules N, and the channel memory are adjusted according to Table 2.2. For instance, the channel memory in 2-PPM is 2400 which is twice that of CSK.

In Figure 2.3, BER curves of PPM modulations and CSK for four different t_s values are presented. In this figure, t_s corresponds to the time required to transmit one bit information to the receiver. Therefore, while in CSK each bit transmission



(a) BER for t_s=0.05s

(b) t_s=0.08s

(c) t_s=0.1s

(d) t_s=0.15s

Figure 2.4. BER comparison of the 4-PPM, MCSK, D-MoSK and P-CSK for different t_s. The t_s indicated on each figure belongs to MCSK, P-CSK and 4-PPM and it is adjusted for D-MoSK as $2t_s$ for fair comparison

and hence each slot takes t_s seconds, in 2-PPM and 4-PPM each slot takes $\frac{1}{2}\,t_s$, in 8-PPM $\frac{3}{8}\,t_s$, and in 16-PPM $\frac{4}{16}\,t_s$, as can be seen in Table 2.2. Figure 2.3 depicts some interesting results. Firstly, at t_s=0.15s, as can be seen in Figure 2.3(a), 8-PPM has the best performance while 2-PPM and 4-PPM have the two worst. When t_s is increased to 0.2s, 8-PPM becomes the worst. This is quite interesting, since the increase in t_s leads to worse relative performance of a modulation scheme, which has not been observed in the molecular communications literature before. The reason of this behavior is related to the discussion in Section 2.7. In that section, it is proposed that PPM has the worst performance when highest taps of the channel are equal. In an interval whose limits

are the two points that make two maximum taps equal, there is an optimum operating slot duration of PPM, which is $t_{slot} = 0.693m^2$. For the parameters used in these simulations $k = 0.2806$, hence $t^*_{slot} = 0.0543$s. For the given channel specifications, as can be observed in Figure 2.1, when t_{slot} is smaller than 0.08s, PPM has optimum performance if $t^*_{slot} = 0.0543$s. When bit duration for CSK is t_s=0.15s, the duration of each time slot for 2-PPM and 4-PPM is equal to t_{slot}=0.075s. On the other hand, the duration for 8-PPM is equal to t_{slot}=0.056s. Since 8-PPM is closer to the optimum t^*_{slot}=0.0543s, it has better performance. However, when t_s=0.2s, the duration of each slot in 8-PPM is $\frac{3}{8}\ t_s$=0.075s. Hence as can be seen Figure 2.1, this point is very close to the point that makes the two highest taps equal. Therefore, at t_s=0.2s, 8-PPM has the worst performance and 16-PPM has the best since at t_s=0.2s duration of each slot for 16-PPM is $\frac{4}{16}\ t_s$=0.05s which is closer to the optimum point. When t_s=0.25s and 0.3s, (when the first tap the highest) 4-PPM has the best performance while 2-PPM and 8-PPM potentially beat CSK. In particular, 4-PPM has the lowest BER values as t_s is increased in the region where first channel tap is maximum. The reason of this behavior can be obtained from Table 2.2. 4-PPM has the highest slot duration with 2-PPM among all PPMs and is also allowed to use higher number of molecules to represent bit-1 than 2-PPM. Therefore, it is expected that it should have better performance than 2-PPM. Furthermore, although 8-PPM and 16-PPM use more molecules than 4-PPM, since their slot durations are smaller than 4-PPM, their signal taps are smaller and ISI taps are greater than 4-PPM and therefore they are not as good as 4-PPM in the region that signal tap is the highest.

Once, the comparison of different PPMs among each other and CSK is complete, the next comparison is done with the most popular modulation methods in the literature. These modulations are Molecular Concentration Shift Keying (MCSK) that proposes to use two different types of molecules for two consecutive transmissions [22], D-MoSK modulation that uses two different types of molecules simultaneously as two orthogonal channels [23], and CSK with pre-equalization (P-CSK) that uses two different types of molecules and represents symbols as the difference of the received number of these molecules. In particular, the transmitter first releases type-A molecules and

after a pre-determined time it releases type-B molecules. The determination of the time and number of type-B molecules is chosen to minimize the signal to interference ratio [24]. For this comparison, 4-PPM is chosen to be used as the most successful modulation scheme among PPMs. For the sake of fairness, the duration of each slot and the number of molecules per bit for each modulation should be carefully determined. 4-PPM requires four time slots to encode a 2-bits symbol while MCSK, P-CSK encode each bit with one slot and D-MoSK encodes 2-bit symbol in a single slot. Furthermore, MCSK D-MoSK, and P-CSK use two types of molecules. Therefore for a fair comparison, 4-PPM should use two types of molecules as two orthogonal channels similar to the way D-MoSK does. Under the light of all these factors, the slot duration of MCSK, P-CSK, and 4-PPM should be t_s while in D-MoSK it should be $2t_s$ for equal data rate constraint. Furthermore, since P-CSK uses an additional molecule to minimize SIR, this number should be adjusted with others.

As can be seen in Figure 2.4, 4-PPM outperforms other modulation techniques. The only exception occurs at t_s=0.08s (in this case t_{slot}=t_s, due to two types of molecule usage), which makes two highest taps equal, which has detrimental effect on PPM as discussed before. Furthermore, as in CSK, all these modulations except 4-PPM require decision threshold to be employed. Hence, the exact knowledge of the channel is a must to obtain the best BER for these modulations. The curves of other modulations are obtained with this assumption. On the other hand, PPM modulations only require the knowledge of the position of the maximum tap.

2.10. Concluding Remarks on PPM in Molecular Communication

PPM modulation schemes are applied to overcome the common issues like ISI complexity and CSI requirement in MCvD channels. The decoding method is quite simple and it is based on finding the slot that the receiver absorbed the maximum number of molecules. Since nanomachines have limited computational capability, the simplicity of the decoding scheme is an important asset in molecular communications. Furthermore, as the decoding rule depends on finding the slot that involves the maximum number of molecules, it does not require channel knowledge except for location

of the expected maximum channel tap. Furthermore, if the signal tap is not the maximum channel tap (which can be a typical result when t_s is decreased), the decoding rule can be easily modified by shifting the decision slot. Since ISI is distributed in each slot of PPM and the aim is finding the slot that involves the maximum number of molecules, the effect of ISI is diminished or can even be beneficial on the performance. The analytical error bounds of 2-PPM, 4-PPM, 8-PPM, and 16-PPM are calculated and BER curves are also obtained using Monte Carlo simulations. Then, comparisons among PPM modulations, CSK modulation and other popular modulation methods in the literature are done. According to these comparisons, PPM modulations are more successful than CSK, and 4-PPM is more successful than the other successful methods in the literature. All of these results conclude that PPM is a suitable method for molecular communications due to its simplicity and robustness to ISI.

3. CHANNEL CODING FOR MOLECULAR COMMUNICATION

3.1. Related Literature

In addition to modulation methods, error control coding has also been used in molecular communications. In [29] and [30], the authors show that higher order Hamming codes (e.g., Hamming(31,26)) can be used to improve bit error rate. The same authors proposed to be used self orthogonal convolutional codes (SOCC) in [31]. In [32], Reed Solomon codes are proposed to use for the molecular communications channel. Although error control coding methods show promising results for molecular communications, they are successful especially when the channel is good enough. For more challenging channels that involve longer ranges, lower bit durations, or 3-D channels, error control coding may not be sufficient to outperform uncoded transmission.

Considering the issues mentioned above related to the MCvD channel, we propose a new coding scheme considering the dynamics of the channel. It is well known that, errors due to ISI occur when consecutive molecule releases are done. In particular, assuming on-off keying is used to transmit bit-1 and bit-0, consecutive bit-1s may lead to erroneous decoding due to accumulation of the redundant molecules. Therefore, any sequence that involves consecutive bit-1s will most possibly be affected by ISI. Another important problem in molecular communication is threshold selection. A general approach is choosing the optimum threshold by using channel parameters and/or pilot symbols. On the other hand, using a fixed optimum threshold may result in erroneous decoding for some part of the sequence especially when consecutive bit-1s are transmitted. Therefore, using an adaptive threshold will potentially outperform the optimum but fixed threshold. Considering all these facts, we propose a code family for the molecular communication channel.

3.2. Code Family for Molecular Communication

Our proposed code family does not involve consecutive bit-1s and uses an adaptive threshold which is updated for each codeword (CW) using a analytical formulation. In particular, the proposed CW has the following properties:

- each CW does not include any consecutive bit-1s
- there is at least one bit-1 (for decoding reason)
- each CW starts with 0 to avoid consecutive bit-1s between two neighbour CWs

Let CW_n be the matrix that consists of all possible CWs of length n. For $n = 4$, there are 4 binary sequences that follow the mentioned rule $(0001, 0010, 0100, 0101)$. Therefore CW_4 can be written as

$$CW_4 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}. \tag{3.1}$$

Furthermore for $n = 4$, there are 3 binary sequences $(1000, 1001, 1010)$ that do not have consecutive bit-1s yet start with bit-1. Similarly, W_4 that consists of these sequences can be represented in matrix form as

$$W_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}. \tag{3.2}$$

Using CW_4 and W_4, CW_5 is obtained as follows:

- Add a new column to left side of CW_4 and W_4, and fill the new columns with 0. The new matrices are

$$CW_{4e0} = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix}, \tag{3.3}$$

$$W_{4e0} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{bmatrix}. \tag{3.4}$$

- Combining the CW_{4e0} and W_{4e0} will result in

$$CW_5 = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{bmatrix}. \tag{3.5}$$

- Add a new column to left side of CW_4, this time fill the new column with 1. The new matrix is

$$CW_{4e1} = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 \end{bmatrix}, \tag{3.6}$$

- Add a new row to bottom of CW_{4e1}. Set the first element of that row 1, fill others with 0. This will result in

$$W_5 = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}, \tag{3.7}$$

By repeating these steps successively for $n = 6, 7, ...$, one can find the binary sequences that neither have any consecutive bit-1s nor start with bit-1 for any $n > 4$. Therefore, for any n, the desired CW_n can be iteratively solved as

$$CW_n = \left[\begin{array}{cc} \mathbf{0} & CW_{n-1} \\ \hline \mathbf{0} & W_{n-1} \end{array}\right] \tag{3.8}$$

and corresponding W_n can be obtained as

$$W_n = \left[\begin{array}{cc} \mathbf{1} & CW_{n-1} \\ \hline 1 & \mathbf{0} \end{array}\right]. \tag{3.9}$$

Therefore, using Equation (3.8) and Equation (3.9), CW family with any length can be constructed. As can be obtained from these equations, the number of CWs for any n can be obtained by appending the number rows of CW_{n-1} and W_{n-1}. The number of possible CWs for each n is listed in Table 3.1. Considering this table for $n = 4$, there are 4 possible CWs, hence the maximum possible information sequence that can be encoded with CW_4 is $k = \lfloor \log_2 4 \rfloor = 2$; hence, the code rate is $k/n = 1/2$.

Table 3.1. Number of code sequences of CW_n and W_n

n	CW_n	W_n	max. k	max k/n
4	4	3	2	2/4
5	7	5	2	2/5
6	12	8	3	3/6
7	20	13	4	4/7
8	33	21	5	5/8
9	54	34	5	5/9
10	88	55	6	6/10
11	143	89	7	7/11
12	232	144	7	7/12

3.3. Decoding Stage

The main motivation behind the decoding stage is finding the slots that involve the maximum and the minimum number of molecules. Clearly, since there is at least one bit-1 in each CW, one can expect that, the maximum slot is most likely to be bit-1 and similarly the minimum slot is most likely corresponds to bit-0. Once these parts of the CW are decoded, the rest part of the CW is decoded by using adaptive threshold τ_a. In particular, the whole decoding stage is done as

- Let **M** be the vector that is the number of observed molecules for the generated CW **Z**
- **Z** has at least one bit-1. One of these bit-1s should most probably located at the place where **M** is maximum. Let this maximum value be d_1 and decode this place as bit-1
- Set the neighbours of this bit-1 with bit-0s (since there is not any consecutive bit-1s)
- Find the place in **M** where the molecule number is minimum. Put here a bit-0. Also lets recall this minimum value as d_0.

- The rest part of the CW is decoded using an adaptive threshold $\tau_a = bd_0 + (1-b)d_1$.
- The determination of b is achieved by both analytically and Least Squares (LS) solution sense.

An illustrative example for the decoding stage can be given as

(i) Let $\mathbf{M}$= $\begin{pmatrix} \mathrm{m}_0 & m_1 & m_2 & m_3 & m_4 & m_5 & m_6 \end{pmatrix}$
(ii) Let m_3 correspond to the maximum slot and m_6 be the minimum, hence $d_1 = m_3$ and $d_0 = m_6$.
(iii) Since the first element of each CW is bit-0, $\hat{\mathbf{Z}}$= $\begin{pmatrix} 0 & \mathrm{z}_1 & z_2 & z_3 & z_4 & z_5 & z_6 \end{pmatrix}$
(iv) Place of minimum (m_6) will be bit-0, $\hat{\mathbf{Z}}$= $\begin{pmatrix} 0 & \mathrm{z}_1 & z_2 & z_3 & z_4 & z_5 & 0 \end{pmatrix}$
(v) Place of maximum (m_3) will be bit-1 and its neighbours will be bit-0, $\hat{\mathbf{Z}}$= $\begin{pmatrix} 0 & \mathrm{z}_1 & 0 & 1 & 0 & z_5 & 0 \end{pmatrix}$
(vi) z_1 and z_5 are decoded using adaptive threshold as $z_1, z_5 = \gtrless_0^1 \tau_a = bd_0 + (1-b)d_1$.

3.4. Correction Stage

Once the corresponding decoding scheme presented in Section 3.3 is applied, it is still possible that consecutive bit-1s are decoded. In particular, $\mathbf{M}$ may involve two or more consecutive elements that exceed threshold τ_a. For these cases, we evaluate the probability of all cases and conclude the most possible one.

3.4.1. Correction of Two Consecutive bit-1s

When two consecutive bit-1s are decoded in $\hat{\mathbf{Z}}$, it is clear that there is at least one error; hence, the corrected version of (1 1) can be (1 0), (0 1) or (0 0). Since the previous symbol before (1 1) is bit-0, we can model the ISI for these two slots as $X_3 \sim \mathcal{N}(Np_x^3, Nf(p_x^3))$ and $X_4 \sim \mathcal{N}(Np_x^4, Nf(p_x^4))$, where $p_x^3 = \sum_{x=3}^{\infty} a_x p_x$, $a_x \epsilon \{0,1\}$, $f(p_x^3) = \sum_{x=3}^{\infty} a_x p_x (1-p_x)$, and N is the number of transmitted molecules per bit. Furthermore, let ISI_i be the distribution of ISI coming from the ith tap and as men-

tioned in Section 1.2, $ISI_2 \sim \mathcal{N}(Np_2, Np_2(1-p_2))$. Since decoding of each bit is independent, using these distribution probabilities of the possible sequences can be listed as

(i) $P(11|01) = P(X_3 > \tau_a)$
(ii) $P(11|10) = P(ISI_2 + X_4 > \tau_a)$
(iii) $P(11|00) = P(X_3 > \tau_a)P(ISI_2 + X_4 > \tau_a)$

Comparing these probabilities, since $p_1 > p_2 > p_3...$, it is clear that $P(11|10)$ is the most possible one. Therefore, if (1 1) is decoded, the most possible correction is (1 0).

3.4.2. Three consecutive bit-1s correction

In a similar fashion, when (1 1 1) is received, the possible actual sequences are (1 0 1) (1 0 0), (0 1 0), (0 0 1) and (0 0 0). The probabilities of all possible sequences are

(i) $P(111|101) = P(X_4 + ISI_2 > \tau_a)$
(ii) $P(111|100) = P(111|101)P(X_5 + ISI_3 > \tau_a)$
(iii) $P(111|000) = P(111|101)P(111|100)P(X_3 > \tau_a)$
(iv) $P(111|010) = P(X_3 > \tau_a)P(X_5 + ISI_3 > \tau_a)$
(v) $P(111|001) = P(111|101)P(111|100)$

Comparing these probabilities, one can choose (1 0 1) as the most possible sequence if (1 1 1) is received.

3.5. Adaptive Threshold Selection

We have proposed an adaptive threshold for each received sequence $\mathbf{M}$. Note that, since each $\mathbf{M}$ varies between different d_0 and d_1, the optimum threshold for each sequence can be obtained as $\tau_a = bd_0 + (1-b)d_1$. We have proposed two different methods to obtain τ_a. These solutions are analytical and Least Squares solutions.

3.5.1. Analytical Adaptive Threshold

Since at the decoding stage, we assign bit-0 to the neighbors of the slot with the maximum number of molecules and since at the correcting stage neighbors of the bit-1s can be corrected, the optimum threshold τ_a can be obtained by considering the decoding of the remaining bits. For those bits, the possible highest ISI may come from the two previous bits and can be distributed as $ISI_3 \sim \mathcal{N}(Np_3, Np_3(1-p_3))$. Moreover, the distribution of the molecules if bit-1 is transmitted at the current slot can be modeled as $X_{signal} \sim \mathcal{N}(Np_1, Np_1(1-p_1))$. Obviously, the number of molecules at slot m_x can be represented as $m_x^0 = \sum_{x=3}^{\infty} a_x ISI_i$ if the transmitted symbol is bit-0 and $m_x^1 = X_{signal} + \sum_{x=3}^{\infty} a_x ISI_i$ if the transmitted symbol is bit-0. Our approach to derive an analytical threshold is determining the weakest m_x^1 and the strongest m_x^0 that is the most susceptible case for erroneous decoding. For this case, we derive the probability of error and minimize this function with respect to threshold. In particular, ignoring the higher order ISI terms than ISI_3, let $m_x^0 = ISI_3$ and $m_x^1 = X_{signal}$. The error probability for decoding m_x can be written as

$$P_E = P_0 \quad P\left(\hat{m_x} = 1 | m_x = 0\right) \quad + \quad P_1 \quad P\left(\hat{m_x} = 0 | m_x = 1\right). \tag{3.10}$$

Using the distribution of m_x, Equation (3.10) can be explicitly written as

$$P_E = P_0 Q\left(\frac{\tau_a - Np_3}{\sqrt{Np_3(1-p_3)}}\right) + P_1\left(1 - Q\left(\frac{\tau_a - Np_1}{\sqrt{Np_1(1-p_1)}}\right)\right) \tag{3.11}$$

where $\tau_a = bd_0 + (1-b)d_1$ is the adaptive optimum threshold, P_0 and P_1 are the occurances of bit-0 and bit-1 respectively and varies for different code length n, finally

b is the optimization parameter that need to be solved and clearly it should be between 0 and 1. Taking the derivative of P_E with respect to b and equating it to 0 yields

$$P_1 \frac{e^{\frac{-\left(-Np_1+bd_0+(1-b)d_1\right)^2}{2Np_1(1-p_1)}}}{\sqrt{Np_1(1-p_1)}} = P_0 \frac{e^{\frac{-\left(-Np_3+bd_0+(1-b)d_1\right)^2}{2Np_3(1-p_3)}}}{\sqrt{2Np_3(1-p_3)}}. \tag{3.12}$$

The solution of Equation (3.12) can be obtained as

$$b_a = \frac{Np_1R - Np_3U + d_1U - d_0R - \sqrt{R \quad U\left(Np_1 - Np_3\right)^2 - 2RU\log\sqrt{\frac{U}{R}}(R-U)}}{(U-R)(d_1-d_0)}. \tag{3.13}$$

where $U = Np_1(1-p_1)/P_1$ and $R = Np_3(1-p_3)/P_0$.

3.5.2. LS Adaptive Threshold

In this section, the determination of $\tau_a = bd_0 + (1-b)d_1$ is achieved by LS sense. In particular, let τ_A be the vector that consists of all of the optimum thresholds for the all transmitted sequences. Let $D = [D_0 \quad D_1]$ be the matrix that involves two column vectors D_1 and D_0, which are the maximum and the minimum number of absorbed molecules for all sequences, respectively. Then, our aim is to calculate $b_A = [b \quad 1-b]^T$ using D matrix and the τ_A vector. Using D and τ_A, we can write the expression,

$$Db_A = \tau_A. \tag{3.14}$$

Using Equation (3.14) we can write

$$D_0 b + D_1(1-b) = \tau_A, \tag{3.15}$$

which can be modified as

$$D_0 b - D_1 b = \tau_A - D_1, \tag{3.16}$$

$$(D_0 - D_1)b = \tau_A - D_1. \tag{3.17}$$

Then the LS solution of b, (b_{LS}) can be obtained using Equation (3.17) as

$$b_{LS} = (A^T A)^{-1} A^T (\tau_A - D_1), \tag{3.18}$$

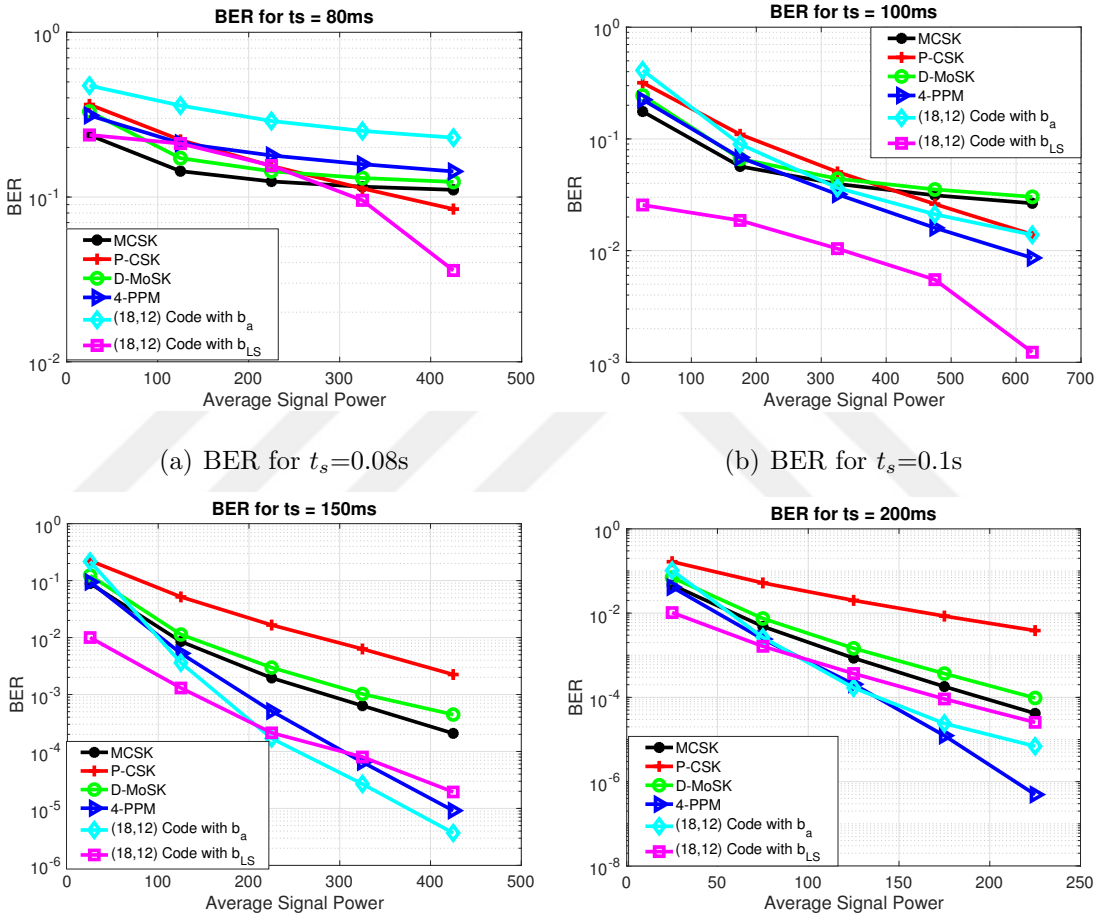
where $A = (D_0 - D_1)$.

