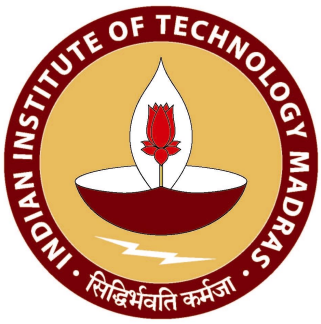
A Study on the modulation techniques and timing synchronization methods in Molecular Communication networks

A Project Report
submitted by

Nandu Sivadasan
EE13B091

in partial fulfilment of the requirements
for the award of the degree of

Dual Degree in Electrical Engineering



Department of Electrical Engineering
Indian Institute of Technology Madras
May, 2018

Department of Electrical Engineering

Indian Institute of Technology, Madras

Certificate

This is to certify that the project titled **"A Study on the modulation techniques and timing synchronization methods in Molecular Communication networks"**, submitted by **"Nandu Sivadasan(EE13B091)"**, to the Indian Institute of Technology, Madras, in partial fulfillment of the requirements for the award of the degree of ***Dual Degree in Electrical Engineering***, is a bona fide record of project work carried out by him in *Department of Electrical Engineering, I.I.T. Madras.* The contents of this project in full or in parts have not been submitted to any other institution for an award of any degree or diploma.

Dr. Arun Pachai Kannu
Project Guide,
Associate Professor,
Department of Electrical Engineering,
Indian Institute of Technology, Madras,
Chennai, Tamil Nadu - 600 036.

Date:
Place:

Acknowledgement

I would like to start by expressing my sincerest gratitude towards my Project guide Dr. Arun Pachai Kannu for giving me an opportunity to do project work under his valuable guidance. He gave me valuable motivation and encouragement throughout my project. He explained me the basic concepts related to the project and also explained the methodology to carry out the project in a stepwise manner. He also taught me about how to present the work as clearly as possible. I am highly indebted for his kind nature and help in both project work and personal life. It was a great honour and privilege to work and study under his valuable guidance.

I am grateful to Dr. Bhaskar Ramamurthy, Director of IIT Madras and Dr. Devendra Jalihal, Head of Electrical Engineering Department for providing me a necessary facilities required for the completion of this project. I also extend my gratitude to all staff members of P.G Senapathy Center for Computing Resource for their co-operation during my simulations.

I am extremely grateful to my parents for what they did for me since from birth. They always do prayers for my better future. They always guide me in every situation of life.

Finally, my thanks go to all the people who have supported me to complete the project work directly or indirectly.

Nandu Sivadasan

Abstract

KEYWORDS: Biological Nanomachines; Concentration Shift Keying; Timing modulation;

Molecular communication is an emerging technology that exploits biological materials or living matter to enable communication among biological nanomachines. Nanomachines are small-scale devices that exist in nature or are artificially synthesized from biological materials. Some examples of nanomachines are biological cells, molecular motors, synthetic molecules, genetically engineered cells, artificial cells, and bio-silicon hybrid devices that are able to process chemical signals.

In this project we study the various existing modulation techniques in molecular communication, in particular we proceed to analyze concentration shift keying(CSK) and timing modulation. Next we see the various timing synchronization techniques used in molecular communication systems. We will see how to use the modulation techniques along with accurate time synchronization in B-MosK environment to establish a reliable communication channel.

Contents

List of Tables

List of Figures

Abbreviations

CSK	Concentration Shift Keying
ML	Maximum Likelihood
MoSK	Molecular Shift Keying
LoC	Lab on Chip
MEMES	Micro Electro Mechanical Systems
CvD	Communication via Diffusion
BCSK	Binary Concentration Shift Keying
QCSK	Quadrature Concentration Shift Keying
LL	Log-Likelihood
IRSK	Isomer based Ratio Shift Keying
PO	Peak Observation
IG	Inverse Gaussian
ISI	Intersymbol interference
AIGN	Additive Inverse Gaussian Noise
Tx	Transmitter
Rx	Receiver

Chapter 1

Introduction

Since Richard Feynman's talk on top-down nanotechnology, people have eagerly pursued practical work on smaller and smaller scales, most notably through nanotechnology. Nanotechnology, recently, opened a new branch of research called nano communication networks (NCNs), which may be realized by several methods. For example, we can rely on traditional RF communication systems. Such a method, however, has to overcome RF device barriers. Therefore, researchers have introduced a new concept utilizing diffusion that is especially useful for short-range communications.[9]

Molecular communication is an emerging communication paradigm for bio-nanomachines (e.g., artificial cells, genetically engineered cells) to perform coordinated actions in an aqueous environment. This interdisciplinary research is considerably different from the traditional communication system, since it utilizes not electromagnetic waves but biological molecules both as carriers and as information. A basic modus-operandi involves one or more transmitter nanmomachines emitting protein/amino acid molecules, which propagates through an mostly aqueous medium and reaches one or more receiver nanomachines.

1.1 Scope of the field

Molecular communication has a variety of potential applications in the biomedical, military, and environmental areas. The most direct and promising applications are in the biomedical field. The advantages provided by molecular communication are from size, biocompatibility, and biostability. Some envisaged applications are drug delivery system (DDS), bio-hybrid implants, and lab-on-a-chip (LoC) systems. Some important applications are : [6]

Micro-electromechanical systems (MEMS): MEMS applies microtechnology to develop a small-scale system such as a lab-on-a-chip (LOC)or a network-onchip (NoC) for on-chip analysis of biological samples (e.g., molecules).

Tissue engineering: Tissue engineering aims to develop a tissue structure from biological cells to restore the lost tissues of a patient's body. In tissue engineering, stem cells (e.g., autologous cells) are extracted from a patient's body and cultured in vitro.

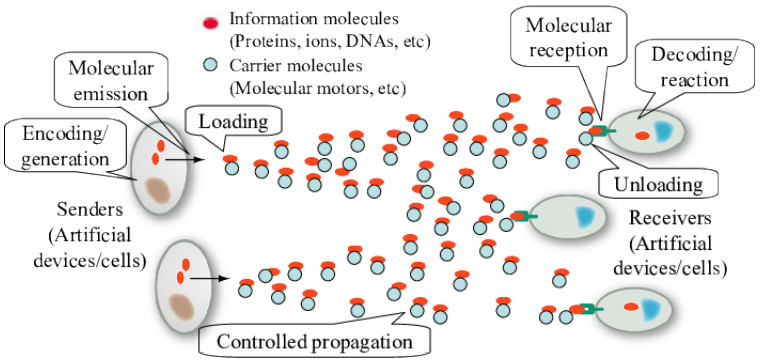


Figure 1.1: A layout of the basic molecular communication system.

1.2 Previous Works

Earlier research work in the field of mathematics by Nyquist and Shannon established information and communication theory. As abstract model in communication these techniques can be used in more general studies of communication such as molecular biological communication. Nonetheless, there has long been interest in information theory as a tool for explaining biological behavior, especially in terms of biomolecular interactions. To the knowledge, the first discussion of information theory in the context of biomolecular interactions occurred in [1] which analyzed the efficiency of the kidney by recognizing its operation in terms of information processing

1.2.1 Modulation schemes

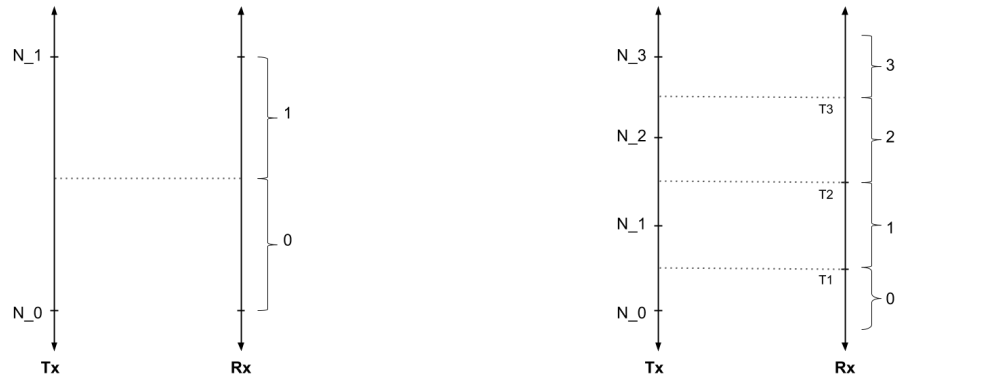
Like in the conventional EM-wave communication methods, carrier modulation can be done, in order to achieve higher data rates. While in radio communication we encode information in the amplitude/phase of the carrier, here we can encode information in the concentration/ release time of molecules.

