MULTIPLE USER AND CHANNEL INTERFERENCE REDUCTION METHODS FOR MOLECULAR COMMUNICATION VIA DIFFUSION SYSTEMS

by

Mustafa Can Gürsoy

B.S., Electrical and Electronics Engineering, Boğaziçi University, 2015

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

Graduate Program in Electrical and Electronics Engineering
Boğaziçi University
2017

ACKNOWLEDGEMENTS

I would like to dedicate this thesis to my grandfather, Ahmet Nurettin Gürsoy, with all my gratitude and respect.

I can't express how grateful I am for having my family, especially my mother and father. Them and their loving hearts are the greatest treasure I have in this world, and I feel blessed to carry their legacy of living peacefully and with dignity.

I owe special thanks and regards for Assoc. Prof. Ali Emre Pusane and Prof. Tuna Tugcu, without whom I could not have accomplished this work. Their endless support and supervision paved my academic path in my Master's studies, and their valuable mentorship helped me become the researcher I am today.

I present my sincere and humble gratitudes to Binnaz Poşul, for all the frowns she turned into smiles, and holding my hand through the challenging periods.

I thank all my friends from within and without the department. I feel very lucky to meet all the members of WCL and BUSIM, especially Alican Gök and Bọlaji Yusuf. My M.Sc. period was a pleasure with you guys. Emre Bayır and Oğuz Gelal, here's to the first Master's thesis among our group. Your brotherhood makes this journey of life enjoyable and fulfilling.

ABSTRACT

MULTIPLE USER AND CHANNEL INTERFERENCE REDUCTION METHODS FOR MOLECULAR COMMUNICATION VIA DIFFUSION SYSTEMS

Molecular Communication via Diffusion (MCvD) is a very efficient way of communicating at the nano-scale, due to its simplicity and low energy cost. However, MCvD systems are subject to severe inter-symbol interference which inhibits high data rates. Overall interference is exacerbated when multiple nano-machines are connected to one another. Stemming from the need to establish good schemes to combat both these sources of interference, the thesis revolves around the idea of incorporating a full communication frame inside a molecule's inner chemical structure. Inspired by the frame structure in traditional networks, the molecule frame consists of an overhead of frame identifier bits, and an information bearing payload. Using the frame identifier bits in a cyclic manner is found to combat the inter-symbol interference very effectively, since it exponentially increases the time interval between cycles. However, due to the constraint on the energy consumption per bit, having a longer overhead causes to transmit fewer molecules per transmission, since it decreases the length of the payload. This, in turn, makes the received signal more noisy. Since the frame length is fixed due to the messenger molecule's chemical structure, a trade-off between the lengths of the overhead and payload arises, which suggests there is an optimal allocation between them. By finding the optimal allocation point, high data rates while preserving low error rates can be achieved. Furthermore, by attaching a header in front of the frame molecules, destination addressing can be provided, and interference among different receivers can be eliminated by creating orthogonal channels between the receivers.

ÖZET

DİFÜZYON İLE İLETİŞİM SİSTEMLERİNDE ÇOKLU KULLANICI VE KANAL GİRİŞİMİNİ AZALTICI YÖNTEMLER

Difüzyon ile iletişim, uygulama kolaylığı ve düşük enerji maliyetinden dolayı nano ölçekte iletişim için etkin bir yöntemdir. Ancak, difüzyon ile iletişim sistemleri yoğun simgeler arası girişime maruz kalmakta ve bu durum yüksek veri hızlarına engel olmaktadır. Çok sayıda nano-makinenin birbiri ile haberleşmesi, sisteme çoklu-kullanıcı girişimi katmakta ve sistemdeki toplam girişimi daha da kötüleştirmektedir. Bu iki girişim kaynağı ile de mücadele etme gereksiniminden yola çıkarak bu tez, bir molekülün kimyasal yapısına bütün bir iletişim çerçevesini kodlamayı ele almaktadır. Geleneksel bilgisayar ağlarındaki çerçeve yapısından yola çıkarak, çerçevenin içine diziliş numarası içeren ek yük ile veriyi içeren yük kısımları yerleştirilmiştir. Diziliş numaralarını döngüsel olarak kullanmanın döngü zaman aralığını üssel olarak artırmasından dolayı simgeler arası girişimi önlemede çok etkin olduğu görülmüştür. Ancak, ek yük kısmının uzun olması, bit başına harcanabilecek enerjinin sabit olmasından dolayı her iletimde gönderilen çerçeve-molekül sayısını azaltmaktadır. Bu durum alıcıya ulaşan işaretin gürültüsünü artırmaktadır. Kullanılan iletişim molekülünün kimyasal yapısından dolayı çerçevenin uzunluğunun sabit olması, ek yük ve yük uzunlarının arasında bir ödünleşme olduğunu göstermekte ve en iyi bir paylaştırma noktası olduğuna işaret etmektedir. En iyi paylaştırma noktasını bularak düşük hata oranlarını kaybetmeden yüksek veri hızlarına erişilebilmektedir. Bunun ötesinde, çerçeve moleküllerin önüne üstbilgi ekleyerek hedef adresleme yapılabilmekte, bu da alıcılar arasında dikgen kanallar oluşturup alıcılar arası girişimi önleyebilmektedir.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

b_f	Number of bits in the frame identifying overhead
b_i	Number of bits in the information bearing payload
C_0	Initial concentration of the messenger molecule
$C(t)$	Concentration of the messenger molecule
d	Closest point of the spherical receiver to the point transmitter
D	Diffusion coefficient of the messenger molecule
E	The enzyme in the degradation reaction
$f_{hit}(t)$	Time distribution of a single molecule's arrival
$F_{hit}(t)$	Total probability of a single molecule's arrival until time t
$f_{hit,d}(t)$	Time distribution of a single molecule's arrival in a channel with degradation
$F_{hit,d}(t)$	Total probability of a single molecule's arrival until time t in a channel with degradation
k_1	Reaction constant of the forward reaction
k_{-1}	Reaction constant of the reverse reaction
k_p	Reaction constant for the decomposition of the enzyme-substrate pair
L	Memory of the MCvD channel
m	Transmitted molecules per bit constraint for a MaaF system
m_f	Number of transmitter frame replicas
n	Total bit slots in a MaaF molecule
N^{Tx}	Number of transmitted molecules
N_k^{Rx}	Total number of received frame molecules that hold symbol k within.
P	The product of the degradation reaction
P_c	Probability of correct frame decoding.
P_e	Exact FER for a MaaF system.
$P_{e,0}$	Probability of error associated with symbol 0.
$P_{e,app}$	Approximate FER for a MaaF system.

p_k	Channel coefficients of an MCvD system
r_0	Point-to-center distance between the point transmitter and the spherical receiver
r_r	Radius of the receiver
R_k	Number of molecule arrivals corresponding to the k^{th} tap
S	The substrate (messenger molecule) in the degradation reaction
t_b	True bit rate constraint of a MaaF system
t_{deg}	Degradation time of a molecule
t_s	Symbol duration
t_f	Frame duration
$t_{f_{\text{eff}}}$	Effective frame duration of a MaaF system with cyclic re-use
T	Absolute temperature in Kelvins
X_1	The corresponding symbol for the payload string of the intended frame
X_k	The corresponding symbol for the payload string of the k^{th} frame
Δt	The incremental time step of an MCvD simulation
λ	Rate of messenger molecule degradation
$\Lambda_{(1/2)}$	Half-life of the messenger molecule
μ_{R_k}	Mean of R_k
σ_{R_k}	Standard deviation of R_k
τ	Threshold for CSK and MCSK

LIST OF ACRONYMS/ABBREVIATIONS

3D	Three Dimensional
BER	Bit Error Rate
CCI	Co-channel Interference
CDF	Cumulative Distribution Function
CSK	Concentration Shift Keying
D-MoSK	Depleted-Molecule Shift Keying
FER	Frame Error and Erasure Rate
FIR	Finite Impulse Response
FSK	Frequency Shift Keying
ISI	Inter-symbol Interference
MaaF	Molecule-as-a-Frame
MAP	Maximum Aposteriori Probability
MCSK	Molecular Concentration Shift Keying
MCvD	Molecular Communication via Diffusion
MMSE	Minimum Mean Squared Error
MoSK	Molecule Shift Keying
PAM	Pulse Amplitude Modulation
PDF	Probability Density Function

1. INTRODUCTION

Fields of bioengineering and nanotechnology provides the ability to fabricate and manipulate devices in the nano-scale. These devices may be engineered simple organisms like bacteria, or completely human-made machines [1]. As an umbrella term, these nano-scale entities are referred to as nano-machines, and they are regarded as the smallest functional unit for nano-scale sensing, computation, and actuation.

Due to their sizes, nano-machines have limited computation capacity, and are capable of performing simple tasks. As the tasks become more elaborate, however, communication and co-operation among the nano-machines is required for organized actuation. Unfortunately, traditional electromagnetic communication paradigms fail to suffice in the nano-scale, due to their limitations on antenna size and directionality [2]. This hints for the need to come up with a new method of communication that is suitable for the specific scale and constraints [3].

With its bio-inspired approach, molecular communications is proposed as a method of communications that aims to achieve information transfer at the nano-scale using chemical signals. It is envisioned to connect the aforementioned nano-machines by the transmission and reception of information-embedded molecular signals, and provide co-operation between them. Such molecular communication links can be achieved by engineering biological phenomena including directed active transport motions over micro-tubules [4], calcium signaling [5], bacterial movement and density sensing behavior [6–8], or utilizing the diffusive nature of the fluid environment that the communication is taking place in [9]. Among these methods, Molecular Communication via Diffusion (MCvD) is considered as a very plausible option to achieve nano-scale connectivity due to its simplicity and low energy consumption.

In molecular communication systems, the message is encoded in physical and/or chemical properties of the messenger molecules. For MCvD systems, [10] introduced the quantity (Concentration Shift Keying - CSK) and type (Molecule Shift Keying -

MoSK) modulations into the molecular communications literature, which correspond to Pulse Amplitude Modulation (PAM) and Frequency Shift Keying (FSK) in digital communication literature, respectively [11]. In CSK, the information is encoded in the number of transmitted molecules, and detection is done via the help of a threshold. In MoSK, the transmitted bits are represented by different molecule types, and the receiver performs maximum detection among the molecule types to decide on the transmitted bit. Furthermore, position-based modulation schemes that modulate the molecule release time to differentiate between symbols are addressed in [12]. Constructing more advanced methods that utilize the combinations of these physical properties are also possible. [13] suggests the use of two molecule types to generate two orthogonal channels, and divide the time slots in the receiver end for detection. To give another example, [14] extends this idea by transmitting separate data in every orthogonal channel in a parallelized manner.

In an MCvD channel, the messenger molecules exhibit Brownian motion and randomly propagate in the environment [15]. This random propagation of molecules cause some molecules arrive at the receiver after their intended symbol intervals. In such a scenario, some residual molecules may arrive at the symbol intervals after their intended one, causing inter-symbol interference (ISI). ISI is a major issue in MCvD systems as it inhibits high data rates, causes error floors, and decreases overall capacity. Motivated by this fact, various methods are proposed to combat ISI. As receiver based methods, [16] addresses the use of maximum aposteriori probability (MAP) sequence detection and introduces a minimum mean squared error (MMSE) equalizer to mitigate ISI. Furthermore, [17] considers to oversample the molecule count, and perform decoding by utilizing the extra information residing in the channel memory. To exploit the information residing in the difference between two consecutive time slots, [18] proposes an adaptive thresholding algorithm. Transmitter based methods can also be employed to mitigate ISI. Considering the error propagation problems in decision feedback equalizers, [19] proposes a power adjustment method to control the average ISI in the channel. Furthermore, [20] theorizes to perform pre-equalization using the help of another molecule type. Is it noteworthy that the modulations proposed in [12–14] can also be considered transmitter based methods.

Due to their very natures, some molecules are capable of containing multiple information bits in their chemical structures. As first mentioned by [10], small hydrofluorocarbons can carry two bits inside of them, which makes the molecule suitable for 4-MoSK. Considering the potential health hazards of hydrofluorocarbons for in-vivo applications, [21] considers the use of aldohexose sugars as messenger molecules for higher order MoSK modulations. Aldohexose sugars have 4 chiral carbon atoms subject to optical isomery, which makes them able to carry information in 4-bit blocks, and perform 16-MoSK. [22] generalizes the argument by stating that one can think of the bits as *tokens*, and suggests the possibility of using amino-acid chains to store multiple bits of information.

As DNA based encoding and storage technologies develop [23], consideration of nucleotide chains for communications also become possible. Study in [24] suggests to manipulate the use of DNA when establishing communication links among bio-engineered cells. [1] and [25] also propose to use DNA as messenger molecules. Even though DNA strands are capable of transmitting large amounts of information at once, using them in MCvD systems becomes unreliable as the string gets longer. This is due to the fact that DNA molecules' diffusion coefficients decrease as the string length increases [26]. To solve this issue, [27] proposes to fragment the strand into smaller strings, and try to reassemble them at the receiver end. Furthermore, [28] considers channel coding for DNA-based communication.

Whether the molecule is a hydrofluorocarbons [10] proposed, organic sugars as considered in [21], DNA-strands, or amino acid chains, the chemical structure of the molecule allows for a fixed number of bits to be embedded in. Acknowledging the ISI problem in MCvD systems, this thesis considers the utilization of the inner *bit slots* of the molecule to encode a full communication frame in the molecule. The frame molecule is equipped with an overhead of frame identifier bits, and an information bearing payload. Overhead bits of the frame work as sequence numbers, and identify the release instant of a frame in continuous frame transmission scenarios. Stemmed from the way it considers the messenger molecule, the overall scheme is called a Molecule-as-a-Frame (MaaF) system.

Since the number of bits in a frame is fixed and limited due to the frame molecules' chemical structures, the allocatable bits for frame identification and information payload are limited in a MaaF system. This inhibits perfect frame identification, and requires the re-use of the limited number of sequence numbers in a cyclic manner. For consecutive frame transmission scenarios, this introduces ISI among the frames that hold the same sequence number. Parallel to intuition, a longer overhead is found to combat ISI better, since the cycle size increases with the number of frame identifier bits. However, this improvement in ISI combating comes at a cost of having fewer information bits in the frame. Since molecular communication systems' energy consumption is directly proportional to the number of molecules transmitted per bit [29], a transmitted molecules per bit constraint m is imposed on molecular communication systems. This makes having fewer information bits in the frame decrease the number of transmitter frame replicas m_f. This, in turn, makes the signal more noisy at the receiver end.

In scenarios with multiple transmitters and receivers, standard modulations face heavy co-channel interference (CCI) [30]. Fortunately, the MaaF approach can also be extended to work in multiple receiver scenarios as well, with the implementation of an external header to the frame molecule, similar to the approach presented in [4]. The header works as a destination address to identify between the different transmitter-receiver links, creating orthogonal channels for each transmitter-receiver link, which eliminates CCI completely.

Overall, with the goal of introducing a frame-based channel and multiple user interference reduction approach to the molecular communication literature, this thesis

- introduces the concept of a frame-based MaaF scheme,
- emphasizes on the existence of a trade-off between the lengths of the overhead and payload for a MaaF system,
- shows the existence of an optimal allocation point between them that minimizes the frame error and erasure rate (FER),
- provides a theoretical formula and an approximate yet more tractable expression

for calculating FER,

- discusses the method's relationship to other ISI mitigating modulations such as MoSK and D-MoSK under equal system complexities,
- examines the scheme's performance under different and channel conditions,
- and proposes an extension to the scheme to mitigate CCI for multiple receiver scenarios.

The rest of this thesis manuscript is organized as follows: Chapter 2 discusses the MCvD channel model of interest and the existing related modulation schemes. The ISI reducing MaaF concept and its main constraints are introduced in Chapter 3. Chapter 4 addresses the error performances of MaaF schemes and the factors that change them. Furthermore, Chapter 4 also proposes theoretical approaches to find the error rate. MaaF's extension that supports multiple user MCvD systems is explained in Chapter 5. Chapter 6 examines the robustness of the new scheme under different adverse channel conditions, and Chapter 7 concludes the thesis.