3.6. Performance Evaluation

We have tested our codeword family and compared it with the other proposed methods in the literature using Monte Carlo simulations. Furthermore, we have also applied it to a a real time molecular communication system and compared it with the existing code used for this system.

3.7. Performance Evaluation Using Monte Carlo Simulation

In this section, we investigate the performance of our proposed code with analytical and LS threshold methods denoted with b_a and b_{LS} respectively. The performance evaluation is done for different bit durations and different signal powers and compared with the other proposed methods, as well as 4-PPM presented in Chapter 2.



(a) BER for t_s=0.08s

(b) BER for t_s=0.1s

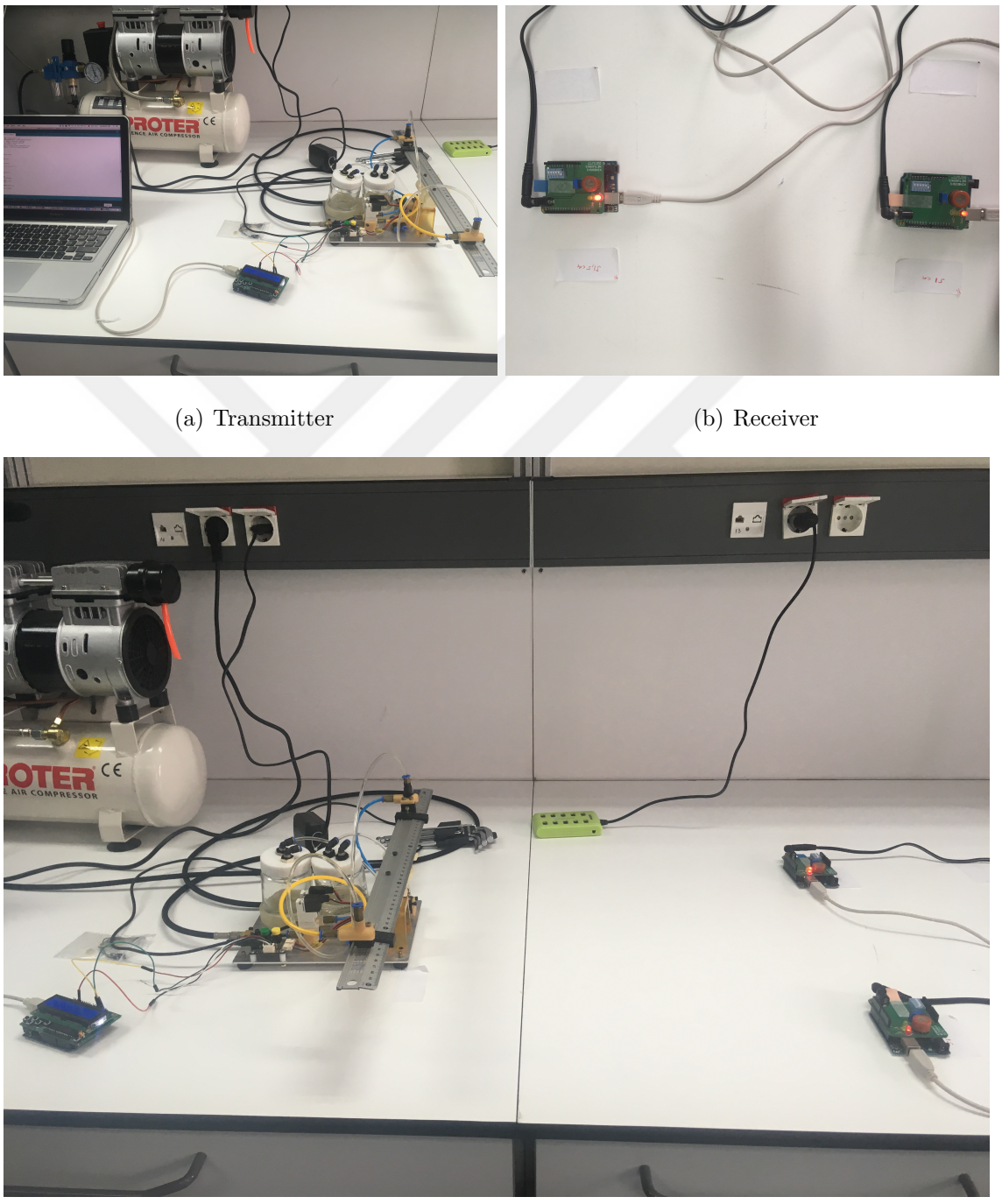
(c) BER for t_s=0.15s

(d) BER for t_s=0.2s

Figure 3.1. BER comparison of the our proposed coding method with 4-PPM, MCSK, D-MoSK and P-CSK for different t_s values. The t_s indicated on each figure belongs to MCSK, P-CSK and 4-PPM and it is adjusted for D-MoSK as $2t_s$ for fair comparison

As seen from Figure 3.1, our proposed coding method outperforms the other proposed methods in the literature and it is even better than PPM modulation proposed in Chapter 3. Furthermore analytical solution is better than LS solution for higher bit durations and LS has better performance than the analytical solution for lower bit durations. This is expected, since, as bit duration decreases, the channel taps that

corresponds to ISI increases; hence, the simplifications done for the analytical approach may not be applicable.



(a) Transmitter

(b) Receiver

(c) Whole system

Figure 3.2. Molecular MIMO testbed

.

3.8. Performance Evaluation Using Real Time Testbed

In this section we have applied our proposed code family to a real time molecular communication system called molecular multiple input multiple output (MIMO) and presented in [33]. We have set up this testbed in our laboratory with the help of the designers of the testbed. As can be seen in Figure 3.2, the testbed has two transmitters that transmit binary data by spraying alcohol once in a symbol duration to transmit bit-1; for bit-0, they do not spray. The emitted molecules diffuse in the air and arrive at the receivers, which are chemical sensors that converts molecule concentrations to electrical signals.

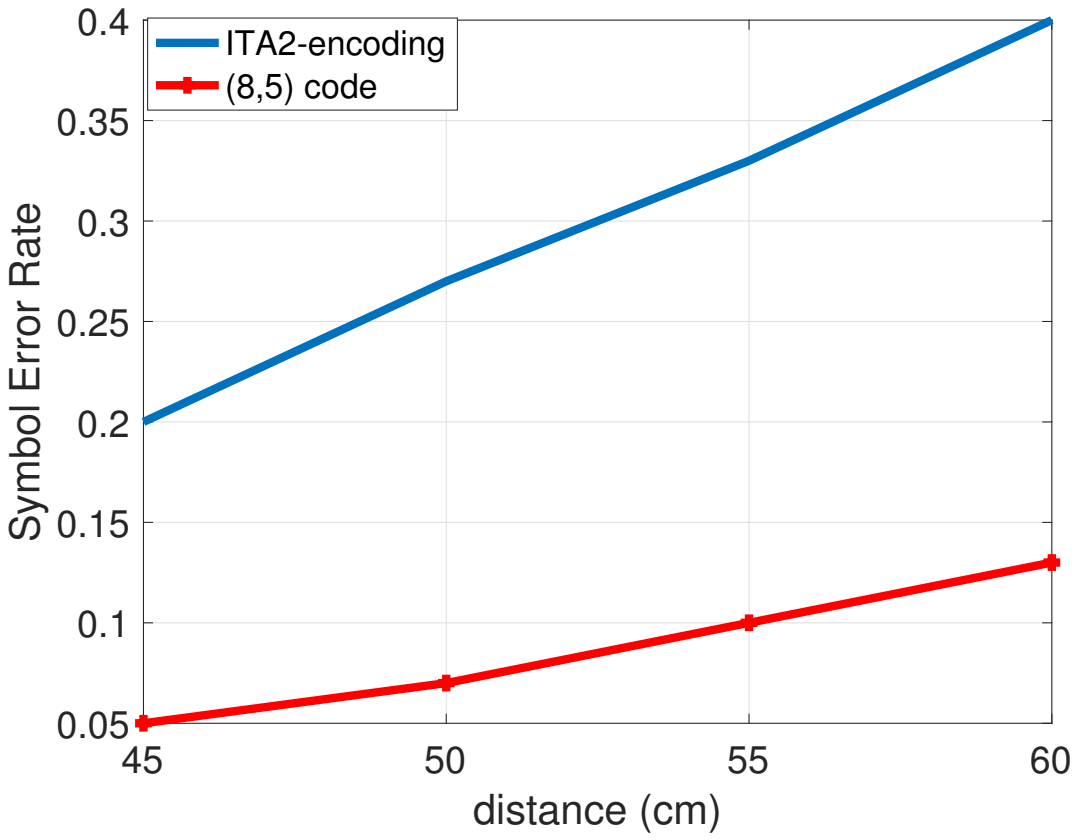


Figure 3.3. Comparison of our proposed codeword with ITA2-encoding scheme using testbed

In [33], the authors proposed to use ITA2-encoding to transmit 26 letters with 5-bit binary sequences. We have tested our proposed code family to transmit these 26 letters in this testbed and compare it with the performance of ITA2-encoding.

Since there are 26 letters, considering Table 3.1, codelength $n = 8$ is sufficient to encode these letters with our codeword family. Therefore, for the sake of simplicity, we applied our proposed codewords with $n = 8$ to the testbed and compare it with ITA2-encoding. Since ITA2-encoding uses 5 bits, we chose bit duration t_s as 4 seconds and 6.4 seconds for our proposed method and ITA2-encoding, respectively, in order to do a fair comparison in terms of data rate[1] .

As can be seen in Figure 3.3, the proposed codeword is quite better than the ITA2-encoding. Since ITA2-encoding allows consecutive bit-1s and its threshold is not adaptive, our proposed method outperforms the existing communication scheme presented in [33].

[1]In ITA2-encoding one letter is transmitted with 5 bits; hence, one letter is transmitted in 5x6.4=32s, in our proposed method, one letter is transmitted with 8 bits; hence, one letter is transmitted in 8x4=32s

4. OPTIMAL RECEPTION DELAY IN DIFFUSION-BASED MOLECULAR COMMUNICATION

4.1. Related Literature

In order to eliminate ISI in molecular communication, various equalization methods are proposed. Equalization can be categorized into two groups as transmitter-based and receiver-based. One of the transmitter-based equalization method is proposed in [24]. In this work, the authors propose sending A-type molecules followed by B-type molecules after a pre-determined delay. The receiver decodes the symbol by the difference of the numbers of A-type and B-type molecules that are received in the symbol slot. The release-time and number of B-type molecules are optimized to maximize the signal to interference ratio (SIR). Another transmitter-based method is proposed in [34], which uses ISI as a constructive component by adapting the number of released molecules so that the residual molecules lead to a beneficial effect on decoding of the following symbols. As a receiver-based equalization method, in [35], minimum mean square error (MMSE) and decision feedback equalizers as well as maximum likelihood sequence estimation methods are proposed for MCvD channel. In [36], decoding is achieved by considering the previous observed samples and increasing the sampling rate of the receiver to improve the performance.

In addition to these methods to reduce ISI, various different approaches have been proposed to improve system performance. For instance, if a flow is available in the channel, then the speed of the molecules is increased, which leads to a reduction in the number of the redundant molecules. In [37] and [38], receiver designs are presented in diffusive channels with flow. In [39], symbol synchronization, which is generally assumed to be perfect in the literature, is achieved using an additional type of molecule and optimal symbol interval is determined using the observed samples using Maximum likelihood and some other sub-optimal solutions.

The methods mentioned above to mitigate ISI either require different types of molecules (and hence different types of receptors), additional computational complexity, and memory at the nanomachines or are proposed for specific channels. In this work, we propose a novel method, called shift-τ method, that does not require any additional cost over the conventional CSK modulation. The basic principle behind this method is as follows: at the beginning of each symbol slot, the transmitter releases molecules to transmit the intended symbol to receiver. In conventional CSK, during this symbol slot, the receiver counts the number of observed molecules and then makes a hard decision. On the other hand, some of these molecules comes from the previous symbols and leads to ISI. Furthermore, due to the diffusive nature of the channel, it can be expected that the molecules absorbed at the beginning of each symbol slot belong to the previous symbols. This is the incentive for our proposed method. In this method, the receiver starts the absorption interval of molecules with a pre-determined delay, and it diminishes the effects of ISI on the current symbol by eliminating the molecules absorbed at the beginning of the time slot in CSK. To achieve this, we propose to find the pre-determined shifting time both numerically and analytically by optimizing different objective functions, and then compare it with conventional CSK modulation. Due to the simplicity of this method, it can be adapted to other proposed schemes in the literature to improve their performance.

4.2. Proposed Method

The channel model considered in this work is a 3D diffusive channel without drift and involves a point transmitter and a spherical receiver with fully absorbing capability. However, the proposed shift-τ method can be used for other channel models. The movement of a molecule under diffusion process can be modeled by Fick's second law. Using this concept, in [27], the probability of a molecule, released from a point transmitter, reaching a spherical receiver whose radius is r_r and distance between its center and the transmitter is r_0 by time t is given by

$$F_{hit}(t) = \frac{r_r}{r_0}\text{erfc}\left[\frac{r_0 - r_r}{\sqrt{4Dt}}\right] \tag{4.1}$$

where D is the diffusion coefficient, which depends on the temperature, the viscosity of the fluid, and Stokes' radius of molecules.

On-off keying, which is a special case of CSK, is used for the proposed system. On-off keying is achieved by releasing M molecules at the beginning of the symbol slot for sending bit-1 and sending nothing for bit-0. Once the molecules are released, the receiver counts the number of absorbed molecules by its receptors during the symbol slot (t_s) and does a hard decision on this count using a pre-determined threshold.

In this method, we propose to shift the decoding slot of the receiver for each symbol by a pre-determined time to decrease ISI in CSK. Let τ be the shifting amount of time for the absorption interval. The transmission and reception intervals of this method and CSK are summarized in Table 4.1. For demonstration of the motivation of the proposed method, we present the bit error rates (BERs) with respect to τ values, obtained via computer simulations, for different symbol slots $t_s = 0.2s$ and $t_s = 0.3s$ with $r_r = 5\mu m$, $r_0 = 10\mu m$, $D = 80\mu m^2/s$ in Figure 4.1. The number of molecules released by the transmitter to transmit bit-1 (M) is chosen as 500 and no molecule is released to transmit bit-0.

Table 4.1. Time interval that receiver counts the observed molecules

Release time in Tx	Rx in CSK	Rx in proposed method
0	$[0, t_s]$	$[\tau, (t_s + \tau)]$
t_s	$[t_s, 2t_s]$	$[(t_s + \tau), (2t_s + \tau)]$
$2t_s$	$[2t_s, 3t_s]$	$[(2t_s + \tau), (3t_s + \tau)]$

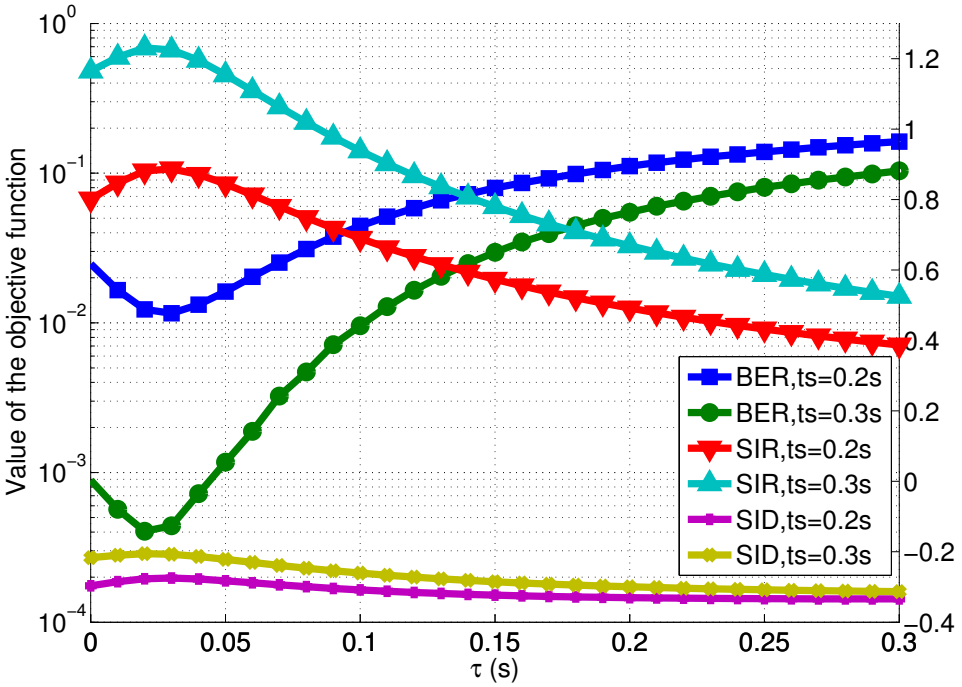


Figure 4.1. Effect of shifting the decoding process by τ seconds on BER and other objective functions. The left hand side of the y-axis belongs to BER curves while right hand side of the y axis belongs to other objective functions SID and SIR.

As seen in Figure 4.1, BER values are reduced (compared to the $\tau = 0$ case) by a range of τ values. In these figures, at $\tau = 0$, the system is equivalent to conventional CSK. When τ is slightly increased, the molecules at the beginning of the symbol slot, which possibly belong to the previous transmissions, are eliminated. On the other hand, if τ is excessively increased, molecules related to the intended symbol will be taken into account for the decoding of the previous symbol, and communication will be even worse than CSK. This tradeoff can be observed in Figure 4.1. We, therefore, analytically and numerically find τ values that increase the signal and decrease ISI to obtain the lowest BER. Firstly, in order to demonstrate the benefit of the proposed method and optimize τ, the channel taps of the system should be defined. If t_s is the symbol slot that corresponds to the time allowed for transmission and decision of one symbol, the channel taps in conventional CSK modulation can be computed as

$$p_k = F_{hit}(kt_s) - F_{hit}((k-1)t_s), \tag{4.2}$$

where $k \, \epsilon \, Z^+$. For the proposed system, the new channel taps, h_ks, are calculated as

$$h_k = F_{hit}(kt_s + \tau) - F_{hit}((k-1)t_s + \tau). \tag{4.3}$$

It is important to note that h_1 and p_1 correspond to the intended symbol and give the ratio of the number of absorbed molecules by the receiver at the current transmission slot to the number of all released molecules in this slot. Similarly, other taps give the ratio of molecules absorbed in future slots to the number of molecules released in the current slot, hence they contribute to ISI. It is assumed that first taps are greater than the other taps which is the most typical case in the literature. The analytical formula of BER in CSK can be obtained by considering each possible transmission case. However, this approach can be useful for memoryless channels as proposed in [28], which may not be always possible in the molecular communication literature. When the channel has memory, the analytical formulation of BER will be more complex and one should consider all possible cases to obtain a closed form solution. Furthermore, analytical minimization of such a BER function will not be tractable. We, therefore, minimize BER via computer simulations to obtain optimal τ. As shown in Figure 4.1, we propose other objective functions to solve this problem both analytically and numerically and compare the results with optimum BER values obtained by exhaustive computer simulations. In the molecular communication literature, BER is directly related to the taps of the channel. The first tap determines the strength of the intended symbol at the receiver while other taps are related to ISI. Therefore, as proposed in [24], SIR $= \frac{h_1}{\sum_{k=2}^{K} h_k}$ can be chosen as the objective function, where K is the memory length of the channel. Then the optimization problem becomes

$$\tau = \arg \max_{0 \leq \hat{\tau} \leq t_s} \frac{h_1(\hat{\tau})}{\sum_{k=2}^{K} h_k(\hat{\tau})}. \tag{4.4}$$

Since it is assumed that the first tap is the greatest tap, interval of the constraint is chosen as $0 \leq \hat{\tau} \leq t_s$ in order not to miss the most important portion of the signal. Furthermore, one can observe from Figure 4.1 that $\hat{\tau}$ is closer to 0 than t_s. This is expected since, although the number of absorbed molecules for the intended symbol will most possibly be the highest in the current slot due to the first tap being the greatest assumption, at the earlier part of the current slot, absorbed molecules most possibly belong to previous symbols. Unfortunately, an analytical solution of Equation (4.4) is not available. In [24], the authors propose to obtain the optimum SIR via simulation. In this letter, we propose an alternative objective function for the SIR to obtain an analytical solution. Maximizing the signal tap h_1 and minimizing the other taps that correspond to ISI is the main motivation to find the optimal τ. Therefore, we can rearrange the terms as the difference of the signal tap to the sum of the ISI taps, which can be referred to as Signal to Interference Difference (SID), as

$$\text{SID} = h_1(\hat{\tau}) - \sum_{k=2}^{K} h_k(\hat{\tau}), \tag{4.5}$$

and the corresponding optimization problem is written as

$$\tau = \arg \max_{0 \leq \hat{\tau} \leq t_s} \left[h_1(\hat{\tau}) - \sum_{k=2}^{K} h_k(\hat{\tau}) \right]. \tag{4.6}$$

The objective function in Equation (4.6) can be rewritten using Equation (4.3) as $h_1 - \sum_{k=2}^{K} h_k = 2F_{hit}(t_s+\tau) - F_{hit}(\tau) - F_{hit}(Kt_s+\tau)$. Since $F_{hit}(Kt_s+\tau) \approx F_{hit}(Kt_s)$, if K is large enough (in a realistic case K is infinite), then this term can be considered as an independent term from τ and can be ignored to solve the optimization problem as

$$\tau = \arg \max_{0 \leq \hat{\tau} \leq t_s} \left[2F_{hit}(t_s+\tau) - F_{hit}(\tau) \right]. \tag{4.7}$$

Accordingly, the explicit version of Equation (4.7) can be obtained using Equation (4.1) as,

$$\tau = \arg \max_{0 \leq \hat{\tau} \leq t_s} 2\frac{r_r}{r_0}\text{erfc}\left[\frac{r_0 - r_r}{\sqrt{4D(t_s+\hat{\tau})}}\right] - \frac{r_r}{r_0}\text{erfc}\left[\frac{r_0 - r_r}{\sqrt{4D\hat{\tau}}}\right]. \tag{4.8}$$