Modulation schemes based on the concentration of molecules, for short to medium range Communication via Diffusion (CvD) system was proposed in [2].

Molecular Concentration based - Concentration Shift Keying(CSK)

When the concentration of messenger molecules is used to differentiate symbols, the technique is known as concentration based modulation [2]. In Concentration Shift Keying(CSK), the transmitter emits different number of molecules for each value of symbol at the beginning of symbol duration. The receiver interprets the number of molecules as the amplitude of the signal and does a threshold decoding. i.e. for a binary symbol case if the number of molecules received during a particular symbol duration in greater than τ it will be read as some symbol $S0$ otherwise $S1$.

CSK can be implemented in practice as BCSK (Binary CSK) or QCSK (Quadruple CSK) see figure 1.2a and 1.2b, depending on the bits per symbol rate. As the modulation order increases, however, so does the error probability since the minimum distance between two neighboring thresholds decreases, in the case where the number of molecules



(a) Binary - CSK modulation (b) Quaternary - CSK modulation

Figure 1.2: An illustration of CSK Modulation scheme

available to transmit are limited, analogous to the power constrain in conventional communication.

Molecular Type based - Molecular Shift Keying(MosK)

When different types of molecules represent different symbols, the technique is known as molecular-type-based modulation, referred to as molecule shift keying (MoSK) [2]. For the transmission of n information bits in one symbol, 2^n different molecules are utilized,e, each representing a combination of the 2^n different n bit sequences.

The transmitter releases one of these molecules based on the current intended symbol. The receiver decodes the intended symbol based on the type and the concentration of the molecule received during a time slot. If the concentration of a single molecule type exceeds the threshold τ at the receiver, the symbol is decoded based on the bit sequence corresponding to this molecule type. On the other hand, an error is assumed, if the concentration of any molecule types does not exceed the threshold or the concentration of more than one molecule type exceeds the threshold. AN illustration for a binary - MoSK scheme is given in figure 1.3 The MoSK system requires only one threshold to be detected by a receiver nanomachine, which makes it simpler than the system.[2]

Molecular Ratio Based

In addition to the concentration-based and the molecular type-based modulations,[3] also suggests a new modulation technique as ratio-based modulation that can be referred to as isomer-based ratio shift keying (IRSK). IRSK encodes the information based on the ratios of messenger molecules. Deploying this technique yields several benefits. First, it can have a high, theoretically infinite, modulation order. Moreover, we need only two types of molecules in the simplest system though it is possible to deploy many types of molecules for more complex systems.

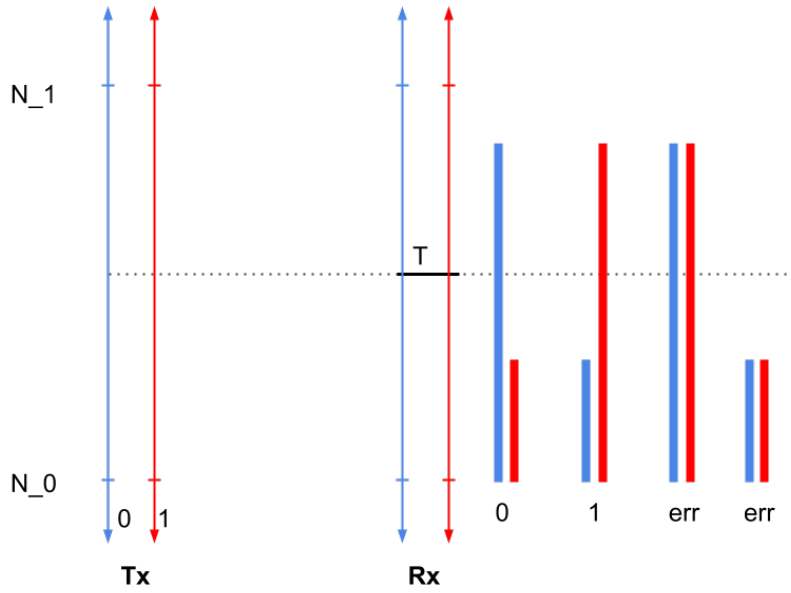


Figure 1.3: An illustration of the Binary MosK modulation scheme

Timing based

When the information is encoded in the release time of the molecules, it is refereed to as timing based modulation [4]. The main kind of noise imbibed within the modulation scheme is the propagation delay between the transmitter and the receiver due to Brownian motion, which leads to different random arrival times at receiver. Since the modulation scheme depends on the release time, achieving higher data rates are comparatively difficult.

1.2.2 Timing Synchronization Methods

One of the crucial requirements for establishing a reliable communication link is symbol synchronization where the start of a symbol interval is determined at the receiver. Most modulation schemes in the previous subsection assume perfect symbol synchronization for data detection. We try to implement the timing detections mentioned in [5]. One of the inherent assumption of the system model used is that the transmitter is not necessarily equipped with an internal clock nor restricted to release the molecules with a constant frequency. To enable symbol synchronization, we use two types of molecules, one for synchronization and one for data transmission. The two methods proposed are.

- Optimal Maximum Likelihood (ML) symbol synchronization scheme as a performance upper bound for the proposed synchronization framework.
- Since ML synchronization entails high complexity, [5] also propose two low-complexity synchronization schemes, namely a peak observation-based (PO) scheme and a threshold-trigger (TT) scheme.

We study the Optimal Likelihood symbol synchronization and one of the suboptimal methods, Peak Observation based method, later in chapter 4.

1.3 Overview

The rest of the thesis is organized as follows. Chapter 2 discusses the mathematical preliminaries of Brownian motion and Wiener process[6]. Chapter 3 describes in detail the system model, medium environment and the modulation schemes used. 4 is a detailed discussion on the theories and our simulations on the existing timing modulation techniques applied to our 1-D environment.

Chapter 2

Brownian motion

The molecules released by the transmitter nanomachine into the medium undergoes Brownian motion. Brownian motion is the random movement of particles in a fluid due to their collisions with other atoms or molecules. Brownian motion is also known as pedesis, which comes from the Greek word for "leaping". Even though a particle may be large compared with the size of atoms and molecules in the surrounding medium, it can be moved by the impact with many tiny, fast-moving masses. Brownian motion may be considered a macroscopic (visible) picture of a particle influenced by many microscopic random effects.[6]

2.1 Wiener Process

The Weiner process is a simple physical model for Brownian motion that is appropriate when friction is minimal. Let $W(t)$ represent the position of a particle in Brownian motion at time $t \geq 0$, where $W(0)$ represents the initial position. Then $W(t)$ is a one dimensional Wiener process if:

1. Foe any times t_1 and $t_2(t_2 > t_1 \geq 0)$, and some constant σ^2,

$$W(t_2) - W(t_1) \sim N(0, \sigma^2(t_2 - t_1)) \tag{2.1.1}$$

2. For two intervals $[t_1, t_2]$ and $[t_3, t_4]$, the increments $W(t_4) - W(t_3)$ and $W(t_2) - W(t_1)$ are statistically independent if the intervals do not overlap.