2. CHANNEL MODEL AND MODULATION SCHEMES IN MCvD SYSTEMS

2.1. Modeling the MCvD Channel

In MCvD systems, the information to be transmitted is encoded by modulating a physical property of the messenger molecule. After the modulation, the messenger molecules are released to the environment. Recalling from Chapter 1, messenger molecules exhibit Brownian motion and randomly propagate in the communication channel after their release. This random propagation is due to Fick's Second Law of Diffusion [31], which suggests that a substance will move towards the gradient of its concentration profile in a fluid medium. When looked at individual molecules, the movement of a molecule is modeled by a 3D random walk with Gaussian distributed *steps* in each discrete incremental time. For each axis, the mean of the Gaussian distribution is zero in an environment without drift, and the variance is represented by $2D\Delta t$, where D is the diffusion coefficient of the messenger molecule in the environment and Δt is the incremental time step of the simulation model [32]. The value of D is affected by the Stokes' radius of the messenger molecule, the viscosity of the communication medium, and the absolute temperature [33]. It is also worth noting that the collisions between the messenger molecules is neglected in the channel, due to the fact that their total number is substantially low relative to the number of other molecules in the environment [34]. Under such conditions, when modeling an MCvD system, movement in each axis is calculated as

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

In the literature, MCvD receivers are considered under two main categories: passive and absorbing receivers [35]. Spherical passive receivers are envisioned as spheres that are perfectly permeable to messenger molecules that count the number of molecules residing at a given time instant. On the other hand, spherical absorbing receivers perfectly absorb any molecule that arrives to them, removing them from the communication channel [36]. Stemming from the findings of [37], this thesis deals with an MCvD scenario between a single point transmitter and a single spherical and perfectly absorbing receiver in a 3D, driftless, and unbounded environment. Figure 2.1 illustrates the considered single transmitter-single receiver scenario. The closest point of the receiver

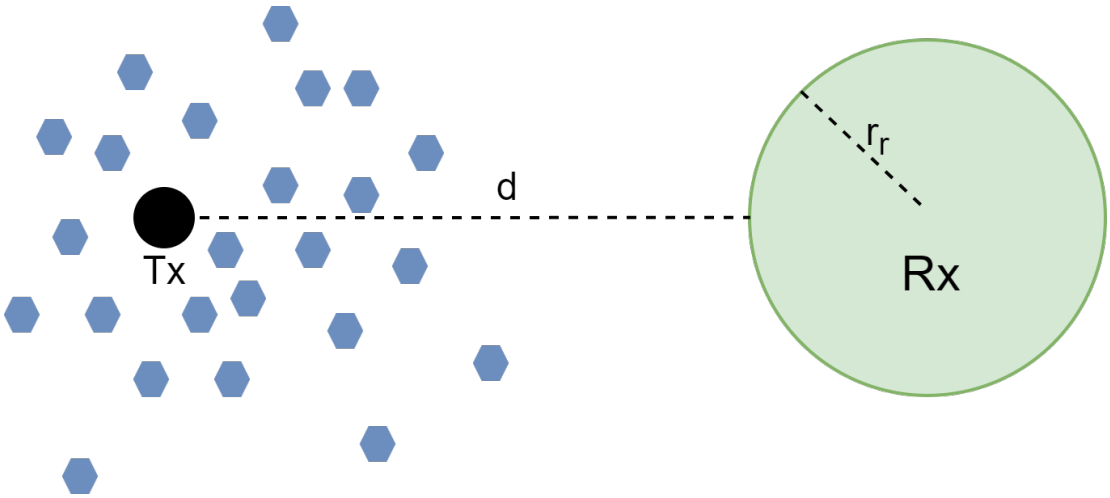


Figure 2.1. The MCvD scenario of interest. Single point transmitter and a spherical receiver with radius r_r that is d away from the transmission point.

to the transmitter point is denoted by d, and the radius of the spherical receiver is represented by r_r. Combining these parameters, the point-to-center distance between the transmitter and the spherical receiver is found by $r_0 = d + r_r$.

Since molecules move in the 3D channel according to (2.1), their arrival times at the receiver are also random. For the MCvD scenario of interest, [37] finds the time distribution of a single molecule's arrival as

$$f_{hit}(t) = \frac{r_r}{r_0\sqrt{4\pi D t}} \frac{1}{t} \frac{r_r - r_0}{t} e^{-\frac{(r_r - r_0)^2}{4Dt}}. \tag{2.2}$$

The expression that (2.2) provides can be thought of representing the probability density of arrivals with respect to time. The density is heavy-tailed with respect to time, which suggests some molecules take considerably longer times to reach the receiver. Figure 2.2 shows the general of an $f_{hit}(t)$ curve for an MCvD system.

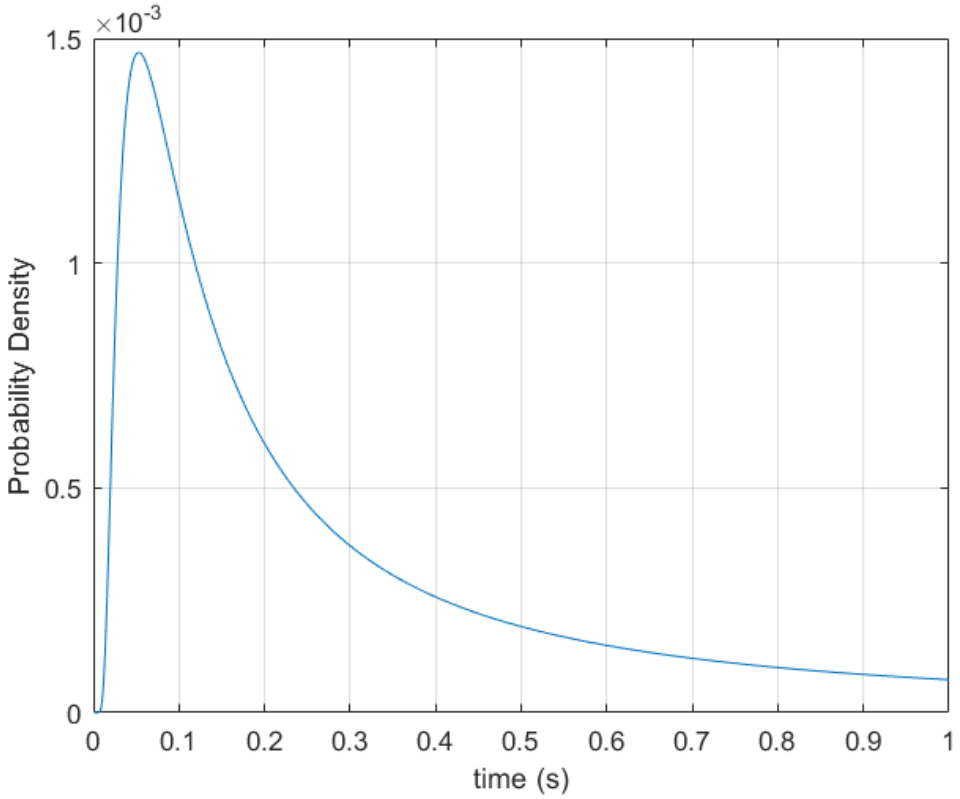


Figure 2.2. The $f_{hit}(t)$ curve of a point transmitter-spherical receiver MCvD system with $r_r = 5\mu$m, $r_0 = 10\mu$m, and $D = 79.4\frac{\mu m^2}{s}$.

Taking the time integral of (2.2), a single molecule's probability of arrival from its release until time t is found by

$$\begin{aligned} F_{hit}(t) &= \int_0^t f_{hit}(u)du \\ &= \frac{r_r}{r_0}\text{erfc}\left(\frac{r_0 - r_r}{\sqrt{4Dt}}\right). \end{aligned} \tag{2.3}$$

Eq. (2.3) can also be interpreted as the fraction of arriving molecules. It is noteworthy that since erfc$(\cdot)$ is bounded between 0 and 1 for positive arguments, the largest value $F_{hit}(t)$ can take is $\frac{r_r}{r_0}$. This suggests that for any other case than the trivial case of

$d = 0$, some molecules may never arrive at the receiver, even after an infinitely long time. For this reason, even though they represent the probability of arrivals, (2.2) and (2.3) are not technically a PDF and CDF, respectively.

Furthermore, the differences between the values of (2.3) on consecutive integer multiples of t_s gives the probability of a single molecule's arrival in consecutive symbol durations as

$$p_k = F_{hit}\big(kt_s\big) - F_{hit}\big((k-1)t_s\big) \tag{2.4}$$

where $k = 1, ..., L$. In this notation, the p_k expressions are called the FIR channel coefficients (taps) of an MCvD system, and L denotes the total number of memory the channel has, including the intended symbol. Since (2.3) has a heavy right tail, the probability of a molecule arrival never becomes zero in the scenario of interest. This makes L infinite, theoretically. However, for practical purposes, picking a large enough L provides sufficient accuracy for modeling the channel. Note that the magnitudes of p_k values directly affect the amount of ISI an MCvD system faces. Since p_2 corresponds to the first ISI tap for consecutive bit transmissions scenarios, it has the most significance among ISI taps [38].

Considering that p_k values correspond to the probabilities of a *single* molecule's arrival at the spherical receiver, the number of arriving molecules when N^{Tx} individual molecules are released can be represented by a binomial distribution as

$$R_k = \text{Binom}(N^{Tx}, p_k) \tag{2.5}$$

R_k denotes the number of molecule arrivals corresponding to the k^{th} tap, where $k = 1, 2, ..., L$.

2.2. Related Modulation Schemes

In molecular communication systems, the information is encoded in a physical property of a messenger molecule, such as its quantity or type. Similar to traditional communication systems, this encoding is referred to as a modulation. In this section, modulation schemes that are related to the discussion in this thesis is explained in further detail.

- Concentration Shift Keying (CSK): CSK is the simplest form of a quantity modulation in molecular communications. In binary CSK, a bit-1 is transmitted by releasing N^{Tx} molecules into the channel, and a bit-0 is transmitted by releasing 0 molecules. During each symboling interval, the receiver counts the number of molecule arrivals, and decides on the transmitted bit by comparing the received number with τ. There may also be multiple thresholds for higher orders.
- Molecule Shift Keying (MoSK): MoSK is the most basic type modulation in molecular communications. In binary MoSK, a bit-1 is transmitted by releasing a molecule of type-A into the channel, whilst a bit-0 is encoded to a type-B molecule. During each symboling interval, the receiver counts the number of molecule arrivals for both these types, and decides on the transmitted bit by performing maximum detection. In the presence of more molecule types, MoSK can also be utilized in higher orders to increase channel capacity.
- Molecular Concentration Shift Keying (MCSK): MCSK uses the aforementioned two molecule types in a different manner. Rather than encoding the bits into the types of the messenger molecules, MCSK uses the two molecule types to create orthogonal channels, and combats ISI by doing so. It does so by transmitting type-A molecules on odd time slots, and type-B molecules on even time slots, while relying on the aforementioned binary CSK for information encoding. The receiver only *listens* listens to type-A molecules on odd time slots, but captures type-B molecules as well. Since the captured type-B molecules correspond to the previous even time slot, capturing them helps ISI combating for the scheme. If the scheme has more molecule types, it can increase the cycle size, and combat

ISI even further.

- Depleted-Molecule Shift Keying (D-MoSK): In essence, D-MoSK is an extended and enhanced version of MCSK. It works in a similar manner to MCSK, but instead of omitting the type-B molecules on odd time slots, it counts them and adds them towards the arrival count of the previous even time slot, and postpones the decision until the even time slot comes. This increases the number of received molecules, and decreases error probability by reducing the signal noise. Similar to MCSK, the scheme can increase the cycle size and combat ISI more if it has access to more than two molecule types.

3. MAAF SYSTEM MODEL CHARACTERISTICS

3.1. The General Structure of a MaaF Molecule

With the goal of ISI reduction in mind, this thesis mainly considers encoding bits within messenger molecules' inner chemical structures. The key idea is the ability to encode multiple bits in a single molecule utilizes the channel more efficiently per a single use. This efficiency is a valuable resource, and it is found that it can be allocated to fight off ISI, or create less noisy channels which in turn increases channel capacity per use.

Multiple bit encoding in a single molecule may be accomplished by the transmitter synthesizing molecules corresponding to the bit string to be sent as discussed in [10] and [25], or utilizing other properties such as optical isomeries in a molecule family [21]. However the encoding method may be, it is possible to generalize a molecule with multiple bits as in Figure 3.1.

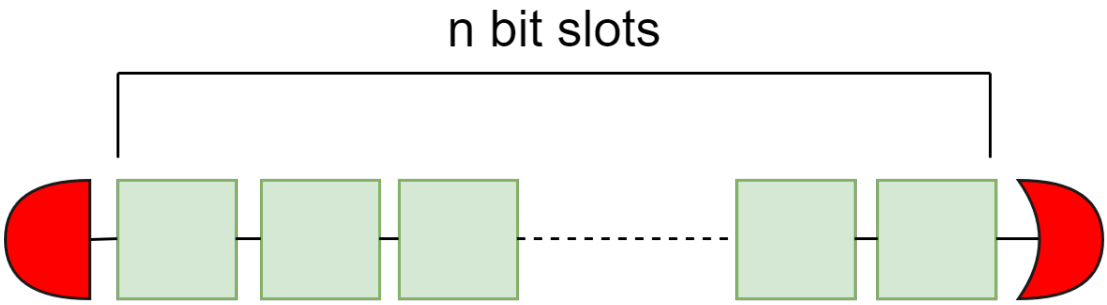


Figure 3.1. A general structure of a MaaF molecule containing n bit slots.

There are n total bit slots for the molecule in Figure 3.1, which are adjustable and encodable by the transmitter. Note that Figure 3.1 can be interpreted for any molecule family holding the possibility to encode multiple bits, including hydrofluorocarbons discussed in [10], aldohexose sugars that are proposed in [21], amino-acid chains mentioned in [22], and nucleic acids addressed in [4, 25]. It is also noteworthy that the head and the foot of the molecule is represented by different atoms/atomic bonds, which makes the receiver to distinguish between the beginning and the end of

the bit string residing in the molecule.

To give an example, considering the aldohexose molecule family in [21], the beginning of the molecule is represented by the CHO, and the end of the molecule is represented by CH_2OH. The connected H and OH in the middle four carbons in the chain is prone to optical isomery. According to their different realizations, different bit strings can be obtained. As presented in Figure 3.2 L-Glucose is a certain realization of this, and represents the bit string 1011.

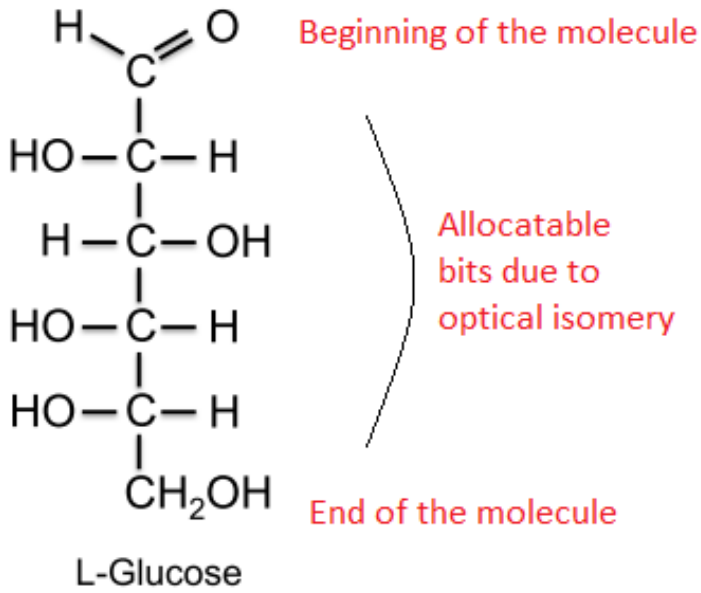


Figure 3.2. An L-Glucose molecule, corresponding to bit string 1011.

Recalling from Section 2.2, MoSK uses every available bit to increase the information payload in one transmission, given the possibility of encoding multiple bits. However, even though it is less than CSK, MoSK still faces harsh ISI when the data rate is to be increased. Stemmed by the need and possibility for ISI combating in multiple bit-carrying molecules, this thesis approaches the messenger molecule as a full communication frame in the chemical structure of a single molecule. The overall scheme is named MaaF, accordingly. MaaF introduces an overhead with frame identifier bits, with the goal of improving molecule identification at the receiver end, and combat ISI.

By introducing frame identifier bits that work as sequence numbers, this work combats ISI by exponentially increasing the effective symbol duration while preserving the bit rate of the communication. This helps the system in terms of FER, especially when the communication is done at high bit rates that greatly surpass the data rate boundaries of the existing technologies.

3.2. The Constraints and the Main Trade-Off of a MaaF System

3.2.1. The Limitation on the Overhead and Payload Lengths

A frame with n available bit slots means there are n bits to allocate between the frame identifier and the information bits . Denoting b_f as the number of bits in the frame identifying overhead, b_i as the number of bits in the information bearing payload, and assuming error-free encoding of the molecule, the total number of bits in a MaaF system can be expressed as

$$b_f + b_i = n. \tag{3.1}$$

The constraint presented in (3.1) is the main constraint governing a MaaF system. It states that there is a fixed limit on the sum of overhead and payload lengths, which is due to the chemical structure of the messenger molecule. This fixed limit means if one wishes to carry more information in a single frame, one has to sacrifice from the already limited number of frame identifier bits, thus decrease sequence numbers.

3.2.2. The Constraint on the Bit Rate of an MCvD System

To compare MCvD systems on a fair basis, two different constraints need to be imposed: data rate and energy consumption. These two constraints need to normalized on a per bit basis, since the most basic unit of transmission is a bit. For a MaaF system, the fixed frame length n suggests there are different b_f-b_i allocations. For each allocation the true bit rate should remain equal, as it directly represents the actual

data rate of the communication.