As can be shown in [40, p. 760], the expression in Equation (4.8) is convex for $0 < \tau < \frac{2m^2}{3}$ where $m = \frac{r_0 - r_r}{\sqrt{4D}}$. To find the maximum point in this interval, the derivative of the objective function is obtained and equated to 0. We calculate the logarithm of this equality to arrive at expression

$$\ln(2)\left[(t_s+\hat{\tau})\hat{\tau}\right] + m^2 t_s = \frac{3}{2}(t_s+\hat{\tau})\,\hat{\tau}\ln(\frac{t_s+\hat{\tau}}{\hat{\tau}}). \tag{4.9}$$

In Equation (4.9), the $\ln\left(\frac{t_s+\hat{\tau}}{\hat{\tau}}\right)$ term is a challenge to process. In order to tackle this challenge, we first rewrite it as $\ln\left(\frac{t_s+\hat{\tau}}{\hat{\tau}}\right) = \ln(1+\frac{\hat{\tau}}{t_s}) - \ln(\frac{\hat{\tau}}{t_s})$. According to Pade

Approximants [41], the first order approximations of $\ln(1+x) = \frac{x(6+x)}{6+4x}$ and $x\ln(x) = \frac{-20x(1-x)}{7+15x}$. Using these expressions, Equation (4.9) can be rewritten as

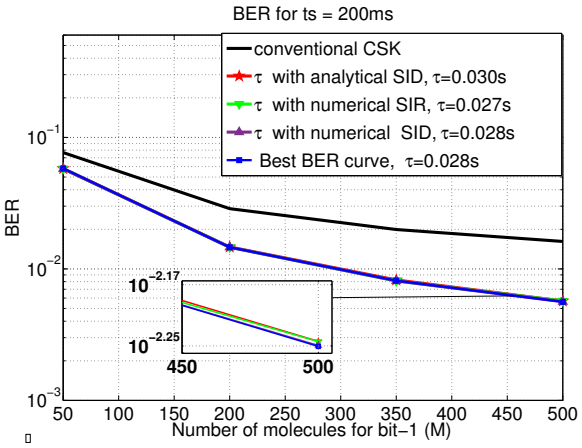
$$\begin{aligned} 15\hat{\tau}^5 + 4.27{t_s}\hat{\tau}^4 + 63.25{t_s}^2\hat{\tau}^3 + \left(-40m^2{t_s}^2 + 48.06{t_s}^3\right)\hat{\tau}^2 + \\ (-78.67m^2{t_s}^3 + 100.60{t_s}^4)\hat{\tau} - 28m^2{t_s}^4 = 0. \end{aligned} \tag{4.10}$$

In Equation (4.10), one of the five roots, which needs to be solved, is in the range of $0 < \hat{\tau} < \frac{2m^2}{3}$. Since we assume that the first tap is the greatest, we can conclude $m^2 < t_s$ using Equation (4.1) and Equation (4.2). Using these two inequalities and considering that $\hat{\tau}$ is closer to 0 than t_s, $\hat{\tau}$ terms whose order is higher than one can be neglected. Thus, Equation (4.10) can be approximately solved as

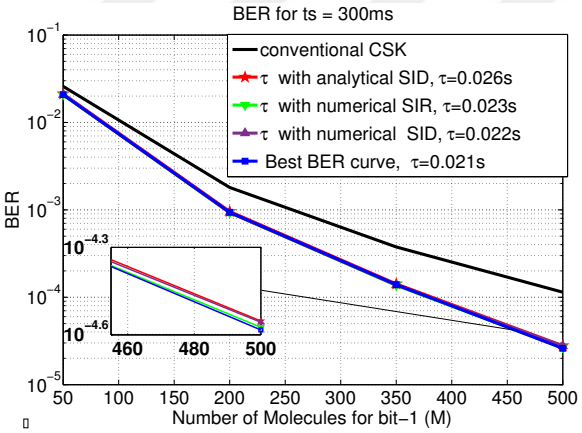
$$\hat{\tau} \approx \frac{84m^2t_s}{305t_s - 236m^2}. \tag{4.11}$$

4.3. Performance Evaluation

The performance evaluation of the proposed method in terms of BER is demonstrated using Monte Carlo simulations. The number of observed molecules at the receiver is modelled by the Gaussian distribution model given in [28]. The BER curves are obtained for five cases: (i) conventional CSK; (ii) using τ obtained from Equation (4.11) (τ with analytical SID); (iii) using τ obtained by solving Equation (4.6) using exhaustive simulations (τ with numerical SID); (iv) using τ obtained by solving Equation (4.4) using exhaustive simulations (τ with numerical SIR); and (v) the best BER curve obtained with this method using exhaustive simulations. For each case, the threshold is optimized to obtain the best performance of the systems. For all simulations, 10^5 consecutive bits are transmitted using the mentioned methods in a channel,



(a) Error performance for proposed method with $t_s = 0.2$s, $r_r = 5\mu m$, $r_0 = 10\mu m$ and $D = 80\mu m^2/s$



(b) Error performance for proposed method with $t_s = 0.3$s, $r_r = 5\mu m$, $r_0 = 10\mu m$ and $D = 80\mu m^2/s$

Figure 4.2. Performance comparison of the analytical and numerical τ optimization for proposed method with conventional CSK and the best BER curve.

whose memory is $K = 20$ for $t_s = 0.3$s, and modified for other scenario accordingly.

As can be seen in Figure 4.2, τ values calculated analytically give almost the same BER curves with τ values calculated numerically. As expected, the proposed method outperforms CSK. In the same figure, one can easily observe that different methods applied to find τ give almost the same BERs, and these values are almost equal to the best BER curve that can be obtained with this method. This justifies that SIR and SID can be good alternatives that can be used for minimization of BER in the molecular communication literature. Furthermore, minimization of SID can also be

achieved analytically, which is an important asset.

4.4. Concluding Remarks on Optimal Reception Delay in Diffusion-based Molecular Communication

In this letter, we propose a simple and novel receiver-based equalization method to mitigate ISI in a MCvD channel for CSK modulation. First, we observe that molecules absorbed by the receiver until τ seconds from the beginning of the transmission slot most probably belong to the previous symbols. Therefore, these molecules do not give any information about the intended symbol. Due to this observation, we propose to shift the decoding slot by τ seconds to have an incremental effect on the performance of CSK. We, then, define the new channel taps for proposed system and minimize SIR numerically and SID, both numerically and analytically, to find the optimal τ. Moreover, best BER values that can be achieved using this method are also determined to compare the performance of the proposed objective functions. Monte Carlo simulations are performed to demonstrate the superior performance of the proposed method with respect to CSK. We also demonstrate that proposed objective functions succesfully achieve almost lowest BER curves that are obtained by exhaustive computer simulations. The main benefits of this method are its simplicity, (it does not require any additional molecule type or computational burden) and it can be combined with other equalization and modulation methods to achieve lower BERs.

5. MOLECULAR SIGNAL MODELING OF A PARTIALLY COUNTING ABSORBING SPHERICAL RECEIVER

In addition to modulation and equalization methods, it is possible to design the receiver considering the MCvD channel properties. Due to the nature of diffusion, it can be expected that the molecules received in the back lobe of the receiver will most possibly take longer time to reach that point than the molecules received in the front lobe. Therefore, the molecules absorbed in the back lobe most likely belong to the previous transmitted symbols. Thus, they contribute to ISI. Inspired from this idea, we model the received molecular signal for a partially counting absorbing receiver. That is, the receiver absorbs all hitting molecules but those counted are only the ones hitting at a specific site. Modeling the time-dependent received signal for such a system is an open issue and has the potential to enhance the communication system performance without any significant additional cost. Most of the received molecules are absorbed from the surface area facing towards the transmitter side. As the path to the back side of the receiver is longer, the receptors on that side are more likely to receive the interference molecules. Therefore, limiting the counting area to the front side with a limited surface area enhances the signal quality. The main contributions of this work are listed as follows:

- The derivation of the joint distribution of the received molecules with respect to time and angle
- The modeling the received signal
- The investigating of the signal properties, and
- The finding of the optimal region for counting

for a partially counting and absorbing spherical receiver in a diffusion-based MC system.

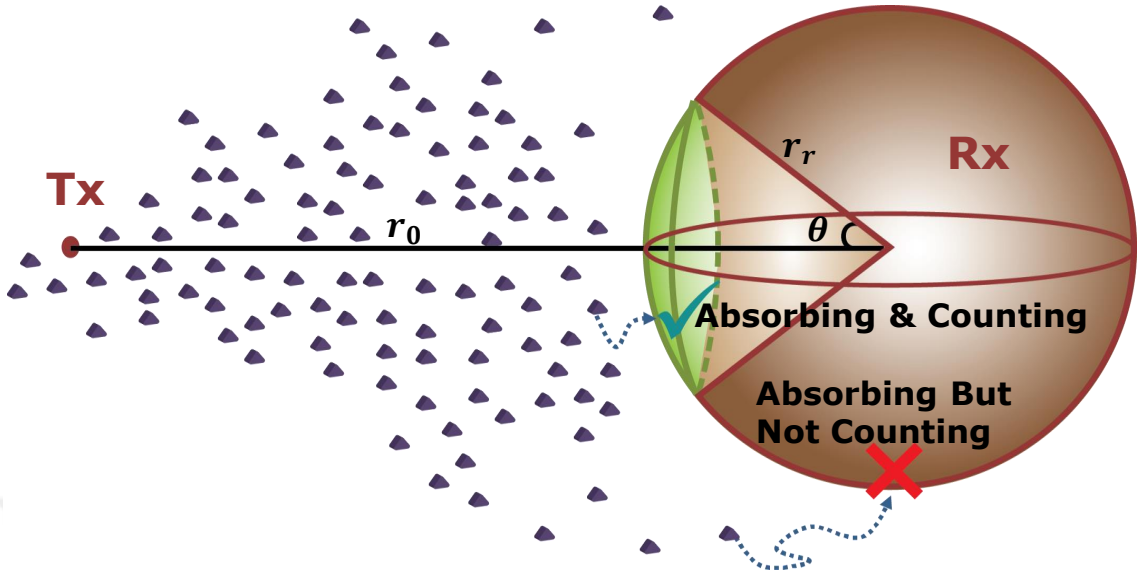


Figure 5.1. System model of a diffusion-based MC with a point transmitter and a partially counting absorbing spherical receiver.

5.1. System Model

The received signal in a diffusion-based MC system is affected by three main processes: emission, propagation, and reception. Analytical derivations for the channel model depend on the emitter, the receiver, the environment, and the propagation dynamics. Therefore, we give the details of the system before deriving the channel model.

5.1.1. Topology Model

We consider a diffusion-based MC system with one point transmitter and one spherical receiver in a 3-D environment (Figure 5.1). Novel feature of the receiver (Rx) is the ability to count the molecules absorbed only through a specific region. This feature complicates the modeling procedure of the received signal. In Figure 5.1, the circular cap facing towards the transmitter node (Tx) counts the absorbed molecules while the rest of the surface area absorbs but does not count the molecules.

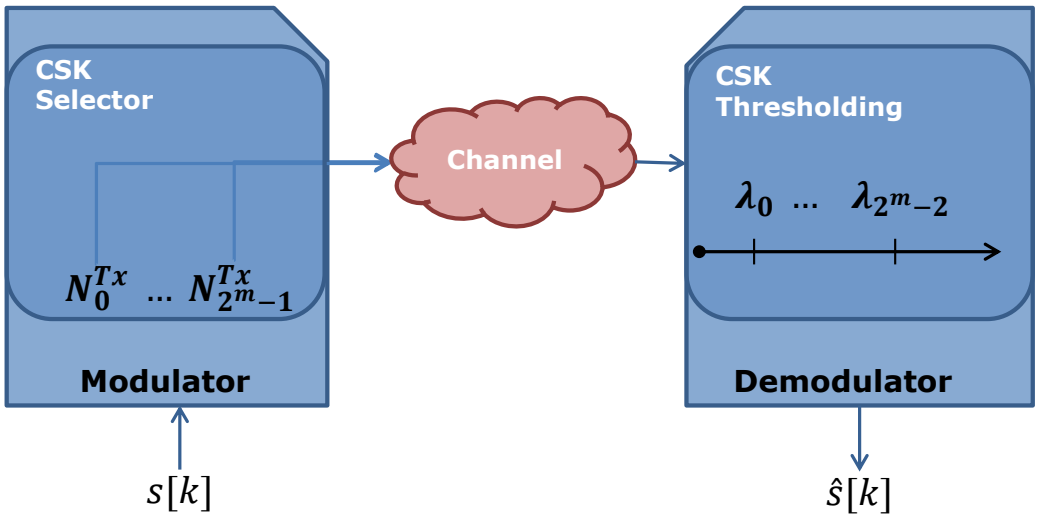


Figure 5.2. Concentration-based modulator and demodulator.

As shown in Figure 5.1, the molecules propagate by the diffusion process when they are emitted from the Tx point. The distance between the emission point and the center of the receiver is denoted by r_0 and the radius of the absorbing receiver is denoted by r_r. The circular cap that counts is determined by the θ angle, which we name as the *counting region*.

It is assumed that Tx and Rx nodes are fully synchronized in the time domain, and the interactions between diffusing molecules are ignored. No environmental or counting circuit noise is considered; only the diffusion noise is considered to isolate the signaling gain due to partial counting receiver. Furthermore, it is assumed that a mechanism in Rx node determines the direction of Tx and aligns its *counting region* facing towards Tx.

5.1.2. Modulation and Demodulation

In this paper, concentration shift keying (CSK) based modulation technique is used. General form of CSK is introduced in [6, 42]. In CSK based modulation techniques, the information is modulated on the amount of the transmitted molecules at the start of each symbol duration (t_s).

General structure of CSK based modulations is depicted in Figure 5.2. For the k^{th} symbol $s[k]$, the modulator maps the symbol to the amount of molecules to emit (i.e., i^{th} symbol is mapped to N_i^{Tx}) at the start of the k^{th} symbol duration. Depending on the modulation order (m), the number of possible symbols is determined and equals to 2^m. In this paper, we use binary CSK where $m = 1$, i.e., it has two symbols s_0 and s_1 which represent bit-0 and bit-1, respectively. After CSK selector maps the symbol to the amount, the molecules are emitted to the channel and they propagate by diffusion. During the symbol duration, the arriving molecules are absorbed and counted according to the counting logic. At the end of the k^{th} symbol duration, the final value is thresholded for obtaining the detected symbol $\hat{s}[k]$ for the k^{th} symbol.

5.2. Channel Model for Partial Counting Receiver

The joint cumulative angle and time distribution of absorbed molecules at the spherical receiver that are released by a point transmitter deserves an analytical derivation. This distribution function is utilized to determine the channel taps analytically for the proposed receiver.

In the literature, marginal cumulative distribution with respect to time is derived for a fully absorbing spherical receiver and introduced to the MC domain from a communication perspective [13] as

$$F_{\mathrm{hit}}(t, r_0, r_r) = \frac{r_r}{r_0} \operatorname{erfc}\left(\frac{r_0 - r_r}{\sqrt{4Dt}}\right), \tag{5.1}$$

where $\operatorname{erfc}(\cdot)$ is the complementary error function. Furthermore, the marginal angular distribution of the received molecules is given in [15] (6.3.3a) for a 3-D medium when the time goes to infinity as

$$p(\theta) = 2\pi r_r^2 \sin\theta \; \epsilon(\theta)\,, \tag{5.2}$$

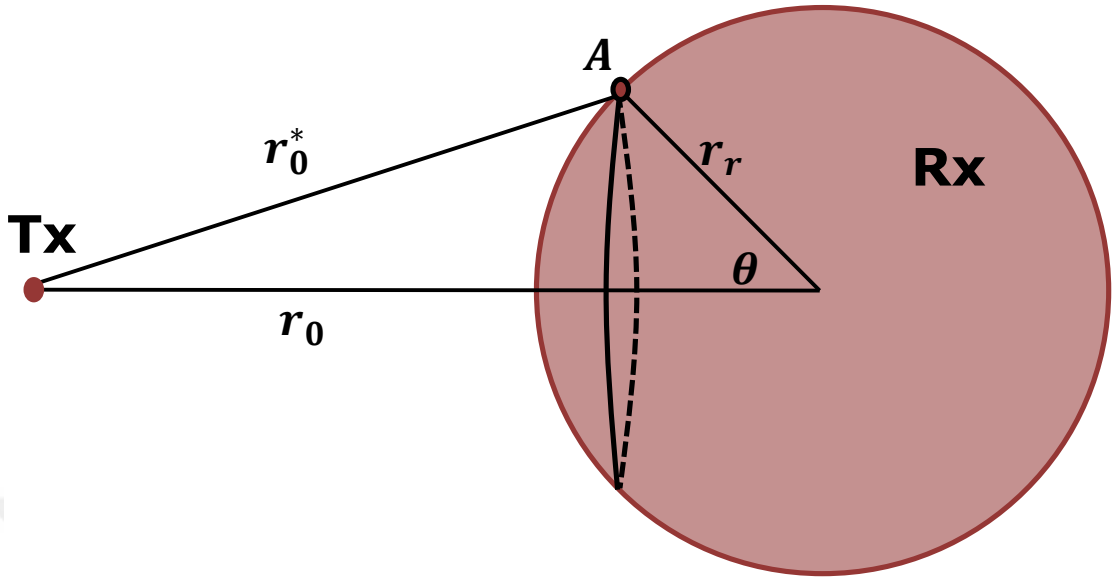


Figure 5.3. An infinitesimally small sphere over the circular region on the surface of the sphere. The circular region is determined by the angle θ.

where

$$\epsilon(\theta) = \frac{\left(1 - \frac{r_r^2}{r_0^2}\right)}{4\pi r_r r_0 \left(1 - \frac{2r_r}{r_0}\cos\theta + \frac{r_r^2}{r_0^2}\right)^{3/2}}. \tag{5.3}$$

In particular, $p(\theta)$ in Equation (5.2) gives the distribution of the molecules absorbed by the cap that is defined by the angle θ which is presented in Figure 5.3. This function is plotted for different parameters in Figure 5.4. As can be seen, the probability of absorption has a peak between $\theta = 0°$ and $\theta = 90°$. Furthermore, it is zero for $\theta = 0°$ and $\theta = 180°$. These are not surprising since $\theta = 0°$ and $\theta = 180°$ represent only a point on the surface. Therefore, the probability of absorption in these regions is zero although $\theta = 0°$ is the closest point to the transmitter. If we increase θ, we expect to have more received molecules since the circular region gets bigger. However, after some point, the rate of increase is not enough compared to the decrease in the hitting rate that can also be observed in Figure 5.4.

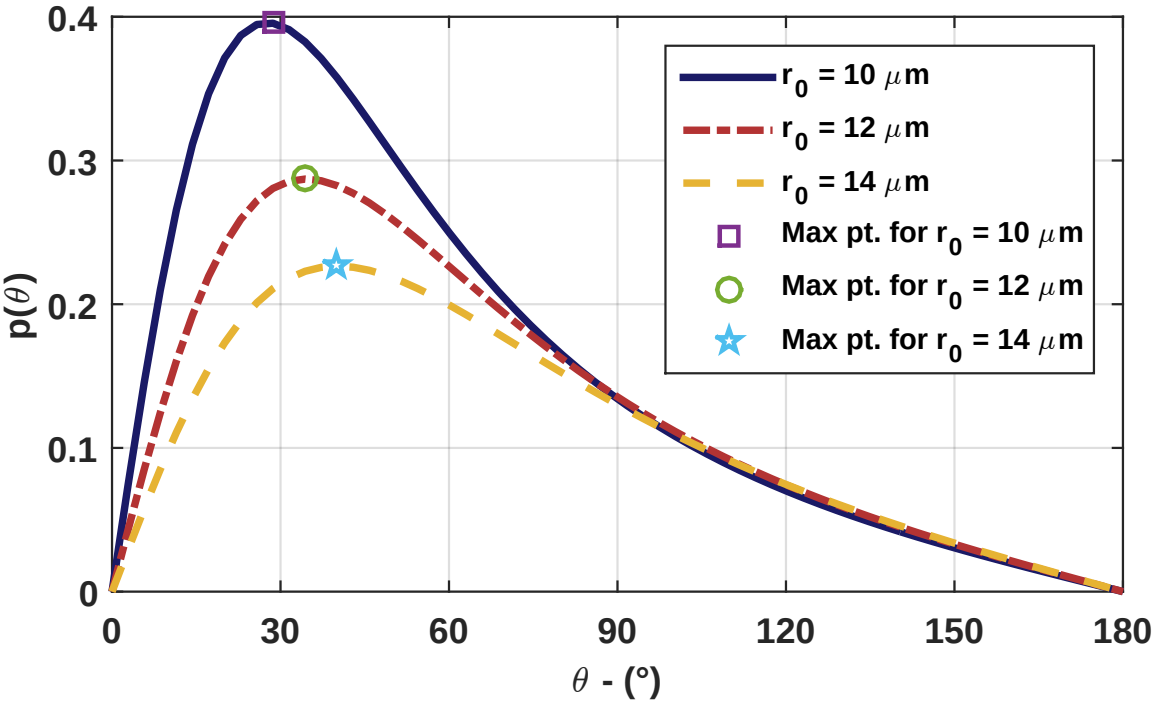


Figure 5.4. Theta versus $p(\theta)$ curves for different r_0 values ($r_r = 5\,\mu\text{m}$ and $D = 80\,\mu\text{m}^2/\text{s}$). Maximum values are attained at 28.6°, 34.3°, and 40.1°.

Since the communication process occurs in a limited time, we need to obtain the joint distribution of absorbed molecules with respect to time and angle to apply partially counting receiver system in MC. To the best of our knowledge, the joint distribution with respect to time and angle has not been derived yet. By utilizing Equation (5.1) and Equation (5.2), we find an approximate analytical closed-form expression for the joint cumulative distribution with respect to time and angle. The main concept of our approach is to cover the desired region on the surface of the spherical receiver with infinitesimally small spheres and to evaluate the absorption probability of these spheres. In other words, each small patch of the spherical receiver is represented by an infinitesimally small sphere. First, we evaluate the probability of absorption of a molecule by a small sphere placed with an angle θ, as shown in Figure 5.3. Accordingly, we add up the probability of absorption of the all small spheres that are placed with the same angle, as shown in Figure 5.5, to obtain the overall probability of absorption for a given θ angle.

When we consider an infinitesimally small sphere with radius dr placed at the surface making an angle of θ with the center of the sphere as shown in Figure 5.3, the distance of this arbitrarily placed sphere to the point transmitter can be calculated using Cosine rule as

$$r_0^* = \sqrt{r_0^2 + r_r^2 - 2r_0 r_r \cos\theta}. \tag{5.4}$$

Considering Figure 5.3, if only the small sphere at point A is available in the environment, the probability of absorption of molecules until time t could be obtained using Equation (5.1) as

$$F_{\text{hit}}^A(t) = F_{\text{hit}}(t, r_0^*, dr) = \frac{dr}{r_0^*} \operatorname{erfc}\left(\frac{r_0^* - dr}{\sqrt{4Dt}}\right). \tag{5.5}$$

Since $r_0^* \gg dr$, we have $r_0^* - dr \approx r_0^*$ and, therefore, Equation (5.5) can be rewritten as

$$F_{\text{hit}}^A(t) \approx \frac{dr}{r_0^*} \operatorname{erfc}\left(\frac{r_0^*}{\sqrt{4Dt}}\right). \tag{5.6}$$

In particular, Equation (5.6) is not valid when the small sphere at point A is placed on the surface of the spherical receiver (also called the big sphere). Since some part of the small sphere lies behind the surface of the big sphere, this part does not act as a receiver. Furthermore, the receptors of the big sphere are also active; hence, this event should also be taken into account. Let A^* be the region of the active receptors of the small sphere placed at point A. Clearly, our aim is to derive the probability

of absorption of a molecule until time t with the active regions of the small sphere at point A, which is $F_{\text{hit}}^{A^*}(t)$.