In the above equation if $t_1 = 0$ and $t_2 = k$ then $W(k)$, the position at time k, is a Gaussian random variable with distribution.

$$W(k) \sim N(0, \sigma^2 k) \tag{2.1.2}$$

In the above equation if $t_1 = k.\Delta t$ and $t_2 = (k+1).\Delta t$ for some finite time interval $0 < \Delta t << 1$ then the increment between successive views of the process $W(k+1) - W(k)$, has the distribution.

$$W((k+1).\Delta t) - W(k.\Delta t) \sim N(0, \sigma^2 \Delta t) \tag{2.1.3}$$

which is independent of k, since the time intervals do not overlap. Thus we have a process whose instantaneous variance increases over time, but the increments are independent and identically distributed.

In a physical Brownian motion, the variance parameter σ^2 is given by

$$\sigma^2 = \alpha D \tag{2.1.4}$$

where D is the free diffusion coefficient of the molecule propagating in the given medium and $\alpha = 2, 4, 6$ if the system is 1-,2- or 3-dimensional, receptively. The value of D is given by

$$D = \frac{k_B T}{6\pi\eta R_H} \tag{2.1.5}$$

where $k_B = 1.38 \times 10^{-23} J/K$ is the Boltzmann constant, T is the temperature, η is the dynamic viscosity of the fluid, and R_H is the hydraulic radius of the molecule.

2.2 Markov Property

The term Markov property refers to the memoryless property of a stochastic process. Consider a sequence of random variables$x_1, x_2, ..., x_m$; the variables are said to follow Markov property if

$$f_{X_1,x_2,....,X_m}(x_1, x_2, x_3, ...x_m) = f_{X_1}(x_1) \prod_{i=2}^{m} f_{X_i|X_{i-1}}(x_i|x_{i-1}) \tag{2.2.1}$$

so that given x_{i-1} each variable x_i is conditionally independent of the past $x_1, x_2, ..., x_{i-2}$. This conditional independence can be used to express the independent increments of a Wiener process. Let $t_1, t_2, ...t_k$ represent a series of observation times, where $0 < t_1 < t_2 < ... < t_k$, and let $W(t_1), W(t_2), ..., W(tk)$ be the corresponding points of a Wiener process (further, let $t_0 = 0$, and let the initial point be $W(0) = W(t0) = 0$). Using basic probability decomposition we can write.

$$f_{W(t_1),W(t_2),...W(t_k)}(w_1, w_2, ...w_k) \tag{2.2.2}$$

$$= f_{W(t_1}(w_1) \prod_{i=2}^{k} f_{W(t_i)|W(t_{i-1}),W(t_{i-2})...W_1}(w_i|w_{i-1}, w_{i-2}..., w_1) \tag{2.2.3}$$

For each $i \in \{1, 2, 3...k\}$ we can write

$$W(t_i) = W(t_i) - W(t_{i-1}) + W(t_{i-1}) \tag{2.2.4}$$

Thus the increment $W(t_i) - W(t_{i-1})$ is independent of any other increments prior to t_{i-1}, so given $W(t_1), W(t_2)...W(t_{i-1})$, $W_(t_i)$ is only dependent on $W(t_{i-1})$. in other words

$$f_{W(t_i)|W(t_{i-1}),W(t_{i-2})...W_1}(w_i|w_{i-1}, w_{i-2}..., w_1) = f_{W(t_i)|W(t_{i-1})}(w_i|w_{i-1}) \tag{2.2.5}$$

for all i, which implies that the Wiener process satisfies Markov property.

2.3 Wiener Process with Drift

Consider a bio-nanomachine traversing a capillary: if the bio-nanomachine released a signal molecule into the bloodstream, that molecule's Brownian motion would be biased in the direction of the blood flow. Thus, we would need to consider the Brownian motion with drift.

Wiener process derived in Section 2.1 was free of drift. That is, the instantaneous average of the Wiener process is always equal to its initial point: $E[B(t)] = B(0)$ for all $t \geq 0$). (To see this, note from (2.2) that the expected value of each increment is zero.) In a Wiener process with drift, the increment distribution in section 2.1 is replaced with

$$W(t_2) - W(t_1) \sim N(v(t_2 - t_1), \sigma^2(t_2 - t_1)) \tag{2.3.1}$$

where v is the drift velocity. Aside from the expected value, the statistical properties of the Wiener process with drift, including the Markov property, are identical to the Wiener process without drift. This is because the drift component is deterministic, and we may subtract it from the motion: for instance, if $W(t)$ is a Wiener process with drift velocity v, then we may form a process $W'(t)$, given by

$$W'(t) = W(t) - vt \tag{2.3.2}$$

which is a Wiener process without drift.

2.4 First Arrival Time

Consider Brownian motion as a Wiener process $W(x)$ in a one dimensional environment, in which the Transmitter is placed at origin and the Receiver is placed at $x = l$(our intended destination). The first arrival time T, in this case is given by

$$T = \min(t : W(t) \geq l) \tag{2.4.1}$$

It is to be noted that if $v > 0$ the pdf of the first arrival time T, denoted by $f_T(n)$, is given by the IG distribution

$$f_T(t) = \begin{cases} \sqrt{\frac{\lambda}{2\pi t^3}} \exp(-\frac{\lambda(t-\mu)^2}{2\mu^2 t}) & \text{if } t > 0 \\ 0 & \text{if } t \leq 0 \end{cases} \tag{2.4.2}$$

where μ and λ are given by

$$\begin{aligned} \mu &= \tfrac{l}{v} \\ \lambda &= \tfrac{l^2}{\sigma^2} \end{aligned} \tag{2.4.3}$$

where the mean and the variance of T are given by $E(T) = \mu$ and $Var(T) = \frac{\mu^3}{\lambda}$, respectively. We will use as shorthand $IG(\mu, \lambda)$ for this distribution; i.e.$T \sim IG(\mu, \lambda)$, implies the relation 2.4.2. It is important to note that if , the distribution of is not IG.

Chapter 3

System model

Suppose that there are two nanomachines, the transmitter and the receiver, in a one-dimensional fluid environment, as shown in figure 3.1. The transmitter is potioned at $x = 0$ and the receiver is at $x = l$. The distance between the two nanomachines is denoted as d [7]. The drift velocity of the fluid is assumed to be constant and denoted as v. The model can be considered an idealization of blood vessel models/capillary motion [6].

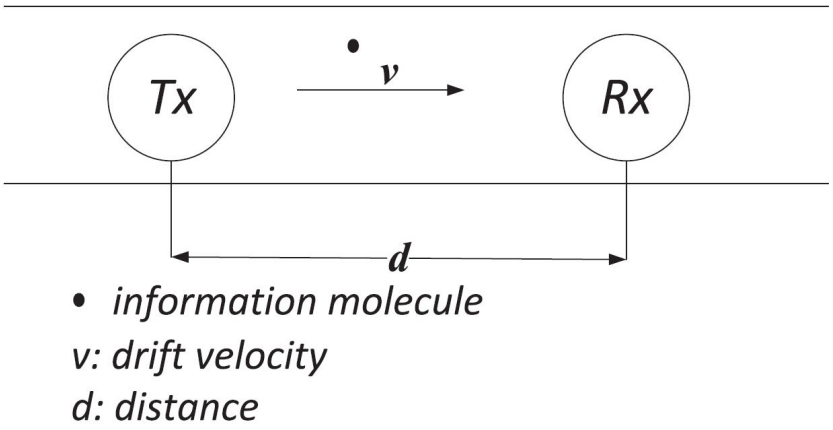


Figure 3.1: Diffusive Communication model with drift

The transmitter releases a number of molecules at a time that is used as the carrier of information. The molecule propagates in the fluid v.i.a Brownian motion from the transmitter to the receiver. It is affected by both diffusion and drift velocity.