To achieve transmission at a fixed bit rate, a MaaF system has to conserve a bit duration constraint of t_b that needs to be satisfied in all allocation scenarios. This satisfies a true bit rate of $1/t_b$. Since the frames have b_i information bits, transmitting the frames at a rate of

$$t_f = b_i t_b \tag{3.2}$$

is sufficient.

Recalling from (3.1), the frame is equipped with b_f frame identifier bits alongside the information bearing payload of length b_i. Having b_f frame identifier bits provides 2^{b_f} different combinations, which can be used as frame identifying sequence numbers. Since there are a finite number of sequence numbers available to a MaaF system, the transmitter sends the frames by cyclically re-using the 2^{b_f} sequence numbers.

At the receiver end, the receiver nano-machine is synchronized to the transmitter as presented in [39]. While the transmitter sends consecutive frames by using the sequence numbers cyclically, the receiver can count until the next frame with the same sequence number is transmitted. This exponentially increases the *effective* frame duration $t_{f_{\text{eff}}}$, which can be expressed as

$$t_{f_{\text{eff}}} = 2^{b_f} t_f. \tag{3.3}$$

Recalling (3.2) and (3.1), substituting $(n - b_f)t_b$ in terms of t_f in 3.3 yields

$$t_{f_{\text{eff}}}(b_f) = 2^{b_f}(n - b_f)t_b. \tag{3.4}$$

The exponential increase in $t_{f_{\text{eff}}}$ is the main source of ISI combating that MaaF provides. By evaluating (2.3) at integer multiples of $t_{f_{\text{eff}}}$ and doing the appropriate

subtractions of (2.4), FIR channel taps for a MaaF system can be found. Since the combination of (2.3) and (2.4) suggest that ISI decreases as $t_{f_{\text{eff}}}$ increases, the b_f value that maximizes $t_{f_{\text{eff}}}$ is the allocation point of best ISI combating. Maximizing $t_{f_{\text{eff}}}$ requires finding the diminishing point of the derivative of (3.4). Taking the derivative of (**??**) yields

$$\begin{aligned} \frac{\partial t_{f_{\text{eff}}}(b_f)}{\partial b_f} &= t_b\big(2^{b_f}\ln(2)(n-b_f) - 2^{b_f}\big) \\ &= 2^{b_f} t_b\big(\ln(2)(n-b_f) - 1\big) \end{aligned} \tag{3.5}$$

which has the root $b_{f_{\text{ISI}}} = \lceil n - \frac{1}{\ln 2} \rceil$ since b_f is an integer, other than the trivial solution of $b_{f_{\text{ISI}}} = n$.

3.2.3. Constraint on the Energy Consumption Per Bit

At the beginning of each frame slot, A MaaF system synthesizes multiple frame replicas with the appropriate frame identifiers and payload, and sends multiple replicas of the same frame molecule to the channel. This is for increasing reliability, since a single molecule replica may never arrive at the receiver due to the implication of (2.3). This can be interpreted as equivalent to signal power in traditional communication systems [40].

The findings of [29] state that the total energy required to synthesize the messenger molecules is linearly proportional to the number of molecules to be synthesized. As discussed in Subsection 3.2.2, the energy consumption of a MaaF system should be normalized for every b_f-b_i allocation. Hence, this imposes another constraint on transmitting a fixed number of frame molecules per bit, which is denoted by m. The presence of this constraint suggests that the more information bits a frame holds, the more replicas of the frame molecule the transmitter is allowed to generate. Hence, the number of transmitted frame replicas m_f relates to b_i by

$$m_f = m b_i. \tag{3.6}$$

Recalling the main constraint expression (3.1), the relationship of m_f with b_f is simply

$$m_f = m(n - b_f). \tag{3.7}$$

As discussed in Section 2.1's (2.5), the arrivals of frame molecule replicas are binomially distributed for a MaaF system. For a binomial distribution with m_f trials and a success probability of p_k, the mean of the distribution is $\mu_{R_k} = m_f p_k$, and the variance is $\sigma^2_{R_k} = m_f\big(p_k(1-p_k)\big)$ [41]. Thus, the mean of the arrival distribution in the intended frame's slot is $\mu_{R_1} = m_f p_1$ for a MaaF system. Similarly, the variance of R_1 is $\sigma^2_{R_1} = m_f\big(p_1(1-p_1)\big)$. Both these quantities are linearly related to m_f, which suggests that the *relative* variance of the intended arrival distribution decreases as m_f increases. From a communications engineering standpoint, this corresponds to having a less noisy received signal at the receiver end. Furthermore, when m is small, frame erasures may be prominent for certain allocations. Having a large m_f greatly decreases the probability of an erasure.

3.2.4. The Allocation Trade-Off Between the Overhead and Payload

Subsections 3.2.2 and 3.2.3 introduce the equations governing $t_{f_{\text{eff}}}$ and m_f in terms of b_f, respectively. Eq. (3.4) suggests that $t_{f_{\text{eff}}}$ is an increasing function of b_f, whilst (3.7) shows that m_f is inversely related to b_f. This hints to a trade-off between this hints a trade-off between ISI reduction performance and noise and erasure combating performance of a MaaF system. With the true objective of minimizing FER, the presence of this trade-off implies the existence of an optimum point of a b_f-b_i allocation.

3.3. Receiver Characteristics

After the synchronization between the MaaF transmitter and the receiver is established, the consecutive frame transmission starts. In the transmission, the 2^{b_f} frame sequence numbers are cyclically re-used. A synchronized receiver can count until the next transmission of the same frame sequence number, and provide ISI reduction at

a cost of delayed detection. At the end of the counting interval marked by the transmission instant of another frame with the same sequence number, the receiver does maximum detection among the information payloads of the arrived frame molecules with appropriate sequence number.

In such a system with maximum detection, frame errors occur when frames containing of other information strings than the intended one end up being received more than the transmitted sequence. Moreover, when all R_k's from $k = 1, 2, 3, ..., L$ are zero, no information string is present for decoding. This is called a frame erasure in a MaaF system. The FER expression contains both of these sources of information loss, as it is defined as the sum of the rates of occurrence for both incorrectly decoded and erased frames. It is worth mentioning that when a frame error/erasure occurs, it is assumed that all of the payload content is erroneously decoded/deleted. Therefore, the term FER becomes identical to and interchangeable with the bit error rate (BER) as well. Also note that this assumption is the upper bound of BER for a MaaF system, and it is imposed on the performance metric for ensuring the superiority of MaaF even for the worst case.

Figure 3.3 shows a sample diagram regarding the counting intervals and decision instants of a MaaF receiver with $b_f = 2$. Every small square corresponds to t_f. Note that the receiver simultaneously counts for all sequence numbers, but performs detection at the end of each sequence number's interval.

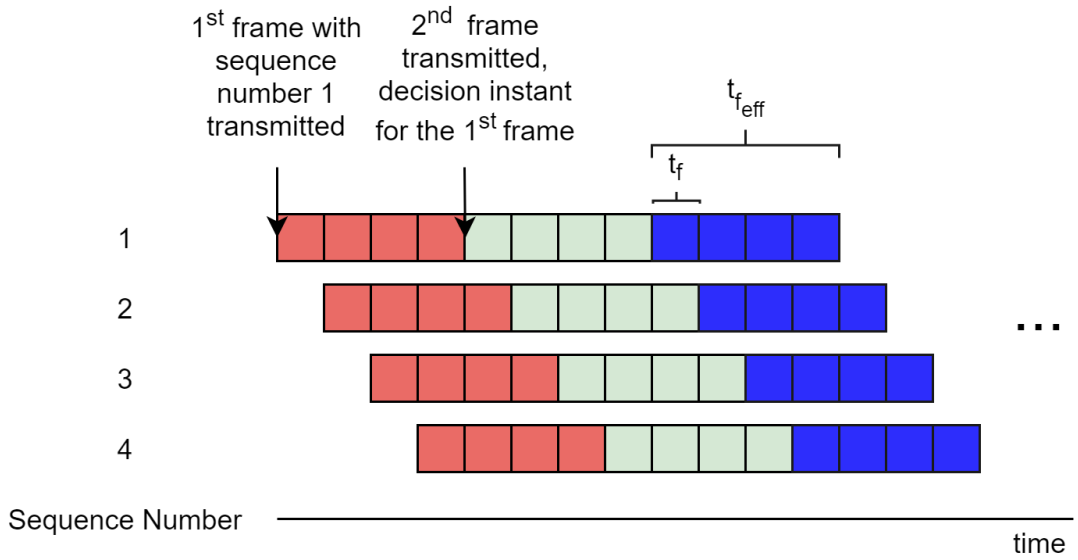


Figure 3.3. MaaF receiver's counting intervals and detection instants. $b_f = 2$.

4. ERROR ANALYSIS OF MAAF SYSTEMS

4.1. Experimental Error Analysis

In this section, FER performances of MaaF systems are analyzed by performing Monte Carlo simulations in single point transmitter/single spherical receiver scenarios in 3D unbounded MCvD environments. The simulations' goals are to verify the existence of the trade-off between the overhead and payload lengths, locate the optimum points experimentally, and observe the shift of the optimum point under different system parameters.

Table 4.1. System and Channel Parameters for Figure 4.1

Symbol	**Parameter Name**	**Value**
r_r	Receiver radius	$5\mu m$
r_0	Point to center distance between Tx & Rx	$10\mu m$
D	Diffusion coefficient of the frame molecule	$79.4\frac{\mu m^2}{s}$
L	Channel memory	30
n	Frame size	15bits
m	Molecules sent (per bit)	3
t_b	Bit duration	$10ms$

To present the existence of the trade-off for a MaaF system, a Monte Carlo simulation with parameters presented in Table 4.1 is conducted. The point-to-center distance between the point transmitter and the spherical receiver with radius 5μm is 10μm. One thing to bear in mind is that this thesis focuses on the characteristics and ISI reduction prospects of MaaF systems. Although acknowledging simple and complex sugars [21], amino acid chains [22], nucleic acids [25], etc. can hypothetically be considered good as candidates for being frame molecules, finding an actual suitable frame molecule is just an application of MaaF. This thesis focuses on the technical approach of the scheme and its mathematical background, while hinting the potential use of

the aforementioned molecules in appropriate applications. Because of this reason, the diffusivity of the insulin molecule ($D = 79.4\frac{\mu m^2}{s}$) is being used in all simulations due to the molecule's widespread use as a benchmark molecule in the literature. Additionally, the channel memory L is 30, which is sufficient for the system of interest. It is also worth noting that the values for r_r, r_0, D, and L are the same for all simulations in this thesis, except for those in Subsection 4.1.4. Figure 4.1 shows the resultant FER vs b_f curve for the MaaF scheme with parameters presented in Table 4.1.

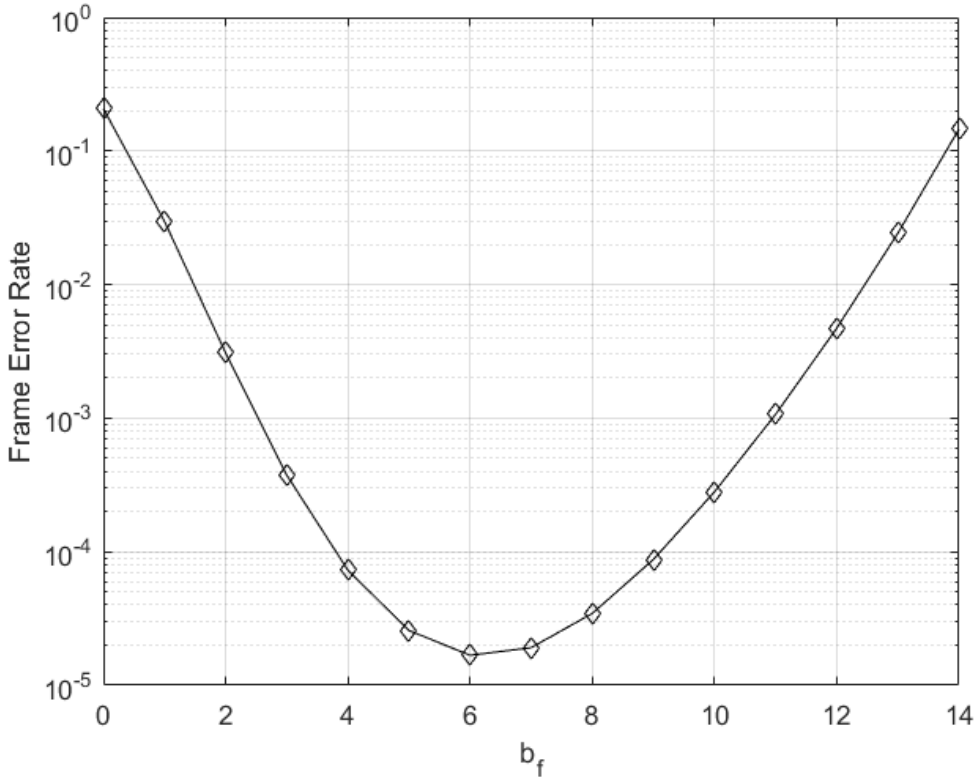


Figure 4.1. FER curve for a MaaF system with $n = 15$, $m = 3$, and $t_b = 10ms$.

The most important finding of Figure 4.1 is the shape of the FER vs. b_f curve. The U-shaped, convex nature of the curve shows the existence of the trade-off between ISI and noise combating, and verifies that this trade off is reflected in the b_f-b_i trade-off. Furthermore, for a MaaF system with given parameters, it suggests that allocating $b_f = 6$ frame identifier bits yields the lowest FER among all b_f-b_i allocations. For $n = 15$, allocating $b_f = 6$ requires $b_i = 9$, which makes $m_f = 27$ molecules. Considering the fact that the channel response of a 3D MCvD system inhibits high data rates,

yielding an error rate of 1.7×10^{-5} when communicating at a rate of 100bits/sec is very desirable for an MCvD scheme [19]. Accomplishing this while only transmitting $m_f = 27$ molecules per transmission proves the ISI reduction power of the frame overhead introduced by MaaF. MaaF purifies the signal from ISI with its cyclic use of frame identifying sequence number, and allows for faster communication with low error rates.

It is also worth mentioning that for the MaaF scheme with n available bits, the b_f axis spans from $b_f = 0$ to $b_f = n-1$, ignoring $b_f = n$. This is done to ensure having at least one information bit in the payload. Since a MaaF scheme does not have any information bits left in the frame for the special case of $b_f = n$, it has to fall back on quantity modulation. In fact, this special case is identical to performing D-MoSK modulation with a cycle size of 2^n. The other extreme where $b_f = 0$ implies $b_i = n$. This is another special case which is identical to 2^n-MoSK.

4.1.1. The Effect of the Energy Consumption Per Bit Constraint

Recalling from Subsection 3.2.3 that the energy consumption per bit constraint is represented by the number of molecules transmitted per bit, it can be inferred that m is an important parameter in a MaaF system's FER performance. 3.7 suggests that m_f increases for the same allocation point if m increases. Having a larger m_f decreases the noise associated with the intended frame arrival term R_1, reducing FER as a result. To analyze the effect of m on a MaaF system's FER performance, Figure 4.2 is presented.

Firstly, Figure 4.2 verifies the prediction which states that a MaaF system yields lower FERs with increasing m. This is due to the fact that the received signal is less noisy for larger m. Furthermore, the FER curves for different m values are observed to come closer to each other on the leftmost side. Considering the leftmost side of the curves represent the case for $b_f = 0$, the reason of this phenomenon can be address to the poor ISI combating of the $b_f = 0$.

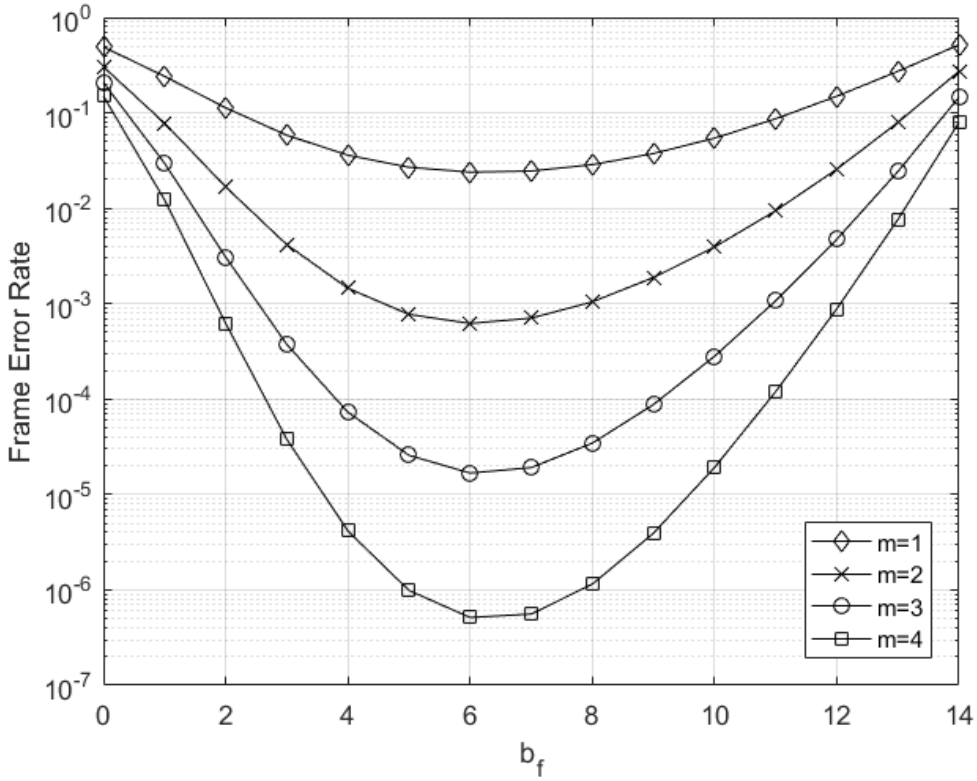


Figure 4.2. FER Curves of a MaaF system with $n = 15$, $t_b = 10$ms, for $m = 1$, 2, 3, and 4.