It is important to note that for different θ angles, the orientation of the active regions of the small sphere is different, which can be observed from Figure 5.6. This implies that, for different θ angles, $F_{\text{hit}}^{A^*}(t)$ should have a different formula that determines the orientation of the active regions. Therefore, in order to derive $F_{\text{hit}}^{A^*}(t)$, an angle factor $T(\theta)$ and a time factor $\phi(t)$ are multiplied by $F_{\text{hit}}^{A}(t)$ as the adjusting factors to determine this difference. These adjusting factors will be derived using marginal cases later on. Therefore, $F_{\text{hit}}^{A^*}(t)$ can be represented as

$$F_{\text{hit}}^{A^*}(t) = K\, F_{\text{hit}}^{A}(t)\, T(\theta)\phi(t), \tag{5.7}$$

where K is the probability that a molecule is not absorbed by the small spheres other than the small sphere at point A until time t. Since the probability of absorption of a molecule by the whole part of the receiver is $F_{\text{hit}}(t, r_0, r_r)$, and since the probability of absorbing of a molecule by the small sphere at point A is $F_{\text{hit}}^{A^*}(t)$, K can be written as $K = 1 - (F_{\text{hit}}(t, r_0, r_r) - F_{\text{hit}}^{A^*}(t))$. Thus, $F_{\text{hit}}^{A^*}(t)$ can be written as

$$\begin{aligned} F_{\text{hit}}^{A^*}(t) &= [1 - F_{\text{hit}}(t, r_0, r_r) + F_{\text{hit}}^{A^*}(t)]\, F_{\text{hit}}^{A}(t)\, T(\theta)\phi(t) \\ &\approx [1 - F_{\text{hit}}(t, r_0, r_r)] \frac{dr}{r_0^*} \operatorname{erfc}\left(\frac{r_0^*}{\sqrt{4Dt}}\right) T(\theta)\phi(t), \end{aligned} \tag{5.8}$$

since $F_{\text{hit}}^{A^*}(t) \ll F_{\text{hit}}(t, r_0, r_r)$.

After approximating $F_{\text{hit}}^{A^*}(t)$, for a small sphere placed at point A that makes θ angle as shown in Figure 5.3, the next step is to find the total number of small spheres that have same θ angle. These small spheres are lined up on a circle with radius $r_r \sin(\theta)$, as shown in Figure 5.5. Since the radius of these spheres are infinitesimal, the total number of spheres on this circle can be calculated by dividing the circumference of the circle to the diameter of the small sphere as $N_\theta = \dfrac{2\pi r_r \sin(\theta)}{2dr}$. Using this N_θ,

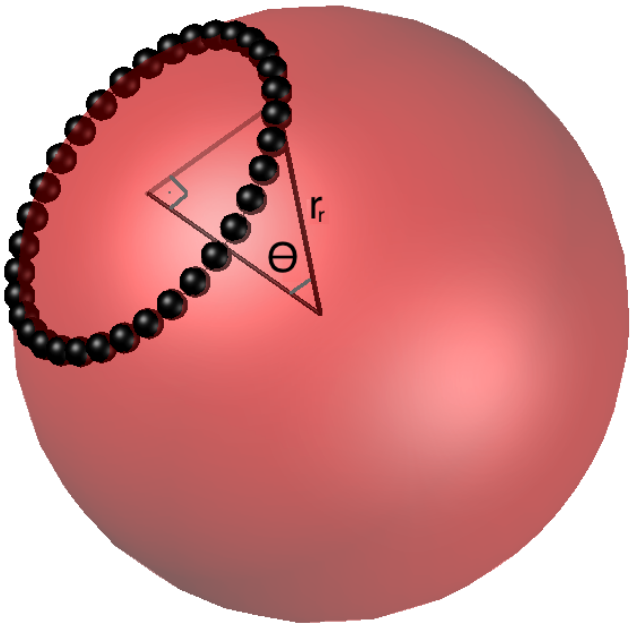


Figure 5.5. Small spheres that make θ angle with the center of the big sphere. Note that, these spheres are lined up on a circle whose radius is $r_r sin(\theta)$.

the probability of absorption of a molecule until time t by any small sphere that makes same θ angle as the sphere at point A can be obtained by

$$\begin{aligned} p(\theta,t) &= N_\theta F_{\text{hit}}^{A^*}(t) \\ &= \pi r_r \sin(\theta)[1-F_{\text{hit}}(t,r_0,r_r)]\frac{\text{erfc}\left(\frac{r_0^*}{\sqrt{4Dt}}\right)}{r_0^*}T(\theta)\phi(t). \end{aligned} \tag{5.9}$$

When t goes to infinity, Equation (5.9) becomes equal to Equation (5.2), where $\lim_{t\to\infty} p(\theta,t)$ can be obtained as

$$\lim_{t\to\infty} p(\theta,t) = \pi r_r \sin\theta(1-\frac{r_r}{r_0})\frac{1}{r_0^*}T(\theta)\lim_{t\to\infty}\phi(t). \tag{5.10}$$

Hence, equalizing Equation (5.2) and Equation (5.10) gives us $T(\theta)$ as

$$T(\theta) = \frac{2r_r r_0^*\,\epsilon(\theta)}{\left(1-\frac{r_r}{r_0}\right)\lim_{t\to\infty}\phi(t)}. \tag{5.11}$$

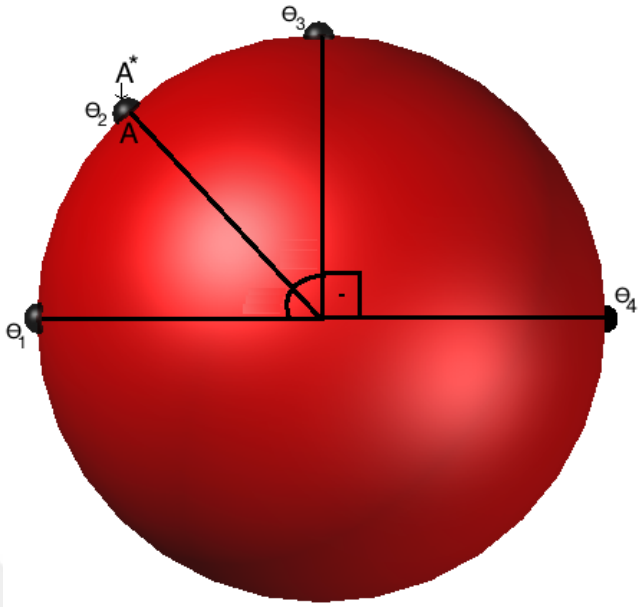


Figure 5.6. Demonstration for the orientation of the active receptors of the small sphere (A^*) for different θ values. For the small sphere with $\theta_1 = 0$, the active receptors are towards to the direction of the transmitter. As θ changes, the orientation of the active receptors changes hence this change should be taken into account.

Although Equation (5.11) contains $\lim_{t\to\infty} \phi(t)$, in the following steps this term is canceled out and $p(\theta, t)$ does not involve any limit term.

The next step is deriving the other compensation function, $\phi(t)$. Note that $p(\theta, t)$ gives the distribution of molecules with respect to angle θ until time t. Therefore, taking the integral of $p(\theta, t)$ with respect to θ from $\theta = 0$ to an arbitrary angle α, gives the cumulative distribution of molecules at the receiver with respect to time and angle as

$$F(\alpha, t) = \int_0^{\alpha} p(\theta, t)\, d\theta. \tag{5.12}$$

In Equation (5.12), one can easily observe that, when $\alpha = \pi$, all of the surface of the receiver is absorbing. Therefore, $F(\pi, t)$ is equal to the marginal cumulative function given in Equation (5.1) as $F(\pi, t) = F_{\text{hit}}(t, r_0, r_r)$. By using this equality, we can obtain

$\phi(t)$ as

$$\phi(t) = \frac{F_{\text{hit}}(t, r_0, r_r)}{\pi r_r[1 - F_{\text{hit}}(t, r_0, r_r)]S_\pi}, \tag{5.13}$$

where

$$S_\pi = \int_0^\pi \frac{\sin\theta}{r_0^*} \operatorname{erfc}\left(\frac{r_0^*}{\sqrt{4Dt}}\right) T(\theta)\, d\theta. \tag{5.14}$$

Note that the denominator of $\phi(t)$ contains $T(\theta)$. Since $\lim_{t\to\infty} \phi(t)$ term in this integral can be taken outside of the integral, we conclude that $\phi(t)$ involves $\lim_{t\to\infty} \phi(t)$ in the numerator while $T(\theta)$ involves this term in the denominator. Therefore, multiplying these two compensation functions together cancels $\lim_{t\to\infty} \phi(t)$ terms in $p(\theta, t)$.

After finding $\phi(t)$, we can write $p(\theta, t)$ as

$$p(\theta, t) = \frac{\sin\theta \operatorname{erfc}\left(\frac{r_0^*}{\sqrt{4Dt}}\right) F_{\text{hit}}(t, r_0, r_r)}{(1 - \frac{2r_r}{r_0}\cos\theta + \frac{r_r^2}{r_0^2})^{\frac{3}{2}} \int_0^\pi \frac{\sin\theta' \operatorname{erfc}\left(\frac{r_0^*}{\sqrt{4Dt}}\right)}{(1 - \frac{2r_r}{r_0}\cos\theta' + \frac{r_r^2}{r_0^2})^{\frac{3}{2}}}\, d\theta'}. \tag{5.15}$$

Figure 5.7 is the heat map of $p(\theta, t)$ that gives the angular distribution of a molecule until time t. Considering this figure, some interesting inferences can be obtained. Firstly, the molecules will accumulate less at the higher angles compared to the lower angles. This is expected since as the angle increases, the distance also increases, which leads to the diminishing of the probability of absorption. Second and more interestingly, at very small angles around zero that can also be considered as the line of sight angles, the probability of absorption is even lower compared to other angles. This is a consequence of the fact that the number of small spheres is very limited for these angles (when $\theta = 0°$, there is only one point and consequently one small sphere); hence, the probability of absorption at these angles is quite low.

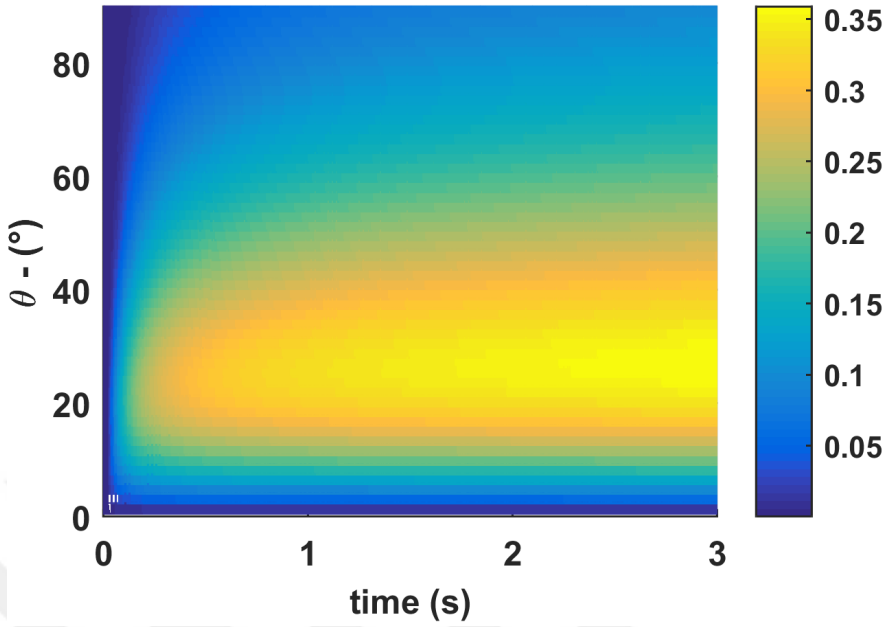


Figure 5.7. $p(\theta, t)$ heat map for $r_r = 5\,\mu\text{m}$, $r_0 = 10\,\mu\text{m}$, and $D = 80\,\mu\text{m}^2/\text{s}$.

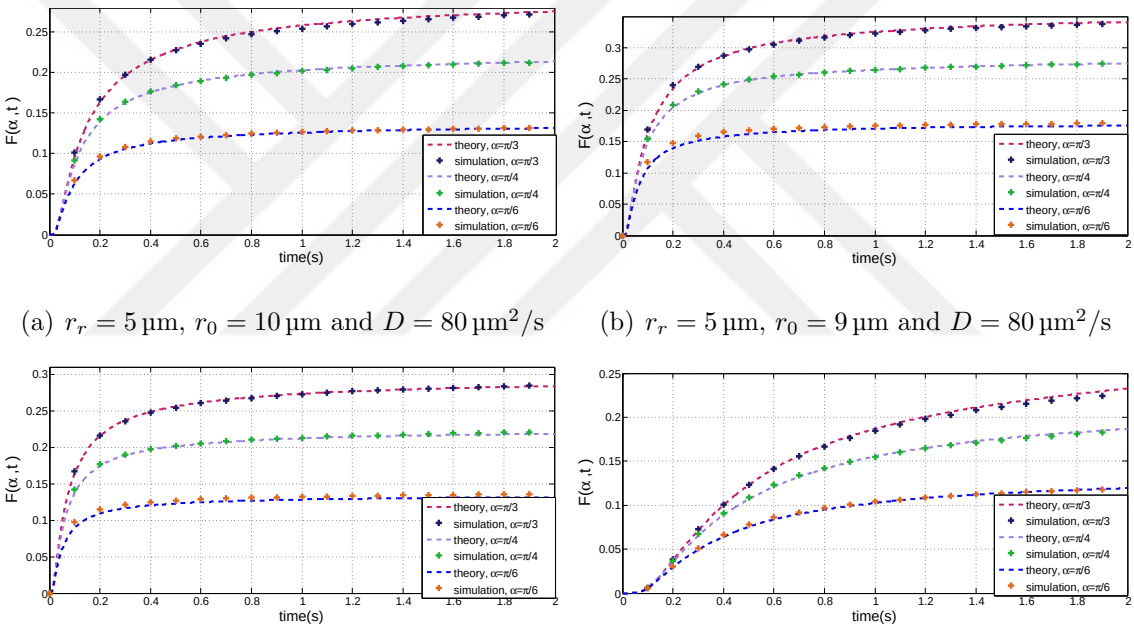
$$F(\alpha,t) = \text{erfc}\left(\frac{(r_0 - r_r)}{\sqrt{Dt}}\right)\sqrt{Dt}\times$$
$$\frac{\left(\text{erfc}\left(\frac{\sqrt{r_0^2-2r_0r_r+r_r^2}}{\sqrt{4Dt}}\right)\sqrt{Dt} + \frac{\sqrt{r_0^2-2r_0r_r+r_r^2}}{2\sqrt{\pi}}Ei\left(-\frac{\left(r_0^2-2r_0r_r+r_r^2\right)}{4Dt}\right)\right.}{U(t)\quad Dt\sqrt{\frac{r_0^2-2r_0r_r+r_r^2}{r_0^2}}} - \tag{5.16}$$
$$\frac{\left.\text{erfc}\left(\frac{\sqrt{r_0^2-2r_0r_r\cos(\alpha)+r_r^2}}{\sqrt{4Dt}}\right)\sqrt{Dt} + \frac{\sqrt{r_0^2-2r_0r_r\cos(\alpha)+r_r^2}}{2\sqrt{\pi}}Ei\left(-\frac{\left(r_0^2-2r_0r_r\cos(\alpha)+r_r^2\right)}{4Dt}\right)\right)}{U(t)\quad Dt\sqrt{\frac{r_0^2-2r_0r_r\cos(\alpha)+r_r^2}{r_0^2}}}$$

Note that $p(\theta, t)$ is used for calculating $F(\alpha, t)$ that is given in Equation (5.12) and results in Equation (5.16) where $Ei(.)$ is an exponential integral function and $U(t)$ is given in Equation (5.17).

Once $F(\alpha, t)$ is obtained, the channel tap for the n^{th} symbol duration, p_n, can be obtained (for a given counting region (α) and symbol duration t_s) as

$$p_n(\alpha) = F(\alpha, nt_s) - F(\alpha, (n-1)t_s)\,. \tag{5.18}$$

$$U(t) = \int_0^{\pi} \operatorname{erfc}\left(\frac{\sqrt{r_0^2 + r_r^2 - 2r_0 r_r \cos\theta}}{\sqrt{4Dt}}\right)\left(1 - \frac{2r_r}{r_0}\cos\theta + \frac{r_r^2}{r_0^2}\right)^{-3/2} \sin\theta \quad d\theta = \frac{{r_0}^2\left(Dt \quad \operatorname{erfc}\left(\frac{r_0 - r_r}{\sqrt{4Dt}}\right) + \sqrt{Dt}\,(r_0 - r_r)\,Ei\left(-\frac{(r_0 - r_r)^2}{4Dt}\right)\right)}{Dr_r t(r_0 - r_r)} - \frac{{r_0}^2\left(Dt \quad \operatorname{erfc}\left(\frac{r_0 + r_r}{\sqrt{4Dt}}\right) + \sqrt{Dt}\,(r_0 + r_r)\,Ei\left(-\frac{(r_0 + r_r)^2}{4Dt}\right)\right)}{Dr_r t(r_0 + r_r)} \tag{5.17}$$



(a) $r_r = 5\,\mu\text{m}$, $r_0 = 10\,\mu\text{m}$ and $D = 80\,\mu\text{m}^2/\text{s}$ (b) $r_r = 5\,\mu\text{m}$, $r_0 = 9\,\mu\text{m}$ and $D = 80\,\mu\text{m}^2/\text{s}$

(c) $r_r = 5\,\mu\text{m}$, $r_0 = 10\,\mu\text{m}$ and $D = 160\,\mu\text{m}^2/\text{s}$ (d) $r_r = 10\,\mu\text{m}$, $r_0 = 20\,\mu\text{m}$ and $D = 80\,\mu\text{m}^2/\text{s}$

Figure 5.8. Comparison of the derived analytical cumulative function $F(\alpha, t)$ with simulation results for $\alpha = \pi/3$, $\pi/4$ and $\pi/6$ from top to bottom.

5.3. Channel Model Validation and Molecular Signal Properties

5.3.1. Received Signal Validation

Once the analytical distribution of the received molecules for partially counting system is obtained, the next step is to compare it with the simulation results obtained by using Equation (1.1). As can be seen in Figure 5.8, validation is done for various parameters with different α values, and simulation and theoretical results are coherent.

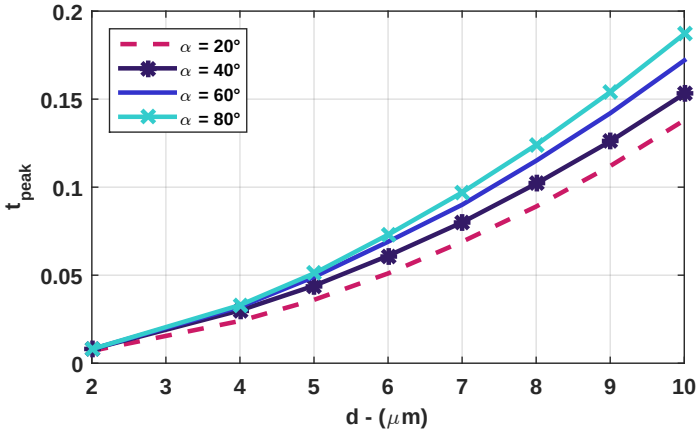


Figure 5.9. t_{peak} vs d curves for $r_r = 5\,\mu\text{m}$, $D = 80\,\mu\text{m}^2/\text{s}$.

5.3.2. Peak Time

The communication literature considers the peak time, t_{peak}, to be a crucial property for characterizing the channel, and defines it as the time that the received signal makes a peak at the receiver. In [13], it is concluded that, for the fully absorbing receiver, t_{peak} is proportional with the square of $d = r_0 - r_r$, which is the shortest distance from the transmitter to the receiver's surface. This is a major drawback in molecular communication via diffusion (MCvD) channels since as d increases, the data rate exponentially decreases to capture the signal until its peak, while in electro-magnetic communication this decrement is linear. We evaluate t_{peak} by taking the derivative of $F(\alpha, t)$ with respect to time, which is the hitting rate of the molecules for a given α, and examine its maximum value. As can be seen in Figure 5.3.1, t_{peak} is still directly proportional with d^2, which is the same with the fully absorbing receiver case.