In a binary- Molecular shift Keying(B-MosK), earlier discussed in section 1.2.1, the transmitter releases two types of molecules, namely type-A and type-B molecules, where type-A molecules are used for information transmission whereas type-B molecules are employed for synchronization. The released molecules diffuse through the fluid medium between the transmitter and the receiver. We use the ideal channel model assumptions.

- The transmitter perfectly controls the release time and the number of molecules released to represent a particular message.Once released the transmitter doesnot interact with the molecules.
- Perfect absorption happens at the receiver and no molecules bounces back into the medium.

- molecules do not change identity or disappear while propagation i.e, the molecules do not react or chemically interact with the medium.
- The trajectories of individual information carrying molecules are independent.

Depending on the modulation scheme, receiver treats the time-varying numbers of bound type-A and type-B molecules as the received signals for data detection and synchronization, respectively.

The MC channel is characterized by the following two quantities.

- The expected number of type-x molecules bound to the corresponding receptors at the receiver at time t due to the release of molecules by the transmitter in one symbol interval starting at t = 0, which is denoted by $P_x(t) \in \{A, B\}$
- The expected number of external noise molecules bound to the receptors at any given time, denoted by $z_x, x \in \{A, B\}$. We ignore stray molecules for the time.

In order to simulate Brownian motion we use the *Weiner process* approximation, in the case where friction is negligible. Let t_k denote the k'th time instant, then as discussed in section 2.1 the position of a molecule at t_k, propagating v.i.a Brownian motion, can be represented as a Weiner process $W(t_k)$.

3.1 Modulation scheme

3.1.1 Concentration shift Keying

In order to represent different values in symbols, the transmitter releases different number of molecules for each value the symbol can represent. A binary example would be : for "0" the transmitter releases no molecules whereas for "1", n1 molecules are released.

CSK is analogous to Amplitude Shift Keying(ASK) in classical communication. Instead of using two n values eg. 0 and n1, the symbols can be tailored to represent b bits using 2^b different values with $2^b - 1$ threshold levels. See figure 1.2a and 1.2b

In general for concentration shift keying, assume that the transmitter needs to convey one of the N messages; the message alphabet would be $\chi = \{x_1, ..., x_N\}$with uniform prior probability $Pr\{x_i\} = p_i$. With $p_i = 1/N$, the transmitter could convey up to $\log(N)$ nats per channel use.// Corresponding to the i'th message,$1 \leq i \leq t$, the transmitter releases x_i molecules into the fluid medium, at once, at the start of the symbol. The receiver counts the number of molecules arrived during the i'th symbol duration y_i and computes an estimate $\hat{x}$ of the transmitted message from y_i. Notice that all the x_i molecules transmitted during at the beginning of the i'th symbol need not arrive during the current symbol duration, some can interfere into the next symbol or can get lost.The transmission is successful when $\hat{X} = x_i$, and we declare an error when $\hat{X} \neq x_i$.

3.1.2 Timing modulation

We consider molecular communication with timing modulation for channel A; i.e., information is encoded in the release times of molecules into the fluid medium.

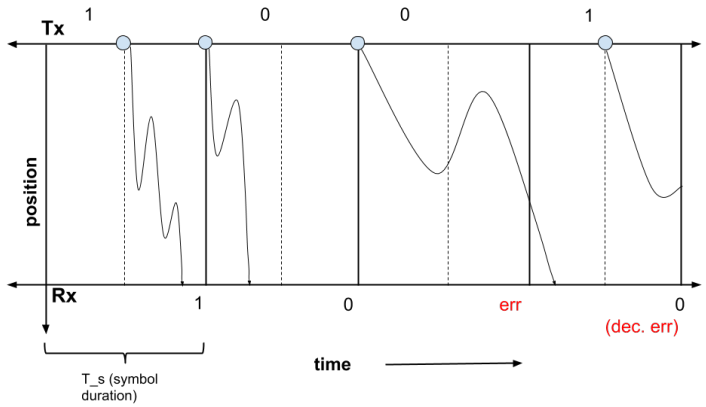


Figure 3.2: An illustration of a naive Timing modulation done with one message molecule for the sequence 1001

As an example of a molecular communication system with timing modulation as in [4], assume that the transmitter needs to convey one of the 2^b messages; the message alphabet would be $\chi = \{X_0, ..., X_{2^b-1}\}$with $Pr\{x_i\} = p_i$. Usually, if b bits is to be transmitted per symbol, then the total symbol duration is divided into 2^b equal intervals, the start of each subinterval denoted as x_k for $k \in \{0, 1, .., 2^b - 1\}$. With $p_i = 1/t$, the transmitter could convey up to $\log(t)$ nats per channel use. Corresponding to the i'th message,$1 \leq i \leq t$, the transmitter releases a molecule into the fluid medium at time x_i.

For the receiver, the transmitted message is a discrete random variable X with alphabet χ and it observes $Y = x_i + N$ where N is the time taken by the molecule to arrive at the receiver. The receiver computes an estimate $\hat{x}$ of the transmitted message from Y, making use of other information such as the pdf of N and the a priori probabilities of the messages. From section 2.4 it can be seen that N has an Inverse Gaussian distribution. The transmission is successful when $\hat{X} = x_i$, and we declare an error when $\hat{X} \neq x_i$.

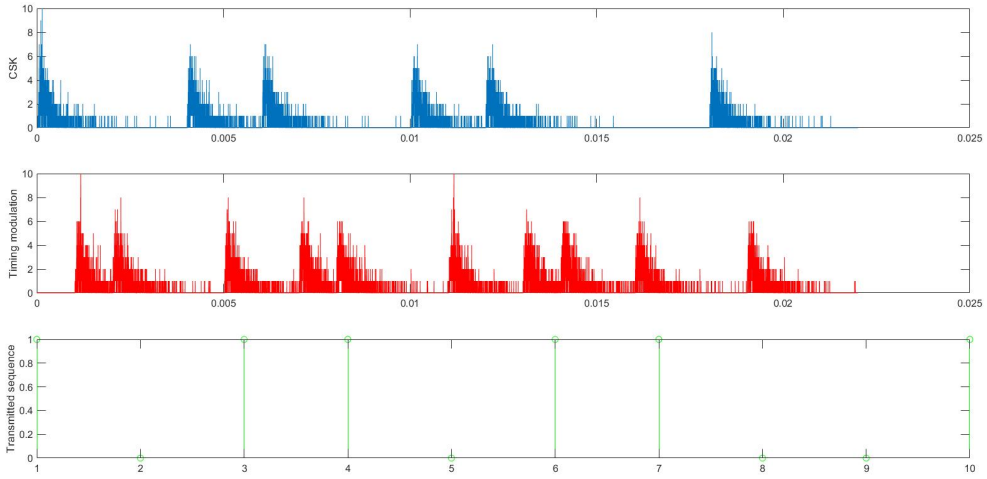


Figure 3.3: An illustration of CSK and Timing modulation environment

Chapter 4

Timing Synchronization

4.1 Synchronization framework

We assume that the receiver knows an approximate range of values of symbol duration. Let $t_s[k] \in \tau[k]$ specify the start of the k-th symbol interval where $\tau[k]$is given as [5]

$$\tau[k] = [t_s[k-1] + T_{min}, ts[k-1] + T_{max}] \tag{4.1.1}$$

T_{min} and T_{max} are in fact the minimum and maximum possible lengths of a symbol interval, respectively. In other words, the length of symbol interval is in $[Tmin, Tmax]$.

To establish symbol synchronization, at the beginning of each symbol interval, the transmitter releases N_B molecules (used for synchronization). Also, depending on whether the value taken by the a[k] and the kind of modulation performed, the transmitter releases either N_A molecules (used for information exchange).