A more important finding resides in the optimal allocation points of the curves. Figure 4.2 finds that the optimum point does not change for different m values, which provides the possibility to perform adjustment on m without losing optimality. This is especially useful for situations where the transmitter nano-machine has energy shortage.

4.1.2. The Effect of the Bit Rate Constraint

Since the main goal of MaaF systems is to combat ISI by incorporating a frame identifying overhead, an analysis on their FER performances for varying bit rate constraints is a beneficial way to gain insight. Recall that Subsection 3.2.2 states that the bit rate constraint is represented by the t_b parameter for an MCvD scheme. To see the direct effect of t_b on the ISI, Figure 4.3 is presented.

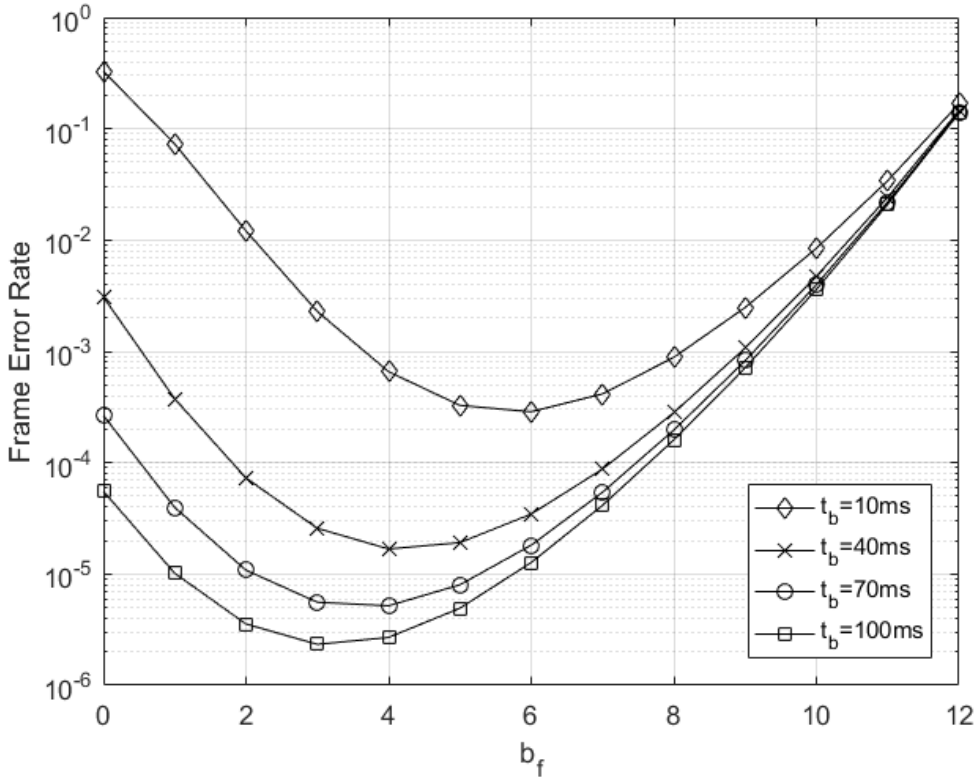


Figure 4.3. FER Curves of a MaaF system with $n = 13$, $m = 3$, for $t_b = 10, 40, 70,$ and 100ms.

First and the most straightforward finding of Figure 4.3 is that FER decreases with increasing t_b. Parallel to intuition, this is due to the inherent ISI reduction caused by an increase in t_b. Additionally, all curves converge to the same point at $b_f = n - 1$. At the rightmost point of $b_f = n - 1$, $m_f = m$ for all curves, since there is only 1 bit left in the payload. Moreover, at the point $b_f = n - 1$, $t_{f_{\text{eff}}}$ becomes so large that all channels converge to nearly ISI-free channels. Due to their equivalences in the channel behavior and m_f values, the curves converge to each other.

Figure 4.3 shows that the optimum point shifts to the left for increasing t_b. This makes sense, since increasing t_b leads to achieving the same $t_{f_{\text{eff}}}$ by incorporating a shorter overhead. Since the system can sufficiently overcome the heavy right tail of (2.3) with fewer frame identifier bits, investing the other bits for information to increase m_f becomes a much better option.

4.1.3. FER vs. m Analysis at the Optimum Allocation Point

Recall that 4.1.1 finds that changing m does not change the point of optimum allocation, but it does change the FER value on the optimum point. To gain additional insight on the amount of ISI at optimum allocation, a Monte Carlo simulation on FER vs. m is conducted, and presented in Figure 4.4.

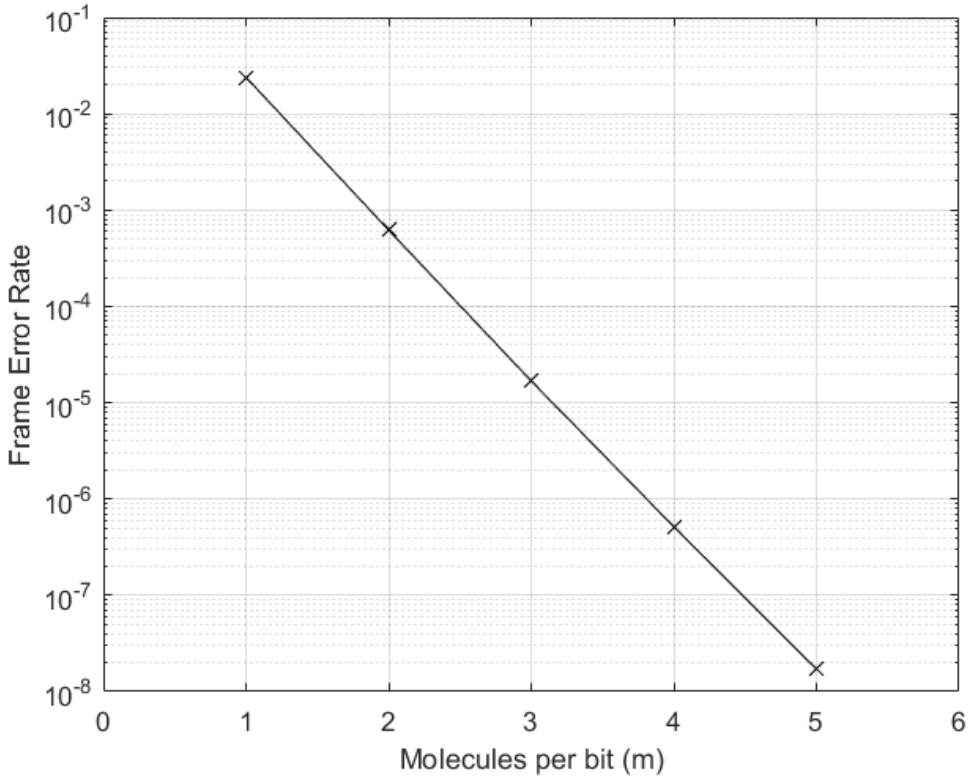


Figure 4.4. FER vs. m at the optimum allocation: $n = 15$, $t_b = 10$ms, $b_f = 6$.

In a communication system with ISI, an error floor is observed in its BER curve [40]. Hence, the insight that Figure 4.4 gives lies in its linear nature: It suggests that ISI is practically non-existent at the optimum point of a MaaF system with $n = 15$. This is, of course, achieved with the help of the frame identification bits that MaaF introduces. Exponential increase in $t_{f_{\text{eff}}}$ due to the introduction of sequence numbers nearly eliminates ISI at the optimal allocation point, removing the error floor for MaaF schemes with sufficient n.

4.1.4. The Effect of the Transmitter-Receiver Distance

Up to this point on the thesis, r_0 was kept constant at 10μm. However, changing r_0 directly changes (2.3) and the FIR channel coefficients, which affects the overall error performance of any MCvD scheme. To show the effect of r_0 on the value and location at the optimal point, Figure 4.5 is presented.

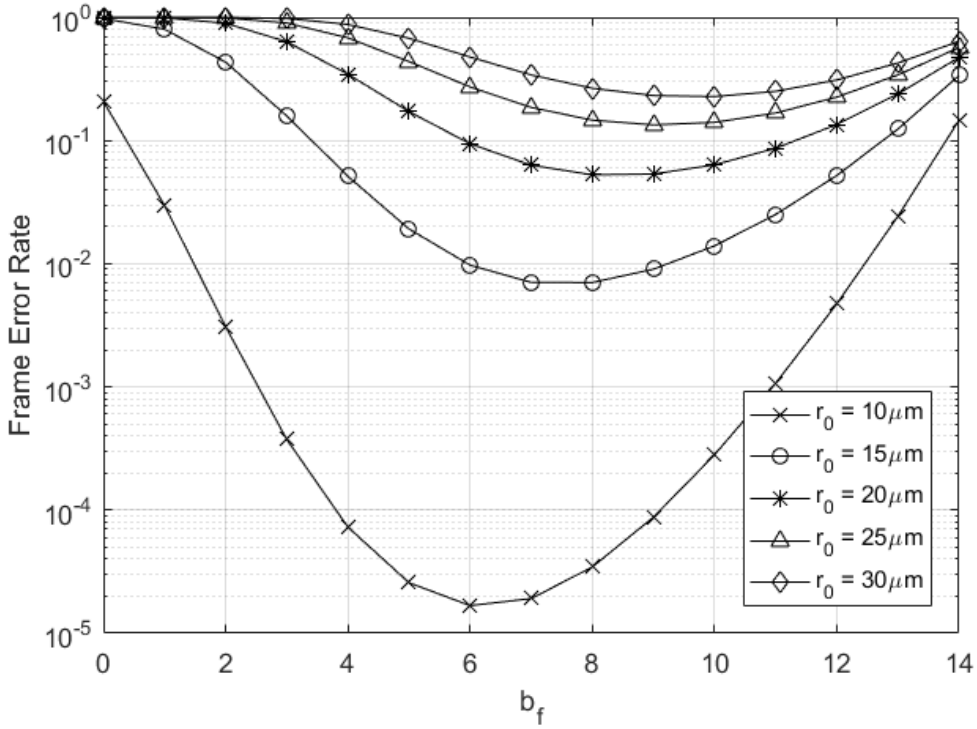


Figure 4.5. FER curves of a MaaF system with $n = 15$, $m = 3$, $t_b = 10$ms, for $r_0 = 10, 15, 20, 25$, and 30μm.

Overall, (2.3) states that fewer molecules are able to arrive at the receiver in the intended period as r_0 increases. Naturally, this leads to an increase in overall FER. Moreover, since transmitted molecules need longer times to reach to the receiver for higher r_0, further increasing $t_{f_{\text{eff}}}$ becomes a necessity to capture sufficient area under (2.2). Hence, the optimal point shifts rightward as r_0 increases.

Even when working at the optimum allocation point, the FER for $m = 3$ at $r_0 = 30\mu$m is 0.22, which is quite high. Increase in error rates as r_0 increases is valid for

every MCvD scheme due to the channel response, and MaaF is no different. However, unlike other modulations discussed in Section 2.2, MaaF can come up with a solution to communicate over longer distances without increasing t_b. Due to the findings of 4.1.1, the system can do significantly better by simply increasing the transmitted molecules per bit. if it needs to communicate over longer distances. Without losing optimality in terms of the b_f-b_i allocation, a MaaF system can reliably communicate over longer distances, at a cost of energy per bit. To visualize MaaF's efficiency in this regard, an analysis on the FER performance of a MaaF scenario with $n = 15$, $r_0 = 30\mu$m, $t_b = 10$ms with respect to m is presented in Figure 4.6.

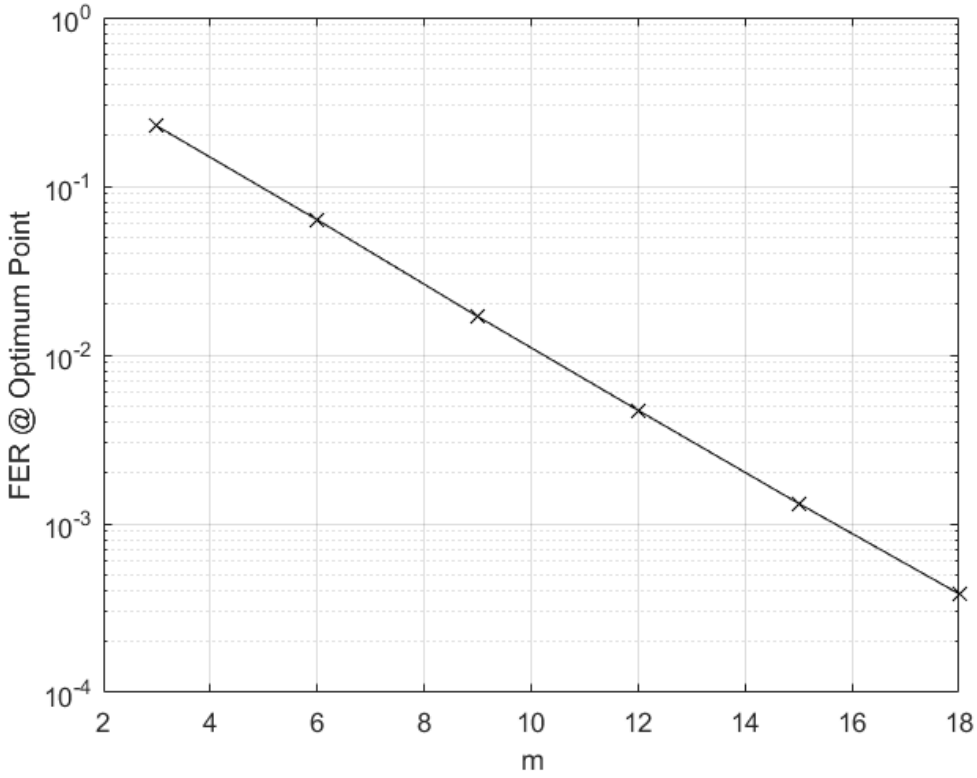


Figure 4.6. FER vs. m at the optimum allocation: $n = 15$, $r_0 = 30\mu$m, $t_b = 10$ms.

Figure 4.6 tells a similar story as Figure 4.4, even though r_0 is larger. Even for the case of $r_0 = 30\mu$m, MaaF combats ISI very effectively with the sequencing mechanism. Figure 4.6 shows that even in a channel with $r_0 = 30\mu$m, a FER of 3.8×10^{-4} can be obtained with $m = 18$, while still preserving a very high bit rate of 100bits/sec. Referring to Figure 4.5 for the location of the optimal allocation point, this corresponds

to sending 90 frame molecule replicas per transmission, since $b_i = n - b_f = 5$. Note that transmitting 90 molecules per channel use is still comparable and relatively low when compared to CSK, MoSK, MCSK, and D-MoSK [13, 14]. Even though it comes at a cost of increased energy consumption due to a larger m_f, MaaF can open the gates of establishing MCvD communication at a high bit rate over distances that are considered too long for MCvD.

4.2. Theoretical Error Analysis

4.2.1. Exact Theoretical FER Expression

Since the MaaF system is operating in an MCvD channel, the amount of ISI it faces differs according to the payload of consecutively transmitted frames. This is due to an accumulation effect which occurs when the replicas containing the same sequence number and information string get transmitted in different sequence cycles. Since both the overhead and payload are exactly the same in such cases, two molecules transmitted for different frames are identical in their chemical structures. Because of its nature, the receiver cannot differentiate between these two molecules. Thus, if they arrive in the same interval, the receiver counts both of them, rather than excluding the one transmitted for the previous frame. This effect may be constructive if the information bit string inside the intended frame is the same as the interferer's. However, it may also be destructive if multiple interferer frames have the same information bit string, and the intended frame holds a different one. Under such a system model, an exact FER expression is found by averaging the FER over all possible ISI symbol combinations [38].