5.3.3. Optimum α for the Given Channel Parameters

The optimum reception angle, α^*, of the receiver in terms of bit error rate (BER) is determined by finding the position of the global minimum of BER formula of the CSK modulation with respect to α. On the other hand, closed-form of the BER formula in CSK is not a tractable function if the number of the channel taps is high. Therefore, we use an alternative objective function whose argument of the global maximum is

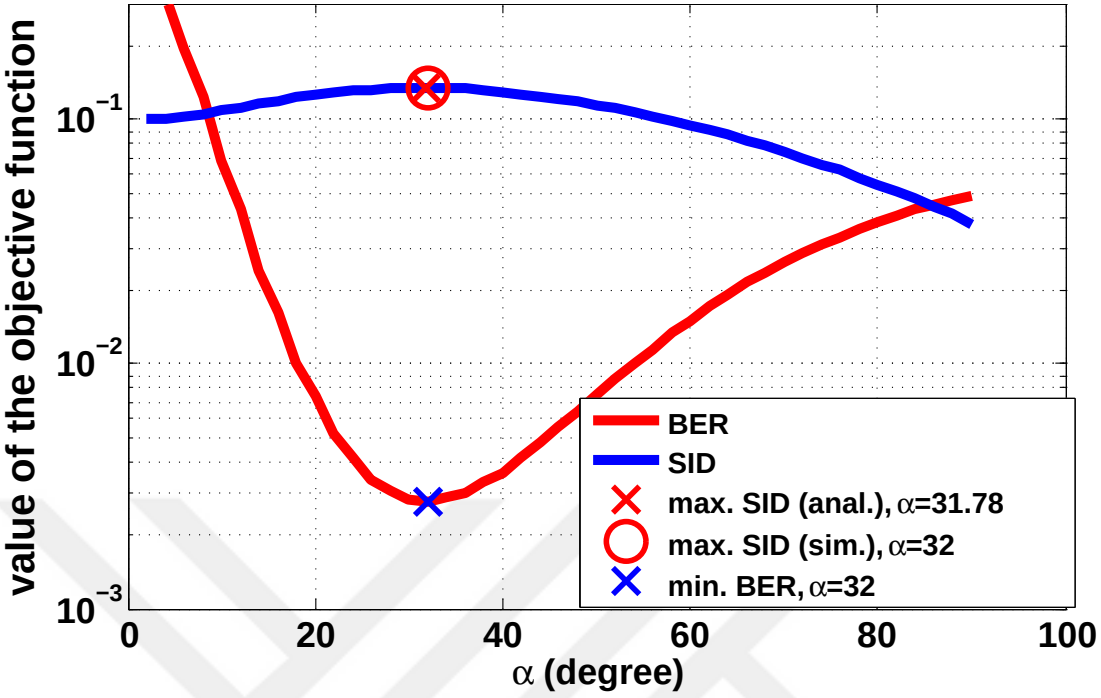


Figure 5.10. BER vs α curves and corresponding SID curves for $r_r = 5\,\mu\text{m}$, $r_0 = 10\,\mu\text{m}$, $D = 80\,\mu\text{m}^2/\text{s}$, $t_s = 150ms$ with minimum point of BER function obtained via computer simulations and maximum of the SID function obtained with both simulation and analytical solution.

almost the same as the argument of the global minimum of BER as proposed in [43]. This function is named as the signal to interference difference (SID), and gives the difference between the first tap and the sum of the other taps:

$$\text{SID} = p_1(\alpha) - \sum_{n=2}^{\infty} p_k(\alpha). \tag{5.19}$$

As shown in Figure 5.10, the argument of the global maximum of this function is very close to the argument of the global minimum of BER. Using SID, the corresponding optimization problem is written as

$$\alpha^* = \arg\max_{0\leq\alpha\leq\pi} \left[p_1(\alpha) - \sum_{k=2}^{\infty} p_k(\alpha) \right]. \tag{5.20}$$

Since $\sum_{k=2}^{\infty} p_k(\alpha) = F(\alpha, \infty) - p_1(\alpha)$, the optimization problem can be rewritten as

$$\alpha^* = \arg\max_{0 \le \alpha \le \pi} \left[2F(\alpha, t_s) - F(\alpha, \infty) \right]. \tag{5.21}$$

The solution of the optimization problem in Equation (5.21) can be solved by taking the derivative of the objective function with respect to α and equating it to zero with reasonable simplifications. The objective function $\text{SID} = [2F(\alpha, t_s) - F(\alpha, \infty)]$ can be written explicitly using Equation (5.16) and discarding the α independent terms as

$$\text{SID} = \frac{Y\left(a\text{erfc}\left(\frac{\sqrt{x}}{\sqrt{4a}}\right) + \frac{1}{\sqrt{2\pi}}\sqrt{a}\sqrt{x}Ei\left(-\frac{(x)}{4a}\right)\right)}{a\sqrt{x}} + \frac{M}{\sqrt{x}} \tag{5.22}$$

$$= \text{SID}_1 + \frac{M}{\sqrt{x}}, \tag{5.23}$$

where $x = r_0^2 - 2r_0 r_r \cos(\alpha) + r_r^2$, $a = Dt_s$, $Y = \frac{-2r_r F_{\text{hit}}(t_s, r_0, r_r)}{U(t_s)}$, and $M = -\frac{r_0^2 - r_r^2}{2r_0}$. Before taking the derivative of SID with respect to α, it needs to be simplified. SID_1 can be expanded to the series around $x = 0$, and, since x is on the order 10^{-12}, higher order terms can be neglected. Therefore, SID_1 can be written as $\text{SID}_1 \approx \frac{Y}{\sqrt{x}} + \frac{Y \log x}{2\sqrt{\pi a}}$. Using this approximation, SID can be written as

$$\text{SID} \approx \frac{Y}{\sqrt{x}} + \frac{Y \log x}{2\sqrt{\pi a}} + \frac{M}{\sqrt{x}}. \tag{5.24}$$

Taking the derivative of SID with respect to α and equating it to 0, we can arrive at

$$\frac{Y}{\sqrt{\pi a}(r_0^2 - 2r_0 r_r \cos(\alpha) + r_r^2)} = \frac{M+Y}{(r_0^2 - 2r_0 r_r \cos(\alpha) + r_r^2)^{1.5}}. \tag{5.25}$$

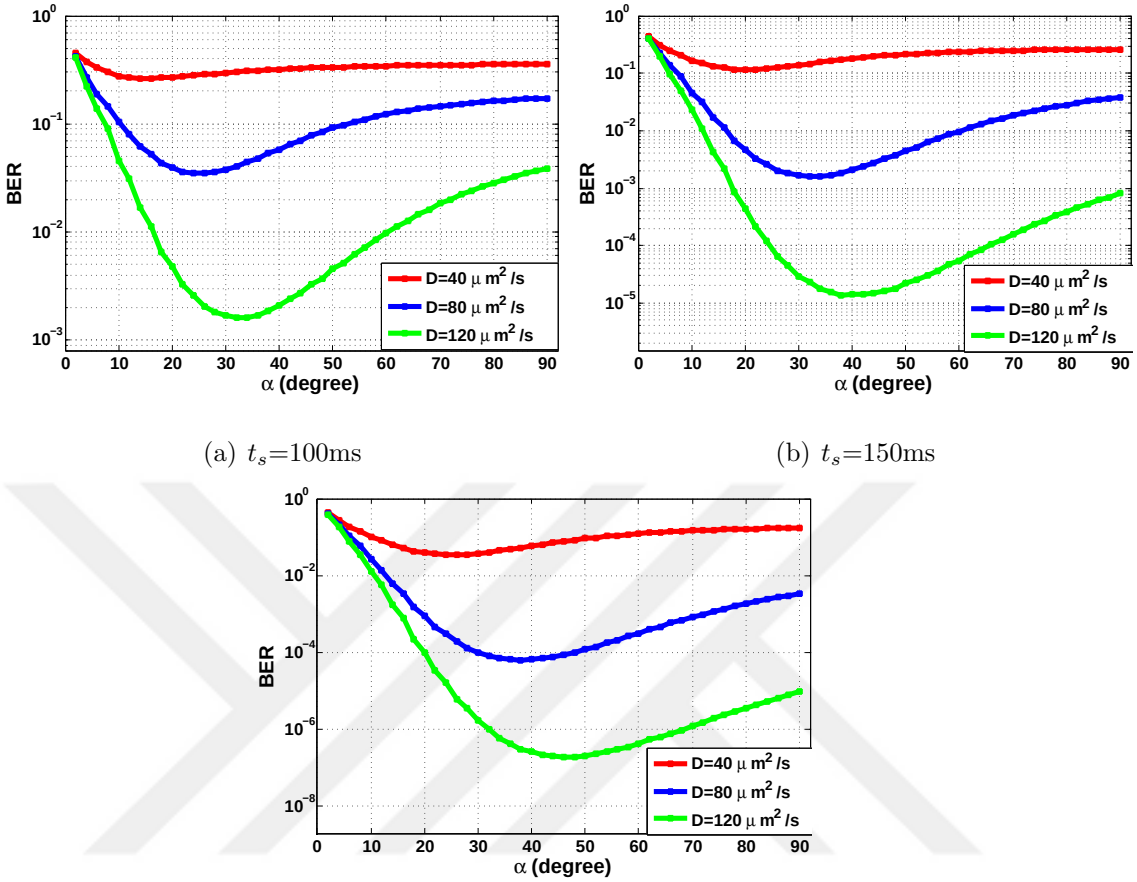
Hence, α^* can be obtained by solving Equation (5.25) as

$$\alpha^* = \cos^{-1}\left(\frac{r_0^2 + r_r^2 - \left(\frac{\sqrt{2a}(Y+M)}{Y}\right)^2}{2r_0 r_r}\right). \tag{5.26}$$

where $a = Dt_s$, $Y = \frac{-2r_r F_{\text{hit}}(t_s, r_0, r_r)}{U(t_s)}$, and $M = -\frac{r_0^2 - r_r^2}{2r_0}$.

5.4. Performance Analysis

In this section the performance analysis of the proposed system is examined for different parameters. We mainly evaluate the performance of the system in terms of BER. These evaluations are done using channel taps, both obtained analytically by using Equation (5.18) and Monte Carlo simulations by releasing 10^5 molecules from the transmitter and recording their arrival times and angles at the receiver. For all simulations, binary CSK is used. In particular, for transmitting the k^{th} symbol $s\,[k]\,\epsilon\,\{0,1\}$, the transmitter releases N_1^{Tx} molecules for $s\,[k] = 1$, and no molecule is released for $s\,[k] = 0$. Then, the number of absorbed molecules among those released molecules is determined using the channel taps and the Gaussian distribution assumption, as stated in [44]. At the end of the k^{th} symbol interval, the number of absorbed molecules is determined. Finally, the number is thresholded and $\hat{s}[k]$ is obtained. The threshold is always chosen to obtain the lowest BER to determine ultimate performance of the



(a) t_s=100ms (b) t_s=150ms

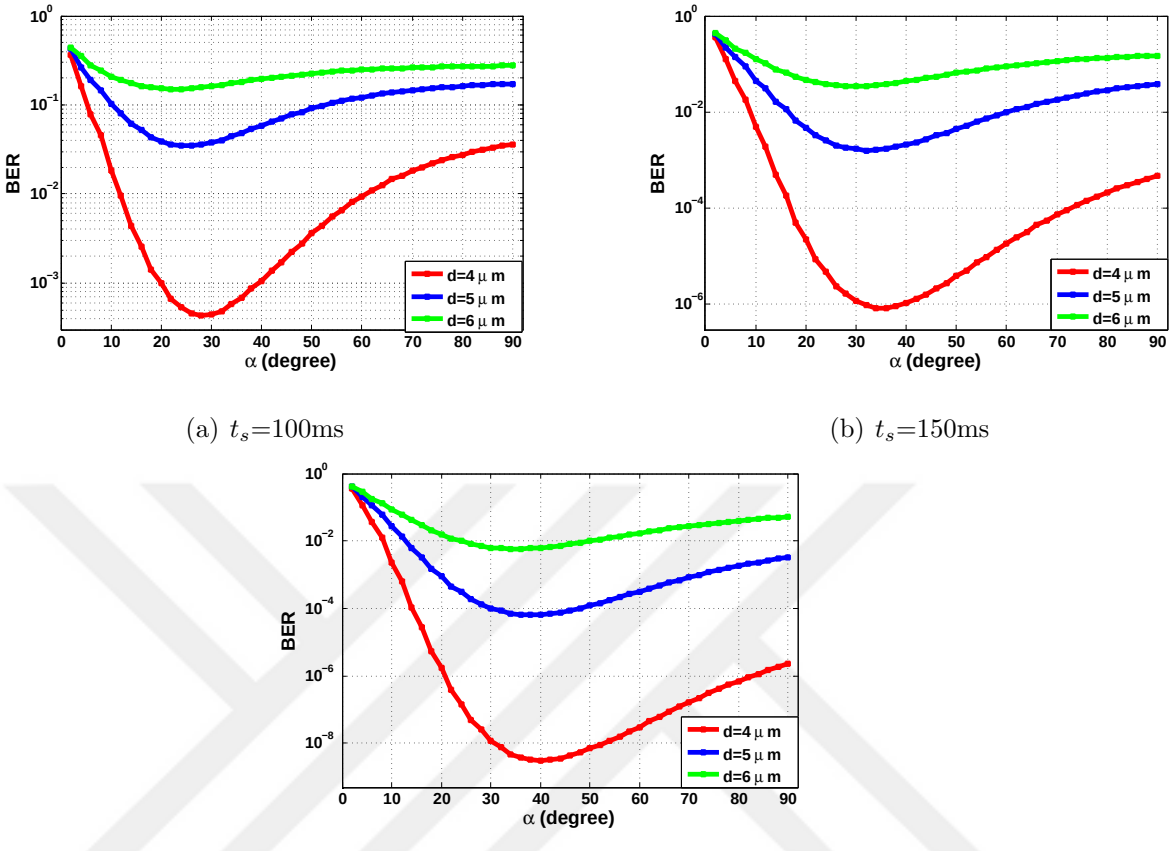
(c) t_s=200ms

Figure 5.11. BER vs α curves for $r_r = 5\,\mu\text{m}$, $r_0 = 10\,\mu\text{m}$ and $N_1^{\text{Tx}} = 500$ with different diffusion coefficient (D) values

.

systems. The performed simulations are conducted for 10^4 times, and for each simulation, 10^5 consecutive binary symbols are transmitted in a channel with memory is 100 symbols.

We firstly evaluate the performance of the proposed system with respect to d and the diffusion coefficient (D). As can be observed from Figure 5.11, the optimum α in terms of BER increases as D increases. This is expected since, as the molecules move faster, they can readily reach the further part of the receiver; hence, α should be increased in order not to miss the molecules coming during the current symbol slot. Similar results can be observed from Figure 5.12 where optimum α increases as the distance between the transmitter and the receiver decreases. Especially in the current



(a) t_s=100ms

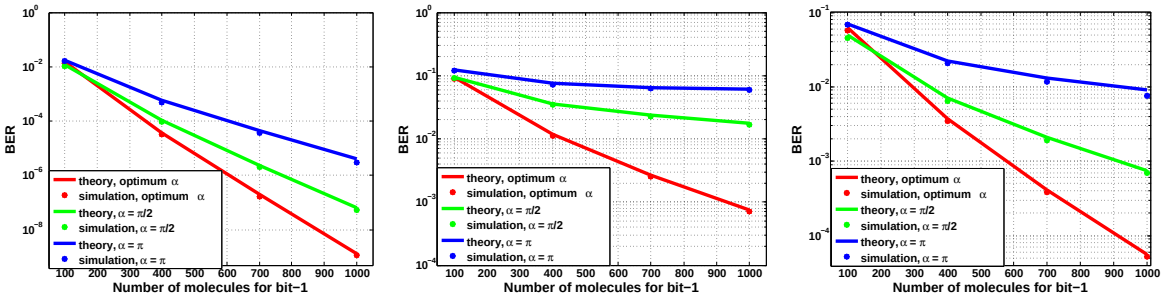
(b) t_s=150ms

(c) t_s=200ms

Figure 5.12. BER vs α curves for $r_r = 5\,\mu\text{m}$, $D = 80\,\mu\text{m}^2/\text{s}$ and $N_1^{\text{Tx}} = 500$ with different d=r_0-r_r values.

time slot, the molecules can move towards the further parts of the spherical receiver as the distance decreases or D increases. Therefore, the relative gain of the first tap compared to other taps increases by increasing α when the distance is shorter or D is higher. Furthermore, one can deduced from Figure 5.11 and 5.12, as t_s is increased the optimum α will also increases. This is also expected since optimum α will be 180° when t_s approaches to infinity.

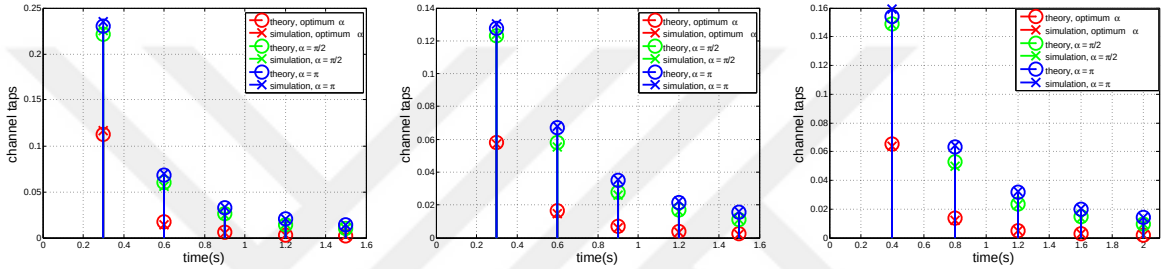
In Figure 5.13, we present the BER curves of three systems; $\alpha = \pi$ (conventional receiver), $\alpha = \pi/2$ (half sphere), and $\alpha = \alpha^*$ as well as their corresponding channel taps for both simulation and analytical results. Considering this figure, it can be concluded that the performance of the system will be significantly improved (BER is 10^2-10^4 times lower than the conventional receiver) if α is chosen properly. Although the signal tap is also decreased with this method, due to the decrease in ISI, this reduction is



(a) BER curves for t_s=300ms and $d = 5\,\mu\text{m}$

(b) BER curves for t_s=300ms and $d = 7\,\mu\text{m}$

(c) BER curves for t_s=400ms and $d = 7\,\mu\text{m}$



(d) channel taps for t_s=300ms and $d = 5\,\mu\text{m}$

(e) channel taps for t_s=300ms and $d = 7\,\mu\text{m}$

(f) channel taps for t_s=400ms and $d = 7\,\mu\text{m}$

Figure 5.13. Top: BER vs number of molecules per bit-1 (N_1^{Tx}) curves for $r_r = 5\,\mu\text{m}$, $D = 80\,\mu\text{m}^2/\text{s}$ with different $d = r_0$-r_r and t_s values with receiver $\alpha = 180°$ (convetional CSK), $\alpha = 90°$ and optimum α to obtain lowest BER. Bottom: corresponding channel taps of the communication systems.

compensated. Furthermore, it can also be seen that analytical solutions using Equation (5.16) and simulations are coherent.

5.5. Concluding Remarks

In this work, it has been confirmed that a partially counting absorbing receiver demonstrates a significant improvement over the conventional fully absorbing one. Considering the movement of molecules under diffusion process, we have proposed a counting region on the spherical receiver surface that faces towards the transmitter node. In order to justify this idea, we have derived the joint cumulative angle and the time distribution of the absorbed molecules at the receiver surface that had yet to be derived in the literature. Using this function and simulations, we have observed that the

molecules are likely to be accumulated with a certain range of angles, which satisfies our claim. We, then, have examined the received signal model for various parameters. The optimum counting region to obtain the lowest BER was also derived. We have presented here evidence of the improved performance of the proposed system.

6. NANONETWORKS

In this section we propose network coding algorithms for single-hop and multi-hop nanonetworks.

6.1. Single-hop Molecular Nanonetworks

Although there are several modulation and equalization techniques presented in the literature to tackle the ISI [34, 45], the effectiveness of these techniques diminish when the distance between the transmitter-receiver pair increases. Similar to the conventional wireless communication systems, using a molecular relay (nanorelay) is an elegant solution to this problem. The nanorelay senses the messenger molecules from individual nanomachines, decodes their messages, and transmits the decoded messages to the intended recipients.

In the conventional wireless communication systems, network coding [46] is frequently employed to increase the overall spectral efficiency [47]. In nontrivial applications, rather than simply amplifying the received signal, the relay is also responsible for decoding and combining the information it receives to maximize efficiency. Network coding is therefore a very good candidate for the diffusion-based molecular nanonetworks, where multiple nanomachines communicate with the help of a nanorelay. This is a very recent research area that has received considerable attention in the literature. In [48], the rate-delay tradeoff of employing a network coding scheme is investigated. Similarly, in [49], different relaying mechanisms are considered and it is concluded that the decode-and-forward approach has a great potential for diffusion-based molecular nanonetworks. In [50], a detailed analysis of a two-hop molecular communication network is provided. The authors consider one-way communications between a source and a destination through a relay and evaluate the error performance of half-duplex and full-duplex communication schemes.

In [1], the conventional network coding approach is applied to molecular nanonetworks and the error performance of a molecular nanonetwork providing two-way communications is analyzed. The closed form error probabilities are obtained and compared for different communication scenarios. The nanonetwork is assumed to consist of two nanomachines communicating via a centrally-placed nanorelay. In the first two symbol intervals, two nanomachines communicate their messages to the nanorelay one-by-one. Then, the nanorelay decodes both messages and uses the third symbol interval to XOR these messages and transmit simultaneously the resultant message to both nanomachines. Since individual nanomachines already know their own messages, they decode the intended message by a very simple processing of the received combined message. The overall communication of messages among nanomachines takes three time slots.

As alternative to conventional network coding, we propose two different network coding approaches. These approaches are more suitable for molecular nanonetworks and very useful for mitigating ISI and enabling a higher range communication. We first observe that a nanorelay only needs to be able to decode the combined message (XORed), rather than individual nanomachine messages. This allows the nanomachines to simultaneously utilize the channel. This is the common starting point of our approaches. Compared to the method in [1], our first approach, namely half-duplex network coding, needs two time slots, one for sending the nanomachines' symbols to the nanorelay and one for sending nanorelay's combined symbol to the nanomachines.

If the use of multiple types of molecules is allowed, the nanorelay can use a second type of messenger molecules to transmit the previous combined message to the nanomachines, while still receiving the current combined message from the nanomachines. Under these assumptions, the overall communication of all messages takes part in only one symbol interval. We call this approach full duplex network coding.

This gain in the number of required symbol intervals can be used to reduce the ISI significantly by allowing the proposed system to utilize the channel for three symbol intervals. We analyze the error performance of the proposed communication system using the analysis of [1] and demonstrate the improved performance. Note that the

error analysis given in [1] considers the transmission of a single bit and is therefore not suitable for evaluating the effects of ISI when a sequence of bits is transmitted. Therefore in addition to single bit transmission, we evaluate this realistic scenario using computer simulations and further demonstrate the advantages of a significantly reduced ISI component in the form of excellent error performance.

6.1.1. Network Model

Consider a basic molecular nanonetwork consisting of two spherical nanomachines (A and B) and a spherical nanorelay (R) placed in an unbounded 3-dimensional (3D) environment as shown in Figure 6.1. We assume that the distance between A and B, d_{AB}, is larger than the distances between the nanomachines and the relay, d_{AR} and d_{BR}, i.e., A and B are not in the communication range of each other but are both in the communication range of R.

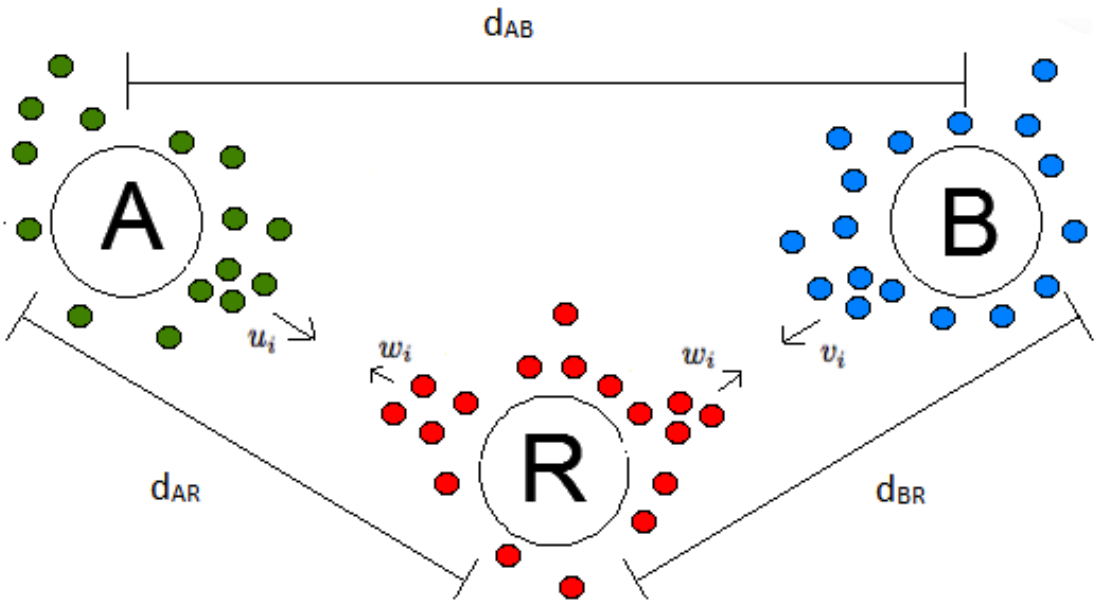


Figure 6.1. Basic communication scheme, A and B want to transmit u_i and v_i to each other by using R. R transmit $w_i = \widehat{u_i \oplus v_i}$ to A and B

Communication between the nanomachines and the relay is achieved using Concentration Shift Keying (CSK) modulation [51], where each bit is transmitted using a

different number of messenger molecules and is decoded using a pre-determined threshold. In particular, we denote the number of messenger molecules transmitted by the nanomachines for a bit-0 by Q_0 and for a bit-1 by Q_1. In practice, these values are chosen far apart from each other ($|Q_1 - Q_0| \gg 0$) in order to improve the decoding performance. It is beneficial to choose $Q_0 = 0$ since it leads to using fewer molecules for communication.