The release/ arrival process of molecules is modeled as a Poisson process when the number of trials in high and the probability of success is small as in [8].

$$r_x(t_n) = \mathcal{P}(\bar{r_x}(t_n)), x \in \{A, B\} \tag{4.1.2}$$

where t_n is the n^{th} time instant for simulation step Δ. $\bar{r_x}(t_n) = \mathrm{E}\{r_x(t_n)\}$, the expected number of arrivals(avg rate) at time t_n is given by

$$\bar{r_x}(t_n) = \sum_{\forall k | t_s[k] \leq t_n} P_x(t_n - t_s[k]) \tag{4.1.3}$$

where $t_s[k]$ is the starting time of the k'th symbol and $P_x(t_n)$ is the expected number of type X molecules bound at the corresponding receptor at time t_n if emission occurred at $t = 0$.

4.2 Optimal ML Synchronization

Let us first define T_{ow} as the size of the observation window used to compute the ML metric for each hypothesis time t for $t_s[k]$, i.e., observation samples $t_n \in [t, t + Tow]$ are used for hypothesis test t.

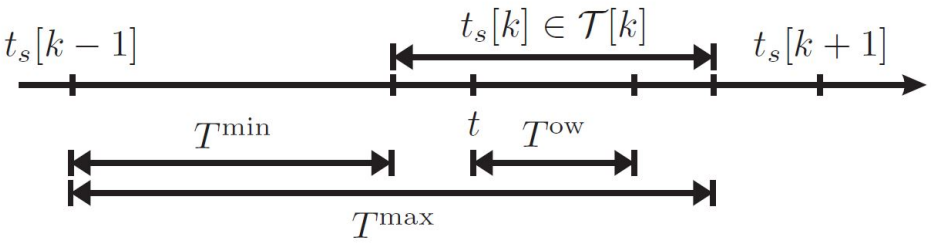


Figure 4.1: Illustration of assumptions adopted for the development of the symbol-by-symbol ML synchronization problem.

From sec.2.4 the arrival time at Rx, in a 1-d environment, for a particle released at time $t = t_s[k]$, can be modeled as an inverse Gaussian random variable. The probability density is expressed as

$$f((t - t_s[k]); \mu, \lambda) = \Big(\frac{\lambda}{2\pi(t - t_s[k])^3}\Big)^{1/2} exp\frac{-\lambda((t - t_s[k]) - \mu)^2}{2\mu^2(t - t_s[k])} \tag{4.2.1}$$

where μ is the mean and λ is the shape parameter. Also sec.2.4 has given the relationship among μ, λ and the communication channel parameters as

$$\mu = \frac{d}{v} \tag{4.2.2}$$

$$\lambda = \frac{d^2}{2D} \tag{4.2.3}$$

where d is the distance between the transmitter and the receiver; v is the medium velocity; and D is the diffusion coefficient of the Brownian particle.

If the Transmitter transmits N_x number of molecules at the beginning of the k'th symbol $t_s[k]$, then the expected number of arrivals during a time interval $[t, t + \delta t], (t > t_s[k])$ is

$$\bar{r_x} = N_x \times f((t - t_s[k]); \mu, \lambda)\delta t \tag{4.2.4}$$

Consider all the previous synchronization symbols we can expand 4.1.3 as:

$$\begin{aligned} \bar{r_x}(t_n) &= \sum_{\forall k | t_s[k] \le t_n} P_x(t_n - t_s[k]) \\ &= \sum_{\forall k | t_s[k] \le t_n} N_x \times f((t - t_s[k]); \mu, \lambda)\delta t \\ \implies \bar{r_x}(t_n) &= \sum_{\forall k | t_s[k] \le t_n} N_x \times \Big(\frac{\lambda}{2pi(t - t_s[k])^3}\Big)^{1/2} exp\frac{-\lambda((t - t_s[k]) - \mu)^2}{2\mu^2(t - t_s[k])}\delta t \end{aligned} \tag{4.2.5}$$

With the assumptions in chapter 3, the ML problem can be mathematically formulated

as

$$\hat{t}_s^{ML}[k] = arg \max_{\forall t \in \tau[k]} \prod_{t_n=t}^{t+T_{ow}} f_P(r_x(t_n), \bar{r}_x(t_n) | t_s[k] = t) \tag{4.2.6}$$

$$= arg \max_{\forall t \in \tau[k]} \Lambda_x^{ML}(t) \tag{4.2.7}$$

Maximising $\Lambda_x^{ML}(t)$ is the same as maximizing $ln(\Lambda_x^{ML}(t))$

$$\hat{t}_s^{ML}[k] = arg \max_{\forall t \in \tau[k]} ln(\Lambda_x^{ML}(t)) \tag{4.2.8}$$

$$= arg \max_{\forall t \in \tau[k]} \sum_{t_n=t}^{t+T_{ow}} [r_x(t_n) ln(\bar{r}_x(t_n)) - \bar{r}_x(t_n) - ln(r_x(t_n)!)] \tag{4.2.9}$$

4.3 Peak Observation-based Synchronization

To formally present the proposed PO synchronization scheme,let us first define constant t_p as

$$t_p = arg \max_{t \geq 0} P_x(t) \tag{4.3.1}$$

In our case for a 1-D brownian motion(wiener process), see section 2.4, $P_x(t)$ is given by the sum of IG distributions $f_T(t - t[k])$'s

$$t_p = \quad arg \max_{t \geq 0} f_T(t) \tag{4.3.2}$$

$$\implies t_p = \quad arg \max_{t \geq 0} \sqrt{\frac{\lambda}{2\pi t^3}} \exp(-\frac{\lambda(t-\mu)^2}{2\mu^2 t}) \tag{4.3.3}$$

The PO symbol synchronization scheme which estimates the start of the symbol intervals as follows

$$\hat{t}_s^P = arg \max_{t_n \in \tau^P[k]} (r_b(t_n)] - t_p \tag{4.3.4}$$

where $\tau^P[k] = [t_s[k-1] + Tmin + t_p, t_s[k-1] + Tmax + t_p]$.

4.4 Simulations

In this section, we provide simulation results to evaluate the effectiveness of the proposed synchronization schemes with the default parameters in Table 4.1 . For simplicity, we assume instantaneous molecule release and a point source transmitter and point reactive receiver model recently developed in [7] and $P_x(t), x \in A, B$ is modeled as $IG(\mu, \lambda)$ as given in 4.3.2. Furthermore, we consider blocks of $K = 100$ symbol intervals. The default values of the system parameters given in Table is used. In order to compare the performances of the considered synchronization schemes, we define the normalized synchronization error(NSE) as

$$\bar{e}_t^2[k] = \frac{\hat{t}_s[k] - t_s[k]}{\bar{T}_{symb}} \tag{4.4.1}$$

Table 4.1: Default Simulation parameters

Parameter	Definition	Value
N_B	Number of type-B molecules released	2×10^3
D	Diffusion Coeff of Type-B molecules	$5 \times 10^{-9} m^2 s^{-1}$
d	Distance between Tx and Rx	$2\mu m$
Δ	Sampling time at Rx	$1\mu s$

An illustration of a 10 symbol detection window for a symbol duration of $2ms$ of the different detection schemes with the Likelihood value for the operation window is show in figure 4.2.