Firstly, it is noteworthy that binary CSK has an alphabet of cardinality 2. For a MaaF scheme, the frame errors occur between the frames with same sequence numbers, due to the discussion made in Subsection 3.3. This makes a MaaF system have an alphabet of cardinality 2^{b_i}. Stemming from this, every payload bit combination is represented by a symbol between 0 and $2^{b_i} - 1$ in this derivation.

Suppose that X_1, the corresponding symbol for the payload string of the intended frame, is conditioned to be $X_1 = 0$ and the ISI frames are conditioned to hold a certain $X_{2:L}$ combination. Under these conditionings, the conditional FER for symbol 0 can be found as

$$P_{e,0|X_1=0,X_{2:L}} = P\Big(\max(N_1^{Rx}, ..., N_{2^{b_i}-1}^{Rx}) \geq N_0^{Rx} | X_1 = 0, X_{2:L}\Big) \tag{4.1}$$

where N_k^{Rx} is the total number of received frame molecules in the last L slots that hold symbol k within. It is noteworthy that the equality case is considered erroneous to account for erasures as well. Moving on to find the average FER over all possible symbol combinations, the conditionings need to be removed from (4.1). Thus, the expected probability of error for symbol 0 can be found as

$$P_{e,0} = \sum_{\forall X_{2:L}} \Bigg[P\Big(\max(N_1^{Rx}, ..., N_{2^{b_i}-1}^{Rx}) \geq N_0^{Rx} | X_1 = 0, X_{2:L}\Big) P(X_1 = 0) P(X_{2:L}) \Bigg] \tag{4.2}$$

where $P(X_1 = 0) = \frac{1}{2^{b_i}}$ and $P(X_{2:L}) = \frac{1}{(2^{b_i})^{L-1}}$. Note that (4.2) is the frame error probability corresponding to the payload combination with symbol 0. Since the error probability derivation is symmetric among all symbols, the overall FER for all symbols can be expressed simply as $FER = 2^{b_i} P_{e,0}$, which yields

$$P_e = \frac{1}{2^{b_i(L-1)}} \sum_{\forall X_{2:L}} \Bigg[P\Big(\max(N_1^{Rx}, ..., N_{2^{b_i}-1}^{Rx}) \geq N_0^{Rx} | X_1 = 0, X_{2:L}\Big) \Bigg]. \tag{4.3}$$

Even though it theoretically gives the exact FER for a MaaF system, there are a few shortcomings of (4.3), unfortunately. Firstly, the N_k^{Rx} expressions are represented by

$$N_k^{Rx} = \sum_{j=1}^{L} A(k, j) \text{Binom}(m_f, p_j) \tag{4.4}$$

where $A(k,j)$ is a piecewise defined function that is 1 if the k^{th} payload symbol is sent in the j^{th} memory slot, and is 0 otherwise. Since the events of $N_k^{Rx} \geq N_0^{Rx}$ for $k = 1, ..., 2^{b_i} - 1$ are not statistically independent from each other, the maximum operation cannot be turned into simpler operations [40], which makes evaluating (4.1) cumbersome.

A more important problem of evaluating (4.1) lies in its complexity. Note that the true FER is found by evaluating every single payload combination for the L consecutive frame transmissions. This requires evaluation of all (2^{b_i}) symbol combinations for every slot in an L-tap channel, resulting in a total of $(2^{b_i})^{L-1}$ evaluations. Since L is practically infinite, choosing a large L is a necessity for accurate representation of an MCvD channel. This is especially valid for MaaF, since $t_{f_{\text{eff}}}$ can be really small when evaluating cases with small b_f. Recalling from Table 4.1, $L = 30$ is chosen for this thesis for modeling the MCvD channel with sufficient reliability.

When theoretically calculating the error performance of a MaaF system shown in Figure 4.1, (4.3) needs to evaluate $2^{(15-6)^{30-1}} = 2^{261}$ different scenarios to find the FER at the optimum point. Thus, the exponential growth in complexity with respect to b_i and L can hinder the mere possibility of finding the result for most MaaF systems. Motivated by the shortcomings of the exact FER expression, the need for an approximate theoretical FER expression that is more mathematically tractable arises.

4.2.2. Approximate Theoretical FER Expression

Since the encoded payload bits are assumed to be statistically independent to one another, different frames with the same sequence number may contain same information bit payload with probability $\frac{1}{2^{b_i}}$. This causes the accumulation effect discussed in 4.2.1, and is more prominent when both b_i and $t_{f_{\text{eff}}}$ are small. When b_i is small, the probability of having identical information bit strings is higher. Furthermore, having a small $t_{f_{\text{eff}}}$ means the ISI is present and effective, which causes a non-negligible number of interferer frame molecules to arrive at the receiver.

Both of b_i and $t_{f_{\text{eff}}}$ are small only when the frame length constraint n is small. For sufficiently large n values, b_i and $t_{f_{\text{eff}}}$ cannot be both small at the same time, which inhibits the occurrence of the accumulation effect discussed in Subsection 4.2.1. For a large b_i-small b_f pair, the probability of re-occurrence of the same information string that correspond to different sequence cycles is substantially low.

On the other hand, for a small b_i-large b_f pair, the FIR channel coefficients for interferer frames are practically zero. Note that under such circumstances, even though there are a lot of string re-occurrences, they not affect FER since their probability of arriving in a different cycle than their own is approximately zero. Thus, for sufficiently large n, it is found that the accumulation of identical molecules can be disregarded. Rather, the approximate assumption of considering every string being different in content can be made safely. When the information content is different for each frame, comparison between N_k^{Rx}s become identical to the comparison between R_1 and $\max(R_2, R_3, ..., R_L)$. For the approximation, errors occur when R_1 fails to be strictly greater than $\max(R_2, R_3, ..., R_L)$. Hence, the approximate error probability FER_{app} can be expressed as

$$\text{FER}_{app} = P\big(\max(R_2, R_3, ..., R_L) \geq R_1\big). \tag{4.5}$$

Considering (4.5), the approximate probability of error and erasure can be calculated by $FER_{app} = 1 - P_c$, where P_c is the probability of correct frame decoding. P_c can be expressed as

$$P_c = P\big(\max(R_2, R_3, ..., R_L) < R_1\big) \tag{4.6}$$

where L denotes the total number of FIR channel coefficients. Note that R_ks are non-negative integers, so (4.6) also accounts for the erasures.

Since all $R_1, ..., R_L$ are assumed to come from different payload strings, they correspond to different transmissions and form orthogonal channels. This makes them

statistically independent. For a fixed value of r, the probability of the maximum of statistically independent random variables $R_2, ..., R_L$ being smaller than r can be obtained as

$$P\big(\max(R_2, R_3, ..., R_L) < r\big) = \prod_{i=2}^{L} P(R_i < r). \tag{4.7}$$

The probability of correct frame decoding can be calculated by averaging the expression on the right hand side of (4.7), over the distribution of r. Thus, P_c becomes

$$P_c = \int_{-\infty}^{\infty} \left[\prod_{i=2}^{L} P(R_i < r)\right] f_{R_1}(r)\ dr\ . \tag{4.8}$$

Since r is from the binomial distribution associated with R_1, P_c is written as

$$P_c = \sum_{r=0}^{m_f} \left[\prod_{i=2}^{L} P(R_i < r)\right] \binom{m_f}{r} p_1^r (1-p_1)^{m_f - r} \tag{4.9}$$

where $\binom{m_f}{r}$ is the number of r element combinations in a set of m_f elements. Considering that $R_2, ..., R_L$ are also binomial random variables with discrete support, $P(R_i < r)$ expressions are equal to their corresponding binomial CDFs' value at $r-1$. After representing each probability by the corresponding binomial CDF, the approximate probability of error and erasure $P_{e,app}$ becomes

$$P_{e,app} = 1 - \sum_{r=0}^{m_f} \left(\left[\prod_{i=2}^{L} 1 - I_{p_i}(r, m_f - r + 1)\right] \binom{m_f}{r} p_1^r (1-p_1)^{m_f - r} \right) \tag{4.10}$$

where $I_p(\cdot)$ is the regularized incomplete beta function.

Note that m_f is a function of b_f. Additionally, $p_1, ..., p_L$ are functions of $t_{f_{\text{eff}}}$, which is a function of b_f as well. To find the optimum allocation, (4.10) needs to be differentiated with respect to b_f. Unfortunately, this operation is not easily tractable.

However, after having an idea on the channel state information and MaaF system parameters, (4.10) can be numerically evaluated to find the minimum point. After finding the minimum point, which corresponds to the optimal allocation, a protocol regarding the b_f-b_i allocation can be established between the transmitter and the receiver.

4.2.3. Verification of the Approximate FER Formula

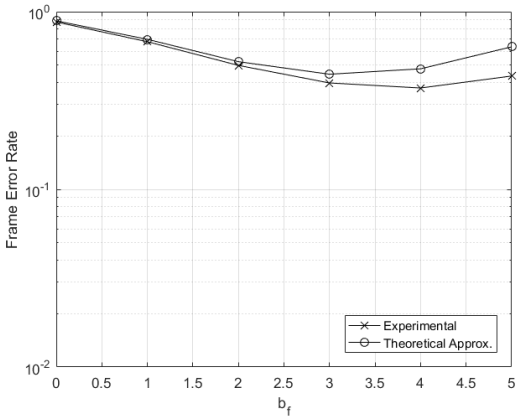
To graphically visualize the accumulation effect discussed in 4.2.1 and validate the accuracy of the theoretical FER approximation found in (4.10), FERs of the simulations and theoretical approximations for different n values are shown and compared in Figure 4.7.

As can be seen from Figure 4.7, the experimental and theoretical curves match much better as n increases. As explained in Subsection 4.2.1, this is due to larger n eliminating the accumulation effect. The parameters b_i and b_f cannot be both small for large n, which causes either not having identical bit strings or having practically no interference between consecutive frame transmissions.

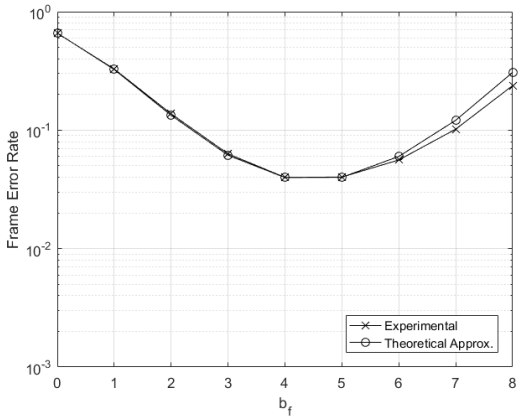
Figure 4.7 shows that the deviations between the experimental and theoretical curves are found to be more pronounced for larger b_f and smaller b_i. This is an excellent sign of the accumulation effect, as the probability of re-occurence $\frac{1}{2^{b_i}}$ becomes non-negligible for small b_i. Furthermore, it is also found that for the considered cases with large b_f, the experimental FER is found to be slightly lower than what (4.10) yields. Recalling the accumulation effect mentioned in Subsection 4.2.1, this suggests that the accumulation effect is mostly constructive rather than destructive for this scenario, since the experimental FER in small b_i region of the curve is lower than the theoretical approximation.

4.3. MaaF's Relation to MoSK and D-MoSK

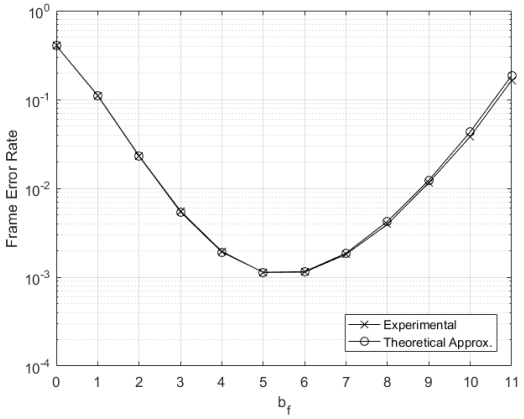
For a MaaF system with the frame length constraint of $n = 1$ bit, there are two options for allocation of the single available bit. First one is to make the single



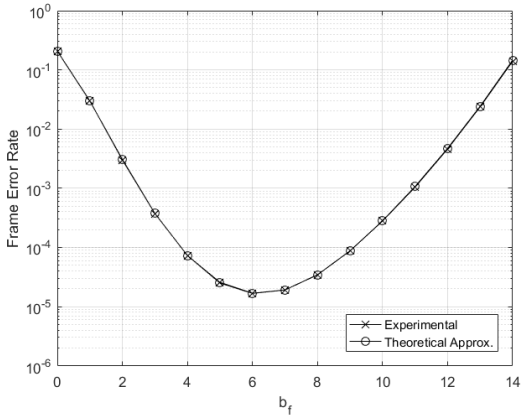
(a) $m = 3$, $t_b = 10$ms, and $n = 6$.



(b) $m = 3$, $t_b = 10$ms, and $n = 9$.



(c) $m = 3$, $t_b = 10$ms, and $n = 12$.



(d) $m = 3$, $t_b = 10$ms, and $n = 15$.

Figure 4.7. Experimental results vs. theoretical approximation of FER for a MaaF scenario with different n. $m = 3$, $t_b = 10$ms, and all other channel parameters are as shown in Table 4.1.

bit an information bit. There are two options to encode in a single bit slot (a bit-1 and a bit-0), as there are two possible distinguishable chemical structures. As it is done in [10], one can call these two different structures type-A and type-B molecules. Encoding a bit-1 with type-A molecules and bit-0 with with type-B molecules is the exact definition of a binary MoSK (2-MoSK) modulation described in Section 2.2.

As another way, the single available bit in the MaaF molecule can be used to provide frame identification. The use of the single slot for frame identification is equivalent to having a cycle size of 2, recalling from Section 3.2.2. However, using the only available bit for sequencing depletes the MaaF molecule from information bit strings,

making $b_i = 0$. In such a case, the system has to fall back on a binary CSK-like quantity modulation to encode the information. Cyclically re-using the two available molecule *types* for frame sequencing and performing binary CSK for communication is exactly what a D-MoSK modulation with 2 orthogonal channel does, as described in [14].

Generalizing the argument for larger n, two extreme ends of the MaaF allocation problem, $b_i = n$ and $b_f = n$, correspond to 2^n-MoSK and D-MoSK with a cycle size of 2^n (2^n-ary D-MoSK as it is called in [14]), respectively. Thus, the popular MoSK and D-MoSK modulations are found to be special allocation cases of MaaF. Even though MaaF looks at things from a communication frame perspective similar to [42], it provides a good compromise between 2^n-MoSK and 2^n-ary D-MoSK for reducing FER. The MaaF approach also shows a solid generalization between the two modulations, and suggests that a synthesis of them works best to minimize the error, rather than the pure versions of the modulations at the extreme ends.

To compare the error performance of MoSK, D-MoSK, and MaaF, the experiment in Figure 4.1 is recalled in Figure 4.8 with the inclusion of $b_f = n$ as an extension. It is noteworthy that Figure 4.8 considers the $b_f = n$ case, unlike other curves in this thesis, which exclude the $b_f = n$ case to avoid putting a quantity modulation result in a curve of frame based type modulations. This exception is only present in this section, which is done to show all MoSK, D-MoSK, and MaaF on the same plot and making the case for MaaF's better error performance. Furthermore, the point corresponding to D-MoSK represents the BER value, since the modulation uses binary CSK to communicate. The argument in Subsection 3.3 also suggests that FER and BER can be used interchangeably, since all bits are assumed to be erroneous in the presence of a frame error.

Figure 4.8 shows that the modulation 2^{15}-ary D-MoSK yields an error rate of 6.3×10^{-2}, under the imposed constraints of $m = 3$ and $t_b = 10$ms. Combined with the error rate for 2^{15}-MoSK of 0.21 under the same constraints, Figure 4.8 states that MaaF's error performance significantly surpasses both modulations, yielding a FER of 1.7×10^{-5} at the optimum point. Thus, given a fixed number of allocatable bits, a

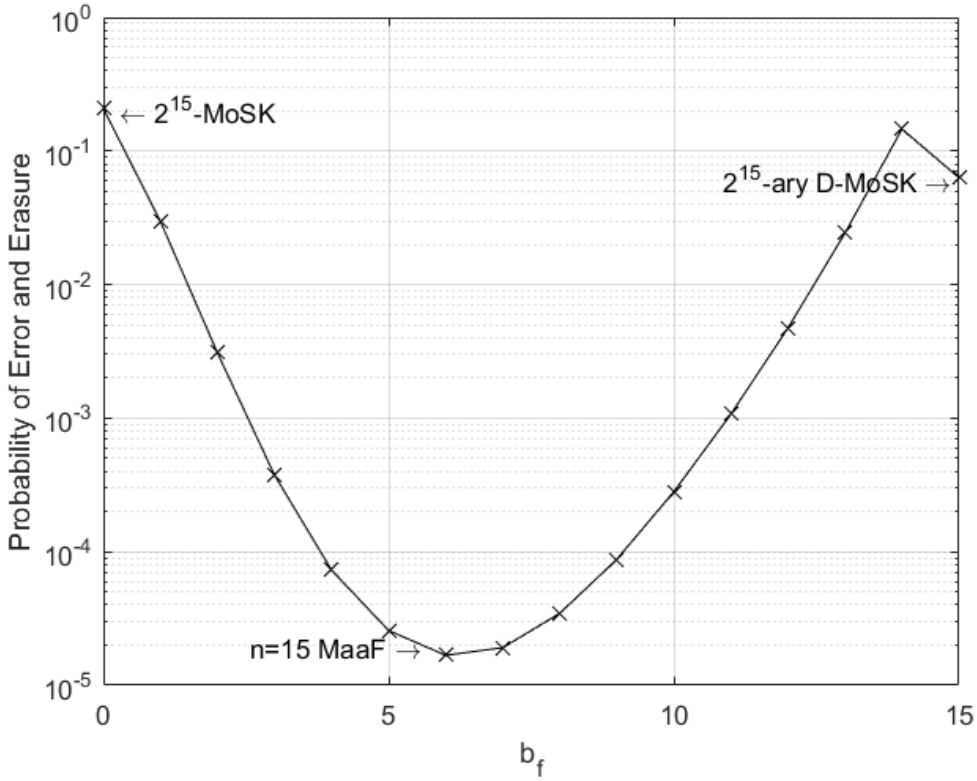


Figure 4.8. Error Rate Comparison for 2^{15}-MoSK, 2^{15}-ary D-MoSK, and $n = 15$ MaaF. $t_b = 10$ms, $m = 3$. d and r_r are as shown in Table 4.1.