In order to evaluate the performance of the aforementioned methods, first hitting probability of the molecules, which is defined as the probability of a molecule to reach a receiver at a distance r from the transmitter at time t, can be utilized. The analytical expressions for modeling channels with 3D spherical transmitters are an open problem. Thus, we use Monte-Carlo simulations to obtain the approximation of the hitting probability $f(d, r_r, t)$. The first hitting PDF approximation for a receiver radius $r_r = 5\mu m$ is simulated for different center to center distance d and transmission slot symbol duration T_s. The simulation has been run for 2 seconds with 10^5 molecules and the number of received molecules has been counted during the transmission slots using the same methodology used as in [52].

One of the major problems in molecular communication is ISI. If the symbol duration Ts is kept long enough then there will be a smaller number of residual molecules and hence less ISI. On the other hand such an approach may lead to reduce the data rate. Therefore there is a trade-off between the data rate and the ISI.

Another problem in molecular communication is the distance between the receiver and transmitter. Eventually, as the distance increases, the number of received molecules decreases which leads to errors in communication. Using a relay between receiver and the transmitter is one solution that can be borrowed from the wireless communication literature. Another solution is keeping symbol duration long enough in order to have enough molecules arriving at the receiver but, as in the ISI case, this reduces the data rate.

For a given transmitter-receiver distance r, the probability of a molecule reaching the receiver in the first symbol interval $[0, T_s)$ is obtained as

$$p_1 = \int_0^{Ts} f(d, r_r, t)dt. \tag{6.1}$$

Assuming the release of Q messenger molecules at the transmitter at $t = 0$, the number of molecules absorbed at the receiver can be modeled using Binomial distribution with mean Qp_1 and variance $Qp_1(1-p_1)$. For a large number of messenger molecules (a large Q), this can be approximated by the Normal distribution $C^a_{signal} \sim \mathcal{N}(aQp_1, aQp_1(1-p_1))$ where $a \in \{0, 1\}$ [53]. In this representation a is an indicator to symbolize the intended symbol. Therefore, if 0 is sent, C^0_{signal}=0.

Note that, only a portion of the originally released messenger molecules manage to arrive at the receiver in the first symbol interval. The rest of the messenger molecules are called residual molecules and they cause ISI. The ISI component can be calculated as

$$p_{\text{ISI}} = \int_{Ts}^{\infty} f(d, r_r, t)dt. \tag{6.2}$$

The probability of receiving a molecule in the second symbol interval, p_2, usually dominates the ISI if Ts is chosen large enough and therefore can be used to approximate the ISI as

$$p_{\text{ISI}} \approx p_2 = \int_{Ts}^{2Ts} f(d, r_r, t)dt. \tag{6.3}$$

Assuming that Q messenger molecules are released from the transmitter at $t = 0$ and equally likely bits, the concentration of the residual molecules can be modeled as a binomial distribution with mean $0.5Qp_2$ and variance $0.5Qp_2(1-p_2)$. For a large Q, C_{ISI} can be approximated by the normal distribution $C_{\text{ISI}} \sim \mathcal{N}(0.5Qp_2, 0.5Qp_2(1-p_2))$. It is important to note that the variance of the noise in the environment is directly

Table 6.1. Schedule of the conventional network coding approach presented in [1]

Time Slot	$A \rightarrow R$	$B \rightarrow R$	$R \rightarrow A, B$
$0 - Ts$	u_i		
$Ts - 2Ts$		v_i	
$2Ts - 3Ts$			$\hat{u}_i \oplus \hat{v}_i$
$3Ts - 4Ts$	u_{i+1}		
$4Ts - 5Ts$		v_{i+1}	
$5Ts - 6Ts$			$\hat{u}_{i+1} \oplus \hat{v}_{i+1}$

proportional to the amount of molecule sent, unlike the case for conventional wireless communication systems.

Three different methods are presented for communication of A and B. The first one is proposed in [1], which is based on conventional network coding and the others are our proposed approaches to increase spectral efficiency. These three methods are summarized in Table 6.1, 6.2, and 6.3, respectively. In all of these methods, nanomachine A wants to transmit message $u_i \in \{0, 1\}$ to nanomachine B, and B wants to transmit $v_i \in \{0, 1\}$ to A. In order to achive this, the nanorelay R is used and these nanomachines send their messages to R. At R, these two messages are combined (XORed) and sent back to nanomachines. This combined message is denoted as w_i. Since nanomachine A knows its own intended bit $u_i \in \{0, 1\}$, it can decode the message from B using the relation $\hat{v}_i = u_i \oplus \hat{w}_i$. Similar approach is applied for decoding $u_i \in \{0, 1\}$ by B.

6.1.2. Conventional Network Coding

In [1], communication among nanomachines A and B using the nanorelay R is split into time slots consisting of $3T_s$ seconds each. During the i^{th} time slot, it is assumed that nanomachine A wants to transmit message $u_i \in \{0, 1\}$ to nanomachine B, and B wants to transmit $v_i \in \{0, 1\}$ to A. First, nanomachine A transmits u_i to nanorelay R in $[0 - Ts]$. Then, nanomachine B transmits v_i to the nanorelay in $[Ts - 2Ts]$. The nanorelay then decodes u_i as $\hat{u}_i$ and v_i as $\hat{v}_i$, and transmits a combined message

$w_i = \hat{u}_i \oplus \hat{v}_i$ to both nanomachines in $[2Ts - 3Ts]$. Nanomachines A and B estimate the intended message by XOR'ing the decoded message with their original messages.

During this communication, an error in the decoding of v_i by A or of u_i by B occurs in one of the following two cases: i) when the combined message is erroneous, i.e., $w_i \neq u_i \oplus v_i$, but at least one of the nanomachines decodes $\hat{w}_i$ correctly, ii) when the combined message is correct, i.e., $w_i = u_i \oplus v_i$, but at least one of the nanomachines fails to decode $\hat{w}_i$ correctly. Therefore, assuming that P_E^R and P_E^M denote the probabilities of decoding error at the relay and nanomachines, respectively, the total probability of error, P_E can be obtained as

$$\begin{aligned} P_E = P_E^R \left[2P_E^M(1 - P_E^M) + (1 - P_E^M)^2\right] + \\ \left(1 - P_E^R\right) \left[2P_E^M(1 - P_E^M) + (P_E^M)^2\right]. \end{aligned} \tag{6.4}$$

For any practical communication system, P_E^R and P_E^M are expected to be small numbers. Therefore, (6.4) can also be upper bounded by ignoring the quadratic terms as

$$P_E \leq P_E^R + 2P_E^M. \tag{6.5}$$

The upper bound given in Equation (6.5) also represents an upper bound for our proposed methods. The difference is the calculation of P_E^R and $2P_E^M$ which is calculated for each cases individually.

For conventional network coding, P_E^R occurs only when either of the symbols u_i or v_i is decoded erroneously which has a probability

$$P_E^R = 2P_E^M(1 - P_E^M). \tag{6.6}$$

Therefore, the upper bound for probability of error in conventional network coding can be rewritten as

$$P_E \leq 2P_E^M(1 - P_E^M) + 2P_E^M. \tag{6.7}$$

Since P_E^M denotes the decoding error at nanorelay, it can be obtained as considering the decoding rule at the nanorelay

$$\hat{w}_i = \begin{cases} 0 & C_{signal}^a + C_{\text{ISI}} < \tau, \\ 1 & \text{otherwise.} \end{cases} \tag{6.8}$$

Here, C_{signal}^a and C_{ISI} are normally distributed, thus their addition is approximated as; $C^a = C_{signal}^a + C_{\text{ISI}} \sim \mathcal{N}(aQp_1 + 0.5Qp_2, aQp_1(1-p_1) + 0.5Qp_2(1-p_2)) = \mathcal{N}(\mu_a, \sigma_a^2)$.

Table 6.2. Schedule of the half duplex network coding approach

Time Slot	$A \rightarrow R$	$B \rightarrow R$	$R \rightarrow A, B$
$0 - 1.5Ts$	u_i	v_i	
$1.5Ts - 3Ts$			$w_i = \widehat{u_i \oplus v_i}$
$3Ts - 4.5Ts$	u_{i+1}	v_{i+1}	
$4.5Ts - 6Ts$			$w_{i+1} = \widehat{u_{i+1} \oplus v_{i+1}}$

Considering the decoding rule, $P_E^M = \frac{1}{2}\left[Pr(C^0 > \tau) + Pr(C^1 < \tau)\right]$. Since C^a has a Gaussian distribution we can write P_E^M explicitly using Q functions as,

$$P_E^M = \mathcal{Q}\left(\frac{\tau - \mu_0}{\sigma_0}\right) + 1 - \mathcal{Q}\left(\frac{\tau - \mu_1}{\sigma_1}\right). \tag{6.9}$$

6.1.3. Half Duplex Network Coding

In the proposed half-duplex communication system, nanomachines simultaneously use the channel. The scheduling table of half duplex communication is presented in Table 6.2. For the ease of comparison with the conventional system, the total time allocated to a time slot is chosen as $3T_\mathrm{s}$. During the first symbol interval which can be considered as a period between 0 to $1.5T_\mathrm{s}$, nanomachines A and B simultaneously transmit u_i and v_i, respectively and between $1.5T_\mathrm{s}$-$3T_\mathrm{s}$ the nanorelay transmits w_i to nanomachines. Since the nanorelay is only interested in estimating $u_i \oplus v_i$ (as opposed to estimating u_i and v_i separately), the introduced interference has a minor effect on the performance[2] .

The decoding at the relay is nontrivial due to the simultaneous transmission of messenger molecules from both nanomachines. Taking into account the increased

[2]Estimating the summation under interference, as will be explained later, has a negative impact on the communication performance, but reducing the need for a separate symbol interval allows using larger values for the symbol interval and quickly overcomes the performance loss.

symbol intervals of $1.5T_{\rm s}$, the number of molecules at the receiver is obtained as $C_R^{ab} \sim \mathcal{N}((aQ_A + bQ_B)p_1^*, (aQ_A + bQ_B)p_1^*(1 - p_1^*))$, where $a, b \in \{0, 1\}$',

$$p_1^* = \int_0^{1.5Ts} f(d, r_r, t)dt \tag{6.10}$$

and

$$p_2^* = \int_{1.5Ts}^{3Ts} f(d, r_r, t)dt \tag{6.11}$$

represent the arrival probabilities of messenger and the residual molecules at the nanorelay, respectively, and Q_A and Q_B represent the number of molecules transmitted by nanomachines A and B corresponding to the values of u_i and v_i, respectively.

Due to the interference of the messenger molecules, decoding at the nanorelay can be given as

$$w_i = \begin{cases} 1 & \tau_1 < C_R^{ab} + C_{\rm ISI} < \tau_2, \\ 0 & \text{otherwise} \end{cases} \tag{6.12}$$

where $C_{\rm ISI}$ is redefined using the adjusted p_1^* and p_2^* values and $Q = Q_A + Q_B$ as $C_{\rm ISI} \sim \mathcal{N}(0.5Qp_2^*, 0.5Qp_2^*(1 - p_2^*))$.

Considering the $|Q_1 - Q_0| \gg 0$ assumption, if the number of received molecules at the nanorelay is between two predetermined thresholds τ_1 and τ_2, it is assumed that only one of the nanomachines has transmitted Q_1 molecules, and $w_i = 1$. If the received molecule concentration is very low, then both nanomachines are assumed to have transmitted bit-0s ($w_i = 0$), and when the received molecule concentration is very high, both nanomachines are assumed to have transmitted bit-1s ($w_i = 0$).

The overall error probability P_E^R can be obtained by splitting it into three cases of $[u_i = v_i = 0]$ (case 1), $[u_i \neq v_i]$ (case 2), and $[u_i = v_i = 1]$ (case 3) with probabilities 1/4, 1/2, 1/4, respectively (assuming equally likely messages). For each

Table 6.3. Schedule of the full duplex network coding approach

Time Slot	$A \rightarrow R$	$B \rightarrow R$	$R \rightarrow A, B$
$0 - Ts$	u_i	v_i	$w_{i-1} = \widehat{u_{i-1} \oplus v_{i-1}}$
$Ts - 2Ts$			
$2Ts - 3Ts$			
$3Ts - 4Ts$	u_{i+1}	v_{i+1}	$w_i = \widehat{u_i \oplus v_i}$
$4Ts - 5Ts$			
$5Ts - 6Ts$			

of these cases, the random variable $\mathcal{C}_j = C_R^{ab} + C_{\text{ISI}}$, $j = a + b$. Since $\mathcal{C}_j$ is addition of two Gaussian distributed random variables, it can be modeled as $\mathcal{C}_j \sim \mathcal{N}((aQ_A + bQ_B)p_1^* + 0.5Qp_2^*, 0.5Qp_2^*(1 - p_2^* + (aQ_A + bQ_B)p_1^*(1 - p_1^*)) = \mathcal{N}(\mu_{a+b}, \sigma_{a+b}^2)$. Considering these three cases, the probability of error can be obtained as

$$\begin{aligned} P_E^R = \frac{1}{2}\left[Pr(\mathcal{C}_1 < \tau_1) + Pr(\mathcal{C}_1 > \tau_2)\right] + \\ \frac{1}{4}\left[Pr(\tau_1 < \mathcal{C}_0 < \tau_2)\right] + \frac{1}{4}\left[Pr(\tau_1 < \mathcal{C}_2 < \tau_2)\right]. \end{aligned} \tag{6.13}$$

Since all $\mathcal{C}_j$s have Gaussian distributions we can write P_E^R using Q functions.

$$\begin{aligned} P_E^R = \frac{1}{2}\left[1 - \mathcal{Q}\left(\frac{\tau_1 - \mu_1}{\sigma_1}\right) + \mathcal{Q}\left(\frac{\tau_2 - \mu_1}{\sigma_1}\right)\right] + \\ \frac{1}{4}\left[\mathcal{Q}\left(\frac{\tau_1 - \mu_0}{\sigma_0}\right) - \mathcal{Q}\left(\frac{\tau_2 - \mu_0}{\sigma_0}\right)\right] + \\ \frac{1}{4}\left[\mathcal{Q}\left(\frac{\tau_1 - \mu_2}{\sigma_2}\right) - \mathcal{Q}\left(\frac{\tau_2 - \mu_2}{\sigma_2}\right)\right]. \end{aligned} \tag{6.14}$$

For the detection of molecules at the nanomachines, the situation is much simpler. Assuming that the molecule concentration at the nanomachines is $\mathcal{C}_a = C_{signal}^a + C_{\text{ISI}}$, we arrive at the decision rule given by,

$$\hat{w_i} = \begin{cases} 0 & C_a < \tau_3, \\ 1 & \text{otherwise}, \end{cases} \tag{6.15}$$

The decoding process at the nanomachines, defined by Equation (6.15), is much more easier than the one at the nanorelay due to the use of a single threshold. The overall error probability P_E^M can be obtained by splitting it into two equally likely cases of $[w_i = 0]$ (case 1) and $[w_i = 1]$ (case 2). The random variable $\mathcal{C}_a = C + C_{\text{ISI}}$ for $a \in \{0, 1\}$, is once again modeled with the normal distribution as $C_a = C^a_{signal} + C_{\text{ISI}} \sim \mathcal{N}(aQp_1^* + 0.5Qp_2^*, aQp_1^*(1 - p_1^*) + 0.5Qp_2^*(1 - p_2^*)) = \mathcal{N}(\mu_a, \sigma_a^2)$ and these random variables are used to calculate the overall error probability P_E^M as

$$P_E^M = \frac{1}{2}\left[Pr(\mathcal{C}_1 < \tau_3) + Pr(\mathcal{C}_0 > \tau_3)\right]. \tag{6.16}$$

Using Q functions, P_E^M can be rewritten as

$$P_E^M = \frac{1}{2}\left[1 - \mathcal{Q}\left(\frac{\tau_3 - \mu_1}{\sigma_1}\right) + \mathcal{Q}\left(\frac{\tau_3 - \mu_0}{\sigma_0}\right)\right]. \tag{6.17}$$

6.1.4. Full Duplex Network Coding

In the proposed full-duplex communication system, all nodes in the nanonetwork – nanorelay and both nanomachines – are allowed to simultaneously use the channel. For the sake of comparison, the total time allocated to a time slot is again chosen as $3T_{\text{s}}$.

As indicated in Table 6.3, during the i^{th} time slot, nanomachines A and B simultaneously transmit u_i and v_i, respectively, while the nanorelay transmits w_{i-1} using a different type of molecule. Since two different types of molecules coexisting in the channel are assumed not to interact with each other, this can be considered as an orthogonalization of the channel. The other speed increase comes from allowing A and B to transmit their messages simultaneously. Since these nanomachines do use the same molecule type, there is interference at the nanorelay. However, the nanorelay is only interested in estimating $u_i \oplus v_i$ (as opposed to estimating u_i and v_i separately), and the introduced interference has limited effect on the performance even compared to the half duplex approach. Let p_1^{**} and p_2^{**} denote the arrival probabilities of messenger and the residual molecules at the nanorelay, respectively. These probabilities can be computed as

$$p_1^{**} = \int_0^{3Ts} f(d, r_r, t)dt \tag{6.18}$$

and

$$p_2^{**} = \int_{3Ts}^{6Ts} f(d, r_r, t)dt \tag{6.19}$$

where one can easily observe $p_1^{**} > p_1^{*} > p_1$ and $p_2 > p_2^{*} > p_2^{**}$ if T_{s} is chosen large enough. Therefore if the nanomachines and nanorelays can be use different molecules for transmission and have different receptors for transmitted molecules, it is possible to increase the weight of the first channel tap, that corresponds to the signal and decrease the second channel tap that leads to ISI.

The decoding scheme and error scenarios of full duplex are similar to the half duplex so they are excluded from the text to aviod repetition. The only difference from the half duplex case is channel taps and therefore C_R^{ab} and C_{ISI} should be modified using these new channel taps.

6.1.5. Performance Evaluation

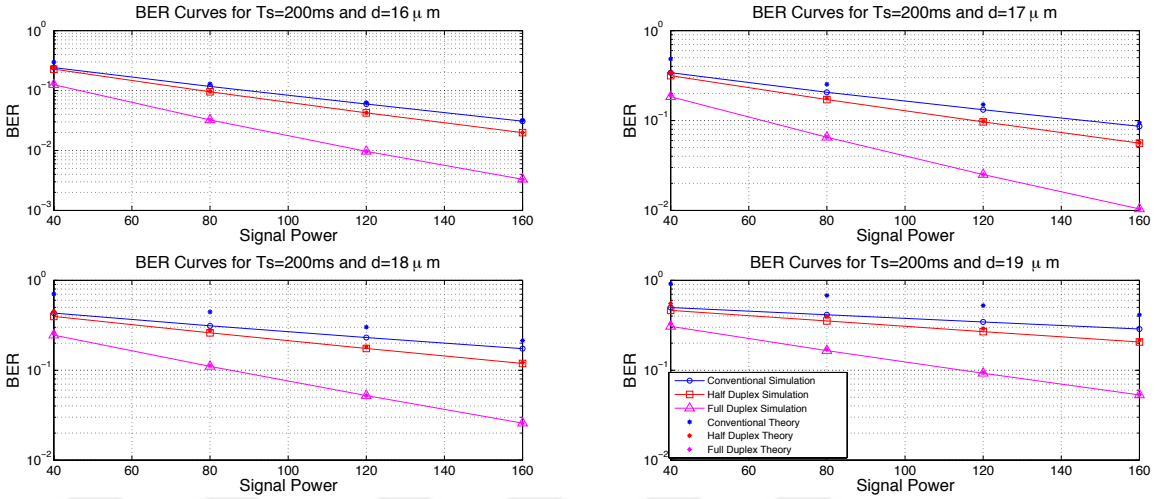


Figure 6.2. Bit error rate comparison of the proposed methods with the conventional network coding proposed in [1] for $T_s = 200ms$, for different distances for a single transmission in 3D medium.

In this section, analytical error probability expressions as well as computer simulation results are obtained for the proposed communication systems. In order to minimize the error probability due to unnecessary ISI, $Q_0 = 0$ and $Q_1 = M$ are chosen for all transmissions. The error probabilities are presented for various values of M. We note that since d_{AB}, d_{AR}, d_{BR} are not necessarily equal, each channel will have its own probability distributions. For modeling the diffusion channel, the diffusion coefficient is chosen as $D = 79.4\mu m^2/s$ which typically equivalent to diffusivity of insulin-like molecules at human blood at 310K. The simulations are done for the 3D unbounded environment.

The thresholds are chosen by considering the midpoint of the number of received molecules for the corresponding cases shifted by the average ISI terms. In particular, in order to determine τ_1, we consider the boundary between the cases when both nanomachines release $Q_0 = 0$ molecules and when one of them releases $Q_0 = 0$ molecules and the other $Q_1 = M$ molecules. τ_1 in this case would be equal to $\frac{M}{2}p_1^* + Mp_2^*$, and, similarly $\tau_2 = \frac{3M}{2}p_1^* + Mp_2^*$. The calculation of τ_3 is simpler: since there is only one transmitter, τ_3 is simply obtained as $\frac{M}{2}p_1^* + \frac{M}{2}p_2^*$.