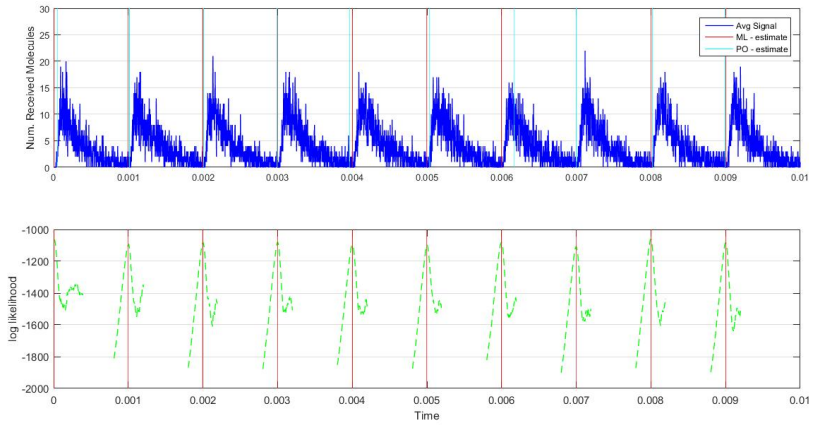


Figure 4.2: Illustration of the proposed ML and PO synchronization scheme

This section aims to test the effects of the symbol duration on timing synchronization(Optimal ML and PO synchronization). The estimates will become inaccurate at higher symbol rates due to ISI. The results for the same are tabulated in table 4.2

Table 4.2: Effect of symbol duration on timing synchronization

Symbol	ML based			PO based		
	Offset			Offset		
Duration(ms)	Timing(s)	Sample	NSE	Timing(s)	Sample	NSE
0.2	0.001881	1881	9.40679	0.0001281	128	0.6409
0.4	3.590e-05	36	0.08960	2.710e-05	27	0.06780
0.6	1.130e-05	11	0.01890	3.650e-05	37	0.06085
0.8	5.530e-06	6	0.00692	3.400e-05	34	0.04253
1.0	3.380e-06	3	0.00338	3.360e-05	34	0.03356
1.2	3.350e-06	3	0.00280	3.330e-05	33	0.02778
1.4	3.150e-06	3	0.00225	3.430e-05	34	0.02450
1.6	2.850e-06	3	0.00178	3.400e-05	34	0.02123
1.8	2.800e-06	3	0.00155	4.190e-05	42	0.02325
2.0	2.260e-06	2	0.00113	3.870e-05	39	0.01937

Chapter 5

Receiver design

5.1 Concentration shift Keying

We assume perfect symbol symbol synchronization and that the receiver knows the start of the k-th symbol interval. At the beginning of each symbol interval, the transmitter releases $N_t x$ type-A molecules, depending on whether the value taken by the the value of the symbol x[k].

Since the molecules propagating through the environment exhibit Brownian motion, a single molecule has a certain hitting probability at the receiver during the current symbol duration. Let this probability, which depends on the distance between the transmitter and the receiver, drift velocity, diffusion coefficient and the symbol duration, be denoted as P_{hit}. See figure 5.1 .

Assuming intra-molecule collisions have negligible effect on the molecule's movement, if N_{emit} molecules are sent in a symbol duration, the number of molecules received within a symbol duration t_s among these molecules is a random variable following a binomial distribution ($\mathcal{B} \rightarrow Binomial$).

$$N_c \sim \mathcal{B}(N_{emit}, P_{hit}(t_s)) \tag{5.1.1}$$

As discussed in section 3.1.1 the CSK uses a threshold based decoding on the total number of molecules received N_R during the current symbol duration. The sources contributing to the N_R are : N_c the molecules belonging to the current symbol duration, N_p the residue molecules from the previous symbol and the molecules from other sources/channels which can be summed up as noise N_n. Here we consider the effect of only one of the previous symbols, hence N_p can be calculated as the difference of two random variables.

$$N_p \sim \mathcal{B}(N_{emit}, P_{hit}(2t_s)) - \mathcal{B}(N_{emit}, P_{hit}(t_s)) \tag{5.1.2}$$

We assume the noise in this communication system is Additive Gaussian White Noise (AGWN). Thus, the N_n is also a random variable following a normal distribution with zero mean and σ variance

$$N_n \sim \mathcal{N}(0, \sigma) \tag{5.1.3}$$

A Binomial distribution $\mathcal{B}(n, p)$ can be approximated with a normal distribution $\mathcal{N}(np, np(1-p))$ when p is not close to one or zero and np is large enough. Using

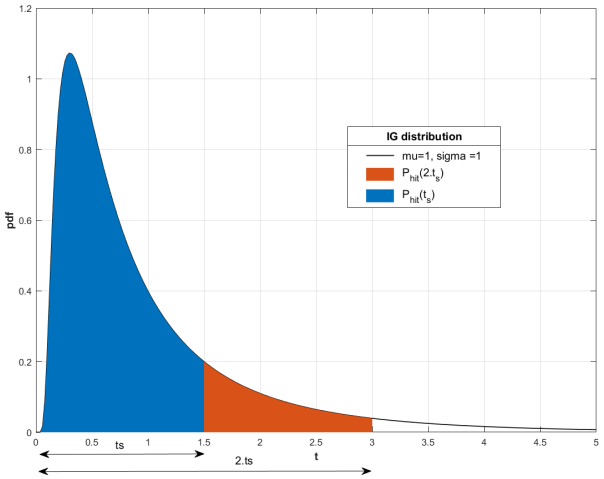


Figure 5.1: An illustration of Phit values for IG distribution

this approximation, we can find N_R as the addition of three normal distributions.

$$N_R \sim N_c + N_p + N_n \tag{5.1.4}$$

where the approximated random variables are,

$$\begin{aligned} N_c &\sim \mathcal{N}(N_{emit}.P_{hit}(t_s), N_{emit}.P_{hit}(t_s).(1 - P_{hit}(t_s)) \\ N_p &\sim \mathcal{N}(N_{emit}.P_{hit}(2t_s), N_{emit}.P_{hit}(2t_s).(1 - P_{hit}(2t_s)) \\ N_n &\sim \mathcal{N}(0, \sigma) \end{aligned} \tag{5.1.5}$$

The receiver decodes the symbol by comparing these N_R values with the threshold values of the modulation technique used. IN B-CSK there exist one threshold which can be calculated based on likelihood estimates using equation 5.1.4.[2]

5.2 Timing Modulation

We now discuss receivers for the AIGN channel by recovering the transmitted message (transmission time) from the times the molecules are received. We implement the ML estimator and the ML detector.

The transmitter releases Type-A molecules simultaneously to convey one of messages$X \in \{x_1, ..., x_t\}$. In the case of simultaneous transmissions, the receiver observes N_A mutually independent arrival times [4]

$$Y_j = X + N_j, \qquad \forall j = 1, .., N_A \tag{5.2.1}$$

where $N_j's$ are iid with $N_j \sim IG(\mu, \lambda), j = 1, ..., N_A$, see section 2.4.

We first consider ML detection of the symbol when multiple molecules are used. Assuming that the receiver knows the values of and through an earlier training phase, it can use the multiple observations $Y_j, j = 1, ..., N_A$ to obtain $\hat{X_{ML}}$.

From 5.2.1 the pdfs $f_{Y_j|X}(y_j|X = x_i), j = 1, .., N_A$are i.i.d. with given by.

$$f_{Y|X}(y, X = x) = \begin{cases} \sqrt{\frac{\lambda}{2pi(y-x)^3}} \exp \frac{-\lambda((y-x)-\mu)^2}{2\mu^2(y-x)} & \text{if } y > x \\ 0 & \text{if } y \leq x \end{cases} \tag{5.2.2}$$

The ML estimate, in this case, is given by

$$\hat{X_{ML}} = arg \max_{x_i} \prod_{j=1}^{N_A} f_{Y_j|X}(y_j|X = x_i) \tag{5.2.3}$$

$$= arg \max_{x_i} \prod_{j=1}^{N_A} (y_j - x_i)^{-3/2} \exp \left(\frac{-\lambda}{2\mu^2} \sum_{j=1}^{N_A} \frac{((y_j - x_i) - \mu)^2}{(y_j - x_i)} \right) \tag{5.2.4}$$

Simplifying the above equation, the ML estimate can be expressed as

$$\hat{X_{ML}} = arg \max_{x_i} \Lambda_{NA}(x_i) \tag{5.2.5}$$

where,

$$\Lambda_{NA}(x_i) = -\frac{3}{2} \sum_{j=1}^{N_A} \log(y_j - x_i) - \frac{\lambda}{2\mu^2} \sum_{j=1}^{N_A} \frac{((y_j - x_i) - \mu)^2}{(y_j - x_i)} \tag{5.2.6}$$

5.3 Simulation

Almost all simulation were run on VIRGO cluster, P.G Senapathy Centre for Computing Resource, IIT Madras. The remaining was run on an intel-i58250U supported by an Nvidia MX150 graphics processor.