MaaF system can perform good ISI reduction while also controlling the noise imposed by the MCvD channel. As a result, the enhanced mechanism reduces the overall error rate.

Additionally, note that the difference in error rates between 2^{15}-ary D-MoSK and $b_f = 14$ MaaF is pronounced and in favor of 2^{15}-ary D-MoSK. Note that the case for $b_f = 14$ still relies on type modulation while sending only 3 frame replicas per transmission. For the receiver to come up with a detection, it needs to collect at least one frame replica. Since only 3 frame replicas is sent per transmission, the necessity of a single molecule's arrival causes a lot of erasures, which increases the probability of error and erasure. However, since 2^{15}-ary D-MoSK is a quantity modulation, the threshold can be set to $\tau = 0.5$, stemming from the nearly one-tap nature of the channel due to high $t_{f_{\text{eff}}}$. However, erasures are not applicable for a D-MoSK modulation, since receiving zero molecules is decoded as a bit-0. Because of this reason, 2^{15}-ary D-MoSK

has a slightly lower error rate than a MaaF with $b_f = 14$ when m is as low as 3.

Since the channel is very close to a single tap channel for 2^{15}-ary D-MoSK, ISI is nearly completely eliminated, which makes the probability of and error for a bit-0 transmission substantially low. In such a case, error events can be thought to occur only for bit-1 transmissions, which happens when all the transmitted $m_f = 3$ molecules fail to arrive at the receiver. As $t_{f_{\text{eff}}}$ is extremely high in this scenario, the single channel coefficient converges to $p_1 \approx \frac{r_r}{r_0}$, as suggested by (2.3). This makes the probability of the a single molecule's failing to arrive at the receiver to be $1 - \frac{r_r}{r_0}$. Considering $r_r = 5\mu$m, $r_0 = 10\mu$m, and $m_f = m = 3$ for 2^{15}-ary D-MoSK, the probability of all m_f molecules failing to arrive at the receiver is found by $(1 - \frac{r_r}{r_0})^{m_f} = (1 - \frac{5\mu m}{10\mu m})^3 = 0.125$. Thus, the error probability of 2^{15}-ary D-MoSK is roughly $6.3 \times 10^{-2} = 0.063$, the joint probability of transmitting a bit-1 (0.5) and the probability of all m_f molecules failing to arrive at the receiver (0.125).

5. EXTENSION OF MAAF FOR MULTIPLE RECEIVER SYSTEMS

Nano-networks are thought to be consisting of multiple nano-machines communicating to one another. In an MCvD scenario with multiple nano-machines, there exists numerous communication links between multiple transmitters or receivers. Ideally, these links should operate in complete orthogonality to each other. However, utilizing CSK-like methods imply the use of a common resource that is the channel, introducing CCI to the MCvD system. Unfortunately, the existence of CCI is a major issue to be dealt with to establish nano-networks [43].

Motivated by the importance of CCI on multi-user nano-networks, [44] derives an analytical expression for the fraction of arriving molecules for multiple transmitter-single receiver scenarios using stochastic geometry. Based on the findings of [44], [45] studies the severity and effects of the CCI in multiple transmitter-single receiver scenarios. Inspired by CDMA methods in traditional communications [46], works in [47, 48] discuss CDMA-like schemes to overcome CCI. Furthermore, [49], considers transmission rate optimization to maximize communication efficiency in presence of CCI.

In this chapter, the case with a single transmitter and multiple receivers is addressed. When the transmitted messenger is different between all transmitter-links, using traditional modulation methods such as CSK, MoSK, MCSK, etc. suffer from high CCI. This phenomenon is due to the fact that the receivers cannot differentiate which messenger molecule is being sent to which receiver, since they all share the same propagation medium. To overcome this, [30] suggests using different types of molecules for each receiver connection to create orthogonal channels, while relying on CSK in each link.

Although the method proposed in [30] is a very good approach in terms of eliminating CCI completely, it brings another question to the table regarding the absorption

of other connections' molecules. Is it more beneficial for the system if the other receivers absorb and ignore the molecules that are not destined for them with appropriate receptors, or is it better if they simply reflect messenger molecules that belong to other connections?

Absorbing or reflecting other connections' molecules has different pros and cons in terms of channel properties. The first approach, absorbing, is good for reducing ISI. ISI-causing molecules are usually the molecules that do not follow the line-of-sight path, and take longer paths to arrive at the intended receiver. Presence of other absorbing receivers capture and remove these astray molecules from the channel, reducing ISI. However, under the absorbing receiver assumption, molecules from the channel are completely removed whenever a molecule arrives at any one of the receivers [50, 51]. The absorbing option reduces the total number of arriving molecules at the receiver, which introduces a drawback regarding an increase the relative variance of the arriving molecules. This corresponds to an increase in the noise of an MCvD system [38]. The other option, reflecting the molecules, provides a higher number of received molecules than the absorbing approach. However, due to the fact that it allows more astray molecules to reach the receiver, reflecting the molecules result in a channel with more ISI.

5.1. Analysis on the Channel Coefficients for Multiple Receiver MCvD Scenarios

Throughout this chapter, the considered multiple user MCvD scenario consists of a single point transmitter and four identical and equidistant spherical receivers, and a 3D unbounded environment. Figure 5.1 visually presents the scenario of interest.

For the scenario presented in Figure 5.1, using four different molecules for each transmitter-receiver connection eliminates CCI by creating four orthogonal channels [30]. With the orthogonalized channel consideration, Monte Carlo simulations are performed to generate FIR channel coefficients for the intended transmitter-receiver connection in the presence of three interferer receivers, in a manner similar to pre-

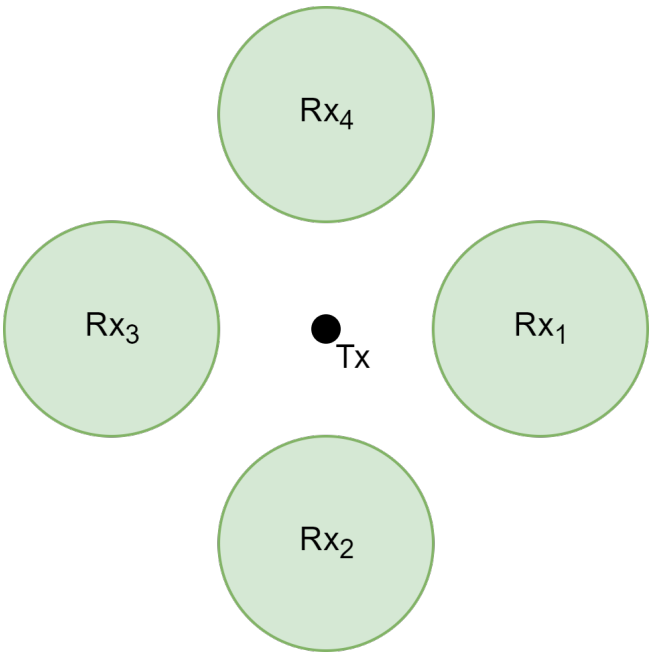


Figure 5.1. A molecular communication system with a single point transmitter and four equidistant and identical receivers.

sented in [50]. Note that the approach is also described in Section 2.1 for a single transmitter-single receiver scenario. The simulations are done separately for both absorbing or reflecting interferer receiver approaches. When simulating the absorbing case, all receivers are assumed to perfectly and instantly absorb every molecule that arrives. On the other hand, simulating the reflecting case requires to consider the interferer receivers to elastically reflect the molecules that hit their surface. In both of the simulations, the diffusion coefficient of the messenger molecule is $D = 79.4\frac{\mu m^2}{s}$. Furthermore, all receivers' centers are $r_0 = 10\mu m$ away from the point transmitter, and their radii are $r_r = 5\mu m$.

For a $t_s = 200$ms that is selected for demonstrative purposes, the first four channel coefficients $(p_1, ..., p_4)$ for each option are shown in Table 5.1 as simulation results. Additionally, the single transmitter-single receiver scenario with the same parameters is also presented for benchmarking.

The results of Table 5.1 agrees with the aforementioned discussion on the absorbing and reflecting option. For example, the ratios $\frac{p_1}{p_i}$ for $i = 2, 3, 4$ are found to

Table 5.1. FIR channel coefficients for different MCvD multi-user approaches for $t_s = 200ms$, $r_r = 5\mu m$, $r_0 = 10\mu m$, and $D = 79.4\frac{\mu m^2}{s}$.

Scenario	p_1	p_2	p_3	p_4
1 Receiver	0.1875	0.0777	0.0390	0.0171
4 Receivers, Absorbing	0.1541	0.0410	0.0120	0.0029
4 Receivers, Reflecting	0.1908	0.0899	0.0448	0.0267

be much higher for the absorbing case, compared to the reflecting scenario. In fact, the absorbing case is even better than the single receiver scenario in this regard. This translates to having a lower ISI, and helping the overall channel response greatly. Additionally, the absolute magnitude of p_1 is smallest for the absorbing case, which can be explained by the interferer receivers *stealing* some molecules originally transmitted for the intended link.

5.2. Error Performance of CSK under Orthogonalized Multiple Receiver MCvD Scenarios

The channel coefficients presented in Table 5.1 agrees with trade-off between ISI combating and received signal noise combating between different cases. To compare the true BER performances of the approaches, an MCvD communication scenario with CSK modulation is imposed on all of them. The BER simulation of binary CSK for a single transmitter-receiver connection in an environment as shown in Figure 5.1 is presented in Figure 5.2. For benchmarking purposes, the simulation also includes the single receiver scenario.

Figure 5.2 shows that when communicating at a t_s of 200ms, it is better to equip every receiver with every molecule's appropriate absorbing receptor, and ignoring the absorbed molecules. The absorbing option combats ISI, and decreases the overall BER.

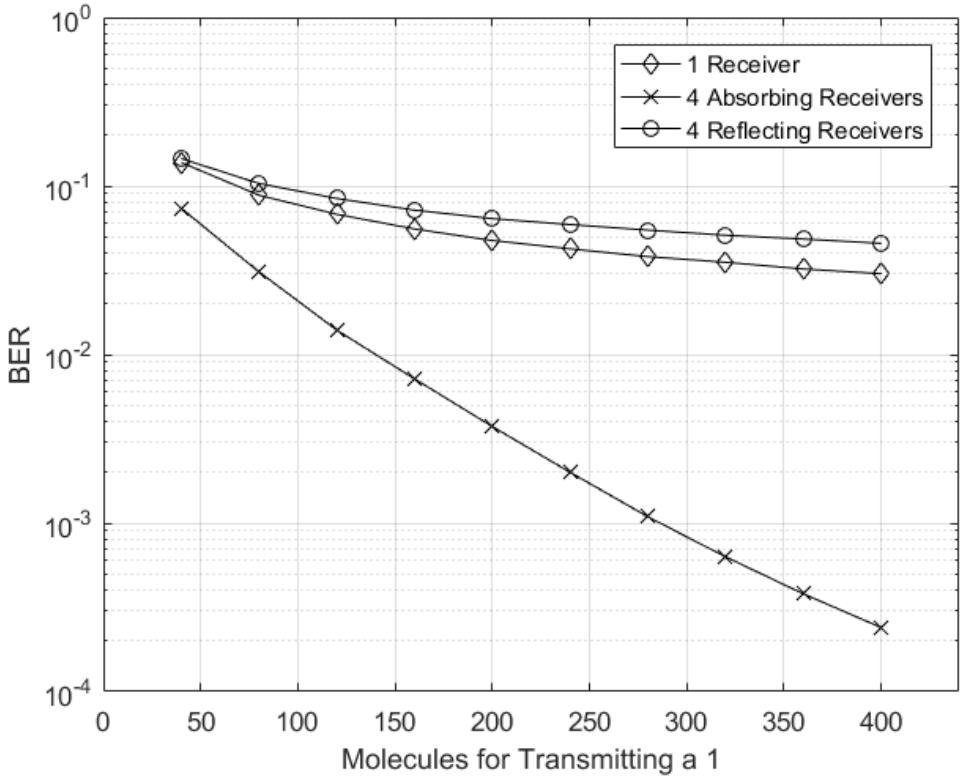


Figure 5.2. BER vs. Transmission Power of BCSK for a transmitter-receiver connection, in a channel with 3 other receivers. $t_s = 200$ms, $L = 30$, $r_0 = 10\mu m$, $r_r = 5\mu m$, and $D = 79.4\frac{\mu m^2}{s}$.

5.3. Extending MaaF to Support Multiple Receiver MCvD Scenarios

In molecule chains, attaching a separate header prefix in front the frame molecule in equivalent to using different molecules and associated molecules in terms of receiver specification [4]. For such an approach, the attached header should be specific for the intended receiver. This exclusiveness in the header inhibits a molecule's absorption from different receivers, unless those receivers are also equipped with those receptors (like in the absorbing receiver case). This way, the header can operate as a destination address in a molecular communication scenario with multiple receivers. Figure 5.3 shows the general destination addressing scheme of a MaaF molecule and its header prefix.

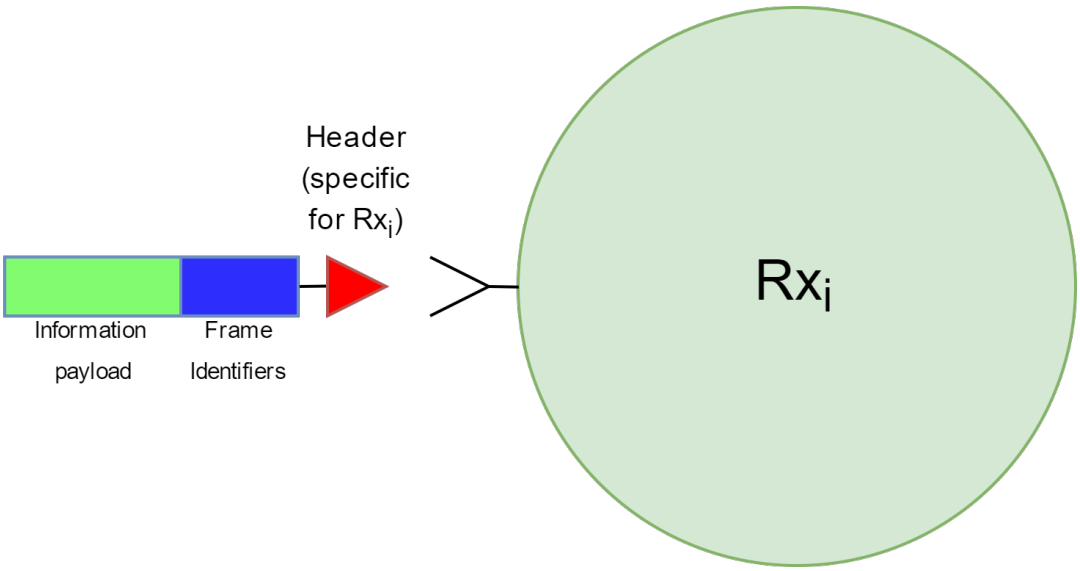


Figure 5.3. General hypothetical structure of a MaaF molecule and its attached header.

Similar to the discussion in Subsection 5.1, the question of whether or not equipping the interferer receivers with the receptors of the intended connection's header prefix is present in MaaF systems as well. Hence, similar to the analysis done in Section 5.2, a FER comparison among different approaches is needed. Considering the transmitter-receiver connection is orthogonalized via the attached header prefix, Figure 5.4 shows the FER curve comparison for the absorbing and reflecting options for the scenario considered in Figure 5.1. Furthermore, the simulation also includes the single receiver scenario for benchmarking purposes. Similar to Table 4.1's parameters, all receivers have $r_0 = 10\mu m$ and $r_r = 5\mu m$. There are $L = 30$ FIR channel coefficients, and the diffusion coefficient $D = 79.4\frac{\mu m^2}{s}$.