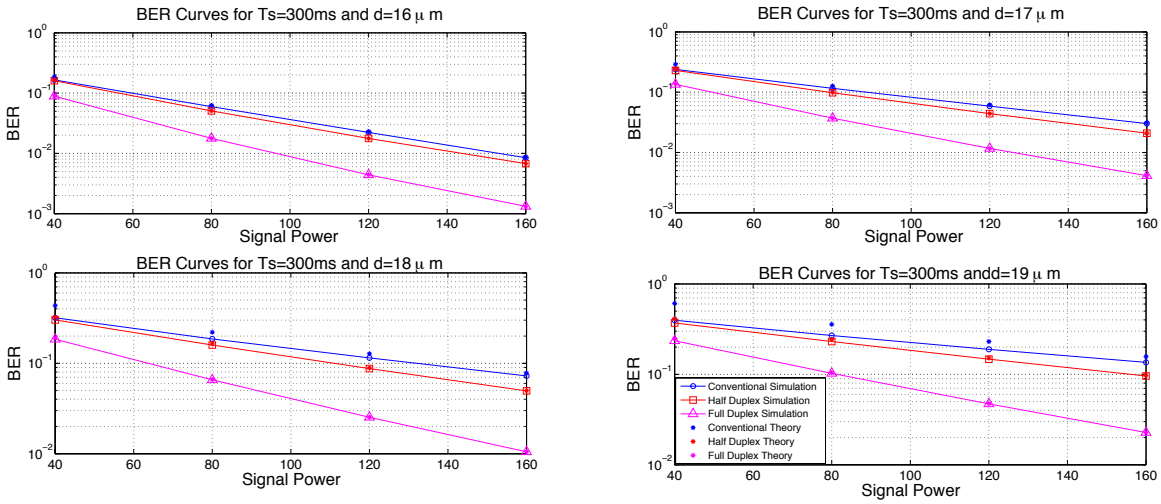


Figure 6.3. Bit error rate comparison of the proposed methods with the conventional network coding proposed in [1] for $T_s = 300ms$, for different distances for a single transmission in 3D medium.

The radius of the nanomachines and nanorelays are choosen $r_r = 5\mu m$. Four different center to center distance between nanomachines and nanorelay are chosen as r=$d_{\mathrm{AR}} = d_{\mathrm{BR}} = 16\mu m$, $17\mu m$, $18\mu m$ and $19\mu m$ to observe the effects of the distance in performance. The distance between nanomachine A and B is chosen as $2r$. The symbol interval T_{s} varies between $200ms$ to $300ms$ to evaluate the performances in different data rates. Analytical error probability expressions as well as the computer simulation results are presented in Figure 6.2 and 6.3 as a function of the signal power, which is defined as the number of molecules transmitted for a bit-1.

As shown in Figures 6.2 and 6.3 the proposed methods outperform the conventional network coding approach especially for higher range of transmission. Furthermore for lower T_s values , the ISI in the channel is more severe, and our proposed approaches still maintain a low BER due to use of longer symbol intervals (for an equal data rate comparison).

The analytical error expressions and computer simulations we have presented so far only consider transmission over a single time slot. This is reflected in the modeling of the ISI component at every stage. Up to now, we consider 2-tap channel for 1 bit transmission. When consecutive transmissions are considered, 2-tap assumption

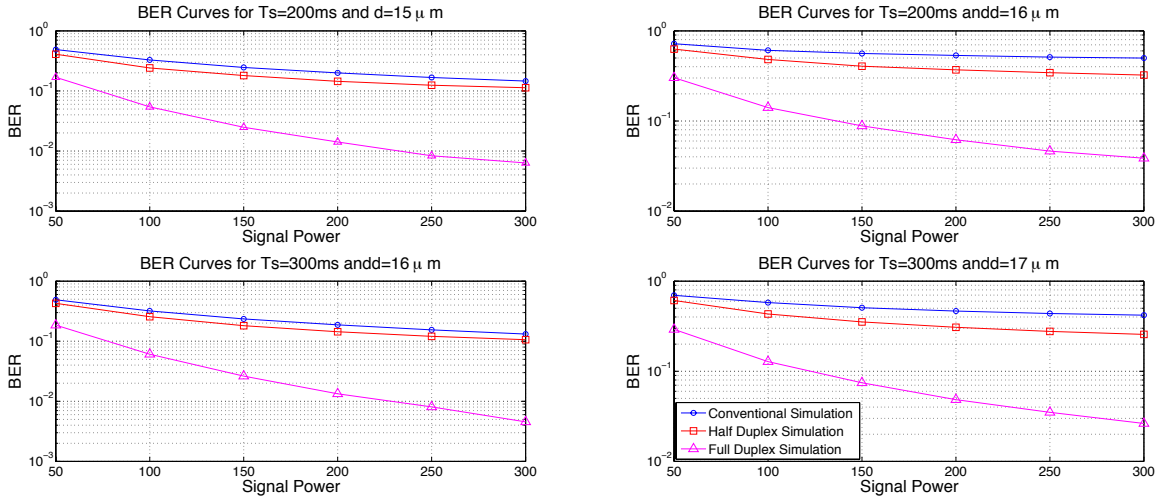


Figure 6.4. Bit error rate comparison of the proposed method with the one proposed in [1] for different T_s and d values for 1000 consecutive transmissions in 3D medium.

may not be sufficient. In order to evaluate the performance of all aforementioned methods. When longer transmissions are considered, the ISI term gets more and more difficult to derive, and we therefore rely on computer simulations to present this case. The simulation results for the all discussed network coding approaches for 1000 bits transmission in a 5-tap channel for 3D medium are presented in Figure 6.4.

According to these results, the error performance gap between the conventional and the proposed systems increases considerably when multiple transmissions are considered. This is due to the excessive ISI that exists in these systems. Since the proposed full duplex system is allowed to use $3T_{\mathrm{s}}$ seconds for transmissions (in order to keep the same data rate), the relative increase in the ISI by considering multiple transmissions is almost negligible. The conventional system on the other hand is severely affected by the increased ISI. On the other the hand half duplex method is more sensitive to ISI since it uses $1.5T_{\mathrm{s}}$ seconds for transmissions (in order to keep the same data rate) but it still outperform the conventional method.

6.1.6. Concluding Remarks on Single-hop Molecular Nanonetworks

As the distance between the transmitter and the receiver, using relay is one of the most efficient solutions in molecular communication. Network coding will lead

an opportunity to increase spectral efficiency further when a relay is used. We have proposed two novel network coding approaches as an alternative to conventional one for diffusion-based molecular nanonetworks. Unlike conventional network coding, we propose decoding the combined message rather than decoding individual messages to obtain the combined message. This allows us to increase the data rate or achieve more reliable communication for same data rate. Analytical error probability expressions are derived for the single transmission case and the error performance of the proposed system is demonstrated via these expressions as well as computer simulations. For the single transmission case, the error performance of the proposed system is shown to be superior to that of the conventional approach. It can be also concluded that our proposed methods are more applicable for higher range communications. This performance gain is shown to be even higher for the case of multiple transmission when the ISI becomes more severe. This is due to the robustness of the proposed system to ISI due to the increased symbol intervals and the resulting better channels with low ISI.

6.2. Multi-hop nanonetworks

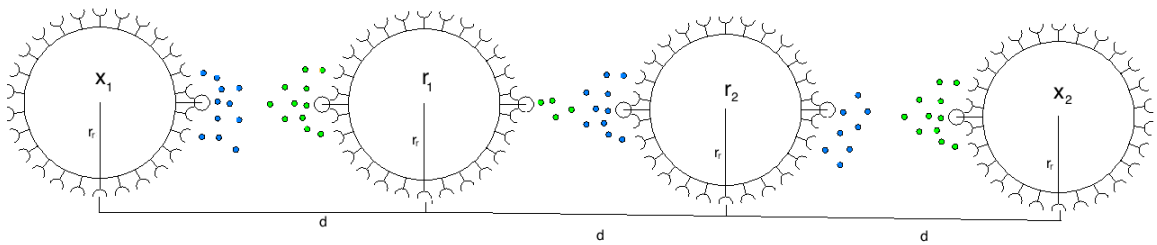


Figure 6.5. Basic communication scheme three-hop nanonetwork where x_1 and x_2 communicate with each other

Up to now, all proposed network coding approaches for molecular communication only consider one-hop networks and these works do not give any insight to easily expand the proposed methods to multi-hop networks. However, the distance between the communicating nanomachines can be so far that a single nanorelay may not be sufficient. We, therefore, propose a method to overcome this problem by considering network coding in a multi-hop network inspired by our method proposed in Section 6.1.4. Our proposed method is based on XORing the previously received messages from

neighbor nanomachines and conveying this XORed message to all of its neighbors. We, then, derive the generalized decoding rule for the intended message transmitted from source to sink. Finally, the performance of the proposed method is compared with the decode and forward method and improved performance of the proposed system is demonstrated via Monte Carlo simulations.

6.2.1. Network Model

Consider a basic molecular nanonetwork consisting of two spherical nanomachines (x_1 and x_2), and n spherical nanorelays (r_1, r_2... r_n) placed evenly in an unbounded 3-dimensional (3D) environment. For ease of presentation, the nanomachines are placed along a line, but the proposed method works for every formation. For the sake of a simpler explanation, we first examine our method for the $n = 2$ case, as shown in Figure 6.5, and then present its generic form. Assuming that only two types of molecules are allowed to be used for communication, each nanomachine and nanorelay pair is assumed to be using different types of molecules for transmission and reception to eliminate self-interference.

Communication between each node pair is achieved using a specialized version of Concentration Shift Keying (CSK) modulation [51] called On-Off Keying (OOK). In particular, bit-1 is represented by releasing M molecules and bit-0 is represented by releasing nothing.

In order to determine the performance of any proposed method, first hitting probability of a released molecule ($f(d, r_r, t)$) should be determined. In particular, $f(d, r_r, t)$ gives the probability density of a molecule to reach the receiver with radius r_r at time t and distance d. The reception model for sphere to sphere communication is already available in the literature [54]. Using $f(d, r_r, t)$ for a given transmitter-receiver distance d, the probability of a molecule reaching the receiver in the kth symbol interval $[(k-1)T_s, kT_s)$ is obtained as

$$p_k = \int_{(k-1)T_s}^{kT_s} f(d, r_r, t)dt, \tag{6.20}$$

and p_1 can be interpreted as the intended signal tap. Since only a portion of the originally released messenger molecules can arrive at the receiver in the first symbol interval, the rest of the messenger molecules are residual molecules and they cause ISI. Therefore, the fraction of molecules that cause ISI can be calculated as

$$p_{\text{ISI}} = \int_{T_s}^{\infty} f(d, r_r, t)dt, \tag{6.21}$$

where p_{ISI} is equal to the sum of the ISI taps and can also be represented as $p_{\text{ISI}}{=}\sum_{i=2}^{\infty} p_i$.

In this section, the proposed network coding method for the communication between x_1 and x_2 is presented. Furthermore, the decode and forward method, which decodes the incoming message and forwards this decoded message directly to the next node, is also presented for comparison. For both methods, two different types of molecules are used and, as stated before, each node uses different type of molecules for transmission and reception, as shown in Figure 6.5.

6.2.2. Decode and Forward Method

In the decode and forward method, each nanomachine releases its own message to its neighbor nanorelay. Once the nanorelay decodes this incoming message, it conveys the message to the next nanorelay, which repeats the same process until the intended symbol information reaches the destination (as can be seen in Table 6.4). In particular, let $x_1^k \in \{0, 1\}$ be the kth symbol of the nanomachine x_1. Then, x_1 releases molecules that represent x_1^k to r_1. Nanorelay r_1 aims to decode x_1^k using hard decision rule given as

$$r_1^k = \hat{x}_1^k = \begin{cases} 0 & C_{x_1^k}^{r_1} < \tau, \\ 1 & \text{otherwise}, \end{cases} \tag{6.22}$$

where $C_{x_1^k}^{r_1}$ corresponds to the number of observed molecules at nanorelay r_1 released from x_1 for transmitting x_1^k, and τ is a pre-determined threshold. Using the same methodology, r_1^k is decoded at nanorelay r_1 as

$$r_2^k = \hat{r}_1^k = \begin{cases} 0 & C_{r_1^k}^{r_2} < \tau, \\ 1 & \text{otherwise}. \end{cases} \tag{6.23}$$

Assuming that n nanorelays are placed between the two nanomachines, then decoding of the kth symbol is achieved by using the number of molecules released from the nth nanorelay as

$$\hat{x}_1^k = \hat{r}_n^k = \begin{cases} 0 & C_{r_n^k}^{x_2} < \tau, \\ 1 & \text{otherwise}. \end{cases} \tag{6.24}$$

As seen from Table 6.4, the required time to transmit a symbol from source to sink is always $2T_{\text{s}}$ regardless of the number of nanorelays (n). It is clear that, the same decoding rule is valid for decoding of x_2^k at nanomachine x_1.

Table 6.4. Schedule of the decode and forward approach where $r_1^k = \hat{x}_1^k$ and $r_2^k = \hat{x}_2^k$

Timeslot	$x_1 \to r_1$	$x_2 \to r_2$	$r_1 \to r_2$	$r_2 \to r_1$	$r_1 \to x_1$	$r_2 \to x_2$
$0 - T_\mathrm{s}$	x_1^1	x_2^1				
$T_\mathrm{s} - 2T_\mathrm{s}$			r_1^1	r_2^1		
$2T_\mathrm{s} - 3T_\mathrm{s}$	x_1^2	x_2^2			$\hat{r}_2^1$	$\hat{r}_1^1$
$3T_\mathrm{s} - 4T_\mathrm{s}$			r_1^2	r_2^2		
$4T_\mathrm{s} - 5T_\mathrm{s}$	x_1^3	x_2^3			$\hat{r}_2^2$	$\hat{r}_1^2$
$5T_\mathrm{s} - 6T_\mathrm{s}$			r_1^3	r_2^3		
$6T_\mathrm{s} - 7T_\mathrm{s}$	x_1^4	x_2^4			$\hat{r}_2^3$	$\hat{r}_1^3$

Since there are two identical communication schemes in two-way communication, we can examine one-way communication from x_1 to x_2 to determine overall error probability, which is twice of the error probability of one-way communications. Note that, in one-way communications, all channels are identical except for the channel from r_n to x_2 that does not include interference at even taps compared to other channels. (This can also be deduced from Table 6.4.) Let P_M and P_S be the probability of error for decoding any transmitted symbol from r_n to x_2 and from any two other neighbor nanomachines, respectively. Note that $P_S > P_M$ due to the lower interference at P_M. In order for x_2 to decode x_1^k erroneously, there should be odd number of incorrect decodings in the network. Therefore, the probability of incorrect decoding of x_1^k by x_2 is obtained as

$$\begin{aligned} P_E = \sum_{i=0}^{n} \binom{n}{2i} P_M (1-P_S)^{n-2i} P_S^{2i} + \\ \binom{n}{2i+1} (1-P_M) P_S^{2i+1} (1-P_S)^{n-2i-1}. \end{aligned} \tag{6.25}$$

Since the probability of occurence of one error is much higher than occurence of other odd number of errors, in (6.25), P_E can be tightly bounded by ignoring the terms

other than $i = 0$ as

$$P_E > P_M(1 - P_S)^n + \binom{n}{1}(1 - P_M)\, P_S(1 - P_S)^{n-1}. \tag{6.26}$$

Assuming that P_M and P_S are sufficiently small (which is a usual case in practical communication systems), the higher order terms can be neglected and P_E can also be bounded as

$$P_E > P_M + \binom{n}{1} P_S. \tag{6.27}$$

Therefore, as can be seen in Equation (6.27), P_E increases as the number of nanorelay increases. Note that it is assumed that the distance between any two neighbors is fixed, and as n increases, the distance from the source to the sink increases. Therefore, it is expected that for longer hop networks the error probability increases.

6.2.3. Network coding method

As can be deduced from Table 6.4, $2T_{\mathrm{s}}$ is required to transmit a symbol from the source to the destination in decode and forward method. Although this method can be useful to enable communication over a longer range, network coding potentially promises to reduce the required transmission time per symbol. The proposed network coding scheme for $n = 2$ nanorelays can be achieved as follows: For each symbol duration, r_1 decodes the XOR of the incoming symbols coming from r_2 and x_1 as $r_1^k = \widehat{r_2^{k-1} \oplus x_1^{k-1}}$. As suggested in [55], instead of individual decoding of the symbols to find XOR of x_1^{k-1} and r_2^{k-1}, it is suggested to decode r_1^k directly as

$$r_1^k = \begin{cases} 1 & \tau_1 < C_k^{r_1} < \tau_2, \\ 0 & \text{otherwise} \end{cases} \tag{6.28}$$

where $C_k^{r_1}$ is the number of observed molecules at the kth slot, and τ_1 and τ_2 are pre-determined thresholds.

Table 6.5. Schedule of the network coding approach where $r_1^k = r_2^{k-1} \oplus x_1^{k-1}$ and $r_2^k = r_1^{k-1} \oplus x_2^{k-1}$

$Timeslot$	$x_1 \to r_1$	$x_2 \to r_2$	$r_1 \to r_2, x_1$	$r_2 \to r_1, x_2$
$0 - T_\text{s}$	x_1^1	x_2^1		
$T_\text{s} - 2T_\text{s}$	x_1^2	x_2^2	r_1^1	r_2^1
$2T_\text{s} - 3T_\text{s}$	x_1^3	x_2^3	r_1^2	r_2^2
$3T_\text{s} - 4T_\text{s}$	x_1^4	x_2^4	r_1^3	r_2^3
$4T_\text{s} - 5T_\text{s}$	x_1^5	x_2^5	r_1^4	r_2^4
$5T_\text{s} - 6T_\text{s}$	x_1^6	x_2^6	r_1^5	r_2^5
$6T_\text{s} - 7T_\text{s}$	x_1^7	x_2^7	r_1^6	r_2^6

In particular, if the number of the absorbed molecules is between two pre-determined thresholds, this means that one of the transmitted symbols is bit-1 and the other one is bit-0. If the number of the absorbed molecules is higher than τ_2 or lower than τ_1, both of the transmitted symbols are regarded as bit-1 or bit-0 respectively. Similarly, $r_2^k = \widehat{r_1^{k-1} \oplus x_2^{k-1}}$ and it is obtained as

$$r_2^k = \begin{cases} 1 & \tau_1 < C_k^{r_2} < \tau_2, \\ 0 & \text{otherwise.} \end{cases} \tag{6.29}$$

Obviously, x_1 decodes r_1^k using simple binary thresholding as

$$\hat{r}_1^k = \begin{cases} 0 & C_{r_1^k}^{x_1} < \tau, \\ 1 & \text{otherwise.} \end{cases} \tag{6.30}$$

Similarly x_2 decodes r_2^k as

$$\hat{r}_2^k = \begin{cases} 0 & C_{r_2^k}^{x_2} < \tau, \\ 1 & \text{otherwise.} \end{cases} \tag{6.31}$$

As can be deduced from Table 6.5, both nanorelays absorb the incoming molecules from their neigbors simultaneously. For decoding, we need to use the following equations when $n = 2$:

$$\begin{aligned} r_1^k &= r_2^{k-1} \oplus x_1^{k-1}, \\ r_2^k &= r_1^{k-1} \oplus x_2^{k-1}. \end{aligned} \tag{6.32}$$

Once nanomachine x_1 receives r_1^k, using the first equation in Equation (6.32) and its own symbol x_1^{k-1}, it can decode r_2^{k-1} and using this it can decode x_2^{k-2} as

$$\begin{aligned} r_2^{k-1} &= r_1^k \oplus x_1^{k-1}, \\ x_2^{k-2} &= r_2^{k-1} \oplus r_1^{k-2}. \end{aligned} \tag{6.33}$$

Since x_1 knows x_1^{k-1} and it has previously decoded r_1^{k-2}, it can easily find that x_2^{k-2}=$r_1^k \oplus x_1^{k-1} \oplus r_1^{k-2}$. Simultaneously, x_2 does similar operations to decode x_1^{k-2}.

Considering Table 6.4 and 6.5, one can easily conclude that, for the decode and forward method, each symbol coming from the other nanomachine is decoded in $2T_\mathrm{s}$ while in network coding method each symbol is decoded in T_s. Therefore, for a fair comparison between these two methods, the symbol duration of the method based on network coding can be doubled. This will also change the channel taps for network coding as

$$p_1^* = \int_0^{2T_\mathrm{s}} f(d, r_r, t)dt \tag{6.34}$$

and

$$p_{\mathrm{ISI}}^* = \int_{2T_\mathrm{s}}^{\infty} f(d, r_r, t)dt, \tag{6.35}$$

where one can easily observe $p_1^* > p_1$ and $p_{\mathrm{ISI}} > p_{\mathrm{ISI}}^*$, if T_s is chosen sufficiently large. Therefore, network coding methods may potentially improve the performance of the system.

6.2.4. Decoding rule for n-hop nanonetworks

Once the decoding rule of the two-hop network is defined, the general decoding rule of the proposed method for an arbitrary n-hop network can be presented. In order to achieve this, nanorelay equations should be defined as done in Equation (6.32) as

$$\begin{aligned} r_1^k &= r_2^{k-1} \oplus x_1^{k-1}, \\ r_2^k &= r_1^{k-1} \oplus r_3^{k-1}, \\ r_3^k &= r_2^{k-1} \oplus r_4^{k-1}, \\ &\vdots \\ r_n^k &= r_{n-1}^{k-1} \oplus x_2^{k-1}. \end{aligned} \tag{6.36}$$

Using the equation system in Equation (6.36), x_1^{k-n} and x_2^{k-n} can be decoded by nanomachines x_2 and x_1, respectively. In particular, since nanomachine x_1 can only receive symbols coming from r_1, the corresponding decoding formula should only involve the estimated symbols coming from r_1 and x_1's own symbols. We propose a simple procedure to obtain the corresponding decoding formula without solving these equations. We will present the decoding of x_2^{k-n} by x_1, which follows the identical procedure with decoding of x_1^{k-n} by x_2 as follows:

(i) Construct a $(n+2) \times (n+1)$ matrix in the form of A= $\begin{bmatrix} x_1^k & x_1^{k-1} & ... & x_1^{k-n} \\ r_1^k & r_1^{k-1} & ... & r_1^{k-n} \\ \vdots & \vdots & \vdots & \vdots \\ r_n^k & r_n^{k-1} & ... & r_n^{k-n} \\ x_2^k & x_2^{k-1} & ... & x_2^{k-n} \end{bmatrix}$.

(ii) Let $a_{i,j}$ be the element of the matrix at the i-th row and j-th column. Then, note that the matrix is constructed using $a_{i,j}$=$a_{i-1,j+1} \oplus a_{i+1,j+1}$.

(iii) Using this rule, expand r_1^k into its components and label these components.

(iv) Unlabel each expanded element.

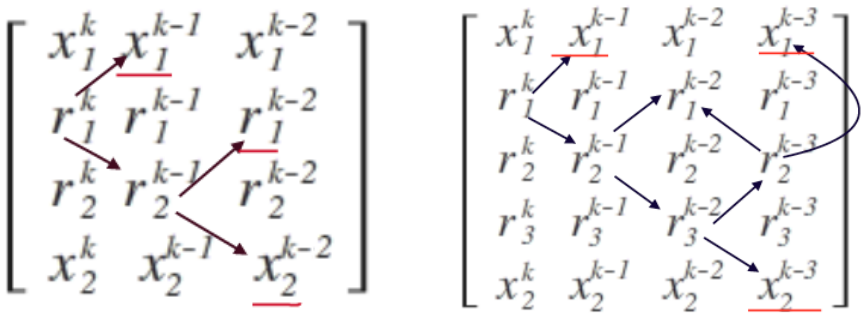
(v) If any component is labelled twice (or factor of two), then it is regarded as unlabelled.