This part of the project aims to charactarize the effects of the following on Symbol error Rate(SER)

- **Symbol duration** : In case of CSK and timing modulation the symbol duration decreases the symbol constellation comes closer, error probability should increase. The reverse effect seen when symbol duration increases. For the default parameters the SER vs Symbol duration is plotted in figure 5.2.

  In Concentration shift keying, the effect will not be as dominant as former, but as duration further decreases,ISI comes into picture and the effect will be visible.

- **Distance Between Rx and Tx** : As the distance increases the arrival times will be spread out with large variance which can make timing synchronization more difficult. Also since the underlying mechanism Brownian motion occurs for a longer time, the randomness of the observations increases. The reverse effect should seen when symbol duration increases, the arrival process should see sharp peaks and less outliers. With smaller distances higher data rates should be achievable. The variation of SER vs drift velocity is shown in figure 5.3a and figure 5.3b supports the prognosis.

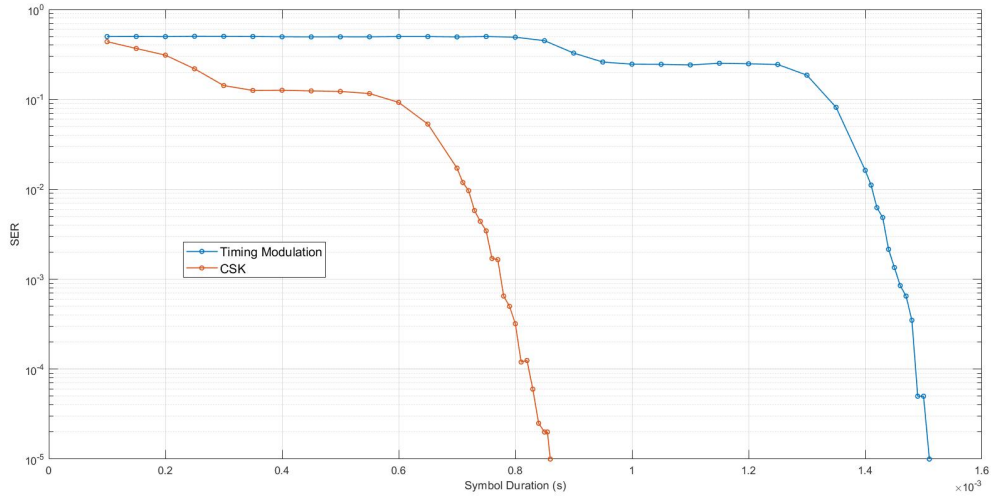
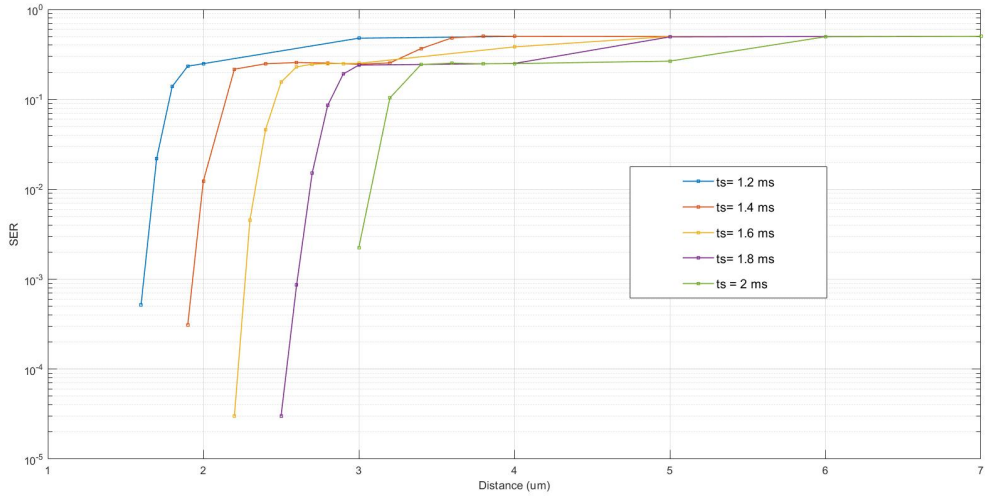


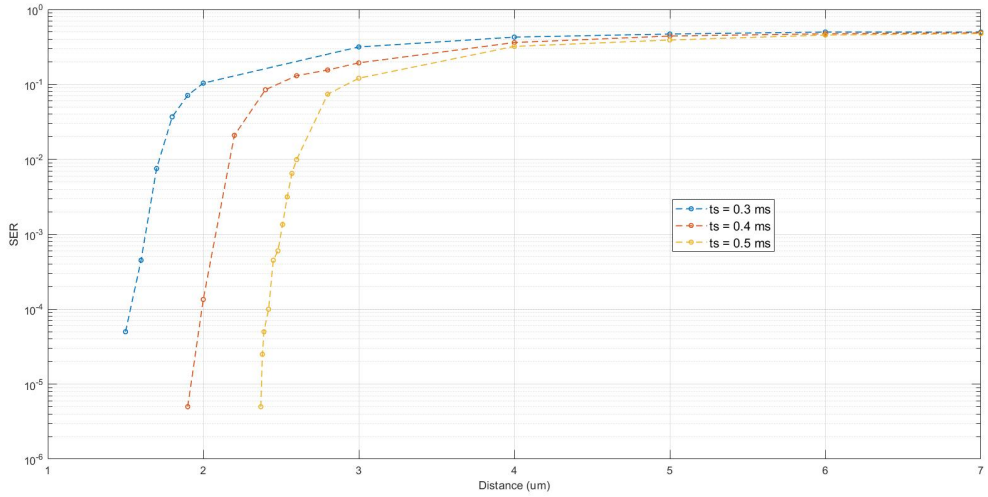
Figure 5.2: SER vs Symbol duration with default parameters as in table 4.1

- **Drift velocity** : As the drift velocity increases, the propagation delay of the molecules decreases, i.e. the peak of the arrivals should be close to the beginning of the symbol. The variation of SER vs drift velocity is shown in figure 5.4a and 5.4b. It can be observed that the variation with drift velocity is converse of that of with distance.

- **Diffusion co-efficient** : As the diffusion coefficient increases, the spread of arrivals is expected to increase. From figure 5.5a and 5.5b, it can be seen that at very low and high diffusion coefficients the SER is low,i.e there is a range of values where transmission is undesirable.