In parallel with Figure 5.2, the absorbing option in Figure 5.4 yields the lowest FER on the left-hand side, where b_f is small. This is due to the fact that $t_{f_{\text{eff}}}$ in very low for small b_f, since the system is operating at 100bits/sec. Since the ISI reduction is not efficient for low b_f, getting additional help from the absorbing interferer receivers benefits the system performance. However, comparing the optimum points of both the options (which are also the desired operation points), the reflecting receiver option surpasses the absorbing option by a large margin. Furthermore, it even performs better than the single transmitter-single receiver benchmark scenario. The explanation of

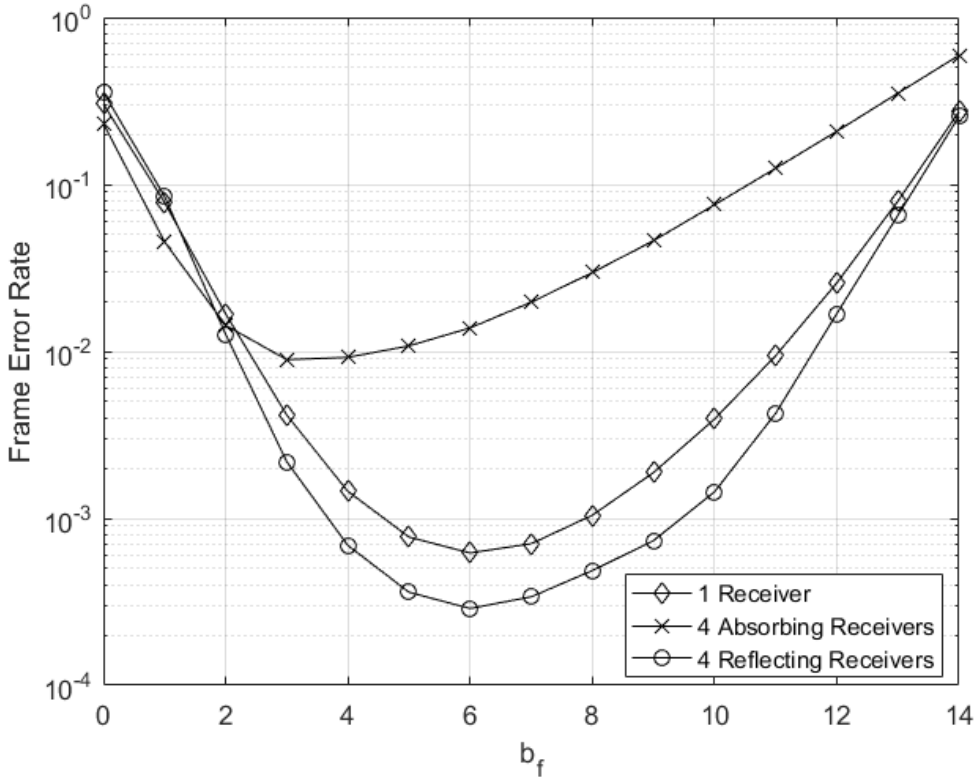


Figure 5.4. FER vs. b_f of a MaaF system for a transmitter-receiver connection, in a channel with 3 other receivers. $n = 15$, $m = 2$, and $t_b = 10$ms.

this phenomenon lies in the ISI reduction efficiency of the sequencing mechanism of MaaF. In a MaaF system operating at the optimal point, $t_{f_{\text{eff}}}$ is already sufficiently high. This provides enough ISI reduction for the MaaF system. For such a system, reducing ISI even further at a price of having lower number of arriving molecules (as it is done by the implementing absorbing interferer receivers) becomes a bad option. A MaaF system enjoys the increased number of expected frame molecule arrivals by implementing reflecting interferer receivers, without losing sufficiency in ISI reduction. Thus, it can be concluded that using receivers equipped only with their own receptors are better in terms of the frame error rate, given the MaaF system is operating at the optimal allocation point.

6. ROBUSTNESS ANALYSIS UNDER ADVERSE CHANNEL CONDITIONS

Generally speaking, the messenger molecules are assumed to freely diffuse with a constant variance per incremental step, as described in (2.1). Furthermore, the molecules are assumed to stay in the channel without decomposing into other molecules at all. However, in real-life applications, the molecules' diffusivity may change with fluctuating channel temperature [52]. The messenger molecules are also organic molecules propagating in a fluid environment, possibly in an in-vivo application. In such a scenario, the messenger molecules may also be subject to degradation before their arrivals to the receiver [53]. For any scheme that should operate under such channel conditions, robustness to these adverse effects is a very desirable feature to have. Stemmed from this argument, this chapter deals with the robustness analysis of MaaF under molecule degradation channels and channels with fluctuating temperature.

6.1. Robustness of MaaF in a Channel with Molecule Degradation

6.1.1. The Channel Response

As discussed in [53, 54], messenger molecules are subject to decomposing after their transmissions in the presence of an enzyme or an inhibitor molecule that reacts with them. This decomposition makes them useless for communication, as the receiver receptors do not recognize them as messenger molecules anymore. That is why the messenger molecules are said to *degrade*. Generally, the presence of the enzyme is a natural phenomenon. However, considering degradation affects the channel capacity [55], deliberate degradation may also be induced by the transmitter for pulse shaping or pre-equalization [20, 56].

As discussed in [53], the concentration of the substrate and its degraded form are governed by

$$\mathrm{E} + \mathrm{S} \underset{\mathrm{k}_{-1}}{\overset{\mathrm{k}_1}{\rightleftharpoons}} \mathrm{ES} \xrightarrow{\mathrm{k}_p} \mathrm{E} + \mathrm{P} \tag{6.1}$$

where E, S, and P denote the enzyme, substrate (the messenger molecule) and the end product, respectively. k_1 and k_{-1} are the reaction constants of the forward and reverse reaction, and k_p is the reaction constant for the formation of the end product [57].

Note that if k_p is sufficiently large, the second part of the reaction becomes so fast, causing the concentration of the ES complex to stay very low. In such a case, the reaction in (6.1) becomes one-sided, ignoring the equilibrium in the first step. This causes the messenger molecule to exhibit exponential decay, where the concentration of the messenger molecule is governed by

$$C(t) = C_0 e^{\lambda t} \tag{6.2}$$

where C_0 is the initial concentration, and λ is the rate of degradation and is equal to $\frac{\ln(2)}{\Lambda_{1/2}}$. $\Lambda_{1/2}$ denotes the half-life of the molecule in the environment.

When using messenger molecules that exponentially decay in a driftless MCvD channel, the fraction of molecules that are absorbed by the receiver is found by considering both arrival to the receiver and survival until doing so. That is to say, for a messenger molecule to reach the receiver, it also needs to avoid getting degraded in the communication environment as well. Hence, the time distribution of absorption is found by integrating the joint probability of the two events. Note that the first event has the time distribution (2.2), and the second event is the tail distribution of 6.2. Furthermore, since the arrival and survival events are statistically independent to one another,

$$f_{hit,d} = \int_0^\infty f_{hit}(t) P(t_{\mathrm{deg}} > t) dt \tag{6.3}$$

where $P(t_{\text{deg}} > t)$ denotes the probability for the degradation time t_{deg} being greater than time t. $f_{hit,d}$ gives the time distribution of the received molecules. Hence, its integral yields $F_{hit,d}$, the fraction of arriving molecules until time t [53] for the scenario of interest.

To visualize the effect of an MCvD channel with messenger molecule degradation on MaaF, Figure 6.1 is presented. It is worth noting that all chemical combinations of the frame molecule family are assumed to have identical half-lives and decay profiles.

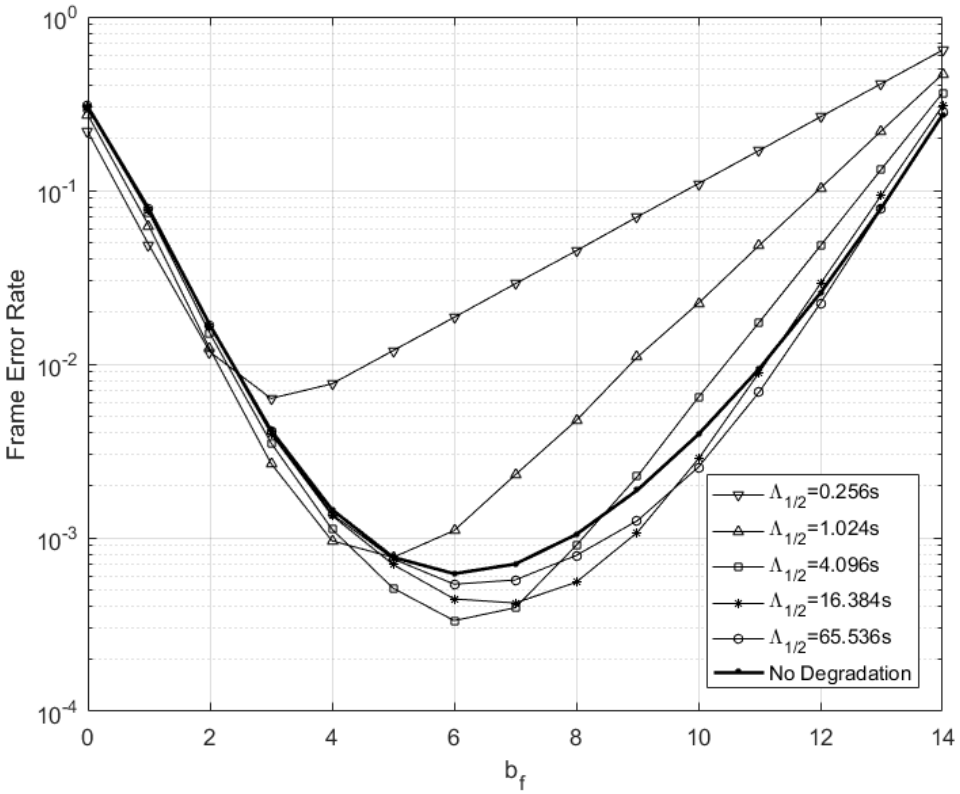


Figure 6.1. FER curves for different $\Lambda_{1/2}$ half-life values. $n = 15$, $t_b = 10$ms, and $m = 2$. All other channel parameters are the same as in Table 4.1.

Figure 6.1 shows that a MaaF system's optimal allocation point and the FER at that point are affected differently in different regions of the half-life parameter $\Lambda_{1/2}$.

- For very low $\Lambda_{1/2}$ values, there is extreme degradation in the channel. The fast exponential decay decreases the arrival probability of the molecules that take longer paths to reach the receiver, since some of them may not survive until

reception. Consequently, the FIR channel coefficients converge to a single-tap channel's, which makes a MaaF system need less ISI combating, therefore reduce b_f at the optimal allocation point. Furthermore, allocating more bits for b_i allows for a larger m_f to increase the amount of frame molecules that arrive at the receiver end, which is especially vital to combat extreme degradation. Thus, the optimal allocation point is observed to shift to the right. In addition, the overall FER is found to increase for very low $\Lambda_{1/2}$ values, since a large proportion of the transmitted frame molecules degrade almost instantly.

- As $\Lambda_{1/2}$ increases, $t_{f_{\text{eff}}}$ at the optimum point becomes comparable to the half-life of the molecule. For this example, the optimum point for the MaaF system in a channel without degradation is at $b_f = 6$, which makes $t_{f_{\text{eff}}} = 5.76$s. When $\Lambda_{1/2}$ starts to get comparable to $t_{f_{\text{eff}}}$, the exponential decay starts to capture the ISI-causing intervals more and more, which actually helps a MaaF system. Comparatively analyzing the FER curves for the cases with $\Lambda_{1/2} = 4.096$ and no degradation, it can be seen that the cases with $\Lambda_{1/2} = 4.096$ yields a lower FER at the optimum point. Additionally, as $\Lambda_{1/2}$ increases, the optimum point shift to the right due to increased need in ISI reduction and decreased need in a larger m_f.
- When $\Lambda_{1/2}$ is increased even further, $t_{f_{\text{eff}}}$ becomes relatively too small compared to the decay rate of the frame molecules, and degradation's effect starts to flatten. As a result, FER slowly converges to the FER curve for the channel without degradation.

Considering a MaaF system gets affected by $\Lambda_{1/2}$ differently for different regions of it, Figure 6.2 is presented to visualize FER vs. $\Lambda_{1/2}$ at the optimal allocation points.

Figure 6.2 tells the same story as 6.1 does. If the exponential decay takes place too fast, molecules get instantly decomposed and become useless for communication. The faster the degradation is, the more molecules get lost in the channel. This, in turn, yields relatively increased FER values. For the region where $\Lambda_{1/2}$ is comparable to $t_{f_{\text{eff}}}$,

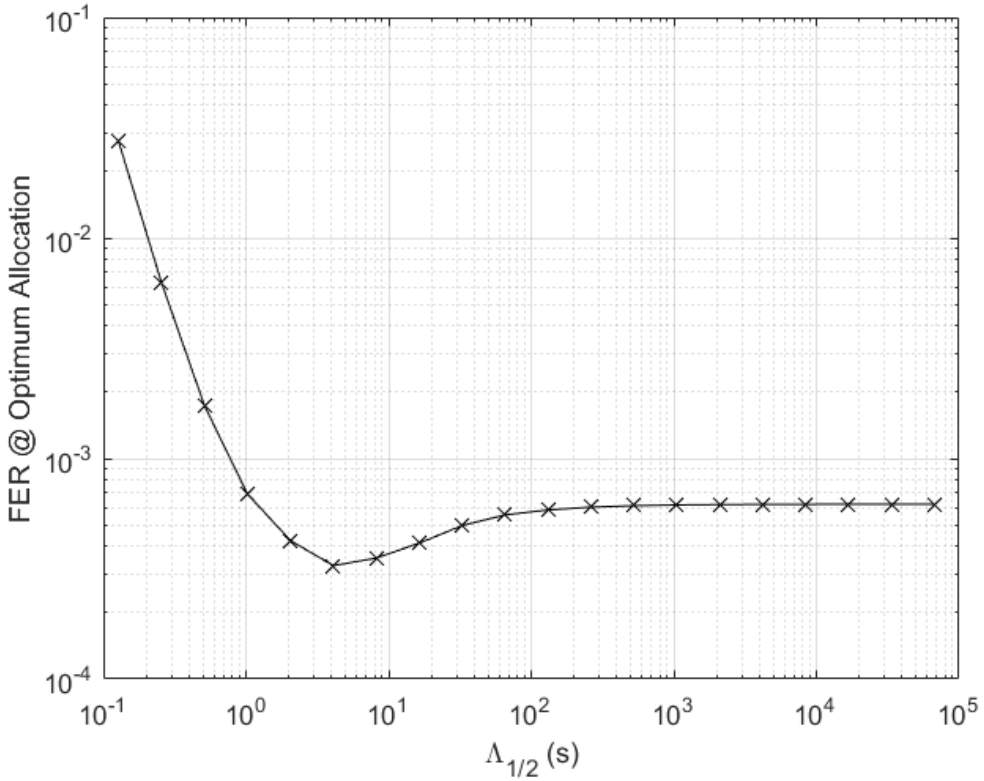


Figure 6.2. Optimal FER vs. $\Lambda_{1/2}$ curve. $n = 15$, $t_b = 10$ms, and $m = 2$. All other channel parameters are the same as in Table 4.1.

FER of a degradation channel is actually better than the degradation-free channel, due to the aforementioned phenomenon. Finally, when $\Lambda_{1/2}$ is increased even further, the channel converges to a regular MCvD channel without degradation as $\Lambda_{1/2} \to \infty$.

6.2. Robustness of MaaF Under Varying Diffusivity

As mentioned in [52], the temperature in the MCvD environment varies over time. In most of the cases, this variation is only slight, but the findings of [52] show that these slight variations still manage to cause performance decrease for CSK. Furthermore, [58] shows that the human body temperature follows a Gaussian distribution, verifying [58] in this regard. Considering molecular communications' prospects to establish communication in nano-scale in-vivo conditions, performance evaluation of MaaF in temperature-varying channels is beneficial.

To test the performance in a channel with time-varying temperature, a control group with a fixed temperature is used for comparison. For the benchmark scenario, insulin messenger molecules at a fixed temperature of 310K is used, which yields a D value of $D = 79.4\frac{\mu m^2}{s}$. The experiment group consists of a scenario where the absolute temperature T is normally distributed with mean $\mu_T = 310$K and standard deviation $\sigma_T = 25$K. Note that the standard deviation is chosen abnormally large to test MaaF's performance under extreme temperature changes. In reality, a standard deviation of $\sigma_T = 25$K is inapplicable to in-vivo systems, but may come in extreme scenarios of other applications.

To generate Figure 6.3, a random temperature is drawn, and 10,000 frames are transmitted for each trial corresponding to the random temperature. Channel temperature is assumed to stay constant throughout the trial. The simulation is repeated 10^5 times to find the best and worst frame error rates for each allocation to come up with reliable empirical bounds on FER.