(vi) Continue until all components of r_1^k are described (labelled components) with the first two rows (note that first row is known and second row comes from r_1) and the last row's last element x_2^{k-n} (that needs to be estimated)

Using this procedure, the derived formula of x_2^{k-n} for different n values is given in Table 6.6, and the examples are presented in Figure 6.6.

Table 6.6. Required r_1^{k-i} and x_1^{k-i} to decode x_2^{k-n} using n nanorelays

n (number of nanorelays)	r_1^{k-i}	x_1^{k-i}
1	0	1
2	0,2	1
3	0	1,3
4	0,2,4	1
5	0,4	1,3,5
6	0,2,6	1,5
7	0	1,3,7
8	0,2,4,8	1
9	0,4,8	1,3,5,9
10	0,2,6,8,10	1,5,9
11	0,2,8	1,7,9,11



(a) Decoding example for $n = 2$, $r_1^k = r_1^{k-2} \oplus x_1^{k-1} \oplus x_2^{k-2}$

(b) Decoding example for $n = 3$, $r_1^k = x_1^{k-1} \oplus x_1^{k-3} \oplus x_2^{k-3}$

Figure 6.6. Decoding example for different n values. Note that underlined terms are required terms with r_1^k for decoding x_2^{k-n}

6.2.5. Probability of error

In order to derive the general error formula of the proposed system, it is beneficial to consider the $n = 2$ case first by defining P_{R_1} and P_M as probabilities of error of the presented decoding schemes in Equation (6.29) and Equation (6.31), respectively. Since the decoding of x_2^{k-n} is achieved using r_1^ks, error occurs when r_1^k is erroneously decoded by x_1 or it was already erroneous due to previous transmissions and decoded correctly by x_1. Therefore, probability of error P_E can be written as

$$P_E = P_{R_1}(1 - P_M) + P_M(1 - P_{R_1}). \tag{6.37}$$

Note that P_{R_1} involves two terms that represent two different cases: (i) when the conveyed messages from neighbor nodes are correct, but r_1 decodes it erroneously, (ii) and when conveyed messages from neighbor nodes are incorrect and r_1 decodes the XOR of these incorrect messages correctly. Therefore, P_{R_1} can be expanded as

$$P_{R_1} = P_R(1 - P_R) + (1 - P_R)P_R, \tag{6.38}$$

where P_R is probability of error for the decoding scheme in Equation (6.29) assuming transmitted symbols to nanorelay are correct. Ignoring higher order terms in Equation (6.38), P_{R_1} can be bounded as

$$P_{R_1} > P_R + P_R. \tag{6.39}$$

Applying this procedure to the general n nanorelay case, we arrive at

$$P_{R_1} > nP_R. \tag{6.40}$$

Therefore, P_E in Equation (6.37) can be bounded as

$$P_E > nP_R + P_M. \tag{6.41}$$

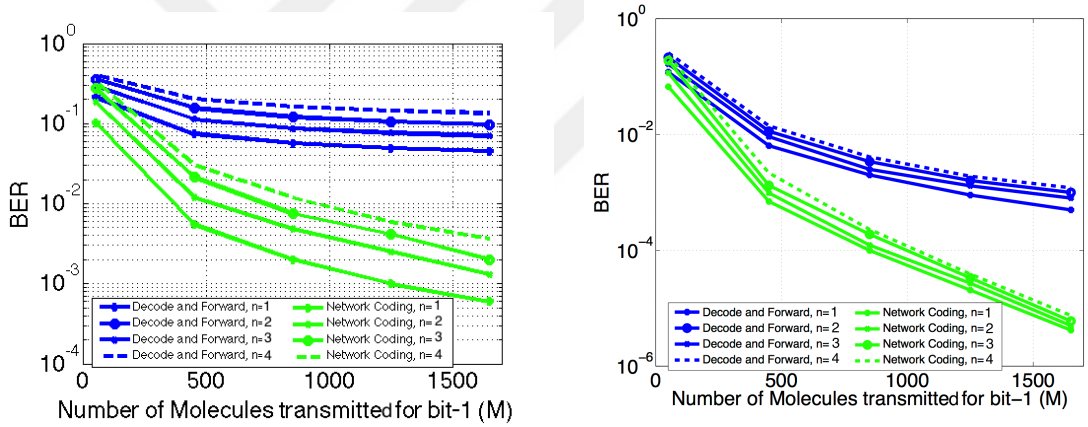
Comparing the probability of error functions of decode and forward and network coding methods presented in Equation (6.27) and Equation (6.41), it can be concluded that P_E increases as n increases. On the other hand, since there is not an explicit form of P_R, P_M, and P_S, and since for a fair comparison T_s is doubled at network coding case, we need to use Monte Carlo simulations to compare the performances of these methods.

6.2.6. Performance Evaluation

In this section, the performances of the two methods are compared. The performances are evaluated using bit error rate (BER) which is obtained using Monte Carlo simulations. BERs are obtained for various values of M, and for modeling the diffusion channel, the diffusion coefficient is chosen as $D = 79.4\mu m^2/s$. The radius of the nanomachines and nanorelays are choosen as $r_r = 5\mu m$ and the center to center distance between each neighbor pair is chosen as $d = 15\mu m$. The simulations are done for different values of time slot T_s. The thresholds are chosen by considering the midpoint of the number of received molecules for the corresponding cases shifted by the average ISI terms. For the network coding case, two thresholds are determined in the

same manner.

As can be seen in Figure 6.7, the proposed network coding approach has better BER performance than the decode and forward approach. The main reason of this observation is the increased signalling time in network coding approach. Since for a fair comparison the symbol duration in network coding approach is twice of the symbol duration of the decode and forward approach, it promises more power on signal and less power on ISI. Even though using two thresholds is more susceptible for error, its T_s advantage compensate that drawback. Furthermore, as can be seen from this figure, BER increases as the number of nanorelays (n) increases for a fixed neighbor node pair distance.



(a) Error performance for the proposed method with $T_s = 0.15$s, $r_r = 5\mu m$, $r_0 = 10\mu m$, and $D = 80\mu m^2/s$

(b) Error performance for the proposed method with $T_s = 0.2$s, $r_r = 5\mu m$, $r_0 = 10\mu m$ and $D = 80\mu m^2/s$

Figure 6.7. BER comparison of two methods for $T_s = 0.15s$ and $T_s = 0.2s$, with different number of nanorelays.

6.2.7. Concluding Remarks on Multi-hop Nanonetworks

Inspiring from the idea proposed for one-hop nanonetworks, a network coding approach for a multi-hop nanonetwork is investigated. As the distance between two nanomachines increases, the probability of transmitting molecules from source to sink drastically reduces. Therefore, nanorelays are placed between the nanomachines to enable communication over longer ranges. In the decode and forward method, each

nanomachine conveys its message to its neighbor nanorelay, then the nanorelay forwards its decoded message to the next nanorelay, and this process continues until it reaches the destination. It is shown that, each message is conveyed to the target nanomachine in two transmission slots. Our proposed network coding method is based on listening to two neighbor nodes' messages simultaneously and deciding on the XOR of their messages. Once the XOR of neighbors' messages is decoded, it is transmitted to the neighbor nodes. This process continues from the source to the sink, and at the sink, the receiver decodes the message that comes from the source using the decoding equations of the network. The whole decoding process of a symbol takes one transmission slot in network coding approach. For a fair comparison, the time slot is doubled in network coding approach which leads to an increase in the signal power and reduces the ISI power. Monte Carlo simulations demonstrate the improved performance of the network coding approach compared to the decode and forward approach.

7. CONCLUSION

This thesis is concerned with proposing effective solutions to mitigate ISI in diffusion based molecular communication channels. MCvD is a very good candidate for communication of nanodevices due to its biocompatability asset. In MCvD channels, the information is encoded by releasing molecules to the fluidic environment and the movement of molecules are governed by Brownian Motion. This leads to slow and random movement of molecules in the channel resulting in excessive ISI. In this thesis, modulation equalization and coding method for MCvD channel is proposed considering the limited computational capabilities of nanomachines.

In MCvD channels, ISI cumulatively increases as the consecutive molecule releases occurrs. Furthermore, determining an optimal threshold is problematic, since the receiver needs to know the channel parameters and these parameters should be time independent; however, sometimes it may not be possible. Considering these issues, pulse position modulation is adopted to molecular communication. PPM is very good candidate for molecular communication due to its sparsity; hence, non-consecutive bit-1 sequences. The decision of the transmitted symbol is achieved by determining the slot that involves the maximum number of molecules. Therefore, there is no need to determine an optimal threshold and no need to know the channel parameters. Furthermore, the proposed decoding scheme in PPM is robust to time-varying channels and compared with the other proposed modulation schemes in molecular communication channel, the performance is considerably increased.

Considering the dynamics of the molecular communication channel, we have proposed a specific code family that eliminates ISI. Our proposed code family avoids consecutive bit-1s and uses adaptive threshold for decoding. We have presented the general rule for the derivation of the codeword family with any code length. Accordingly, decoding and correcting stages are presented. Finally, the performance comparison is done with the other methods and observed that it is even better than PPM modulation.

Due to the nature of diffusion, it takes some time for the molecules to reach the receiver. Considering this fact, we have proposed an equalization method by shifting the detection interval by an optimized delay time. We have proposed an objective function called SID for MCvD channels as a tractable alternative for BER function. The Monte Carlo simulations demonstrates that the proposed method increases the BER performance and proposed objective function yields same argument with BER.

Considering the diffusion process, one can expect that molecules absorbed by the receiver are mostly accumulated at the part of the receiver that is close to the transmitter. We, therefore, propose to use a directed receiver to increase the communication performance. In order to determine the optimum aperture of the receiver, the joint angle-time distribution of absorbed molecules is derived and using this distribution objective function SID is used. The improved performance of the channel is presented using Monte Carlo simulations.

As the distance of the receiver and the transmitter increases, using a relay is one of the most conventional solutions. Recently, network coding gained importance in network literature for improving spectral efficiency. In molecular communication literature, network coding has been adapted to molecular communication channel. In this thesis we have proposed a special XOR operation for network coding in molecular communication channel. Since our proposed XOR operation requires less time compared to conventional methods, the required time for decoding can be increased in our case; hence, performance of the system increases.

In order to propose effective methods for a specific channel, the examination of the dynamics of the channel is essential. In molecular communication via diffusion channel, the fundamental problem is slow and random movement of molecules resulting in ISI. Therefore, sparse communication schemes are preferable for molecular communication in order to diminish ISI. Furthermore, since nanomachines may not have adequate computational capabilities like conventional communication devices, low complexity solutions are welcomed for molecular communications.

REFERENCES

1. Aijaz, A., A. Aghvami and M. Nakhai, "On Error Performance of Network Coding in Diffusion-Based Molecular Nanonetworks", *IEEE Transactions on Nanotechnology*, Vol. 13, No. 5, pp. 871–874, 2014.

2. Akyildiz, I. F., F. Brunetti and C. Blazquez, "Nanonetworks: A new communication paradigm", *Elsevier Comput. Netw.*, Vol. 52, No. 12, pp. 2260–2279, Aug. 2008.

3. Farsad, N., H. B. Yilmaz, A. Eckford, C.-B. Chae and W. Guo, "A comprehensive survey of recent advancements in molecular communication", *IEEE Commun. Surveys Tuts.*, Vol. 18, No. 3, pp. 1887–1919, 2016.

4. Guo, W., C. Mias, N. Farsad and J.-L. Wu, "Molecular versus electromagnetic wave propagation loss in macro-scale environments", *IEEE Trans. Molecular, Bio. and Multi-Scale Commun.*, Vol. 1, No. 1, pp. 18–25, 2015.

5. Genc, G., Y. E. Kara, H. B. Yilmaz and T. Tugcu, "ISI-Aware Modeling and Achievable Rate Analysis of the Diffusion Channel", *IEEE Commun. Lett.*, Vol. 20, No. 9, pp. 1729–1732, 2016.

6. Kuran, M. S., H. B. Yilmaz, T. Tugcu and I. F. Akyildiz, "Modulation Techniques for Communication via Diffusion in Nanonetworks", *Proc. IEEE Int. Conf. on Commun. (ICC)*, pp. 1–5, Jun. 2011.

7. Noel, A., K. Cheung and R. Schober, "Improving Receiver Performance of Diffusive Molecular Communication With Enzymes", *IEEE Trans. NanoBiosci.*, Vol. 13, No. 1, pp. 31–43, Mar. 2014.

8. Guo, W., T. Asyhari, N. Farsad, H. B. Yilmaz, A. Eckford and C.-B. Chae, "Molecular communications: channel model and physical layer techniques", *IEEE Wire-*

less Commun. Mag., Vol. 23, No. 4, pp. 120–127, Aug. 2016.

9. Pierobon, M. and I. F. Akyildiz, "A physical end-to-end model for molecular communication in nanonetworks", *IEEE J. Sel. Areas Commun.*, Vol. 28, No. 4, 2010.

10. Kilinc, D. and O. B. Akan, "Receiver design for molecular communication", *IEEE J. Sel. Areas Commun.*, Vol. 31, No. 12, pp. 705–714, 2013.

11. Srinivas, K. V., A. W. Eckford and R. S. Adve, "Molecular communication in fluid media: The additive inverse Gaussian noise channel", Vol. 58, No. 7, pp. 4678–4692, Jul. 2012.

12. Nakano, T., Y. Okaie and J.-Q. Liu, "Channel Model and Capacity Analysis of Molecular Communication with Brownian Motion", *IEEE Commun. Lett.*, Vol. 16, No. 6, pp. 797–800, Jun. 2012.

13. Yilmaz, H. B., A. C. Heren, T. Tugcu and C.-B. Chae, "Three-Dimensional Channel Characteristics for Molecular Communications With an Absorbing Receiver", Vol. 18, No. 6, pp. 929–932, Jun. 2014.

14. Yilmaz, H. B., G.-Y. Suk and C.-B. Chae, "Chemical propagation pattern in molecular communications", *IEEE Wireless Commun. Lett.*, Vol. 6, No. 2, pp. 226–229, April 2017.

15. Redner, S., *A Guide to First-passage Processes*, Cambridge University Press, Cambridge UK, 2001.

16. Akkaya, A., H. B. Yilmaz, C.-B. Chae and T. Tugcu, "Effect of Receptor Density and Size on Signal Reception in Molecular Communication via Diffusion with an Absorbing Receiver", *IEEE Commun. Lett.*, Vol. 19, No. 2, pp. 155–158, Feb. 2015.

17. Yilmaz, H. B., C. Lee, Y. J. Cho and C.-B. Chae, "A Machine Learning Approach to Model the Received Signal in Molecular Communications", *Proc. IEEE Int.*

Black Sea Conf. on Commun. and Netw. (BlackSeaCom), June 2017.

18. Lee, C., H. B. Yilmaz, C.-B. Chae, N. Farsad and A. Goldsmith, "Machine Learning based Channel Modeling for Molecular MIMO Communications", *Proc. IEEE Workshop on Signal Process. Adv. in Wireless Commun. (SPAWC)*, July 2017.

19. Koo, B., C. Lee, H. B. Yilmaz, N. Farsad, A. Eckford and C.-B. Chae, "Molecular MIMO: from theory to practice", *IEEE J. Sel. Areas Commun.*, Vol. 34, No. 3, pp. 600–614, March 2016.

20. Noel, A., D. Makrakis and A. Hafid, "Channel Impulse Responses in Diffusive Molecular Communication with Spherical Transmitters", *arXiv preprint arXiv:1604.04684*, 2016.

21. Kuran, M. S., H. B. Yilmaz, T. Tugcu and I. F. Akyildiz, "Modulation techniques for communication via diffusion in nanonetworks", *Communications (ICC), 2011 IEEE International Conference on*, pp. 1–5, IEEE, 2011.

22. Arjmandi, H., 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.

23. Kabir, M. H., S. R. Islam and K. S. Kwak, "D-MoSK Modulation in Molecular Communications", *IEEE Transactions on Nanobioscience*, Vol. 14, No. 6, pp. 680–683, 2015.

24. Tepekule, B., A. E. Pusane, M. Ş. Kuran and T. Tugcu, "A novel pre-equalization method for molecular communication via diffusion in nanonetworks", *IEEE Communications Letters*, Vol. 19, No. 8, pp. 1311–1314, 2015.

25. Garralda, N., 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.

26. Garralda, N., I. Llatser, A. Cabellos-Aparicio and M. Pierobon, "Simulation-based evaluation of the diffusion-based physical channel in molecular nanonetworks", *Computer Communications Workshops (INFOCOM WKSHPS), 2011 IEEE Conference on*, pp. 443–448, IEEE, 2011.

27. Yilmaz, H. B., A. C. Heren, T. Tugcu and C.-B. Chae, "Three-Dimensional Channel Characteristics for Molecular Communications With an Absorbing Receiver", *IEEE Communications Letters*, Vol. 18, No. 6, pp. 929–932.

28. Yilmaz, H. B. and C.-B. Chae, "Arrival modelling for molecular communication via diffusion", *Electronics Letters*, Vol. 50, No. 23, pp. 1667–1669, 2014.

29. Leeson, M. S. and M. D. Higgins, "Forward error correction for molecular communications", *Nano Communication Networks*, Vol. 3, No. 3, pp. 161–167, 2012.

30. Lu, Y., M. D. Higgins and M. S. Leeson, "Diffusion based molecular communications system enhancement using high order hamming codes", *Communication Systems, Networks & Digital Signal Processing (CSNDSP), 2014 9th International Symposium on*, pp. 438–442, IEEE, 2014.

31. Lu, Y., M. D. Higgins and M. S. Leeson, "Self-orthogonal convolutional codes (SOCCs) for diffusion-based molecular communication systems", *Communications (ICC), 2015 IEEE International Conference on*, pp. 1049–1053, IEEE, 2015.

32. Dissanayake, M. B., Y. Deng, A. Nallanathan, N. Ekanayake and M. Elkashlan, "Reed Solomon Codes for Molecular Communication with a Full Absorption Receiver", *IEEE Communications Letters*, 2017.

33. Koo, B.-H., 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.

34. Tepekule, B., A. E. Pusane, H. B. Yilmaz, C.-B. Chae and T. Tugcu, "ISI mitiga-

tion techniques in molecular communication", *IEEE Transactions on Molecular, Biological and Multi-Scale Communications*, Vol. 1, No. 2, pp. 202–216, 2015.

35. Kilinc, D. and O. B. Akan, "Receiver design for molecular communication", *IEEE Journal on Selected Areas in Communications*, Vol. 31, No. 12, pp. 705–714, 2013.

36. Mosayebi, R., 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.

37. ShahMohammadian, H., G. G. Messier and S. Magierowski, "Nano-machine molecular communication over a moving propagation medium", *Nano Communication Networks*, Vol. 4, No. 3, pp. 142–153, 2013.

38. Noel, A., 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.

39. Wicke, W., A. Ahmadzadeh, V. Jamali, H. Unterweger, C. Alexiou and R. Schober, "Molecular Communication using Magnetic Nanoparticles", *arXiv preprint arXiv:1704.04206*, 2017.

40. Vaidyanathan, P., S.-M. Phoong and Y.-P. Lin, *Signal processing and optimization for transceiver systems*, Cambridge University Press, 2010.

41. Baker, G. A. and P. Graves-Morris, "Pade approximants. Part 2: Extensions and applications", *Encyclopedia of Mathematics and its applications, Reading, Mass.: Addison-Wesley, 1981*, Vol. 1, 1981.

42. Kim, N.-R. and C.-B. Chae, "Novel modulation techniques using isomers as messenger molecules for nano communication networks via diffusion", *IEEE J. Sel. Areas Commun.*, Vol. 31, No. 12, pp. 847–856, Dec. 2013.

43. Akdeniz, B. C., A. E. Pusane and T. Tugcu, "Optimal Reception Delay in Diffusion-Based Molecular Communication", *IEEE Communications Letters*, Vol. 22, No. 1, pp. 57–60, 2018.

44. Yilmaz, H. B., C.-B. Chae, B. Tepekule and A. E. Pusane, "Arrival modeling and error analysis for molecular communication via diffusion with drift", *Proceedings of the Second Annual International Conference on Nanoscale Computing and Communication*, p. 26, ACM, 2015.

45. Yilmaz, H. B., N.-R. Kim and C.-B. Chae, "Effect of ISI mitigation on modulation techniques in molecular communication via diffusion", *Proceedings of ACM The First Annual International Conference on Nanoscale Computing and Communication*, p. 3, ACM, 2014.

46. Li, S. Y., R. W. Yeung and N. Cai, "Linear Network Coding", *IEEE Trans. Inf. Theor.*, Vol. 49, No. 2, pp. 371–381, 2003.

47. Peng, M., C. Yang, Z. Zhao, W. Wang and H.-H. Chen, "Cooperative network coding in relay-based IMT-advanced systems", *IEEE Communications Magazine*, Vol. 50, No. 4, pp. 76–84, 2012.

48. Unluturk, B., D. Malak and O. Akan, "Rate-Delay Tradeoff With Network Coding in Molecular Nanonetworks", *IEEE Transactions on Nanotechnology*, Vol. 12, No. 2, pp. 120–128, 2013.

49. Einolghozati, A., M. Sardari and F. Fekri, "Relaying in diffusion-based molecular communication", *IEEE International Symposium on Information Theory Proceedings (ISIT)*, pp. 1844–1848, 2013.

50. Ahmadzadeh, A., A. Noel and R. Schober, "Analysis and design of two-hop diffusion-based molecular communication networks", *IEEE Global Communications Conference (GLOBECOM)*, pp. 2820–2825, IEEE, 2014.

51. Kuran, M., H. Yilmaz, T. Tugcu and I. Akyildiz, "Modulation Techniques for Communication via Diffusion in Nanonetworks", *IEEE International Conference on Communications (ICC)*, pp. 1–5, 2011.

52. Tepekule, B., A. E. Pusane, H. B. Yilmaz and T. Tugcu, "Energy efficient ISI mitigation for communication via diffusion", *IEEE International Black Sea Conference on Communications and Networking, BlackSeaCom 2014, Odessa, Ukraine, May 27-30, 2014*, pp. 33–37, 2014, http://dx.doi.org/10.1109/BlackSeaCom.2014.6848999.

53. Yilmaz, H. B., C.-B. Chae, B. Tepekule and A. E. Pusane, "Arrival Modeling and Error Analysis for Molecular Communication via Diffusion with Drift", *Proceedings of the Second Annual International Conference on Nanoscale Computing and Communication*, NANOCOM' 15, pp. 26:1–26:6, ACM, New York, NY, USA, 2015, http://doi.acm.org/10.1145/2800795.2800816.

54. Genc, G., Y. E. Kara, T. Tugcu and A. E. Pusane, "Reception modeling of sphere-to-sphere molecular communication via diffusion", *Nano Communication Networks*, 2018.

55. Akdeniz, B. C., B. Tepekule, A. E. Pusane and T. Tuğcu, "Novel network coding approaches for diffusion-based molecular nanonetworks", *Transactions on Emerging Telecommunications Technologies*, Vol. 28, No. 7, 2017.