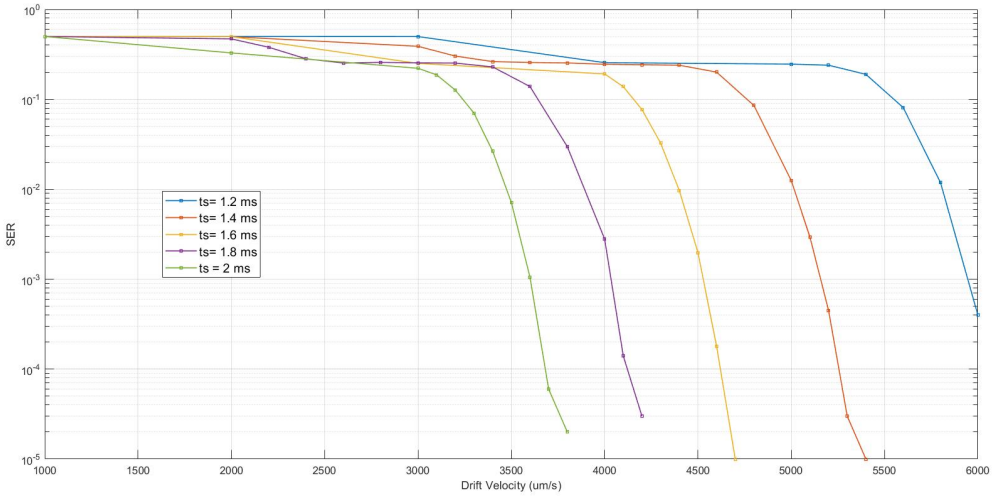


(a) Timing modulation

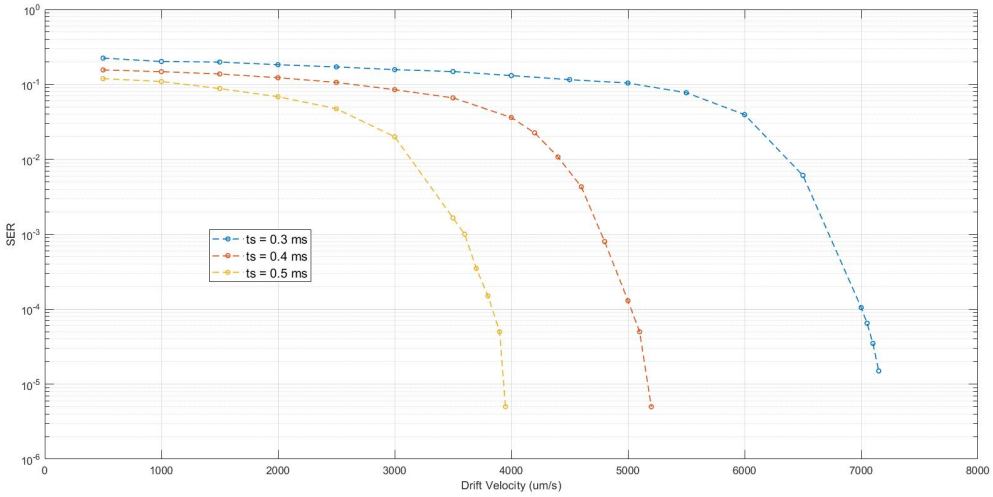


(b) Concentration shift Keying

Figure 5.3: SER vs Distance with default parameters as in table for the different modulation schemes

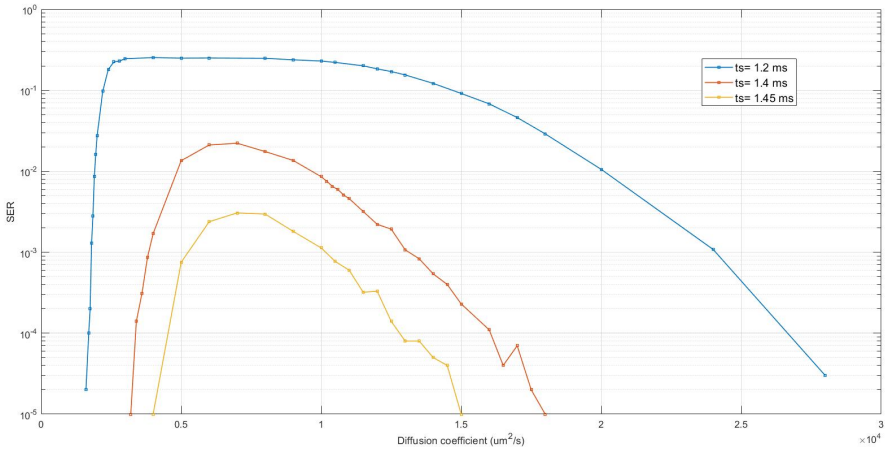


(a) Timing modulation

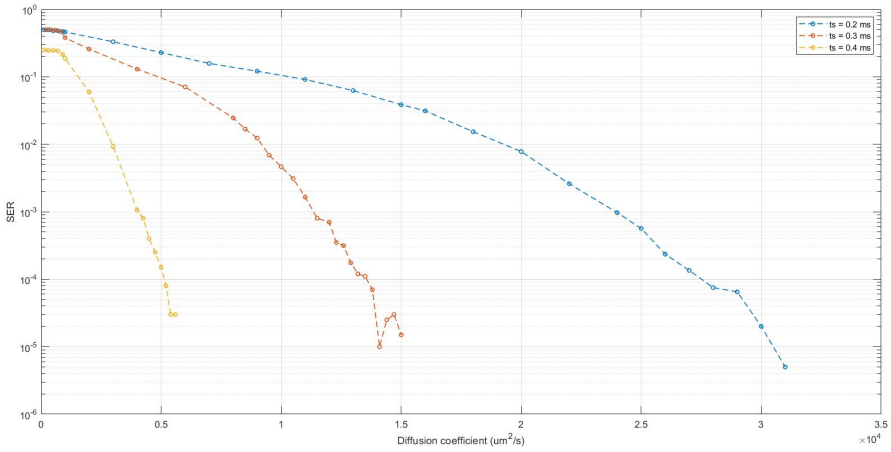


(b) Concentration shift Keying

Figure 5.4: SER vs Drift velocity with default parameters as in table for the different modulation schemes



(a) Timing modulation



(b) Concentration shift Keying

Figure 5.5: SER vs Diffusion coefficient with default parameters as in table for the different modulation schemes

Bibliography

[1] H. A. Johnson and K. D. Knudsen, "Renal efficiency and information theory", *Nature*, vol.206, pp.930931, 1965

[2] M. S. Kuran, H. B. Yilmaz, T. Tugcu, I. F. Akyildiz, "Modulation Techniques for Communication via Diffusion in Nanonetworks", *IEEE Communications Society subject matter experts for publication in the IEEE ICC 2011*

[3] Na-Rae Kim, Chan-Byoung Chae , "Novel Modulation Techniques using Isomers as Messenger Molecules for Nano Communication Networks via Diffusion", *IEEE journal in selected areas in communications/supplement — Part 2, Vol. 31, No. 12, December 2013.*

[4] K. V. Srinivas, Andrew W. Eckford, "Molecular Communication : The Additive Inverse Gaussian Noise Channels", *IEEE Transactions on Information Theory Vol.58, No.7*, 2012

[5] Vahid Jamali, Arman Ahmadzadeh, and Robert Schober , "Symbol Synchronization for Diffusive Molecular Communication Systems", *IEEE ICC 2017 SAC Symposium Molecular, Biological, and Multi-Scale Communications Track.*

[6] Tadashi Nakano, Andrew W Eckford, Tokuko Haragachi, "Molecular Communication", *Cambridge University Press*

[7] Lin Lin, Jiali Zhang, Maode Ma, and Hao Yan , "Time Synchronization for Molecular Communication With Drift", *IEEE Communication Letters, Vol. 21, No. 3, March 2017.*

[8] V. Jamali, A. Ahmadzadeh, C. Jardin, C, Sticht and R. Schrober "Channel Estimation for a Diffusive Molecular Environment" *IEEE Trans. Commun.*, vol.13, no. 3, pp. 4238-4252, Oct 2016

[9] "IEEE ComSoc Technology News Special Issue on Molecular Communication" *CTN Issue: March 2014*