Even with a high standard deviation of $\sigma_T = 25$K, the frame error rate performance has a low variance, which is desirable. Under the sample size of 10^5, the standard deviation of the FER distribution is found as 1.01×10^{-4}. The reason for this performance is due to the significantly low ISI when operating at the optimal allocation point.

Combining the findings of [38] on the sensitivity of the threshold τ on CSK's BER performance and [52]'s analysis on temperature fluctuations, in can be inferred that threshold-based quantity modulations are not resistant to temperature changes [59]. Since all transmitted frames are affected equally by the temperature changes in the channel, the differences in their arrival counts is a relatively more stable variable than its absolute quantity. Hence, MaaF's robustness under temperature fluctuations is also due to the maximum detection instead of a threshold based approach.

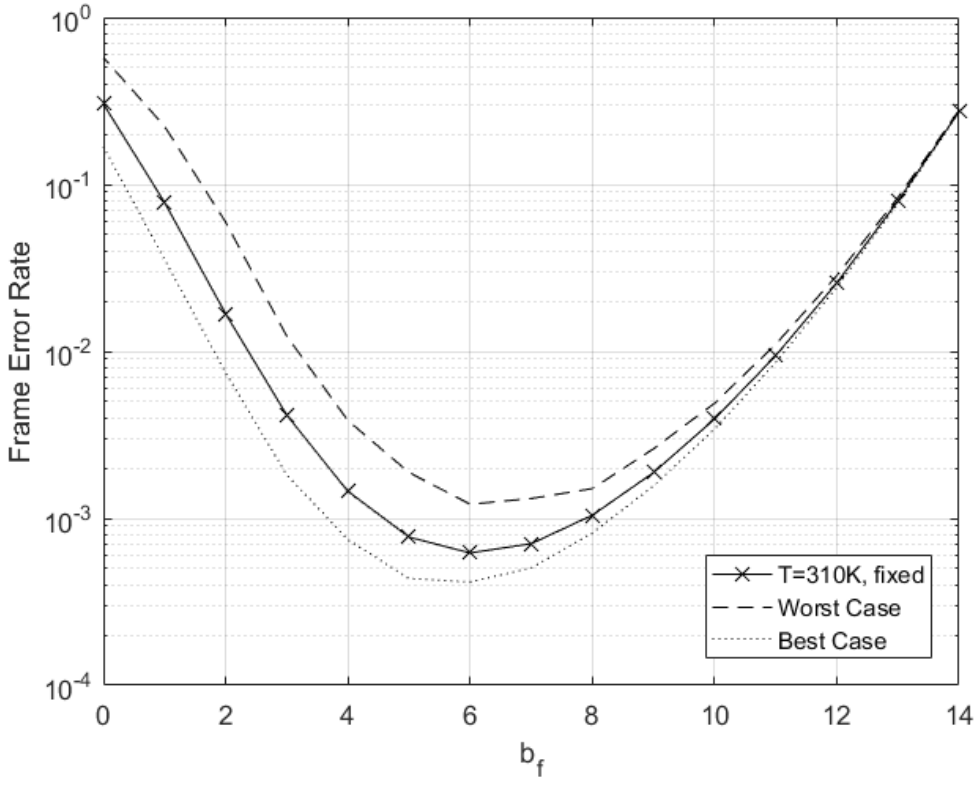


Figure 6.3. Empirical FER bounds for $m = 2$, $n = 15$, $t_b = 10$ms, and $T \sim N(\mu_T, \sigma_T^2)$. $\mu_T = 310$K, $\sigma_T = 25$K.

7. CONCLUSIONS

Motivated by finding reliable methods to reduce ISI in MCvD systems, this thesis revolves around the idea of the possibility to encode bits within the chemical structures of messenger molecules. Inspired by the frame structure from traditional computer networks, a messenger molecule's inner structure is considered a full communication frame. In the frame, an overhead with frame identifiers and a payload with information bits can be encoded. The frame identifier bits work as sequence numbers and are limited in number due to the fixed frame length constraint imposed by the frame molecule's chemical structure. This arises the need to cyclically re-use sequence numbers. This cyclic re-use exponentially increases the effective frame duration, helping a MaaF system combat ISI very effectively without sacrificing from the bit rate.

As a consequence of the constant bit rate and transmitted molecules per bit constraints imposed on all molecular communication schemes, a trade-off between the amounts of frame identifier and information bits is found. It is mentioned that the trade-off can also be thought as a trade-off between ISI combating and noise combating. This suggests the existence of an optimal allocation point between these parameters, which yields good FER values while maintaining high bit rates. Furthermore, extreme ends of the allocation options of a MaaF system represent higher orders of two well-known modulations in the molecular communication literature. MaaF also represents a synthesis of MoSK and D-MoSK. It is realized by experimental and theoretical results that MaaF performs much better by combining these approaches in an optimized manner.

By attaching a header extension in front of the MaaF molecule, individual receiver specification and destination addressing can be accomplished. This eliminates CCI and paves the way for multi-user nano-networks. For modulations that suffer from ISI, equipping the other receivers with receptors for the intended link and telling them to ignore the molecules help the system in terms of decreasing BER. However, MaaF systems work better if every unintended receiver is reflecting instead of absorbing,

due to its already strong ISI reduction mechanism. It is also found that MaaF systems perform well under temperature-varying channels and channels with degradation. Considering the in-vivo MCvD applications, robustness to temperature fluctuations is a very desirable property. Additionally, MaaF systems are robust against messenger molecule degradation, unless the degradation profile is extreme.

Throughout this thesis, the frame identifier bits and the information content is assumed to be encoded and decoded perfectly. Due to this, an error control coding approach to MaaF systems is a possible future work. Since the benefits of establishing bit reliability in the overhead and payload are of different importance, it is expected that the problem leads towards an unequal error protection approach. Furthermore, it is believed that the discussion of MaaF on multiple transmitter/receiver scenarios is a broad subject that holds more questions. Hence, this topic requires specific attention as another future work.

REFERENCES

1. Nakano, T., A. W. Eckford and T. Haraguchi, *Molecular communication*, Cambridge University Press, 2013.

2. Farsad, N., W. Guo and A. W. Eckford, "Tabletop molecular communication: Text messages through chemical signals", *PloS one*, Vol. 8, No. 12, p. e82935, 2013.

3. Akyildiz, I. F., F. Brunetti and C. Blázquez, "Nanonetworks: A new communication paradigm", *Computer Networks*, Vol. 52, No. 12, pp. 2260–2279, August 2008.

4. Farsad, N., H. B. Yilmaz, C.-B. Chae and A. Goldsmith, "Energy model for vesicle-based active transport molecular communication", *Proc. of 2016 IEEE International Conference on Communications (ICC)*, July 2016.

5. Eckford, A. W., N. Farsad, S. Hiyama and Y. Moritani, "Microchannel molecular communication with nanoscale carriers: Brownian motion versus active transport", *10th IEEE International Conference on Nanotechnology*, pp. 854–858, August 2010.

6. Gregori, M., I. Llatser, A. Cabellos-Aparicio and E. Alarcón, "Physical channel characterization for medium-range nanonetworks using flagellated bacteria", *Computer Networks*, Vol. 55, No. 3, pp. 779–791, 2011.

7. Einolghozati, A., M. Sardari and F. Fekri, "Decode and forward relaying in diffusion-based molecular communication between two populations of biological agents", *2014 IEEE International Conference on Communications (ICC)*, pp. 3975–3980, June 2014.

8. Einolghozati, A., M. Sardari and F. Fekri, "Design and Analysis of Wireless Communication Systems Using Diffusion-Based Molecular Communication Among Bac-

teria", *IEEE Transactions on Wireless Communications*, Vol. 12, No. 12, pp. 6096–6105, December 2013.

9. Einolghozati, A., M. Sardari, A. Beirami and F. Fekri, "Capacity of discrete molecular diffusion channels", *2011 IEEE International Symposium on Information Theory Proceedings*, pp. 723–727, July 2011.

10. Kuran, M. S., H. B. Yilmaz, T. Tugcu and I. F. Akyildiz, "Modulation techniques for communication via diffusion in nanonetworks", *Proc. of IEEE International Conference on Communications (ICC)*, April 2011.

11. Proakis, J. G., M. Salehi, N. Zhou and X. Li, *Communication systems engineering*, Vol. 2, Prentice Hall New Jersey, 1994.

12. 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.

13. 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, March 2013.

14. Kabir, M. H., S. M. R. Islam and K. S. Kwak, "D-MoSK modulation in molecular communications", *IEEE Transactions on NanoBioscience*, Vol. 14, No. 6, pp. 680–683, September 2015.

15. Berg, H. C., *Random walks in biology*, Princeton University Press, 1993.

16. 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, December 2013.

17. 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, December 2014.

18. Damrath, M. and P. A. Hoeher, "Low-Complexity Adaptive Threshold Detection for Molecular Communication", *IEEE Transactions on NanoBioscience*, Vol. 15, No. 3, pp. 200–208, April 2016.

19. Tepekule, B., A. E. Pusane, H. B. Yilmaz, C.-B. Chae and T. Tugcu, "ISI mitigation techniques in molecular communication", *IEEE Transactions on Molecular, Biological and Multi-Scale Communications*, Vol. 1, No. 2, pp. 202–216, June 2015.

20. Tepekule, B., A. E. Pusane, M. . 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, August 2015.

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

22. Rose, C. and I. S. Mian, "Inscribed Matter Communication: Part I", *IEEE Transactions on Molecular, Biological and Multi-Scale Communications*, Vol. 2, No. 2, pp. 209–227, December 2016.

23. Yazdi, S. M. H. T., H. M. Kiah, E. Garcia-Ruiz, J. Ma, H. Zhao and O. Milenkovic, "DNA-Based Storage: Trends and Methods", *IEEE Transactions on Molecular, Biological and Multi-Scale Communications*, Vol. 1, No. 3, pp. 230–248, September 2015.

24. Ortiz, M. E. and D. Endy, "Engineered cell-cell communication via DNA messaging", *Journal of Biological Engineering*, Vol. 6, No. 1, p. 16, 2012.

25. Eckford, A. W., T. Furubayashi and T. Nakano, "RNA as a nanoscale data trans-

mission medium: error analysis", *Proc. of IEEE 16th International Conference on Nanotechnology (IEEE-NANO)*, August 2016.

26. Smith, D. E., T. T. Perkins and S. Chu, "Dynamical scaling of DNA diffusion coefficients", *Macromolecules*, Vol. 29, No. 4, pp. 1372–1373, February 1996.

27. Furubayashi, T., T. Nakano, A. Eckford, Y. Okaie and T. Yomo, "Packet fragmentation and reassembly in molecular communication", *IEEE Transactions on Nanobioscience*, Vol. 15, No. 3, pp. 284–288, April 2016.

28. Doty, D. and A. Winslow, "Design of Geometric Molecular Bonds", *IEEE Transactions on Molecular, Biological and Multi-Scale Communications*, Vol. 3, No. 1, pp. 13–23, March 2017.

29. Kuran, M. Ş., H. B. Yilmaz, T. Tugcu and B. Özerman, "Energy model for communication via diffusion in nanonetworks", *Nano Communication Networks*, Vol. 1, No. 2, pp. 86–95, July 2010.

30. Akgül, Ö. U. and B. Canberk, "An interference-free and simultaneous molecular transmission model for multi-user nanonetworks", *Nano Communication Networks*, Vol. 5, No. 4, pp. 83–96, December 2014.

31. Schulten, K. and I. Kosztin, "Lectures in theoretical biophysics", *University of Illinois*, Vol. 117, 2000.

32. Deng, Y., A. Noel, M. Elkashlan, A. Nallanathan and K. C. Cheung, "Modeling and Simulation of Molecular Communication Systems With a Reversible Adsorption Receiver", *IEEE Transactions on Molecular, Biological and Multi-Scale Communications*, Vol. 1, No. 4, pp. 347–362, December 2015.

33. Tyrrell, H. J. V. and K. Harris, *Diffusion in liquids: a theoretical and experimental study*, Butterworth-Heinemann, 2013.

34. Moore, M. J., T. Suda and K. Oiwa, "Molecular Communication: Modeling Noise Effects on Information Rate", *IEEE Transactions on NanoBioscience*, Vol. 8, No. 2, pp. 169–180, June 2009.

35. Noel, A., Y. Deng, D. Makrakis and A. Hafid, "Active versus Passive: Receiver Model Transforms for Diffusive Molecular Communication", *2016 IEEE Global Communications Conference (GLOBECOM)*, pp. 1–6, December 2016.

36. 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 Communications Letters*, Vol. 19, No. 2, pp. 155–158, February 2015.

37. 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, June 2014.

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

39. Moore, M. J. and T. Nakano, "Oscillation and synchronization of molecular machines by the diffusion of inhibitory molecules", *IEEE Transactions on Nanotechnology*, Vol. 12, No. 4, pp. 601–608, 2013.

40. Goldsmith, A., *Wireless communications*, Cambridge University Press, 2005.

41. Yates, R. D. and D. J. Goodman, *Probability and stochastic processes: a friendly introduction for electrical and computer engineers*, Vol. 41, John Wiley & Sons Hoboken, NJ, 2005.

42. Cerf, V. and R. Kahn, "A protocol for packet network intercommunication",

Vol. 22, No. 5, pp. 637–648, May 1974.

43. Akyildiz, I. F. and J. M. Jornet, "The Internet of nano-things", *IEEE Wireless Communications*, Vol. 17, No. 6, pp. 58–63, December 2010.

44. Deng, Y., A. Noel, W. Guo, A. Nallanathan and M. Elkashlan, "Analyzing Large-Scale Multiuser Molecular Communication via 3-D Stochastic Geometry", *IEEE Transactions on Molecular, Biological and Multi-Scale Communications*, Vol. 3, No. 2, pp. 118–133, June 2017.

45. Dinc, E. and O. B. Akan, "Theoretical Limits on Multiuser Molecular Communication in Internet of Nano-Bio Things", *IEEE Transactions on NanoBioscience*, Vol. 16, No. 4, pp. 266–270, June 2017.

46. Viterbi, A. J. and A. J. Viterbi, *CDMA: principles of spread spectrum communication*, Vol. 122, Addison-Wesley Reading, MA, 1995.

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

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

49. Nakano, T., Y. Okaie and A. V. Vasilakos, "Transmission Rate Control for Molecular Communication among Biological Nanomachines", *IEEE Journal on Selected Areas in Communications*, Vol. 31, No. 12, pp. 835–846, December 2013.

50. Arifler, D. and D. Arifler, "Monte Carlo Analysis of Molecule Absorption Probabilities in Diffusion-Based Nanoscale Communication Systems with Multiple Receivers", *IEEE Transactions on NanoBioscience*, Vol. 16, No. 3, pp. 157–165, April 2017.

51. Arifler, D. and D. Arifler, "Monte Carlo Analysis of Molecule Absorption Probabilities in Diffusion-Based Nanoscale Communication Systems with Multiple Receivers", *IEEE Transactions on NanoBioscience*, Vol. 16, No. 3, pp. 157–165, April 2017.

52. Qiu, S., T. Asyhari, W. Guo, S. Wang, B. Li, C. Zhao and M. Leeson, "Molecular channel fading due to diffusivity fluctuations", *IEEE Communications Letters*, Vol. 21, No. 3, pp. 676–679, March 2017.

53. Heren, A. C., H. B. Yilmaz, C. B. Chae and T. Tugcu, "Effect of Degradation in Molecular Communication: Impairment or Enhancement?", *IEEE Transactions on Molecular, Biological and Multi-Scale Communications*, Vol. 1, No. 2, pp. 217–229, June 2015.

54. Noel, A., K. C. Cheung and R. Schober, "Improving Receiver Performance of Diffusive Molecular Communication With Enzymes", *IEEE Transactions on NanoBioscience*, Vol. 13, No. 1, pp. 31–43, March 2014.

55. Nakano, T., Y. Okaie and J. Q. Liu, "Channel Model and Capacity Analysis of Molecular Communication with Brownian Motion", *IEEE Communications Letters*, Vol. 16, No. 6, pp. 797–800, June 2012.

56. Wang, S., W. Guo and M. D. McDonnell, "Transmit pulse shaping for molecular communications", *2014 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, April 2014.

57. Segel, I. H. *et al.*, *Enzyme kinetics*, Vol. 957, Wiley, New York, 1975.

58. Mackowiak, P. A., S. S. Wasserman and M. M. Levine, "A critical appraisal of 98.6 F, the upper limit of the normal body temperature, and other legacies of Carl Reinhold August Wunderlich", *The Journal of the American Medical Association (JAMA)*, Vol. 268, No. 12, pp. 1578–1580, September 1992.

59. Akdeniz, B. C., M. C. Gürsoy, A. E. Pusane and T. Tuğcu, "On the performance of the modulation methods in time-varying molecular communication channels", *2017 40th International Conference on Telecommunications and Signal Processing (TSP)*, pp. 128–131, July 2017.